

CS 766 Project Proposal

Chen Wang, Xiahe Liu

February 15, 2017

1 Introduction and Motivation

Semantic segmentation is a hot topic recently in the field of computer vision. The main task of semantic segmentation [Tho16] is to cluster parts of the images together which belong to the same object class. Semantic segmentation has a broad range of applications, for example, the autonomous driving car system will use this technology to distinguish different road signs [MBLAGJ⁺07]; Medical systems will use this technology to detect tumors from brains [MBVLG02], etc. And these interesting apps motivate us to investigate this topic. Semantic segmentation is definitely not a new topic and people have studied it decades ago, developing various of different methods revolving around this topic. Traditional segmentation methods includes unsupervised clustering algorithm [CLP98], random decision forests [SJC08], SVMs [YHRF12], Markov Random Fields [ZBS01] and Conditional Random Fields [RKT⁺09]. Recently, more and more researchers use approaches with convolutional neural networks to solve semantic segmentation problems and have got quite big advance in accuracy. [KSH12] firstly proposes a way to use CNNs to do image classifications on imageNet and develops a very famous neural networks called AlexNet. Since that, multiple advances has been made to build a better CNNs to do image classification or semantic segmentation tasks. Some noticeable work includes [LSD15], which applies the complete network as an image filter and [LSvdHR16] which combines CRF and CNNs together to make a better result.

Since there are so many interesting use-cases and approaches related to semantic segmentation, we would like to explore this topic more. Our goal is to re-implement the state of the art approaches in this field and try to make a thorough comparison between different approaches. Finally, we would like to propose some interesting applications based on these approaches and see if we could make any improvement to current state of the arts.

2 Milestones

2.1 Learn CNNs and build up small demos

Since we are pretty new to this area and we don't have any previous experience with CNNs, we plan to learn more about this advanced technology in the first period of our project. At the same time, we would like to build up some small demos with popular deep learning frameworks like Tensorflow, Torch, etc. to do some pretty simple image classification and segmentation tasks. By doing this, we will familiarize ourselves with the basic background knowledge.

2.2 Build CNNs and train

The second milestone is to actually build up our CNNs based on the method proposed in [LSD15] and [LSvdHR16]. For the data set, we would like to use PASCAL VOC 2012, which is a very famous and popular dataset which is broadly used in the area of semantic segmentation. We will report our accuracy in this milestone, comparing to the accuracy reported in the paper and trying to figure out the reason if there is a difference.

2.3 Compare with traditional approaches and more

In this milestone, we would like to compare the state-of-the-arts CNNs approach with some traditional approaches in the area of semantic segmentation. During the process, we would very likely



Figure 1: Example of semantic segmentation.

to think about how to combine traditional methods with CNNs, just like [LSvdHR16], which combines the traditional CRF methods with CNNs, to make our result better. We are also thinking about trying to make use of the technology we have learned and explored to make some interesting applications on mobile or web.

References

- [CLP98] Chang Wen Chen, Jiebo Luo, and Kevin J Parker. Image segmentation via adaptive k-mean clustering and knowledge-based morphological operations with biomedical applications. *IEEE transactions on image processing*, 7(12):1673–1683, 1998.
- [KSH12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [LSD15] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015.
- [LSvdHR16] Guosheng Lin, Chunhua Shen, Anton van den Hengel, and Ian Reid. Efficient piecewise training of deep structured models for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3194–3203, 2016.
- [MBLAGJ⁺07] Saturnino Maldonado-Bascon, Sergio Lafuente-Arroyo, Pedro Gil-Jimenez, Hilario Gomez-Moreno, and Francisco López-Ferreras. Road-sign detection and recognition based on support vector machines. *IEEE transactions on intelligent transportation systems*, 8(2):264–278, 2007.
- [MBVLG02] Nathan Moon, Elizabeth Bullitt, Koen Van Leemput, and Guido Gerig. Automatic brain and tumor segmentation. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 372–379. Springer, 2002.
- [RKT⁺09] Chris Russell, Pushmeet Kohli, Philip HS Torr, et al. Associative hierarchical crfs for object class image segmentation. In *Computer Vision, 2009 IEEE 12th International Conference on*, pages 739–746. IEEE, 2009.
- [SJC08] Jamie Shotton, Matthew Johnson, and Roberto Cipolla. Semantic texton forests for image categorization and segmentation. In *Computer vision and pattern recognition, 2008. CVPR 2008. IEEE Conference on*, pages 1–8. IEEE, 2008.
- [Tho16] Martin Thoma. A survey of semantic segmentation. *arXiv preprint arXiv:1602.06541*, 2016.
- [YHRF12] Yi Yang, Sam Hallman, Deva Ramanan, and Charles C Fowlkes. Layered object models for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(9):1731–1743, 2012.

- [ZBS01] Yongyue Zhang, Michael Brady, and Stephen Smith. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging*, 20(1):45–57, 2001.