UNIVERSITAT DE GIRONA

MEDICAL IMAGING ANALYSIS

LABORATORY REPORT

Image segmentation and evaluation

Author
Mohit VAISHNAV

Supervisor
Dr. Robert MARTI
Tutor
Richa AGARWAL

Submission Date 4^{th}, April

1 Introduction

Image segmentation acts as an important step in diagnosis framework which can be used to further implement machine learning processing on a lesion location. For this particular lab, we have used dermatoscopic images where lesion in each image has to be delineated. Mainly three typical issues are found in segmentation of medical images known as *Ground truth generation, segmentation evaluation, segmentation evaluation*.

2 Implementation

We are given three types of algorithms to be compared namely *PDF based, Fuzzy C-means, Level Set* segmentation technique.

2.1 Ground Truth Generation

There are several ways to generate the ground truth segmented images of the original image which are placed in folder *imagesgt*. We have used three different techniques to obtain the segmented images which have to be combined in such a way to get a common region from all those regions. To obtain such a common image I have compared those images with the majority of the experts opinion. If more than one expert marks the image region of the same type then I chose to be with the majority, it is commonly known as *Majority Voting Scheme (STAPLE)*. All the images have been given the same name as they had while saving those fused images together.

3 Segmentation

3.1 Histogram based segmentation

Thresholding is the most commonly used segmentation method and also the simplest one [1, 2]. *Histogram* as a basis for such a thing is one such way of achieving our result. The algorithm to chose a single threshold is referred to as **intermeans** algorithm. Initially a guess is made at one possible value for the threshold which divided the pixels in two category of values. From this mean values of the two category are again calculated and re positioned in such a way that they lie exactly at half way between the two means. Mean value is again calculated and threshold is obtained. This process is repeated until threshold value stops changing.

3.2 Level Set Method

Segmentation problem is to get the region of interest which is usually obtained using control points. There are many issues in obtaining the control points, like if two separate close curves needed to be merge into one or vice-verse, how and from where this has to take place. Such kind of difficulties are alleviated using *level set* method [3, 4]. Idea here is to imbed a curve within a surface. The user initially guess the contour which is moved to the edges of the respective object by image driven forces. In such a method, two types of forces are considered, internal and external. *Internal* forces plays a role to keep the model smooth in the deformation process whereas *external* forces are useful in moving the model towards object boundary.

Deformable models are of two forms. In *Parametric* form explicit representation of the parameters are used, also known as *Snakes* which is compact and robust to boundary gaps and noise. Along with these advantages some of its drawbacks are restrictions in topological adaptability of the model, explicitly where splitting and merging are involved. Another form, *implicit active contour or level sets* are designed to tackle these topological changes naturally however they are not robust to boundary gaps.

3.3 Fuzzy Clustering

Generally there are two types of classification of image segmentation and pattern recognition: *supervised and unsupervised*. Fuzzy-C means clustering is one of the form of unsupervised clustering widely used in computer vision area. Fuzzy clustering also known as soft clustering [5] is a form of clustering where each data point belong to more than one cluster. It is used to segment the image having similar or near similar values of pixels in different clusters. In traditional approach each pixel of an image belong to exactly one single class and is known as *hard clustering* while in case of fuzzy theory, an idea was proposed of partial membership. Conventional fuzzy algorithms usually works in a noise free images which is a limitation of this technique. It does not incorporate any information about the spatial context which makes it sensitive to noise.

4 Segmentation Evaluation

Using the evaluation code provided in the lab, I ran it for all the techniques provided to us as mentioned above. Evaluation has been done over various parameters such as *Sensitivity, Specificity, Precision, Dice Coefficient, Negative predicted values, Hausdorff Distance*.

4.1 Dice Coefficient [6]

Given segmented image A and ground truth image B, dice coefficient is defined as

$$C_{dic} = \frac{2|A \cap B|}{|A| + |B|}$$

Dice coefficient represents the ratio of overlapped region between segmented region and the truth region. Maximum value, 1, represents identical value with the truth region, whereas the minimum value 0 shows the region totally missing.

4.2 Sensitivity

True positive rate, also called sensitivity and recall measures the positive pixels in the ground truth image which are positively evaluated by the segmentation technique.

$$Recall = Sensitivity = TPR = \frac{TruePositive}{TruePositive + FalseNegative}$$

4.3 Specificity

True Negative rate, also known as Specificity is the number of normal cases correctly identified as negative divided by the total number of normal cases (relates to the test's ability to identify negative results).

$$Specificity = 1 - FalsePositiverate = \frac{TrueNegative}{FalsePositive + TrueNegative}$$

4.4 Precision

Precision [7] tells us about the probability of the subject with a positive screening test that truly have a disease.

$$Precision = PositivePredictiveValue = \frac{TruePositive}{TruePositive + FalsePositive}$$

4.5 Negative Predictive Value

It is the probability that a subject with a negative screening test don't have any disease.

$$NegativePredictiveValue = \frac{TruePositive}{TruePositive + FalsePositive}$$

4.6 Hausdorff Distance

Hausdorff Distance between two finite point sets A and B is defined as

$$HD(A,B) = max(h(A,B), h(B,A))$$

where $h(A,B)$ is given by the maximum of the minimum euclidean distance between two set of points. It is usually sensitive to outliers, so it is recommended not to use it directly. Another option could be to use Average Hausdorff Distance, averaged over all points. It is computationally intensive for large images.

5 Own Implementation

Three of the top segmentation technique has already been implemented which is as described in above section. So I preferred to use the most basic of the segmentation algorithm for the comparison purpose with the existing ones. Using *Otsu's* method, I have done the thresholing of the given images and get the segmented images. Algorithm could be as described below:

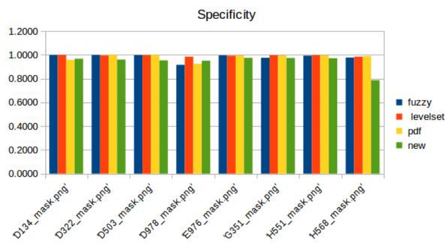
$$T = graythresh(im)\% findthethresholdforinputimage$$

$$segImg = im2bw(im, T);$$

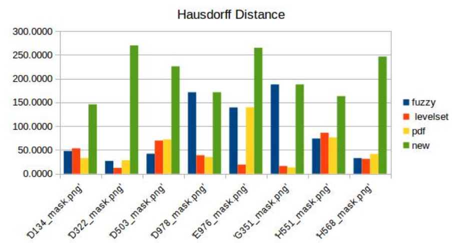
This *segImg* is in the inverted form with the binary values in a reversed way than as desired. So while saving the image, I swiped the values and then saved the image as a binary image in *png* format.

6 Results Obtained

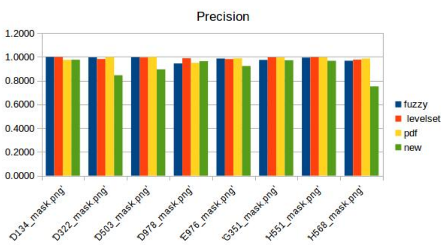
After obtaining the six evaluation parameters as discussed above, I have saved the parameter in a *dat* file format with file name **out_write.dat** which



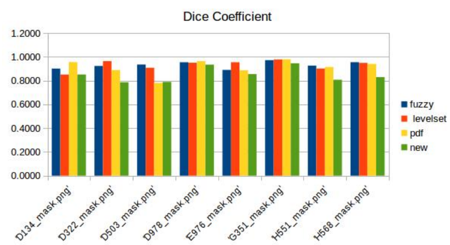
(a) Specificity



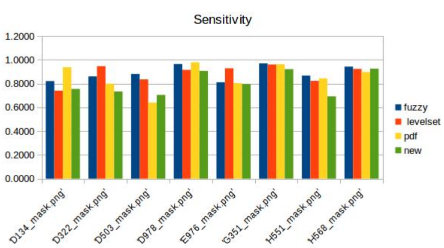
(b) Hausdorff Distance



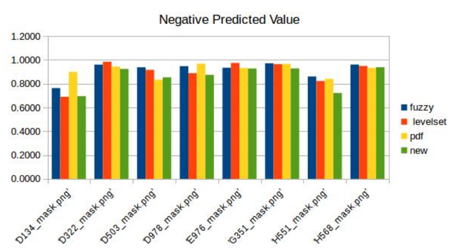
(c) Precision



(d) Dice Coefficient



(e) Sensitivity



(f) Negative Predicted Value

Figure 1: Comparison of Fusion image comparing all parameters

could be easily accessible in most of the data handling software commonly available. Dimension of the matrix is:

$$length = Numberofparameters(6) \times Numberofalgorithmsused(4) = 24$$

$$breadth = NumberofGTImages(8) \times Numberofexpertsopinion(4) = 32$$

For comparison purpose, I have preferred to use fused image and displayed in Fig. 1 each individual parameter. As expected the most basic algorithm proposed should have the least desirable results. Because we are not using any sophisticated algorithm to make our algorithm robust. Most deviation amongst all the data-set images is observed in the *H568.png* because of the least difference between the lesion and nearby region thereby making it difficult to distinguish between them. While in few images most of the algorithms have performed well because of the clearly delineated lesion region. In those images all type of algorithm shows the same kind of performance. Most deviation in between the images are observable using the *Hausdorff* distance where even the slightest of the deviation makes a big observable difference in output. Another point to be noted is from the *Negative values predicted*, where the proposed method has a similar performance wrt all the given state of the art algorithm. In terms of *specificity and precision*, my proposed method has to similar or better performance from *Fuzzy and PDF* based algorithm for few data set images.

7 Conclusion

It could be said that segmentation is a very big and emerging research area in medical domain. There are various method to segment the images and find the desired parts from the whole images where the defect is present. Due to umpteen types of the skin, texture, color etc, we cannot say if there is a particular method which could be universal. Even the algorithm I have proposed to use in this lab is to show this characteristic nature of the methods used. As observed in 1 over various parameters compared wrt different algorithms, we can say the same statement with the most variation observable in *Hausdorff Distance*. Hence there is trend in increasing deep learning methods widely applied to solve the problems in medical domain to have an accurate and precise results.

References

[1] Tutorial ntu. http://disp.ee.ntu.edu.tw/meeting/%E6%98%B1%E7%BF%94/Segmentation%20tutorial.pdf.

[2] Biomathematics and statistics scotland. http://www.bioss.ac.uk/people/chris/ch4.pdf.

[3] Levelset segmentation. https://computation.llnl.gov/projects/sapphire/levelsets/levelsets.html.

[4] Level set tutorial. http://profs.etsmtl.ca/hlombaert/levelset/.

[5] Levelset segmentation. https://en.wikipedia.org/wiki/Fuzzy_clustering.

[6] Dice coefficient. https://aapm.onlinelibrary.wiley.com/doi/pdf/10.1120/jacmp.v17i2.5820.

[7] Precision. http://sphweb.bumc.bu.edu/otlt/MPH-Modules/EP/EP713_Screening/EP713_Screening5.html.