ABSTRACT
Digital Image Processing (DIP) provisions robust research platform in areas of epidermis, dermis, and subcutaneous tissues. The skin is the principal organ of the human body, containing blood vessels, lymphatic vessels, nerves, and muscles, which can perspire, perceive the external temperature, and protect the body which can be faced larger problem directed by any skin disease. This research deals with various image processing techniques, image segmentation shows a vital role in step to analyze the given image and has become a prominent objective in computer vision. This work deals on the basic principles on the methods used to segment the infected part in an image and pre-processing of images to enhance the quality on the four diseases namely: Seborrhoeic Dermatitis, Diabetic Foot Ulcer, Impetigo, and Melanoma. Here, three segmentation methods for the given four diseases are evaluated for the efficient use for the medical purpose.

Keywords
Skin disease, Segmentation, K-Means, Marker-controlled Watershed, Otsu thresholding, Jaccard Index, Dice Coefficient

1. INTRODUCTION
Image heightening and image based pattern recognition is working on the various image processing techniques, image segmentation is the fundamental step to analyze images and extract data from them. Separation algorithms are based on two properties similarity and discontinuity tranquil different skin disease segmentation processes are used by many authors but this system is implemented on the four diseases namely: Seborrhoeic Dermatitis, Diabetic Foot Ulcer, Impetigo, and Melanoma. Impetigo is a common and highly contagious skin infection by which infants and children are mainly infected. Melanoma, also known as malignant melanoma, is a type of cancer that is developed from the pigment-containing cells known as melanocytes. A diabetic ulcer is a major complication of diabetes mellitus and probably the major component of the diabetic foot. Seborrhoeic Dermatitis is a common skin condition that mainly infects our scalp. Patients’ lives are seriously affected by all these four diseases. More than a cosmetic nuisance, anxiety, depression, and other psychological problems can be produced by a skin disease that affects patients’ lives in ways comparable to arthritis or other disabling illnesses [1].

There has also been an increasing interest in applying soft segmentation algorithms, where a pixel may be classified partially into multiple classes, for MR images segmentation. Clustering is a method of grouping a set of patterns into a number of clusters such that similar patterns are assigned to one cluster. Each pattern can be represented by a vector having many attributes. The main advantage of clustering is that interesting patterns and structures can be found directly from very large data sets with little or none of the background knowledge. The cluster results are subjective and implementation dependent. K-means clustering is a simple clustering method with low computational complexity. The clusters produced by K-means clustering do not overlap.

In computer vision and image processing, Otsu's method is used to perform automatic image thresholding. In the simplest form, the algorithm returns a single intensity threshold that separate pixels into two classes, foreground and background.

The method involves iterating through all the possible threshold values and calculating a measure of spread for the pixel levels each side of the threshold, i.e. the pixels that either fall in foreground or background. In this paper, image processing is used to pre-process and segment four different skin disease images that occur in humans. Finally, the segmentation methods are compared using Jaccard Index and Dice Coefficient to find out which method is best for which particular disease. All the systems developed till now have worked on segmenting similar types of diseases (similarity based on color, shape etc.) but not on different diseases. There are few works on Impetigo and Diabetic Foot Ulcer diseases. These things motivated to implement the proposed system.

The rest of this paper is organized as follows. Sections 2 and 3 describe the literature review and proposed solution. Section 4 provides details of the implementations of the pre-processing, segmentation and comparison parts. Section 5 describes the result analysis. Finally, section 6 and 7 draw the conclusion of this paper with a few comments and suggestions on future research.

2. LITERATURE REVIEW
2.1 Image Segmentation
Image segmentation algorithms are classified into two types, supervised and unsupervised. Unsupervised algorithms are fully automatic and partition the regions in feature space with high density. The different unsupervised algorithms are Feature-Space Based Techniques, Clustering (K-means algorithm, C-means algorithm), Histogram thresholding, Image-Domain or Region Based Techniques (Split-and-merge techniques, Region growing techniques, Neural network based techniques, Edge Detection Technique), Fuzzy Techniques. The watershed segmentation technique has been widely used.
in medical image segmentation. Watershed transform is used to segment gray matter, white matter and cerebrospinal fluid from Magnetic Resonance (MR) brain images. The method originated from mathematical morphology that deals with the topographic representation of an image. Watersheds are one of the typical regions in the field of topology. A drop of the water falling it flows down until it reaches the bottom of the region. Monochrome image is considered to be a height surface in which high-altitude pixels correspond to ridges and low altitude pixels correspond to valleys. This suggestion says if we have a minima point, by falling water, region and the frontier can be achieved.

2.2 Disease Detection

ISIC has worked on developing algorithms for automated diagnosis of melanoma, the most lethal skin cancer. Thresholded Jaccard, Balanced Thresholded Jaccard, multipartition test sets, etc. has been used. Li-sheng Wei et al. has proposed a recognition method on texture and color features and used vertical image segmentation to identify three common skin diseases Paederus, herpes, Psoriasis [1].

PSL images analysis based on texture and morphological features of the images has been done by Lakshay Bajaj. The disease region is segmented using the active contour and preprocessing algorithm. Both the methods are compared and give equal results but it is found that active contour has a flaw; the user needs to mark the region around the infected area. Therefore, Pre-Processing Algorithm is given preference over active contour because of its autonomous nature and reducing the involvement of users [1] [2] [3].

There are many proposed algorithms that enhances the image by focusing on parameters like contrast, brightness adjustment. On different study it is found a novel method where they have used DCT (Discrete Cosine Transform), DWT (Discrete Wavelet Transform) and SVD (Singular Value Decomposition). Some of the researcher have used median filtering for removing noises and the Maximum Entropy Thresholding method for segmentation [4] [5] [6].

2.3 Enhance the Image

Image are converted a greyscale image, sharpening filter, median filter, smooth filter, binary mask, RGB extraction, histogram and Sobel operator [6]. Comparison of K-means clustering and marker-controlled watershed algorithm with Fuzzy C-means clustering and marker-controlled watershed algorithm has been done. The integration of K-means clustering with the marker-controlled watershed algorithm gives better segmentation [7]. The current practices used by dermatologists include biopsy, scrapings, diascopy, patch testing and prick Test which are invasive methods of detection. In patch testing and prick test, the allergen is directly applied to the infected area. The skin might take a long time even several days to react to the allergen [8].

The work on Impetigo and Diabetic Foot Ulcer are low. Marker controlled watershed and the K-Means algorithm have been used by few authors. There is no comparison of segmentation methods to clarify which segmentation gives the best result. Boundary regioning, differentiating pixel distribution, waterbag or pus validation in Diabetic Foot Ulcer are not done by any author. Some research article shows that the uses of dull razor for median filtering of the noises like hairs and pigment. Also for segmentation purpose, region growing has been used which can localize suspicious lesion region in PSL images [9] [10] [11].

Modified Sobel operator based on color instead of gray before segmentation of images is called color gradient generation. After color gradient generation, a threshold is applied and k-means clustering is performed on the color gradient. Then morphological closing is performed on the clusters to obtain the binary mask and by applying the mask, segmentation of the diseased part from the healthy skin has been done. After that different machine learning process have been used for segmentation. Firstly, the input images are enhanced for better processing then, the lesion portion is segmented from the enhanced image by two methods: Otsu thresholding and Morphological operations. The authors have done the preprocessing and segmentation in this manner. Input skin images are resized for optimal data resolution and processing speeds. Artifacts on the skin; i.e., hair is isolated, filtered and subjected to illumination equalization. Otsu histogram-based thresholding is utilized for isolating the lesion from the surrounding skin for analysis [12] [13] [14][15] [16] [17].

Segmentation is performed using otsu thresholding method which will convert binary image. After Otsu thresholding, edges of the output image become irregular. The authors also use contour based image segmentation algorithm using morphological edge detection. K-means, FCM, IFCM, Otsu’s method and Active Contour techniques are applied [18][19][20][21].

2.4 Image Restoration

We have found image enhancement, image restoration and hair removal techniques through applying thresholding, color-based segmentation algorithm, discontinuity-based segmentation, edge-based segmentation and soft computing.

Filtering is performed on image which is a non-linear process used for enhancing the overall image by preserving the edges of the image. Otsu method is performed where image is segmented into 3 levels using IM Quantize with 2 threshold level. Different machine learning algorithms are also implemented for segmentation [22] [23] [24].

There are also certain disadvantages of the current image processing techniques used for skin disease segmentation. Works have been done on some types of diseases either based on color or shape etc. The work on Impetigo and Diabetic Foot Ulcer are low. Marker controlled watershed and the KMeans algorithm have been used by few authors. There is no comparison of segmentation methods to clarify which segmentation gives the best result. Boundary regioning, differentiating pixel distribution, water bag or pus validation in Diabetic Foot Ulcer are not done by any author.

3. PROPOSED METHODOLOGY

In our model, the images of the different skin diseases will be given as input and the model will segment the infected part of the four diseases - Impetigo, Melanoma, Diabetic Foot Ulcer, and Seborrheic Dermatitis. But to get the output there is some process which needs to be done.

3.1 Iseases being considered

There are many differences between the four diseases considered. Seborrheic dermatitis infects the scalp, central face, and anterior chest. Several factors have been associated with seborrhea that could play a role in its etiology, including hormone levels, fungal infections, nutritional deficits, and neurogenic factors. In infants, seborrheic dermatitis might present as thick, greasy scales on the vertex of the scalp (i.e,
cradle cap). The condition is not pruritic in infants as with older children and adults. The scales can vary in color, appearing white, off-white, or yellow. In adolescents and adults, it often presents as scalp scaling, i.e. dandruff. Along with the scalp signs, seborrheic dermatitis might also cause mild to marked reddenning of the nasolabial fold, often with greasy scaling [25]. The clinical picture, in this case, is not typical of impetigo. Impetigo is a highly contagious infection of the superficial epidermis; it is most common in young children but can affect any age group. It is spread through direct contact, and the incidence is increased during the summer months. Impetigo presents as a single red macule or papule that quickly becomes a vesicle. Melanoma is a type of skin cancer that can begin as a mole or wart. It kills more people than any other form of skin cancer and can spread to other areas of the body. On some people, a melanoma may look like seborrheic dermatitis. People with a history of seborrheic dermatitis might not notice melanoma in its early stages if they are accustomed to unusual skin growths [26]. Melanomas can develop anywhere on the body. They most often develop in areas that have had exposure to the sun, such as our back, legs, arms, and face [27].

Diabetic ulcer only occurs in the foot. Impetigo sores can appear anywhere on the body, but children tend to get them on their faces. Sometimes they show up on their arms or legs. The infected areas range from a dime to a quarter size. They start as tiny blisters that break and reveal moist, red skin. After a few days, it gets covered with a grainy, golden crust that gradually spreads at the edges [28]. Diabetes ulcer is a metabolic disorder and hence the defects observed in diabetic wound healing are thought to be the result of altered protein and lipid metabolism and thereby abnormal granulation tissue formation. Increased glucose levels in the body end up in uncontrolled covalent bonding of aldose sugars to a protein or lipid without any normal glycosylation enzymes [29]. Thus the 4 diseases are different from one another by their causes to have happened and other features.

3.2 Block Diagram

Figure 5: Block diagram of the proposed system

As can be seen from the block diagram, the proposed system consists of three phases: pre-processing, segmentation and comparison. Before any analysis of the input image is done, image pre-processing is done so that all images are consistent in desired characteristics. Resolution matching is done on the image so that they are all the same size (600x450). 600x450 is chosen so that the processing time is less. Furthermore, in this phase, the images are resized, then the contrast is enhanced, the histogram is calculated, also the comparison of histograms have been done to differentiate the distribution of pixels, again waterbag or pus validation of diabetic foot ulcer is implemented. In Melanoma images the hairs are present as noise so these are removed using the Dull Razor algorithm. The boundary regioning has also been done. In the segmentation phase, the marker-controlled watershed algorithm with clustering and K-means algorithm and Otsu thresholding have been used. In the comparison phase, the segmentation processes are compared using the Jaccard index and Dice coefficient.

4. IMPLEMENTATION

For implementation, 5279 images of Melanoma, 5025 images of Impetigo, 5019 images of Seborrheic Dermatitis and 5048 images of diabetic Foot Ulcer are collected and then all the images are resized using Plastiliq Image Resizer, after that the contrast is enhanced and the histogram analysis is done. Finally, three different methods: Marker controlled Watershed algorithm, K-means algorithm and Otsu Thresholding are used separately to segment the infected region.

4.1 Pre-processing

4.1.1 Contrast Enhancement

Contrast is an important factor in any subjective evaluation of image quality. Contrast is created by the difference in luminance reflected from two adjacent surfaces. In other words, contrast is the difference in visual properties that makes an object distinguishable from other objects and the background. In visual perception, contrast is determined by the difference in the color and brightness of the object with other objects. Low contrast image values concentrated near a narrow range (mostly dark, or mostly bright, or mostly medium values). Contrast enhancement changes the image value distribution to cover a wide range. The contrast of an image can be revealed by its histogram. Contrast enhancement in this system is used for doing better segmentation [30]. The visual representation of contrast enhancement is below:
4.1.2 Histogram
An image histogram represents the distribution of image intensity values for an input digital image. Histogram manipulation is often used to modify image contrast or for image segmentation when the range of values for the desired feature is clearly definable. In an image histogram, the x-axis shows the gray level intensities and the y-axis shows the frequency of these intensities. Histogram is used for visualizing the distribution of pixels in the images. Furthermore, the histogram of different diseases has differences and this help to differentiate between the diseases properly.

4.1.3 Removal of Hairs
Some images of Melanoma included artifacts, mostly hair; these artifacts are misleading for the segmentation algorithms and give over-segmentation. To solve this problem, the Dull Razor algorithm is used so that the hairs could be removed. The Dull Razor technique, an artifact removal pre-processing technique, deals well with hair and other artifacts. However, it tends to erase the details of the image by making the pigmented network unclear. The results are much more interesting than those achieved by the median filtering [31]. Firstly, all images are turned into gray images. Then closing operation is applied on the gray image for removing the hairs. After removing hair, the image is sharpened and smoothed. In this regard bilateral filtering and median filtering are used correspondingly.

4.1.4 Boundary Regioning
Size () function return a row vector which is converted into arrays of zeros using zeros () function. Then loops and conditions are used to create a new image, where if the pixels are below 75 in the input image then the new image contains pixels of value 0 (black), if the pixels are below 100 in the input image then the new image contains pixels of value 132, if the pixels are below 135 in the input image then the new image contains pixels of value 200 and the rest with value 255(white). The output of boundary regioning is shown below:

4.1.5 Waterbag Validation
To check whether there is any waterbag or pus in the hole of a diabetic foot ulcer, the image is resized first, and then the ‘Salt-and-Paper’ noise is added. Then by median filtering, the noise is removed, after that bilateral filtering is done and at last the edges are detected by edge base segmentation, which is done by Sobel operator. The output of waterbag or pus validation:

4.1.6 Comparison and Analysis of Histogram
Comparison of histograms of the four different diseases has been done. The images of four diseases differ in the distribution of pixels.
4.2 Segmentation

4.2.1 Marker Controlled Watershed Algorithm
First of all, an input color image is converted into a gray image. Adapthisteq function is used to enhance the contrast of the grayscale image. Graythresh computes a global threshold level that can be used to convert an intensity image to a binary image with imbinarize. The gray thresholding function uses Otsu’s method, which chooses the threshold to minimize the intraclass variance of the black and white pixels. So, this way after converting into a binary image, the morphological opening is performed on the binarized image. The morphological open operation is an erosion followed by a dilation. Bwareaopen is used to reduce noise. Bwperim calculated the perimeter of the objects. Then, regional maxima are calculated, which is actually the foreground marker. But some shadowed images are not marked and go directly to the edges. That is why the edge of the marker is cleaned and shrunk. That was implemented by doing closing by erosion which created some isolated pixels which are removed by bwareaopen. Bwdist computes the Euclidean distance transform of the binary image. Watershed returns a label matrix L that identifies the watershed regions. The watershed transforms find the "watershed ridge lines" in an image by treating it as a surface where light pixels represent high elevations and dark pixels represent low elevations, and this is the background marker. After that, the gradient magnitude of the gray image is calculated. Imimposemin calculates the regional minimal of the gradient image at the foreground and background pixels. Finally, watershed segments the infected part. The output of marker-controlled Watershed:

![Image](image1.png)

**Figure 12**: (a) Input Image, (b) Background-Foreground Markers and Object Boundaries Superimposed

![Image](image2.png)

**Figure 13**: Different phases of Marker Controlled Watershed Algorithm

4.2.2 Otsu Thresholding
In image processing, Otsu’s method, named after Nobuyuki Otsu (O Tsu Nobuyuki), is used to perform automatic image thresholding. In the simplest form, the algorithm returns a single intensity threshold that separates pixels into two classes, foreground, and background. By using Otsu thresholding, the infected and non-infected region can be detected properly. As the intensity of the infected region is more than the non-infected, so it will become black and the rest of the portion remain white. Thus, Otsu helps to differentiate between the infected and non-infected areas. Here the visual representation of Otsu thresholding:

![Image](image3.png)

**Figure 14**: (a) Input image of Melanoma, b) Output after Otsu Thresholding

4.2.3 K-means clustering
Clustering algorithms are unsupervised algorithms but are similar to Classification algorithms but the basis is different. K-Means clustering algorithm is an unsupervised algorithm and it is used to segment the interest area from the background. It clusters or partitions the given data into K-clusters or parts based on the K-centroids. In the proposed system, the images are clustered by the intensity of the infected area present in those images. First of all, a centroid is chosen and clusters are made. The pixel values which are near the centroid create one cluster and the others are in another cluster. Then the centroid value is also updated. The process repeats 5 times and within this iteration, the desired output is found. The output of K-means clustering for the below image:

![Image](image4.png)

**Figure 15**: Input Image of Melanoma

![Image](image5.png)

**Figure 16**: K-means Algorithm for 5 clusters of the Melanoma image

4.3 Comparison of segmentation processes
This system used three types of segmentation processes for segmenting the infected parts in the images. In this regard, two parameters are used to measure the performance of the segmentation processes. One is Jaccard index and the other is Dice Coefficient. The results from both the parameters show which process is better for segmentation for which particular disease among the four.
Figure 17: Comparison of segmentation processes by Jaccard Index

Figure 18: Comparison of segmentation processes by Dice Coefficient

5. RESULT ANALYSIS
There are buttons labeled with disease names which lead to the subGUIs where further image pre-processing and segmentation are done. The sub GUI is shown below:
There is also a button labeled with comparison that leads to a GUI where histogram analysis and comparison of the four different diseases are done.

5.1 Pre-Processing

In the pre-processing part, resizing, contrast enhancement, calculation of histogram, boundary regioning, waterbag/pus validation of Diabetic Foot Ulcer, histogram analysis of different diseases and hair noise removal are done. For contrast enhancement and histogram calculation, an image is loaded first into the sub GUI:

5.2 Marker controlled Watershed Algorithm

Using this algorithm, the infected part is segmented and to avoid over segmentation, two markers are calculated - the foreground marker and the background marker. But in case of Diabetic Foot Ulcer and Melanoma, it gave a slight over segmentation.

5.3 Otsu Thresholding

In case of Diabetic Foot Ulcer, it gave a slight over segmentation, but for rest of the three diseases, it gave a very good segmentation.

5.4 K-means algorithm with 5 clusters

This method gave a good result for all the four diseases but at different cluster numbers. Melanoma is segmented best at cluster 3, same goes for Diabetic Foot Ulcer. Seborrheic Dermatitis is best segmented at cluster 2 and finally Impetigo also at cluster 2. In case of more different images are given as input, cluster 4 and 5 are calculated for more perfection.
5.5 Comparison of segmentation processes

With the help of Jaccard Index and Dice Coefficient, it is found that Melanoma images are better segmented using Otsu Thresholding, Impetigo images using Marker Controlled Watershed algorithm and the rest two diseases, which are Seborrheic Dermatitis and Diabetic Foot Ulcer using K-Means algorithm.

6. CONCLUSION

The newborns and infants get infected with the deadly Impetigo disease. Again, Melanoma, the most lethal skin cancer can be the cause of death. Furthermore, with the increase of Diabetics patients, disease like Diabetic Foot Ulcer is also expanding rapidly. As the result, image segmentation is affected by lots of factors, such as homogeneity of images, spatial characteristics of the image continuity, texture and image content. In this work, the system with three techniques of image segmentation have been discussed which is implemented on four different diseases.

7. FUTURE SCOPE

This system can be used for feature extraction and classification of these four diseases, also for different diseases using machine learning. Further decreases the buffer time before the patient can reach the dermatologist. It can further be improved by providing precautions and immediate relief measures which the patient can follow so as to take care not to aggravate the disease.

8. REFERENCES


