BOĞAZİÇİ UNIVERSITY

CMPE 492

REPORT

WEEK 3

BARAN DENIZ KORKMAZ

DOĞUKAN KALKAN

FALL 2020

1 Weekly Summary

1.1 Work Done

1.1.1 FaceForensics++

1. The FaceForensics++ dataset has been carefully analyzed. The dataset consists of manipulated video sequences using manipulation techniques of Deepfakes, Face2Face, FaceShifter, FaceSwap, and NeuralTextures. The dataset also contains the **DeepFakeDetection** data.
2. At the initial stage, the images compressed in c40 technique (LQ - $\sim 5GB$) are downloaded from both **FaceForensics and DeepFakeDetection**.
3. The png formatted video frames are extracted from videos using the script written by the collaborators.

1.1.2 GPU

1. **Google Cloud:**
   - We made two more quota requests.
   - As we explained, Google Cloud is not very good at helping its customers through e mail. So we decided to talk to them in live chat. We talked to Sales Team, which is the team in charge of granting us quota.
2. **AWS:**
   - Another cloud computing option that we can use during our project is AWS. Their AWS Educate platform allows students to create different types of instances. It is a bit more complicated than Google Cloud, and setting up the environment is difficult as well.
   - Even though it is difficult to set up, we wanted to try to use AWS Educate and tried to create some instances.

1.2 Learning Outcomes

1.2.1 FaceForensics++

1. In conclusion to the work done related to the dataset, the collaborators are ready to test the general pipeline using the dataset of images extracted from the low-quality videos (compressed in c40).
2. Our aim is to begin with the smaller subset of large FaceForensics++ database which accounts for approximately 2GB of raw extracted images. With a smaller subset of 5GB images from **FaceForensics++ and DeepFakeDetection**, we can much quickly test the general pipeline we have constructed previous week in transfer learning implementation using **VGG19**.

2 Challenges

2.1 FaceForensics++

- At initial stage, starting the training with the entire dataset can be hard to tackle if any error occurs. Therefore we decided to work with a smaller subset of low quality compressed images to verify the integrity of general pipeline.

2.2 GPU

1. We completed the transfer learning implementation last week and we also obtained the dataset with which we are going to train our model.
2. However, we are told by Google Cloud Team and Sales Team that they cannot grant students GPU quota.
3. So, we shifted our attention to AWS Educate. However, AWS Educate does not allow us to create GPU instances, either.

3 What's Next?: Upcoming Week

- **We have an urgent need for a GPU instance to run our current implementations.**