

EARIN Lab 5 Report

Krzysztof Rudnicki, 307585
Jakub Kliszko, 303866

May 16, 2023

1 Exercise Variant 2

Use MIST dataset. Evaluate at least 3 different numbers/values/types of:

- learning rate
- mini-batch size (including batch containing only 1 example)
- number of hidden layers (including 0 hidden layers - linear model)
- width (number of neurons in hidden layers)
- optimizer type (e.g., SGD, SGD with momentum, Adam)

2 Implementation

Program can be ran by installing python, moving to project directory and issuing command:

```
python main.py
```

Results will be displayed on three 2d scatter plots.

Plots Title is filled with parameters used abbreviated for space sake

Abbreviations meaning:

- lr - Learning Rate
- bs - Batch Size
- hl - Number of Hidden Layers
- w - Width
- Adam or SGD or SGD_Momentum - Optimizer type

Plot types:

- Loss value for learning step
- Train Accuracy for epoch
- Validation Accuracy for epoch

Figure 1: Exemplary plot of loss for learning Step
lr0.001-bs64-hl1-w128-Adam

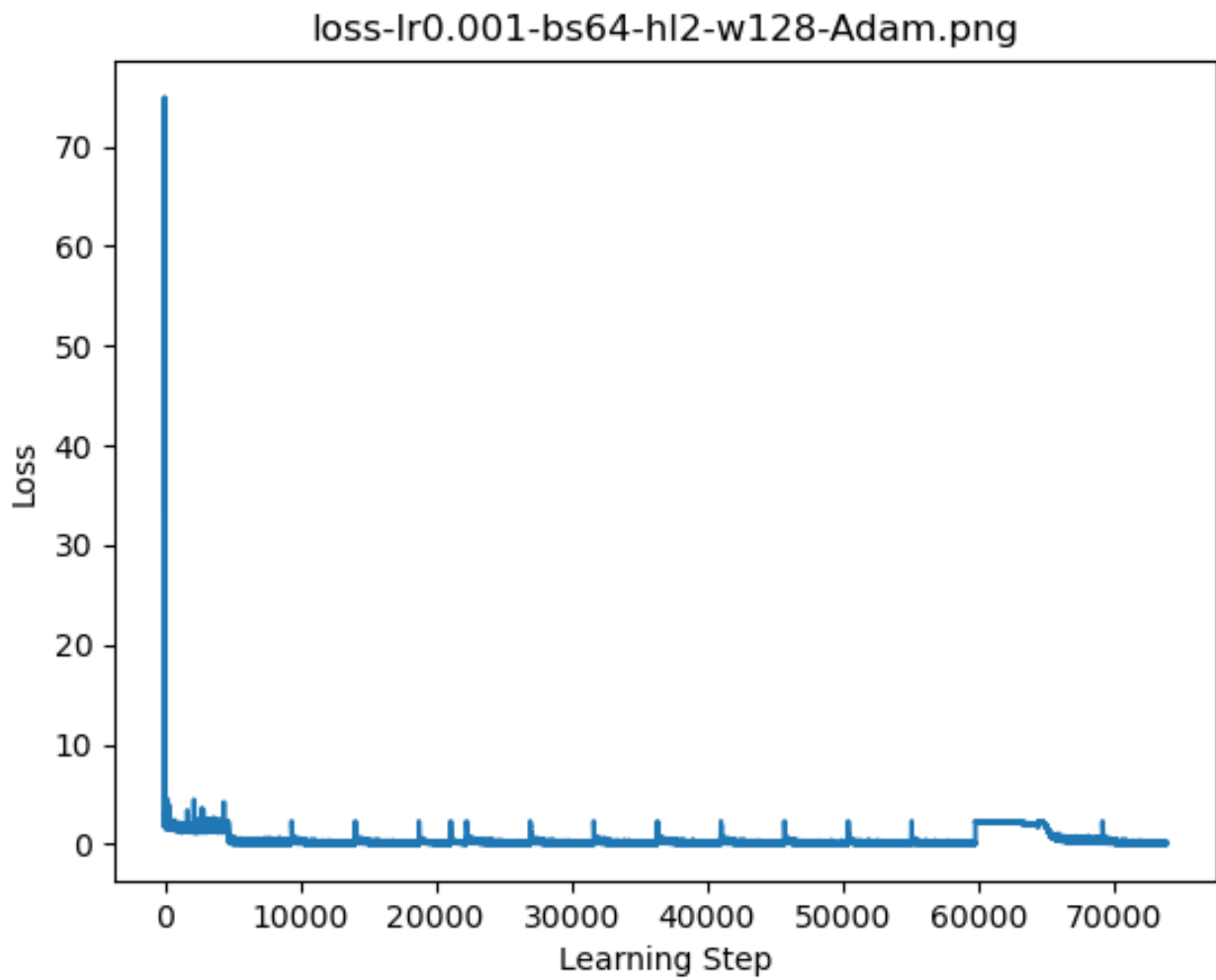


Figure 2: Exemplary plot of Validation Accuracy for epoch
lr0.001-bs64-hl1-w128-Adam

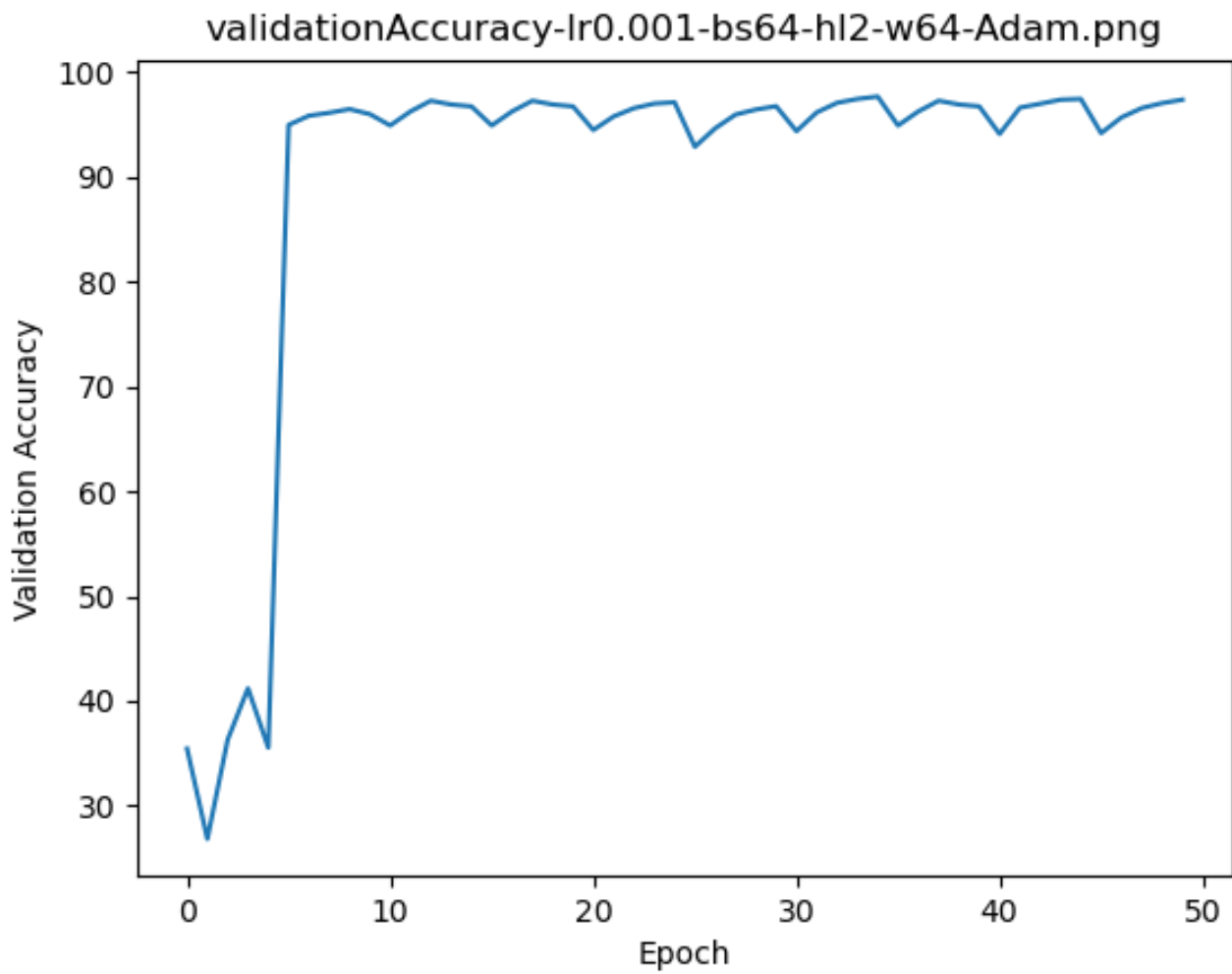
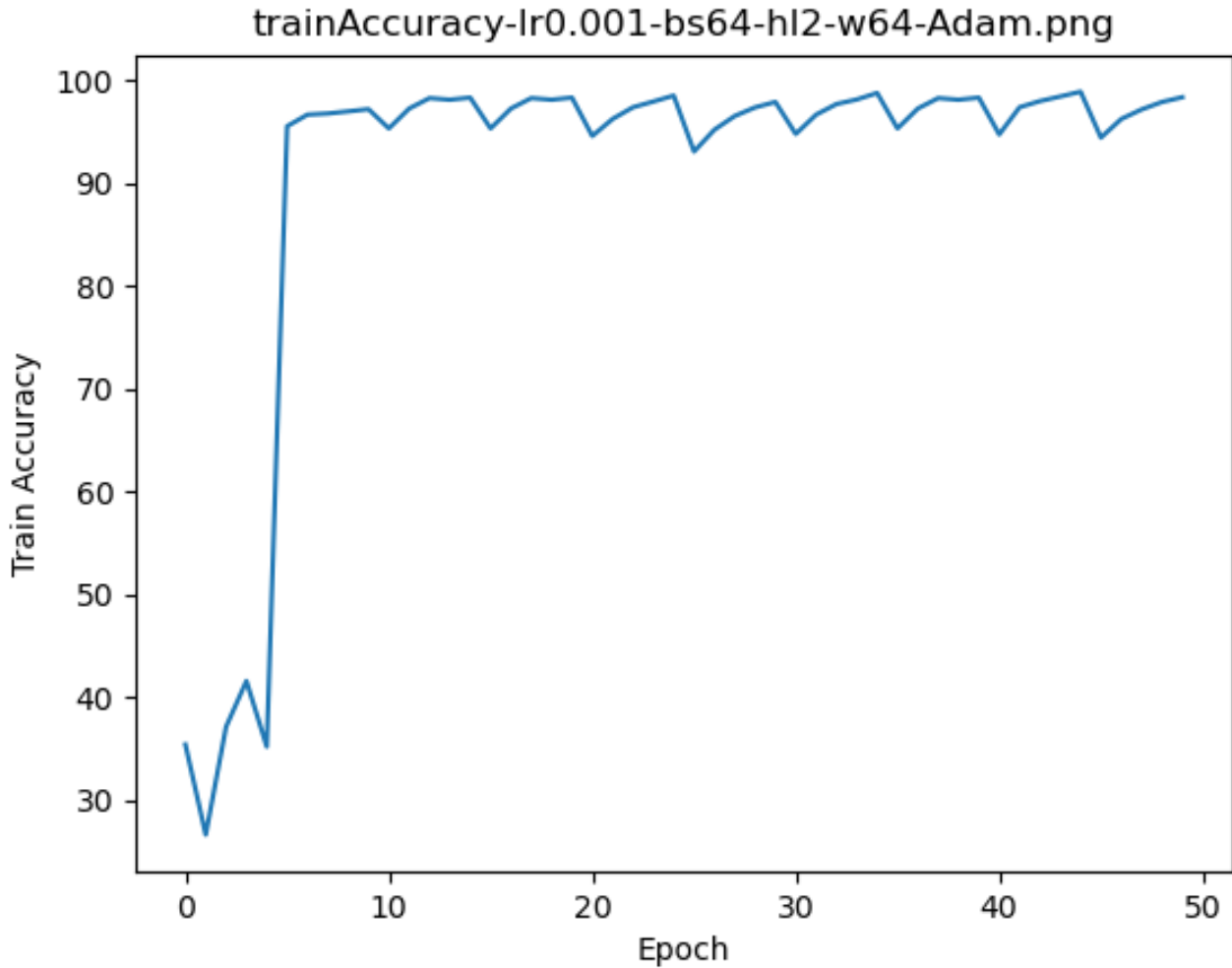


Figure 3: Exemplary plot of Train Accuracy for epoch
lr0.001-bs64-hl1-w128-Adam



Results will be displayed and saved in the same folder as code directory for further inspection, with file name containing information about input parameters

Additionally speed it took for neural network to run will be saved to results.txt along with what parameters were used

We decided to run 19 tests in total:

1. learning rate [0.1, 0.01, 0.001] (3 tests)
2. mini-batch size [1, 64, 128, 256] (4 tests)
3. number of hidden layers [0, 1, 2, 3] (4 tests)
4. width [64, 128, 256, 512, 1024] (5 tests)
5. optimizer type [SGD, SGD_Momentum, Adam] (3 tests)

3 Results

We have successfully implemented network analyzing MNIST dataset

3.1 Loss Graphs

3.1.1 Learning Rate

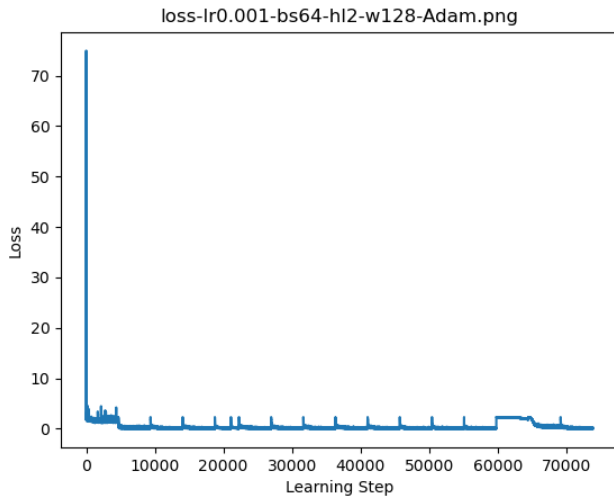


Figure 4: Default settings + learning rate = 0.001

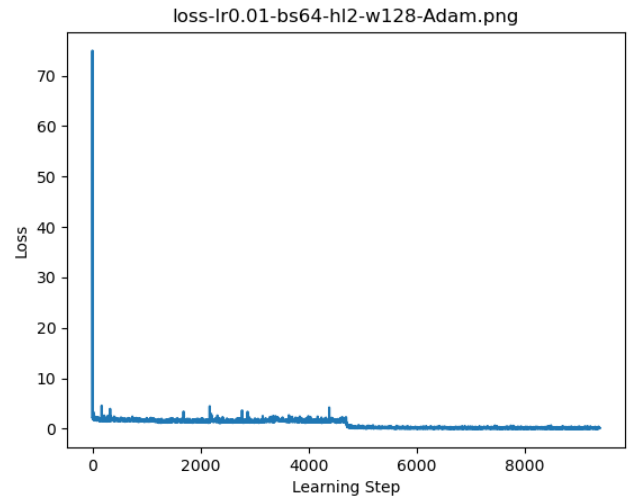


Figure 5: Default settings + learning rate = 0.01

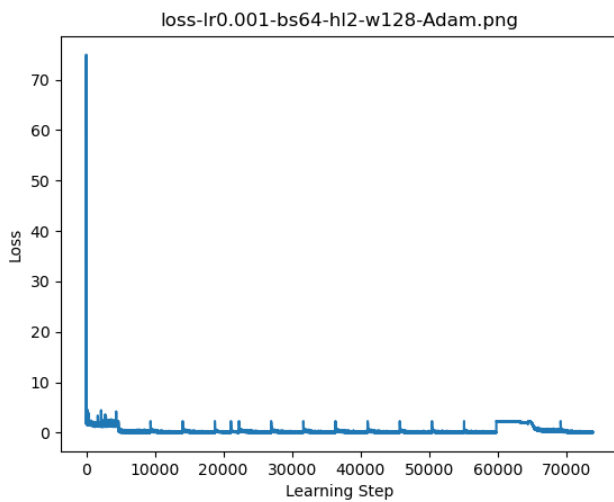


Figure 6: Default settings + learning rate = 0.1

3.1.2 Mini-Batch size

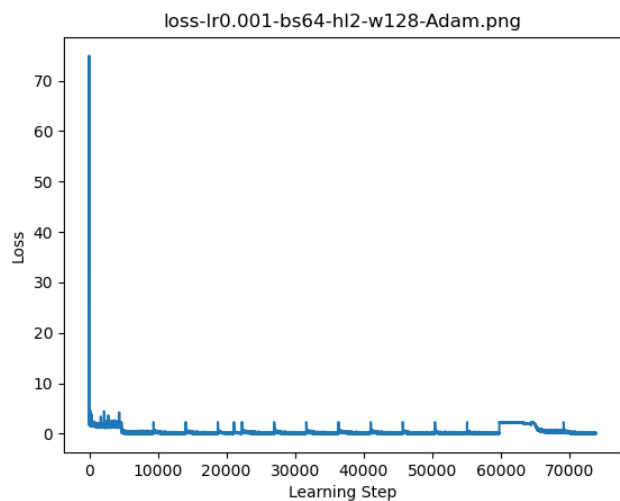


Figure 7: Default settings + batching size = 64

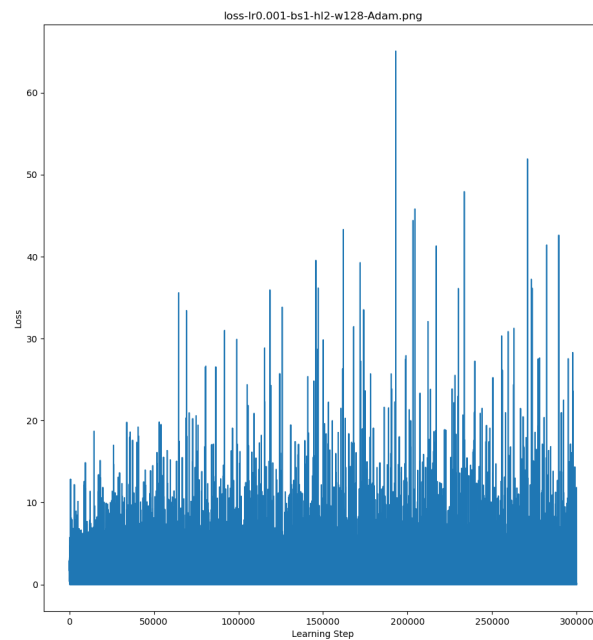


Figure 8: Default settings + batching size = 1

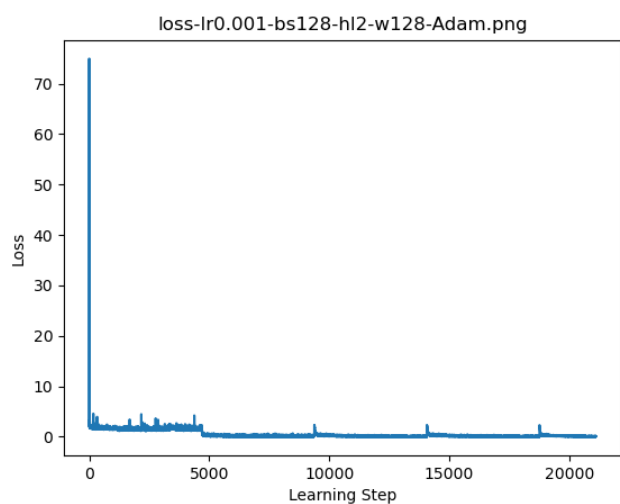


Figure 9: Default settings + batching size = 128

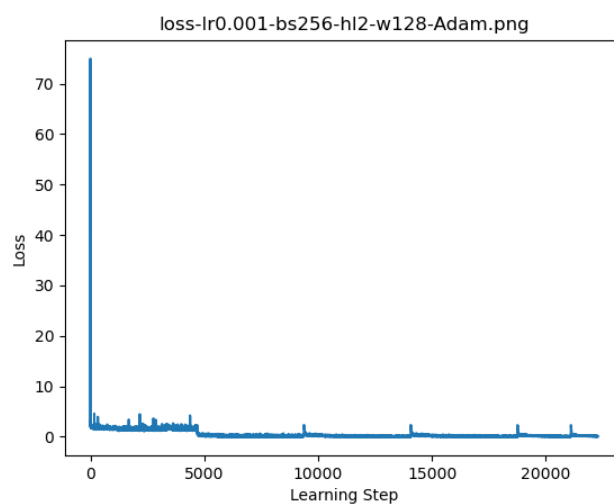


Figure 10: Default settings + batching size = 256

3.1.3 Number of Hidden Layers

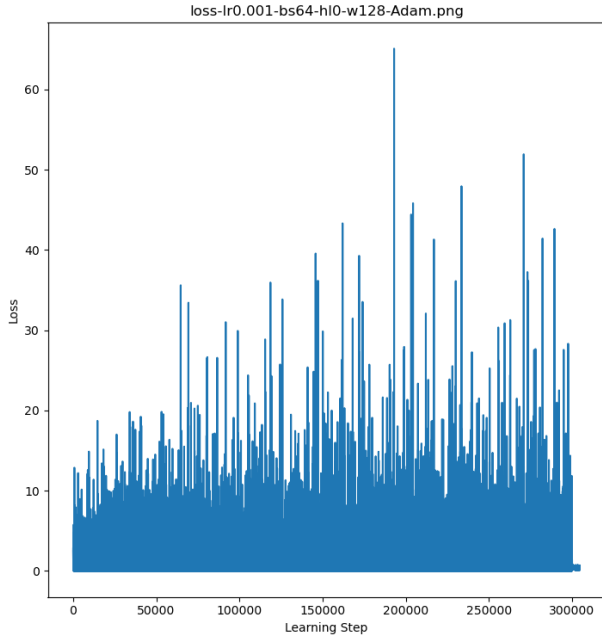


Figure 11: Default settings + hidden layers = 0

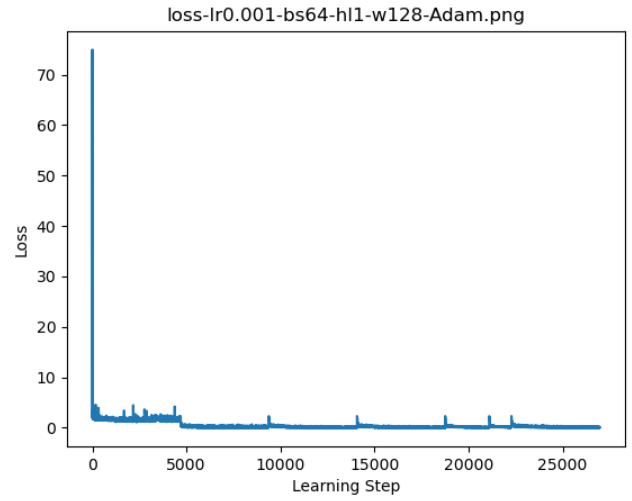


Figure 12: Default settings + hidden layers = 1

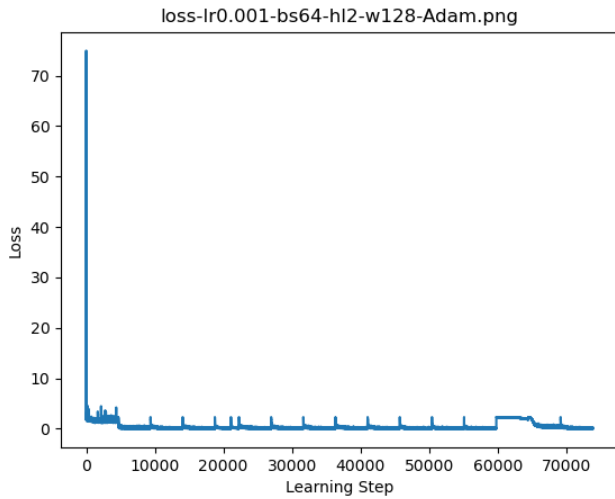


Figure 13: Default settings + hidden layers = 2

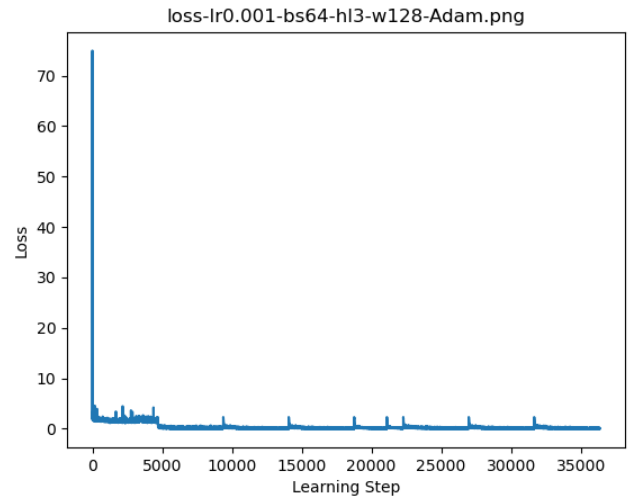


Figure 14: Default settings + hidden layers = 3

3.1.4 Width

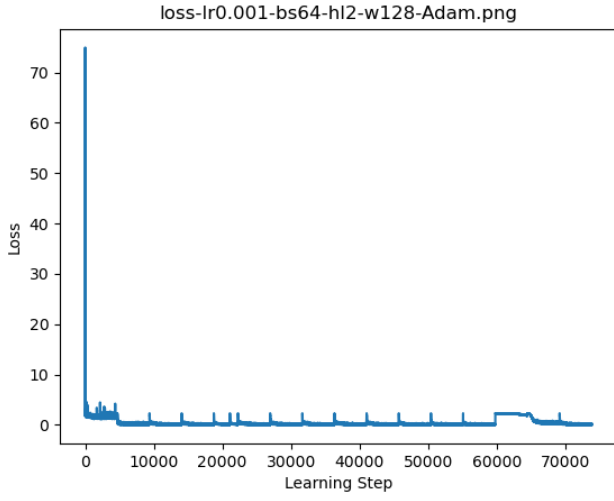


Figure 15: Default settings + width = 128

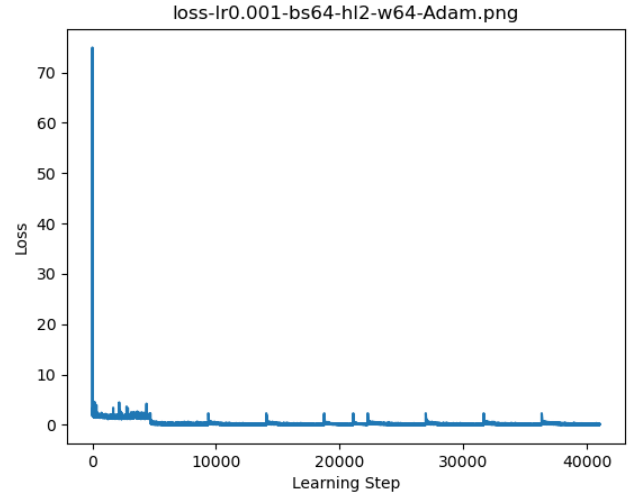


Figure 16: Default settings + width = 64

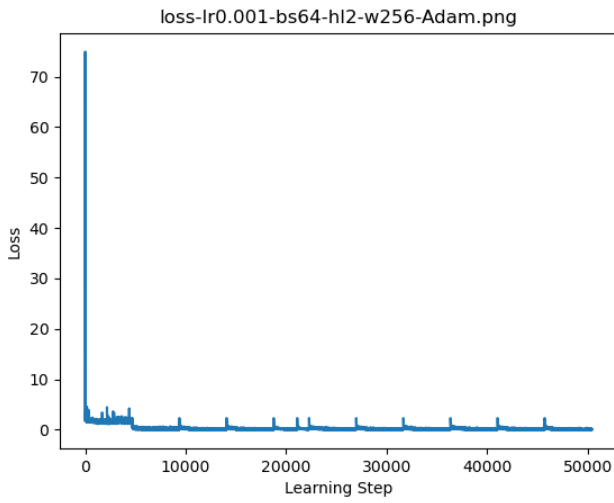


Figure 17: Default settings + width = 256

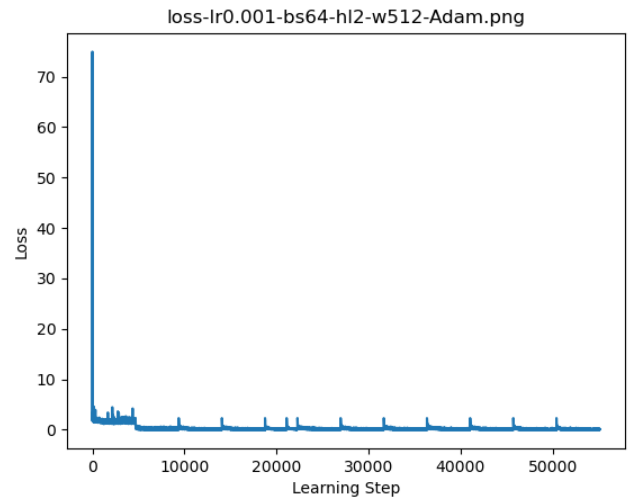


Figure 18: Default settings + width = 512

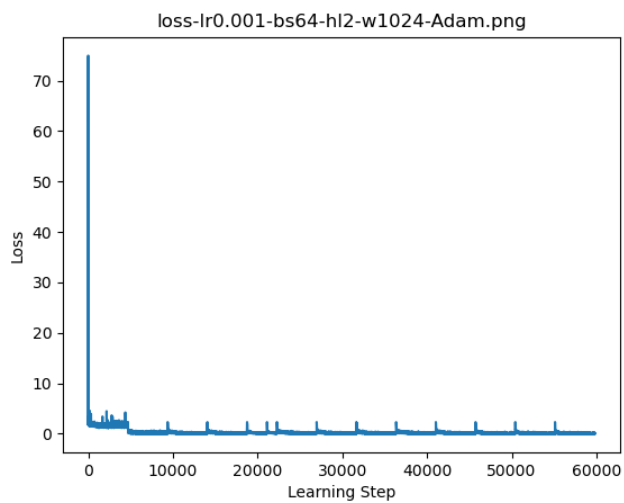


Figure 19: Default settings + width = 1024

3.1.5 Optimizer Type

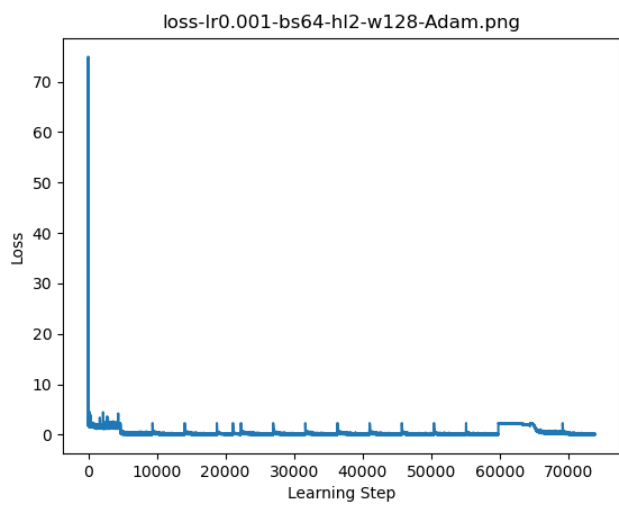


Figure 20: Default settings + Adam optimizer

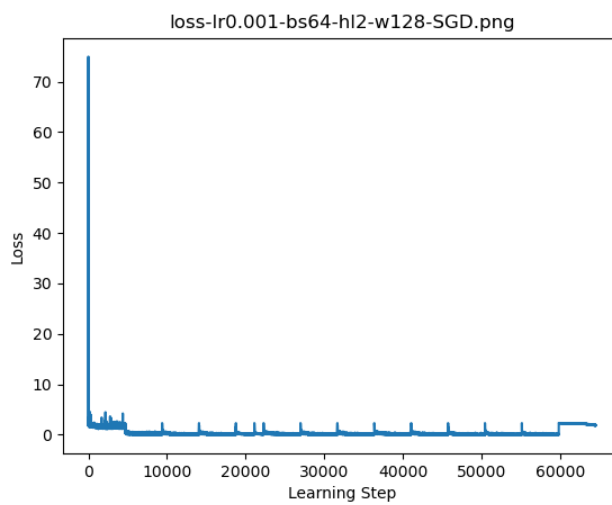


Figure 21: Default settings + learning rate = 0.01

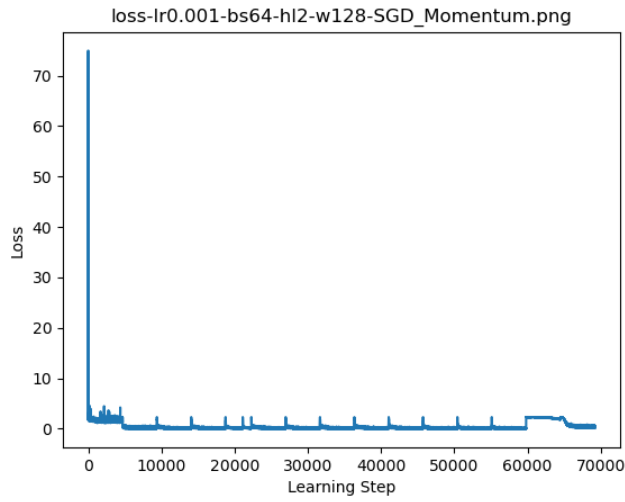


Figure 22: Default settings + SGD_Momentum optimizer

3.2 Train Accuracy Graphs

3.2.1 Learning Rate

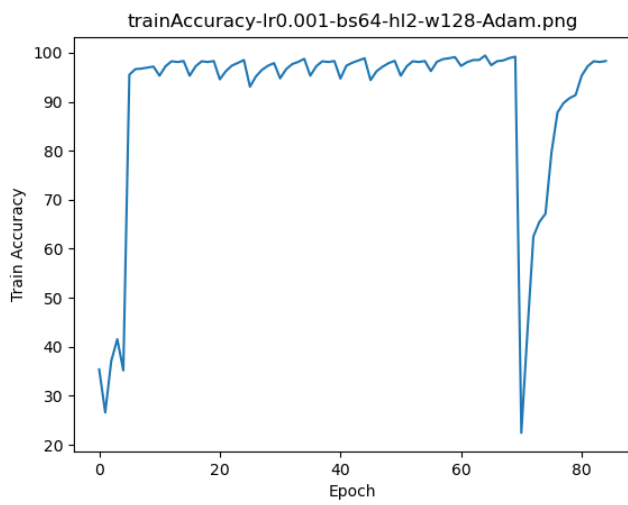


Figure 23: Default settings + learning rate = 0.001

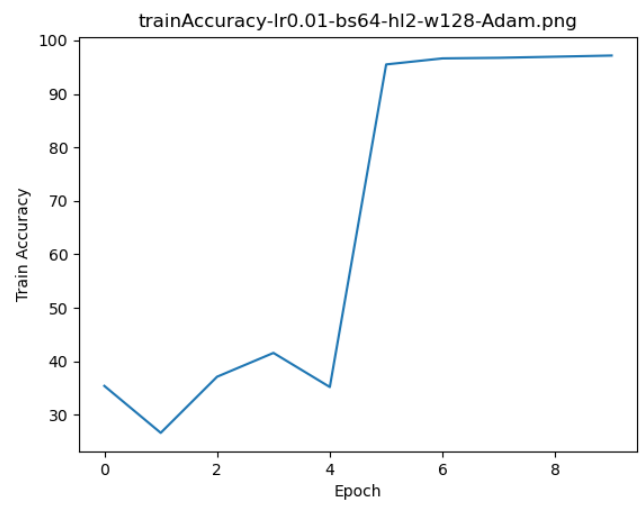


Figure 24: Default settings + learning rate = 0.01

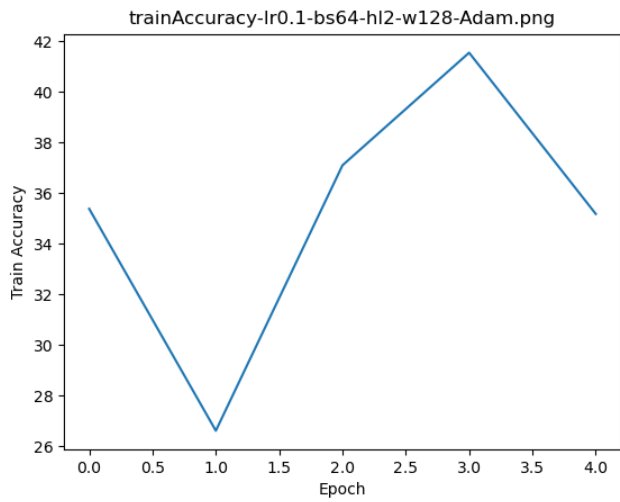


Figure 25: Default settings + learning rate = 0.1

3.2.2 Mini-Batch size

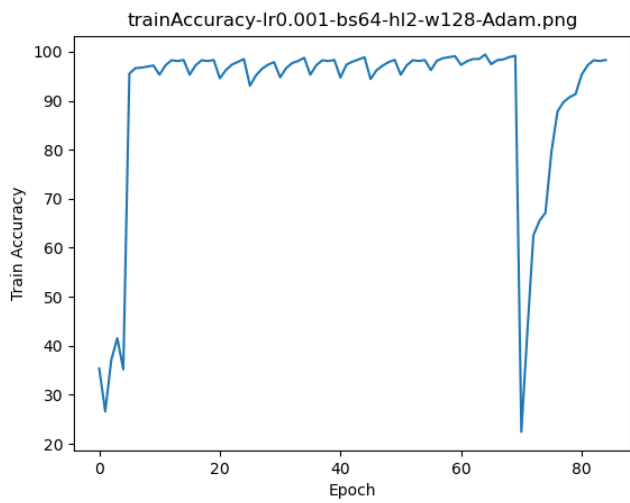


Figure 26: Default settings + batching size = 64

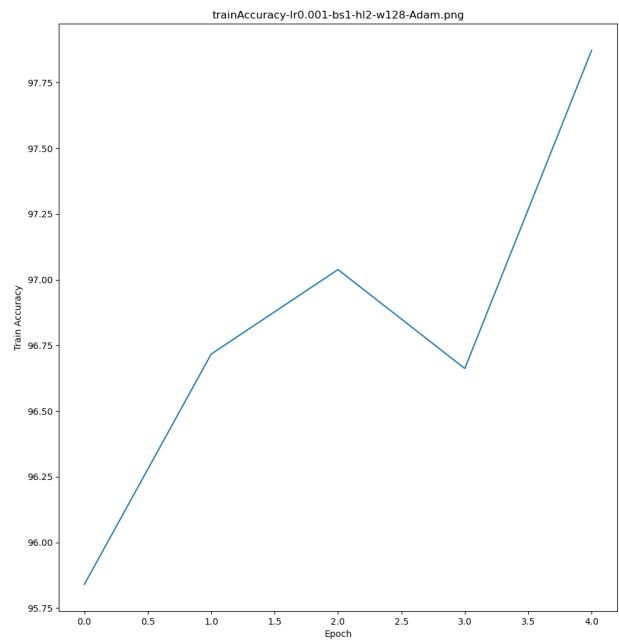


Figure 27: Default settings + batching size = 1

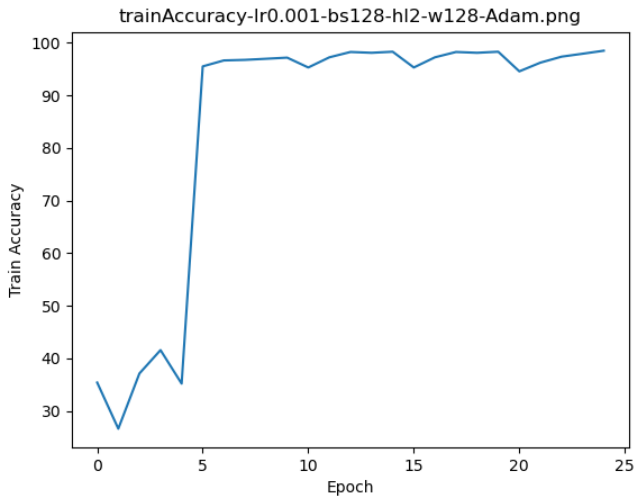


Figure 28: Default settings + batching size = 128

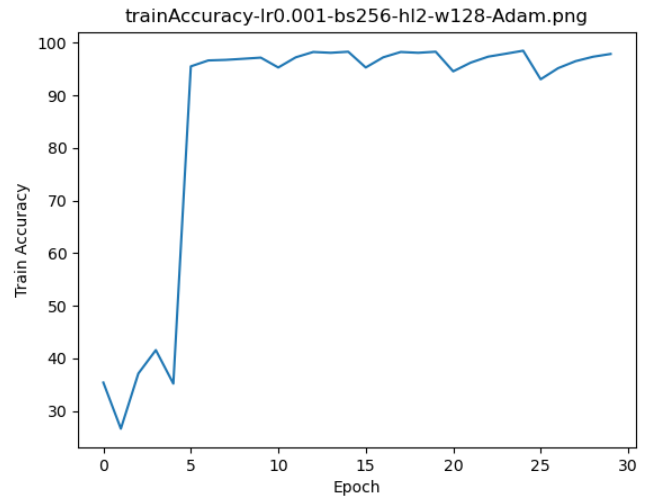


Figure 29: Default settings + batching size = 256

3.2.3 Number of Hidden Layers

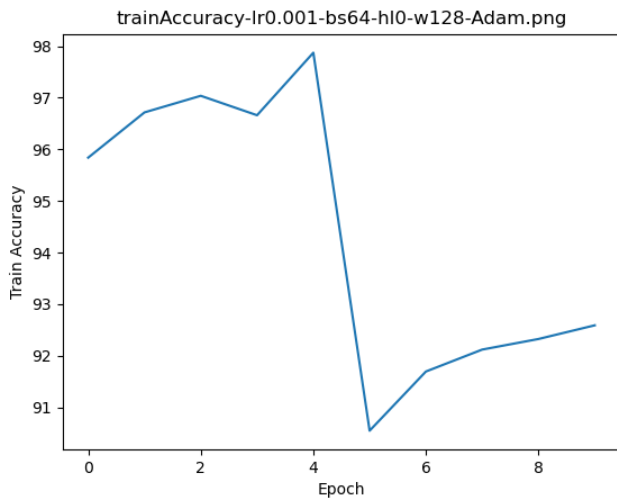


Figure 30: Default settings + hidden layers = 0

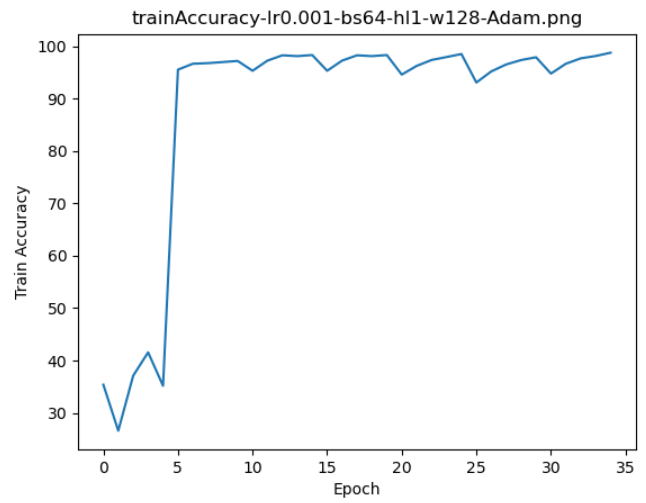


Figure 31: Default settings + hidden layers = 1

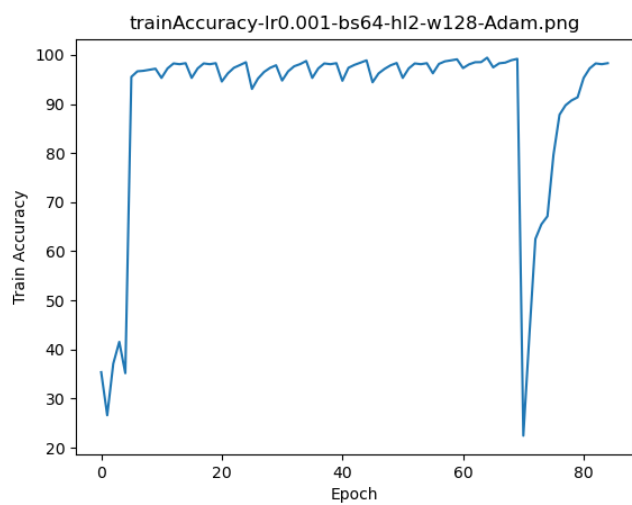


Figure 32: Default settings + hidden layers = 2

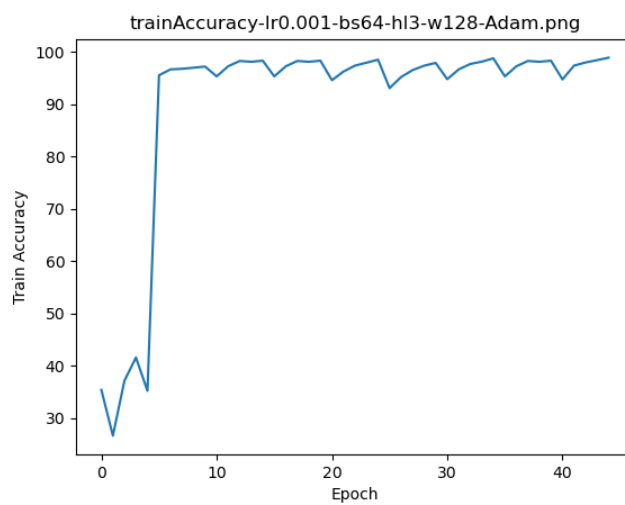


Figure 33: Default settings + hidden layers = 3

3.2.4 Width

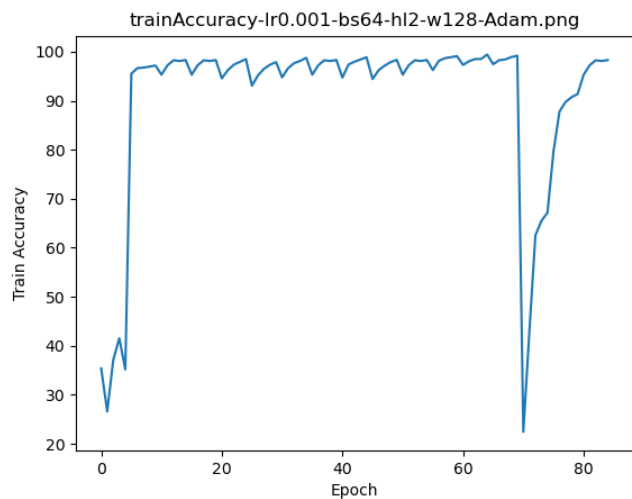


Figure 34: Default settings + width = 128

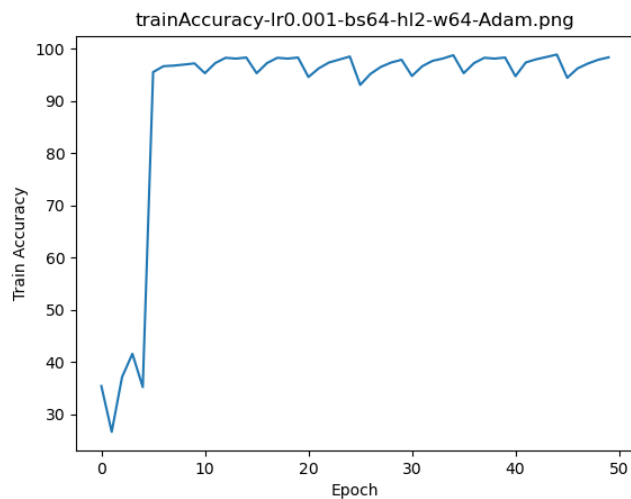


Figure 35: Default settings + width = 64

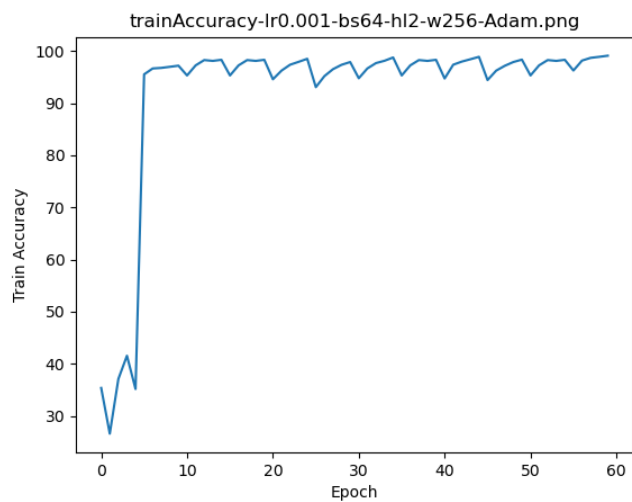


Figure 36: Default settings + width = 256

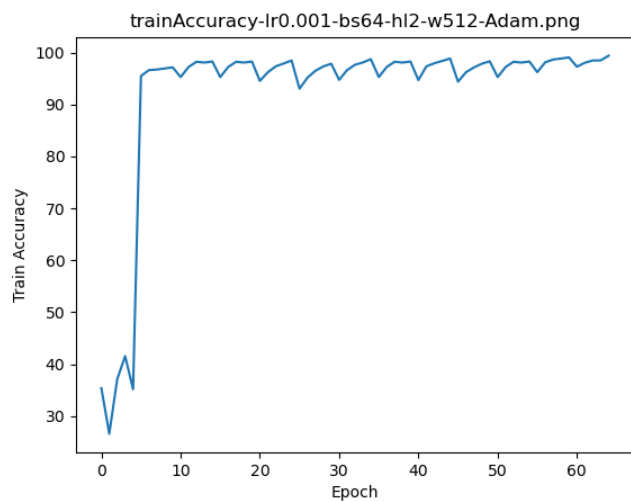


Figure 37: Default settings + width = 512

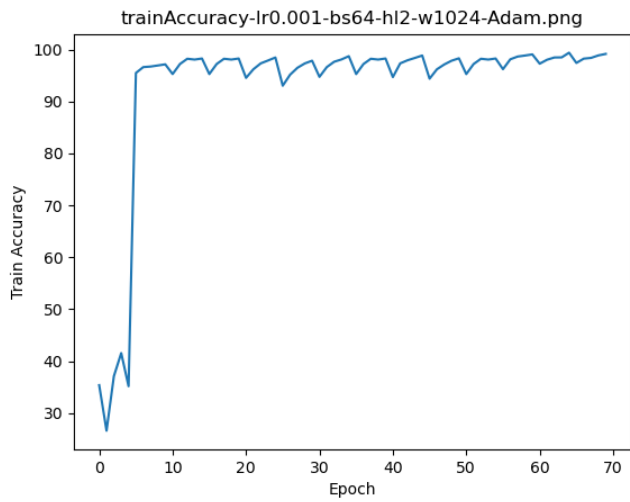


Figure 38: Default settings + width = 1024

3.2.5 Optimizer Type

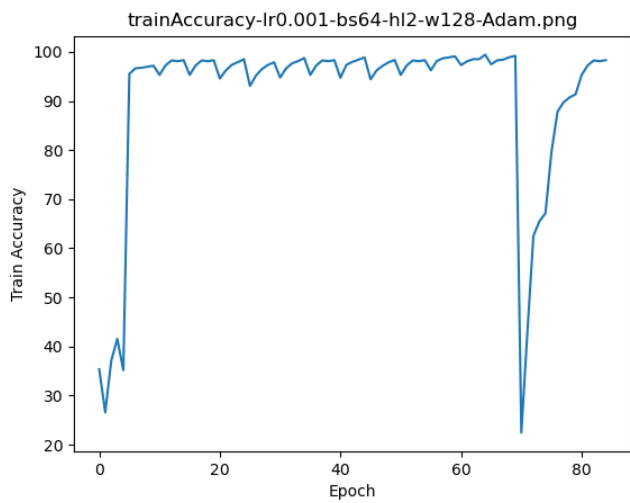


Figure 39: Default settings + Adam optimizer

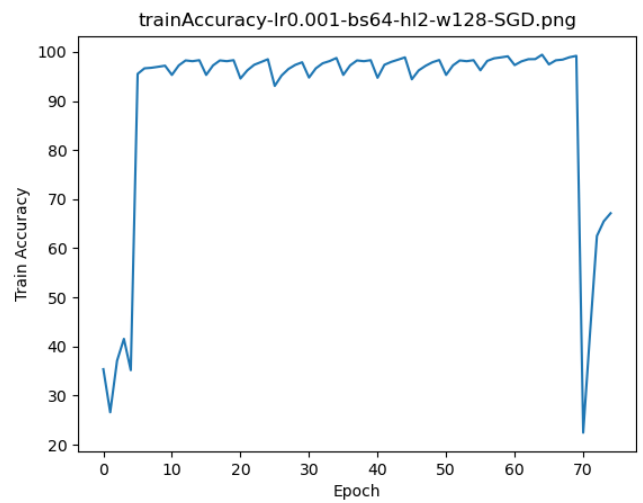


Figure 40: Default settings + learning rate = 0.01

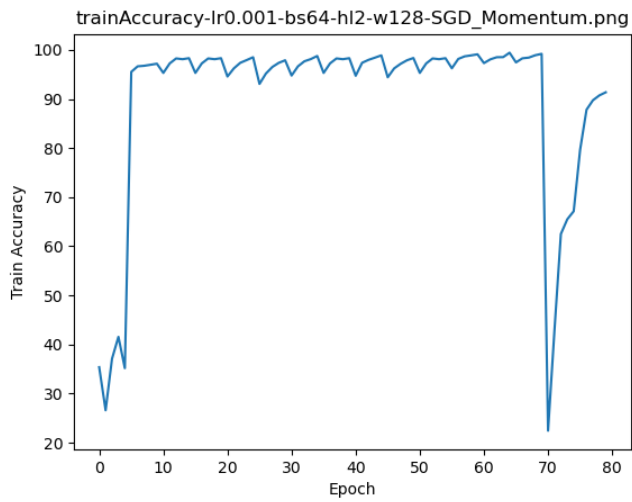


Figure 41: Default settings + SGD_Momentum optimizer

3.3 Validation Accuracy Graphs

3.3.1 Learning Rate

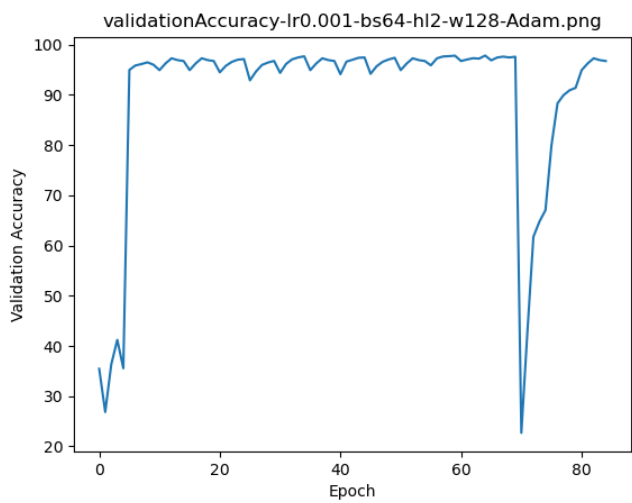


Figure 42: Default settings + learning rate = 0.001

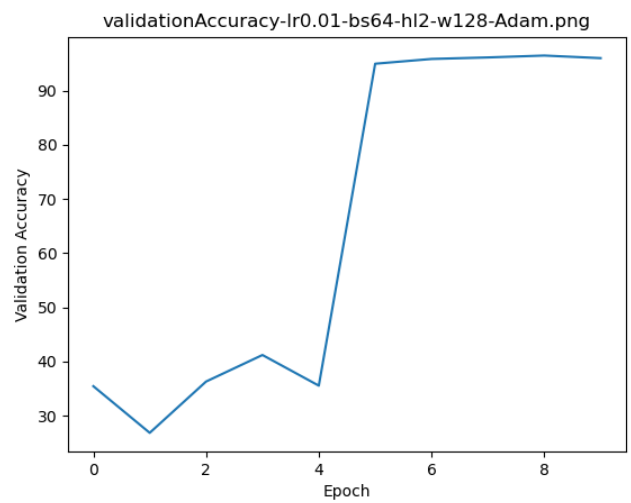


Figure 43: Default settings + learning rate = 0.01

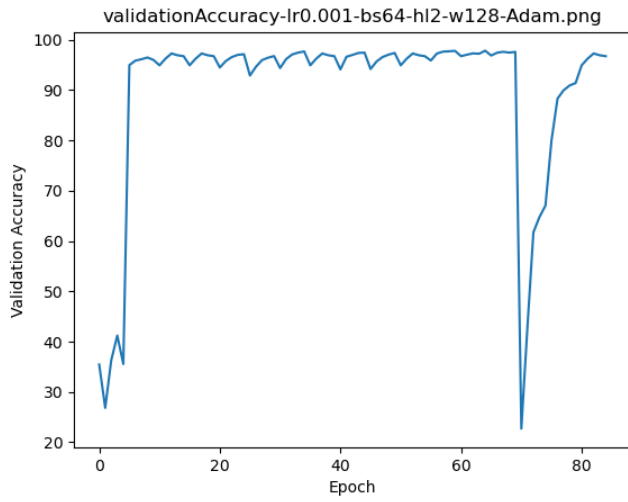


Figure 44: Default settings + learning rate = 0.1

3.3.2 Mini-Batch size

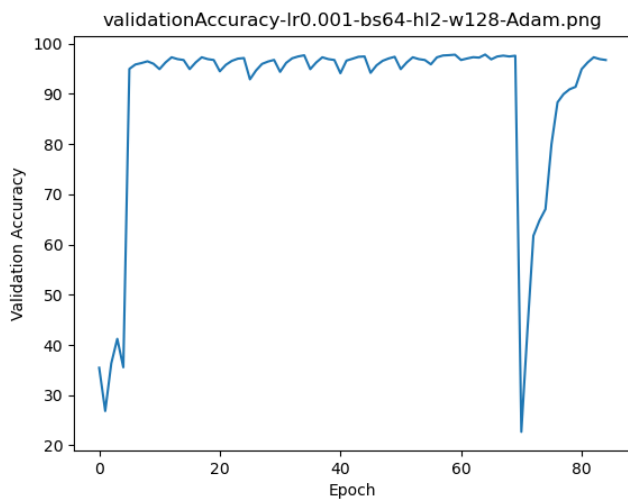


Figure 45: Default settings + batching size = 64

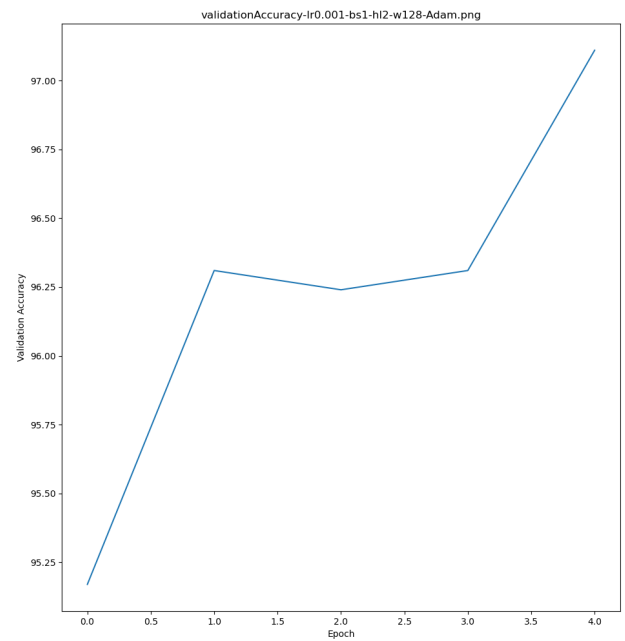


Figure 46: Default settings + batching size = 1

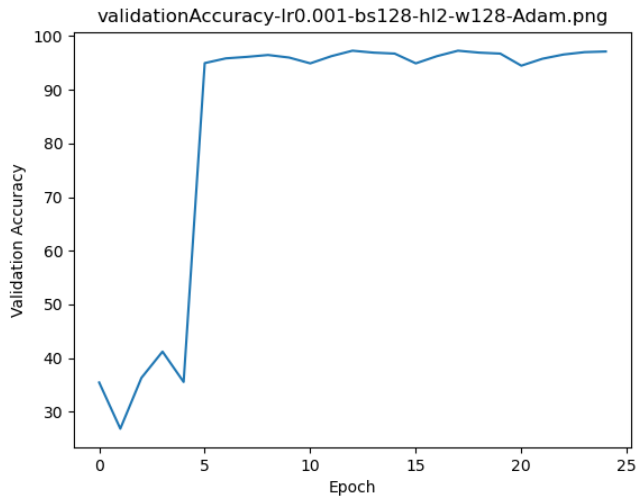


Figure 47: Default settings + batching size = 128

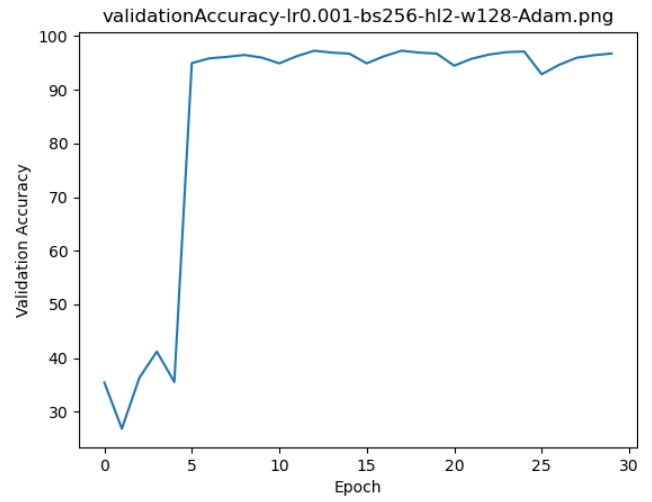


Figure 48: Default settings + batching size = 256

3.3.3 Number of Hidden Layers

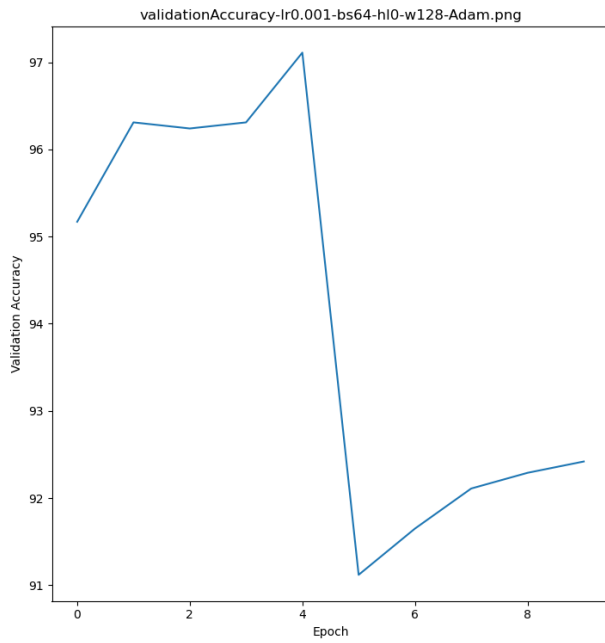


Figure 49: Default settings + hidden layers = 0

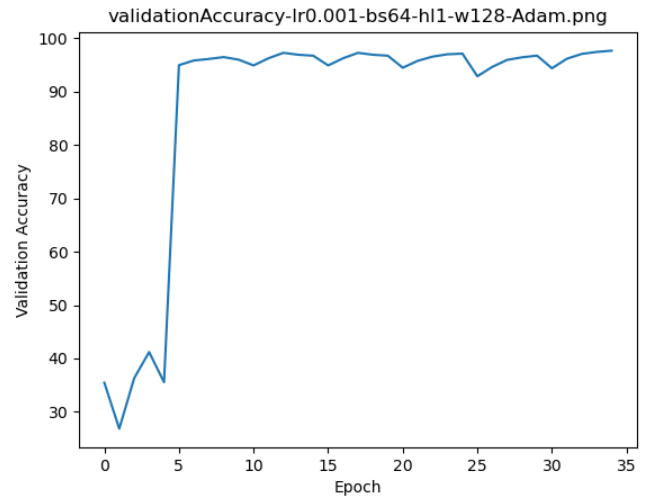


Figure 50: Default settings + hidden layers = 1

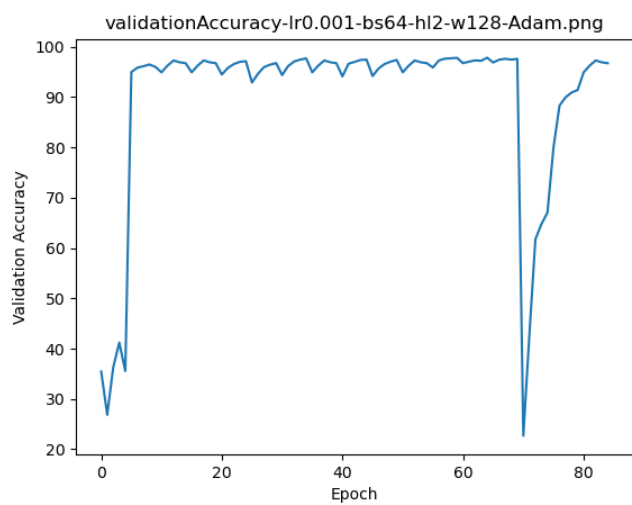


Figure 51: Default settings + hidden layers = 2

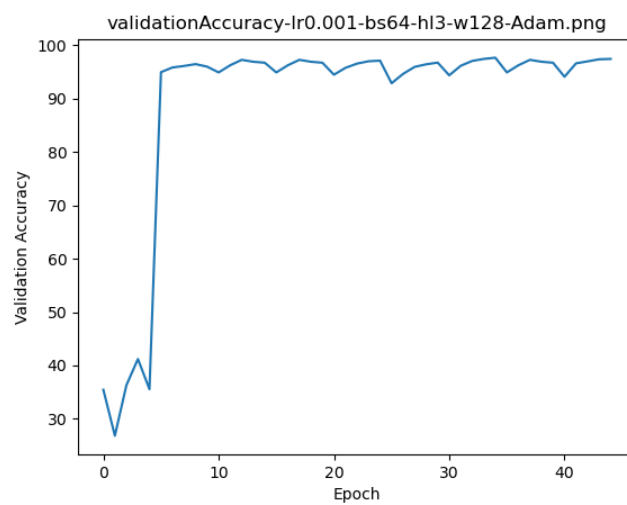


Figure 52: Default settings + hidden layers = 3

3.3.4 Width

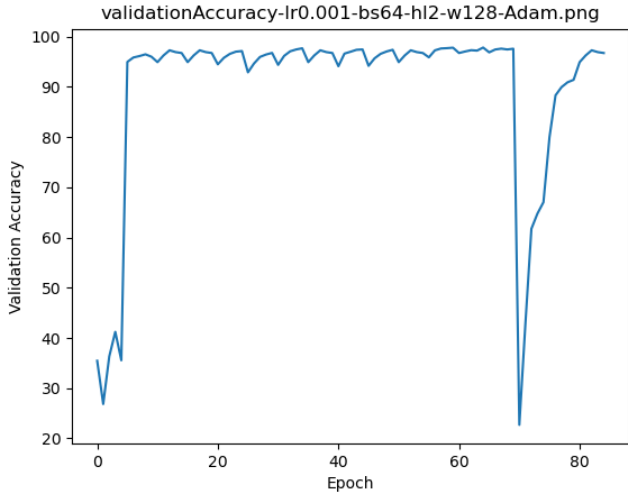


Figure 53: Default settings + width = 128

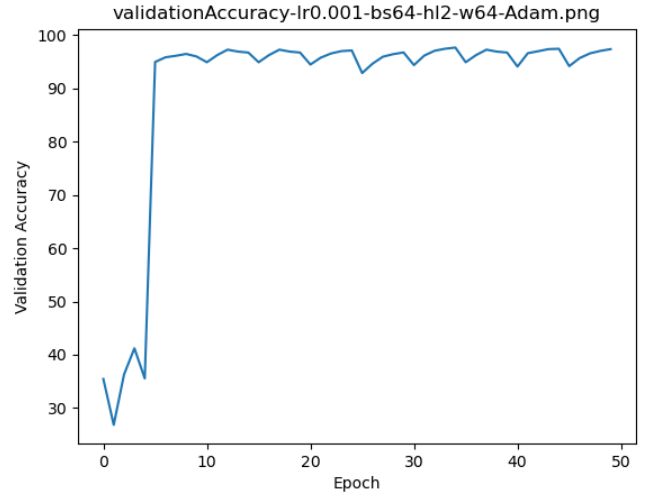


Figure 54: Default settings + width = 64

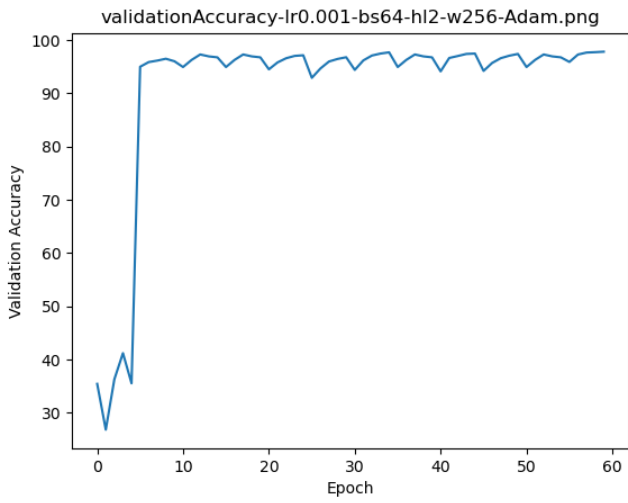


Figure 55: Default settings + width = 256

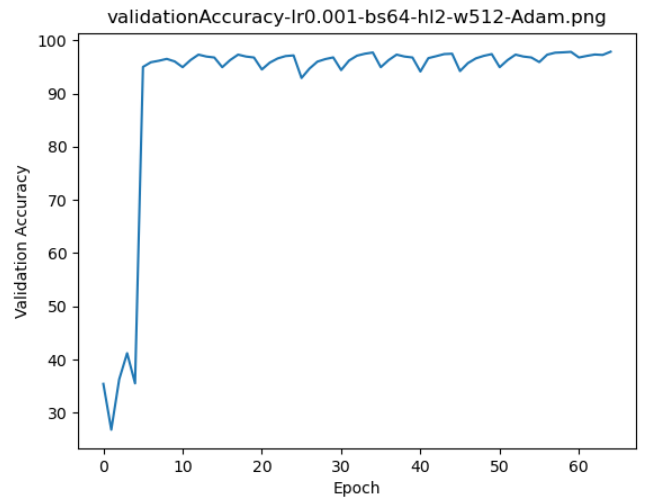


Figure 56: Default settings + width = 512

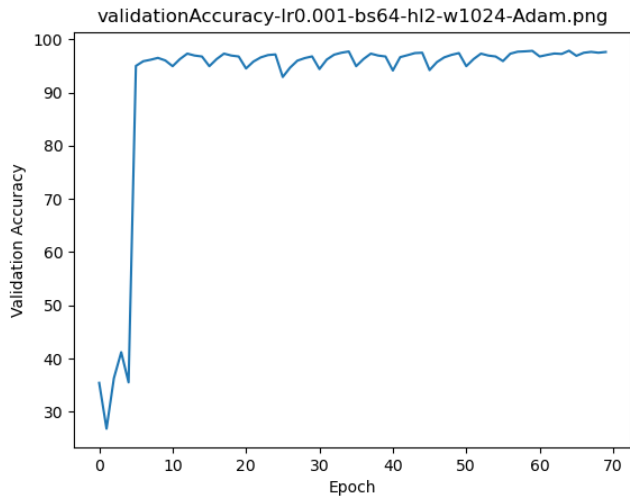


Figure 57: Default settings + width = 1024

3.3.5 Optimizer Type

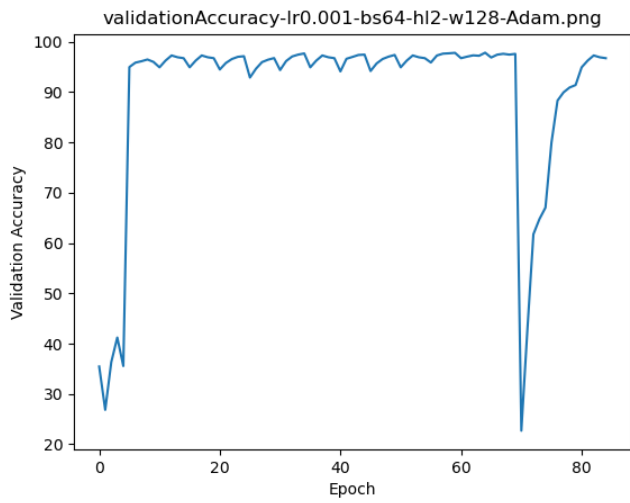


Figure 58: Default settings + Adam optimizer

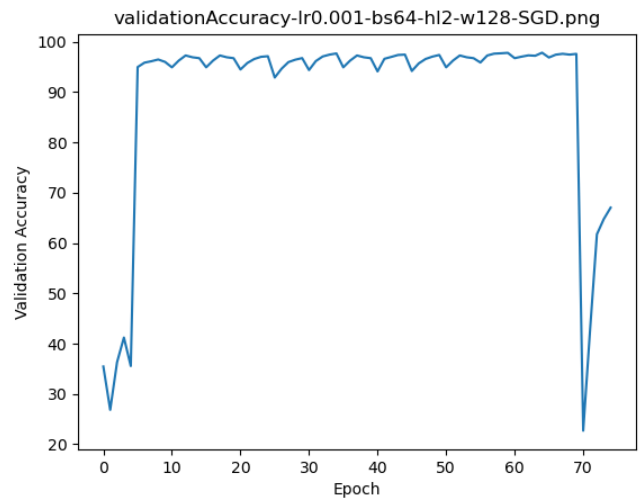


Figure 59: Default settings + learning rate = 0.01

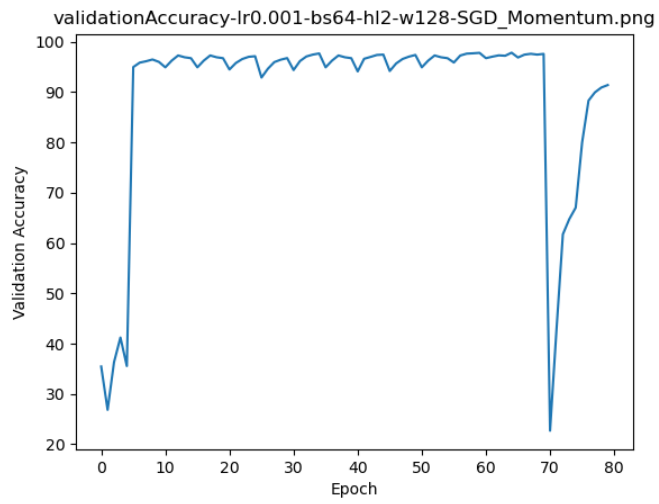


Figure 60: Default settings + SGD_Momentum optimizer