CS584 – Machine Learning Project

Project title

Aerial Drone 2D Image Semantic Segmentation

Team members

1) Prashant Jagwani (A20484911)
2) Vibhanshu Jain (20493731)

Description of the problem.

Semantic segmentation refers to the process of linking each pixel in an image to a class label. These labels could include a person, car, flower, piece of furniture, etc., just to mention a few.
We can think of semantic segmentation as image classification at a pixel level. For example, in an image that has many cars, segmentation will label all the objects as car objects. However, a separate class of models known as instance segmentation is able to label the separate instances where an object appears in an image. This kind of segmentation can be very useful in applications that are used to count the number of objects, such as counting the amount of foot traffic in a mall.
Implementing a Encoder-Decoder architecture (U-Net) for a Semantic segmentation of Aerial Drone images for increasing the safety of autonomous drone flight and landing procedures.

A brief survey of what has been done and how the proposed work is different.

The paper ‘A Survey on Deep Learning-based Architectures for Semantic Segmentation on 2D images’ surveys the evolution of semantic segmentation architectures and presents a chronological order of techniques in the context of new challenges. The field of semantic segmentation has seen a rapid increase in interest and a number of survey studies have been published, but many lack an overarching vision or necessary depth of analysis regarding deep learning-based methods. Recent reviews have provided comprehensive surveys on the subject, but their categorization is often coarse and they do not discuss the chronological evolution of techniques. They provide a table of related

methods, explain them briefly in chronological order, and provide their metric performance and computational efficiency. The authors believe that this survey will better help readers understand the evolution, current state-of-the-art, and future directions of 2D semantic segmentation.

Preliminary plan (milestones)

The frameworks we are going to use for the projects are :-

1. OpenCV – Image Processing
2. Pytorch – Deep Learning Library
3. Albumentations – Image Augmentations

First we will use U-Net which is an Encoder/decoder structure with skip connections that connect the same levels of ED and final input-sized classification layer. This leads to Efficient computation load due to no fully connected layers or a refinement block.

We will make use of different architectures for comparison:

1) Mobilenetv2
2) Resnet
3) Effecientnet
4) Resnext
5) Xceptionet

Reference (Papers and Dataset)

1. Olaf Ronneberger, Philipp Fischer, Thomas Brox: U-Net: Convolutional Networks for Biomedical Image Segmentation (2015)

2. Irem Ulku, Erdem Akagunduz: A Survey on Deep Learning-based Architectures for Semantic Segmentation on 2D images (2019)

3. Dataset Name: Semantic Drone Dataset
   Dataset Link: http://dronedataset.icg.tugraz.at