Object recognition and computer vision 2023/2024

Gül Varol, Jean Ponce, Armand Joulin, Josef Sivic, Ivan Laptev, Cordelia Schmid and Mathieu Aubry

Assignment 3: Image classification

(Adapted from Rob Fergus and Soumith Chintala)

Goal

This assignment requires you to participate in a Kaggle competition with the rest of the class on the classifysketch dataset. The objective is to produce a model that gives the highest possible accuracy on a test dataset containing the same categories.

Guidelines

You should first clone the Github repository https://github.com/willowsierra/recvis23_a3. The *main.py* file contains code for training and evaluating your models. Once training is completed, it produces a file *kaggle.csv* that lists the IDs of the test set images, along with their predicted label. This file should be uploaded to the Kaggle webpage, which will then produce a test accuracy score. You can register for the competition using the private link: https://www.kaggle.com/c/mva-recvis-2023/.

What to hand-in

You should write a 1-page, double-column report in CVPR'24 format briefly presenting your approach and obtained results. This report should be uploaded in pdf format to the course Google Classroom by November 28th 2022.

The model architecture is specified in *model.py*. Currently a simple baseline model is provided. You can use this model or create your own. You are free to implement any approach covered in class, or in the research literature. Of course, tricks that you devise yourself are also encouraged. The test and training data are provided in the Kaggle competition data download page. You should unzip the images.

External dataset policy

1. Solutions should not use the original dataset classifysketch for training. Students should use the training set provided on the Kaggle competition.
2. Solutions should not use pre-trained models learned on datasets in 1.
3. Solutions can use any datasets or self-collected images without annotation.
4. Solutions can use ImageNet pre-trained checkpoints.