THE EFFECT OF HAIR ARTIFACTS ON SKIN IMAGE DATASETS

A DISSERTATION SUBMITTED TO MANCHESTER METROPOLITAN UNIVERSITY
FOR THE DEGREE OF MASTER OF SCIENCE
IN THE FACULTY OF SCIENCE AND ENGINEERING

2023

By
Jesty Sebastian
Department of Computing and Mathematics

Contents

List of Tables

List of Figures

Abstract

Skin lesion classification, vital for the early detection and diagnosis of skin diseases, particularly melanoma, can be significantly impacted by the presence of artifacts, such as hair, in dermatoscopic images. In this study, a meticulously curated and balanced dataset, originating from the extensive ISIC dataset, was utilized to train a robust skin lesion classification model. Furthermore, a specialized skin image dataset, also derived from ISIC data, was employed for comprehensive testing and evaluation, encompassing dermatoscopic images augmented intentionally with artificially added hair patterns of varying densities. A machine learning model, based on ResNet50, was employed for skin lesion classification using this dataset. Subsequently, the application of hair artifact removal techniques to test images allowed for the model's performance evaluation through rigorous probability distribution analysis.

The findings revealed that as hair density decreases, there is a consistent reduction in the probability of melanoma classification following hair artifact removal. Despite occasional misclassifications, it is evident that the elimination of hair artifacts substantially enhances classification accuracy. This research underscores the critical importance of hair artifact removal in elevating the accuracy of skin lesion classification models. The methodology and insights derived from this study hold the potential to contribute significantly to the development of more robust and precise dermatology diagnostic tools, ultimately leading to improved patient outcomes within the field of dermatological practice.

Declaration

No part of this project has been submitted in support of an application for any other degree or qualification at this or any other institute of learning. Apart from those parts of the project containing citations to the work of others, this project is my own unaided work. This work has been carried out in accordance with the Manchester Metropolitan University research ethics procedures, and has received ethical approval number 56595.

Signed: Jesty Sebastian

Date: 29th Sep 2023

Acknowledgements

I would like to express my gratitude to my supervisor, Prof. Moi Hoon Yap, for her support during the course of this study.

Abbreviations

ANN	Artificial Neural Networks
CAD	Computer-Aided Diagnosis
CNN	Convolutional Neural Network
DHS	Digital Hair Removal
DL	Deep Learning
GAN	Generative Adversarial Network
ISIC	International Standard Industrial Classification of All Economic Activities
MF	Median Filtering
NN	Neural Network
PCA	Principal Component Analysis
PDE	Partial Differential Equations
PSNR	Peak Signal-to-Noise Ratio
ResNet	Residual Network
ROC	Receiver Operator Characteristic
SSIM	Structural Similarity Index Measure
SVM	Support Vector Machine
UV	Ultra Violet

Chapter 1

Introduction

1.1 Problem Statement

Skin cancer is a significant public health concern, characterized by the abnormal growth of skin cells triggered by exposure to ultraviolet (UV) radiation from the sun, tanning beds, or other sources. It is one of the most common forms of cancer globally, with several distinct types, the most prevalent being basal cell carcinoma, squamous cell carcinoma, and melanoma (**Narayanamurthy V et al**; **2018**). Skin cancer typically manifests as changes in the appearance of moles, freckles, or other skin lesions, and its incidence continues to rise, posing a substantial threat to individuals of all ages. While it is often highly treatable when detected early, neglecting skin cancer can lead to severe consequences, including disfigurement and even fatality. Thus, understanding the basics of skin cancer, its risk factors, and preventive measures is crucial for maintaining skin health and reducing the impact of this disease.

The early detection of probable skin cancers is greatly aided by visual inspection, which is a primary technique for identifying skin diseases. Dermatologists in particular rely on their skill to visually examine skin abnormalities, looking for evidence of irregularity including changes in size, shape, color, texture, and symmetry in moles, freckles, and other lesions. This approach, however, has inherent drawbacks. False positives or false negatives may occur from subjective interpretation, which can cause variation in diagnoses. Visual inspection may also miss small changes over time or the full depth of a lesion, making it difficult to spot malignancies which develop slowly(**Dinnes J et al**; **2018**). Additionally, factors like patient variability and the time and resource intensiveness of comprehensive examinations should be considered. Despite these limitations, visual inspection remains a valuable and accessible tool for initial skin lesion assessment, but it should be supplemented with additional diagnostic techniques when warranted to ensure accurate diagnoses and timely interventions.

Deep learning has revolutionized the field of skin cancer classification by offering increased accuracy and efficiency in diagnosing this potentially life-threatening disease. Specifically, convolutional neural networks (CNNs) have played an important role in this advancement. These sophisticated algorithms excel at automatically learning and extracting intricate features from skin images, enabling them to distinguish between benign and malignant lesions, including various types of melanomas, with remarkable precision (**A. K. Gairola et al. 2022**). Their ability to process large amounts of data and identify complex patterns has significantly improved diagnostic accuracy and speed. However, the success of deep learning in skin cancer classification is connected inherently to the quality and diversity of the training datasets. Continuous efforts to curate comprehensive and representative datasets, coupled with ongoing research into model interpretability and generalizability, are essential in harnessing the full potential of deep learning for early skin cancer detection and treatment.

The impact of artifacts in skin image datasets is a critical challenge that significantly influences the performance of convolutional neural network (CNN) models in dermatology. These artifacts, including borders, hair, measurement devices, air pockets, and clinical markings, often infiltrate skin image datasets and introduce biases that can undermine the reliability of machine learning algorithms. As highlighted by (**Pewton**; **S.W. and Yap**; **M.H.**; **2022**), the presence of these artifacts can distort image features and lead to erroneous classification results, particularly in skin cancer diagnosis. Among these artifacts, hair artifacts are of particular concern, as they can obscure important details and texture information within skin images, making it challenging for CNN models to discern the true characteristics of lesions. Additionally, as noted by (**Sultana et al. 2014**) in 2014, these artifacts can complicate the segmentation and classification of images, making it more challenging for CNN models to accurately identify and differentiate between benign and malignant lesions. Therefore, addressing artifact-related issues through careful data curation and advanced preprocessing techniques, with a specific focus on hair artifact removal, is essential to enhance the robustness and effectiveness of CNN models in dermatological image analysis, ultimately improving diagnostic accuracy and patient care.

This project builds upon prior research that has examined the influence of the skin hair artifact within skin image datasets, specifically focusing on its impact on melanoma classification. Previous studies, such as the work conducted by (**Cassidy et al. 2022**), laid the foundation by introducing a well-curated, balanced dataset comprising melanoma and other skin lesion images. Additionally, the investigation by (**Pewton and Yap 2022**) examined the presence of dark corners as an artifact within the same dataset. This project takes a significant step forward by narrowing its attention to the skin hair artifact and seeks to comprehensively understand how the presence of skin hair affects the performance of CNN models in melanoma classification. By concentrating on this specific artifact, the research aims to uncover the unique challenges it poses and the potential biases it introduces into the classification process. This in-depth exploration will contribute vital insights to enhance the accuracy and reliability of CNN-based melanoma diagnosis, ultimately advancing the field of dermatological image analysis.

1.2 Aim

The aim of this project is to investigate and quantify the impact of the skin hair artifact within skin image datasets on the performance of convolutional neural network (CNN) models for melanoma classification. Through a systematic and comprehensive analysis, we seek to determine how the presence of skin hair influences the accuracy, robustness, and reliability of CNN-based melanoma diagnostic systems. By identifying the challenges introduced by this specific artifact, this project aims to develop strategies and insights that can lead to more effective preprocessing techniques and model adaptations, ultimately improving the accuracy of melanoma classification and contributing to the advancement of dermatological image analysis.

1.3 Objectives

The objective of this project is to enhance the accuracy, interpretability, and reliability of skin cancer classification models by addressing artifacts in skin image datasets. This will be achieved by:

- **Reviewing Existing Literature**: Conduct an in-depth review of the current literature on skin hair artifacts in skin image datasets to establish a strong foundational understanding of these anomalies.

- **Adapting and Implementing the Pre-trained Model**: Adapt and implement a pre-trained ResNet model for detecting artifacts, with a focus on ensuring effective artifact recognition.

- **Analyzing the Impact of Hair Artifacts**: Perform a comprehensive analysis to quantify the influence of hair artifacts on the performance of skin cancer classification models, providing valuable insights into the extent of their impact.
- **Understanding the Importance of Hair Artifact Removal**: Explore methods to detect and reduce hair artifacts, recognizing their vital role in improving the accuracy of deep learning-based skin cancer diagnosis.
- **Application of Hair Removal**: Implement hair artifact removal techniques on dermatological images and assess their impact on classification accuracy.

These objectives collectively contribute to a thorough exploration of the role of hair artifacts in skin cancer classification, highlighting the importance of hair removal mechanisms in advancing the field of dermatological image analysis and skin cancer diagnosis.

Chapter 2

Literature Review

2.1 Melanoma Classification and Skin Lesion Detection

Melanoma classification has undergone a transformative evolution as a result of advances in medical imaging, computer technology, and machine learning. From the days when dermatologists relied solely on visual inspection, the introduction of dermoscopy provided a breakthrough by allowing magnified examination and improved lesion assessment. Subsequent developments saw the emergence of rule-based systems and the introduction of Computer-Aided Diagnosis (CAD) systems, which used algorithms to analyze images and suggest diagnoses. The development of computer-aided diagnosis (CAD) systems for cancer detection has involved the integration of image processing, pattern recognition, and neural networks (H. Lee and Chen 2014b). These CAD systems use algorithms to detect potentially cancerous regions in medical images. Textures and shapes are extracted to depict these areas, and distinct patterns are learned from sample images by a supervised classifier for precise diagnosis (H. Lee and Chen 2014a). The importance of image segmentation accuracy, which is critical for locating tumors in treatment planning, was emphasized by (**S.C cheng et al. 2013**). But the automation of CAD systems to analyze large image datasets for accurate diagnoses while accounting for human biology variations was a challenge (H. Lee and Chen 2014a).

Texture analysis combined with machine learning techniques such as Artificial Neural Networks (ANNs) have shown promise in skin cancer detection, owing to their nonlinear prediction capabilities and structural similarity to the human brain. ANNs are made up of input, hidden, and output layers, and computations are learned using techniques such as backpropagation. They've been widely used in skin cancer detection systems to classify extracted features. For example, (F. Xie, Fan, et al. 2016) developed a system that outperformed other classifiers by using a self-generating NN to extract lesions, followed by feature extraction and classification using an ensemble NN model. (Masood, Al-Jumaily, and Adnan 2014) proposed an ANN-based diagnostic system that outperformed other learning algorithms in terms of specificity and sensitivity. (Aswin, Jaleel, and Salim 2014) investigated the use of ANN and genetic algorithms for skin cancer detection, yielding an overall accuracy of 88%. These ANN-based systems provide efficient and accurate methods of categorizing skin lesions and have the potential to aid in the early detection of skin cancer. (Alquran et al. 2017) proposed a combined image processing and machine learning technique. Preprocessed and segmented dermoscopy images are used to isolate lesion regions. Dermoscopy's ABCD rule is used to analyse extracted features such as colour and texture. Following accurate lesion classification with Support Vector Machine (SVM), Principal Component Analysis (PCA) improves feature efficiency. This method provides a comprehensive approach for detecting and classifying melanoma.

Traditional techniques such as decision trees and support vector machines have been supplemented by cutting-edge techniques. Deep learning, particularly Convolutional Neural Networks (CNNs), ushered in a new era by enabling automated feature extraction from images. Early work by (Haggenmüller et al. 2021) demonstrated the effectiveness of CNNs in binary classification, outperforming dermatologists in distinguishing between melanoma and melanocytic nevi. (Marchetti et al. 2018) established ensemble approaches that outperformed dermatologists' expertise in malignancy classification. CNNs also demonstrated their prowess in multiclass tasks, even when integrating different image types. Furthermore, CNN-based classifiers trained on clinical images outperformed dermatologists who had not been trained on clinical images. (Gairola, Kumar, and Sahoo 2022), compared different convolutional neural network models and investigated how well CNN's models perform in diagnosing melanoma and non-melanoma lesions. These advancements not only improved diagnostic accuracy, but also paved the way for real-time mobile apps and telemedicine, making melanoma detection more accessible and efficient.

2.2 Factors Affecting Melanoma Classification

Skin image datasets frequently include a variety of artifacts, such as Borders, Hair, Measurement Device, Air Pocket(s), and Clinical Markings, which can introduce biases and affect the performance of CNN models (Pewton and Yap 2022). Due to these artifacts, segmentation and classification of images also become more challenging. To develop accurate and comprehensible skin cancer classification models, it is crucial to comprehend and minimize the effects of artifacts on skin image datasets (Abbas, Celebi, and Garcıa 2011). By recognizing and removing artifacts, the models can concentrate on learning relevant features and making precise predictions based on the actual properties of skin lesions.

artifacts such as borders, hair, measurement devices, air pockets, and clinical markings can all have an impact on skin image datasets. These artifacts have the potential to introduce biases and have an impact on CNN model performance (Pewton and Yap 2022). Borders, for example, have the ability to obscure the edges of skin lesions, making identification difficult for the model. Similarly, hair may mask skin lesions, and the presence of measurement devices can create shadows that make it difficult to distinguish between different types of skin lesions. Air pockets may also cause artifacts by altering the perceived size of skin lesions. Clinical markings, such as moles or birthmarks, may also be classified as artifacts because the model may misinterpret them (Sultana, Dumitrache, et al. 2014).

Understanding and minimising the effects of artifacts on skin image datasets is critical for developing accurate and understandable skin cancer classification models. One method is to use a large and diverse dataset Mustafa (2017). A large dataset will assist the model in learning to differentiate between different types of skin lesions, including those with artifacts. A diverse dataset will aid the model's learning to generalise to new cases, including those with varying imaging conditions or artifacts. Unfortunately, big data is difficult to access, particularly for medical applications (K. W. Lee and Chin 2020).

Using data augmentation techniques is another way to reduce the impact of artifacts. Data augmentation is a technique for creating artificial variations on existing data. This can aid the model's learning to recognise features relevant to skin cancer classification, even if those features are obscured by artifacts (K. W. Lee and Chin 2020). For example, images can be flipped, rotated, or cropped to create new variations.

Finally, a solid image pre-processing technique is required. artifacts can be removed or reduced with image pre-processing Madooei et al. (2012). For example, borders can be removed, and hair and measurement devices can be masked.

2.3 Influence of Hair Artifacts on Melanoma Classification

Skin image datasets are important resources for training and evaluating models that classify various skin conditions, such as melanoma. The presence of hair artifacts in these datasets, on the other hand, poses a significant challenge. Hair artifacts can obscure important features of skin lesions, making accurate classification more difficult (Abbas, Celebi, and Garcıa 2011). The complication stems from the possibility of misdirecting the model's attention, as it may prioritize hair over the actual lesion. Addressing the effect of hair artifacts on skin image datasets is critical for improving the reliability and precision of automated systems designed for melanoma classification. Preprocessing techniques, data augmentation, and collaboration with medical experts are frequently used to ensure the effectiveness of classification models in real-world scenarios.

In a study conducted by (Li and Shen 2018), the influence of hair removal on melanoma classification in dermoscopic images was explored. Their research involved employing a digital hair removal (DHS) algorithm to eliminate hair from the images. Subsequently, they evaluated the classification performance of a deep learning model trained on both the original images and the hair-removed images. Notably, the outcomes demonstrated a noteworthy enhancement in performance for the model trained on hair-removed images (94.1%), surpassing the performance of the model trained on original images (88.9%).

Similarly, (F. Xie, Yang, et al. 2020) delved into the significance of hair in the context of melanoma classification. Their approach centered on utilizing a deep learning model to distinguish the lesion from the surrounding hair in dermoscopic images. By isolating the segmented lesion area, they trained a model to classify melanoma. Impressively, the model trained on the segmented lesion area exhibited superior performance in comparison to the model trained on the unaltered images (88.6%).

(Pour and Seker 2020) contributed to this area by employing a convolutional neural network (CNN) to classify melanoma in dermoscopic images. Their investigation encompassed a comparison between the classification performance of the CNN trained on original images and that of the CNN trained on images with hair digitally removed. Their findings underscored the effectiveness of the CNN trained on hair-free images, which achieved a higher accuracy rate (93.3%) relative to the CNN trained on the original images (90.6%).

According to these findings, hair can have a negative impact on melanoma classification in dermoscopic images. This is due to the fact that hair can obscure the features of the lesion, making identification difficult for the classifier. Hair removal can improve melanoma classifier performance, but it is important to note that hair removal can introduce artifacts into the images, which can also affect classification performance.

2.4 Hair Removal Mechanisms and their Application

In the field of dermatology and medical imaging, addressing hair occlusion in skin images is a critical challenge that directly impacts the accuracy of subsequent analysis algorithms. Traditional methods of manual shaving to remove hair from the skin before capturing images have proven to be time-consuming, expensive, and impractical for total-body nevus imaging. In response, various software-based solutions have been proposed to enhance the quality of skin image datasets by mitigating the hair occlusion problem.

In **Lee et al. -1997** introduced the DullRazor algorithm, which uses image processing techniques to identify and replace hair pixels with nearby non-hair pixels, significantly improving segmentation results and enabling accurate lesion identification. Later an improved version based on the morphological closing operation to specific color components that specifically targeting on dark hairs is proposed by P. Schmid-Saugeon et al. in 2003. Meanwhile. In 2008, (Zhou et al. 2008) proposed a mechanism for detecting and modelling explicit curvilinear structures. It replaces artifact pixels with feature-guided exemplar-based inpainting, preserving underlying lesion features while removing distracting artifacts. Hearn (2008) utilizes the fast marching method to enhance the segmentation process, particularly for the purpose of eliminating hair artifacts from dermoscopic images.

(Wighton, T. K. Lee, and Atkins 2008) presented an alternative approach using inpainting to fill occluded regions with estimated colors, outperforming DullRazor's linear interpolation and demonstrating superior accuracy and consistency. In 2009, (F.-Y. Xie et al. 2009) combined a morphological closing-based top-hat operator for hair detection with a partial differential equation-based image inpainting algorithm for hair-occluded information restoration, resulting in improved texture and accuracy in repaired skin regions.

(Nguyen, T. K. Lee, and Atkins 2010) addressed the limitations of existing hair segmentation methods and proposed a novel approach involving hair identification and verification/reconstruction stages, successfully segmenting light and dark hair without prior knowledge of the hair type.

(Kiani and Sharafat 2011) introduced the E-shaver method, utilizing edge detection and interpolation techniques to detect and remove both dark and light-colored hairs more efficiently than traditional methods. Additionally, (Fiorese, Peserico, and Silletti 2011) introduced VirtualShave, which combines segmentation and inpainting modules to create images closely resembling naturally hairless skin through quantitative evaluations, showcasing superior performance compared to existing techniques. These software-based approaches collectively contribute to improved accuracy and reliability in dermatoscopic image analysis and diagnosis, by effectively addressing the hair occlusion challenge.

In 2011,Abbas, Celebi, and Garcıa (2011) introduced an automated algorithm designed for hair detection and correction. The algorithm employs matched filtering technique to identify hair, followed by the application of edge-based methods, morphological operations, and thresholding to improve the appearance of non-skin regions within the image. In 2012, A (Sultana, Ciuc, et al. 2012) combined Top Hat Transform with bicubic interpolation. These methods effectively remove hair artifacts, enhancing the quality of dermatoscopic images and facilitating subsequent lesion detection and diagnosis.

A new mechanism proposed by (Toossi et al. 2013) that involves enhancing image contrast with Principal Component Analysis (PCA) and employing adaptive Canny edge detection. Morphological operations are then used to precisely segment the detected hairs. During the hair repair step, a multi-resolution coherence transport inpainting technique replaces hair pixels with patches resembling nearby skin pixels, improving overall image quality for precise skin condition diagnosis. (Satheesha, Satyanarayana, and Giriprasad 2014) used pixel interpolation to remove hair from skin lesion images. The quadratic Radon transform is used in this technique to effectively detect and remove hairs. A. (Nasonova et al. 2014) introduced an innovative approach where hair detection is accomplished through the precise identification of hair regions using Gabor filters. Subsequently, the regions containing removed hair are subject to restoration using an algorithm based on Partial Differential Equations (PDE). To further improve the quality of the results, a two-dimensional image warping technique is applied to sharpen the edges of the reconstructed image.

(Mirzaalian, T. K. Lee, and Hamarneh 2014) present a novel method for extracting hair features based on quaternion-based tubularness filters and Markov random fields. A realistic hair-on-skin simulator is also included to validate the algorithm. Additionally, dual-channel matched filters are proposed to enhance light and dark hairs separately. In 2015 (D. A. F. Mahmood and H. A. Mahmood 2015) proposed a two-step based method. In first step, an adaptive Canny edge detector and morphological operators are used to identify and delineate hair in dermoscopic images and in second step the masked region is replaced with neighboring pixel information by inpainting.

In 2016, (Koehoorn et al. 2016) introduced a technique aimed at effectively eliminating both long and short hairs from images while also reducing the occurrence of false positives. This was achieved through the utilization of multi-level thresholding and inpainting techniques in their approach. In 2017, Zaqout (2017) introduced a block-based approach to hair detection that involves using a bottom hat operation in a colour space. The image is divided into non-overlapping blocks during the inpainting phase, and each block is processed using histogram functions and morphological operations. (Salido and Ruiz Jr 2018) identified and eliminated the hair using median filtering (MF) and morphological bottom-hat filtering. The lost pieces are then restored using the harmonic inpainting method. In 2019, (Talavera-Martınez, Bibiloni, and González-Hidalgo 2019) developed a benchmark for hair removal using six cutting-edge algorithms, each using a unique strategy to segment and inpaint the hair pixels. These techniques include greyscale closing, greyscale and RGB top hat transformation, Gaussian derivative, matched filters, and canny edge detector.

In contrast to past methods, deep learning (DL) has lately become a potent technique for image analysis. In a number of medical applications, convolutional neural networks (CNNs) in particular have displayed amazing abilities, showcasing their superior performance. Deep learning algorithms have been used successfully for tasks including filling in blank areas in images, smoothing blurry images, and reducing noise levels in images. In order to eliminate noise from images and fill in missing areas, Xie et al. (2012) developed a method that depends on sparsity and leverages deep networks that have been trained beforehand.

D. Kim et al. (2021) present an unsupervised algorithm for removing hair artifacts from dermoscopic images of skin lesions using a Generative Adversarial Network (GAN). Even in difficult scenarios involving different hair densities, lengths, and colours, the algorithm effectively identifies and eliminates hair features. Notably, the algorithm improves melanoma classification accuracy by significantly improving classification results after hair removal. Similarly, W El-Shafai et al. (2023) proposed a method to iteratively eliminate hair artifacts from photos using a Generative Adversarial Network (GAN) combined with Convolutional Neural Networks (CNNs). These studies show how GAN-based approaches can improve medical image quality and diagnostic outcomes in dermatology. In 2021, Wei Li et al. (2021) suggested an approach that combines hair segmentation using a U-Net-based architecture, hair gap inpainting by transfer learning, and evaluation using a unique measure Intra-SSIM. By successfully removing hair artifacts, this method seeks to improve the quality of skin lesion images for accurate medical analysis.

2.5 Hair Removal's Impact on Classification

Studies have repeatedly shown that removing hair artifacts from skin lesion images considerably increases the accuracy of melanoma classification algorithms (D. Kim et al., 2021). The presence of hair in these images can contribute unwanted noise, obstruct feature extraction, and result in incorrect classification (Van der Mei et al., 2002). The implementation of hair removal algorithms successfully improves the signal-to-noise ratio (W El-Shafai et al. 2023), allowing classifiers to focus on the actual traits of the lesions themselves.

A notable benefit of using hair removal techniques is the decrease in false positive classifications (Victor, A. and Ghalib, M.R., 2017). Sometimes, hair strands resemble the characteristics of melanoma, which can lead to erroneous positive diagnoses. Models are less likely to mistake benign tumors for malignant ones by getting rid of hair artifacts. The removal of hair artifacts also makes it easier to extract superior features from skin lesion images (M. A. Rahman, 2017), which increases the prominence and informativeness of features related to lesion color, shape, and texture. These improved feature extraction capabilities make a significant contribution to the overall improvement in model performance.

Additionally, using hair removal methods helps to achieve results that are more consistent and repeatable in a variety of datasets and situations. Given that hair artifacts can have a wide range of appearances and densities; their presence could potentially weaken the accuracy of melanoma categorization. As a result, the elimination of these artifacts strengthens the validity of classification models (Abbas Q et al., 2011). Most significantly, the improved precision made possible by hair removal techniques has important therapeutic implications. The accuracy of melanoma diagnosis is crucial for patient outcomes, and the decrease in false negatives brought on by these techniques has the potential to help with the early detection of malignant lesions, thus saving lives.

2.5.1 Impact on Diagnosis

Even though hair artifacts do not obscure melanoma features, they can nonetheless affect how accurately melanoma classification algorithms perform. These artifacts add extra textures and color variations to the images, which could cause misclassification. Determining whether the presence of hair artifacts significantly impacts diagnosis accuracy is therefore critical.

2.5.2 Resource Allocation

The availability of resources, including processing power and time, should be considered while deciding whether to remove hair artifacts. Researchers must balance the price and resource limits of some hair removal methods against the advantages of greater accuracy.

2.5.3 Data Quality Considerations

It is crucial to evaluate the dataset's overall quality. Hair removal may be considered to improve dataset quality and ensure reliable model training if the dataset has a significant proportion of hair artifacts and these artifacts negatively affect classification results.

2.5.4 Clinical Relevance

The clinical relevance of classification results should also be taken into account. Hair artifacts may cause misdiagnoses or have an adverse effect on patient outcomes. The potential clinical effects of hair artifact removal may therefore have an impact on the choice.

2.5.5 Expertise and Ethical Considerations

Successful implementation of hair removal methods requires expertise in image processing. In order to ensure that data processing complies with privacy and ethical standards, researchers and practitioners need also be aware of regulatory requirements and ethical considerations.

2.5.6 User Preferences

In the decision-making process, user preferences, particularly those of dermatologists and physicians who interpret the data, are taken into consideration. In order to facilitate interpretation, some users could prefer images without hair artifacts, whilst others might feel at ease working with images that have.

2.6 Summary

The reviewed research makes it clear that the presence of hair might present a number of difficulties and complexities for the precise diagnosis of melanoma lesions. These difficulties include everything from a rise in false positives and negatives to a possible masking of crucial diagnostic traits.

Furthermore, the investigations have shown that there isn't a single, universal approach to the problem of hair artifacts in melanoma classification. The choice of whether to use a hair removal method should be carefully considered, taking into account different aspects such as the dataset's properties, the kind of machine learning model used, and the final objective of the melanoma detection system.

Additionally, it is crucial to understand that, even while hair removal might occasionally enhance the effectiveness of melanoma detection systems, it may not always be essential or advantageous when the presence of hair does not obscure the lesion's sight. The development of reliable algorithms that can accurately identify between benign and malignant tumors despite the presence of hair should be the main goal in such circumstances.

To better comprehend the trade-offs associated with image preprocessing for melanoma classification, it is crucial to investigate the requirement of a hair removal strategy. It is clear that not all hair in skin images is problematic for correct diagnosis. Therefore, further study should concentrate on developing standards and guidelines for determining when hair removal is actually necessary.

This study could be looking into how hair artifacts affects melanoma classification algorithms, trying to answer questions like: Does hair provide a considerable amount of noise that impairs classification accuracy? Are there any particular machine learning models or preprocessing methods that may successfully handle hair without the requirement for removal?

By addressing these questions, this research can help to clarify how hair artifacts play a part in melanoma classification. This will make it possible to create more context-aware algorithms that can adjust to the presence of hair, thereby enhancing the effectiveness of melanoma detection systems.

Chapter 3

Methodology

The purpose of this chapter is to explain the ideas and concepts and procedures employed within this research. In this study, a comprehensive methodology has been developed for assessing the performance of a ResNet model in the classification of skin images for melanoma detection, with a focus on the consideration of the presence or absence of hair artifacts. The objective of this methodology is to explore whether there is a noticeable influence on the accuracy and dependability of the model in melanoma classification by hair artifacts..

3.1 Dataset Description

3.1.1 Curated Balanced Dataset for Train and Validation

The curated balanced dataset proposed by (Cassidy et al. 2022) is used in this study for train and validate the model. The dataset contains a total of 9,810 images with equal number of melanoma images and other category skin lesion images. The authors splitted the dataset into 7848 training images and 1962 validation images. This dataset is widely accepted for skin lesion classification tasks.

The dataset includes images from ISIC datasets from the year of 2016 to 2020, with ground truth labeling. (Cassidy et al. 2022) performed comprehensive steps to ensure the data is clean and suitable for classification tasks during the creation of curated balanced dataset.

The elimination of all superpixel images from the ISIC 2017 training and test sets was the first step in creating this curated balanced dataset. These superpixel images

Dataset	Train	Test	Total
ISIC 2016	900	379	1279
ISIC 2017	2000	600	2600
ISIC 2018	10015	1512	11527
ISIC 2019	25331	8238	33569
ISIC 2020	33126	10982	44108

Table 3.1: Summary of the ISIC 2016 - 2020 datasets. (Cassidy et al. 2022)

were judged unnecessary because they serve largely as masks for the ISIC segmentation challenge and are unrelated to classification tasks.In particular, 2,000 superpixel images were deleted from the training set, whereas 600 were removed from the test set.

Following that, a thorough duplicate removal strategy was implemented. This involved looking for binary duplicates across all training and test sets. The analysis identified 12,039 binary identical duplicate photos in training sets and 1,592 in test sets.

Notably, the 2019 training set contained 2,074 downsampled image files. Given their distinct resolution, these images were not subject to a comprehensive binary identicality inspection. However, further analysis involving the ISIC code within the filenames led to the identification of 2,263 duplicate image files within the downsampled set. Also, while evaluating both within and across sets, a total of 13,967 binary identical duplicate photos were found and eliminated.

In total, 14,310 duplicate picture files were successfully removed from all training sets, yielding an effectively curated dataset of 56,987 image files. The training set in this dataset included 45,590 photos, including 3,924 melanoma images and 41,666 non-melanoma images. Similarly, the validation set included 11,397 photos, 981 of which were melanoma images and 10,416 of which were not. Importantly, the classes in this dataset were harmonized to produce the curated balanced dataset, ensuring that all skin lesion classifications are represented fairly. The class distribution of the images in the two datasets is shown in Table 3.1.

Set	Curated		Curated Balanced	
	Melanoma	Non-Melanoma	Melanoma	Non-Melanoma
Training	3,924	41,666	3,924	3,924
Validation	981	10,416	981	981
Total	56,987		9,810	

Table 3.2: Summary of the dataset counts(Cassidy et al. 2022)

To ensure the dataset's integrity, a range of similarity algorithms was employed to check for any potential duplicates within it. Interestingly, no true positives were discovered by any of the algorithms, indicating that the dataset's cleanliness and reliability have been confirmed.

The curated balanced dataset was chosen as the main dataset for this study because, the dataset contains a sizable number of images and also offers a wide range of instances, with a focus on images with hair artifacts. This diversity enables researchers to thoroughly test methodologies and critically assess findings. The dataset also stands out for having a balanced distribution of classes, which considerably lowers bias in any deep learning models created for this study.

The curated balanced dataset, which represents careful curation, class balance, and a variety of images to develop strong and objective skin lesion classification, essentially acts as the cornerstone of this research. Its use in this work advances the accuracy of skin lesion diagnosis through machine learning and provides a useful resource for both the ongoing study and the larger field of medical image analysis.

Artifacts in dataset

The curate balanced dataset contains various artifacts that can be affected by the quality and interpretation of skin lesion images. These artifacts are important to consider when working with medical image data, particularly for tasks like skin lesion classification. Notable artifacts have been identified in the dataset:

Cropped Lesions: In some images, heavily cropped lesions are shown, where large portions of the lesion or normal skin boundary regions have been removed. The context and completeness of the lesion representation can be impacted.

Dermascope Measurement Overlays: In certain images, dermascope measurement overlays are exhibited, which can obscure parts of the lesion or its boundary. These overlays may contain measurements or annotations that interfere with image analysis.

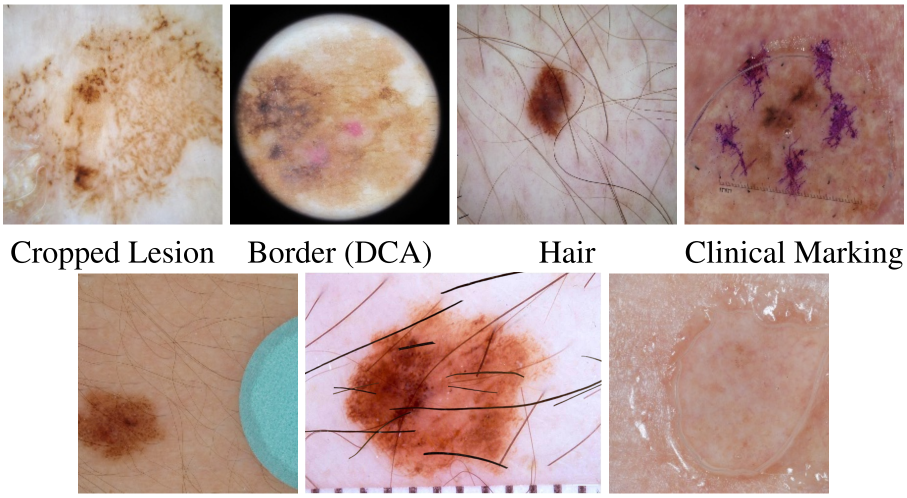
Presence of Hair: Hair artifacts are frequently observed in some images, introducing unwanted noise and complexity to the dataset. Parts of the lesion may be covered by hair, making it challenging to discern important features.

Clinical Markings: In some images, clinical markings around the lesion are present. These markings, while useful for medical practitioners, can introduce additional information that may not be deemed relevant to a machine learning model focused on lesion classification.

Size Reference Stickers: Size reference stickers placed close to the lesion are visible in certain images. These stickers are used for size calibration but can be distracting for image analysis algorithms.

Rulers: In some images, rulers are placed near the lesion. These rulers serve as measurement references but can disrupt the image's visual composition.

Immersion Fluid Artifacts: Images may show pockets of air or other artifacts due to the application of immersion fluid during dermoscopic examinations. These artifacts can cause distortions in the appearance of lesions.



Cropped Lesion Border (DCA) Hair Clinical Marking

Size ref sticker Ruler Fluid Artifact

Figure 3.1: Examples for Skin Image Artifacts

These artifacts underscore the importance of data preprocessing and quality control when working with medical image datasets. These issues must be addressed to improve the reliability and generalizability of machine learning models trained on such data. Consideration of these artifacts during dataset curation and analysis is essential to ensure accurate and clinically meaningful results in tasks like skin lesion classification.

It can be observed from the table 3.3 that hair artifacts are the most common category among the various artifact categories found in dermatology images. They are prominently encountered in both the training and validation datasets, appearing in a substantial number of images across different classes. The significance of hair removal techniques is highlighted in this context. The presence of hair artifacts can result in the introduction of noise and unwanted features into images, potentially causing misclassification or a reduction in accuracy in dermatological image analysis tasks. Therefore, the implementation of effective hair removal methods is deemed crucial to enhance the quality of the data used for classification and diagnosis. The removal of hair artifacts

Artifact Category	Subset				Artifact Totals
	Train Mel	Train Oth	Val Mel	Val Oth	
Borders	1721	663	417	179	2980
Hair	2224	2595	560	617	5996
Measurement Device	962	749	202	183	2096
Air Pockets	1129	637	442	142	2350
Clinical Markings	124	90	29	20	263
Other	100	55	55	18	228
No Artifacts	377	616	57	172	122

Table 3.3: Distribution of artifacts in curated balanced dataset

can lead to the creation of cleaner and more accurate image datasets, ultimately resulting in improved model performance and increased reliability in dermatological image analysis.

3.1.2 Dataset for Testing

The test dataset used in this study is the skin hair dataset developed by AGH University of Science and Technology, derived from the ISIC dataset and has been purposefully crafted to explore the impact of hair artifacts in dermoscopic image analysis. This dataset comprises dermoscopic images, some with artificially incorporated hair patterns and corresponding binary masks. It's segmented into two subsets: one with images containing hair and the other without. To create images with hair, the process involves selecting a clean image from ISIC, choosing an image with hair, manually marking hair areas using Photoshop, extracting the hair, and integrating it into the dermoscopic image. To diversify hair patterns, augmentation techniques like random movement, rotation, and color changes were applied. In total, 252 images were generated, each with 84 unique masks, covering different hair types, densities, and colors (light, brown, and dark). This dataset provides valuable insights into the challenges posed by hair artifacts in dermatological image analysis.

Since both the train dataset and test dataset are derived from ISIC datasets, the process of manual inspection was carried out to identify and subsequently remove images that were common between the train and test datasets. This action was taken to prevent the possibility of data leakage and overfitting, with the aim of ensuring that the model's performance evaluation would be conducted on a truly independent set of images. The identified images, which were found to be shared between both sets, were systematically eliminated from the train dataset.

3.2 Research Design

In this study, a research design utilizing supervised deep learning is employed, with a specific focus on the training and evaluation of a ResNet50 (Residual Neural Network) model for melanoma classification and the assessment of the impact of hair in melanoma classification. This design was chosen to harness the capabilities of deep learning for addressing the complex tasks of skin lesion identification and the determination of whether hair removal methods should be employed or not.

The selection of this research design aligns with our research objectives for several reasons:

3.2.1 Skin Lesion Classification

Our primary objective is to develop a robust classification model capable of distinguishing between melanoma and non-melanoma skin lesions. Melanoma is a highly malignant form of skin cancer, and early and accurate detection is crucial for timely diagnosis. Deep learning models, such as ResNet50, have demonstrated remarkable capabilities in image classification tasks, making them well-suited for the challenging task of melanoma identification.

3.2.2 Hair Artifact Analysis

While melanoma classification remains the primary focus, the effect of hair artifacts in skin images is acknowledged. Existing literature highlights that critical features can be obscured, and diagnostic accuracy can potentially be affected by the presence of hair artifacts. Recognizing the significance of this factor, an evaluation of its impact on the classification process is undertaken.

The study intends to examine the effects of different hair artifact densities on the ResNet50 model's ability to classify melanoma. Instead of relying exclusively on conventional measurements, inference analysis and graphical representations are used to acquire a better understanding of how the model responds to various degrees of hair artifacts. The method enables an understanding of the model's sensitivity to hair artifacts and its capability to accurately categorize skin lesions under diverse circumstances.

This investigation seeks to shed light on how hair artifacts affect the classification of dermatological images, offering useful information to both researchers and doctors. It is important to take into account both the existence and the degree of hair artifacts in order to maintain diagnostic accuracy in a variety of situations.

3.2.3 Transfer Learning

The ResNet50 architecture initialized with pre-trained weights on a large dataset is leveraged. This choice harnesses the feature extraction capabilities of a pre-trained model and allows us to adapt it to our specific skin lesion classification tasks. Transfer learning reduces the need for extensive data collection and training time while potentially improving model generalization.

By integrating these concepts, this design aligns seamlessly with our research objectives, aiming to comprehend the necessity of a hair removal mechanism in the context of melanoma classification, thereby enhancing the accuracy and reliability of our classification model for dermatologists and clinicians alike.

3.3 Data Preprocessing

The data preprocessing phase plays a crucial role in preparing the collected data for the effective training and evaluation of the ResNet model. Several essential steps were integrated into the data preprocessing pipeline. First, all skin lesion images underwent resizing, ensuring a consistent dimension of 32x32 pixels across the dataset. This uniformity simplifies the model training process.

Next, pixel values within the images were normalized on a channel-wise basis, using specific mean and standard deviation values for each channel. These values, namely Mean values of [-1.3048, -1.2044, -0.9766] and Standard deviations of [0.6709, 0.6858, 0.6829], were employed to standardize the data and mitigate issues related to varying image intensities.

Ultimately, this comprehensive data preprocessing pipeline results in a clean, standardized dataset, enhancing the ResNet50 model's performance, robustness, and ability to generalize effectively during both training and evaluation phases.

3.4 Model Architecture

ResNet (Residual Network) is a deep convolutional neural network architecture that was first introduced in 2015 and is renowned for its efficiency in deep model training. For learning residual mappings and enhancing gradient flow, it introduces skip connections.

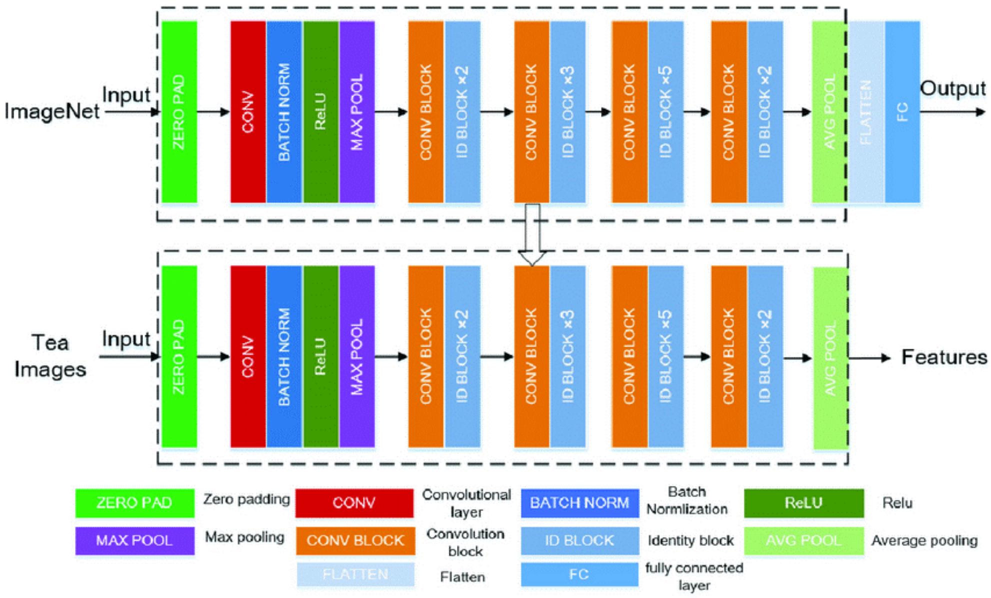


Figure 3.2: ResNet Architecture

The ResNet model was selected for the task due to several compelling reasons:

Deep Network Advantages: ResNet's deep architecture is well-known for its ability to learn complicated image attributes. The introduction of skip connections (residual connections) to mitigate the vanishing gradient problem allows the model to achieve large depth while preserving training stability.

Notable Performance: ResNet architectures have continuously demonstrated remarkable performance across a wide range of computer vision applications, including image categorization. ResNet-50, in particular, strikes a favorable balance between model depth and computational efficiency, making it a popular choice.

Transfer Learning Utility: When pretrained on large datasets, models like ResNet perform excellently as feature extractors. Fine-tuning a pretrained ResNet model can significantly improve performance on specific tasks, eliminating the need for initial training.

The residual block with shortcut connections, which is depicted in the image below, is the fundamental building block of a ResNet architecture.

The model begins with convolutional layers, stacks these building blocks and, concludes with global average pooling and a fully connected layer. Due to its deep structure and skip connections, ResNet has excelled in computer vision tasks.

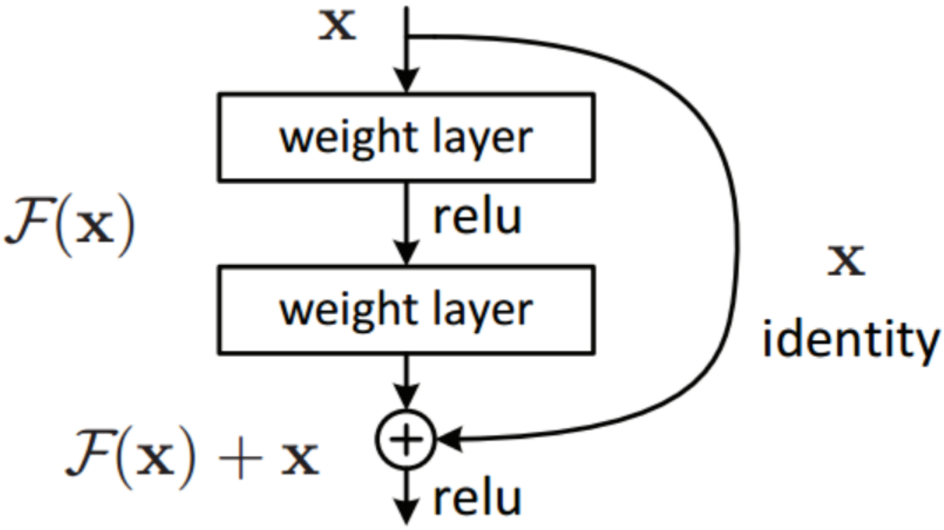


Figure 3.3: Residual Block of ResNet

3.4.1 ResNet-50 Architecture

The ResNet-50 architecture underpins this model. It replaces the pre-trained ResNet-50 model's final fully connected layer with a new fully connected layer with the same number of output features as the number of classes in the dataset. The model can now be fine-tuned for a specific classification task thanks to this modification.

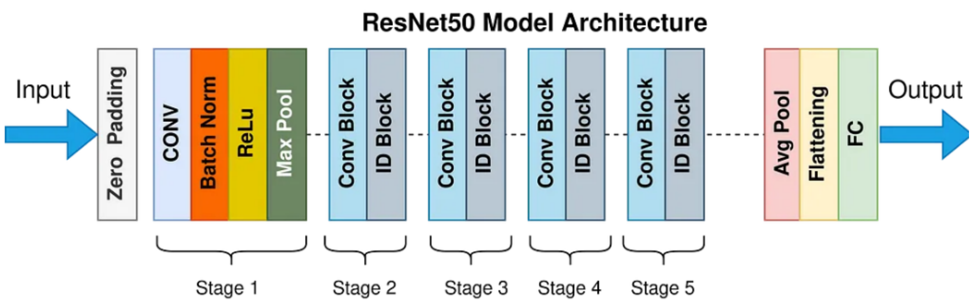


Figure 3.4: ResNet50 Architecture

The ResNet model is built with an initial convolutional layer, followed by batch normalisation, ReLU activation, and max pooling. Then it has four layers, each with a different number of residual blocks, as specified in the ResNetConfig. The residual blocks are made up of convolutional layers, batch normalisation, and skip connections, which allow the model to learn residual functions.

The AdaptiveAvgPool2d layer is used to perform adaptive average pooling, which ensures that the model can handle input images of varying sizes. Finally, the output is passed through the fully connected layer to obtain the predicted class probabilities.

The pre-trained ResNet-50 model is used in this model design, which also initializes the fully connected layer and loads the pre-trained weights. The network can benefit from the features learned from a large dataset by using a pre-trained model, which aids in better feature extraction and classification performance on a specific task.

3.4.2 Motivation Behind Model Design

The ResNet (Residual Network) model was created to address the issue of vanishing gradients and the degradation of accuracy as network depth increases. The key concept is to use skip connections, also known as "shortcut connections," to allow information to flow directly from earlier layers to later layers. These skip connections create shortcut paths for the gradient during backpropagation, making deep network training easier.

ResNet models enable the network to learn residual functions rather than directly fitting the desired underlying mapping by introducing residual blocks, which consist of identity mappings and a residual mapping. This method allows the model to concentrate on learning residual details, which are often easier to optimise, while identity mappings provide a fallback option for the model to achieve the desired accuracy.

The use of bottleneck layers in the design of ResNet models reduces the computational complexity of deep networks by using 1x1 convolutions to reduce and then increase the number of channels. This design aids in increasing efficiency while maintaining accuracy.

3.5 Training Process

During the training process, a total of 20 epochs were conducted, indicating the completion of one full pass through the entire training dataset. A batch size of 32 was employed, allowing 32 images to be processed simultaneously in each training iteration. The initial learning rate was set at 1e-3 (0.001), and a learning rate scheduler, specifically the OneCycleLR scheduler, was implemented to dynamically adjust the learning rate during training.

For the classification task, the Cross-Entropy Loss function was utilized to quantify the dissimilarity between predicted class probabilities and actual class labels, which is particularly suitable for multi-class classification problems.

Additionally, pretrained weights from a ResNet-50 model pretrained on the ImageNet dataset were used as a form of regularization, providing the model with valuable prior knowledge. Batch normalization layers were also incorporated into the ResNet architecture to standardize activations in each layer, contributing to the stability and efficiency of training.

3.6 Testing and Evaluation

In the testing phase, a distinct dataset containing skin images with various densities of hair was used. Specifically, some particular images with varying hair density were tested using three copies of each image. The trained ResNet50 model was then subjected to this test dataset, and model inference was performed on each image, generating predictions for melanoma classification. To ensure accuracy, certain standards were followed during this examination.

3.6.1 Model Inference

In this study, the application of model inference and probability distribution analysis is employed to comprehensively assess the impact of hair artifacts on skin cancer diagnosis. A distinct dataset containing skin images with varying densities of hair is used, and the trained neural network is subjected to model inference, providing valuable insights into the model's performance and its ability to mitigate the influence of hair artifacts.

3.6.2 Hair Removal

The Telea's inpainting algorithm proposed by (Telea 2004) based on Navier-Stokes equations is implemented as hair removal technique to understand the importance of artifact removal in classification tasks. In this method, a hair mask, which identifies regions containing hair artifacts, is used as a guide to inpaint and fill in these areas in the original image. The inpainting process takes into account the surrounding pixel values and gradients to create a seamless blend between inpainted and non-inpainted regions, effectively removing hair artifacts while preserving image details.

Peak Signal-to-Noise Ratio (PSNR) values is used for evaluating the effectiveness of hair removal using an inpainting method in different examples. PSNR is a metric commonly used to measure the quality of image processing tasks, such as denoising, compression, or inpainting. Higher PSNR values generally indicate better image quality.

3.7 Comparative Analysis

A comparative analysis is conducted following the hair artifacts removal. This step is crucial in addressing the research question regarding whether the model's classification accuracy is affected by the presence and density of hair artifacts.

3.8 Software and Tools

The comprehensive suite of software, libraries, and frameworks leveraged for the implementation of the methodology facilitated various research stages. The primary programming language employed was Python (Version 3.10.12), serving as the foundation for ResNet model development, data preprocessing, and experimental procedures. PyTorch (Version 2.0.1+cu118), a widely adopted deep learning framework, was utilized for constructing, training, and evaluating the ResNet model, offering a flexible and efficient platform for convolutional neural network development.

Additionally, TensorFlow (Version 2.13.0) and Keras (Version 2.13.1) were imported to support experimentation and neural network development. Data handling and preprocessing relied on NumPy (Version 1.23.5) for efficient numerical operations and Pandas (Version 1.5.3) for data manipulation and organization. TorchVision (Version 0.15.2+cu118) facilitated data loading and management, while Matplotlib (Version 3.7.1) played a crucial role in generating visualizations. Scikit-Learn (Version 1.2.2) contributed to machine learning tasks and metrics evaluation, and GPU hardware, such as NVIDIA CUDA, expedited model training and inference.

Google Colab provided an interactive environment for code development and experimentation. This collection of software tools and libraries collectively enabled the effective implementation of the methodology, emphasizing transparency and reproducibility in research endeavors.

3.9 Ethical Considerations

In this research, a comprehensive approach was taken to address several ethical considerations related to data collection and model usage, ensuring the highest ethical standards were maintained. The skin lesion dataset was sourced from publicly available and ethical sources, with a strong emphasis on prioritizing patient privacy and data protection, aligning with relevant ethical guidelines. The model's usage was guided by principles of fairness, responsibility, and bias mitigation, with ongoing evaluations conducted to monitor ethical implications. Robust data security measures were implemented to safeguard data confidentiality and integrity, and diligent data handling practices were observed, strictly for research purposes. The research's unwavering commitment to adhering to established ethical guidelines and principles underscores its ethical utilization of data and AI models, both in execution and beyond.

Chapter 4

Implementation

This section gives a detailed analysis of the studies designed to examine the impact of hair artifacts on the classification of melanoma and evaluate the effectiveness of hair removal methods. Results from notable cases are shown in the presentation; each Experiment consists of three images showing various hair artifact densities.

High Hair Density: This instance represents a scenario where hair artifacts is pronounced and substantial.

Moderate Hair Density: In this variation, the density of hair patterns was intentionally reduced while retaining a discernible level of interference.

Low Hair Density: Here, the density of hair interference was deliberately minimized, approaching a negligible level.

Each set also has an image that corresponds to it that shows how the hair pattern removal worked.

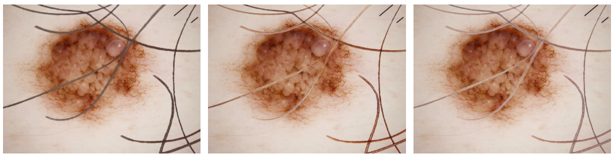
4.1 Model Inference

A pre-trained melanoma classification model is used for inference. The model had been trained on the curated dataset to classify skin lesions as either melanoma or other. The following steps are performed for each image in the experiment:

- Applied the test transformation to the original and hair-removed images.
- Converted the preprocessed images into PyTorch tensors.
- Performed inference using the pre-trained model to predict the class labels and class probabilities.

4.1.1 Experiment 1

For the first experiment, a non-melanoma image was chosen as the test image for analysis.



High Hair Density | Moderate Hair Density | Low Hair Density

Figure 4.1: Skin images with different hair densities - Experiment 1

In Figure 4.1,three different hair density variations of a non-melanoma image are depicted.

```
1 predicted_labels
```
```
[0, 0, 0]
```
```
1 predicted_probs
```
```
[array([0.71746576, 0.2825342 ], dtype=float32),
 array([0.65511394, 0.34488612], dtype=float32),
 array([0.54104394, 0.4589561 ], dtype=float32)]
```

Figure 4.2: Predicted Probability - Experiment 1

For the first image, which likely has high hair density, the model assigns a relatively higher probability to melanoma with a probability distribution of (0.7175, 0.2825). This suggests a misclassification with a higher confidence in melanoma.

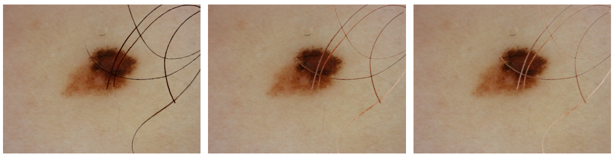
In the case of the image with moderate hair density, the model still predicts melanoma with a relatively high probability, as seen in the distribution of (0.6551, 0.3449).

Even for the image with low hair density, the model continues to predict melanoma, with a distribution of (0.5410, 0.4590).

Despite being classified as melanoma, there is a diminishing trend in the model's probability of melanoma classification as the hair density decreases.

4.1.2 Experiment 2

An image from the melanoma class was chosen as the analysis' input for the second experiment.



High Hair Density | Moderate Hair Density | Low Hair Density

Figure 4.3: Skin images with different hair densities - Experiment 2

In Figure 4.3,three different hair density variations of a melanoma image are depicted.

```
1 predicted_labels
```

```
[1, 1, 1]
```

```
1 predicted_probs
```

```
[array([0.4211048, 0.5788952], dtype=float32),
 array([0.34048298, 0.659517  ], dtype=float32),
 array([0.32342368, 0.6765763 ], dtype=float32)]
```

Figure 4.4: Predicted Probability - Experiment 2

For the given melanoma input, the model consistently predicts non-melanoma for all three images, despite the actual melanoma nature of the input. Notably, the model's probability of melanoma is influenced by the density of hair in the image:

In images with high hair density, the probability for melanoma is 0.42110. This suggests that higher hair density leads to a stronger confidence in melanoma classification, even though it is an incorrect prediction.

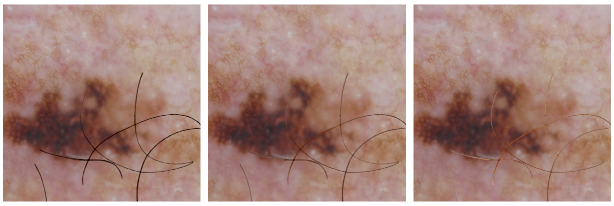
Similarly, for images with moderate hair density, the probability for melanoma is decreased to 0.34048 . That means, the presence of moderate hair density leads to a lower likelihood of melanoma.

Even when presented with images featuring low hair density, the probability for melanoma again decreased to 0.3234. This demonstrates that in the absence of hair or with lower hair density, the model tends to decline the probability of melanaoma.

Despite the initial classification as non-melanoma, as observed in the previous Experiment, there is a noticeable decrease in the model's probability of melanoma classification as the hair density diminishes

4.1.3 Experiment 3

For the third experiment, a non-melanoma image was utilized as the test image for the analysis.



High Hair Density | Moderate Hair Density | Low Hair Density

Figure 4.5: Skin images with different hair densities - Experiment 3

Figure 4.5,illustrates three distinct variations of hair density in a non-melanoma image.

```
1 predicted_labels
```

```
[1, 1, 1]
```

```
1 predicted_probs
```

```
[array([0.4019897 , 0.59801024], dtype=float32),
 array([0.3980556, 0.6019444], dtype=float32),
 array([0.39325514, 0.6067448 ], dtype=float32)]
```

Figure 4.6: Predicted Probability - Experiment 3

For the given non-melanoma images, the model consistently and correctly predicts non-melanoma for all three images. Interestingly, even in these cases where the correct classification is non-melanoma, we can observe a subtle relationship between the probability for melanoma and the density of hair:

In images with high hair density, the model assigns a probability of melanoma (0.40199) that is relatively closer to the non-melanoma probability (0.5980). This suggests that higher hair density leads to an increase in the model's uncertainty.

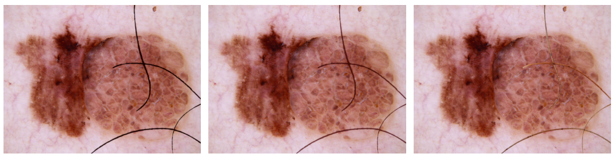
Similarly, for images with moderate hair density, the model exhibits a similar pattern of assigning a melanoma (0.3981) probability that is relatively close to the non-melanoma probability (0.6019). This indicates a consistent trend where moderate hair density introduces a slightly lower degree of uncertainty in the model's decision.

Even in images with low hair density, the model continues to predict non-melanoma correctly, with probabilities showing a similar pattern of decrease in melanoma probability (0.3933) but still leaning towards non-melanoma (0.6067).

Even when correctly classifying non-melanoma images, the model's probability for melanoma exhibits a decreasing trend based on the density of hair. This suggests that higher hair density may introduce uncertainty in classification.

4.1.4 Experiment 4

In the fourth experiment, an image from the melanoma class was selected as the input for the analysis



High Hair Density | Moderate Hair Density | Low Hair Density

Figure 4.7: Skin images with different hair densities - Experiment 4

Figure 4.7, illustrates three distinct hair density variations in a melanoma image.

```
1 predicted_labels
```

```
[0, 0, 0]
```

```
1 predicted_probs
```

```
[array([0.88196963, 0.11803041], dtype=float32),
 array([0.8689467 , 0.13105331], dtype=float32),
 array([0.8450331 , 0.15496695], dtype=float32)]
```

Figure 4.8: Predicted Probability - Experiment 4

For the given melanoma images, the model consistently and correctly predicts melanoma for all three images. Interestingly, even in these cases where the correct classification is melanoma, a relationship can be observed between the probability for melanoma and the density of hair.

In images with high hair density, the model assigns a probability of melanoma (0.8819) with a relatively higher value. This suggests that higher hair density is associated with a higher likelihood of melanoma classification.

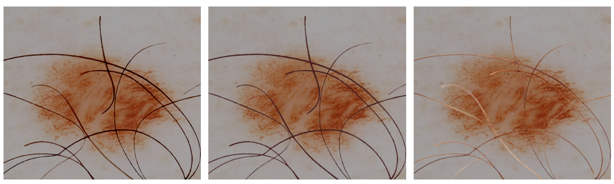
Similarly, for images with moderate hair density, the model maintains a correct melanoma prediction (0.8689) while exhibiting a slightly lower probability for melanoma compared to high hair density images. This indicates a trend where moderate hair density introduces a slight decrease in the probability of melanoma but remains below the threshold for melanoma classification.

Even in images with low hair density, the model continues to predict melanoma (0.8450) correctly, with probabilities showing a further decrease in the probability for melanoma. However, the probability for melanoma in low hair density images remains below the threshold for melanoma classification.

In this experiment, the model's probability for melanoma exhibits a relationship with the density of hair. Higher hair density is associated with a higher likelihood of melanoma classification, while lower hair density leads to a gradual decrease in the probability for melanoma, although it remains below the classification threshold in all cases.

4.1.5 Experiment 5

In the fifth experiment, testing was performed using a non-melanoma image as the input. Notably, this experiment produced varying predictions that were dependent on the density of hair.



High Hair Density | Moderate Hair Density | Low Hair Density

Figure 4.9: Skin images with different hair densities - Experiment 5

Figure 4.9, illustrates three distinct hair density variations in a non-melanoma image.

```
predicted_labels
```

```
[0, 0, 1]
```

```
predicted_probs
```

```
[array([0.5829259, 0.4170741], dtype=float32),
 array([0.534852  , 0.46514797], dtype=float32),
 array([0.47780868, 0.52219135], dtype=float32)]
```

Figure 4.10: Predicted Probability - Experiment 5

For the three images tested, the model correctly predicted melanoma for the third image, while predicting non-melanoma for the first two images. Examining the class probabilities in relation to hair density reveals the following trends:

In the presence of high hair density, the model assigns a relatively higher probability of melanoma classification for the first image, with a distribution of [0.5829, 0.4171].

Similarly, for the image with moderate hair density, the model maintains probability of melanoma, reflected in a distribution of [0.5349, 0.4651].

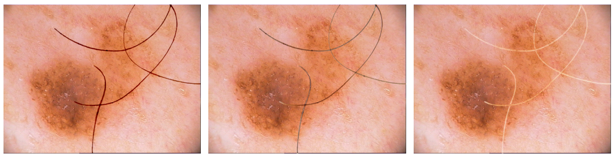
Conversely, when presented with an image featuring low hair density, the model exhibits lowest probability of melanoma classification, observed in a distribution of [0.4778, 0.5222].

These findings underscore a consistent trend: as the density of hair decreases, there is a corresponding decrease in the model's probability of melanoma classification.

4.1.6 Experiment 6

Experiment six involved using a non-melanoma image as the input for testing. Notably, this experiment yielded differing predictions, and these predictions were influenced by the level of hair density in the images.

Figure 4.11, illustrates three distinct hair density variations in a non-melanoma image.



High Hair Density | Moderate Hair Density | Low Hair Density

Figure 4.11: Skin images with different hair densities - Experiment 6

```
1 predicted_labels
```

```
[0, 1, 1]
```

```
1 predicted_probs
```

```
[array([0.5114559 , 0.48854405], dtype=float32),
 array([0.4892782 , 0.51072174], dtype=float32),
 array([0.47504038, 0.5249596 ], dtype=float32)]
```

Figure 4.12: Predicted Probability - Experiment 6

For the given non-melanoma images, the model generates varying predictions, indicating a level of uncertainty in the classification. Let's examine how the probability for melanoma relates to the density of hair:

In the first image, which likely has high hair density, the model assigns a probability of melanoma (0.5115) that is close to non-melanoma (0.4885). This suggests a relatively balanced probability distribution, indicating uncertainty in the classification. The presence of high hair density appears to introduce this uncertainty.

In the case of the second image with moderate hair density, the model predicts correctly by assigning a probability for melanoma (0.4893) that is lower than non-melanoma (0.5107). This indicates a leaning toward non-melanoma classification.

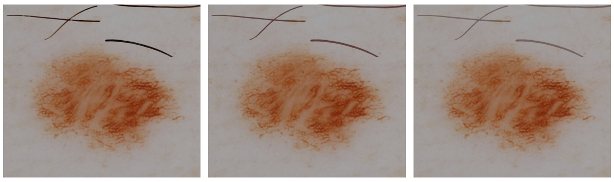
For the third image with low hair density, the model once again predicts non-melanoma. Notably, the probability for melanoma (0.4750) is lower than the probability for non-melanoma (0.5249), reinforcing the consistent trend that as the density of hair decreases, the probability of melanoma classification also decreases. This observation aligns with the concept that lower hair density levels are associated with a diminishing likelihood of the model classifying an image as melanoma.

4.2 Hair Artifact Removal using Telea

The Telea's inpainting algorithm is implemented to understand the trend in probability.

4.2.1 Hair Removal - Example 1

A non-melanoma image is selected for the hair removal task in this first example.



High Hair Density	Moderate Hair Density	Low Hair Density

Figure 4.13: Skin images with different hair densities - Hair Removal Example 1

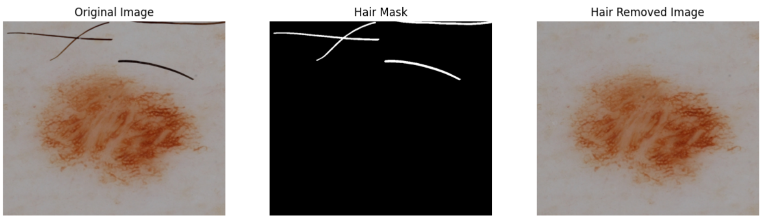


Figure 4.14: Hair Removal Example 1

```
predicted_labels
```

```
[0, 0, 0, 0]
```

```
predicted_probs
```

```
[array([0.63281786, 0.3671822 ], dtype=float32),
 array([0.60226107, 0.39773896], dtype=float32),
 array([0.5738065 , 0.42619348], dtype=float32),
 array([0.51289654, 0.4871034 ], dtype=float32)]
```

Figure 4.15: Predicted Probabilities - Example 1

Despite displaying an incorrect prediction, a subsequent decrease in the probability for melanoma can be observed in figure 4.15.

4.2.2 Hair Removal - Example 2

Also in this second example, a non-melanoma image is chosen for the hair removal task.

Figure 4.16: Skin images with different hair densities - Hair Removal Example 2

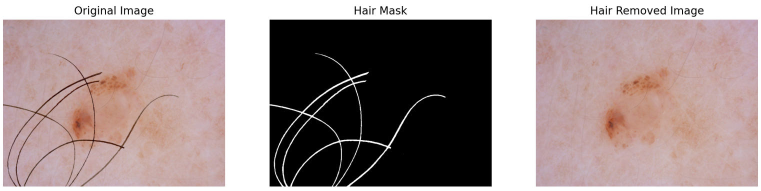


Figure 4.17: Hair Removal Example 2

```
1 predicted_labels
```

```
[1, 1, 1, 1]
```

```
1 predicted_probs
```

```
[array([0.3553163, 0.6446837], dtype=float32),
 array([0.36063948, 0.6393605 ], dtype=float32),
 array([0.3443188, 0.6556812], dtype=float32),
 array([0.32935634, 0.6706436 ], dtype=float32)]
```

Figure 4.18: Predicted Probabilities - Example 2

In this example, correct predictions are made by the model, and a consistent decrease in the probability for melanoma can be observed in the figure 4.18.

4.2.3 Hair Removal - Example 3

In this third case too, a non-melanoma image is chosen for the hair removal task.

High Hair Density | Moderate Hair Density | Low Hair Density

Figure 4.19: Skin images with different hair densities - Hair Removal Example 3

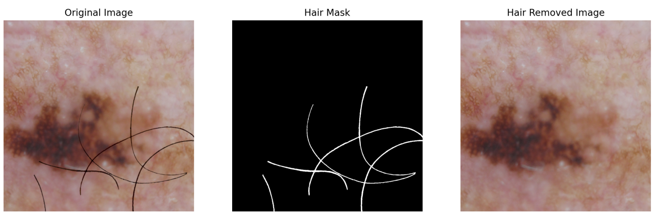


Figure 4.20: Hair Removal Example 3

```
1 predicted_labels
```

```
[1, 1, 1, 1]
```

```
1 predicted_probs
```

```
[array([0.4019897 , 0.59801024], dtype=float32),
 array([0.3980556, 0.6019444], dtype=float32),
 array([0.39325514, 0.6067448 ], dtype=float32),
 array([0.38712516, 0.61287487], dtype=float32)]
```

Figure 4.21: Predicted Probabilities - Example 3

The model makes valid predictions, as in Example 2. The figure 4.21 depicts a constant decrease in the probability of melanoma in this example.

After applying the hair removal process, the inference for melanoma classification was performed on the hair-removed image. Interestingly, in some cases, the probability for melanoma decreased following the removal of hair artifacts. This phenomenon can be attributed to the fact that hair artifacts may introduce noise or unwanted features into the image, leading to misinterpretations by the classification model. By removing these artifacts, the model focuses more on the genuine skin characteristics related to melanoma, resulting in a potentially more accurate classification. This observation underscores the importance of artifact removal techniques in improving the reliability of skin cancer classification models when dealing with images that contain various artifacts like hair.

In summary, these experiments consistently demonstrated a relationship between the density of hair artifacts and the model's probability of melanoma classification. Higher hair density often led to higher probabilities of melanoma, even in cases of incorrect classification. Conversely, lower hair density was associated with decreased probabilities of melanoma, indicating increased uncertainty in the classification. These findings highlight the importance of considering hair artifacts in melanoma classification and the potential need for improved hair removal techniques in dermatology imaging.

Chapter 5

Evaluation

The primary objective of the experiments conducted in this study was to investigate the influence of hair artifacts on the classification of skin lesions, with a specific focus on melanoma. Melanoma, a type of skin cancer, is a condition in which early and accurate diagnosis is deemed crucial for effective treatment. Challenges posed by the presence of hair artifacts in dermatology images to the accurate classification of skin lesions, potentially leading to misdiagnosis or increased uncertainty, were examined.

An assessment was made regarding how the density of hair artifacts impacts the classification of skin lesions, with particular emphasis on the differentiation between melanoma and non-melanoma cases.

The significance of evaluating the impact of hair artifacts on melanoma classification cannot be understated. In clinical settings, imaging techniques are relied upon by dermatologists to aid in the diagnosis of skin lesions, including melanoma. However, the presence of hair can obscure important features of these lesions, potentially leading to misinterpretation. Therefore, understanding the effect of hair artifacts is considered essential for the improvement of automated melanoma classification systems' accuracy, ultimately enhancing patient care.

In the following sections, the results of the experiments will be presented, the findings will be analyzed, and their implications for melanoma classification and the broader field of dermatology imaging will be discussed.

5.1 Recap of Experiments

In the implementation chapter, a series of experiments were conducted to assess the impact of hair artifacts on skin lesion classification, specifically focusing on melanoma.

These experiments utilized a dataset comprising non-melanoma and melanoma images with varying levels of hair density. The key observations from each experiment are summarized as follows:

Experiment 1

Input Type: Non-melanoma image.

Predicted Class Label: Melanoma

Observations: The model consistently predicted melanoma, with a diminishing trend in melanoma probability as hair density decreased.

Experiment 2

Input Type: Melanoma image.

Predicted Class Label: Non-Melanoma

Observations: The model consistently predicted non-melanoma, with higher hair density leading to a stronger confidence in melanoma classification, even though it was incorrect.

Experiment 3

Input Type: Non-melanoma image.

Predicted Class Label: Non-Melanoma

Observations: The model correctly predicted non-melanoma for all images. There was a subtle relationship between hair density and the probability for melanoma, with higher hair density introducing uncertainty.

Experiment 4

Input Type: Melanoma image.

Predicted Class Label: Melanoma

Observations: The model correctly predicted melanoma. Higher hair density was associated with a higher likelihood of melanoma classification, while lower hair density introduced uncertainty.

Experiment 5

Input Type: Non-melanoma image.

Predicted Class Labels: Varying

Observations: The model correctly predicted melanoma for two images and non-melanoma for one. The presence of high hair density increased the probability of melanoma classification.

Experiment 6

Input Type: Non-melanoma image.

Predicted Class Label: Varying

Observations: The model generated varying predictions for non-melanoma images, indicating uncertainty in classification. High hair density led to balanced probabilities.

These experiments collectively demonstrate a consistent trend: the density of hair artifacts in skin lesion images influences the model's probability of melanoma classification. Higher hair density often results in higher probabilities of melanoma, even when the classification is incorrect, while lower hair density is associated with decreased probabilities and increased uncertainty. These findings underscore the importance of considering and mitigating the impact of hair artifacts in dermatology imaging for accurate skin lesion classification.

In the following section, the evaluation of these results will be presented, providing a deeper understanding of the implications for melanoma classification and the broader field of dermatology imaging.

5.2 Evaluation Metrics

In this study, the evaluation process diverges from traditional evaluation metrics such as accuracy, F1-score, or the area under the ROC curve. Instead, an approach emphasizing inference and the examination of probability distributions generated by the model is employed for the investigation of the impact of hair artifacts on melanoma classification. Below, the methodology for evaluation is outlined:

Inference and Probability Examination: The primary evaluation method encompasses the following steps:

5.2.1 Model Inference

A pre-trained melanoma classification model is utilized to perform inference on a diverse set of skin lesion images. These images encompass both non-melanoma and melanoma cases, each displaying varying levels of hair density artifacts.

5.2.2 Probability Distribution Analysis

Rather than relying on traditional metrics, the focus is on the examination of probability distributions produced by the model for each image. Specifically, attention is directed towards the probabilities assigned to melanoma and non-melanoma classifications. This facilitates the discernment of patterns and trends related to hair density and its impact on the model's predictions.

5.2.3 Key Observations

Through this evaluation approach, the following key observations are made:

Influence of Hair density

It is observed that the density of hair artifacts within the images has a notable influence on the model's probability distribution. Higher hair density often results in a stronger probability of melanoma classification, even in cases where the classification is incorrect. Conversely, as the density of hair diminishes, there is a consistent decrease in the probability assigned to melanoma classification.

Impact on Classification

The presence of hair artifacts can affect the model's classification outcomes, thereby highlighting the necessity for robust techniques to mitigate this impact for accurate melanoma diagnosis.

While the study does not rely on traditional evaluation metrics, this approach offers valuable insights into the complex relationship between hair artifacts and melanoma classification. Subsequent sections will delve into the implications of these findings and their significance in the context of dermatology imaging and clinical practice.

5.3 Results Analysis

In this section, the results of each experiment are presented, encompassing the model's predictions and class probabilities. The analysis revolves around the influence of hair artifacts on melanoma classification, with an emphasis on understanding notable trends, variations in predictions contingent on hair density, instances of misclassifications, and uncertainty. Visual aids, including tables are employed to enhance the clarity of the results.

Experiment	**Hair Density**	**Original Label**	**Predicted Class Label**	**Predicted Class Probabilities**	
				Melanoma	**Others**
1	High	Others	Melanoma	0.7175	0.2825
	Moderate	Others	Melanoma	0.6551	0.3449
	Low	Others	Melanoma	0.541	0.4589
2	High	Melanoma	Others	0.4211	0.5788
	Moderate	Melanoma	Others	0.3405	0.6595
	Low	Melanoma	Others	0.3234	0.6765
3	High	Others	Others	0.4019	0.598
	Moderate	Others	Others	0.3981	0.6019
	Low	Others	Others	0.3933	0.6067
4	High	Melanoma	Melanoma	0.8819	0.118
	Moderate	Melanoma	Melanoma	0.8689	0.1311
	Low	Melanoma	Melanoma	0.845	0.1549
5	High	Others	Melanoma	0.5829	0.4171
	Moderate	Others	Melanoma	0.5348	0.4651
	Low	Others	Others	0.4778	0.5222
6	High	Others	Melanoma	0.5114	0.4885
	Moderate	Others	Others	0.4893	0.5107
	Low	Others	Others	0.475	0.5249

Table 5.1: Experiment Results

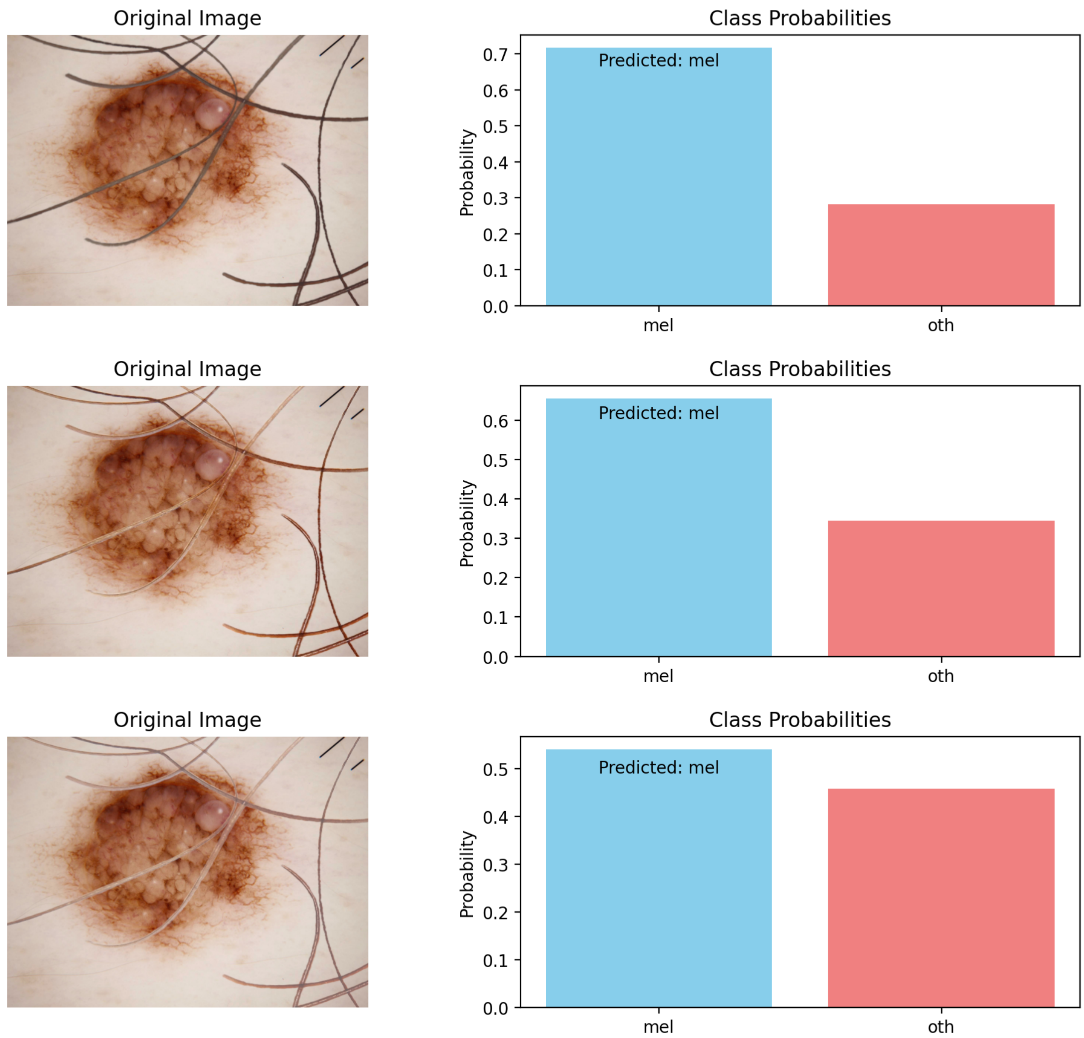


Figure 5.1: Class Probabilities - Experiment 1

The Figure 5.1, displays the class probabilities for each image in Experiment 1.

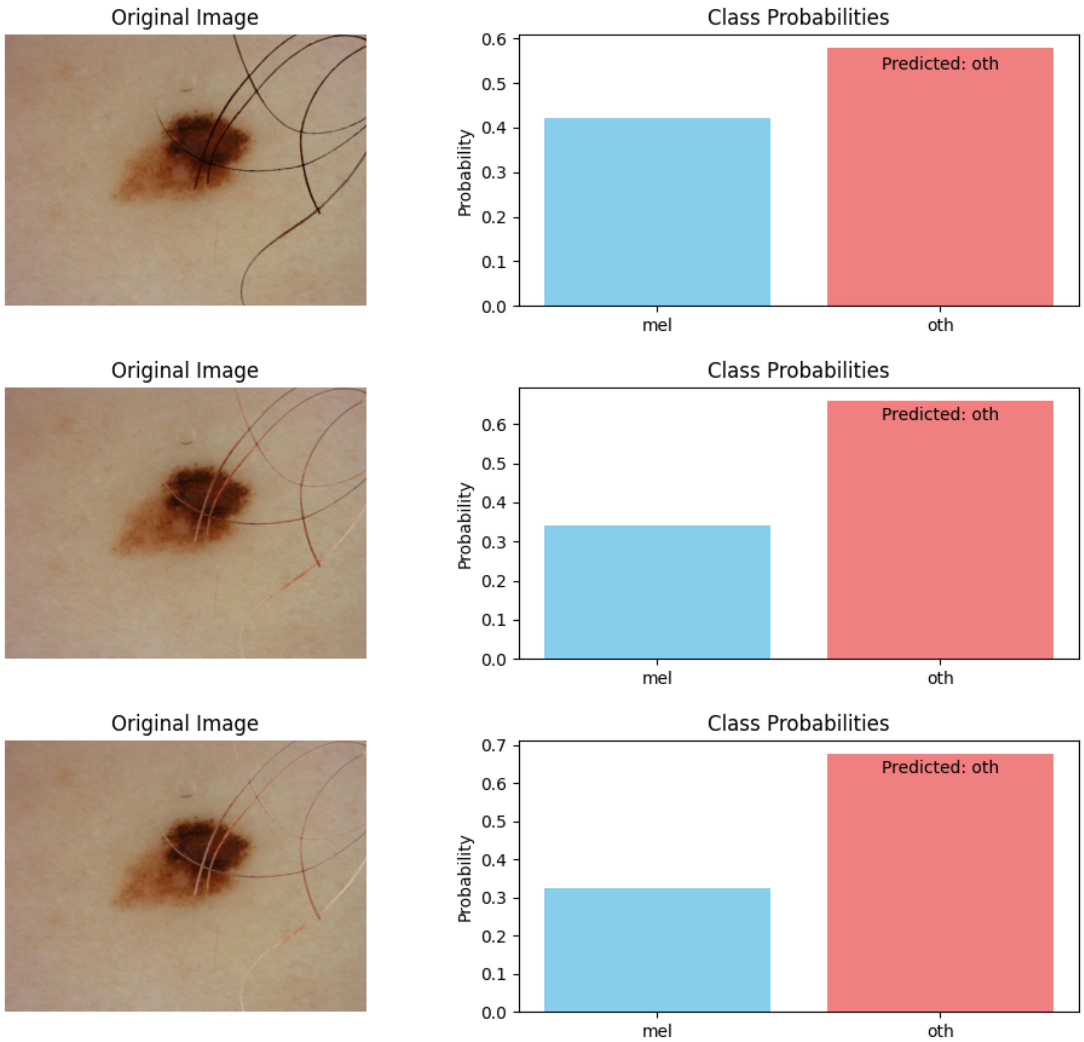


Figure 5.2: Class Probabilities - Experiment 2

The Figure 5.2, presents the class probabilities for each image in Experiment 2.

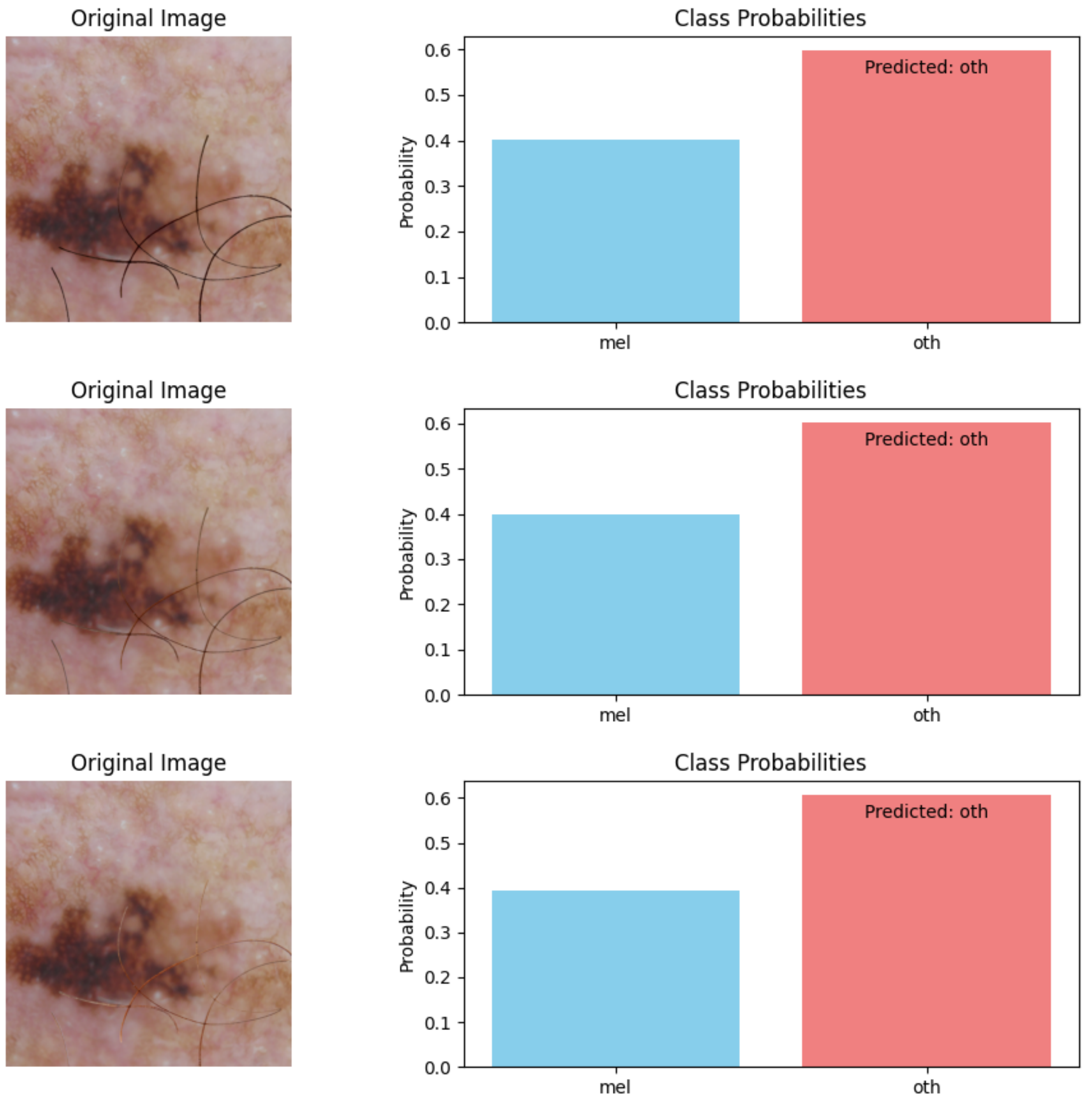


Figure 5.3: Class Probabilities - Experiment 3

Figure 5.3, showcases the class probabilities for each image in Experiment 3.

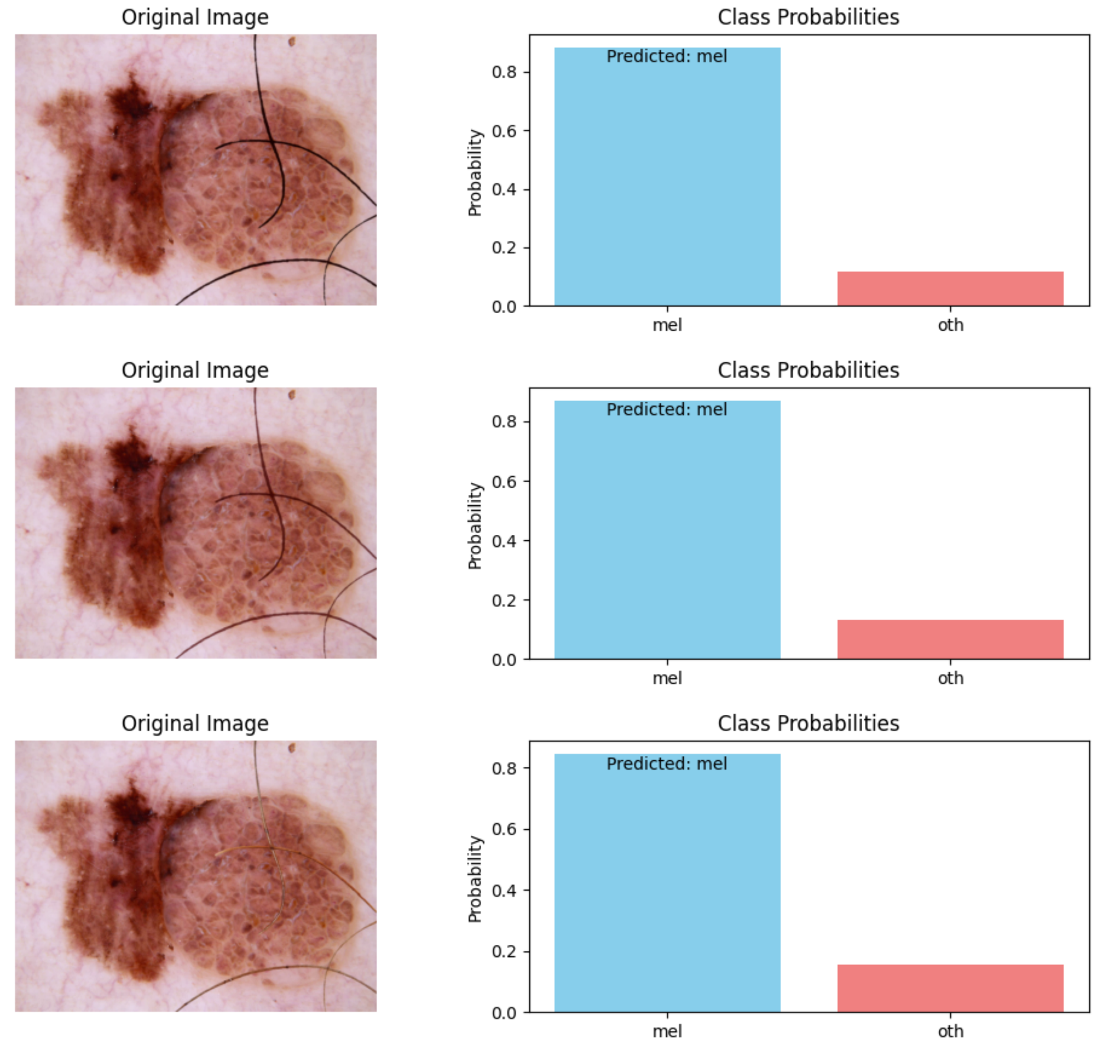


Figure 5.4: Class Probabilities - Experiment 4

Figure 5.4, displays class probabilities for each image in Experiment 4.

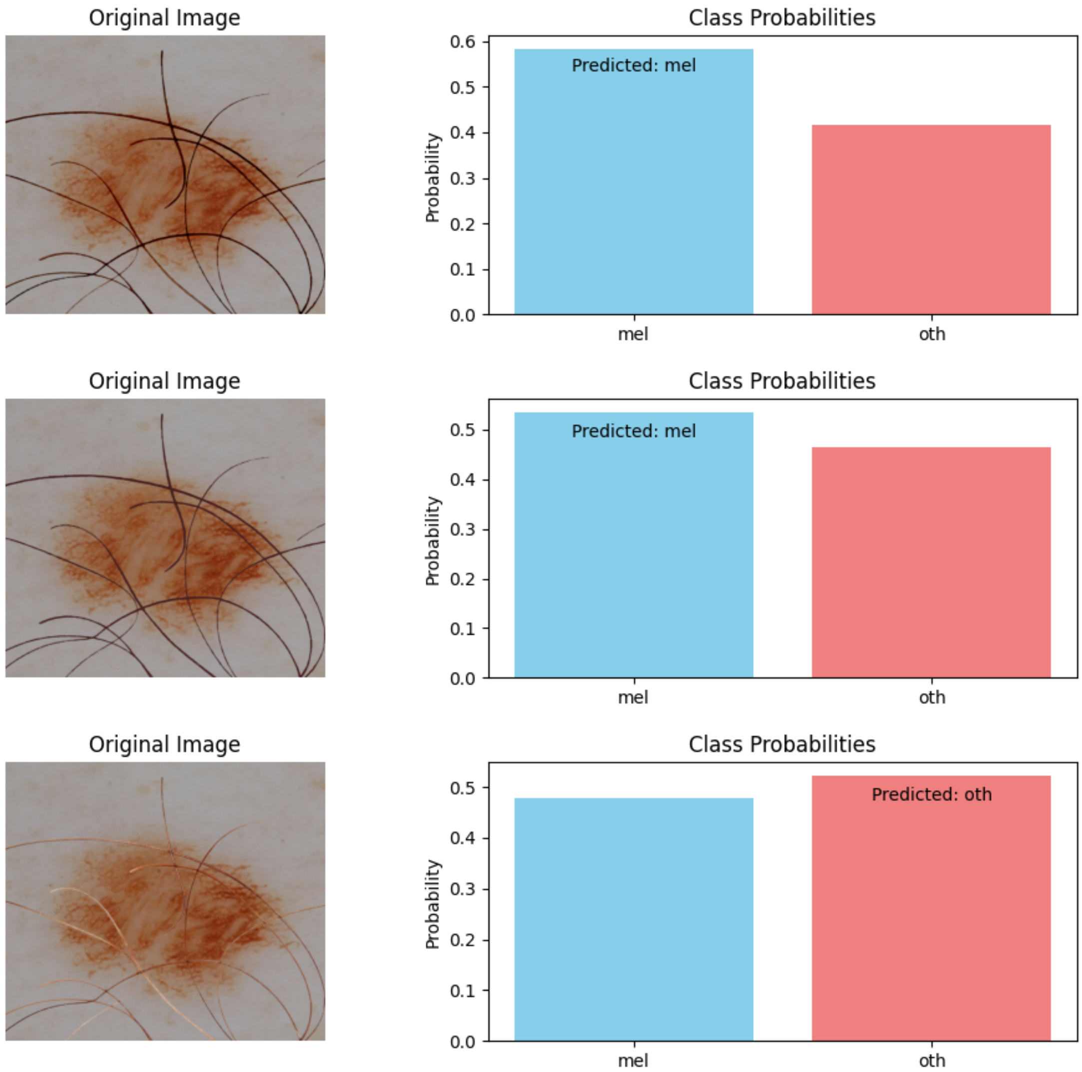


Figure 5.5: Class Probabilities - Experiment 5

Figure 5.5, displays class probabilities for each image in Experiment 5.

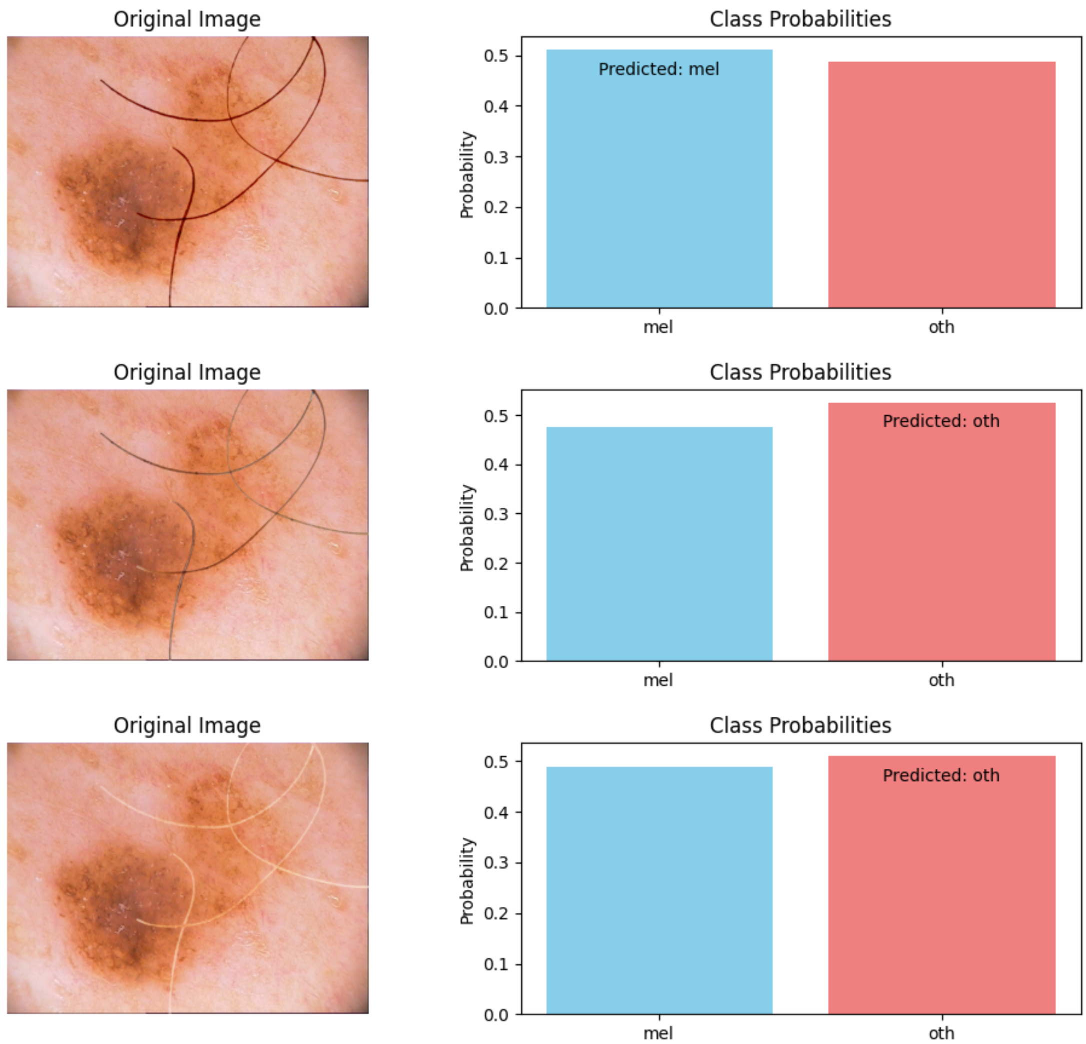


Figure 5.6: Class Probabilities - Experiment 6

Figure 5.6, displays class probabilities for each image in Experiment 6.

The results displayed in above figures collectively highlight a consistent relationship between the density of hair artifacts and the model's probability of melanoma classification. Higher hair density tends to increase the probability of melanoma, even in instances of misclassification, while lower hair density introduces uncertainty and decreases the probability of melanoma classification.

5.3.1 Impact of Hair Artifacts

The results strongly suggest that hair artifacts have a significant impact on the model's predictions. When hair artifacts are present at high or moderate densities, the model often misclassifies images as ”Melanoma,” leading to false positives.

5.3.2 Better Diagnosis with Hair Removal Mechanisms

Given the clear relationship between lower hair density and a decrease in melanoma predictions, it becomes evident that implementing hair removal mechanisms is crucial in melanoma diagnosis. Hair removal mechanisms can significantly improve the accuracy of melanoma diagnosis. These mechanisms enable the model to focus on the actual skin features related to melanoma rather than being misled by hair artifacts.

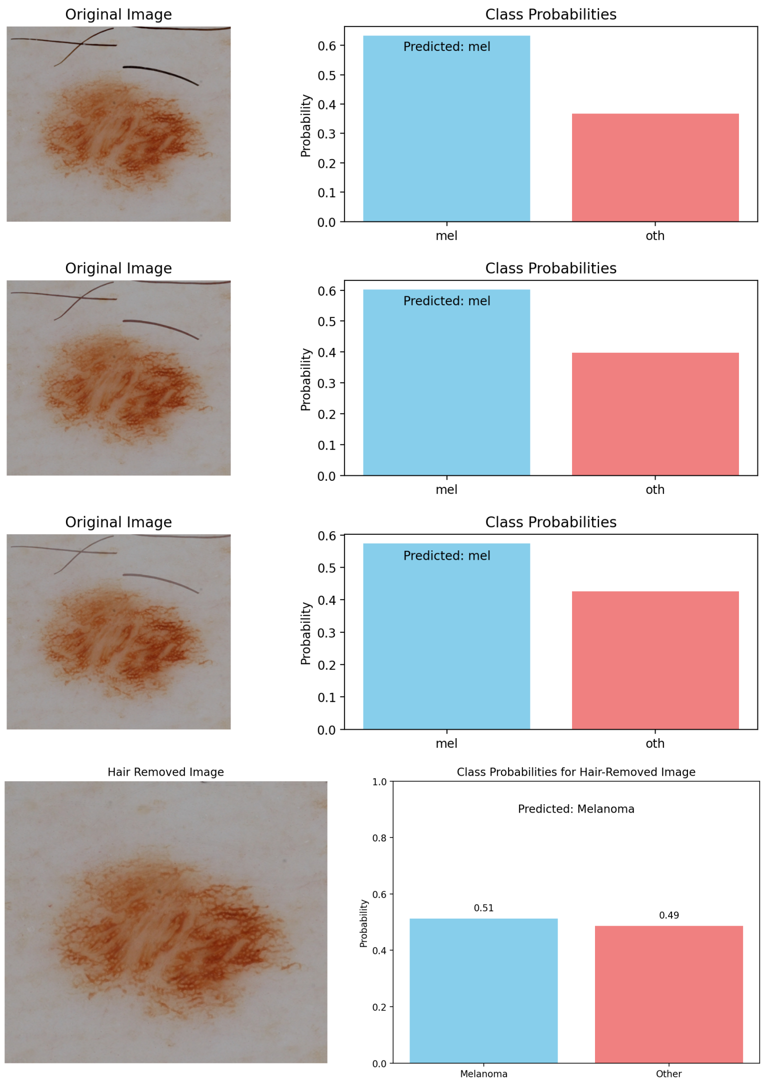


Figure 5.7: Class Probabilities - Example 1

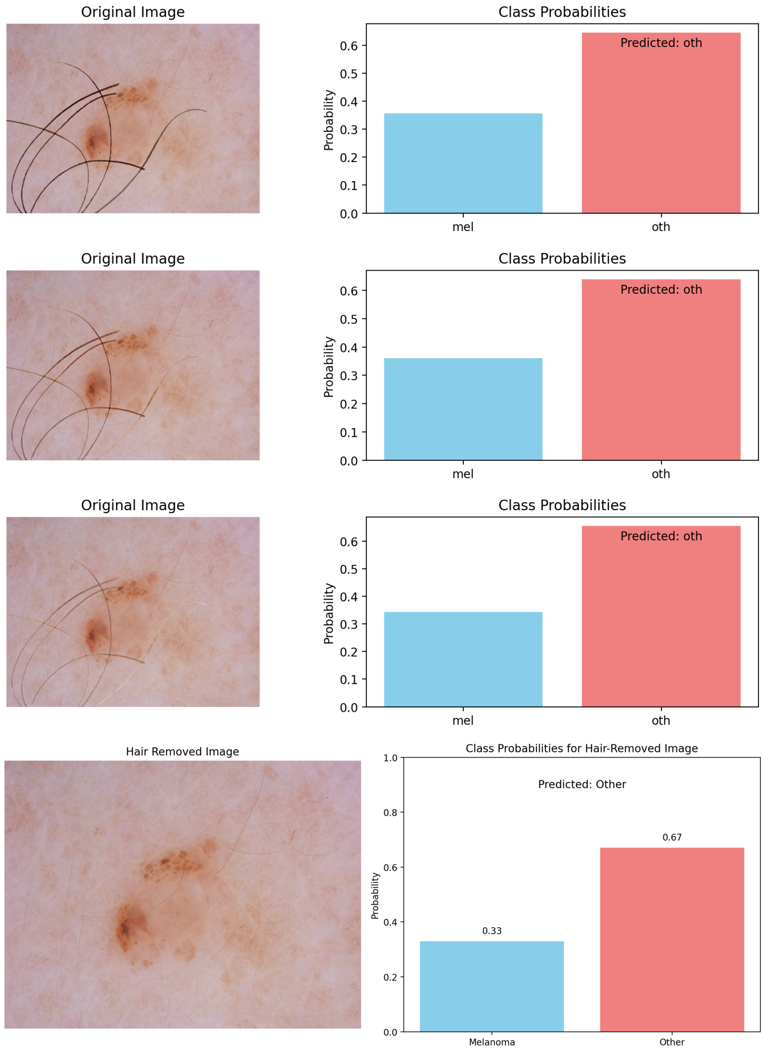


Figure 5.8: Class Probabilities - Example 2

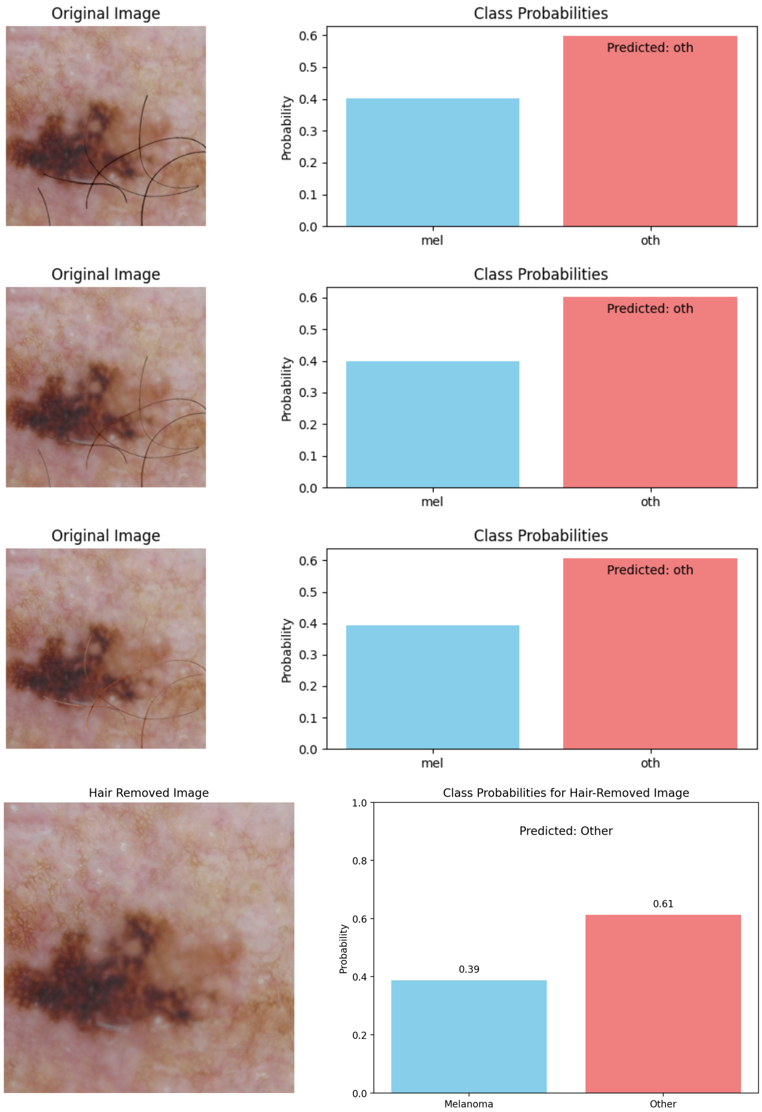


Figure 5.9: Class Probabilities - Example 3

The above figures 5.7, 5.8, and 5.9 illustrate a consistent trend in the reduction of melanoma probability as hair density decreases, as observed in experiments 1 to 6. When hair removal is applied, a further decrease in melanoma probability can be seen in these figures. All of these images belong to the non-melanoma class. The first image is incorrectly predicted as melanoma, while the remaining two are predicted correctly. Despite the incorrect prediction in the first case, the probability of melanoma decreases with decreasing hair density in all instances. This incorrect prediction may be influenced by other factors in the image, as classification relies on more than just the presence of hair artifacts. Therefore, these observations clearly emphasize the importance of removing hair artifacts in classification tasks to enhance model accuracy.

5.3.3 Evaluation of Hair Removal

Example	PSNR
1	22.38
2	22.29
3	25.42

Table 5.2: Evaluation of hair removal using PSNR

The PSNR values in Table 5.2 suggests that the inpainted images are relatively close in quality to the original image

5.3.4 Clinical Significance

In a clinical setting, where accurate melanoma diagnosis is essential for patient care, these findings underscore the importance of addressing the issue of hair artifacts. Implementing hair removal techniques as part of the diagnostic process can enhance the reliability of the model's predictions and reduce the likelihood of unnecessary concerns caused by hair artifacts.

In summary, the results highlight a critical relationship between the presence and density of hair artifacts and the accuracy of melanoma classification. Lower hair density is associated with more reliable predictions of melanoma conditions. Therefore, implementing effective hair removal mechanisms as a standard part of melanoma diagnosis can lead to more accurate and clinically meaningful results, ultimately improving patient outcomes and reducing false positive diagnoses.

Chapter 6

Conclusion

In this final chapter, the project is summarized, the overall findings are evaluated, the finished work is examined, and prospective areas for project improvement and enhancement beyond its current state are suggested.

6.1 Assessment of Outcome

The objective of this project was to examine and measure the influence of skin hair artifacts present in skin image datasets on the effectiveness of convolutional neural network (CNN) models when used for melanoma classification.

Throughout the course of this project, extensive research was carried out to gain a deep comprehension of the subject matter, specifically examining the impact of skin hair artifacts found in skin image datasets on the performance of convolutional neural network (CNN) models when applied to melanoma classification. This research encompassed a comprehensive literature review, which provided a strong foundation of knowledge in the field.

After completing the literature review, a thorough investigation was conducted to assess the likelihood of melanoma classification in the presence of varying densities of hair in skin images. Additionally, methods for removing these artifacts were implemented to mitigate their influence on the classification process. The project then involved the evaluation of probability variations and hair removal techniques in the context of image classification tasks.

This study has provided valuable insights into the impact of hair artifacts on the accuracy and reliability of deep learning-based skin cancer diagnosis. Through a comprehensive analysis of various experiments, a consistent relationship between the density of hair artifacts and the model's probability of melanoma classification is observed . Notably, higher hair density often led to an increased likelihood of melanoma predictions, even if some misclassifications occurred, while lower hair density introduced uncertainty and decreased the probability of melanoma classification.

These findings underscore the critical importance of addressing hair artifacts in dermatological image analysis. The study emphasizes that implementing effective hair removal mechanisms, such as preprocessing techniques or dedicated artifact detection and removal methods, can significantly enhance the accuracy of melanoma diagnosis. By doing so, these mechanisms enable the model to focus on genuine skin features associated with melanoma, thus reducing the likelihood of false positives and improving diagnostic reliability.

In a clinical context where accurate melanoma diagnosis is paramount for patient care, these insights hold immense significance. The integration of hair removal techniques into the diagnostic process emerges as a vital step toward enhancing the clinical relevance of deep learning models in dermatology. It can lead to more precise and trustworthy results, ultimately benefiting patient outcomes by minimizing unnecessary concerns stemming from hair artifacts.

In summary, this study underscores the pivotal role of addressing hair artifacts in advancing the field of dermatological image analysis. By implementing robust hair removal mechanisms, we can unlock the full potential of deep learning-based skin cancer diagnosis, contributing to improved patient care, reduced false positives, and more accurate diagnoses in the realm of skin cancer.

6.2 Limitations

Several limitations were encountered during the course of this research project. Firstly, it was observed that the effectiveness of hair removal techniques was dependent on the quality of the hair masks used. In cases where the hair masks were inaccurate or incomplete, the inpainting process may not have been entirely successful, potentially leading to residual artifacts. Secondly, the study primarily focused on the impact of hair artifacts on melanoma classification, but did not extensively explore the effects on other types of skin lesions. Future research may need to broaden its scope to encompass a wider range of dermatological conditions. Finally, while the study highlighted the importance of addressing hair artifacts, it did not propose novel hair removal methods. Future work in this area could involve the development of innovative and efficient hair removal algorithms tailored to dermatological image analysis.

6.3 Future Work

In light of the identified limitations, there are several promising avenues for future research in the field of dermatological image analysis. One potential area of exploration is the development of advanced hair artifact detection algorithms. These algorithms could automatically identify regions in skin images that are affected by hair artifacts, enabling precise inpainting and artifact removal. Additionally, the integration of artificial intelligence techniques, such as generative adversarial networks (GANs), could be explored to enhance the realism and quality of inpainted regions. Furthermore, future studies could investigate the impact of hair artifacts on the performance of different types of skin lesion classifiers, potentially leading to more specialized models for various dermatological conditions. Qualitative assessments, including expert evaluations and user studies, could be incorporated into the evaluation process to provide a more holistic understanding of the effectiveness of hair removal techniques. Lastly, collaboration with dermatologists and medical practitioners could facilitate the development of practical, clinically relevant solutions for skin image analysis, ultimately improving the accuracy and reliability of diagnostic tools in the field of dermatology.

References

Abbas, Qaisar, M Emre Celebi, and Irene Fondón Garcıa (2011). “Hair removal methods: A comparative study for dermoscopy images”. In: *Biomedical Signal Processing and Control* 6.4, pp. 395–404.

Alquran, Hiam et al. (2017). “The melanoma skin cancer detection and classification using support vector machine”. In: *2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT)*. IEEE, pp. 1–5.

Aswin, RB, J Abdul Jaleel, and Sibi Salim (2014). “Hybrid genetic algorithm—Artificial neural network classifier for skin cancer detection”. In: *2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT)*. IEEE, pp. 1304–1309.

Cassidy, Bill et al. (2022). “Analysis of the ISIC image datasets: Usage, benchmarks and recommendations”. In: *Medical image analysis* 75, p. 102305.

Fiorese, M, Enoch Peserico, and Alberto Silletti (2011). “VirtualShave: automated hair removal from digital dermatoscopic images”. In: *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE, pp. 5145–5148.

Gairola, Ajay Krishan, Vidit Kumar, and Ashok Kumar Sahoo (2022). “Exploring multiple deep learning models for skin cancer classification”. In: *2022 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*. IEEE, pp. 805–810.

Haggenmüller, Sarah et al. (2021). “Skin cancer classification via convolutional neural networks: systematic review of studies involving human experts”. In: *European Journal of Cancer* 156, pp. 202–216.

Hearn, Jonathan (2008). “Competitive medical image segmentation with the fast marching method”. PhD thesis. Case Western Reserve University.

Kiani, Kimia and Ahmad R Sharafat (2011). "E-shaver: An improved DullRazor® for digitally removing dark and light-colored hairs in dermoscopic images". In: *Computers in biology and medicine* 41.3, pp. 139–145.

Koehoorn, Joost et al. (2016). "Effcient and effective automated digital hair removal from dermoscopy images". In: *Mathematical Morphology-Theory and Applications* 1.1.

Lee, Howard and Yi-Ping Phoebe Chen (2014a). "Cell cycle phase detection with cell deformation analysis". In: *Expert systems with applications* 41.6, pp. 2644–2651.

— (2014b). "Skin cancer extraction with optimum fuzzy thresholding technique". In: *Applied intelligence* 40, pp. 415–426.

Lee, Kin Wai and Renee Ka Yin Chin (2020). "The effectiveness of data augmentation for melanoma skin cancer prediction using convolutional neural networks". In: *2020 IEEE 2nd International Conference on Artificial Intelligence in Engineering and Technology (IICAIET)*. IEEE, pp. 1–6.

Li, Yuexiang and Linlin Shen (2018). "Skin lesion analysis towards melanoma detection using deep learning network". In: *Sensors* 18.2, p. 556.

Madooei, Ali et al. (2012). "Automated pre–processing method for dermoscopic images and its application to pigmented skin lesion segmentation". In: *Color and Imaging Conference*. Vol. 20. Society of Imaging Science and Technology, pp. 158–163.

Mahmood, Dr Ahlam F and Hamed Abdulaziz Mahmood (2015). "Artifact Removal from Skin Dermoscopy Imagesto Support Automated Melanoma Diagnosis". In: *Al-Rafidain Engineering Journal (AREJ)* 23.5, pp. 22–30.

Marchetti, Michael A et al. (2018). "Results of the 2016 International Skin Imaging Collaboration International Symposium on Biomedical Imaging challenge: Comparison of the accuracy of computer algorithms to dermatologists for the diagnosis of melanoma from dermoscopic images". In: *Journal of the American Academy of Dermatology* 78.2, pp. 270–277.

Masood, Ammara, Adel Ali Al-Jumaily, and Tariq Adnan (2014). "Development of automated diagnostic system for skin cancer: Performance analysis of neural network learning algorithms for classification". In: *Artificial Neural Networks and Machine Learning–ICANN 2014: 24th International Conference on Artificial Neural Networks, Hamburg, Germany, September 15-19, 2014. Proceedings 24*. Springer, pp. 837–844.

Mirzaalian, Hengameh, Tim K Lee, and Ghassan Hamarneh (2014). “Hair enhancement in dermoscopic images using dual-channel quaternion tubularness filters and MRF-based multilabel optimization”. In: *IEEE Transactions on Image Processing* 23.12, pp. 5486–5496.

Mustafa, Suleiman (2017). “Feature selection using sequential backward method in melanoma recognition”. In: *2017 13th International Conference on Electronics, Computer and Computation (ICECCO)*. IEEE, pp. 1–4.

Nasonova, Alexandra et al. (2014). “Image warping in dermatological image hair removal”. In: *Image Analysis and Recognition: 11th International Conference, ICIAR 2014, Vilamoura, Portugal, October 22-24, 2014, Proceedings, Part II 11*. Springer, pp. 159–166.

Nguyen, Nhi H, Tim K Lee, and M Stella Atkins (2010). “Segmentation of light and dark hair in dermoscopic images: a hybrid approach using a universal kernel”. In: *Medical Imaging 2010: Image Processing*. Vol. 7623. SPIE, pp. 1436–1443.

Pewton, Samuel William and Moi Hoon Yap (2022). “Dark Corner on Skin Lesion Image Dataset: Does it matter?” In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4831–4839.

Pour, Mansoureh Pezhman and Huseyin Seker (2020). “Transform domain representation-driven convolutional neural networks for skin lesion segmentation”. In: *Expert Systems with Applications* 144, p. 113129.

Salido, Julie Ann and CONRADO Ruiz Jr (2018). “Hair artifact removal and skin lesion segmentation of dermoscopy images”. In: *Asian Journal of Pharmaceutical and Clinical Research* 11.3.

Satheesha, TY, D Satyanarayana, and MN Giriprasad (2014). “SVMs classification based on insitu melanoma”. In: *International Conference on Circuits, Communication, Control and Computing*. IEEE, pp. 387–392.

Sultana, Alina, Mihai Ciuc, et al. (2012). “Preliminary work on dermatoscopic lesion segmentation”. In: *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*. IEEE, pp. 2273–2277.

Sultana, Alina, Ioana Dumitrache, et al. (2014). “Removal of artifacts from dermatoscopic images”. In: *2014 10th International Conference on Communications (COMM)*. IEEE, pp. 1–4.

Talavera-Martınez, Lidia, Pedro Bibiloni, and Manuel González-Hidalgo (2019). “Comparative study of dermoscopic hair removal methods”. In: *VipIMAGE 2019: Proceedings of the VII ECCOMAS Thematic Conference on Computational Vision* *and Medical Image Processing, October 16–18, 2019, Porto, Portugal*. Springer, pp. 12–21.

Telea, Alexandru (2004). "An image inpainting technique based on the fast marching method". In: *Journal of graphics tools* 9.1, pp. 23–34.

Toossi, Mohammad Taghi Bahreyni et al. (2013). "An effective hair removal algorithm for dermoscopy images". In: *Skin Research and Technology* 19.3, pp. 230–235.

Wighton, Paul, Tim K Lee, and M Stella Atkins (2008). "Dermascopic hair disocclusion using inpainting". In: *Medical Imaging 2008: Image Processing*. Vol. 6914. SPIE, pp. 735–742.

Xie, Feng-Ying et al. (2009). "PDE-based unsupervised repair of hair-occluded information in dermoscopy images of melanoma". In: *Computerized Medical Imaging and Graphics* 33.4, pp. 275–282.

Xie, Fengying, Haidi Fan, et al. (2016). "Melanoma classification on dermoscopy images using a neural network ensemble model". In: *IEEE transactions on medical imaging* 36.3, pp. 849–858.

Xie, Fengying, Jiawen Yang, et al. (2020). "Skin lesion segmentation using high-resolution convolutional neural network". In: *Computer methods and programs in biomedicine* 186, p. 105241.

Zaqout, Ihab Salah (2017). "An efficient block-based algorithm for hair removal in dermoscopic images". In: 41.4, pp. 521–527.

Zhou, Howard et al. (2008). "Feature-preserving artifact removal from dermoscopy images". In: *Medical Imaging 2008: Image Processing*. Vol. 6914. SPIE, pp. 439–447.

Appendix A

Project Source Code

The source code for the entire project is available at the following link:

Source Code Colab Notebook Link