ML Model Optimization Strategies for Image Datasets

Alexander Peplowski | Azfar Khoja | Daniel Wang

Objectives

Evaluate:

1. Network **Shape** vs Performance
2. Data **Availability** vs Performance
3. Network **Node Count** vs Performance
4. Target Algorithm vs **Convergence Time**
5. **Regularization** vs Performance

Baseline Algorithms

Convolutional Neural Network

- Batch Normalization
- 4 Convolutional Layers:
  - (3x3) Kernel
  - 32,32,64,64 Feature Maps
- Max Pooling x2
- Two FC Layers

Neural Network

- Batch Normalization
- 3 Layers
- 490 Nodes per hidden layer
- **Equal** number of **parameters** as CNN baseline

SVM

- Linear Kernel
- Regularization alpha = 0.01
- Gradient Descent Optimizer

Data Sets

CIFAR-10

- 10 Classes
- 32x32 Image Size
- 3 Color Channels
- 60,000 images

Fashion MNIST

- 10 Classes
- 28x28 Image Size
- Grayscale (1-Channel)
- 70,000 Images

Procedure

1. Network Shape vs Performance

Step	Description
Define Models	Generate models with **equal parameter count** for MLP and CNN depths from1 to 8
Run Experiment	Evaluate validation loss, test accuracy using CIFAR-10 for CNN and F-MNIST for MLP

2. Data Availability vs Performance

Step	Description
Split Training Set	Generate **smaller stratified sets** sampled from CIFAR-10 and F-MNIST
Data Augmentation	Reflect images horizontally to double data set size
Run Experiment	Evaluate test set accuracy using above sets

3. Network Node Count vs Performance

Step	Description
Define Models	Generate models with equal depth and **variable number of neurons**
Run Experiment	Evaluate validation loss, test accuracy using CIFAR-10 and F-MNIST

4. Target Algorithm vs Convergence Time

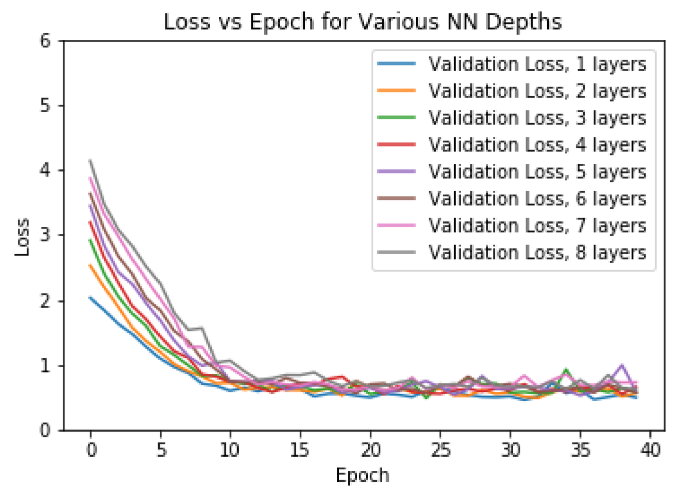
Step	Description
Define Models	Compile **baseline models** with gradient descent optimizer
Run Experiment	Evaluate validation loss, test accuracy using F-MNIST with fixed epoch count

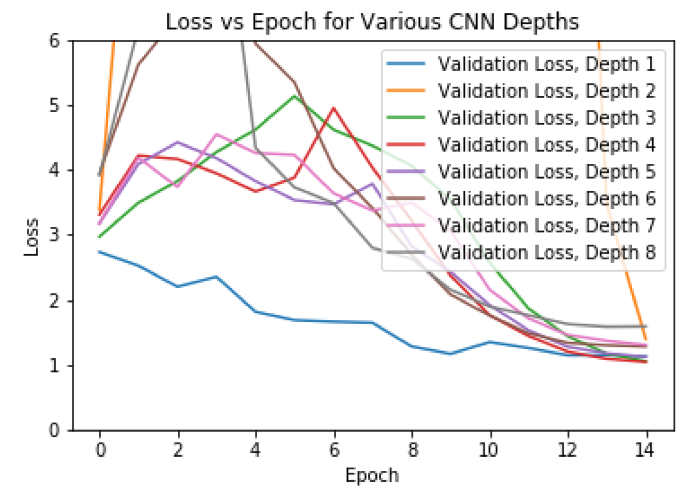
5. Regularization vs Performance

Step	Description
Define Models	Generate CNN, MLP models with **variable l2 regularization values**
Run Experiment	Evaluate validation loss using CIFAR-10 and F-MNIST

Results

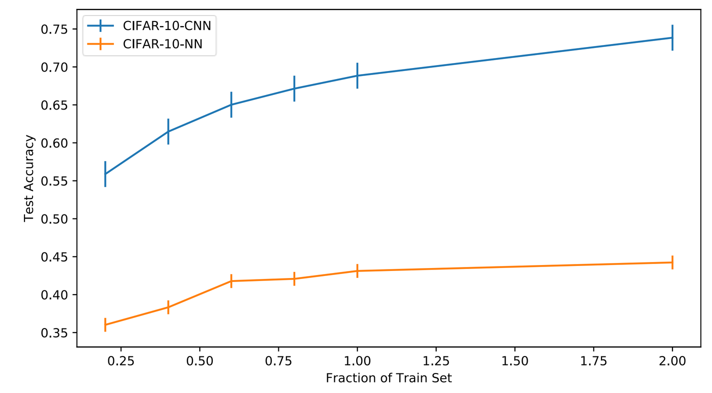
1. Network Shape vs Performance



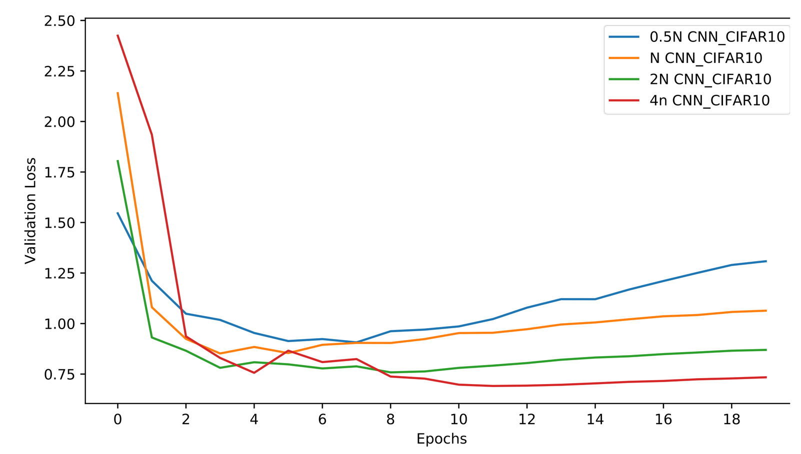


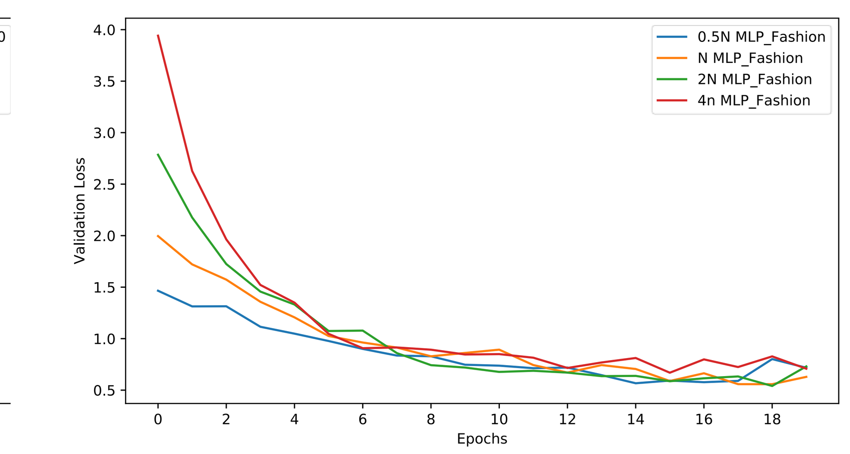
Results

2. Data Availability vs Performance

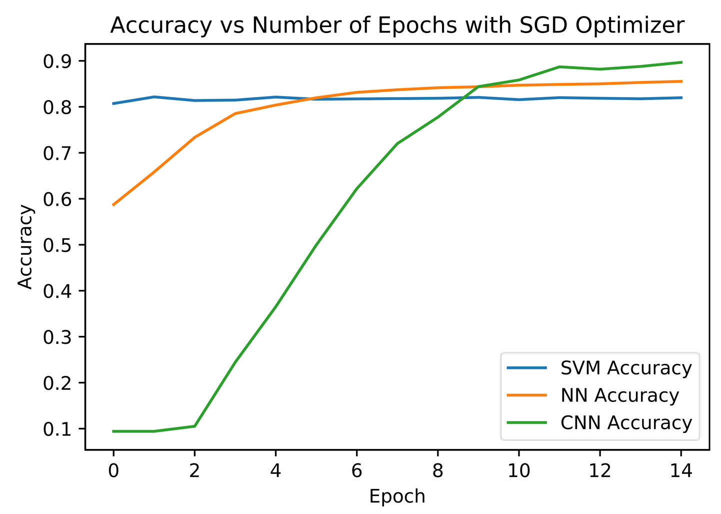


3. Network Node Count vs Performance

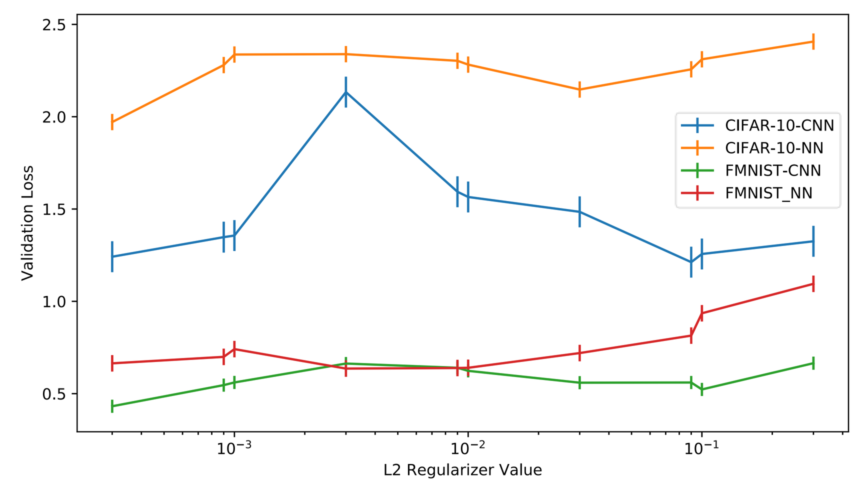




4. Target Algorithm vs Convergence Time



5. Regularization vs Performance



Conclusions

- 1.1: When controlling for number of parameters, MLP network depth **has little impact** on performance for the f-MNIST dataset
- 1.2: When controlling for number of parameters, CNN network depth is **best at depth 4** for the CIFAR-10 dataset
- 2: Increasing data availability has a significant impact on CNN, small impact on NN. **Data augmentation** helps in both cases.
- 3.1: When varying the node count, the **highest capacity** CNN model has the best performance and fastest training time on the CIFAR-10 dataset.
- 3.2: When varying the node count, the capacity of MLP models **does not impact** performance for the F-MNIST data set.
- 4: SVMs with linear kernels have **fast convergence**, but CNNs have the **best performance** on the F-MNIST dataset
- 5: Not clear how to select a general regularizer across datasets, use hyperparameter search!

Mila