Machine Learning II – Final Project
December 2018 – Group 9
Junior Ovince Individual Report

American Sign Language (ASL) Model Classification

Introduction. – The group works on this ASL dataset, found on Kaggle[1], which contains 3000 datapoints for each of the 29 classes/alphabets. We decided to divide the work into three main components: the literature review of models used to classify images of hand gestures, the methods, and the results of our work. I worked mostly on the methods and results.

Description of my work. – For the methods, I followed the guidelines on chapter 22 from the Neural Network Design book. So, I developed three main sections: pretraining, training, post training. The actual results follow these sections. For the pretraining, I developed the whole task of data preprocessing, particularly the split of the sets into training, validation, and test one. I also have a great contribution in developing the training section and I used the following links to develop the codes:

https://github.com/PacktPublishing/Deep-Learning-with-PyTorch (chapters 4 and 5)

https://github.com/kuangliu/pytorch-cifar/blob/master/models/lenet.py

https://github.com/pytorch/vision/blob/master/torchvision/models/alexnet.py

Results. - The results were presented in the post training section and my contribution was also important. I trained the models and summarized the results into the tables. I also participated in the conclusion of the work.

Summary and Conclusion. - Among the three models explored, the baseline and the LeNet ones were promising. A two convolutional layer network is enough to find the ground truth of the ASL image classification problem. We could finetune some hyperparameters using the different models. We could also explore more the feature maps obtained.

Initially, I was thinking to use TensorFlow but the learning curve was steep for me. So, I used Pytorch and managed to connect it through Tensorboard. The capability wwas not great and I will dedicate more time to learn Tensorflow and use Tensorboard as I like the way one can do better graphical analysis with it by comparing different model performances and so one.

References. - The Neural Network Design book, the three links aforementioned, and the codes in the exam 1 were used to develop this project. Around 70% of the codes are from those sources.

---

[1] https://www.kaggle.com/kairess/99-9-asl-alphabet-classification-with-slimcnn/data