CMP 719 - Computer Vision

HACETTEPE UNIVERSITY
Department of Computer Engineering

Programming Assignment
Spring 2024
Dr. Nazlı Ikizler Cinbis

Due Date: 23:59 pm on Sunday, June 2nd, 2024

Image Classification with CNNs and Attention

In this assignment, you will get familiar with image classification by training Convolutional Neural Networks (CNN). You can use PyTorch environment for coding.

Dataset

The dataset includes images of different foods and is available at the following url: https://www.kaggle.com/datasets/harishkumardatalab/food-image-classification-dataset. From this dataset, **you will work on a subset 8 food classes that you select**. For each class, you will also select 200 images at random, so your whole dataset size will be ≈ 1600 images. For the training, validation and test splits, divide the selected subset by $70\% - 10\% - 20\%$ ratios, respectively.

Figure 1: Sample images from the dataset.

PART 1 - Modeling and Training a CNN classifier from Scratch - [40 pts]

In this part, you are expected to model a CNN classifier and train it. You should first define the components of your model. For instance; a simple architecture might contain 3 convolutional layers and 2 fully connected layers. Define your initial network.

- How many layers are there on your initial classifier? What structure do these layers have? Give details of your parameters; number of in channels, out channels, stride, etc. Specify your architecture in detail. Write about your choice of activation function, loss function and optimization algorithm.
- Show your model's behaviour with respect to two different learning rates by plotting the change in the loss value wrt epochs.
- After selecting best hyperparameters, give training, validation, and test accuracy. Plot your model's loss and accuracy change. Comment on your findings.
- Add a regularization method (other than data augmentation) to your training and comment on your findings.

After optimizing the architecture, evaluate it on the test set. What is your model's test set accuracy? Which classes tend to confuse the most? Explain your findings and results.

PART 2 - Transfer Learning with CNNs - [30 pts]

Now, fine-tune the EfficientNet B0 or EfficientNetV2 B0 network pre-trained on ImageNet (Download the pretrained model from internet. The Pytorch version of EfficientNet is available here[1]). Freeze all the layers, except the FC layer and train FC layer for this food classification problem.

[1] https://pytorch.org/vision/main/models/generated/torchvision.models.efficientnet_b0.html

- Explore finetuning different number of layers. Tune your parameters accordingly and give accuracy on validation set. Compare your results.
- Evaluate your best model (the model that gives the best results in the validation set) in the test set. Discuss the results.

Compare the results of Part-1 and Part-2. State your observations clearly and precisely.

PART 3 - CNNs with Soft Attention - [30 pts]

Implement the soft-attention mechanism that is discussed in class and using your best model from Part-1 or Part-2 as your base CNN model, incorporate the attention mechanism to your base network.

- Experiment with using this attention layer in different parts of the network.
- Evaluate your best model in the test set. Compare your base CNN with the attention-added version.

Discuss the models from all the parts clearly and extensively. State all your observations.

What to Hand In

All the documentation of the code, together with the plots, explanations and related discussions should be submitted as a Python Notebook.

Name your file as $< Name\_Surname\_ID.ipynb >$ and submit by emailing to nazli@cs.hacettepe.edu.tr .

Academic Integrity

All work on assignments must be done individually unless stated otherwise. You are encouraged to discuss with your classmates about the given assignments, but these discussions should be carried out in an abstract way. That is, discussions related to a particular solution to a specific problem (either in actual code or in the pseudocode) will not be tolerated. In short, turning in someone else's work, in whole or in part, as your own will be considered as a violation of academic integrity. Please note that the former condition also holds for the material found on the web as everything on the web has been written by someone else.