

Semi-supervised learning using Total variation for biomedical image segmentation

Master Thesis

Prateek Purwar

Department of Electrical Engineering and Information Technology

Advisor: Dr. Gregory Paul
Supervisor: Prof. Dr. Orcun Goksel

Abstract

Acknowledgements

I would like to acknowledge the support of my Semester Thesis advisor Dr. Gregory Paul, Post-Doc. at Computer Vision Laboratory at ETH. He consistently steered me in the right direction whenever he thought I needed it. I would sincerely thank Denis Samuylov, who helped with implementation and coding at each step of my thesis.

Contents

1	Introduction	1
1.1	Focus of this Work	1
1.1.1	Dataset	1
1.2	Thesis Organization	2
2	Related Work	3
3	Method for image segmentation	4
3.1	Bayesian Formulation	4
3.2	Data energy term or Maximum likelihood cost function	4
3.3	Prediction models	5
3.3.1	Random Forests	5
3.3.2	Convolutional Neural Networks	5
3.4	Prior energy term: Variational Image Processing	5
4	Implementation	6
4.1	Solving variational problem using ASB	6
4.2	Random Forest	6
4.3	Architecture of CNN used	6
5	Experiments and Results	7
6	Conclusion	8

List of Figures

Chapter 1

Introduction

1.1 Focus of this Work

The task of image segmentation into binary classes is very useful in different cases in biomedical tasks. It can be used for detection of diseases, shape analysis etc. The methods to solve the segmentation problem has evolved among two lines: 1) level of interaction: from semi-interactive to fully automatic, and 2) level of classification: pixels to complete images. Nowadays, with the use of fully-convolutional networks, the segmentation can be obtained for complete image. This helps in using the local as well as contextual information for segmentation and significantly, improves performance in terms of speed. This speeds up the task of segmentation especially for huge images in cases such as electron microscopic images, 3-D images etc. Currently, the benchmark performance in terms of accuracy is achieved by the use of deep-neural networks. Deep neural networks such as U-net are specially designed for task of segmentation.

The only limitation with use of convolutional neural networks is need of huge training data for training these networks. Earlier networks solved the segmentation problem by training a neural network for classifying each pixel into corresponding classes. These networks were trained using a patch around training pixels. These networks faced problem of speed for cases of huge images where we have to run neural network for total number of pixels. Also, we get limited by the local information provided by the patch surrounding training pixels. This problem gets solved by use of fully-convolutional networks which produces segmentation for the whole input image at once. But we require fully annotated ground truth images for training such networks i.e. label for each pixel in the input image.

1.1.1 Dataset

The dataset or the images which we are trying to segment are electron microscopic images. The microscopic images are huge in size i.e. order of $450 \times 512 \times 512$. This poses a huge problem of creating a fully annotated ground truth. In addition to huge size of images, the images may contain multiple instances of objects to be annotated. For example, the liver tissue contain many vesicles. This makes the task of creating ground truth dataset hard and time consuming. The argument that the annotated dataset can be used again as for object segmentation in natural images also fails here as the objects to be segmented in microscopic images differ completely. This inhibits the experts to invest time to create fully annotated dataset. Further, the manual annotations for a single microscopic image may differ for each expert or even different for a single expert. This arise due to arbitrary shapes and sizes of objects to be segmented. This influences us to design a approach with semi-interactivity possible so that the experts can modify/add information according to

their needs. In summary, we have to consider following scenarios before developing a method to solve our segmentation task,

- High variability between images: The images to be segmented may be entirely different for example of liver and neuron tissue. It is not same as case for natural images.
- High variability between objects to segment: The objects to be segmented may differ completely from being smooth (round vesicles) to branched (neurons)
- High variability between goals:

Therefore, we plan to use an approach that can be modified easily to new images/objects to be segmented. This influences us to use hybrid of data and model driven approach. We decided to use an semi-supervised approach using Maximum-aposteriori inference. The idea is to propagate segmentation label from labeled to unlabeled pixels using an prediction model and a prior, as shown in figure 1.1. The prediction model and prior are combined using Variational image processing methods. As shown in figure 1.1, this can be used to generate segmentation or can be used to train other prediction model if needed. We can observe that the limitation on availability of fully annotated ground truth images inspires us to observe the trade-off between partially labeled data and accuracy of segmented images. This will enable the experts to use their annotation budget i.e. the pixels labeled by experts for training, efficiently giving best results. In addition, we observe effect of different ways of introducing maximum-likelihood and different regularisation parameter for total variation.

1.2 Thesis Organization

Chapter 2

Related Work

Chapter 3

Method for image segmentation

3.1 Bayesian Formulation

We model our image segmentation problem as a Bayesian inference problem. Let us consider an observed image, I and labeled or segmented ground truth, S , the joint probability can be defined as:

$$p(I, S) = p(S)p(I|S) \quad ,$$

and applying Bayes theorem,

$$\begin{aligned} p(S|I) &= \frac{p(S)p(I|S)}{p(I)} \\ &\propto p(S)p(I|S) \end{aligned}$$

The left hand side is the probability of S given the image I , is called the posterior probability. $p(S)$ is the prior probability of S . The Maximum a posteriori (MAP) estimate, S^* can be calculated as follow:

$$S^* = \arg \max_S (p(S)p(I|S)) \quad . \quad (3.1)$$

The above problem can as well be stated as an energy minimization problem by writing Equation 3.1 in terms of energy:

$$\begin{aligned} E(S) &= -\log(p(I, S)) \\ &= -\log(p(I|S)) - \log(p(S)) \\ &= E_d(I, S) + E_r(S) \end{aligned}$$

The total energy, E , that we want to minimize can be considered as linear combination of data term, E_d and prior term (or regularization), E_r . This modifies calculating MAP estimate to:

$$S^* = \arg \min_S (E_d(I, S) + E_r(S)) \quad .$$

3.2 Data energy term or Maximum likelihood cost function

Explain different cost functions

3.3 Prediction models

3.3.1 Random Forests

Explain random forests

3.3.2 Convolutional Neural Networks

Neural networks

Explain transfer learning also.

3.4 Prior energy term using Total Variation

Explain total variation regularisation term.

Chapter 4

Implementation

4.1 Solving variational problem using ASB

Solution of ASB problem

4.2 Random Forest

Feature selection and training of Random forests.

4.3 Architecture of CNN used

Training of CNN

Loss function for scribbles

Chapter 5

Experiments and Results

Chapter 6

Conclusion

Bibliography