CS-7643-O01, OAN, OSZ Final Project Proposal (group submission)

Paresh Upadhyay, Murugan Viswanathan Viswathana

TOTAL POINTS

2 / 2

QUESTION 1

1 Project Proposal Feedback **2 / 2**

  ✓ **- 0 pts** *Correct*

    💬 Proposal looks good! Scope of project is
reasonable for a team of 2.

ılı gradescope

Team: Team Vision

Project Title: Image Segmentation using Deep Learning Network

Project Summary:

Image segmentation is a pivotal and one of the most extensively studied area in image processing and computer vision. Its applications span across diverse fields including enhancing robotic perception, video surveillance and facilitating augmented realty experiences. One of the most promising applications of image segmentation is in the medical imaging domain where it plays a critical role in the precise delineation of anatomical tissues, enabling in-depth quantitative analysis.

In this project, we aim to explore the realm of image segmentation using the latest and popular techniques. Our goal is to understand, implement and experiment with cutting-edge image segmentation methodologies. With this project we hope to hone our practical skills and deepen out insights into the world of computer vision and deep learning.

Approach:

Our approach aims to build a Deep Learning Network for image segmentation task using one of the well-established image segmentation techniques, listed below.

(a) Encoder - Decoder based network: For example, we will explore architectures like SegNet, which would enhance our understanding of the encoder-decoder paradigm in image segmentation

(b) RNN based architecture like ReSeg: We will investigate the use of Recurrent Neural Networks for image segmentation.

(c) Dilated Convolutional Models and DeepLab Family of networks: These models introduce the concept of dilated convolutions for accurate and efficient segmentation. We plan to delve into these models to understand their mechanism.

(d) Attention based models: For example, we will explore models that employ attention mechanisms, to study how attention enhances segmentation performance.

In our project, we aim to implement both (a) the encoder - decoder implementation approach and (b) the RNN based implementation for image segmentation. This dual approach will provide a comprehensive view of different techniques. Our goal is to explore the other mentioned architectures as well. Our implementation will include model training, optimization, and evaluation. We will use the datasets mentioned in the section below to conduct the analysis and evaluate the performance.

As a stretch goal, we aspire to explore a technique called Panoptic Segmentation, a more advanced task that unifies semantic and instance segmentation. This exploration will provide a deeper understanding of segmentation challenges and potential solutions.

Resources / Related work:

Image segmentation is a well-established and studied area in computer vision. There are several state-of-the-art techniques and research works that exist. For out project, we investigated various methods as indicated in the previous section and have referred to valuable resources that will guide us in our project. The following references are some of the key research works that we would be referring to for our project:

[1] *"Image Segmentation Using Deep Learning: A Survey", Minaee et al.*

    - This provides a comprehensive overview of various deep learning approaches for image segmentation.

[2] *"SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation", Badrinarayayan et al.*

    - The SegNet architecture is a fundamental encoder-decoder model for image segmentation, which we may explore.

[3] *"Full Convolutional Networks for Semantic Segmentation", Long et al.*

    - This work presents some fundamental key concepts in fully convolutional networks for image segmentation, providing knowledge for our work.

[4] *"W-Net: A Deep Model for Fully Unsupervised Image Segmentation", Xia et al*

[5] *"Attention is all you need", Vaswani et al.*

[6] *"ReSeg: A recurrent neural network based model for semantic segmentation", Visin et al*

[7] *"Attention to scale: Scale-aware semantic image segmentation", Chel et al*

[8] *"Panoptic Segmentation", Kirillov et al.*

    - Kirillov's work introduces Panoptic Segmentation, which is the challenging task that we aim to explore as a stretch goal.

[9] *"Computer Vision", Szeliski.*

    - Szeliski's book provides a foundational reference for computer vision.

Datasets:

The following datasets are available for our project:

1. PASCAL Visual Object Classes (VOC) (http://host.robots.ox.ac.uk/pascal/VOC/)
2. PASCAL Context (https://cs.stanford.edu/~roozbeh/pascal-context/)
3. Berkeley Segmentation Dataset (BSC) (https://www2.eecs.berkeley.edu/Research/Projects/CS/vision/bsds/)
4. COCO (https://cocodataset.org/#home)

Team Members (listed alphabetically):

1. Murugan Viswanathan
2. Paresh Upadhyay

1 Project Proposal Feedback **2 / 2**

✓ **- 0 pts** *Correct*

💬 Proposal looks good! Scope of project is reasonable for a team of 2.

gradescope