Visual Computing

H.N Srikanth (sm21mtech12012)

CODE IMPLEMENTATION

Read Me

Swin Transformer V2 is implemented with the support of the GitHub repo <https://github.com/SwinTransformer/Swin-Transformer-Object-Detection>

Swin-T variant is implemented because other variants require multiple GPUs.

Primarily the environment setup is made on the Nvidia Tesla V100 GPU server as per directions given in the requirements.txt file.

The above mentioned github repo is cloned and further process is done as explained below.

Swin-T (mask\_rcnn\_swin\_tiny) model is downloaded from the repo in.pth format and uploaded to the server.

Then COCO dataset is downloaded and kept in the data directory as recommended in the repo.

The initial inference is made using swin.py from config files on this dataset.

Later on, the model is trained for ten epochs, and weights are loaded into the checkpoints folder, which is used in further inference.

Working on the implementation of novelty areas that were discussed in 2nd presentation (i.e, Dropouts).

I am implementing this trained Swin Transformer V2 with the POTHOLE dataset.