CS696E Progress Update 3

CK Wijesundara

Where I’m At Now [1/2]

- Brand new face detector!
  - Pre-trained *Dual Shot Face Detection (CVPR 2019)* model
  - Better average precision than previous MTCNN model
    - 0.900 vs. 0.607 on WIDER Face (Hard) dataset
  - Much faster inference time (with the implementation I’m using)
    - MTCNN took 0.969 seconds to detect faces in Lord of the Rings image, compared to 0.119 seconds with DSFD
  - PyTorch optimized model for inference from: https://github.com/hukkelas/DSFD-Pytorch-Inference/

Where I’m At Now [2/2]

- Implemented cropping out faces based on bounding boxes from face detector
- Implemented a pre-trained face encoding model (FaceNet) that converts faces into 512-dimensional encodings
  - Using the FaceNet PyTorch library from here: https://github.com/timesler/facenet-pytorch
- Implemented clustering on those 512-dimensional encodings for grouping similar looking faces
  - Basic experimentation with KNN clustering
  - Full disclosure, the results right now aren’t great
- Age/gender estimation is on hold while I’m working on this

What’s Next?

- Figure out how/if it’s possible to improve the face clustering
  - Try to see if I can get density-based clustering working (where I don’t need to specify a specific number of clusters in advance, the algorithm figures it out)
- Eventually get back to age/gender recognition