Free-hand sketch classification problem

1. Introduction

- Three students at most form a group.
- This project is to build models to do free-head sketch classification. You can choose one or more methods to finish this task, e.g., SVM, and various deep learning algorithms.
- Deadline for this project is June 19, 2022, 23:59. Each group needs to submit a report pdf on Canvas and the source code is also required by providing the link to your github repo.
- This project will be evaluated from workload (25%), model performance (25%), results analysis (30%), report writing (10%) and code (10%).

2. Assignment

Free-hand sketch drawing is a valuable way to convey messages and express emotions throughout human history. Sketches contain not only vivid categorical features of the target objects but also abstract, variant visual appearances. In this project, you are asked to construct and train (deep) models for 28×28 sketch image classification.

Tips: In order to raise the sketch classification accuracy, you may learn better sketch representations by feeding the model with not only the required 28×28 sketch images but also the paired sketch sequences as in [2–4], the related photos as in [5], etc.

3. Dataset

QuickDraw [1] is one of the largest free-hand sketch datasets. It includes 345 categories of common objects, and each one contains 70 thousand training, 2.5 thousand validation and 2.5 thousand test samples. The dataset is available at https://magenta.tensorflow.org/sketch_rnn. The original sketches in QuickDraw are described as vectorized sequences, which can be further translated into sketch images. For example, the "Dataset" section in https://github.com/CMACH508/RPCL-pix2seq offers an approach to create the pixel-formed sketch images.

In this project, we choose 25 categories (cow, panda, lion, tiger, raccoon, monkey, hedgehog, zebra, horse, owl, elephant, squirrel, sheep, dog, bear, kangaroo, whale, crocodile, rhinoceros, penguin, camel, flamingo, giraffe, pig, cat) from QuickDraw for the sketch classification problem. Each sketch individual is translated to a 28×28 sketch image as the model input.

4. Project Report

Each group is required to turn in a project report with your main ideas, utilized methods and algorithms, experimental settings, and finally experimental results. The project report (.pdf) can be written either in English (encouraged) or in Chinese. And the details are in the following:

- **Main Ideas** A brief introduction of your method(s).
- **Methods and algorithms** The utilized method and algorithm for the introduced classification model, including the motivations, the detailed description, etc.
- **Experimental settings** Training details, e.g., network structure, learning rate, data preprocessing method if necessary, such as augmentation, clustering, dimension reduction, etc.
- **Experimental results** The sketch classification accuracy over the entire test set is required (totally 2500×25 sketch samples from 25 categories). And you may demonstrate the model's performance from more perspectives, e.g., robustness, extracted features in latent space, etc.
- **Conclusion** Analysis of the experiment results, e.g., cons and pros of the model, comparison of different models, ablation study of your modification, etc.
- **Group member and contribution** At the end of the report, please attach the contribution of each member as a percentage. For example, A 30%, B 30%, C40%. And work done by each student is needed to be clarified. You are also required to submit the source code of your classification model by providing the link to your github repo in the report. If you do not know how to use github, please visit its tutorial (https://guides.github.com/activities/hello-world/) for some advice.

References

1. Ha D, Eck D. A neural representation of sketch drawings[J]. arXiv preprint arXiv:1704.03477, 2017.
2. Song J, Pang K, Song Y Z, et al. Learning to sketch with shortcut cycle consistency[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 801-810.
3. Xu P, Huang Y, Yuan T, et al. Sketchmate: Deep hashing for million-scale human sketch retrieval[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 8090-8098.
4. Li L, Zou C, Zheng Y, et al. Sketch-r2cnn: An rnn-rasterization-cnn architecture for vector sketch recognition[J]. IEEE transactions on visualization and computer graphics, 2020, 27(9): 3745-3754.
5. Lamb A, Ozair S, Verma V, et al. Sketchtransfer: A new dataset for exploring detail-invariance and the abstractions learned by deep networks[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2020: 963-972.