Project Proposal 8420

Will Sherrer
Clemson University
wsherre@g.clemson.edu

Ashton Sobeck
Clemson University
asobeck@g.clemson.edu

1 Proposal

Our group wants to explore the effectiveness of Convolutional Neural Networks (CNN) when images that have been reduced by Principal Component Analysis (PCA) have been input as training. We aim to use the Canadian Institute for Advanced Research (CIFAR) dataset to conduct PCA on and compare how different amount of features in input images effect the accuracy of a CNN.

2 Motivation

Throughout our class of CPSC 8420, we have learned many things. One of the most crucial concepts of the class is certainly Singluar Value Decomposition (SVD). Using SVD, one can utilize PCA to find the important features of data. CNNs are quite effective in identifying patterns within images through the use of filters and non-linearity within their structure. However, when input images get larger and larger, the network can grow in size and in return, require much more computational power to train in a timely manner.

3 Method

The machine learning techniques we plan to implement are PCA and SVD. These will be used to reduce the dimension of an images which the CNN models will be trained and tested on. A technique we could be improving on is seeing if AI models can maintain performance whilst minimizing input data. We believe this could be a good preliminary study on performance vs training sample data.

4 Intended Experiments

Our experiment we are thinking about is designing 2 CNN models with different architectures. This will just test if one architecture is better than another with testing. We then plan to take the CIFAR images and train the models with the original image as a control. We will record the training loss and accuracy as well as testing loss and accuracy. Then we will use PCA to reduce the data further to 20%, 40%, 60%, and 80% of the original and train both models with the same hyper-parameters for consistency. We will record the training and testing loss and accuracy and plot the data and compare to see if the models are able to learn a reduced image.

5 Conclusion

We found an interesting paper, "Image retrieval method based on CNN and dimension reduction" (https://arxiv.org/pdf/1901.03924.pdf) that also conducts a study on this topic. They found that while CNN's can still have satisfactory performance, further practical and theoretical studies should be conducted to prove the viability and best parameters.

35th Conference on Neural Information Processing Systems (NeurIPS 2021).