MATLAB image processing tool-based GUI for high-throughput image segmentation and analysis to study structure and morphology of skin H&E stained sections

IOP Conf. Series: Materials Science and Engineering **737** (2020) 012232 doi:10.1088/1757-899X/737/1/012232

MATLAB image processing tool-based GUI for high-throughput image segmentation and analysis to study structure and morphology of skin H&E stained sections

O S OSMAN[1,2]

[1] Komar University of Science and Technology, Computer Engineering Department, Sulaimani, KRG-Iraq
[2] University of Sulaimani, Administration Department, Sulaimani, KRG-Iraq

Email: osman.sharif@komar.edu.iq

Abstract: The complex structure of skin tissue can make the analysis of high-throughput data manually inconvenient and leads to inaccurate analysis and time consumption. Therefore, automated system that can segment and detect features which might provide critical information for interesting phenotype is required. User friendly graphical user interface GUI in MATLAB can provide facilities to create a tool to enhance, segment and analyse images without having expert skills in image processing, this can be used in the study of skin morphology phenotyping to find interesting morphological and metabolic phenotypes. Using image processing capability facilitates to develop a tool to analyse a range of different images in term of intensity and quality because of the variation in histology performed in different laboratory. Consequently, develop of automated high-throughput bioimaging tool is considered to be a very important topic in disease diagnosis and drug development. Significant assessment of the morphological features in H&E skin section through the use of GUI MATLAB tool by quantifying all of epidermal and dermal thickness and number and size of adipocyte in subcutaneous. Using our developed tool, we were able to detect interesting epidermis, dermis and adipocyte phenotypes in mice skin sections. The Morphological Bio-imaging Tool provides facilities in the high-throughput analysis of H&E skin section to understand genetic basis of diseases.

Keywords: Morphology, GUI, Phenotype, Histology, Skin layers, high-throughput analysis, Image processing.

1. Introduction

The skin consists of epidermis outer layer, dermis lay underneath and a deeper adipose layer (the subcutis layer). In human body, epidermis thickness depends on body site from 0.8 to 1.5 mm [1]. The dermis is varying in thickness from about 0.6 mm on the eyelids to 3 mm in plantar skin [2]. Subcutis layer lies below the dermis and it contains loose connective tissue and fat cells, the size and number of adipocytes can change in different pathological states, so a decrease or increase in fat can indicate an abnormality e.g. obesity or diabetes [3].

The changes in skin layer morphometry are reflecting critical genetic functions and diseases. Manual quantification and analysis of high-throughput data is a hard task for the observer to determine an objective result, and furthermore, the automated analysis is required to gain accurate analysis and

reduce time taken. The relationship between the thickness of the stratum corneum and the cellular epidermis was investigated to find any correlation between age, gender, body site, pigmentation, blood content, smoking history and skin type [4]. The analysis of relationships between the epidermal sub-layers also papillary dermis and the age of the patient, type of psoriasis, total body surface area involvement, scalp and nail involvement, duration of psoriasis, and family history of the disease was performed [5]. Such study requires automated, unsupervised techniques to assess skin in high throughput histopathological analysis. Adipocyte size can be quantified using efficient and accurate automated methods instead of manual calculation, which is very complicated and time consuming [6]. Adipocyte size in the reproductive fat pad was determined efficiently using computer image analysis by measuring cross-sections of cell [7].

Although, the above studies investigate morphological features in skin section, they are not sufficient to obtain accurate analysis in high-throughput screening. In histology image analysis, common techniques used for segmentation are pixel intensity-based technique (such as thresholding) that depends on the intensity value and colour of the pixels in that image to isolate the foreground and background [8]. This technique is not enough in histology high-throughput analysis because of the variation in tissue processing during specific staining schemes produces variation in the intensity and colour of the target objects. For this reason, automated methods have been created to work adaptively with these variations in different datasets. Techniques used in this tool and work adaptively with different datasets include the active contour- based analysis that starts by initializing a curve automatically from information known about the object itself. This internal control makes the system adapt to the intensity and colour variation in the images.

MATLAB allows creating a graphical user interface GUI to implement methods and algorithms in a convenient way for the user to enhance image, segment and quantify incremental features in accurate manner and time saving. In this research a powerful tool has been created using MATLAB GUI that overcomes the subjectivity of manual process and also to work adaptively with different data sets that are slightly different in intensity because of variation of histology process.

2. Implementation

The main objective of creating this tool is to detect early skin phenotypes in mice by features detection in Haematoxylin and eosin H&E skin section, segment skin layers and quantify each of skin layers thickness and size and number of adipocytes in subcutis layer. The tool created using MATLAB that allows creating a GUI for user who do not need to be familiar with MATLAB environment neither expertise in image analysis. The software interface consists of buttons and axes to implement image processing and analysis algorithms and show the result images followed by quantifying interested morphological features and save the result in an excel sheet. The user interface and example of segmentation is shown in (Figure 1). The analysis can be performed either for individual analysis to segment and analyse one image of the skin tissue or for batch analysis if the user needs to analyse many images of *n* number of biological replicates.

The image processing techniques in this friendly user graphical interface is able to segment image into different regions that are different in there pixel properties utilising a region based active contour [9]. Active contour is easy to implement and is based on global information in the image instead of the gradient or local information in the image and at the end, results to a closed contour as a segmented image as given by the following equation [10]:

$$F(c_1, c_2, C) = \mu L(C) + \nu A(inside(C)) + \lambda_1 \int_{inside(C)} u_0(x, y) - c_1|_2 dxdy\,| + \lambda_2 \int_{\int outside(C)} |u_0(x, y) - c_2|_2 dxdy$$

Where the L is the length of the curve and A is the area of the region inside the curve, μ, ν, λ1 and λ2 are fixed parameters, where μ≥0, ν≥0, λ1 and λ2≥0, c1 and c2 are the means of the region inside and outside C.

2.1. Pre-processing

Prior to segmentation step multiple pre-processing and refinement steps are needed, including contrast enhancement using contrast stretching [11], and using a median filter to remove noise in the image. Additional step used to segment epidermal layer by adding pixel intensity in the image with the corresponding pixel intensity in same image to produce an output image more suitable to detect epidermis in the image by turning epidermis region to darker area than other regions in the output image.

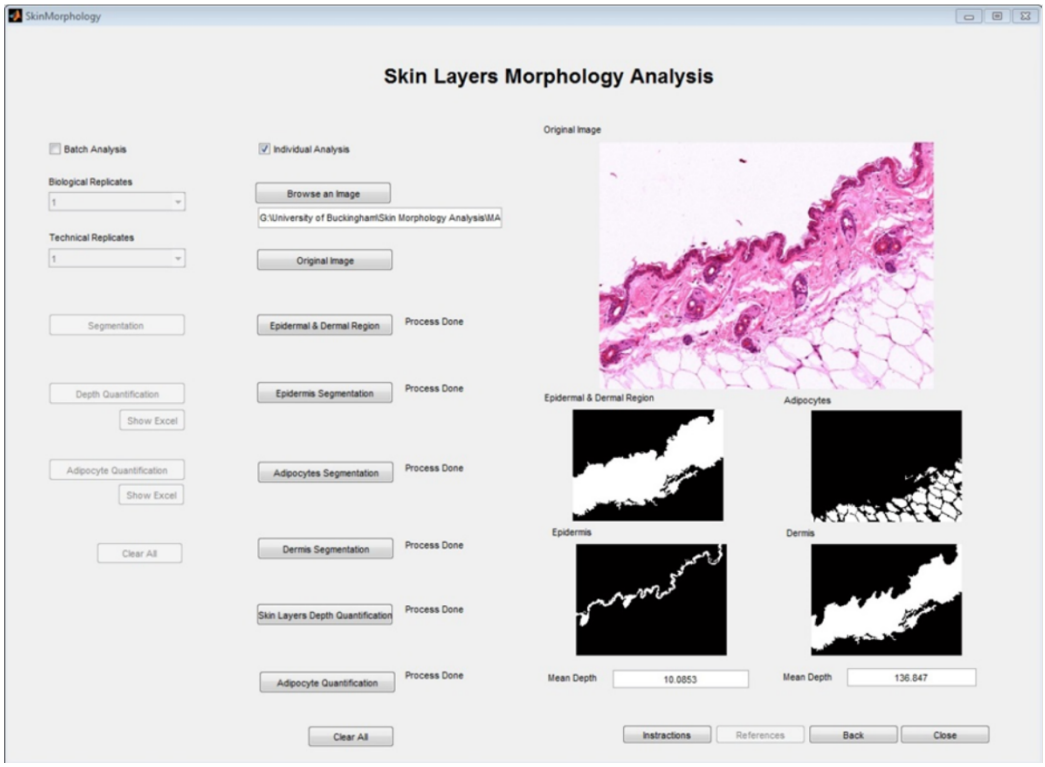

Figure 1. The skin morphology analysis graphical user interface skin layer segmentation.

2.2. Initial curve generation

Pre-processing step followed by generating initial curve as a start point of active contour model. We developed automated initialization curve generated from the ROI itself (shown in Figure 2b) to reduce the number of iterations and develop a fast segmentation independent of the object location by implementing following steps using MATLAB:

- Convert RGB colour space to HSV space.
  `ConvertedImage = rgb2hsv(image);`
- Select the saturation channel from HSV colour space.
  `SelectedImage = ConvertedImage ( : , : , 2);`
- Threshold the saturation image into a binary image using Otsu method to partition the image into two components with white pixels being the region of the epidermis and dermis, and the black pixels being the background.

- Select the largest object in the binary image that represents the initial curve.
  ```
  [L, N] = bwlabel(segmented region);
  D = regionprops(L, 'area');
  w = [D.Area];
  LargestObject = find(w == max(max(w)));
  Segmented_LargestObject = ismember(L, LargestObject);
  ```
- Apply the "open" mathematical morphology operator using the disk structuring element of size 4, to make the initialize curve cover the whole area of the dermis and epidermis in the original image.
  ```
  StructuredElement = strel('disk',4);
  InitilizedCurve= imopen(Segmented_LargestObject,se);
  ```

2.3. Epidermal and dermal region segmentation

Once the initialize curve generated, the RGB colour space converted to HSV colour space, the active contour without edge (Figure 2c) then applied to the selection of the saturation image to segment epidermal and dermal region accurately as show in (Figure 2d and 2e).

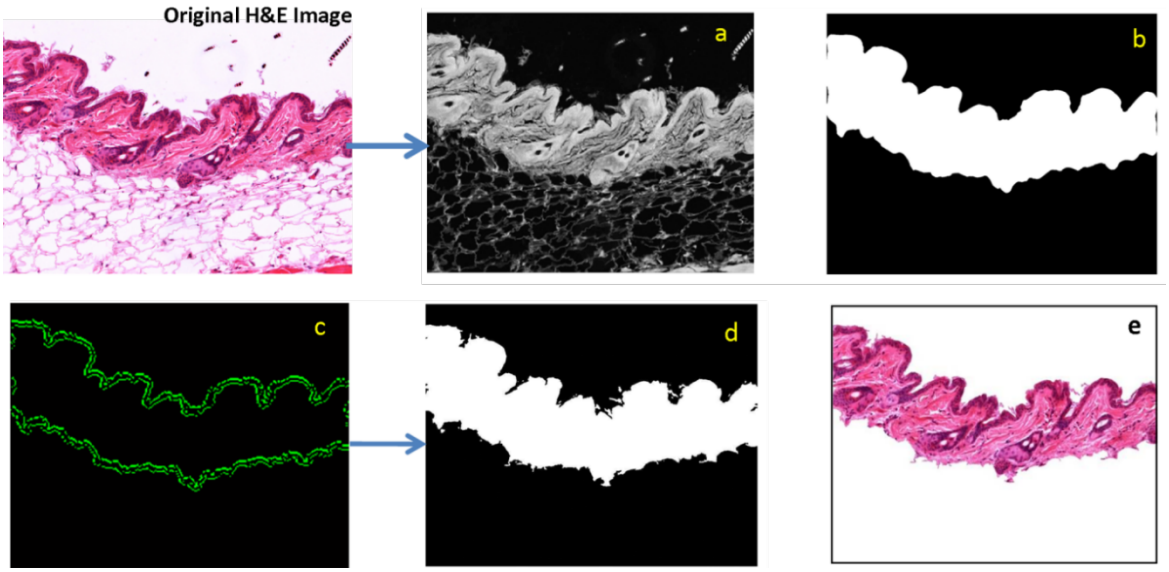

Figure 2. Overview of epidermis and dermis region segmentation using Active contour algorithm. Original H&E skin image in top panel left. a) Saturated channel of the *HSV* colour space, b) initialization start curve, c) curve evolution moves to the object boundary (active contour iterations), d) segmented binary epidermis and dermis, e) superimposed epidermis and dermis.

2.4. Epidermal layer segmentation

To segment epidermis some additional pre-processing step were used in this tool as described previously by adding the segmented epidermis and dermis image to itself to improve epidermis segmentation to enhance contrast of the epidermal region (Figure 3j). This step followed again by applying active contour segmentation to segment this layer precisely. To make the tool adaptive to segment images with intensity variation, another effective pre-processing step was added to the darker and brighter than majority images in the data set depending on the mean and standard deviation of overall pixels intensity in the image. Raise to Power Operator was used to enhance high intensity pixel values in bright images. Using this operator, each pixel intensity value in the output image is equal to a

basis value raised to the value of the corresponding pixel value in the input image; the resulting image is given by:

$$Q(i,j) = cP(i,j)^r$$

Where P is the input and Q is the output image, r >1 and c is the scaling factor. In our tool r=2, and c=2 were selected.

Another additional process was used to detect epidermal region using fuzzy c-means thresholding [12], in the images those have low contrast between epidermis and dermis region.

Using different pre-processing techniques make the tool adaptive to segment skin layers in images in various data sets as shown in (Figure 3b-d).

2.5. Dermal layer segmentation

Once the region that contains epidermis and dermis layer segmented, and also epidermal layer segmented, the dermis layer was then segmented by subtracting the segmented epidermis from the segmented epidermis with dermis image (Figure 3c).

2.6. Adipocytes segmentation

The subcutis layer isolated from saturated image by subtracting the epidermis and dermis region for the whole image (Figure 3f) that follow by the process of adipocytes segmentation. The process starts with background subtraction in the saturated image within the HSV colour space, followed by the use of the piecewise grey level linear transformation to select the intensity range (0.079-0.125) that is equivalent to the intensity of the adipose cells in the image [13]. Pixels in the selected range of intensity were converted to white, and the rest pixels were converted to black. The adipocyte segmentation was followed by fully automated post-processing steps to enhance cell boundaries, and fill holes and close gaps using mathematical morphology as shown in (Figure 3g).

2.7. Morphological feature quantification

This user-friendly morphology analysis GUI able to quantify thickness of the skin layer and also number and size of adipocytes in subcutis layer to evaluate morphological changes in the skin of different gene-knockout mice. The method quantifies segmented adipocytes (Figure 3g, 3h) categorise the cells in 10 categories with calculating number of cells in each size category as shown in Table 1. The depth of each of the segmented epidermal and dermal layers was calculated automatically by generating a skeleton of the binary image of each layer using skeletonization [11], after which the skeleton branches were used to quantify the depth by measuring the length of the branches to yield half the depth of the object. Figure 3l illustrate the result of skeletonization process.

μm^2 ↓ Size Group	No.Of Cells	%Num/Total
<500	56	40.88
500-999	42	30.66
1000-1499	22	16.06
1500-1999	13	9.49
2000-2499	3	2.19
2500-2999	0	0.00
3000-3499	1	0.73
3500-3999	0	0.00
4000-4499	0	0.00
>4500	0	0.00

Table 1. Size categories and quantification result of number and size of adipocyte in one example H&E skin image.

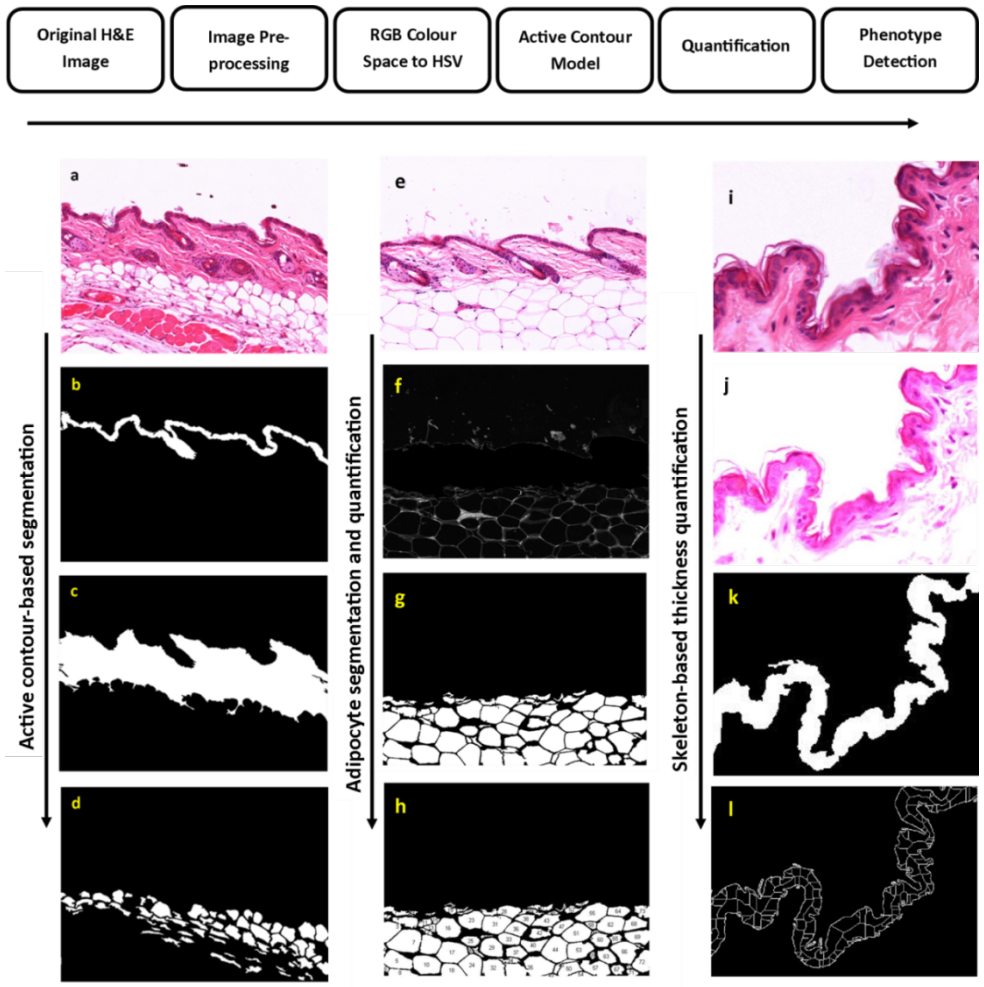

Figure 3. High-throughput cutaneous features detection and quantification. Skin layer segmentation using an active contour method (a-d), quantification of adipocyte size and number (e-h), and epidermal and dermal layer thickness quantification (i-l).

3. Result and discussion

The skin morphology analysis GUI created to evaluate skin morphology in researches of genetic basis of diseases and detecting interesting phenotypes, and also the effect of diet. Utilising the developed tool in the study of mouse genetics, the tool was able to evaluate cutaneous morphology in 475 knockout mouse lines [14]. The tool was able to segment and quantify the epidermis, dermis and subcutis layer morphometry successfully, in a huge number of H&E images, and was able to investigate the effect of different genes or environmental factors such as diet, on the morphology of the epidermis, dermis and subcutis in high-throughput screen. The tool was able to find the effect of diet on epidermal and dermal thickness and adipocyte size in wild-type animals in two genetic backgrounds (B6Brd;B6Dnk;B6N-Tyr<c-Brd>,B6N) as shown in figure 4. Epidermis thickness didn't show significant different between the two different diet groups, but dermis depth showed significant

change in different diet groups. In subcutis we found that chow fed animals contained a higher number of small adipocytes compared to high fat fed animals.

In the research project of evaluating the biological function of genes and their impact of the skin morphology, our tool showed ideal performance to identify 53 interesting adipocyte phenotypes, 18 interesting dermal phenotypes and 3 interesting epidermal phenotypes in 475 gene knockout mice described previously [14].

4. Conclusion

Automated skin morphology analysis can deal with many problems that face high-throughput analysis. The extracted morphological features from skin layers in automated high-throughput screen facilitate researches of understanding of the genetic basis of disease. The use of the active contour algorithm to adaptively segment skin layers in a huge number of H&E stained images has solved the problem of the variation in intensity and colour arising from the use of multiple laboratories and investigators, and between experiments performed in the same laboratory at different times. The novel depth quantification obtained by computing the length of skeleton branches was used successfully to quantify the depth of both the epidermis and dermis to identify interesting skin phenotypes. The automated method also provides accurate measurements of the cross-sectional area of adipocytes in histological sections. The performance of the automated algorithm was consistent compared to the utilized manual methods and, moreover, the accuracy of the unsupervised area measurements is increased. This has enabled the identification of specific gene knockouts that results in a cutaneous phenotype, which would not have been possible (or at least realistic) using existing manual methods.

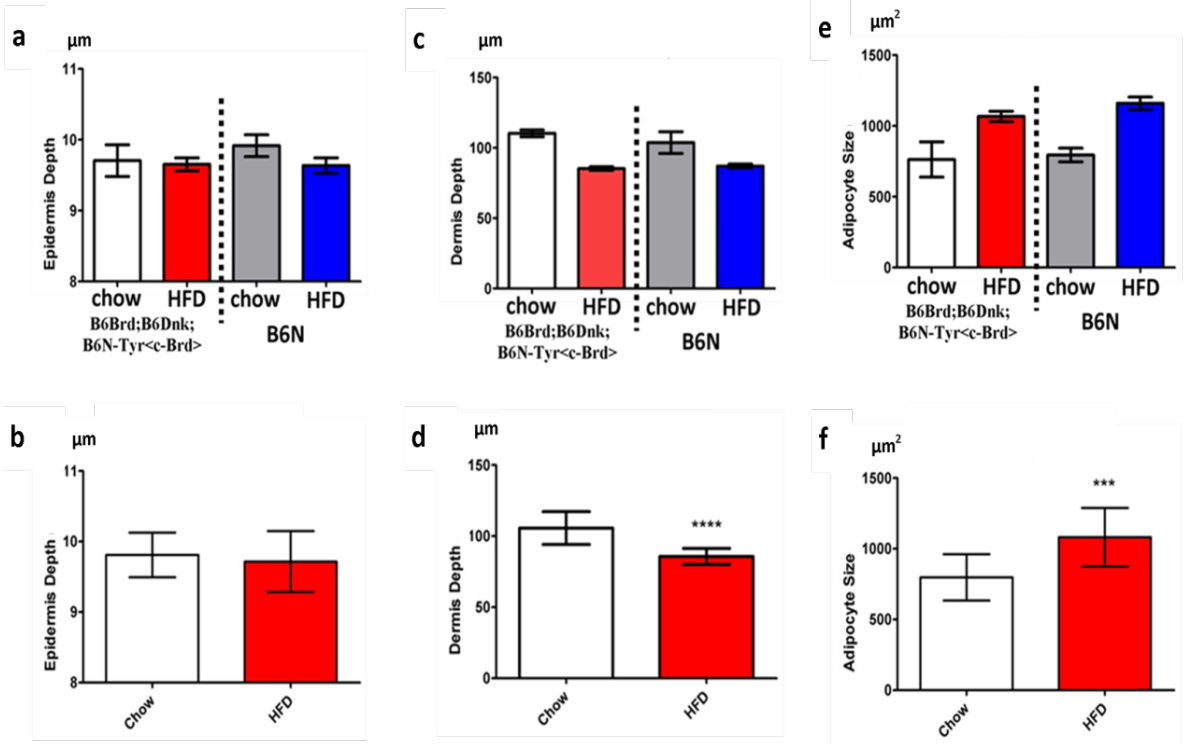

Figure 4. Effect of diet on skin layers in *wild-type* animals. a, b) the influence of diet (chow or high fat diet (HFD) in two genetic backgrounds (B6Brd;B6Dnk;B6N-Tyr<c-Brd>,B6N). c, d) the influence of diet (chow or high fat diet [HFD]), showing significant decrease in depth. e, f) Significant increase in adipocyte size in HFD animals.

References:

[1] Gawkrodger D 1997 *Dermatology: An Illustrated Colour Text Churchill Livingstone.* New York, 1997
[2] Gherardi, A., *A skin surface characterization system based on capacitive image analysis.* 2008.
[3] Jo J, et al., 2009 Hypertrophy and/or Hyperplasia: Dynamics of Adipose Tissue Growth. *PLoS Comput Biol*. **5**(3): p. e1000324
[4] Sandby-Møller, J., 2003 Epidermal Thickness at Different Body Sites: Relationship to Age, Gender, Pigmentation, Blood Content, Skin Type and Smoking Habits *Acta Derm Venereol* pp 410–413
[5] Alper M, et al., 2004 Measurement of epidermal thickness in a patient with psoriasis by computer-supported image analysis. *Braz J Med Biol Res*. **37**(1): pp 111-7
[6] Björnheden T, et al., 2004 Computerized determination of adipocyte size. *Obesity research*,. **12**(1): pp 95-105
[7] Chen H C and R V Farese, Jr., 2002 Determination of adipocyte size by computer image analysis. *J Lipid Res*, 2002. **43**(6): pp 986-9
[8] Otsu N 1975 A threshold selection method from gray-level histograms. *Automatica* **11**(285-296): pp 23-27.
[9] Yun T, et al. 2009 A region-based active contour model for image segmentation. in Computational Intelligence and Security. *CIS'09. International Conference on*. 2009. IEEE.
[10] Chan T F and L A Vese 2001 *Active contours without edges.* IEEE Trans Image Process. **10**(2): pp 266-77
[11] Gonzalez, R C, R E Woods and S L Eddins 2004 *Digital image processing using MATLAB*: Pearson Education India.
[12] Masood A and A A Al-Jumaily 2013 Fuzzy C mean thresholding based level set for automated segmentation of skin lesions. *Journal of signal and information processing,* **4**(03) p. 66
[13] Osman O S, et al. 2013 A novel automated image analysis method for accurate adipocyte quantification. *Adipocyte* **2**(3) pp 160-4
[14] J L Selway, O S Osman, M A Kepczynska, A Al-Habian, T Ditommaso, I Smyth, C Lelliott, J White, D Melvin, S Jassim and K Langlands. A novel automated image analysis method for rapidly segmenting and quantifying cutaneous features: Identification of genotypes with clinical features - *Journal of investigative dermatology*, 2013. p S142-S142