Interview Questions

Deep learning on traffic sign detection using public data.

Introduction

The goal of this test is to let you demonstrate your knowledge in image processing and more specifically on object detection using neural networks. The time is short, and we don't expect you to come with a perfectly trained model and state of the art results (we don't want to select people based on the GPU they have...), the goal is to see how you setup everything, how familiar you are with those tools and obviously how you criticize your results.

Since we have to train you to fit in our team quickly and understand the tools we have. The initial instruction will be provided using python and yolo to train first model.

The subject we choose here is the detection of common seen images, those signs are used for traffic restrictions (like speed limit, height limit...) mostly everywhere around the world (Europe, China, Russia, India, Brazil, Mexico... but not in the US). More details will be provided later.

You will provide us with a written report (2-3 pages is fine) with comments on your choices and results for each part of the question. Any addition on thing you wanted to try but you didn't have the time to get them completed is also welcome. More detailed instruction on the reports as well as the deadlines for each part are given at the end of this document.

Part1: Data Preparation

You will need to get familiar with computer vision and deep learning application. Please take a look at the YOLOv3 model structure first and understand the critical point of its philosophy.
The dataset we want you to use is GTSDB. Your main goal here is just to make it ready for training. Please download GTSDB dataset from http://benchmark.ini.rub.de/?section=gtsdb&subsection=news (Links to an external site.) (get the full version).

Then, please take a look of the yolo github : https://github.com/anliec/darknet. You will find how you have to format the annotation for training. This part is to train your ability to preprocessing data for structured deep learning. The more detailed training process is in the original github(https://github.com/AlexeyAB/darknet).

Please provide detailed steps of your data preparation and specify your training parameter (number of classes, the id of the class...etc).

Part 2: Training and Testing

Please build/train a traffic-sign detection system by Yolov3 using GTSDB. You can use any NN you want as long as explain why you chose it. you can also try multiple of them and compare the results, but do not go for more than 2 or 3, training one is already time consuming.

If you do not have a GPU for training, Google Colab could be an option. Here is an example on how to train YOLO on Google Colab (by "The AI Guy"). You should not treat this as a step by step guide for training that you need to follow, this is included simply as a proof-of-feasibility of using cloud computing for training, how you perform your training is part of the evaluation and is up to you.

https://colab.research.google.com/drive/1Mh2HP_Mfxoao6qNFbhfV3u28tG8jAVGk#scrollTo=sx01fpFEnfNe

After you have finished this part, make sure you can answer the following questions (it would be really cool if you can provide more perspectives):

1. How do you tweak your hyperparameters? What are they all about?

2. What's your training strategy? How can you tell if it works?

3. What's the performance of your detection system? How do you evaluate it?

4. Do you see any **pattern in your FP/FN detection**? Can you **categorize** the errors and propose potential solutions?

5. Did you need to go back to Part1 to improve it?

And again: We understand that it's difficult to train a perfect NN in such a short while, especially when you can't afford to use a really luxury GPUs. So please don't worry, you're not going to be judged by the performance of your result. All we want to see is: you can make it happen, and you know what you're doing.

Now, since you have familiarized with the training process, we would like you to think beyond our existed method and answer the following questions.

6. How does your results compare to the state of the art in sign or object detection? (You are not supposed to be as good as state of the art) What are the main differences with your approach?

7. What is the main challenge you encountered in this project, if there is any, please tell us you thought of how you will overcome it?

Deliverables:

After completing this question set you will send us (james.tsai@ce.gatech.edu , pyu68@gatech.edu and zyy@gatech.edu) a report of around 5 or 6 pages in pdf format with your analysis and comments. You will also send in your code for parts 1 and 2 by giving us the link to your github / gitlab repository. Please include in the repository the dataset an easy instruction on how to set it up (like a bash script) as well as the trained weight of your neural networks (if they are too large you can use GT's Dropbox or any file sharing service you want).

We highly recommend you follow the above instructions carefully.

Deadline:

Sunday Aug 8, 2021 11:59 PM