README

The Carvana Image dataset can be found on Kaggle's Carvana Image Masking Challenge website –
https://www.kaggle.com/c/carvana-image-masking-challenge/data
The external dataset has been shared using Google Drive –
https://drive.google.com/drive/folders/0B0Te0p-dLjuqc1VRZEp0d3k1Q3c?usp=sharing
Trained models have also been shared using Google Drive –
https://drive.google.com/drive/folders/1j6G4kKzwqpQDwgD312os0PS0P1zFBhen?usp=sharing
The code has been developed to run on the Google Cloud Platform. Instructions on setting up GCP account are available here –
https://cloud.google.com/deployment-manager/docs/how-to
GCP command line tools are required –
https://cloud.google.com/sdk/

To run the code on a local machine, the code requires –

- Python 2.7
- Tensorflow (with Python2.7 wrapper)
- OpenCV (with Python2.7 wrapper)
- Other libraries that can be installed using the setup.py script provided with the code

There are 2 sets of code – one for training the models and one for testing. Both training and testing code have been designed to run on the Google Cloud Platform (gcp).
The training code is packaged in the following manner –

```
--trainer
        --model
                --losses.py
                --u_net.py
        --params.py
        --train_cv2.py
        --cloudml-gpu.yaml
--start.sh
--setup.py
```

To run the code on gcp, run the "start.sh" shell script by typing the command –
sh start.sh

Values for arguments in "start.sh" would need to be changed before running it. Inside start.sh, the argument "job-dir" must contain the GCS (Google Cloud Storage) URL which will store logs and the model's weight file. The argument "train-dir" should contain the GCS URL where all the training images and the training mask csv file are stored.
To run the code locally, the script "local.sh" can be executed after giving valid directory paths for "job-dir" and "train-dir" arguments.

The packaging of the testing code is in the following manner –

```
--tester
        --model
                --losses.py
                --u_net.py
        --params.py
        --test_submit_multithreaded.py
        --cloudml-gpu.yaml
--start.sh
--setup.py
```

Again, just type the command “sh start.sh” to run start.sh. To configure the script, 3 directory URLs need to be specified –

-- job-dir – directory containing test data and a csv with image names (like first column of sample_submission.csv provided by Kaggle Carvana Image Challenge)

-- weights-dir – directory containing model weight file

-- output-dir – directory where generated output will be stored

To run locally, use “sh local.sh” after giving valid directory paths for the above 3 arguments.