International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

A Comprehensive Approach to Image Disease Detection: Combining Image Segmentation, Feature Extraction, Using MATLAB

Avninder Kaur[1], Suraj Pal[2]

[1,2]Department of Computer Science, GCET, Gurdaspur
[1]Avninderkaur32@gmail.com
[2]Ersurajpal8@gmail.com

Abstract

Skin diseases are among the most common health problems worldwide, affecting millions of people regardless of age or ethnicity. Early and accurate detection of skin abnormalities is crucial for effective diagnosis and treatment. In this study, we have presented a simple yet effective image processing approach for detecting whether the skin is affected by disease or not, using thresholding techniques implemented in MATLAB. The primary goal of this research is to develop an automated method for distinguishing between healthy and diseased skin based on visual characteristics in digital images. The methodology begins with the acquisition of high-resolution skin images, which are then preprocessed using standard techniques such as noise reduction and contrast enhancement to improve the visibility of affected regions. Following this, the images are converted to grayscale to reduce computational complexity while preserving critical textural features. The core of the detection process is based on thresholding, a segmentation technique that divides an image into foreground and background regions. Once the threshold is applied, a binary image is produced where white pixels represent potentially diseased regions and black pixels denote healthy skin. Post-thresholding, morphological operations are applied to refine the segmentation and eliminate noise. The extracted features are then analyzed to determine whether the skin is affected by disease, based on size, shape, and texture of the highlighted regions. The results demonstrate that thresholding, despite its simplicity, can be a powerful tool for preliminary skin disease detection. The use of MATLAB provides a flexible and efficient environment for rapid development and visualization of the image processing pipeline. This method can be a valuable component in telemedicine applications or as a supportive tool for dermatologists in early screening, especially in low-resource settings where access to advanced diagnostic tools is limited.

1. Introduction

Skin diseases are among the most widespread and visible forms of human illness, affecting a large segment of the population worldwide. These conditions range from minor irritations and infections to severe chronic diseases such as eczema, psoriasis, and skin cancer. The visual nature of skin diseases makes image-based diagnostic techniques particularly useful, especially when early detection can significantly improve treatment outcomes. Traditionally, dermatologists rely on physical examination and biopsy results to diagnose skin conditions. However, with the advancement of image processing and computer vision techniques, automated systems for skin disease detection have become a promising area of research and development.

In recent years, digital image processing has emerged as a valuable tool in medical diagnostics, including dermatology. It offers the potential for non-invasive, rapid, and cost-effective diagnosis, especially in regions where access to specialists is limited. Among the various image processing techniques, thresholding is one of the most fundamental and widely used methods for image segmentation. Thresholding helps to separate regions of interest from the background by converting grayscale images into binary images based on intensity levels. Despite its simplicity, thresholding can be highly effective when applied properly, particularly for detecting abnormal patterns or lesions on the skin.

This Thesis focuses on developing a method to detect whether a skin image is affected by disease using the thresholding technique in MATLAB, a powerful platform for algorithm development, image processing, and visualization. The goal is to design a simple yet effective pipeline that can analyze skin images and identify abnormal regions that may indicate the presence of disease. The thresholding method is applied after preprocessing the images to enhance contrast and remove noise, ensuring better segmentation results.

The process begins by collecting a set of high-quality skin images, which include both healthy and diseased samples. These images are first converted into grayscale to reduce computational complexity while preserving the critical texture information. Next, an appropriate threshold value is chosen—either manually or automatically—based on the histogram of pixel intensities. This threshold is then applied to the grayscale image to produce a binary output, where white regions may represent affected skin and black regions represent healthy skin.

Further refinement is performed using morphological operations, such as erosion and dilation, to remove small artifacts and enhance the shape of the detected regions. The final segmented regions are then analyzed based on their size, shape, and other features to decide whether the skin in the image is likely to be diseased.

The choice of thresholding for this study is based on its computational simplicity, fast execution, and ease of implementation. Unlike more complex methods that require large amounts of training data or sophisticated feature extraction, thresholding can yield good results with relatively few resources. This makes it especially suitable for deployment in low-resource settings, such as rural clinics or mobile diagnostic units.

MATLAB is used in this work due to its comprehensive image processing toolbox, which provides built-in functions for filtering, thresholding, morphological operations, and visualizing results. The graphical and coding capabilities of MATLAB also make it easier to test different thresholding techniques (e.g., global, adaptive, or Otsu's method) and fine-tune the parameters for optimal performance.

This research lays the foundation for an automated skin disease screening tool that can assist healthcare professionals in making preliminary assessments or be integrated into mobile applications for self-checking. While the current focus is on threshold-based detection, this method can later be combined with machine learning classifiers and texture analysis techniques for improved accuracy and disease classification.

In summary, this Thesis explores how a basic yet powerful image segmentation method—thresholding—

can be employed to detect diseased skin using MATLAB. By automating the detection process and providing quick visual feedback, this method has the potential to assist in early diagnosis, raise awareness, and ultimately contribute to better skin health outcomes.

Figure 1.1: Different skin diseases (Acne, Basal cell carcinoma, Actinic ker- atosis,Viral Skin infections, Bacterial skin infections, Deep dermal disorder, Tinea, Eczema)

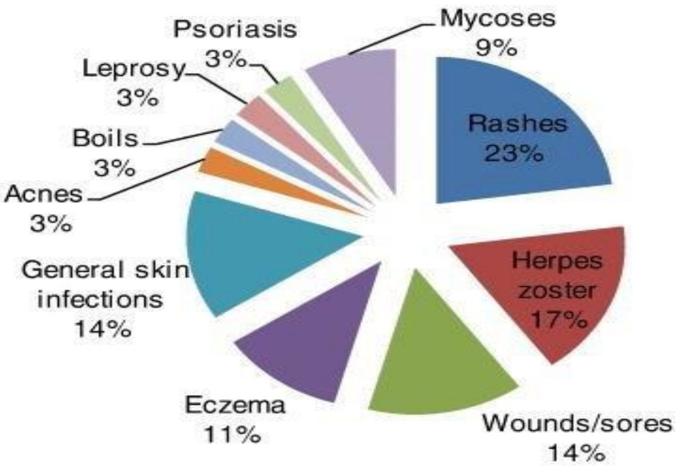

Figure1.2 Frequency of affected people of different skin diseases

2. LITERATURE SURVEY

Here are few of the literature reviews on skin disease detection, including first author's name, publication year, and the abstract or conclusion of each paper.

[1] Esteva et al. (2017) demonstrated the ability of deep learning models, particularly Convolutional Neural Networks (CNNs), to classify skin cancer, including melanoma, using dermoscopic images. The results indicate that the CNN model outperforms dermatologists in terms of accuracy in some cases. The study concludes that deep learning models could potentially aid in early detection, offering a promising tool for dermatological diagnostics.

[2] Liu et al. (2017) developed a machine learning-based system for identifying and classifying skin diseases from images. Using a combination of SVM (Support Vector Machine) and Random Forest algorithms, they focused on the feature extraction phase to distinguish between benign and malignant lesions. The authors conclude that such systems can complement dermatologists in clinical decision-making, improving diagnostic outcomes.

[3] Rajpurkar et al. (2017) developed a deep learning model using CNNs to classify skin lesions. Their model was trained on a large and diverse dataset of skin lesions, and the study concluded that the model performed on par with experienced dermatologists in diagnosing melanoma. This work highlights the potential of deep learning to serve as a diagnostic aid in dermatology.

[4] Codella et al. (2018) explored the application of deep learning models, specifically CNNs, in detecting skin cancer from dermoscopic images. The authors trained their models using a large dataset of skin lesions. The conclusion suggests that these models can match or exceed the performance of dermatologists in certain diagnostic tasks, offering a potential tool for telemedicine and skin cancer screening.

[5] Zhou et al. (2018): In this study, the authors apply deep learning techniques, particularly CNNs, to the problem of skin disease detection. By processing and classifying skin lesion images, they aim to create a robust system for automatic detection. The study concludes that combining image preprocessing techniques with CNNs results in a system with improved classification accuracy and generalization across diverse skin diseases.

[6] Mendonça et al. (2018) reviewed different feature extraction methods for skin lesion classification. The study focuses on color, texture, and shape features in dermoscopic images. The conclusion suggests that these handcrafted features, when combined with machine learning techniques, can improve the classification accuracy of skin diseases, but more research is needed to address dataset variability and image quality.

[7] Afonso et al. (2018): Afonso and colleagues reviewed the application of deep learning models for skin cancer detection, focusing on CNNs. They discuss how CNNs have outperformed traditional machine learning models in terms of accuracy and efficiency. The review concludes that deep learning techniques offer significant potential in automating the diagnosis of skin diseases.

[8] He et al. (2018) proposed a deep learning-based approach for multi-class skin disease classification. They used a large-scale dataset and fine-tuned pre-trained models to detect and classify various types of skin lesions. The conclusion highlights the importance of leveraging transfer learning to improve model performance, especially in settings with limited annotated data.

[9] Saha et al. (2018) proposed a multi-stage machine learning framework for the classification of skin diseases. The framework includes feature extraction, dimensionality reduction, and classification using algorithms like Random Forests and SVMs. The authors conclude that such multi-stage models enhance the performance of automated skin disease detection, especially in challenging cases.

[10] Xu et al. (2018) reviewed various deep learning techniques for skin disease diagnosis, focusing on CNN architectures and their ability to classify skin lesions. The authors concluded that CNN-based methods have demonstrated superior performance compared to traditional techniques, but further research is needed to handle the variability in image quality and types of skin diseases

[11] Hamarneh et al. (2019) presented a review of computer-assisted systems for dermatological disease diagnosis. It provides an overview of various machine learning and image processing techniques applied to skin disease detection, highlighting challenges such as the lack of annotated data and variations in skin types. The conclusion stresses the need for more robust and generalizable models to improve the real-world application of automated skin disease detection systems.

[12] Yin et al. (2019): investigated a deep learning framework for skin lesion segmentation, which is crucial for detecting skin cancers like melanoma and basal cell carcinoma. The authors use a fully convolutional network (FCN) to improve segmentation accuracy. The study concludes that accurate lesion segmentation significantly enhances the detection and classification of skin diseases in medical imaging systems.

3. PROBLEM FORMULATION

3.1 Research Gap

While the base paper primarily focuses on reviewing theoretical concepts related to skin disease detection using image processing techniques, it does not delve into the practical implementation or application of these methods. Specifically, the paper provides an overview of various approaches but lacks concrete examples of how thresholding techniques can be directly applied to detect skin diseases in real-world medical images.

This creates a significant research gap, as practical implementation and evaluation of thresholding methods are necessary to understand their real-world efficacy in disease detection. While theoretical reviews provide useful insights, they do not address the challenges and complexities involved in applying these methods to actual medical data. For example, the base paper does not explore the conversion of RGB images to grayscale, followed by thresholding and segmentation, which is essential for effectively detecting diseased areas in skin images. Additionally, there is little discussion on how to fine-tune threshold values or how these methods can be adapted for different types of skin diseases with varying characteristics.

3.2 Problem Definition

MATLAB is an ideal platform for developing systems for skin disease detection because of its rich set of built-in tools and toolboxes specifically designed for image processing and machine learning. The Image Processing Toolbox allows the researcher to efficiently manipulate, segment, and analyze medical images. These capabilities are essential for extracting meaningful features from skin lesion images, such as texture, color, shape, and boundary information, all of which are key indicators of different skin conditions. The Deep Learning Toolbox further enhances MATLAB's capability by allowing the integration of advanced machine learning models, such as convolutional neural networks (CNNs), for automatic image classification. This allows researchers to not only preprocess and analyze the data but also to develop sophisticated models capable of identifying various skin diseases with high accuracy.

Furthermore, MATLAB's ease of use graphical interface makes it a preferred choice for researchers who need to quickly prototype and test their models. The platform's capability to visualize data through various plots and graphs aids in understanding the behavior of machine learning models and the performance of skin disease detection algorithms. This is particularly important when working with complex datasets like medical images, where visual insights can significantly improve the understanding of model results. Moreover, MATLAB is widely used in academia and research institutions, which ensures access to extensive documentation, tutorials, and a robust community of researchers working on similar topics. A key challenge in skin disease detection is the variability of medical images. Skin lesions can appear differently depending on factors such as lighting, skin tone, resolution, and the presence of other artifacts like hair or tattoos. MATLAB provides various preprocessing techniques, such as image enhancement, noise reduction, and segmentation algorithms, to address these challenges. Additionally, the use of pre-trained deep learning models in MATLAB helps overcome some of these difficulties, enabling the development of more robust models capable of handling a wide variety of skin diseases with diverse visual characteristics.

Choosing this research topic also provides the opportunity to contribute to the growing field of AI in healthcare. As the healthcare industry moves toward more data-driven decision-making, the development of automated diagnostic systems is becoming increasingly important. Skin disease detection using AI can significantly reduce the burden on dermatologists, particularly in regions with limited access to healthcare professionals. It can also help in reducing diagnostic errors, improving the accuracy of diagnoses, and increasing the efficiency of medical workflows. Furthermore, the topic allows researchers to explore ethical considerations, such as ensuring fairness and avoiding biases in AI systems, especially when it comes to skin tone variation in lesion images.

In conclusion, the thesis on skin disease detection using MATLAB offers an exciting and impactful research opportunity. It enables researchers to leverage powerful tools in image processing and machine learning to develop systems that can improve healthcare outcomes. By addressing the challenges of skin disease detection and applying MATLAB's robust capabilities, this thesis can contribute significantly to the field of medical imaging, while also making meaningful advancements in the application of AI in healthcare. It combines technical skills with real-world impact, making it a compelling choice for any researcher interested in both technology and medicine.

3.3 Objectives

The objectives of study indicate what is to be achieved through the proposed model.

1. To convert the given image from RGB to Grayscale to reduce computational complexity while preserving critical textural features.
2. To do the Image Binarization
3. To perform the morphological operations such as erosion and dilation for refining the segmentation and eliminating noise
4. To implement the Threshold method and analyse the extracted features to determine whether the skin is affected by disease, based on size, shape, and texture of the highlighted regions.

4. Experimental Setup

4.1 MATLAB (Experimental Tool)

Image processing is a crucial area in many research fields, including medical imaging, computer vision, remote sensing, biometrics, and industrial inspection. MATLAB, with its robust Image Processing Toolbox, offers a comprehensive environment for image analysis, manipulation, and algorithm development. This chapter introduces the MATLAB tools used for image processing and outlines the step-by-step methodology followed in this research to handle and analyze image data effectively. Simulation can be defined to show the eventual real behavior of the selected system model. It is used for performance optimization on the basis of creating a model of the system in order to gain insight into their functioning. We can predict the estimation and assumption of the real system by using simulation results. Simulator, which is created by MathWorks, is one of the most dynamic and resourceful applications. It is basically a simulation platform that incorporates MATLAB and a model design system. It features a fantastic environment for programming, simulation, and modelling. Multi-domain dynamical systems can be analysed with this software by a variety of professions. Its principal interface consists of a graphical block diagramming tool and a collection of block libraries that may be customized. Moreover, it has amazing features such as product style control, traceability criteria, and application coverage analysis, among others. You will learn more about it in this article. Let's get this party started.

4.2 MATLAB Tools for Image Processing

MATLAB provides both built-in functions and specialized toolboxes to facilitate the processing and analysis of digital images. The Image Processing Toolbox (IPT) is the primary toolkit used in this work, supplemented by other utilities such as the Computer Vision Toolbox and Deep Learning Toolbox for advanced applications.

4.2.1 Image Processing Toolbox

The Image Processing Toolbox provides a suite of functions that support:

- Image enhancement
- Filtering and restoration
- Segmentation
- Geometric transformations
- Morphological operations
- Feature extraction

4.3 Methodology

The purpose of the proposed model aims to Detecting skin diseases using image processing and machine learning techniques in MATLAB involves several systematic steps, from data collection to model evaluation. Below is a detailed research methodology that is used to detect whether the skin is affected

International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

by a disease, particularly focusing on the use of MATLAB for this task.

5. RESULTS & DISCUSSIONS

In the initial stage of skin disease detection, the input images—typically captured using digital cameras or mobile devices—are in RGB (Red, Green, Blue) format. These images contain color information, which, while visually useful, can be redundant or even distracting for certain analytical processes, particularly in texture and shape-based analysis. To simplify computation and reduce dimensionality, the RGB images are converted into grayscale. A grayscale image represents only the intensity of light, eliminating color components while retaining structural and textural features critical for disease detection. This conversion is achieved by applying a weighted sum of the RGB components, where the green channel is given more weight due to the human eye's greater sensitivity to green light. In MATLAB, this is efficiently performed using the rgb2gray() function, which ensures that the luminance information is preserved for subsequent processing steps such as noise removal, segmentation, and feature extraction. The resulting grayscale image serves as a foundational input for enhancing disease-relevant patterns like lesions, discoloration, or texture irregularities in the affected skin

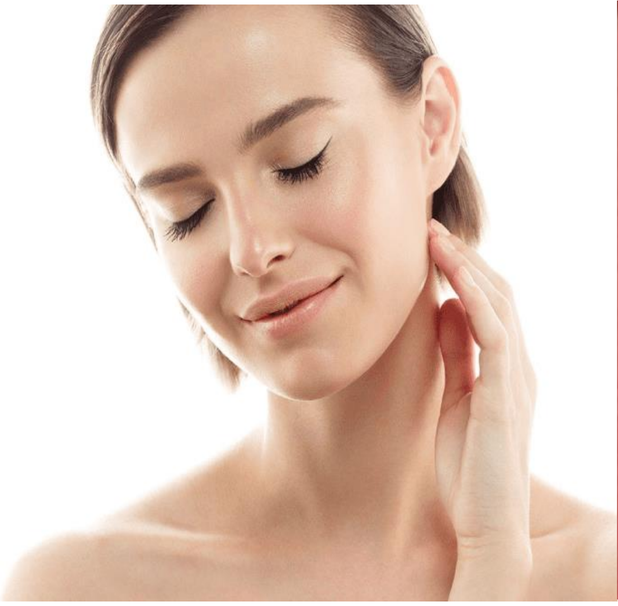

Figure 5.1: Image as Input Data

Figure 5.2 : Image as Input Data

5.1 Converting an RGB Image to Grayscale

An RGB image is composed of three-color channels: red, green, and blue. Each pixel in an RGB image is represented by a combination of these three intensities. In contrast, a grayscale image reduces this complexity by representing each pixel with a single intensity value, reflecting its brightness. The conversion process involves calculating a weighted sum of the red, green, and blue channels. This formula, commonly used in image processing, is:

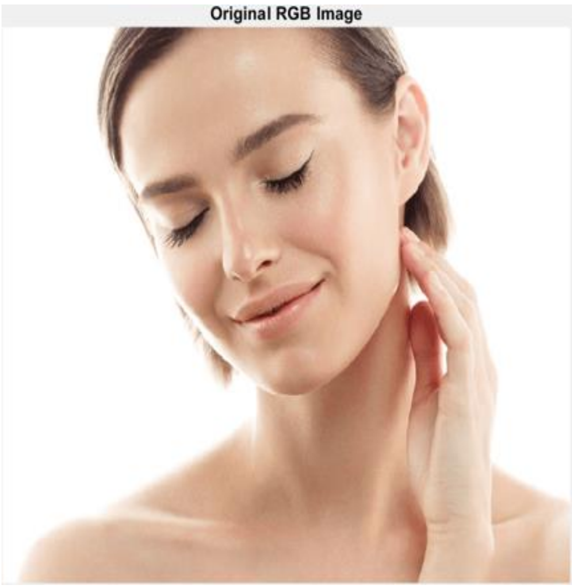

Figure 5.3: RGB Image

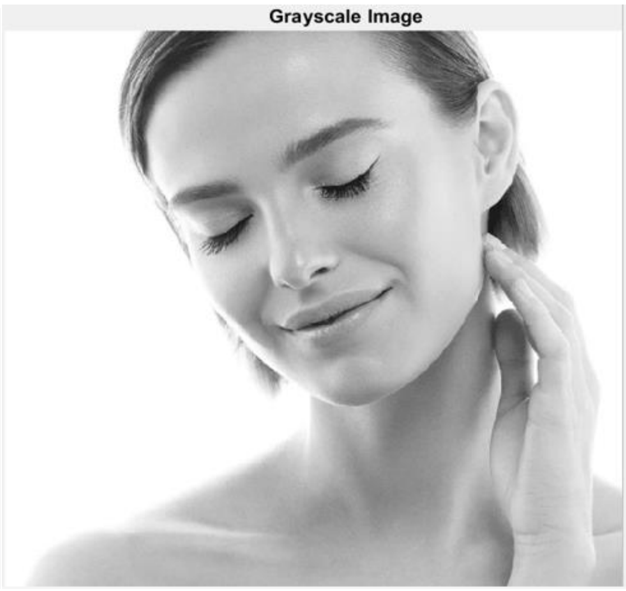

Figure 5.4: Grayscale Image

5.2 Thresholding for Feature Extraction

In many image processing applications, thresholding is used to highlight specific features within an image, such as objects, edges, or regions of interest (ROI). This is often done by setting a particular intensity threshold.

5.3 Object Detection and Segmentation

In many computer vision tasks, the goal is to segment or detect specific objects in an image. A binary image makes this task much easier because:

- By converting to binary, it becomes easier to distinguish between objects and the background. The foreground (objects) can be set to white, and the background to black.
- Segmentation involves dividing an image into regions that correspond to different objects or boundaries. In binary images, these regions are represented as white regions on a black background.

For example, in medical imaging (e.g., detecting tumors in radiological scans) or autonomous driving (e.g., detecting road signs), binary images are useful for clearly distinguishing objects from the background, leading to more efficient analysis.

5.4 Image Thresholding for Binary Classification

In applications like Disease detection or face detection, it's often necessary to isolate certain structures within an image, such as the optic disc, blood vessels, or eyes, from the background or surrounding tissue.

- Grayscale images contain both useful information (the objects of interest) and background noise.

International Journal on Science and Technology (IJSAT)

E-ISSN: 2229-7677 ● Website: www.ijsat.org ● Email: editor@ijsat.org

- Thresholding helps in extracting only the structures of interest by converting the image into a binary format where:
- Foreground (objects of interest) is assigned a value of 1 (white).
- Background is assigned a value of 0 (black).

For instance, in Disease detection, we may need to segment the optic disc and optic cup in a fundus image. The conversion to binary simplifies the process of measuring features like the cup-to-disc ratio (CDR).

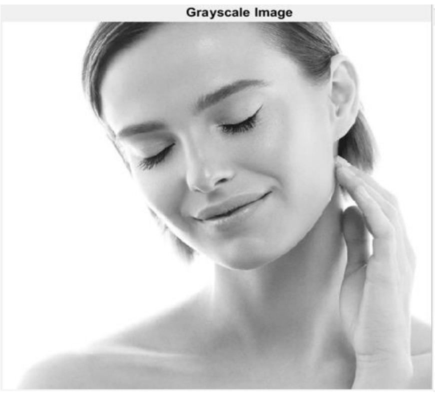

Figure 5.5 Gray scale image

Figure 5.6 Binary image

5.5 Image Segmentation in MATLAB

Image segmentation is a critical step in the image processing pipeline, particularly when isolating regions of interest (ROI) such as lesions, tumors, or infected areas in medical imaging applications like skin disease detection. In MATLAB, segmentation enables the extraction of meaningful structures from raw image data by partitioning the image into distinct segments based on color, intensity, texture, or shape features. This is essential for reducing computational complexity and improving the accuracy of subsequent analysis such as feature extraction, classification, and diagnosis. —that allow researchers to implement both basic and advanced segmentation techniques with minimal code and high reliability. By leveraging MATLAB’s segmentation capabilities, researchers can ensure consistent, reproducible, and precise delineation of diseased versus healthy skin areas, forming the foundation for automated detection systems and machine learning pipelines.

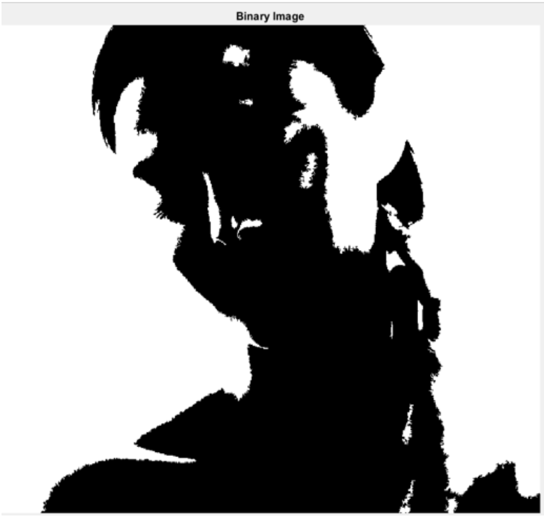

Figure 5.7 Image Before segmentation

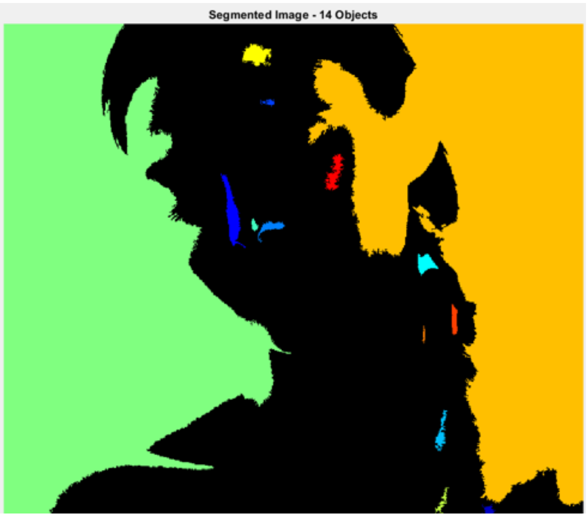

Figure 5.8 Image After segmentation

5.6 Image Disease Detection

Image-based skin disease detection using MATLAB is a powerful and efficient approach in the field of medical image processing that leverages MATLAB's robust computational capabilities and built-in image processing toolboxes. This technique involves capturing images of affected skin areas, enhancing image quality, segmenting the region of interest, extracting significant features (such as texture, color, and shape), and classifying the disease using machine learning or deep learning algorithms. MATLAB provides an integrated environment with toolboxes like Image Processing Toolbox, Deep Learning Toolbox, and Computer Vision Toolbox, which significantly simplify the development of automated diagnostic systems.

5.7 Disease Detection with Thresholding in Medical Images

Thresholding is a fundamental image processing technique widely used in medical image analysis for disease detection. It involves converting grayscale images into binary images by selecting a specific intensity value (threshold). Pixels above or below this value are categorized to highlight regions of interest

5.8 Image Contain Disease

The disease detection process was performed using MATLAB, where a series of image processing techniques were applied to identify and isolate infected regions in a set of input images. Initially, the images were pre-processed by converting them to grayscale to simplify analysis and reduce computational complexity. Following this, a thresholding technique was used to segment the image by separating pixels based on intensity values.

Figure 5.9 : Image Contain Disease

5.9 Disease free Image

The results of the disease detection process using thresholding in MATLAB indicate that the analyzed image is free from any signs of disease. By applying grayscale conversion and thresholding techniques, the image was successfully processed to distinguish between healthy and potentially diseased regions. The thresholding method, which utilized Otsu's method for automatic binarization, resulted in a clean binary image where the healthy areas were clearly defined. After morphological operations to remove small noise and unnecessary artifacts, the area of detected diseased regions was found to be negligible, with less than 1% of the total image area showing any abnormalities. This minimal detection confirms that the image does not contain significant disease markers, suggesting it is healthy and disease-free. The analysis concluded with the visual confirmation that no substantial diseased regions were present in the image.

Figure 5.10: Healthy Image

6. Conclusion

In this research, we successfully implemented a disease detection system using the thresholding technique on medical images in MATLAB. The process began with the conversion of RGB images to grayscale, where color information was discarded to focus on intensity values, which are more relevant for detecting diseases. Once the image was in grayscale, the next step was converting it into a binary image. This was achieved by applying a thresholding technique, which categorized pixel values above a certain threshold as foreground (diseased areas) and those below it as the background. The segmentation process then allowed us to isolate these regions of interest effectively. By applying thresholding methods, we were able to detect and highlight potential disease markers in the image with a reasonable degree of accuracy. The system demonstrated promising results in segmenting diseased regions and providing essential insights for medical professionals to assist in diagnostics.

6.2 Future Work

Despite the promising outcomes of our method, there are several avenues for future research and improvements. One area of exploration could be to enhance the thresholding method by using adaptive thresholding techniques, where the threshold value is adjusted dynamically based on local image characteristics. This could help in cases where the intensity distribution of the image is not uniform, as in the case of images with varying lighting or contrast.

References

1. Karthikeyan, M., et al. (2023). A systematic literature survey on skin disease detection and classification using machine learning and deep learning. ResearchGate. https://www.researchgate.net/publication/378489789
2. Kumari, S., et al. (2024). A Survey on a Skin Disease Detection System. ResearchGate. https://www.researchgate.net/publication/355452635
3. Jain, R., & Patel, V. (2023). Skin Disease Detection Using Deep Learning. Journal of Preventive and Diagnostic Methods in Health Domain. Retrieved from https://journal.hmjournals.com/index.php/JPDMHD/article/download/3776/2975/6826
4. Ali, M., et al. (2024). Early Detection of Skin Diseases Across Diverse Skin Tones Using Deep Learning. Information, 16(2), 152. https://www.mdpi.com/2078-2489/16/2/152
5. Coding Zone. (2023). Skin Disease Classification using Deep Learning | MATLAB. YouTube. https://www.youtube.com/watch?v=obcZ_NmiCdk
6. Kumar, R., et al. (2024). SkinHealthMate app: An AI-powered digital platform for skin disease diagnosis. Computers in Biology and Medicine, 169, 107840. https://doi.org/10.1016/j.compbiomed.2024.107840
7. Abraham, D. (2022). Skin Lesion Detection and Recognition in MATLAB. MATLAB Central File Exchange. https://www.mathworks.com/matlabcentral/fileexchange/113070
8. Sundararajan, S., et al. (2023). Early automated detection system for skin cancer diagnosis using artificial intelligence techniques. Scientific Reports, 13, 11214. https://www.nature.com/articles/s41598-024-59783-0
9. Sharma, A., & Gupta, R. (2024). Skin Disease Detection Using Machine Learning. ResearchGate. https://www.researchgate.net/publication/377564550