Assignment 3 ADRL - 2022

Instructor: Prathosh A.P., IISc

November 14, 2022

1. **Self-Supervised Learning:**
   (a) Consider the CIFAR-10 dataset (https://www.cs.toronto.edu/~kriz/cifar.html). Construct a CNN-based classifier (call it the base-model) and report the accuracy. Now change the definition of the classes in two different ways (one binary classification problem and one 5-class classification problem). Use the pretrained base-model with an additional classification head at the last layer in accordance to the new class definitions. Retrain only the final classification heads and report the accuracies. Plot the t-SNE graphs for all three cases and record your observations.
   (b) Implement MOCO for CIFAR 10. Consider 5 different types of augmentations (rotations, blur, color distortion, cropping and resizing) for defining the positive samples. Use the entire training data to learn the representations. Once the representations are learned, use a linear and logistic layers and retrain with 10-50% of supervised training data and compare the results with the base CNN model in question (a). Experiment with two different sizes for the encoder dictionary.
2. **Few Shot Learning (FSL) :** All the experiments in this section will be on the Omniglot dataset (https://www.kaggle.com/datasets/watesoyan/omniglot). It contains 1623 handwritten characters collected from 50 alphabets of different languages. There are 20 examples associated with each character, where each example is drawn by a different human subject. You should first resize the grayscale images to 28×28 and augment the character classes with rotations in multiples of 90 degrees (4 angles). Use the first 1200 characters plus rotations for training (4,800 classes in total) and the remaining classes, including rotations, for test. You have to consider two scenarios - 1-shot and 5-shot classifications, both with 5-way (or classes). The train test protocol consists of several train-test episodes: every episode is of 5-way, 5 shot, that is, a single training episode is made of 5 examples of 5 randomly sampled classes from the train dataset along with a test episode with 5 examples of 5 randomly sampled classes from the test dataset. The final accuracy should be reported as the average over 1000 randomly generated episodes from the test set. You should use the above setup for both the experiments below.

(a) Implement prototypical networks for the above problem of FSL. Consider a 5-layer CNN for the embedding layer with 32-dimensional embedding. Use Euclidean measure for the distances to the prototypes.

(b) Implement MAML on the same 5-layer CNN as above with a classification head at the end. Compare these results with the pro-typical networks in terms of computational cost, training time and accuracies.