Final Project Proposal

Group 2

Jianing Wang, Li Du, Wen-Chuan Chang

Background and problem statement

In today's world, businesses can no longer turn a blind eye to the opportunity for expansion, by providing goods and services online instead of just in stores. To target specific markets and demographics, implementing the use of A.I. and big data are becoming increasing mainstream.

For example, if we can classify those pictures which have been uploaded by users to social media, then we could know the preference of users and classify the picture. On top of that, we can send different type of advertisements to different groups of users based on their preference, which would decrease the number of aimless advertisements.

This study aims to build a model to classify 100 different types of pictures. To solve this problem, we plan to use different platforms (Pytorch, caffe and Tensorflow) to implement the network. The classification accuracy and F-score of these methods will be compared.

Datasets and work plan

CIFAR-100 dataset will be used to train and test a neural network for this project.

CIFAR-100[1] consists of 60000 32x32 color images in 100 classes, with 600 images in each class. In these 600 images, there are 500 images in the training set and 100 images in the testing set. In addition, the 100 classes were grouped into 20 superclasses. For example, an image labeled as “dolphin” also belongs to the superclass of “aquatic mammals”. This adds one more dimension to the dataset and allows for more potential analyses.

The raw data for Python based modeling are available as “pickled” files, which can be unzipped and extracted using the “pickle” package. The data archive contains images for training grouped into 5 files (i.e. batches) and images for testing grouped into 1 file. The images are provided as numpy arrays with 3072 columns and 10000 rows. Each image is “flattened” into one row with the first 1024 columns containing the red channel values, the next 1024 the green and the final 1024 the blue. The labels of these images are indicated as 10000 numbers ranging from 0 to 9. The crosswalk between the numbers and the actual label is given in a dictionary by a separate file.

Rodrigo Benenson hosts a very comprehensive compilation of the papers and works[2] that explore the best performing approaches using a variety of datasets for deep learning projects.

---

[1] Official hosting page for CIFAR-100 dataset can be found at: https://www.cs.toronto.edu/~kriz/cifar.html

[2] “What is the class of this image?” at http://rodrigob.github.io/are_we_there_yet/build/classification_datasets_results.html#43494641522d313030 accessed Oct 29 2018.

According to his record, the best accuracy score of a deep learning method is approximately 76% and the lowest accuracy score is about 54%. These papers provide good reference information for this project, especially the range of accuracy to expect and target at.

The proposed work plan includes training and testing a multilayer neural network (e.g. CNN) as well as exploring hyperparameters that optimizes the performance and investigating the performance from multiple perspectives (such as intra-class variations of neural network performances and superclass accuracies).

Judgement

We are going to use the computational time and accuracy to judge the performance of the network. By using the system (sys) clock command in Python, computational time could be derived, and we may also implement another approach to enhance the accuracy of our result such confusion matrix method.

Reference Material

Djork-Arné Clevert, Thomas Unterthiner, Sepp Hochreiter (2015) Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs). ICLR 2016.

Yangqing Jia ; Chang Huang ; Trevor Darrell (2012) Beyond spatial pyramids: Receptive field learning for pooled image features. IEEE Conference on Computer Vision and Pattern Recognition.

Schedule