Project Documentation for Mathematics of Neural Networks

Johann Kipping

January 19, 2022

Abstract

The aim of this project is to test the learned things in the first half of the lecture series on Neural Networks on the example of the Fashion-MNIST Dataset

Things that could be done on Fashion-MNIST

- Hyperparameter search
- Activation functions
- Initialization
- Optimizers
- Regularization (Overfitting?) + Generality
- FFT/im2col/Winograd
- Attribution
- Feature Visualization
- Adversarial Attacks
- Going deep?

Section 0: Getting TF to run with a graphics card with compute capability of 3.0

AS this is a documentation of work this will be included here as it took a lot of time

1 Idea collection and planning

This chapter aims to collect ideas to most of the topics that were discussed in the lecture and continues to formulate a plan on how to tackle them in this project work.

Loss functions used is categorical cross entropy. Why not Mean squared or something else

Hyperparameters:
Idea: Try several choices of hyperparameters to get a feel for the effects. Do this in a grid-like fashion.

Activation functions / network structure
Idea: Test small network structures (fast trainable) for effect of structural changes and different choices of activation functions using knowledge from the lecture and Exercise 4 from sheet 6 as starting point for network evaluation.

Initialization
Idea: Verify the validity of the three approaches discussed in the lecture. (Zero/Random/Known data)

Optimizers
Idea: Use Adam for all computations to reduce complexity of project as it incorporates other approaches.

Regularization (Overfitting?) + Generality
Idea: Test for overfitting and generalization with own pictures (for fun). Maybe try regularization if poor performance is observed.

FFT/im2col/Winograd
Idea: Let tensorflow decide on the usage of different approaches as insight was achieved in the exercises and modifications are not cheap.

Progress will be documented with screenshots of the console and explanation. Graphs will be used.

2 Getting to know Neural Networks

2.1 Starting out

As a starting point the network structure from Exercise 4 of sheet 6 was chosen. As described in section 1, Adam was be used throughout this section in order to reduce complexity.

2.2 Network structure

2.3 Initialization experiments

2.4 Hyperparameters

3 Regularization and Attribution

Check for generality and attribute in the same step by using an image from the test set and an image that was taken by me

Using gradient tape -¿ saliency map
Global average pooling -¿ class activation map
Grad-CAM! am ehesten