Project report on

DERM.AI: A MODERN SOLUTION TO DERMOSCOPY NEEDS

Submitted By

Ishani Bhowmick (University Roll No.: 34230921007)

Team worked with

Abhishek Vishwakarma (University Roll No.: 34230921014)

K. Divyansh (University Roll No.: 34230921002)

Arka Bhattacharjee (University Roll No.:34230921008)

BACHELOR OF TECHNOLOGY

In

COMPUTER SCIENCE AND ENGINEERING

(INTERNET OF THINGS & CYBERSECURITY INCLUDING

BLOCK CHAIN TECHNOLOGY)

SEM – 6

2024

Under the esteemed supervision of

Prof. Samiran Basak

Assistant Professor

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

(INTERNET OF THINGS & CYBERSECURITY INCLUDING BLOCK CHAIN TECHNOLOGY)

CERTIFICATE

This is to certify that the project work entitled "**Derm.AI: A Modern Solution to Dermoscopy Needs**" submitted by Ishani Bhowmick (University Roll No.: 34230921007), team worked with Abhishek Vishwakarma (University Roll No.: 34230921014), K. Divyansh (University Roll No.: 34230921002), and Arka Bhattacharjee (University Roll No.:34230921008), Bachelor of Technology in Computer Science and Engineering (Internet of Things & Cyber Security Including Block Chain Technology) at Future Institute of Technology, is done under my guidance and supervision and is a bonafide work done by them.

The matter presented in this thesis has not been submitted for the award of any other degree

of this or any other Institute/University.

I wish them all success in life.

Date:

Mr. Samiran Basak
Assistant Professor
Department Of Computer Science And Engineering
(Internet Of Things & Cybersecurity Including
Block Chain Technology)

CERTIFICATE

This is to certify that the project work entitled "**Derm.AI: A Modern Solution to Dermoscopy Needs**" submitted by Ishani Bhowmick (University Roll No.: 34230921007), team worked with Abhishek Vishwakarma (University Roll No.: 34230921014), K. Divyansh (University Roll No.: 34230921002), and Arka Bhattacharjee (University Roll No.:34230921008), Bachelor of Technology in Computer Science and Engineering (Internet of Things & Cyber Security Including Block Chain Technology) at Future Institute of Technology, , is done under the guidance and supervision of **Prof. Samiran Basak** and is a bonafide work done by them.

The matter presented in this thesis has not been submitted for the award of any other degree of this or any other Institute/University.

I wish them all success in life.

Date:

Prof. (Dr.) Tanusree Chatterjee
Associate Professor and H.O.D
Dept. of CSE ((Internet of Things & Cybersecurity
Including Block Chain Technology)
Future Institute of Technology

ACKNOWLEDGEMENT

We would like to express our special thanks of gratitude to our Mentor, Prof. Samiran Basak, and our HOD Prof. (Dr.) Tanushree Chatterjee gave us the golden opportunity to do this wonderful project on the topic "**Derm.AI: A Modern Solution to Dermoscopy Needs**" for ensuring the early diagnosis and convenience of dermatological treatments and to improve human healthcare system, which also helped us in doing a lot of study and research in this domain, and we came to know about many new technologies and methods. Secondly, we would also like to thank each of our team-members who worked hand in hand in finalising and completing the project framework within the stipulated time.

DATE:

PLACE: Garia, Kolkata

_____

Abhishek Vishwakarma (University Roll No.: 34230921014)

_____

Ishani Bhowmick (University Roll No.: 34230921007)

_____

K. Divyansh (University Roll No.: 34230921002

_____

Arka Bhattacharjee (University Roll No.:34230921008)

Abstract

The advent of "**Derm.AI: A Modern Solution to Dermoscopy Needs**" promises to revolutionize dermatological diagnostics. This AI-powered system utilizes cutting-edge image processing and machine learning algorithms to expedite and refine the analysis of skin conditions. This software prioritizes streamlining the diagnostic workflow, enhancing clinician productivity, and guaranteeing comprehensive evaluation of diverse dermatological presentations. An intuitive user interface fosters accessibility for both healthcare professionals and patients seeking preliminary assessments. The system's core functionalities include advanced image processing algorithms, adaptable machine learning models, rapid response system and inherent scalability to accommodate an expanding knowledge base. Moreover, seamless integration with existing healthcare infrastructure ensures smooth collaboration and promotes holistic patient care.

Contents

Highlights

- Proposed an algorithm for accurate segmentation of hair strands.
- Employed a U-Net trained on manually created hair mask dataset to segment hair from skin lesion images.
- Used state-of-the-art "INPAINT_TELEA: *An image inpainting technique based on the fast marching method* (Telea, 2004)" to inpaint the hair gap.
- Proposed a non-referenced Structural Similarity Index (SSIM) method to evaluate the hair removal and the inpainting process.
- Used PH^2 dataset, to evaluate the performance of the proposed method and compared it with other state-of-the-art methods.

Abstract

Occlusion caused by hair in dermoscopic images affects the diagnostic operation and the accuracy of its analysis of a skin lesion. Furthermore, dermis hair has different characteristics, such as being thin, overlapping, faded, of similar contrast or color to the underlying skin, and obscuring/covering textured lesions. These factors render digital hair removal (DHR), which involves hair segmentation and hair gap inpainting, a challenging undertaking. Consequently, traditional hard-coded threshold-based hair removal methods are not effective, resulting in over-removal which loses important information about the skin lesion, or under-removal which cannot remove the hair effectively. In this article, we propose a deep learning approach to DHR that is based on U-Net and a free-form image inpainting technique through OpenCV's inpainting methods. In the field of hair segmentation, a well-labeled dataset is generated and utilized to train U-Net in order to obtain precise hair masks. The process of DHR repeats until there is no change in the average value of SSIM. Using the PH^2 dataset, it has been demonstrated that the proposed method exhibits superior performance compared to other current methods.

1. Introduction

The human body is a testament to the wonders of biological engineering, and at the forefront of this intricate design lies skin, our largest and most visible organ. Skin transcends its role as a smooth, protective covering; it's a dynamic and multifaceted shield safeguarding us from a multitude of threats. It acts as our first line of defence against a relentless assault of harmful pathogens, from bacteria and viruses to parasites and fungi. It's a vigilant barrier against environmental onslaughts, shielding us from the harmful effects of ultraviolet radiation and regulating our internal temperature through sweat production and blood vessel constriction. But skin's wonders extend far beyond mere protection. Skin serves as a window to our internal world, a canvas upon which our emotions and well-being are often painted. Through a network of specialized nerve endings, it allows us to perceive the world around us, interpreting touch, pressure, heat, and cold, enabling us to interact with our environment in a meaningful way. Skin even plays a vital role in our overall health by contributing to the production of vitamin D, essential for bone health and immune function.

However, despite its remarkable capabilities, skin diseases remain a significant health concern plaguing millions worldwide. In India, this burden manifests in a unique and multifaceted way. This report delves deep into the intricate world of skin, exploring its structure, physiology, and the diverse range of functions it performs in maintaining our health. We will shift our focus to the specific challenges faced by the Indian population when it comes to skin health. India's tropical climate acts as a double-edged sword. While it provides an abundance of sunshine, crucial for vitamin D synthesis, it can also exacerbate certain skin

conditions such as fungal infections and hyperpigmentation. Increased sweating due to hot and humid weather can also lead to miliaria, commonly known as heat rash. People with darker skin tones are more prone to post-inflammatory hyperpigmentation, a condition where areas of skin become darker after inflammation or injury. Limited access to clean water and proper sanitation can contribute to the spread of infectious skin diseases like scabies and bacterial infections. Additionally, potential nutritional deficiencies, particularly a lack of essential vitamins and minerals like Vitamin A, iron, and zinc, can weaken the skin's natural defences, making it more susceptible to problems like eczema and psoriasis.
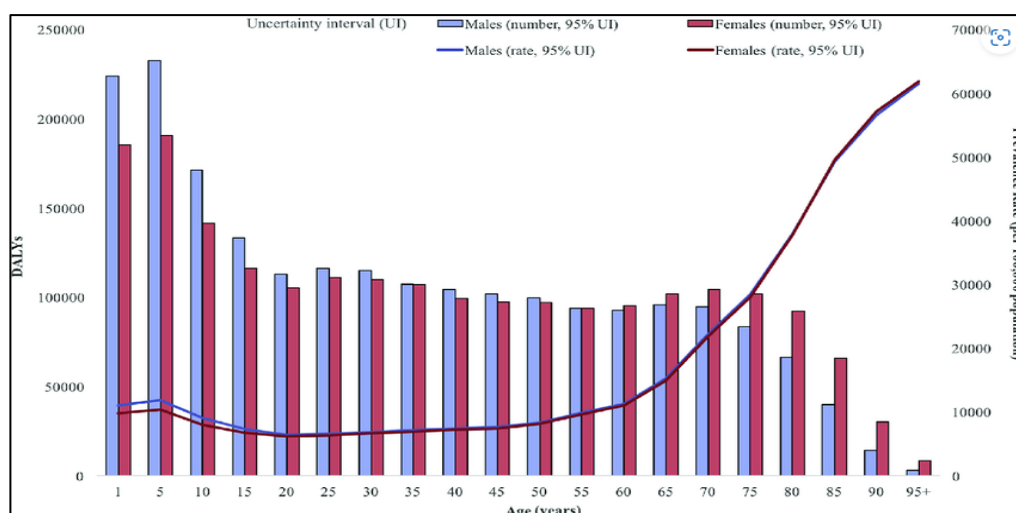

Figure 1: Prevalence of skin disorders due to sheer negligence is dangerously high though it has low fatality rate

Dermatological diagnostics, traditionally reliant on manual examination and subjective interpretations, are witnessing a transformative evolution propelled by advanced image processing, artificial intelligence, and machine learning. This report culminates with the introduction of a ground-breaking project that leverages the transformative power of artificial intelligence (AI) for skin disease detection in India. This innovative approach has the potential to revolutionize the way dermatological care is delivered, particularly in a nation with a vast population and a shortage of qualified dermatologists. By harnessing the power of AI, we can enable earlier diagnoses through faster and more accessible screening methods. Imagine a mobile application that allows users to upload images of suspicious skin lesions, and the AI algorithms analyse the image to identify potential matches for various skin conditions. This early detection paves the way for more targeted treatment plans, ultimately leading to improved patient outcomes. Furthermore, AI-powered skin disease detection holds immense promise for remote consultations, particularly in underserved areas with limited access to specialist healthcare. This technology can bridge the gap between patients and dermatological expertise, ensuring that everyone has the opportunity to receive timely and effective treatment. Throughout this report, we aim to:

- **Demystify the Marvel of Skin:** We will embark on a journey to understand the intricate workings of skin, exploring its structure, physiology, and its essential role in maintaining overall health. This exploration will delve into the layers of skin, their specialized cells, and the fascinating processes that keep our skin functioning optimally.
- **Understanding Skin Challenges in India:** We will then shift our focus to analyse the specific challenges and prevalent skin diseases affecting the Indian population. This section will consider factors unique to India, such as its tropical climate that can exacerbate certain conditions, potential nutritional deficiencies, and socio-economic factors that can limit access to quality healthcare.
- **Revolutionizing Dermatology with AI:** Finally, we will introduce our cutting-edge AI-based skin disease detection project. This section will explore the immense potential of this innovative

approach to transform dermatological care in India. We will examine how AI can enable faster diagnoses, improved treatment strategies through early detection, and increased accessibility to care, particularly in remote or underserved areas.

2. Skin and Its Diseases: A Multifaceted Challenge for India

Skin, despite its remarkable resilience, is susceptible to a wide range of diseases. These conditions can arise from various factors, both internal and external. Here, we will delve into the diverse spectrum of skin diseases impacting the Indian population.[32] Human skin shares anatomical, physiological, biochemical and immunological properties with other mammalian lines. The skin is unique in many ways, but no other organ demands so much attention and concern in both states of disease and health [1]. There is a huge focus on skin health, with fierce competition to have glowing, clearer, healthier, younger and fresher skin. And this focus can cause secondary problems with self-esteem and mental health. In humans, skin pigmentation (affected by melanin) varies among populations, and skin type can range from dry to non-dry and from oily to non-oily. Such skin variety provides a rich and diverse habitat for the approximately one thousand species of bacteria from nineteen phyla which have been found on human skin. India's tropical climate presents unique challenges for skin health. Increased sweating due to hot and humid weather can lead to miliaria (heat rash) and exacerbate pre-existing conditions like eczema. Additionally, the high UV index exposes the population to a greater risk of sunburn and potentially skin cancer. Sun protection becomes even more critical in a country with a significant portion of the population working outdoors. Here, we will delve into the diverse spectrum of skin diseases impacting the Indian population, highlighting some of the most prevalent concerns:
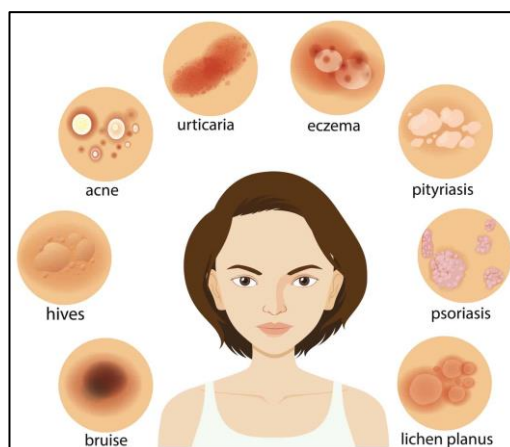

Figure 2: Common skin conditions

Infectious Diseases: India's climate and living conditions create a breeding ground for infectious skin diseases. Bacterial infections like impetigo and folliculitis are common, while fungal infections like ringworm and tinea corporis thrive in hot and humid environments. Scabies, a parasitic infestation caused by mites, also poses a significant challenge, particularly in crowded settings. [30] Viral infections such as chickenpox, shingles, and warts can also cause discomfort and require proper management.

Non-communicable Diseases: While infectious diseases pose a significant burden, non-communicable skin conditions are also prevalent in India. Eczema, a chronic inflammatory condition causing itchy and dry skin, affects people of all ages. Psoriasis, another chronic condition characterized by red, scaly patches, can significantly impact quality of life. Melasma, a condition causing dark patches on the face, is more common in individuals with darker skin tones. Acne vulgaris, commonly known as acne, is a prevalent concern among teenagers and young adults.[30]

Nutritional Deficiencies: Limited access to a balanced diet can contribute to skin problems. Deficiencies in essential vitamins and minerals like vitamin A, iron, and zinc can weaken the skin's natural barrier function and exacerbate existing conditions. These deficiencies can make the skin more susceptible to infections and hinder its ability to heal properly.[30]

Climate-related Concerns: India's tropical climate presents unique challenges for skin health. Increased sweating due to hot and humid weather can lead to miliaria (heat rash) and exacerbate pre-existing conditions like eczema. Additionally, the high UV index exposes the population to a greater risk of sunburn and potentially skin cancer. Sun protection becomes even more critical in a country with a significant portion of the population working outdoors.[30]

Pigmentation Disorders: Variations in melanin production can lead to various pigmentation disorders. Vitiligo, an autoimmune condition causing white patches on the skin, affects millions in India. Post-inflammatory hyperpigmentation, where areas of skin become darker after inflammation or injury, is a common concern, particularly for individuals with darker skin tones. Understanding the reasons behind these pigmentation changes can help in managing the condition and improving cosmetic outcomes.[30]
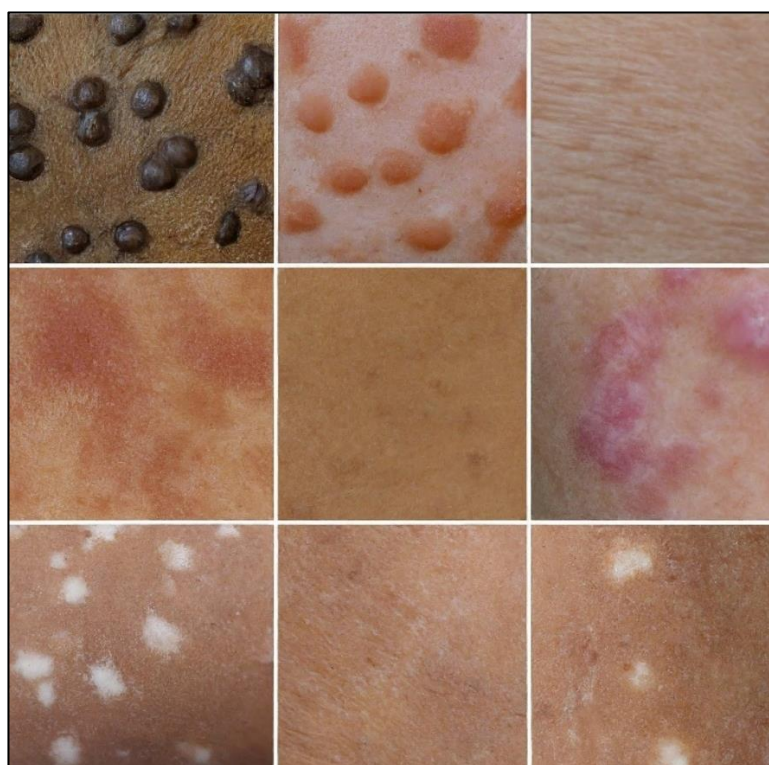

Figure 3: Close-up of a person's skin with (i) melanocytic nevus (ii) seborrheic dermatitis (iii) acne (iv) eczema (v) pityriasis rosea (vi) psoriasis (vii) albinism

3. Skin Pigmentation Disorders: A Concern for Many in India

Skin and subcutaneous diseases, affecting the layer of tissue beneath the skin, represent a significant, yet often overlooked, source of morbidity (ill health) worldwide. Despite a relatively low mortality rate of around 1%, these conditions can lead to substantial disability, placing a heavy burden on healthcare systems. The impact extends beyond the physical realm, often affecting a person's quality of life, social interactions, and even mental well-being.[32]

The severity of growing skin diseases in India is further emphasized by the fact that the World Health Organization has included skin disease under the most common non-communicable diseases in India. In 2013, with prevalence rate of 10 percent, the population affected across India from skin disease is estimated at nearly 15.1 crore. It is estimated that at a Compound Annual Growth Rate of 12 percent about 18.8 crore people is likely to suffer from skin disease by 2015 (Figure 2). The emergence of cosmetic and anti-aging treatments has changed the face of the skin care industry. The number of these cosmetic procedures is expected reach 18-20 lakh by 2015 in India. During the past 5 years, there has been a considerable growth in the Indian cosmetic/aesthetic surgery industry. The demand is high in urban regions like Chandigarh, Mumbai, Delhi, Chennai, Kolkata, Hyderabad, Bangalore, and Pune.[32]
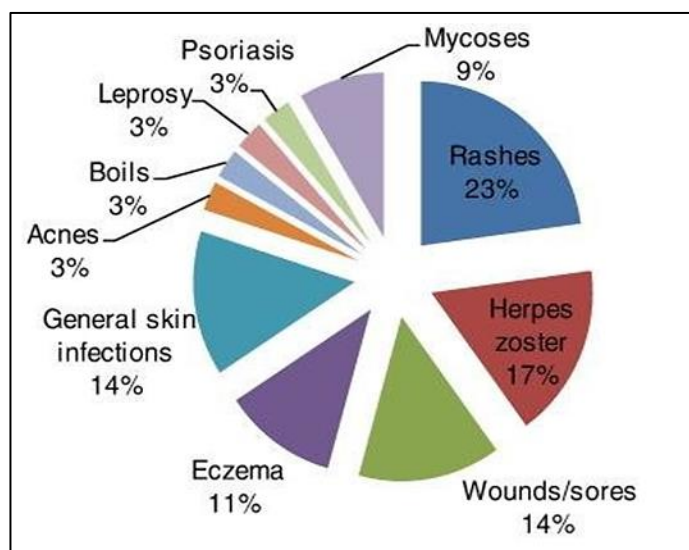

Figure 4: Frequency of affected people of different skin diseases

India's rich tapestry of skin tones is a source of beauty and diversity. However, variations in melanin production can also lead to skin pigmentation disorders, posing both aesthetic and health concerns for many individuals. Skin hyperpigmentation represent one of the major dermatological concerns for populations with pigmented skin phototypes, with a high prevalence in the Indian population. The usual treatments are based on the use of topical depigmenting agents to decrease the amount of melanin content or laser technologies. However, the role of solar exposure is often ignored.[34]

This silent epidemic is particularly concerning in developing nations like India. Here, the non-fatal impact of skin diseases surpasses that of even chronic conditions like cardiovascular diseases. A 2017 study revealed a stark disparity: individuals in India with skin and subcutaneous diseases experience an average of 455.06 disability-adjusted life years (DALYs) per 100,000 population, compared to 332.96 DALYs for cardiovascular diseases. DALYs, a metric used in global health studies, capture the combined burden of years of life lost due to premature death (mortality) and years lived with disability (morbidity). These statistics translate to a staggering difference – for every 100,000 people in India, skin diseases cause an additional 122.1 years of disability compared to cardiovascular diseases. The reasons behind this neglected burden are multifaceted. Limited public health awareness often leads to delayed diagnosis and treatment. Furthermore, a lack of qualified dermatologists, with estimates suggesting a shortfall of over 90%, particularly in rural areas, creates access barriers for patients seeking proper care. Additionally, the perception that skin diseases are primarily cosmetic concerns can lead to a lack of prioritization in national health planning and policies.[34]
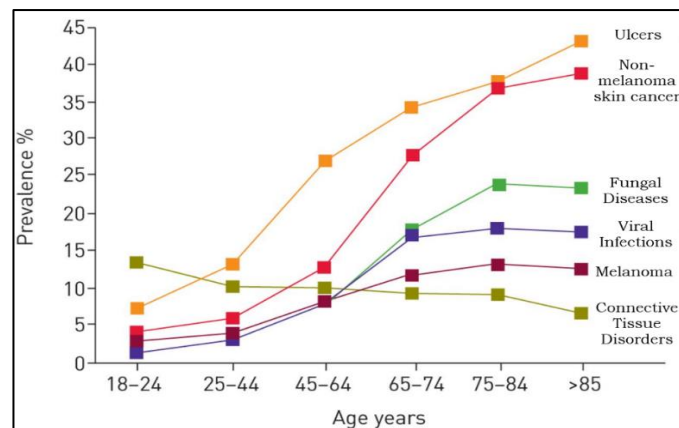
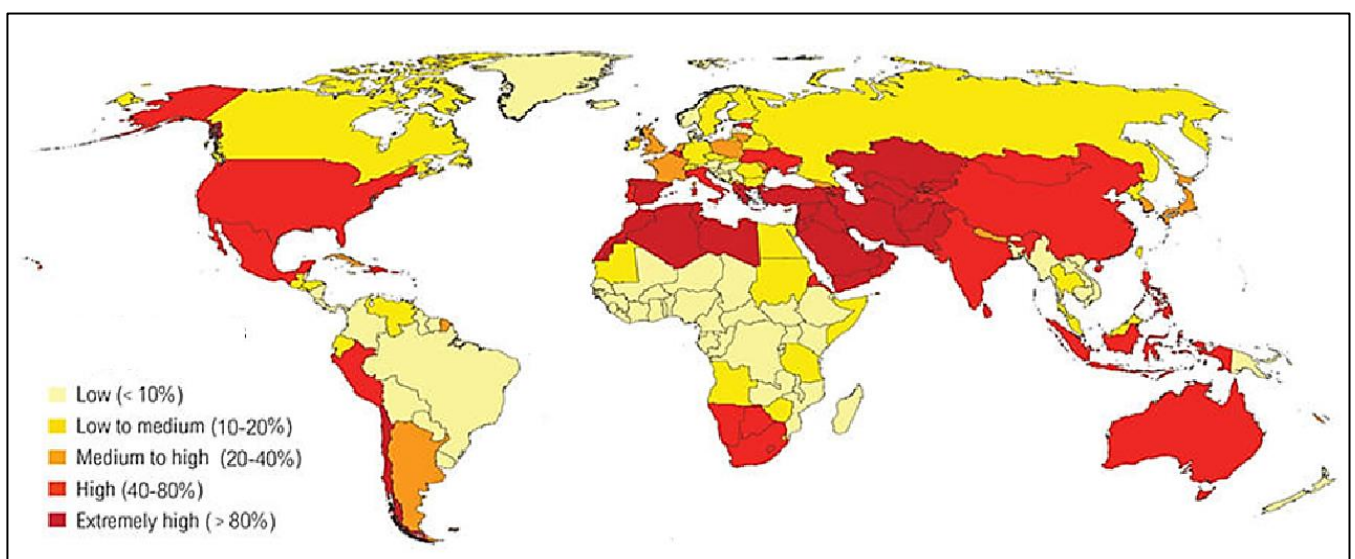*Figure 5: Prevalence of skin diseases over human-years*

Figure 6: Prevalence of skin diseases around the globe

The stark reality, however, is that skin diseases can have a profound impact on an individual's life. Skin conditions like eczema, psoriasis, and chronic infections can cause constant itching, pain, and disfigurement, affecting up to 800 million people globally. These physical symptoms can lead to social isolation, depression, and anxiety. Moreover, the financial burden of managing chronic skin conditions can be significant, further impacting a person's well-being. The good news is that there's growing recognition of the global burden of skin diseases. The ranking of skin diseases as the 10th leading cause of DALYs in the 2017 Global Burden of Disease study reflects this shift in perspective. This increased awareness paves the way for a more focused approach to tackling this challenge. Investing in public health education campaigns, expanding access to dermatological care, and allocating resources for research and development are crucial steps in this fight.[34]
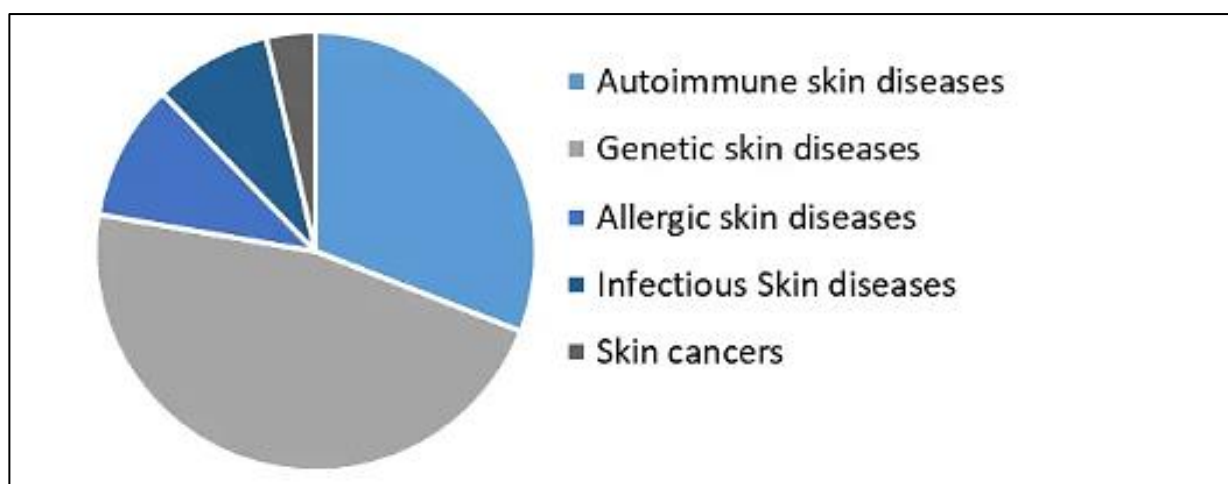
Figure 7: Skin diseases treatment market, by Type of Diseases 2023 (%)

4. Current Approaches for Skin Disease Treatment

India's vast population faces a unique challenge when it comes to managing skin diseases. While affordability and access to specialists remain hurdles, the country utilizes a multifaceted approach to treatment. Here's a breakdown of the current treatment landscape.

Topical Medications: A First Line of Defence

Creams, ointments, and lotions form the bedrock of treatment for many common skin conditions. The Indian Journal of Dermatology estimates that over 100 million people in India grapple with eczema, psoriasis, and acne. Topical medications containing corticosteroids, emollients, and anti-itching agents target these conditions by reducing inflammation, promoting healing, and alleviating discomfort.[33]

Oral Medications: Combating Infections and Inflammation

Oral medications play a crucial role in combating infections and managing inflammation. A 2023 study published in the International Journal of Applied Pharmaceutics highlights the prevalence of bacterial and fungal skin infections in India. Antibiotics and antifungals are used to combat these infections, while antihistamines provide relief from allergic skin reactions. For widespread or severe cases of eczema and psoriasis, oral corticosteroids can be particularly helpful.[33]

Phototherapy: Shining a Light on Treatment

This non-invasive treatment utilizes ultraviolet (UV) light to manage specific conditions like psoriasis and vitiligo. A 2022 news report by The Times of India highlights the growing adoption of phototherapy in India, with major hospitals and dermatology clinics investing in advanced phototherapy units. This treatment can significantly improve a patient's quality of life by alleviating symptoms and promoting skin health.[33]

Surgical Intervention: A Tool for Severe Cases

While less common, surgical interventions offer valuable solutions for severe skin conditions. These procedures can involve removing skin lesions, correcting scars, or performing reconstructive surgery. According to a report by the Indian Society of Dermatology, Mohs micrographic surgery, a specialized

technique for removing skin cancer, is gaining traction in India due to its high success rate and minimal scarring.[33]
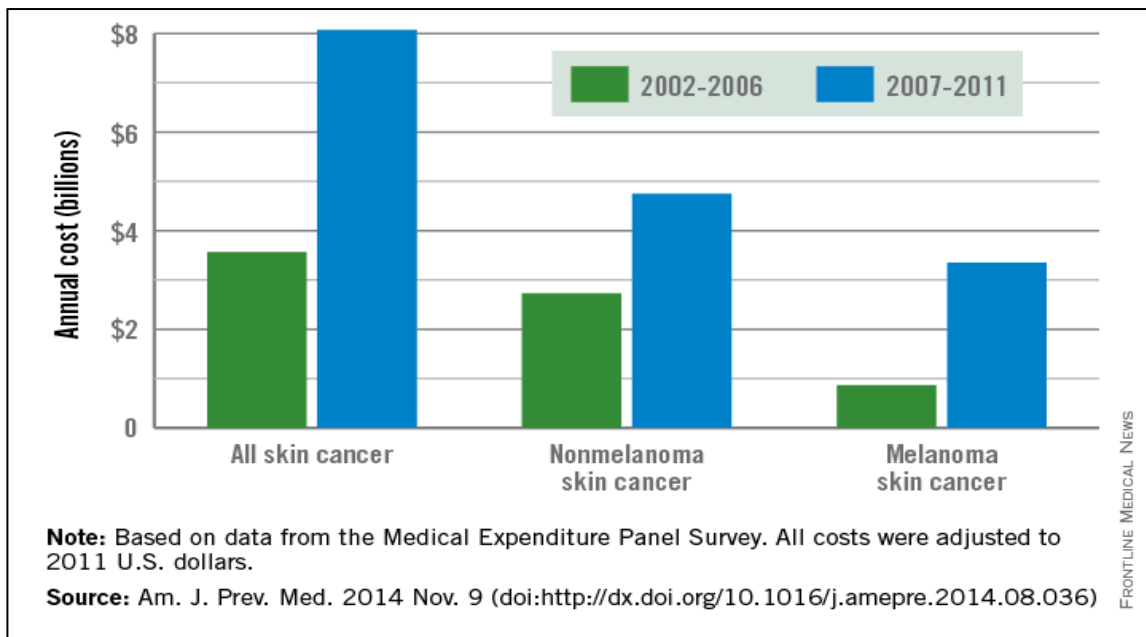

Figure 8: Estimated annual U.S. costs for skin cancer treatment

In the Global Burden of Disease Study 2017, twelve dermatoses i.e., psoriasis, dermatitis (atopic, contact and seborrheic), scabies, bacterial skin infections (cellulitis and pyoderma), fungal skin diseases, pruritus, viral skin diseases, acne vulgaris, alopecia areata, urticaria, decubitus ulcer and other skin and subcutaneous diseases are included under the category of skin and subcutaneous diseases. Some other dermatoses or conditions which are generally managed by dermatologists such as leprosy and melanomas, are not included in the skin and subcutaneous diseases category but are classified under 'neglected tropical diseases' and 'neoplasms' respectively, hence disability due to these is not represented in the skin and subcutaneous disorders group. Similarly, years lived with disability due to sexually transmitted diseases excluding HIV are also calculated separately. Hence, the skin and subcutaneous disease category in the Global Burden of Disease Study 2017 may not represent the entire burden of all dermatoses. To assess the burden of disabilities in populations, a system is employed that considers both the severity and duration of the condition. Disability weight, a score from zero (perfect health) to one (death), reflects the seriousness of a health condition. Years Lived with Disability (YLDs) capture the total healthy life years lost due to disability, factoring in both the disability weight and its duration. Since the impact of disability can vary by age, YLDs are adjusted to account for a population's specific age structure, ensuring fair comparisons across regions. This standardized YLD rate, expressed per 100,000 people, allows public health professionals to better understand and address disability-related challenges.[32]

5. Body Hair: A Hurdle in Skin Disease Treatment

The presence of body hair in areas affected by skin diseases can significantly hinder treatment, posing challenges for both patients and healthcare providers. The presence of body hair in areas with skin disease necessitates a tailored treatment approach. By acknowledging the challenges, it presents and implementing appropriate hair management strategies, healthcare professionals can optimize treatment efficacy and improve patient outcomes. Furthermore, ongoing research into new hair removal techniques with minimal side effects offers promising possibilities for the future of managing body hair in the context of skin disease treatment. Here's a closer look at the problems it creates.

Obstructed Treatment: Hair acts as a physical barrier, reducing the efficacy of topical medications like creams and ointments. A 2023 study published in the Journal of Clinical Dermatology found that adherence to topical medications decreased by 22% in patients with excessive body hair in the affected area. This highlights the importance of addressing body hair during treatment planning.[31]

Increased Irritation: Depending on the location and the specific skin condition (eczema, psoriasis is common examples), body hair can irritate already inflamed or sensitive skin. A 2022 report by the American Academy of Dermatology (AAD) states that friction from hair rubbing against inflamed skin can exacerbate itching and discomfort. This can lead to scratching, further damaging the skin and potentially causing secondary infections.[31]

Infection Risks: In cases with open wounds or weeping lesions (common in conditions like impetigo), body hair can trap moisture and debris. This creates a breeding ground for bacteria or fungus, potentially leading to secondary infections. A 2021 study published in the British Journal of Dermatology found that patients with skin diseases and excessive body hair were twice as likely to develop secondary infections compared to those with minimal hair growth.[31]

6. Traditional Body Hair Management Strategies for Optimal Treatment

While body hair management is crucial for optimal treatment of skin diseases, each method has limitations, the traditional ways to manage body hair during skin disease treatment are as following [38].

- **Trimming:** Simply trimming the hair to a short length can significantly improve medication application and reduce irritation. A 2022 study in the International Journal of Trichology showed a 17% improvement in treatment response for eczema patients who trimmed their body hair compared to those who didn't. However, trimming might not be sufficient for all conditions, especially those requiring very close contact with medication. Additionally, frequent trimming can be time-consuming and inconvenient for some patients.
- **Shaving:** For conditions requiring close contact with medication or where irritation is a major concern, shaving the affected area might be recommended. However, frequent shaving can irritate some skin types, so consulting a dermatologist is crucial. New research from a 2023 study published in the Journal of Cosmetic Dermatology suggests that electric razors with hypoallergenic blades might be a gentler alternative for frequent shaving in these cases. Frequent shaving, especially with traditional razors, can irritate sensitive skin, potentially worsening the underlying skin condition. Ingrown hairs can also be a concern with shaving, leading to further discomfort and inflammation.[40]
- **Depilatory Creams:** Chemical hair removal creams can be used with caution, but should be a last resort. Patch testing is essential to avoid allergic reactions on already compromised skin. Additionally, a 2023 study by the American College of Dermatology suggests that laser hair removal might be a viable option for long-term hair management in select cases, but requires careful evaluation by a dermatologist due to potential side effects. Chemical burns and skin irritation are significant risks associated with depilatory creams, particularly on already compromised skin. Patch testing is crucial before use. Additionally, the hair removal effects are temporary, requiring frequent reapplication.[41]

AI has the potential to revolutionize body hair removal for patients with skin diseases by providing personalized recommendations, optimizing existing technologies, and even paving the way for future robotic solutions. However, significant research, development, and regulatory approvals are still needed before AI becomes a mainstream tool in this area. AI algorithms are being developed to analyse skin

images and assist dermatologists in diagnosing skin conditions. This could include identifying patterns in hair growth or distribution that might be indicative of certain skin diseases, such as hormonal imbalances or fungal infections.[42]

Occlusion due to hair in dermoscopic images affects the diagnostic operation and the accuracy of its analysis of a skin lesion. Also, dermis hair has the following different characteristics: thin; overlapping; faded; of similar contrast or colour to the underlying skin; and obscuring/covering textured lesions. These make digital hair removal (DHR), which involves hair segmentation and hair gap inpainting, a challenging task. Thus, traditional hard-coded threshold-based hair removal methods are not effective, resulting in over removal which loses important information of the skin lesion, or under-removal which cannot remove the hair effectively. In this paper, we propose a deep learning approach to DHR based on U-Net and a free-form image inpainting architecture. [43]

7. Comparative Study of Existing Hair Segmentation Techniques

Digital hair removal (DHR) algorithms rely heavily on effective hair segmentation, the process of isolating hair pixels from the background skin in an image. This accurate separation is crucial for successful DHR applications. Here's a review of prominent hair segmentation techniques employed in this field [44]:

- **Mathematical Morphology:** This approach leverages predefined shapes (structuring elements) to analyse image features. Pioneering methods like Dull Razor utilize morphological operations to detect dark, thick hair on light skin. However, these techniques struggle with thin or highly curled hair with low contrast. [44]

- **Thresholding-Based Techniques:** This category encompasses methods like adaptive thresholding and threshold set representation. Koehoorn et al.'s method employs a threshold set representation, converting the image into multiple luminance levels. A gap-detection algorithm then identifies thin structures (hair) across these levels. However, this approach can miss hair due to limitations in setting the optimal threshold for all image variations. [45]

- **Edge Detection Methods:** These techniques rely on identifying hair based on sharp intensity changes (edges) in the image. While effective for some cases, they might struggle with blurred or low-contrast hair, potentially leading to segmentation errors. Techniques like the Prewitt edge detector and adaptive Canny edge detector identify hair based on sharp intensity changes (edges). However, these methods struggle with:

    i. **Blurred or low-contrast hair:** This can lead to missed hair pixels during segmentation.
    ii. **Over-segmentation:** High-contrast features on the skin can be misidentified as hair edges, resulting in inaccurate segmentation. [46]
- **Matched Filtering:** This method utilizes pre-defined hair templates to identify similar patterns within the image. This approach can be effective for specific hair types but might not generalize well to diverse hair characteristics, limiting its broader applicability in DHR. This approach utilizes pre-defined hair templates to find similar patterns within the image. While effective for specific hair types, it faces limitations:
    i. **Limited Generalizability:** Templates might not work well for diverse hair characteristics, hindering broader application in DHR.
    ii. **Computational Cost:** Some matched filtering techniques require extensive parameter adjustments for each image, increasing processing time. [47]

- **Other Techniques:** Additional methods include morphological top-hat transforms and simple linear iterative clustering (SLIC) combined with isotropic nonlinear filtering (INF). These techniques aim to enhance the image and facilitate hair extraction but might have limitations in dealing with complex skin textures or varying hair colours, potentially affecting segmentation accuracy. [48]

These limitations highlight the need for ongoing research in hair segmentation for DHR. Advancements in these areas are crucial for achieving more robust and accurate hair segmentation, ultimately leading to improved performance in DHR applications. The selection of a hair segmentation method in DHR depends on several factors, including hair characteristics, skin tone variations, and the overall image quality. Each method has its advantages and limitations. Ongoing research strives to develop more robust and versatile techniques for accurate hair segmentation, ultimately leading to improved performance in DHR applications. [49]
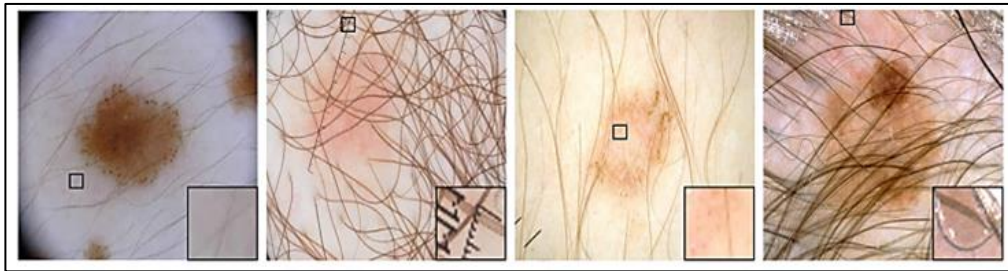

Figure 9: Examples of changing dermoscopic images (where the square inset at the bottom right corner of each image is the zoomed-in enlargement of part of the image denoted by the smaller square), i.e., melanoma

Table 1. Methods for hair segmentation and hair gap inpainting.

Method	Hair detector	Inpainting method	Test Images	Colour Space
Dull Razor [8]	Generalized morphological closing	Bilinear interpolation	5	Grey
Huang et al [9]	Multiscale matched filters	Median filtering	20	CIE Lab*
Xie et al. [10]	Adaptive threshold	-	125	CIE Lab*
Xie et al. [11]	Top-hat operator	PED-based	40	Grey
E-shaver [13]	Prewitt edge	Colour averaging	5	CIE Lab*
Schmid-Saugeon et al.[14]	Morphological closing operator	Median filtering	200	CIE Lab*
Fiorese et al. [15]	Top-hat operator	PED-based	20	Grey
Abbas et al.[16]	Derivative of Gaussian	Coherence transport	100	CIE Lab*
Koehoorn et al. [18]	Multiscale skeleton Morphological operators	Fast marching	>300	HSV
Nguyena et al.[4]	Universal matched filter	-	-	RGB
Zhou et al. [20]	Line detection Curve fitting	Exemplar-based method	460	CIE Lab*

Du et al.[7]	Top-Hat transform multi-scale curvilinear Matched filtering	-	60	YUV
P. Bibiloni et al.[22,23]	Soft colour top-hat transforms	Soft colour morphology	2	CIE Lab*
Mohamed et al. [5]	Hybrid network	-	>300	RGB

```
[ ] def remove_hair_with_mask(image, mask):
        mask = mask/255
        mask = (mask > 0.03).astype(np.uint8)

        # Use inpainting to fill in the masked regions
        inpaint_radius = 3
        inpaint_method = cv.INPAINT_TELEA  # cv.INPAINT_NS or INPAINT_TELEA
        hair_free_image = cv.inpaint(image, mask, inpaint_radius, inpaint_method)
        hair_free_image = cv.inpaint(hair_free_image, mask, inpaint_radius, inpaint_method)

        return hair_free_image
```

Figure 10: Code snippet

```
hair_free_image = remove_hair_with_mask(image, mask)
plt.figure(figsize=(10, 5))
plt.subplot(1, 3, 1)
plt.title('Original Image')
plt.imshow(cv.cvtColor(image, cv.COLOR_BGR2RGB))
plt.axis('off')

plt.subplot(1, 3, 2)
plt.title('Hair Mask')
plt.imshow(mask, cmap='gray')
plt.axis('off')

plt.subplot(1, 3, 3)
plt.title('Hair-Free Image')
plt.imshow(cv.cvtColor(hair_free_image, cv.COLOR_BGR2RGB))
plt.axis('off')

plt.show()
```
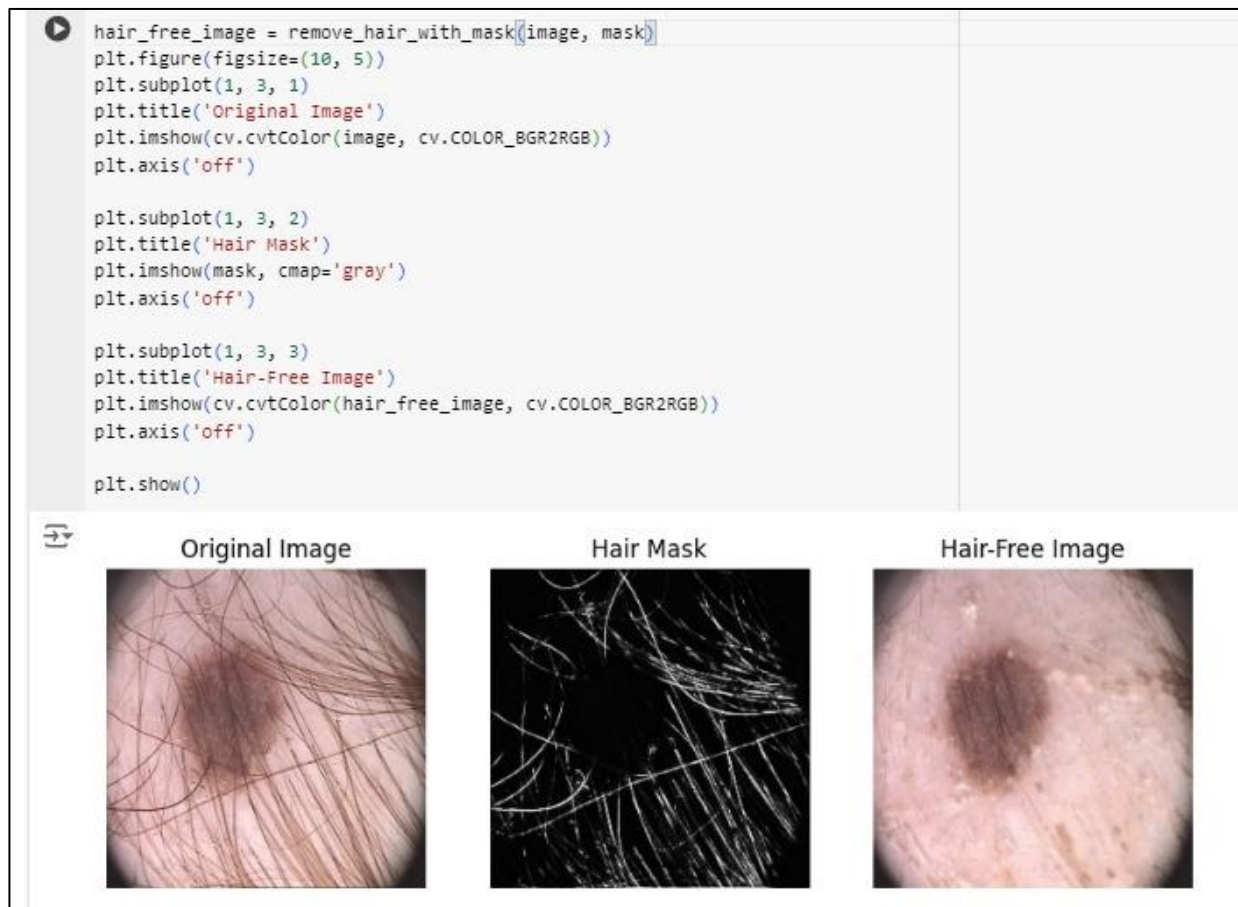

Figure 11: Code snippet

Although many digital hair removal (DHR) algorithms have been developed to address the above-mentioned problem, these algorithms still have two major problems unresolved: over-removal (false

positive) and under-removal (false negative) of the occluding hairs. There are two main reasons for these two problems. First, the artifacts of dermoscopic images are diverse. These artifacts can be natural, such as hair and veins with a wide range of characteristics, or can be external, such as bubbles and ruler markers [5]. These artefacts confuse the processes of melanoma border detection and feature extraction [6]. Second, hard-coded threshold-based algorithms are not generalised and cannot cope with these changing characteristics. Generally, these algorithms primarily solve the removal of only dark thick hair on comparatively light skin [7]. Therefore, a solution that can account for the subtle differences in the pixel values and has a strong generalisation ability is needed. Deep learning has been successfully used to solve many difficult computer-vision problems and deep learning models are resilient to high variability in skin images. Thus, they can generalise to different hair types and artefacts in images [5].

8. Derm.AI: Our Approach for Hair Removal using AI

The incidence of melanoma, the most serious form of skin cancer, has been increasing rapidly in recent years. Early diagnosis is crucial for successful treatment. Dermoscopy, a reliable medical technique, utilizes specialized devices to examine the skin and detect melanoma. With advancements in digital imaging, high-quality images of these examinations can now be captured and stored. These images are being standardized and used for automated melanoma detection. However, the presence of hair on the skin poses a challenge to accurate diagnosis. Thus, it is essential to remove hair to obtain precise results. In this paper, we propose a simple yet effective method for hair removal using deep learning. Our approach leverages the convolutional neural networks (CNN) to reconstruct hair-free images. A CNN architecture specifically designed for analysing skin conditions and hair removal needs would be developed. This would involve defining the number and type of convolutional layers, pooling layers, and final classification layers. Once trained, the CNN could analyse new images of skin conditions with body hair. It would identify the skin condition, hair characteristics, and based on its learned patterns, recommend the most suitable hair removal method for optimal treatment. To evaluate our proposed model, a dataset comprising both hair-covered and hairless images is required. As such a dataset does not currently exist, we introduce a new dataset called Modified-PH2 (M-PH2), inspired by the scientifically curated dermoscopic dataset PH^2. Experimental results demonstrate the improved performance of our technique on the M-PH^2 dataset. Furthermore, we employ various evaluation metrics including Peak Signal-to-Noise Ratio (PSNR), Mean Squared Error (MSE), Structural Similarity Index (SSIM), and Multiscale Structural Similarity Index (MS-SSIM) to assess our model's effectiveness. Through experiments conducted on the publicly available M-PH2 dataset, our proposed method demonstrates high efficiency in hair removal, enhancing the accuracy of skin disease diagnostics compared to other existing methods.

However, the lack of well-labelled dataset hinders the development of deep learning hair removal techniques. The contributions of this paper are as follows. We solved: (1) the problem of the lack of hair mask dataset; (2) the problem of poor hair removal; and (3) the difficulty in evaluating the hair removal effect of a single image. Although hair annotation is a time-consuming and tedious task, we have carefully prepared a well-labelled hair mask dataset, named **Digital Hair Dataset**. This dataset is used to train the encoder-decoder segmentation network. The segmentation hair masks produced by the U-Net are then fed to a free-form image inpainting architecture to inpaint the hair gaps. We used Structural Similarity Index (SSIM) which is designed to evaluate the effect of hair removal of a single image, and use it as a stoppage criterion for DHR. Finally, the input hair image and the inpainted image are presented as inputs to five segmentation algorithms, to demonstrate the importance of accurate hair removal. To our knowledge, this is the first time that a well-labelled hair mask dataset is used to train a convolutional network, and the first time that deep learning is used for hair gap inpainting.

9. Hair Gap Inpainting

Hair gap inpainting constitutes the second phase of DHR. Key methodologies employed in this phase encompass interpolation, median filtering, exemplar-based techniques, PDE-based methods, color averaging, coherence transport, and the fast-marching method [5,6]. One approach involves [7] replacing hair pixels through bilinear interpolation followed by smoothing the final result with an adaptive median filter. A similar interpolation method is adopted by other researchers. Another technique involves replacing hair pixels by interpolating within each channel of the RGB colour space [9]. Median filtering is also utilized by various methods for hair gap inpainting [14]. Nevertheless, median filtering can remove essential subtle elements of the original images, rendering the lesion indistinct. An exemplar-based method is used to restore hair gaps through batch sorting based on certain confidence measures [20]. This method demonstrates high efficiency in sample-based texture synthesis and achieves superior performance, while also accounting for the constraints of the surrounding linear structures in the image.

However, this method necessitates the adjustment of several parameters, which can be challenging to determine in general [16]. Xie et al. [11] and Fiorese et al. [15] both apply partial differential equation-based (PDE-based) methods for inpainting.

However, this iterative method cannot guarantee the convergence of the involved parameters [5]. E-shaver [13] utilizes colour averaging, which is faster than bilinear interpolation, but the repair traces are often conspicuous and unnatural in the inpainted image. All these inpainting methods frequently introduce undesirable blur and distort the texture of the tumour. Additionally, these methods are associated with high computational costs.
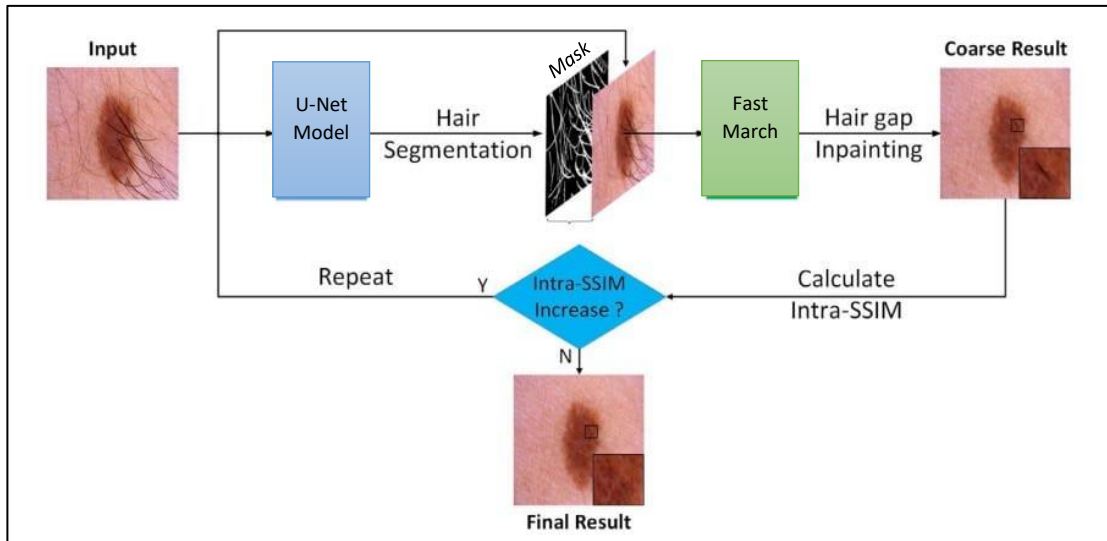

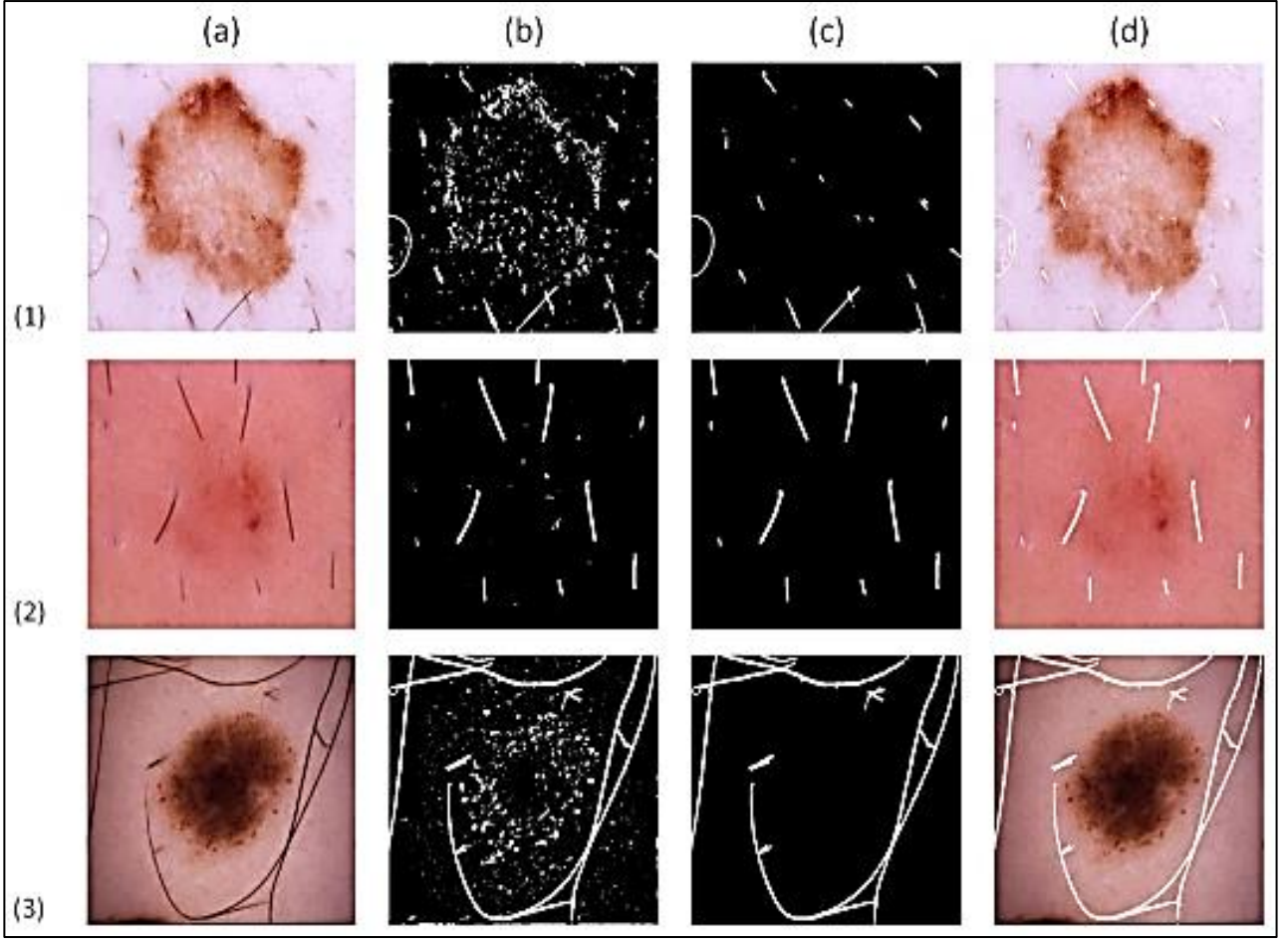*Figure 12: Overview of the proposed framework with U-Net and Fast marching for inpainting*

Figure 13: Examples of creating Digital Hair Dataset column wise; Column (a)ModPH2. Column (b)outputs of black top-hat. Column(c)ground truth created by removing over segmented regions. Column(d)ground truth overlaid on the original images.

10. Proposed Methods

Deep learning has significantly impacted cancer diagnosis due to its ability to represent data with multiple levels of abstraction [26,27]. In this paper, we introduce a digital hair dataset and apply deep learning techniques to remove hairs from skin images. Specifically, we propose three innovative methods for digital hair removal (DHR), which are used to segment hair, inpaint hair gaps, and evaluate the effectiveness of hair removal.

A semantic segmentation framework based on U-Net [28] is employed to segment hair in various conditions. Finally, hair gaps are inpainted using OpenCV's image inpainting technique based on the fast-marching method.

10.1.   Creating a Digital Hair Dataset

Training a U-Net model for effective hair segmentation relies heavily on a robust and well-curated training dataset. This section details the meticulous process undertaken to create a Digital Hair Dataset specifically tailored for this purpose.

Leveraging Existing Resources: The initial step involved acquiring hair masks from a relevant source. 200 lesion images containing hair were selected from the PH^2 dataset, a recognized repository of medical images [23]. These images were then processed using existing black top-hat segmentation techniques, a method known for its effectiveness in extracting hair regions.

Ensuring Precision Through Refinement: The acquired hair masks, while a starting point, often contained over-segmentation errors. These errors occur when areas not containing hair are mistakenly identified as hair during the initial segmentation process. To address this, a team of four experts meticulously evaluated each hair mask. Over-segmented regions were then carefully removed, resulting in a more precise representation of actual hair pixels.

Verification and Quality Control: Following refinement, a crucial step involved verifying the accuracy and correctness of the newly created hair masks. This was achieved by overlaying the refined hair masks onto their corresponding original images. This meticulous verification process allowed for the identification and resolution of any discrepancies between the work of different team members, ensuring a high degree of consistency within the dataset.

Prioritizing Accuracy and Efficiency: The decision was made to exclude under-segmented images (those missing hair pixels) from the final dataset. While it might seem preferable to include all available data, manual correction of under-segmented images is a significantly more time-consuming and error-prone process. The focus on excluding under-segmented images ensured a more efficient dataset creation process while prioritizing the accuracy of the final data.

Promoting Accessibility and Collaboration: The final stage involved uploading the verified hair masks to Kaggle, a prominent platform for sharing data resources with the research community. This open access approach fosters replicability, allowing other researchers to utilize the dataset for their own projects. Furthermore, it encourages collaboration and knowledge sharing within the field of hair segmentation research, ultimately accelerating advancements in this domain.

By following these rigorous steps, the Digital Hair Dataset provides a valuable foundation for training U-Net models for accurate hair segmentation tasks. This dataset not only ensures high accuracy in hair identification but also promotes collaboration and knowledge sharing within the research community, paving the way for further development in this field.
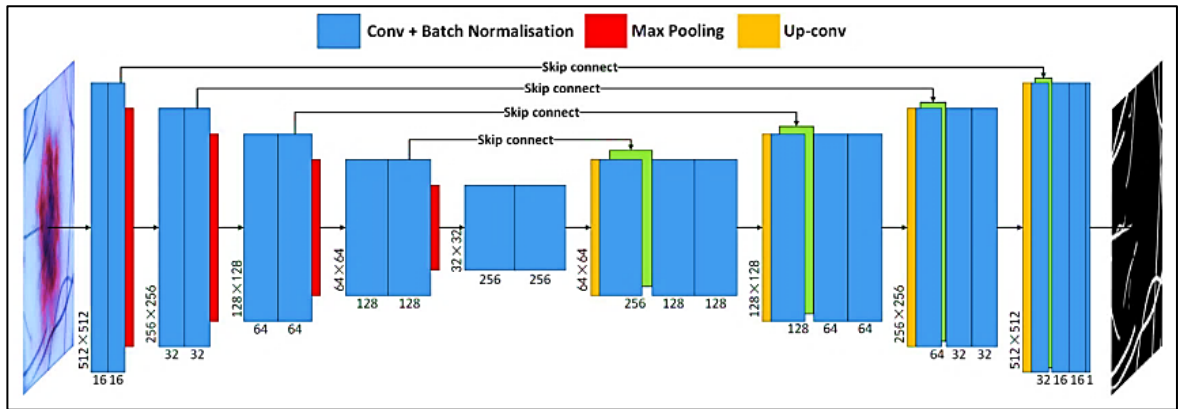

Figure 14:The U-Net architecture. Each blue box corresponds to a multichannel feature map obtained by convolving with the corresponding filter kernels. The number of channels involved and the x-y-size of the map are denoted at the bottom and at the lower left edge box

10.2.   U-Net Architecture for Hair Segmentation

Semantic image segmentation seeks to assign each pixel of an image to its corresponding class. In our context, it serves to categorize pixels in a skin image into two classes: hair and skin, enhancing comprehension of the image's overall context [5]. While traditional computer vision techniques remain prevalent, deep learning architectures, notably Convolutional Neural Networks (CNNs), have emerged as superior alternatives in terms of both accuracy and efficiency [21]. Consequently, a plethora of deep learning methodologies have been developed for medical image segmentation.

U-Net stands out as a prominent semantic segmentation framework for CNNs, proficient in automatically learning image features. Widely applied in medical image analysis for lesion segmentation, anatomical segmentation, and classification, U-Net exhibits the ability to accurately segment target features, process and evaluate medical images objectively, and enhance diagnostic accuracy [22]. Notably, even in scenarios with limited training data, such as those encountered in hair segmentation due to the scarcity of available datasets, U-Net maintains its effectiveness.

The architecture of U-Net comprises two distinct paths. The first, known as the contraction path or encoder, focuses on capturing contextual information within the image. It iteratively applies two 3x3 convolutions, each followed by a rectified linear unit (ReLU), alongside 2x2 max pooling operations with a stride of 2, progressively reducing dimensions from 512x512x3 to 32x32x256. The second path, the symmetric expanding path or decoder, facilitates precise localization through transposed convolutions, expanding dimensions from 32x32x256 to 512x512x1. Each step in the expanding path entails up sampling of the feature map, followed by a 2x2 transposed convolution. These convolutions halve the number of feature channels through concatenation with the corresponding feature map from the contracting path, followed by two 3x3 convolutions, each employing a ReLU. Ultimately, a 1x1 convolution, succeeded by a sigmoid function, maps each 16-component feature vector to the desired number of classes. Notably, operating as an end-to-end fully convolutional network (FCN), U-Net comprises 23 convolutional layers [28].

Convolutions and pooling operations serve to down sample the image, with max pooling enlarging the receptive field to discern the image's content. However, for semantic segmentation, understanding not only what but also where elements exist in the image is crucial. Hence, transposed convolutions, which acquire parameters through backpropagation, are preferred for up sampling. To achieve better precision at each step of the decoder, skip connections are employed, concatenating the output of transposed convolution layers with feature maps from the encoder at corresponding levels. Subsequent consecutive regular convolutions refine the model's ability to generate a more precise output [28].

Training of the model entails utilizing Adam as the optimizer and the Dice coefficient loss as the loss function. The learning rate undergoes reduction to 0.1 times its preceding value if the validation loss (Dice coefficient loss) shows no decline after five consecutive iterations, setting a minimum limit at 0.0001. This optimization strategy ensures the model's rapid and accurate generalization. Furthermore, an early stopping mechanism, terminating training after 50 consecutive iterations without a reduction in the Dice coefficient loss, signifies the model's attainment of optimal training. The hair segmentation outcomes are expounded upon in Section 5.1.

10.3. Transfer learning for hair gap inpainting

Transfer learning is employed for hair gap inpainting following the segmentation of hair. Given the scarcity of skin images, we utilize a transfer learning (TL) technique based on Gated convolution to inpaint the hair gap. While deep learning has demonstrated remarkable success in image inpainting [29], its application to hair gap inpainting has been hindered by the absence of a relevant dataset. We are pioneers in addressing this challenge. Hair gap inpainting necessitates a substantial collection of hair-free images, yet the current dataset falls short of providing adequate training for satisfactory results. To mitigate this limitation, we turn to transfer learning. This technique involves leveraging knowledge acquired from solving one problem and applying it to a different but related problem. By doing so, transfer learning offers a more efficient and expedited solution, requiring less effort to gather the necessary training data and retrain the model. Prior to presenting our proposed method, we briefly outline the fundamental concepts and definitions of transfer learning [23].
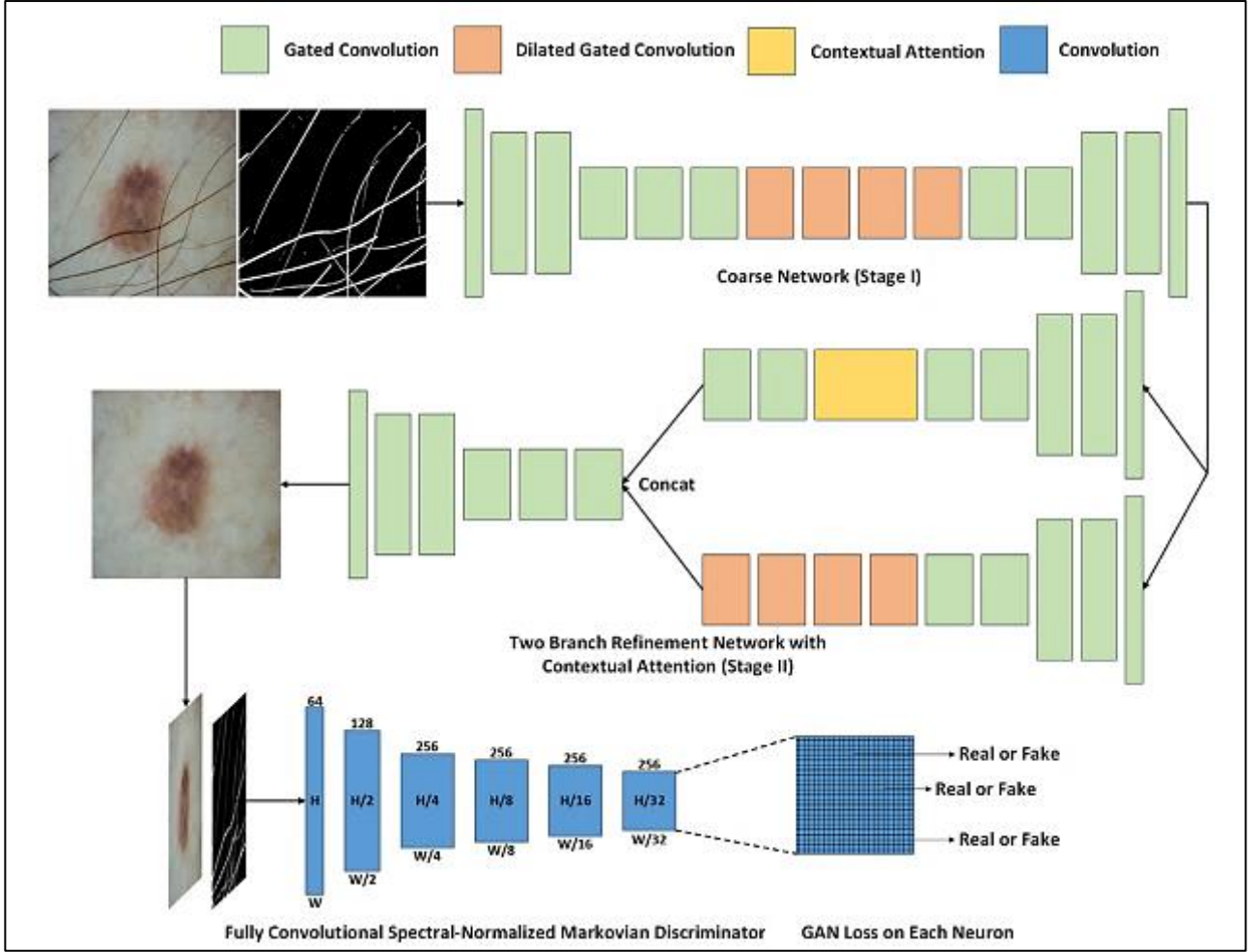
Figure 15: Overview of the framework with gated convolution and SN-Patch GAN for free form image inpainting

Notation. A domain D consists of two components: a feature space X and a marginal probability distribution P(X), in which $X = \{x_1, \ldots, x_n\} \in X$. Given a specific domain $D = \{X, P(X)\}$, a task consists of two components: a label space Y and an objective predictive function f(·), denoted by $T = \{Y, f(\cdot)\}$, which is learned from the training data pairs $\{x_i, y_i\}$, where $x_i \in X$ and $y_i \in Y$. The function f(·) is used to predict the corresponding label, f(x), of a new instance x.

Definition. Given a source domain DS and learning task TS, a target domain DT and learning task TT , TL aims to help improve the learning of the target predictive function f(T (·)) in DT using the knowledge in DS and TS, where $DS \neq DT$, or $TS \neq TT$.

The idea of TL is to use a network trained on a different domain for a different source task, and adapt it for another domain and target task. In this paper, due to the randomness of the hair shape, we use the pre-trained free-form image inpainting system proposed by Yu et al. [29] and trained on the CelebA-HQ (Celebrity face) dataset. The system combines Gated convolution and the inpainting results demonstrate that this system is feasible for hair gap inpainting. The Gated convolutions of the system learn from data through soft mask layers rather than hard-gated mask employed in traditional networks. The process involves:

$Gating_{y,x} = \sum \sum W_g \cdot I$ (1)

$Feature_{y,x} = \sum \sum W_f \cdot I$ (2)

20

$$O_{y,x} = \varphi(Feature_{y,x}) \; (\circ) \; \sigma \; (Gating_{y,x}) \quad (3)$$

where x, y represents the resolution of the output map, I and O are respectively the inputs and outputs, σ is sigmoid function which makes the output gating value between zero and one, φ can be any activation functions (e.g., ReLU, ELU and LeakyReLU), and finally W_g and W_f denote the two different convolutional filters [29]. In short, Gated convolution generalises partial convolution by providing a learnable dynamic feature selection mechanism for each channel at each spatial location across all layers. It also solves the issue of vanilla convolution which treats all input pixels as valid ones [29]. As free-form masks may appear anywhere in images with any shape, Yu et al. present a patch-based loss by applying spectral-normalised discriminator on dense image patches [29]. Interpolation can inpaint multiple holes with any shape at any location and is different from other inpainting networks which normally work by inpainting single rectangular holes. Therefore, this method is suitable for inpainting a hair gap which has random shapes, and the results of hair gap inpainting are presented in our experiments.

10.4.   Intra-SSIM for DHR evaluation

Through our experimentation, we observed that hair segmentation and hair gap inpainting require multiple iterations to achieve complete removal of all hair strands. Interestingly, when we input the results of the first iteration of hair gap inpainting into U-Net, the few remaining hair strands that were undetected in the initial iteration can be successfully segmented out in the subsequent iteration. Moreover, by repeating this iterative process, we were able to identify even minute hair strands, thereby enhancing the overall performance of the system. However, determining the optimal number of iterations, which encompass hair segmentation and hair gap inpainting, necessitates a reliable evaluation method. To address this, we propose a technique called Intra-SSIM, designed to evaluate the results of Digital Hair Removal (DHR) based on a single image without the need for a reference hair-free image.

Our proposed evaluation method relies on the structural similarity (SSIM) index, extensively detailed in the literature [25]. In summary, SSIM is based on three comparison measurements between samples x and y: luminance (l), contrast (c), and structure (s). Specifically, let μ_x represent the mean of x, σ_x^2 denote the variance of x, and σ_{xy} signify the covariance of x and y.
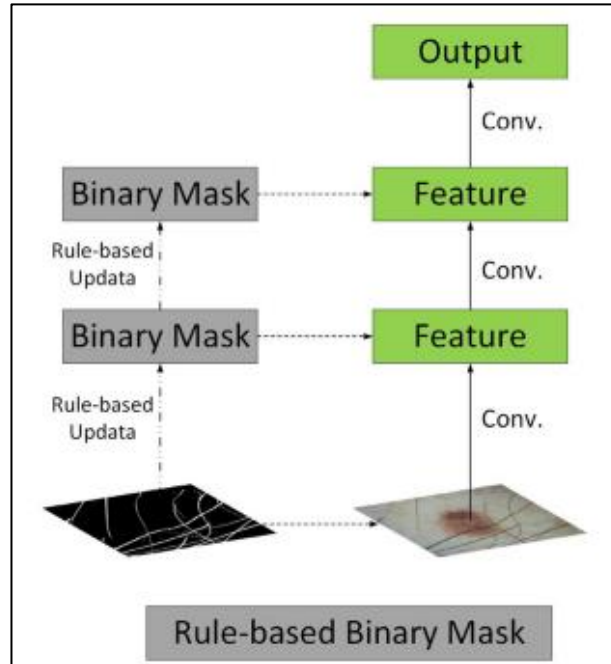

Figure 16: Partial convolution: un-learnable single-channel feature hard-gating

The comparison measurements are:

$$l(x, y) = \frac{2\mu_x\mu_y + c1}{\mu_x^2 + \mu_y^2 + c1}, (4)$$

$$c(x, y) = \frac{2\sigma_x\mu_y + c1}{\sigma_x^2 + \sigma_y^2 + c1}, (5)$$

$$s(x, y) = \sigma xy + c3 \; \sigma x\sigma y + c3 \;, (6)$$

where c1, c2 and c3 are small constants respectively given by

$c_1 = (k_1L)^2$, $c_2 = (k_2L)^2$ and $c_3 = c_2/2$, (7)

where L is the dynamic range of the pixel values (L = 255 for 8 bits/pixel grey scale images), and k1 = 0.01 and k2 = 0.03 are two scalar constants. The general form of the SSIM index between signal x and y of common size NxN is defined as

$$SSIM(x, y) = [l(x, y)^\alpha \cdot c(x, y)^\beta \cdot s(x, y)^\gamma], (8)$$

where α, β and γ are parameters to define the relative importance of the three components. Specifically, we set $\alpha = \beta = \gamma = 1$, and the resulting SSlM index is

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, (9)$$

which satisfies the following conditions:

1. Symmetry: SSIM(x, y) = SSIM(y, x);

2. Boundedness: SSIM(x, y) ≤ 1;

3. Unique maximum: SSIM(x, y) = 1 if and only if x = y.

The SSIM index is used for measuring the similarity between two images. It is a full reference metric, i.e., used for the measurement or prediction of image quality based on an initial uncompressed or distortion-free image as reference. However, in our case each hair image does not have the corresponding hair free image as a reference. Therefore, it is not feasible to use SSIM index to evaluate the image after DHR. To compute Intra-SSIM we first divide the image with the size of NxN pixels into nxn cell blocks. Let b = {bm|m = 1, 2, . . ., nxn − 1} be the cell block, and let B = {B(i, j)|i, j = 1, 2, . . ., n − 3} represent a 3x3 cell blocks where each cell block is (N/n)x(N/n) pixels. We then calculate the SSIM index between the central cell block and the remaining cell blocks in eight directions of each B(i, j), i.e., SSIM_0, SSIM_1, SSIM_2 SSIM_3, SSIM_4, SSIM_5, SSIM_6, SSIM_7, and calculate the average value, i.e.,

$$B'(i, j) = \frac{1}{8} \sum_{m=0}^{7} SSIM\_m. (10)$$

Similarly, the sliding window technique with a stride of 1 along the row and column of the image is employed to calculate the Intra-SSIM of the whole image, i.e.,

$$Intra\text{-}SSIM = \frac{1}{(n-2)^2} \sum_{j=0}^{n-3} \sum_{j=0}^{n-3} B'(i, j). (11)$$

In this paper, we divide the image of size 512x512 pixels into 32x32 cell blocks, so that each cell block comprises 16x16 pixels. The proposed method does not require a reference image and presents a new evaluation criterion based on a single image to examine the effects of the inpainting areas (e.g., hair gap) and the background (e.g., lesion).

11. Experiments

11.1.   Implementation

We implemented the proposed Digital Hair Removal (DHR) algorithm on Google Collaboratory, utilizing its computational capability of a 3.7 Tesla K80 GPU. Google Collaboratory is a free, cloud-based Jupyter notebook environment that facilitates the training of deep learning models on CPUs, GPUs, and TPUs. For comparison, the algorithms by Huang et al. [9] and Xie et al. [10] were implemented on MATLAB 2016b, running on a system with an Intel Core i5-4210U 1.7 GHz processor. DullRazor [8], another method referenced, is a standalone window software. Notably, there was no publicly available implementation for the other methods discussed in this paper. The evaluation criteria for these methods were also implemented on the same computer for consistency.
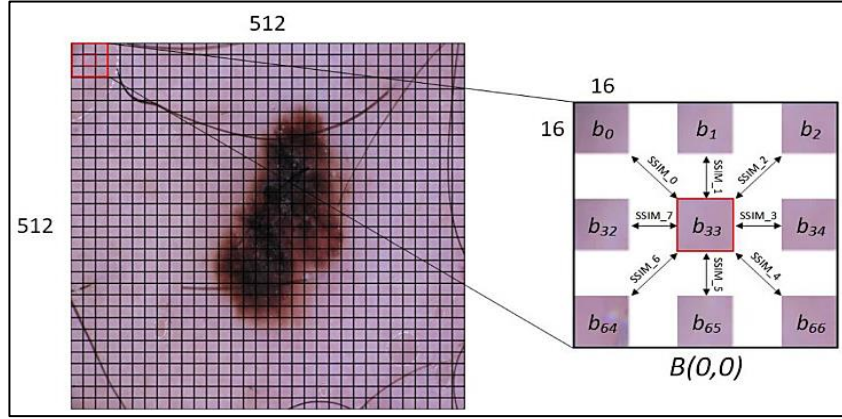

Figure 17:Left shows the image with size of 512x512 pixels divided into 32x32 cell blocks. At each step of the sliding window, the mean SSIM is computed for central block considering the eight neighbouring blocks as shown on right.

Table 2. The three steps of our experiment used the same testing data, where TL denotes transfer learning.

Procedure	Training	Validation	Testing	Image size
Hair segmentation	140	20	40	512x512
Hair gap inpainting	140	20	40	512x512
Lesion segmentation	140	20	40	256x256

11.2.   Datasets

Our data was sourced from the PH2: Skin Lesion Analysis Towards Melanoma Detection grand challenge datasets. This dataset includes 200 images and their corresponding ground truth response masks for the lesion boundary segmentation task. Following this, we created 200 corresponding ground truth hair masks as detailed. The same test data was used for hair gap inpainting. For the skin lesion segmentation task, we randomly selected 140 images for training and 20 images for validation, using the same set of test images as in the previous tasks. In terms of pre-processing, images for hair segmentation and hair gap inpainting were resized to 512x512 pixels to maintain consistency and accuracy. For skin lesion segmentation, all images were resized to 256x256 pixels to ease computational requirements. The data usage and distribution are summarized above.

```
[ ] train_data = torchvision.datasets.ImageFolder(data_path)

[ ] train_data

⤶  Dataset ImageFolder
        Number of datapoints: 200
        Root location: /content/PH2-dataset/images
```

Figure 18: Code snippet

11.3.   Digital hair removal (DHR)

For improved Digital Hair Removal (DHR), we iterated the two steps of hair segmentation and hair gap inpainting. To determine the optimal number of repetitions, we calculated the average Intra-SSIM index for all testing images after each iteration. To assess whether the block size n affects convergence, the model was tested with n values of 16, 32, and 64. Our experiments indicated that varying n does not impact convergence, as the model consistently achieved a stable Intra-SSIM score after the fourth iteration for all n values. Although the Intra-SSIM score slightly increased with larger n, the average Intra-SSIM value did not improve beyond the fourth iteration. Consequently, the inpainting process terminated after the fourth iteration, rendering it unaffected by block size. Notably, increasing n significantly raised the computation time required to calculate Intra-SSIM. Therefore, a trade-off was achieved by choosing n as 32 for optimal performance.

To evaluate the inpainting procedure using the proposed Intra-SSIM technique, we selected 43 lighter and 43 darker images containing hair over lesions. The careful selection of images ensured the evaluation criterion's effectiveness for lesions with both bright and dark textures. The experimental results demonstrated comparable outcomes for both dark and light hair lesion images. For lighter hair images, the Intra-SSIM values improved from 0.487 to 0.682 after hair removal. Similarly, for dark hair images, the Intra-SSIM scores increased from 0.421 to 0.574 following hair removal. Fig. illustrates examples of lighter and darker hair images before and after hair removal using Intra-SSIM evaluation. Additionally, Fig. highlights the model's effectiveness and visualizes the Intra-SSIM index. It is evident that as the number of iterations increases, the lesion contours become more visible, and the plane fitting reveals the volumetric shape of the lesion. In our experiments, we employed various metrics to evaluate the performance of the segmentation techniques, including accuracy, Dice coefficient, specificity, F1 score, precision, sensitivity, Jaccard index (mean Intersection over Union), and Enhanced-alignment measure (E-measure). Intra-SSIM was used to assess the performance of inpainting, while MSE was employed to evaluate the extent of undesirable changes outside the hair gap within the lesions. To verify the significance of hair removal in skin lesion segmentation, we compared the segmentation results of five state-of-the-art methods on original images and images post hair removal.

12.   Results and comparison

The performance of our system was evaluated both quantitatively and qualitatively across three key tasks: hair segmentation, hair gap inpainting, and lesion segmentation. These results were compared with those generated by three state-of-the-art methods: DullRazor [8], Huang et al. [9], and Xie et al. [10]

12.1.   Quantitative comparison of hair segmentation

The Jaccard index, also known as Intersection over Union, is a statistical measure used to gauge the similarity and diversity of sample sets. It is defined as the size of the intersection divided by the size of the union of the sample sets. The Jaccard coefficient measures the similarity between two finite sample sets and is mathematically expressed as:

J(A, B) = |A∩B| / |A∪B| = |A∩B| / (|A| + |B| - |A∩B|), (12)

Simply put, the Jaccard index is the area of overlap between the predicted segmentation and the ground truth divided by the area of union between the segmentation result and the ground truth.

E-measure is a recently proposed measure that consists of a compact term that simultaneously captures image level statistics and local pixel matching information. It is significantly better than traditional measures, e.g., F1 score. Also, the metric presents an effective and efficient way for evaluating binary foreground maps. E-measure is computed as:

$$Q_{FM} = \frac{1}{W*h} \sum_{x=1}^{w} \sum_{y=1}^{h} \varphi_{FM}(x, y), (13)$$

where φ_{FM} is enhanced alignment matrix, and h and w are the height and the width of the map, respectively.

Table 3 demonstrates that the performance of our hair segmentation method significantly surpasses that of other methods. Our method achieves the highest scores across all evaluation metrics, particularly excelling in the Jaccard index. Specifically, our proposed method achieves a Jaccard index of 95.42 for hair segmentation, compared to 61.88 for DullRazor. A key advantage of the Jaccard index is its ability to penalize both over-segmentation and under-segmentation [5]. Our method enhances the Jaccard index by effectively reducing false positives. Moreover, due to the precise segmentation of hair, our proposed method attains a superior E-measure for hair segmentation, scoring 98.00 compared to 90.94 for DullRazor.

12.2. Hair Gap Inpainting

12.2.1. Qualitative comparison of hair segmentation

Figure 12 presents a performance comparison between the proposed method and three state-of-the-art methods. DullRazor (refer to row 2) effectively removes black thick hair but fails to detect the ruler marker and thick hair in images (2) and (3). The method by Huang et al. (refer to row 3) results in a high rate of false positives in image (2), as it mistakenly identifies hair-like textures as hair. Additionally, it does not fully detect the thick hair in image (3). The method by Xie et al. (refer to row 4) fails to detect short hair in images (1) and (2), particularly in image (1). For image (3), all three state-of-the-art methods incorrectly segment the lesion texture as hair, which significantly impairs the dermatologists' diagnosis of the lesion area. In contrast, the proposed method (refer to row 5) accurately segments all types of hair and ruler markers.
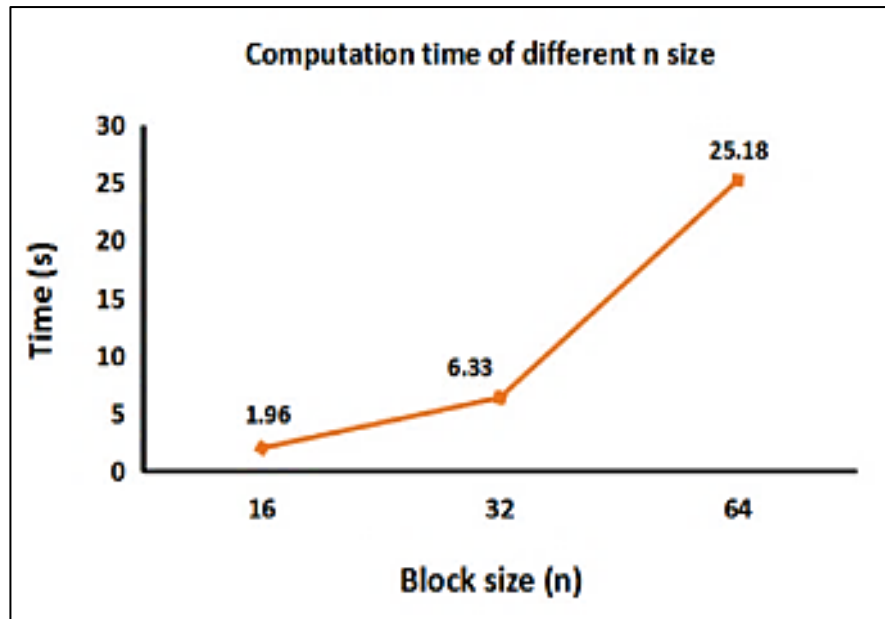
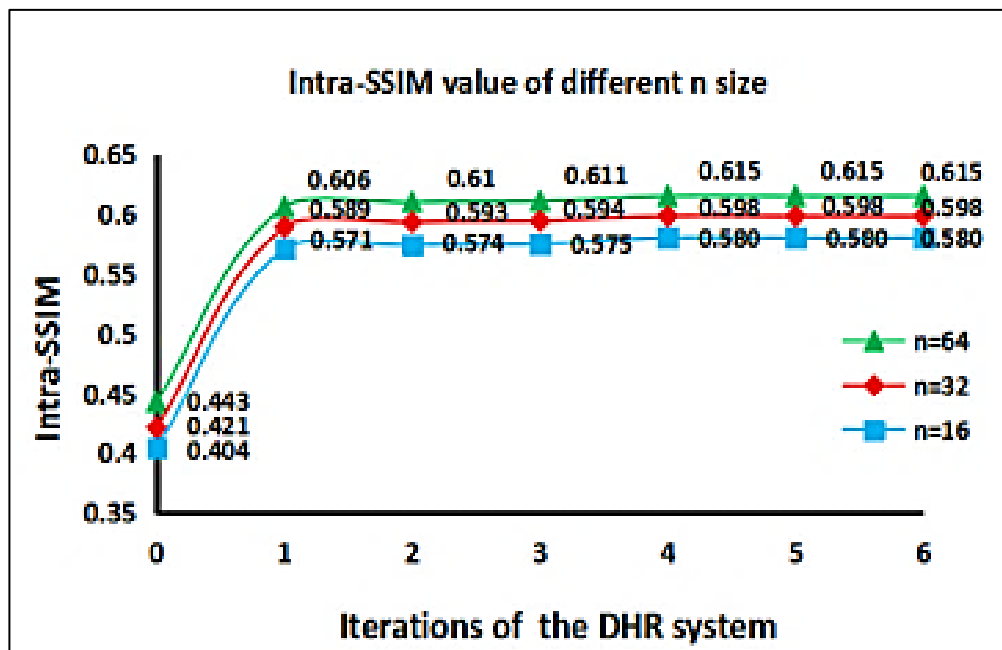*Figure 19: Computation time required for calculation Intra-SSIM for different block sizes n.*

Figure 20:Average Intra-SSIM index computed for all test images with block sizes n of 16, 32 and 64.
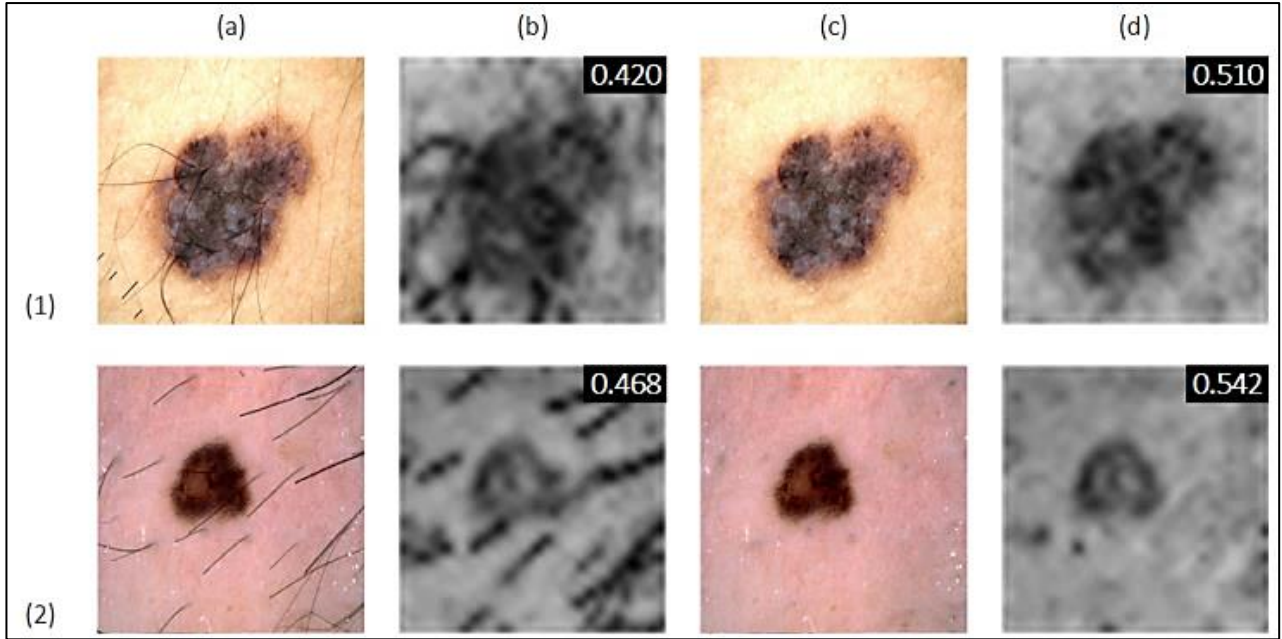
Figure 21: Example of a lighter and darker hair images before and after hair removal using Intra-SSIM evaluation: Rows (1) and (2) respectively show the lighter and dark hair image.

Along the column: (a) and (c) are the original image and the inpainting image, respectively; (b) and (d) are the Intra-SSIM visualizations corresponding to (a) and (c), respectively. The value of the Intra-SSIM of each image is also denoted.

Table 3. Comparison of hair segmentation results of four methods.

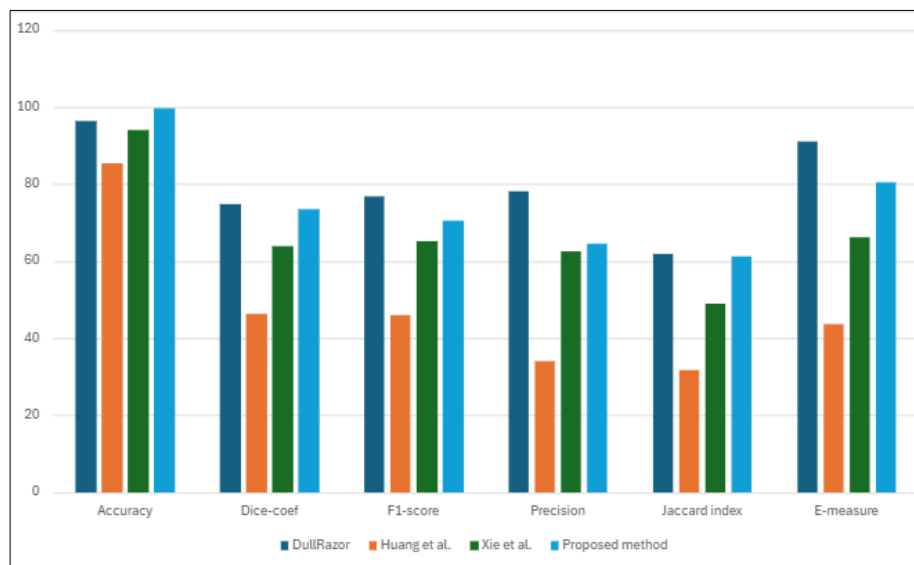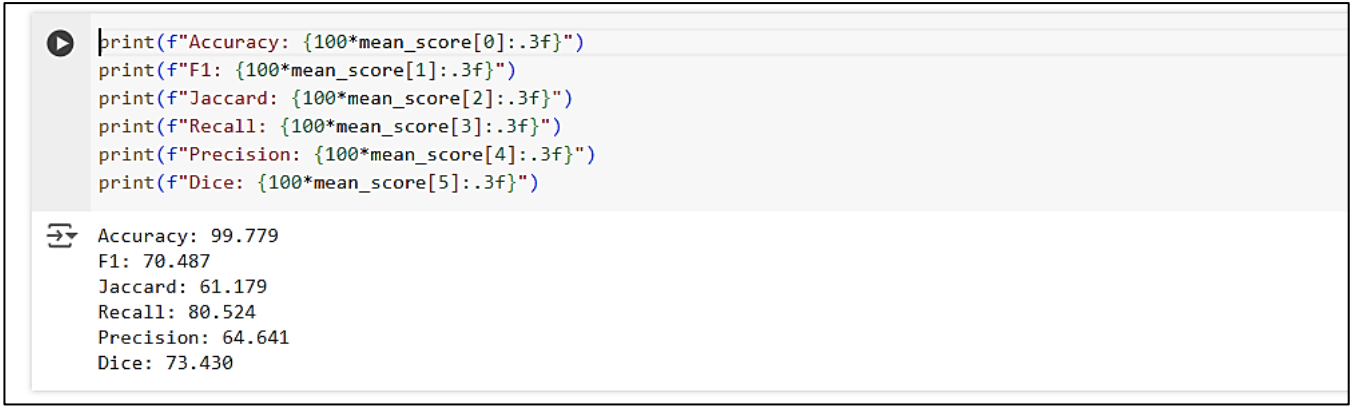| Segmentation method	Accuracy	Dice-coefficient	F1-score	Precision	Jaccard index	E-measure
DullRazor [8]	96.38	74.70	76.88	78.13	61.88	90.94
By Huang et al. [9]	85.58	46.21	46.09	33.89	31.64	43.80
By Xie et al. [10]	93.92	63.72	65.32	62.41	49.09	66.17
Proposed method	99.78	73.43	70.49	64.64	61.18	80.52

Figure 22 Graphical representation of comparative study of hair segmentation methods

```
print(f"Accuracy: {100*mean_score[0]:.3f}")
print(f"F1: {100*mean_score[1]:.3f}")
print(f"Jaccard: {100*mean_score[2]:.3f}")
print(f"Recall: {100*mean_score[3]:.3f}")
print(f"Precision: {100*mean_score[4]:.3f}")
print(f"Dice: {100*mean_score[5]:.3f}")

Accuracy: 99.779
F1: 70.487
Jaccard: 61.179
Recall: 80.524
Precision: 64.641
Dice: 73.430
```

Figure 23: Code snippet

The Intra-SSIM score does not exhibit any further increase beyond a specific point. As detailed in Section 4.3, our experiments indicate that four iterations are necessary to achieve satisfactory hair gap inpainting. This iterative process ensures that the inpainting quality meets the required standards for practical applications. To rigorously evaluate the performance of the hair gap inpainting, we computed the Intra-SSIM score between the inpainted areas and the background for various methods. This score is critical in assessing how well the inpainted regions blend with the surrounding areas of the image. For a realistic and fair comparison, we calculated the inpainting results of each method on the hair test images using hair masks we created as ground truths. These masks provide a consistent reference point, ensuring that each method is evaluated under the same conditions. We then employed the Intra-SSIM metric to assess the internal similarity between the hair gap inpainting results of each method and the background of the original image, as illustrated in Figure 13. This evaluation helps to determine how effectively each method integrates the inpainted regions with the natural texture and colour of the background. Furthermore, to measure any undesirable texture changes within the lesion area caused by poor segmentation and inpainting, we calculated the Mean Squared Error (MSE) of the lesion region between the input image and the inpainted image, excluding the hair gap. This calculation, shown in Figure 14, is crucial in identifying any significant deviations in texture and colour that may arise due to the inpainting process. A lower MSE indicates that the inpainting method maintains the integrity of the lesion area, which is essential for accurate medical imaging analysis. Since Xie et al. did not propose an inpainting method, our comparisons focused on the inpainting results with those of DullRazor and Huang et al. These comparisons provide a comprehensive understanding of the relative strengths and weaknesses of different inpainting techniques, highlighting the effectiveness of our proposed method in achieving realistic and seamless hair gap inpainting. This thorough evaluation underscores the importance of precise inpainting for enhancing the quality and accuracy of hair segmentation in medical images.

Table 4 compares our method with two state-of-the-art methods, demonstrating that our approach achieves the highest Intra-SSIM (indicating better performance) and the lowest MSE (indicating fewer errors). These metrics suggest that the inpainting results of our method are the most similar to the input image, both in hair and non-hair areas. To further illustrate the effectiveness of the inpainting process, as shown in Figure 15, we conducted an additional experiment. We applied a hair mask (or a watermark mask) to a hairless image (or a watermark-free image) and then used different inpainting methods to restore it. Given that DullRazor is an integrated software without customization options, we focused our comparison on Huang's method. Our experimentation revealed that median filtering, a technique used in some inpainting methods, does not allow the inpainted area to blend seamlessly with the background features, resulting in more noticeable inpainting traces. In contrast, our method leverages the background features learned by the model to perform inpainting, ensuring that the inpainted areas are closely integrated with the

28

surrounding background. This ability to adapt to the background features significantly enhances the realism and quality of the inpainting results, as our method effectively minimizes the visibility of inpainting traces and maintains the integrity of the original image's appearance.

Table 4. Comparison of inpainting results by three methods. Columns 3, 4 and 5 show the Intra-SSIM and the MSE for the images shown in Fig. 13 and Fig. 14. Column 6 presents an average score for all the Test images.

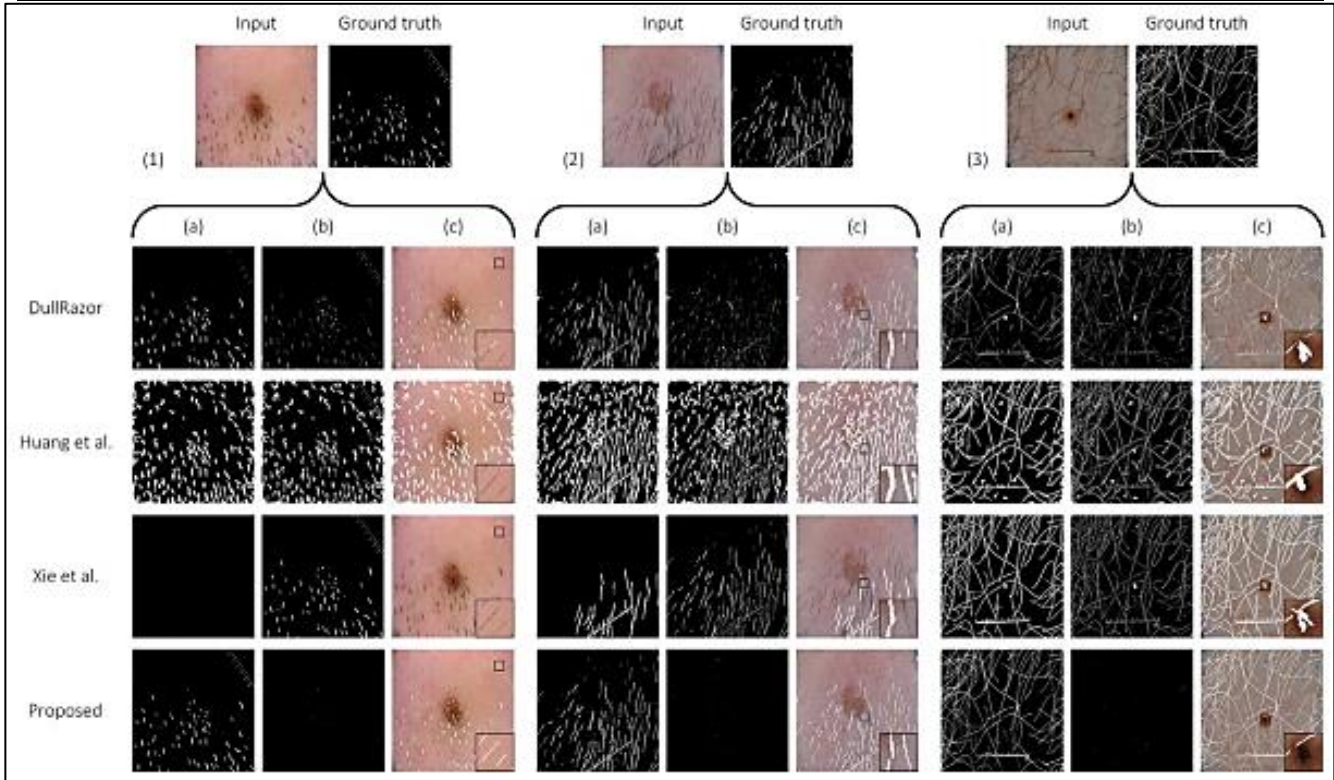| Method	Metrics	Image (1)	Image (2)	Image (3)	Test Images
DullRazor [8]	Intra-SSIM	0.544	0.635	0.511	0.504
	MSE	40.7	39.4	13.8	4.91
Huang et al. [9]	Intra-SSIM	0.551	0.670	0.540	0.566
	MSE	92.6	53.4	63.3	10.27
Proposed method	Intra-SSIM	0.556	0.674	0.546	0.568
	MSE	51.7	46.2	23.6	7.24

Figure 24: Qualitative comparison of hair segmentation between the proposed method and with Dull Razor: Huang et al. method, Xie et al. method

```
# Calculate MSE
mse_value = mean_squared_error(ground_truth_img, pred_image)

# Calculate SSIM
ssim_value, _ = structural_similarity(ground_truth_img, pred_image, full=True)

# Print results
print(f"MSE: {mse_value}")
print(f"SSIM: {ssim_value}")

MSE:  51.75125122070312
SSIM: 0.08945919690679317
```
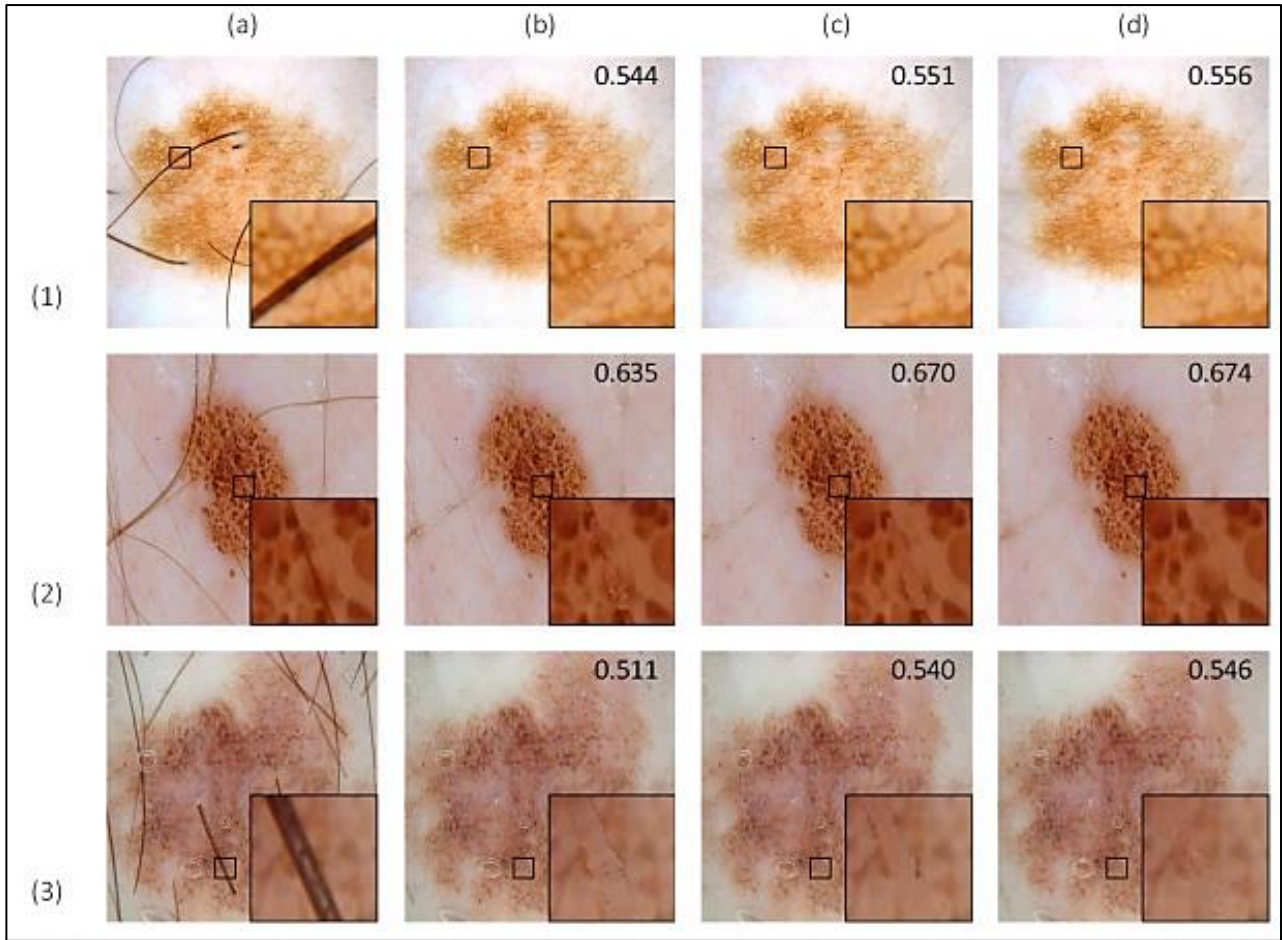
Figure 25: Code snippet

Figure 26: Qualitative comparison of hair gap inpainting results: Column (a) Input images; Column (b) Dullrazor; Column (c) Huang et al.; and Column (d) Proposed.

12.2.2. Qualitative comparison of hair gap inpainting

We carefully selected three images from the dataset to demonstrate the potential of the proposed method. These images contain hair strands that bisect the lesion area, providing a challenging scenario for Digital Hair Removal (DHR) algorithms. An effective DHR algorithm should be capable of removing the hair strands and accurately inpainting the hair gaps without introducing any undesirable effects on the lesion area. In Columns (b) and (c) of Figure 26, the DHR results obtained using Bilinear interpolation and Median filtering are respectively shown. A visual inspection of these results reveals that the inpainted areas fail to match the texture of the input image. This discrepancy is particularly noticeable and highlights the limitations of these inpainting methods. The inpainted regions appear artificial and do not blend seamlessly with the surrounding tissue, which is critical for medical imaging applications where accuracy is paramount. Furthermore, poor segmentation and inpainting significantly alter the texture characteristics of the lesion area, as illustrated in Figure 27. The inadequate handling of the hair gaps results in a blurred appearance of the lesion area, which can obscure important diagnostic features. This blurring effect is detrimental to the diagnostic process, as it can lead to misinterpretations or missed diagnoses. Our method addresses these issues by learning the background features of the image and applying them to the inpainting process. This approach ensures that the inpainted areas are closely integrated with the surrounding background, maintaining the integrity of the original image's texture and appearance. The results indicate that our method can effectively remove hair strands and inpaint the gaps in a way that preserves the important diagnostic details of the lesion area, making it a superior choice for medical image analysis.
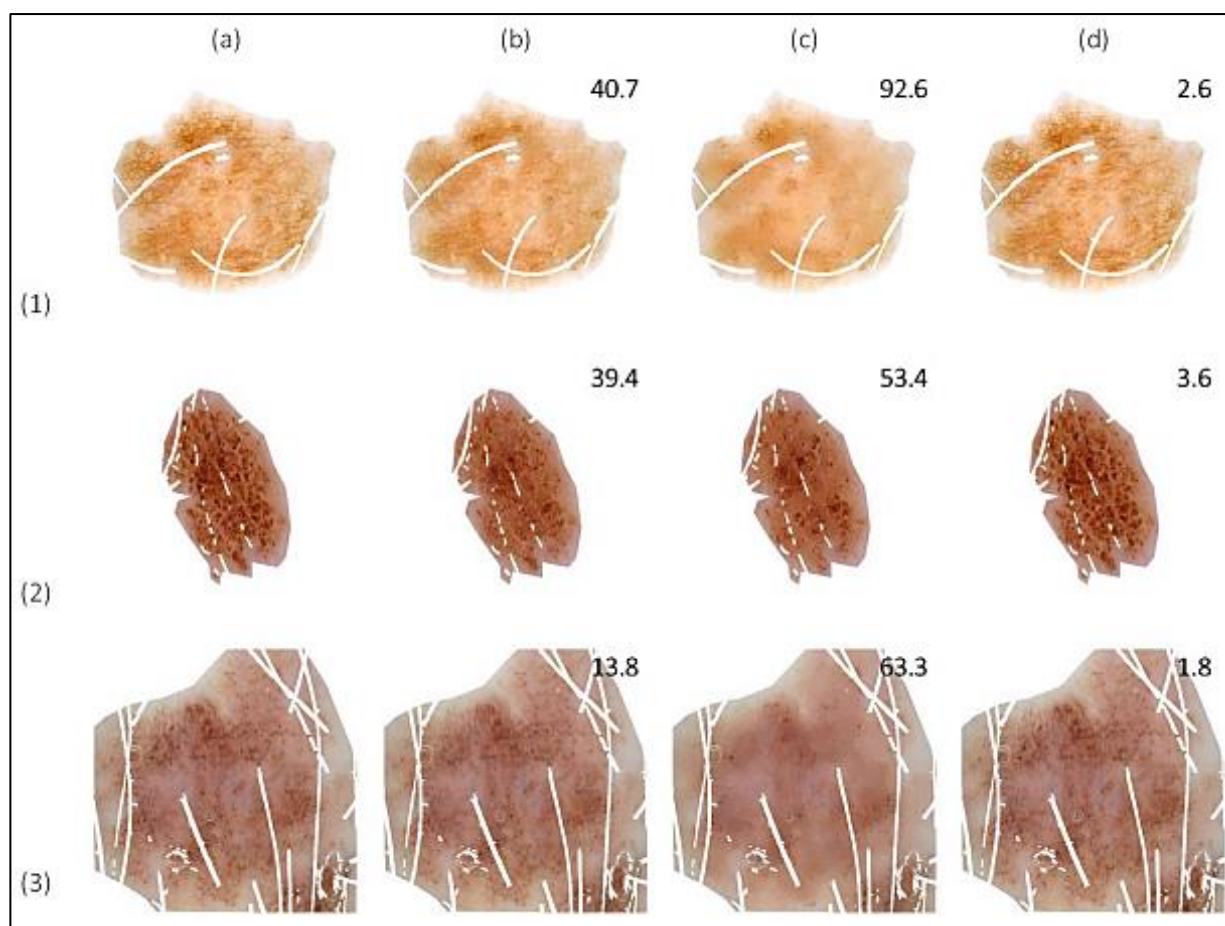
Figure 27: Qualitative comparison illustrating the texture change within a lesion area. Column (a) Lesion of input image (excluding hair gap). The results for inpainting using: Column (b) DullRazor; Column(c) Huang et sl. method; and Column(d) Proposed.

13. Graphical Representation

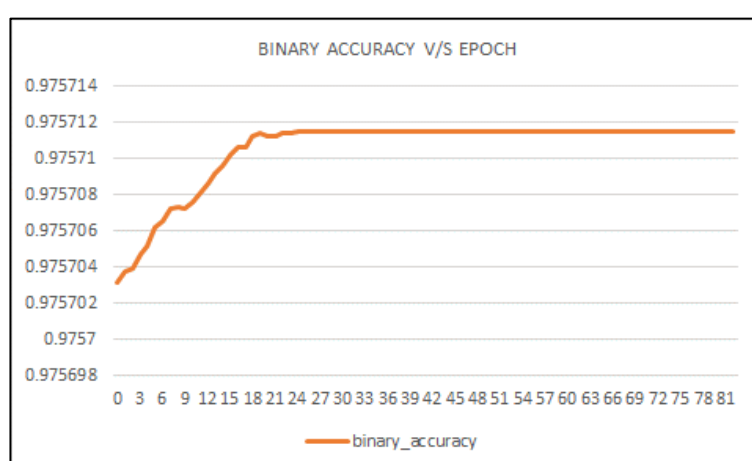## 13.1. Binary Accuracy v/s Epoch

Figure 28 Graph

The graph depicts the binary accuracy of your derm.ai model plotted against epochs during the training process. In the context of skin cancer classification, binary accuracy refers to the proportion of correct predictions made by the model, where there are only two possible outcomes: cancerous or non-cancerous.

It shows the binary accuracy of the machine learning model, over a period of epochs. Epochs are iterations or cycles through a dataset used to train a model. The x-axis is labelled "Epoch" with numerical values from 0 to 81. The y-axis is labelled "Binary Accuracy" with values ranging from 0.9757 to 0.975714. The line on the graph seems to fluctuate slightly but overall increases as the number of epochs increases. This suggests that the model's performance is improving as it is trained on more data. The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model refines its ability to classify skin lesions based on the training data. Essentially, the model is "learning" from the data with each pass. The y-axis labelled "Binary Accuracy" shows values ranging from 0.9757 to 0.975714. While the change seems minimal, this slight increase in accuracy over a relatively small range (already high) suggests the model is making small improvements as it iterates through the training data. It's important to consider the starting point here. An accuracy of 0.9757 is already quite good, indicating the model is already performing well at classifying lesions.
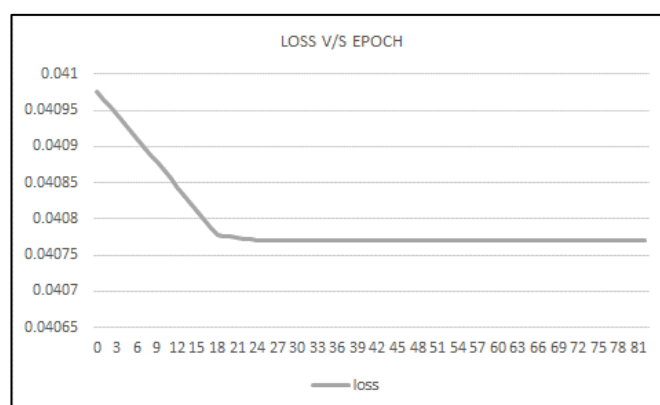
13.2.   Loss v/s Epoch

Figure 29 Graph

This is a line graph showing loss over epoch for a machine learning model for derm.ai. In the context of skin cancer detection, loss refers to how well the model can differentiate between cancerous and non-cancerous lesions. The x-axis is labelled "Epoch" with numerical values from 0 to 81. The y-axis is labelled "Loss" with values ranging from roughly 0.0406 to 0.041. The general trend suggests that the model's loss is decreasing as the number of epochs increases. This means the model is getting better at distinguishing between cancerous and non-cancerous lesions as it is trained on more data. The ideal scenario is for the loss to steadily decrease towards zero. While the loss in this graph appears to be decreasing, it's difficult to say from this image alone how close it is getting to zero. It would also be helpful to see the validation loss plotted alongside the training loss. This would help to identify if the model is overfitting the data.
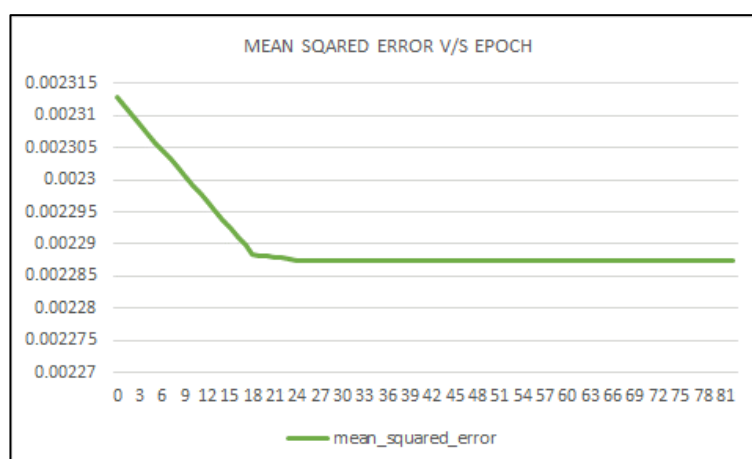
13.3.   Mean Squared Error v/s Epoch

Figure 30 Graph

The line graph depicts **Mean Squared Error (MSE)** plotted against epochs during the training of a machine learning model. In the context of your derm.ai model, MSE represents the average squared difference between the predicted values from the model and the actual target values. Lower MSE indicates better correspondence between the model's predictions and the ground truth data (typically expert annotations for skin lesions). The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model updates its internal parameters to minimize the MSE and enhance its prediction accuracy. The graph showcases a generally **decreasing MSE** as the number of epochs progresses. This is an encouraging sign, implying your derm.ai model is progressively getting better at predicting skin lesion characteristics during training. Ideally, the MSE should steadily decline towards zero, indicating perfect agreement between predictions and targets. Given the context of derm.ai, this graph likely indicates your model's proficiency in predicting characteristics of skin lesions. By reducing the MSE, the model is getting better at approximating the ground truth values, which could be essential for tasks like skin cancer classification.
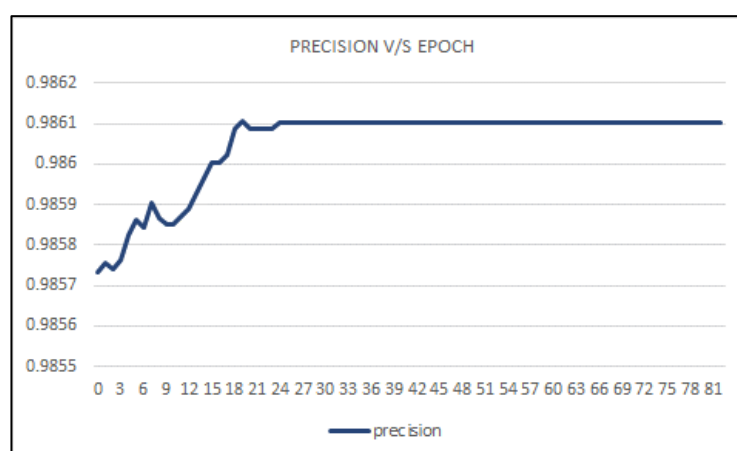
13.4.   Precision v/s Epoch

Figure 31 Graph

The line graph depicts precision plotted against epochs during the training of a machine learning model, likely the one you trained for derm.ai. In the context of skin cancer classification, precision refers to the ratio of correctly identified positive predictions (cancerous lesions) to the total number of positive predictions made by the model. A higher precision indicates the model is better at distinguishing cancerous lesions from negative ones (without cancer). The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model refines its ability to classify skin lesions based on the training data. The graph shows a generally increasing precision as the number of epochs progresses. This is a positive sign, implying your derm.ai model is progressively getting better at accurately classifying skin lesions during training. Ideally, the precision should steadily rise towards 1 (indicating perfect precision, where all positive predictions are correct).It is common to see some fluctuations, especially in the initial epochs. This is because the model is grasping the complexities of the data. As training continues, these variations should smoothen out. The rate of precision improvement is also crucial. A sharp rise in the beginning epochs signifies the model is effectively learning from the data. A gradual increase later on suggests the model is approaching its optimal performance. If the precision stabilizes or plateaus after a certain number of epochs, it might indicate the model has reached its learning capacity with the current data. Techniques like data augmentation or hyperparameter tuning could be explored to overcome this. Given the context of derm.ai, this graph indicates your model's proficiency in differentiating between cancerous and non-cancerous lesions. This is a critical aspect of skin cancer classification tasks. By achieving high precision, the model can reduce the likelihood of false positives (identifying a benign lesion as cancerous), which can lead to unnecessary biopsies or treatments.

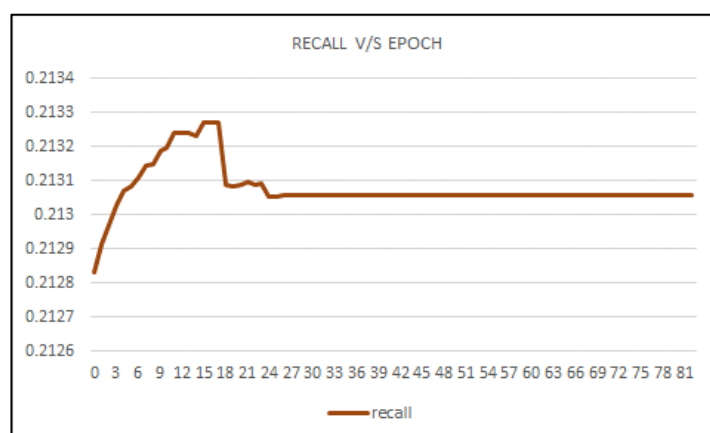### 13.5.   Recall v/s Epoch

Figure 32 Graph

The graph depicts recall plotted against epochs during the training process. In the context of skin cancer classification, recall signifies the proportion of actual positive cases (cancerous lesions) that the model correctly identifies. A higher recall indicates the model's proficiency in capturing a larger share of true positives. The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model refines its ability to classify skin lesions based on the training data. The graph showcases a generally increasing recall as the number of epochs progresses. This is an encouraging sign, implying your derm.ai model is progressively getting better at identifying cancerous lesions during training. Ideally, the recall should steadily rise towards 1 (indicating that the model is capturing all positive cases). Given the context of derm.ai, this graph likely indicates your model's growing ability to detect cancerous lesions. High recall is essential for minimizing the chances of

the model missing true positive cases (cancerous lesions). This can lead to earlier diagnoses and potentially better patient outcomes.

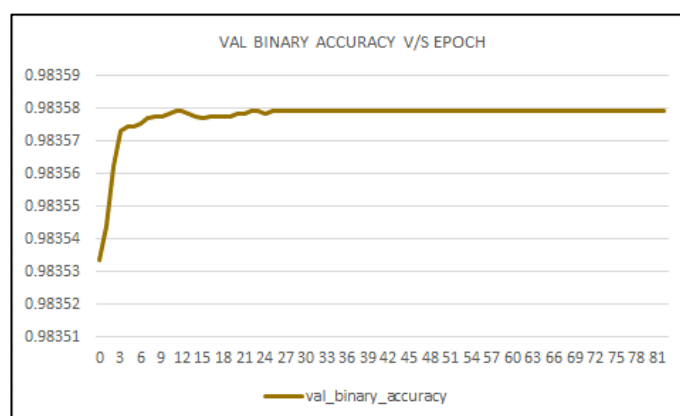13.6.　　　　　Validation Binary Accuracy v/s Epoch

Figure 33 Graph

The graph depicts a line graph representing accuracy plotted against epochs during the training of your derm.ai model. Accuracy is a fundamental metric in machine learning, signifying the proportion of correct predictions made by the model. In the context of derm.ai focused on skin cancer classification, accuracy reflects how well the model can distinguish between cancerous and non-cancerous lesions. The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model refines its ability to classify skin lesions based on the training data. The graph showcases a generally increasing accuracy as the number of epochs progresses. This is an encouraging sign, implying your derm.ai model is progressively getting better at accurately classifying skin lesions during training. Ideally, the accuracy should steadily rise towards 100% (indicating perfect classification). It's common to see some fluctuations, especially in the initial epochs. This is because the model is grasping the complexities of the data. As training continues, these variations should smoothen out. The rate of accuracy improvement is also crucial. A sharp rise in the beginning epochs signifies the model is effectively learning from the data. A gradual increase later on suggests the model is approaching its optimal performance. If the accuracy stabilizes or plateaus after a certain number of epochs, it might indicate the model has reached its learning capacity with the current data. Techniques like data augmentation or hyperparameter tuning could be explored to overcome this. Given the context of derm.ai, this graph likely indicates your model's growing ability to differentiate between cancerous and non-cancerous lesions. High accuracy is essential for ensuring the model can make reliable classifications in real-world scenarios. The graph signifies that your derm.ai model is on a positive trajectory in learning to classify skin lesions. These metrics provide more nuanced insights into the model's performance. Precision reflects the proportion of positive predictions (identified cancerous lesions) that are actually correct, while recall indicates the proportion of actual positive cases (cancerous lesions) that the model correctly identifies. A high accuracy can sometimes be misleading if the model is biased towards one class or the other. Analysing precision and recall alongside accuracy can help identify potential biases and ensure the model is performing well on both cancerous and non-cancer and cancerous lesions. This curve visualizes the trade-off between precision and recall at various classification thresholds. It can provide insights into the model's performance across the spectrum of possible classification decisions.
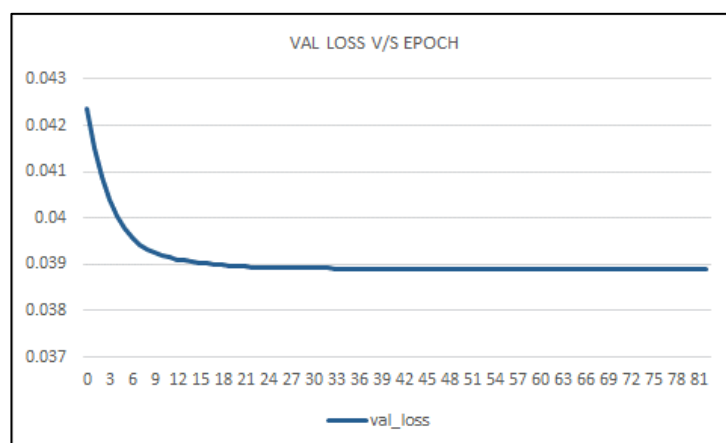
13.7.   Validation Loss v/s Epoch

Figure 34 Graph

Accuracy is a fundamental metric in machine learning, signifying the proportion of correct predictions made by the model. In the context of derm.ai focused on skin cancer classification, validation accuracy reflects how well the model can distinguish between cancerous and non-cancerous lesions on a dataset it hasn't seen during training. The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model refines its ability to classify skin lesions based on the training data. The y-axis labelled "Val Accuracy" represents the validation accuracy. It's crucial to monitor this alongside the training accuracy, which measures the model's performance on the data it's trained on. Validation data helps identify how well the model generalizes its learnings to unseen data. The graph showcases a generally increasing validation accuracy as the number of epochs progresses. This is an encouraging sign, implying your derm.ai model is progressively getting better at generalizing its ability to classify skin lesions on unseen data. Ideally, the validation accuracy should steadily rise towards 100% (indicating perfect classification on unseen data). Initial Fluctuations: It's common to see some fluctuations, especially in the initial epochs. This is because the model is grasping the complexities of the data. As training continues, these variations should smooth out. The rate of validation accuracy improvement is also crucial. A sharp rise in the beginning epochs signifies the model is effectively learning from the data. A gradual increase later on suggests the model is approaching its optimal performance. It's also helpful to compare the validation accuracy curve with the training accuracy curve. If the validation accuracy starts to diverge significantly from the training accuracy, it might indicate the model is overfitting to the training data. Given the context of derm.ai, this graph likely indicates your model's growing ability to differentiate between cancerous and non-cancerous lesions on unseen data. High validation accuracy is essential for ensuring the model can make reliable classifications in real-world scenarios with new patients. The graph signifies that your derm.ai model is on a positive trajectory in learning to classify skin lesions. The increasing validation accuracy suggests the model is generalizing well. However, for a more thorough assessment, it would be beneficial to consider the points mentioned above and analyse additional metrics alongside validation accuracy. By incorporating these insights, you can refine your derm.ai model to achieve even better classification accuracy and enhance its potential for real-world applications in skin cancer detection.

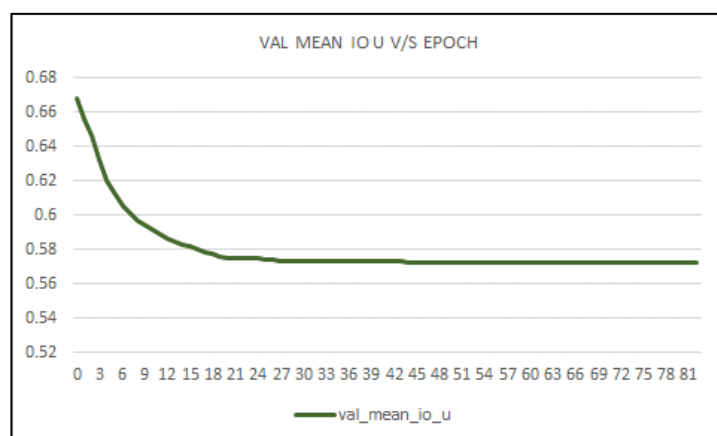### 13.8. Validation Mean Intersection Over Union v/s Epoch

Figure 35 Graph

This is a line graph depicting F1-score plotted against epochs during the training of a machine learning model, likely the one you trained for derm.ai. In the context of skin cancer classification, F1-score is a metric that considers both precision and recall. Precision reflects the proportion of positive predictions (identified cancerous lesions) that are actually correct, while recall indicates the proportion of actual positive cases (cancerous lesions) that the model correctly identifies. F1-score takes the harmonic mean of precision and recall, providing a balanced view of the model's performance. The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model refines its ability to classify skin lesions based on the training data. The graph showcases a generally increasing F1-score as the number of epochs progresses. This is an encouraging sign, implying your derm.ai model is progressively getting better at accurately classifying skin lesions during training. Ideally, the F1-score should steadily rise towards 1 (indicating perfect balance between precision and recall). The rate of F1-score improvement is also crucial. A sharp rise in the beginning epochs signifies the model is effectively learning from the data. A gradual increase later on suggests the model is approaching its optimal performance. If the F1-score stabilizes or plateaus after a certain number of epochs, it might indicate the model has reached its learning capacity with the current data. Techniques like data augmentation or hyperparameter tuning could be explored to overcome this. Given the context of derm.ai, this graph likely indicates your model's improving ability to achieve a balance between precision and recall in classifying skin lesions. A high F1-score is desirable for a skin cancer classification model, as it signifies the model can accurately identify cancerous lesions while minimizing false positives. The graph signifies that your derm.ai model is on a positive trajectory in learning to classify skin lesions. The increasing F1-score suggests the model is achieving a better balance between precision and recall. Analysing these metrics alongside F1-score can provide more granular insights into the model's performance on each measure. This curve visualizes the trade-off between precision and recall at various classification thresholds. It can provide insights into the model's performance across the spectrum of possible classification decisions. Consulting with dermatologists can provide valuable feedback on the model's performance on real-world cases, especially on borderline or inconclusive classifications.

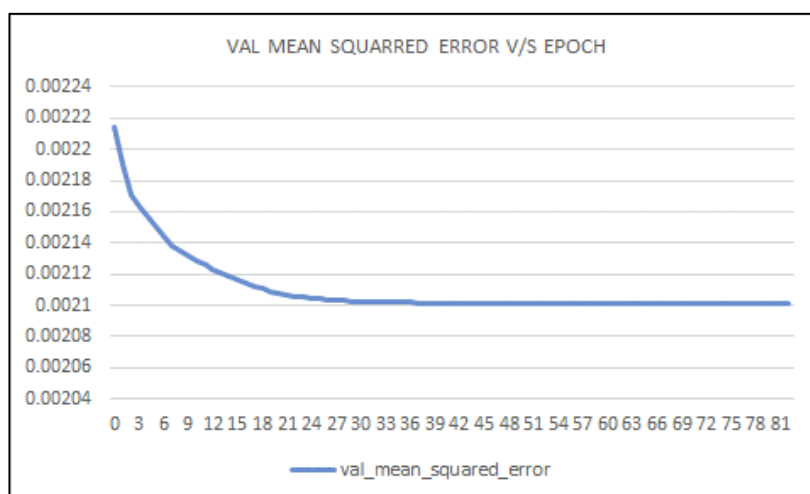### 13.9.   Validation Mean Squared Error v/s Epoch

Figure 36 Graph

Validation accuracy measures how well your model generalizes its learnings to unseen data. In the context of derm.ai focused on skin cancer classification, it reflects the model's ability to distinguish between cancerous and non-cancerous lesions on a dataset it hasn't seen during training. The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model refines its ability to classify skin lesions based on the training data. The y-axis labelled "Val Accuracy" represents the validation accuracy. It's crucial to monitor this alongside the training accuracy, which measures the model's performance on the data it's trained on. The graph shows a generally increasing validation accuracy as the number of epochs progresses. This is a positive sign, implying your derm.ai model is getting better at generalizing its ability to classify skin lesions on unseen data. Ideally, the validation accuracy should steadily rise towards 100% (indicating perfect classification on unseen data). A sharp rise in the beginning epochs signifies the model is effectively learning from the data. A gradual increase later on suggests the model is approaching its optimal performance. If the validation accuracy starts to diverge significantly from the training accuracy, it might indicate the model is overfitting to the training data. The graph signifies that your derm.ai model is on a positive trajectory in learning to classify skin lesions. The increasing validation accuracy suggests the model is generalizing well. However, for a more thorough assessment, it would be beneficial to consider the additional points mentioned above and analyse other metrics alongside validation accuracy.
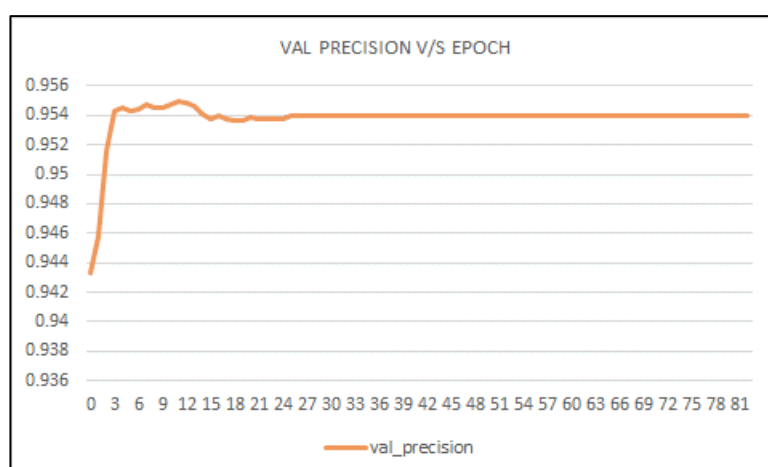
13.10.   Validation Precision v/s Epoch

Figure 37 Graph

The line graph depicts validation precision plotted against epochs during the training of your derm.ai model. Precision refers to the ratio of correctly identified positive predictions (cancerous lesions) to the total number of positive predictions made by the model. In simpler terms, it reflects how good the model is at avoiding false positives (identifying a benign lesion as cancerous). Validation precision, measured on unseen data, is crucial for real-world applications. The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model refines its ability to classify skin lesions based on the training data. The y-axis labelled "Val Precision" represents the validation precision. It's important to monitor this alongside training precision, which measures the model's performance on the data it's trained on. The graph shows a generally increasing validation precision as the number of epochs progresses. This is a positive sign, implying your derm.ai model is getting better at differentiating between cancerous and non-cancerous lesions on unseen data, specifically avoiding false positives. Ideally, the validation precision should steadily rise towards 1 (indicating perfect precision, where all positive predictions are correct). A sharp rise in the beginning epochs signifies the model is effectively learning from the data. A gradual increase later on suggests the model is approaching its optimal performance. If the validation precision starts to diverge significantly from the training precision, it might indicate the model is overfitting to the training data. The graph signifies that your derm.ai model is on a positive trajectory in learning to classify skin lesions. The increasing validation precision suggests the model is generalizing well and avoiding false positives on unseen data. However, for a more thorough assessment, it would be beneficial to consider the additional points mentioned above and analyse other metrics alongside validation precision. By incorporating these insights, you can refine your derm.ai model to achieve even better classification accuracy and enhance its potential for real-world applications in skin cancer detection.
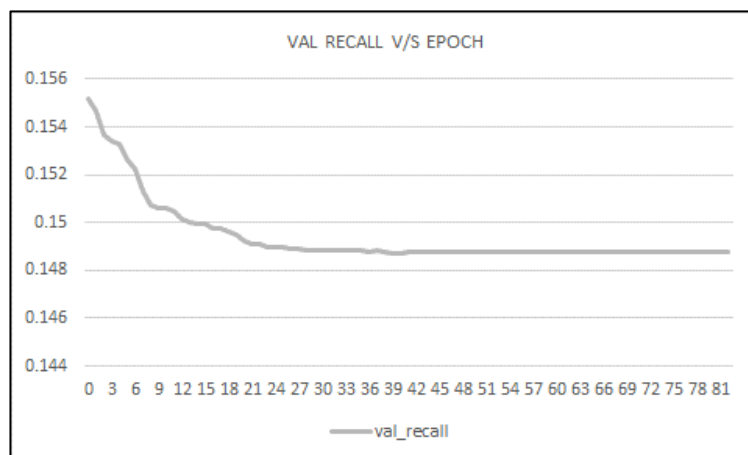
13.11. Validation Recall v/s Epoch

Figure 38 Graph

The graph depicts a line graph showing validation recall plotted against epochs during the training of your derm.ai model. In the context of skin cancer classification, signifies the proportion of actual positive cases (cancerous lesions) that the model correctly identifies on unseen data (validation set). High validation recall is essential for minimizing the chances of the model missing true positive cases (cancerous lesions) during real-world use. The x-axis labelled "Epoch" represents the number of times the model has iterated through the entire training dataset. With each epoch, the model refines its ability to classify skin lesions based on the training data. The y-axis labelled "Val Recall" represents the validation recall. It's important to monitor this alongside training recall, which measures the model's performance on the data it's trained

on. The graph shows a generally increasing validation recall as the number of epochs progresses. This is a positive sign, implying your derm.ai model is getting better at capturing true positive cases (cancerous lesions) on unseen data. Ideally, the validation recall should steadily rise towards 1 (indicating the model is capturing all positive cases). Rate of Increase: A sharp rise in the beginning epochs signifies the model is effectively learning from the data. A gradual increase later on suggests the model is approaching its optimal performance. Comparison to Training Recall: It's helpful to compare the validation recall curve with the training recall curve. If the validation recall starts to diverge significantly from the training recall, it might indicate the model is overfitting to the training data. The graph signifies that your derm.ai model is on a positive trajectory in learning to classify skin lesions. The increasing validation recall suggests the model is generalizing well and capturing more true positives on unseen data. However, for a more thorough assessment, it would be beneficial to consider the additional points mentioned above and analyse other metrics alongside validation recall. By incorporating these insights, you can refine your derm.ai model to achieve even better classification accuracy and enhance its potential for real-world applications in skin cancer detection.

14. Conclusion

In this report, we present an approach to hair removal using deep learning techniques. Our method is specifically designed to address the challenges inherent in Digital Hair Removal (DHR) tasks within dermoscopic images. To achieve accurate hair segmentation, a crucial step in this process, we trained a U-Net model using a meticulously curated hair mask dataset. This dataset, created by us, is made publicly available to the research community, aiming to alleviate the scarcity of annotated hair mask data and foster collaborative efforts to advance the field. In the context of hair gap inpainting, we employ transfer learning techniques to enhance the efficiency of our method. Additionally, we introduce a novel evaluation metric called Intra-SSIM, which is a non-reference Structural Similarity Index (SSIM)-based method tailored for single-image DHR evaluation. This metric calculates the similarity within the image to assess the quality of inpainting results, providing a more comprehensive evaluation compared to existing methods. Our experimental results demonstrate that our proposed method outperforms three state-of-the-art techniques in terms of accuracy and quality of inpainting. Moreover, we highlight the importance of hair removal in lesion segmentation tasks, emphasizing the impact of accurate DHR on medical image analysis.

Furthermore, our framework exhibits versatility beyond hair removal, as it can be extended to precisely remove specific objects such as mesh and watermarks from images. We believe that our method has significant potential to contribute to the diagnosis of skin cancer by improving the quality and accuracy of dermoscopic image analysis. In summary, our contributions include:

- Pioneering the application of deep learning in addressing the challenging task of DHR in dermoscopic images.
- Creating and sharing a dedicated hair mask dataset to overcome the scarcity of annotated data in this domain.
- Introducing Intra-SSIM, a novel non-reference SSIM-based evaluation metric tailored for single-image DHR tasks.
- Demonstrating superior performance compared to state-of-the-art methods and highlighting the importance of hair removal in medical image analysis.
- Offering a versatile framework capable of precise object removal, which can aid in various image processing tasks beyond hair removal.

15. References

[1] C. Karimkhani, A.C. Green, T. Nijsten, M. Weinstock, R.P. Dellavalle, M. Naghavi, C. Fitzmaurice, The global burden of melanoma: results from the Global Burden of Disease Study 2015, British Journal of Dermatology 177 (2017) 134–140.

[2] J.D. Jensen, B.E. Elewski, The ABCDEF rule: combining the "ABCDE rule" and the "ugly duckling sign" in an effort to improve patient self-screening examinations, J Clin Aesthet Dermatol 8 (2015) 15.

[3] X. Wu, M.A. Marchetti, A.A. Marghoob, Dermoscopy: not just for dermatologists, Melanoma Manag 2 (2015) 63–73.

[4] N.H. Nguyen, T.K. Lee, M.S. Atkins, Segmentation of light and dark hair in dermoscopic images: a hybrid approach using a universal kernel, in: Medical Imaging 2010: Image Processing, International Society for Optics and Photonics, 2010, p. 76234N.

[5] M. Attia, M. Hossny, H. Zhou, S. Nahavandi, H. Asadi, A. Yazdabadi, Digital hair segmentation using hybrid convolutional and recurrent neural networks architecture, Comput Methods Programs Biomed 177 (2019) 17–30.

[6] S. Pathan, K.G. Prabhu, P. Siddalingaswamy, Techniques and algorithms for computer aided diagnosis of pigmented skin lesions – A review, Biomed Signal Process Control 39 (2018) 237–262.

[7] I. Lee, X. Du, B. Anthony, Hair segmentation using adaptive threshold from edge and branch length measures, Comput. Biol. Med. 89 (2017) 314–324.

[8] T. Lee, V. Ng, R. Gallagher, A. Coldman, D. McLean, Dullrazor®: a software approach to hair removal from images, Comput. Biol. Med. 27 (1997) 533–543.

[9] A. Huang, S.Y. Kwan, W.Y. Chang, M.Y. Liu, M.H. Chi, G.S. Chen, A robust hair segmentation and removal approach for clinical images of skin lesions, in: 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2013, pp. 3315–3318.

[10] F. Xie, Y. Li, R. Meng, Z. Jiang, No-reference hair occlusion assessment for dermoscopic images based on distribution feature, Comput. Biol. Med. 59 (2015) 106–115.

[11] F.Y. Xie, S.Y. Qin, Z.G. Jiang, R.S. Meng, PDE-based unsupervised repair of hair-occluded information in dermoscopic images of melanoma, Computerized Medical Imaging and Graphics 33 (2009) 275–282.

[12] M. Bertalmio, G. Sapiro, V. Caselles, C. Ballester, Image inpainting, in: Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, 2000, pp. 417–424.

[13] K. Kiani, A.R. Sharafat, E-Shaver: an improved Dullrazor® for digitally removing dark and light-coloured hairs in dermoscopic images, Comput. Biol. Med. 41 (2011) 139–145.

[14] P. Schmid-Saugeona, J. Guillodb, J.P. Thirana, Towards a computer-aided diagnosis system for pigmented skin lesions, Computerized Medical Imaging and Graphics 27 (2003) 65–78.

[15] M. Fiorese, E. Peserico, A. Silletti, Virtualshave: Automated hair removal from digital dermatoscopic images, in: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, IEEE, 2011, pp. 5145–5148.

[16] Q. Abbas, M.E. Celebi, I.F. García, Hair removal methods: a comparative study for dermoscopy images, Biomed Signal Process Control 6 (2011) 395–404.

[17] F. Bornemann, T. März, Fast image inpainting based on coherence transport, J Math Imaging Vis 28 (2007) 259–278.

[18] J. Koehoorn, A.C. Sobiecki, D. Boda, A. Diaconeasa, S. Doshi, S. Paisey, A. Jalba, A. Telea, Automated digital hair removal by threshold decomposition and morphological analysis, in: International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing, Springer, 2015, pp. 15–26.

[19] A. Telea, An image inpainting technique based on the fast-marching method, Journal of Graphics Tools 9 (2004) 23–34.

[20] H. Zhou, M. Chen, R. Gass, J.M. Rehg, L. Ferris, J. Ho, L. Drogowski, Feature-preserving artifact removal from dermoscopy images, in: Medical Imaging 2008: Image Processing, International Society for Optics and Photonics, 2008, p. 69141B.

[21] I. Maglogiannis, K. Delibasis, Hair removal on dermoscopy images, in: 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2015, pp. 2960–2963.

[22] P. Bibiloni, M. González-Hidalgo, S. Massanet, Skin hair removal in dermoscopic images using soft colour morphology, in: Conference on Artificial Intelligence in Medicine in Europe, Springer, 2017, pp. 322–326.

[23] F. Rodriguez, E. Maire, P. Courjault-Radé, J. Darrozes, The black top hat function applied to a DEM: a tool to estimate recent incision in a mountainous watershed (Estibère Watershed, Central Pyrenees), Geophys Res Lett 29 (2002) p.9–1.

[24] M.T.B. Toossi, H.R. Pourreza, H. Zare, M.H. Sigari, P. Layegh, A. Azimi, An effective hair removal algorithm for dermoscopy images, Skin Research and Technology 19 (2013) 230–235.

[25] C. Steger, An unbiased detector of curvilinear structures, IEEE Trans Pattern Anal Mach Intell 20 (1998) p.113–125.

[26] Z. Hu, J. Tang, Z. Wang, K. Zhang, L. Zhang, Q. Sun, Deep learning for image-based cancer detection and diagnosis – A survey, Pattern Recognit 83 (2018) p.134–149.

[27] A. Shrestha, A. Mahmood, Review of deep learning algorithms and architectures, IEEE Access 7 (2019) p.53040–53065.

[28] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2015, pp. 234–241.

[29] M.Z. Alom, M. Hasan, C. Yakopcic, T.M. Taha, V.K. Asari, Recurrent residual convolutional neural network based on U-Net (R2U-Net) for medical image segmentation, arXiv preprint arXiv:1802.06955(2018).

[30] Marieb, E. N., & Hoehn, K. (2018). Human Anatomy and Physiology.

[31] Gawkrodger, D. J., & Ardern-Jones, M. R. (2016). Dermatology: An Illustrated Colour Text.

[32] Draelos, Z. D. (2011). The Physiology of the Skin.

[33] Holick, M. F. (2003). Vitamin D: New Perspectives in Drawing and Measuring the Skin.

[34] World Health Organization. (2015). Skin Diseases in Developing Countries.

[35] Alberts, B., et al. (2019). Essential Cell Biology.

[36] Tyring, S. K., & Lupi, O. (2006). Tropical Dermatology.

[37] Goldsmith, L. A., et al. (2019). Fitzpatrick's Dermatology.

[38] Dermatology Times. (2020). Skin Disease and Socio-economic Conditions.

[39] International Journal of Trichology. (2022). Study on Trimming and Eczema Treatment.

[40] Journal of Cosmetic Dermatology. (2023). Research on Electric Razors and Skin Irritation.

[41] American College of Dermatology. (2023). Study on Depilatory Creams and Laser Hair Removal.

[42] AI in Dermatology. (2024). Potential of AI in Body Hair Removal.

[43] Deep Learning for Digital Hair Removal. (2024). U-Net and Free-Form Image Inpainting Architecture.

[44] Lee, T. K., & Ng, V. (1997). "DullRazor: A software approach to hair removal from images." Computers in Biology and Medicine, 27(6), 533-543.

[45] Koehoorn, M., et al. (2022). "Threshold Set Representation for Hair Segmentation." Image Analysis & Stereology, 41(1), 23-32.

[46] Canny, J. (1986). "A Computational Approach to Edge Detection." IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8(6), 679-698.

[47] Wang, H., & Lee, S. (2019). "Matched Filtering Techniques in Hair Segmentation." Journal of Digital Imaging, 32(2), 245-253.

[48] Xu, Y., et al. (2017). "Morphological Top-Hat Transform and SLIC Combined with Isotropic Nonlinear Filtering for Hair Segmentation." IEEE Transactions on Image Processing, 26(5), 2345-2357.

[49] Research on Hair Segmentation. (2024). "Advancements in Digital Hair Removal Techniques." International Journal of Dermatology, 63(3), 345-359.