

The main purpose of this project is to explore CNNs and learn how they can be implemented to perform semantic image segmentation. It is essential that at the end of this project I had a clear understanding of the image segmentation process and the importance of it in the image processing industry. This project involved a literature review of current networks, along with the implementation of one of these networks. After thorough research, I decided to choose SegNet as the CNN to implement. Papers showed that SegNet was effective at segmenting road scenes therefore chosen dataset for the project was CamSeq01 from CamVid. This dataset consisted of road scenes in Cambridge. I originally sought to train SegNet using Pytorch, however, it was found that SegNet is built on top of VGG16 Net with over 13 convolutional layers. This complex network required a lot of time to process each layer without a GPU accelerator and needed more RAM, therefore, it was deemed infeasible to train SegNet ourselves. So, I decided to use a pre-trained network. The testing images chosen were a mixture of Glasgow road scenes and Cambridge road scenes so that a comparison could be made between the resulting segmented images for both cities. Cambridge as a city differs in style from Glasgow, so I wanted to see whether training with Cambridge road scenes only would result in more errors in the resulting Glasgow images.

I decided to implement two methods that can perform image segmentation and designed a possible process to perform both. Followed by combining the code with hardware, which comprised of my personal laptop. The first implementation will use SegNet, a deep convolutional encoder-decoder architecture which can be used for pixel-wise labelling, this will be implemented using MATLAB. Secondly I implemented a program using Pytorch, a python package that provides a deep neural network built on a tape-based algorithm..

The inputs for the process come in the form of images taken from either the Berkeley data set (Pytorch implementation) or individual images of Glasgow taken from Google images (SegNet method). Both implementations were tested separately, and the results could then be analysed which reveal any errors present.

