Indian Institute of Technology Bombay; September 29, 2023
EE782 Advanced Topics in Machine Learning

Assignment 2: Metric Learning and Generative AI

Instructions:
- Assignment has to be done individually, in Google CoLab using python and pytorch
- Write copious comments using text cells and in the code cells to explain your intent and demonstrate understanding
- End with a cell that lists sources of starter code (e.g., github repo links, or chatGPT) for all modules
- Submit the (1) ipynb file, (2) its py version, (3) link to a 10 minute video explaining the code, your intent, and your observations and decisions on Moodle
- CODE WITHOUT COMMENTS FOR EACH LINE OF CODE WILL NOT BE GRADED
- VIDEO WITHOUT VIEW PERMISSIONS WILL NOT BE GRADED. Test your video by sending link to a friend
- Deadline is Oct 15, 2023, 11:59pm

Objectives:

1. Download labeled faces in the wild (LFW) dataset: http://vis-www.cs.umass.edu/lfw/
2. Get the number of persons who have more than one image

Part-A

3. Split into training, validation, and testing by person (not by image)
4. Start with a network that is pre-trained on ImageNet and appropriate for your computational resources
5. Appropriately crop and resize the images based on your computational resources
6. Setup and train a Siamese network to predict whether a pair of images are of the same person or not using a metric learning scheme (cosine similarity or Euclidean distance; paired with cross-entropy or hinge loss with a margin), and validate on the validation split [3 for basic set up]
   a) Use image augmentation [1 for including this]
   b) Experiment with a regularization technique and its hyper-parameter [1 for including this]
7. Experiment with at least two learning rate schedulers and comment on what works, doesn't work, and potential reason why [1]
8. Experiment with at least two different optimizers and comment on what works, doesn't work, and potential reason why [1]
9. Test on the test split [1]
10. Gather a few images of yourself and your friends (with their consent) and check if the model works well on this data [1]

Part-B

11. Train a generative model for generating face images, using a GAN. The generator takes a Gaussian noise vector as input, and tries to output a face image, while the discriminator distinguishes between real and fake face images. See, for example: https://www.youtube.com/watch?v=_pIMdDWK5sc [2]
12. (Bonus) Modify the GAN to become a conditional GAN, where the condition itself is (features of) a real face image. The CGAN should generate another image of the same person, and you can use the Siamese network from Part-A as an additional discriminator for person matching. You will need to ensure that it does not simply show the same image as the one used for conditional input. See, for instance, https://www.youtube.com/watch?v=Hp-jWm2SzR8 [2]