Capstone Project Proposal

Sanghoon Yoo
1/30/2020

1. What is the problem you want to solve? why is it an interesting problem?

- I would like to solve a semantic segmentation problem which can classify every pixel in an image as a certain class. Basically, it will look like below:

- I think image classification based on this type of semantic segmentation could be useful for various area such as surveillance system, self-driving cars, or any kinds of robots which needs computer vision.

2. What data are you going to use to solve this problem? How will you acquire this data?

- On COCO website(cocodataset.org/#download), there are 2017 Train, Validation and Test images available. These images were used in 2019 COCO competition as well, so I will utilize these image dataset for this semantic segmentation problem.
- COCO provides API to load their dataset including images and annotations files, so I can use its API to acquire this data.

3. In brief, outline your approach to solve this problem

- Semantic segmentation is a supervised classification machine learning problem. Furthermore, each image might have more than one class(or object). For instance, it can have a dog and person in the same image as shown above. Therefore, a model needs to predict probability of all classes per each pixel in order to perform semantic segmentation. For predictors, pixel values will be utilized, but I am not sure if it would be enough to use only grayscale, or I need to use RGB values or other pixel values at this point. To solve this problem, a traditional machine learning approach will not work. I will use deep learning methods.

4. What will be your final deliverable?

- I would like to deploy my model as a web service with an API. However, if I feel everything looks fine and can spend more time, then I may want to build web application too.

5. What computational resources would you need at a minimum to do this project?

- COCO dataset is huge. The size of training images is 18GB. I am not sure if I will use whole dataset at this moment, but I would like to build my model with large dataset since machine learning at scale is one of the skills that employers want to see. My project is all about image processing, so GPU is definitely necessary. My laptop has built-in NVIDIA GPU, but I have no idea this GPU can handle this big dataset, and training.
- There are a bunch of options for cloud resources such as Paperspace(Springboard prefer), Google Cloud, Microsoft Azure, and Amazon Web Services. I think I will go with Google Cloud because it is the most popular platform, and there are $ 300 credits available which is larger compared to AWS service.
- I will follow below procedure to have a concrete idea of memory usage for my project: I will prototype and test my model on my local machine first, and use free cloud tools from Google Cloud(I think Colab is free platform). After that if I need more resources then I will utilize Datalab platform from Google Cloud.