

# Mish

---

A Self Regularized  
Non-Linear Non-  
Monotonous Adaptive  
Neural Activation  
Function.

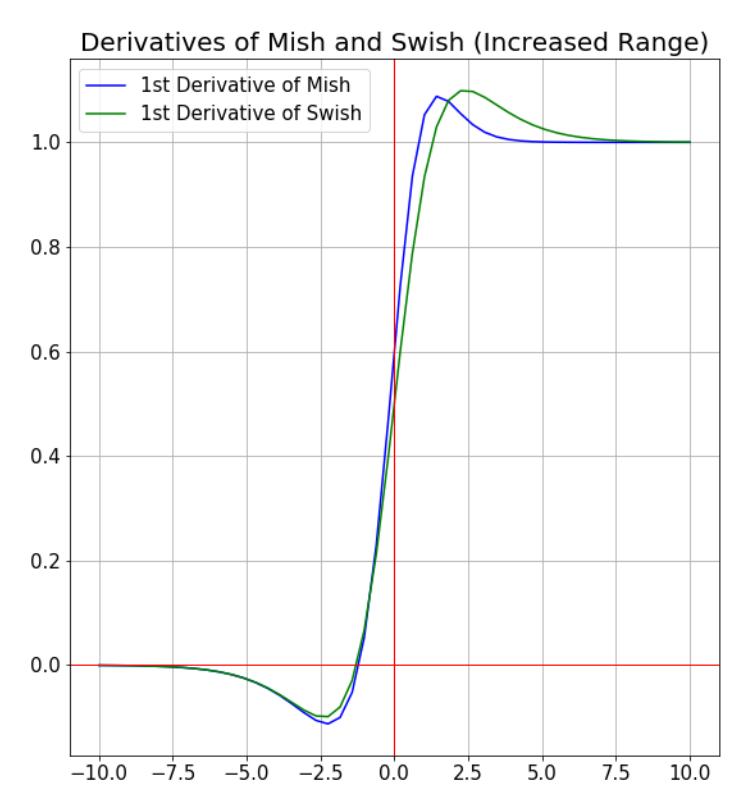
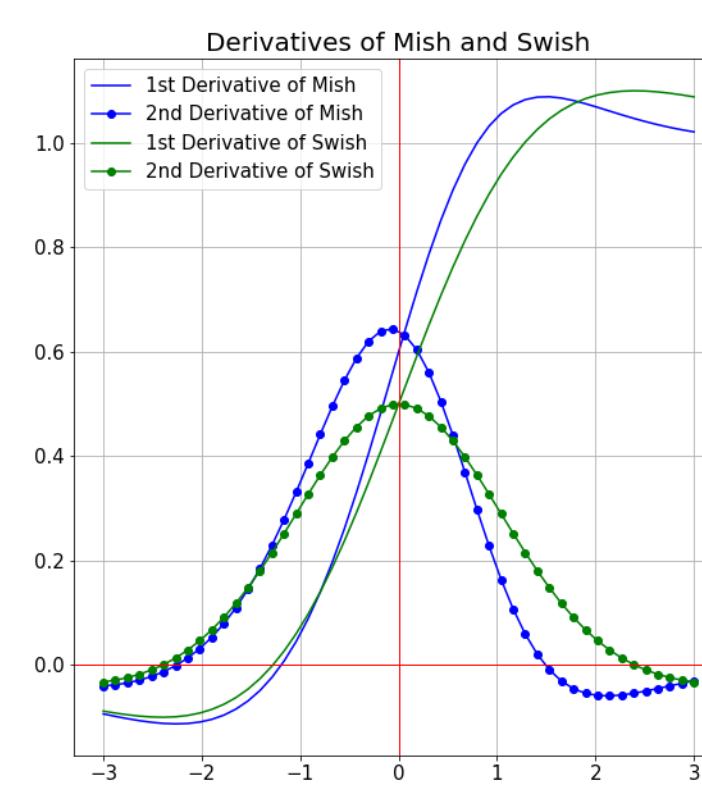
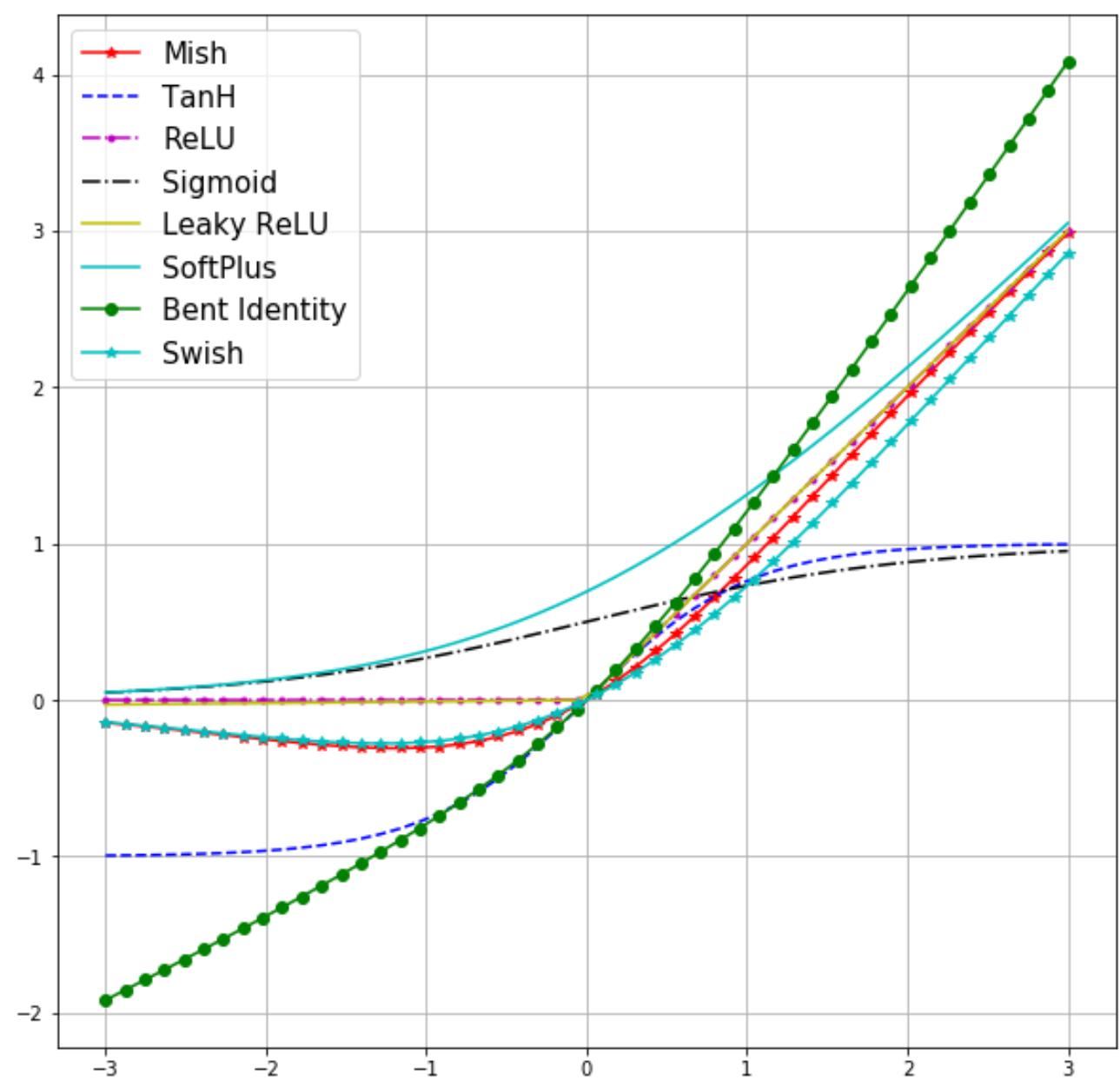
- DIGANTA MISRA



# Introduction

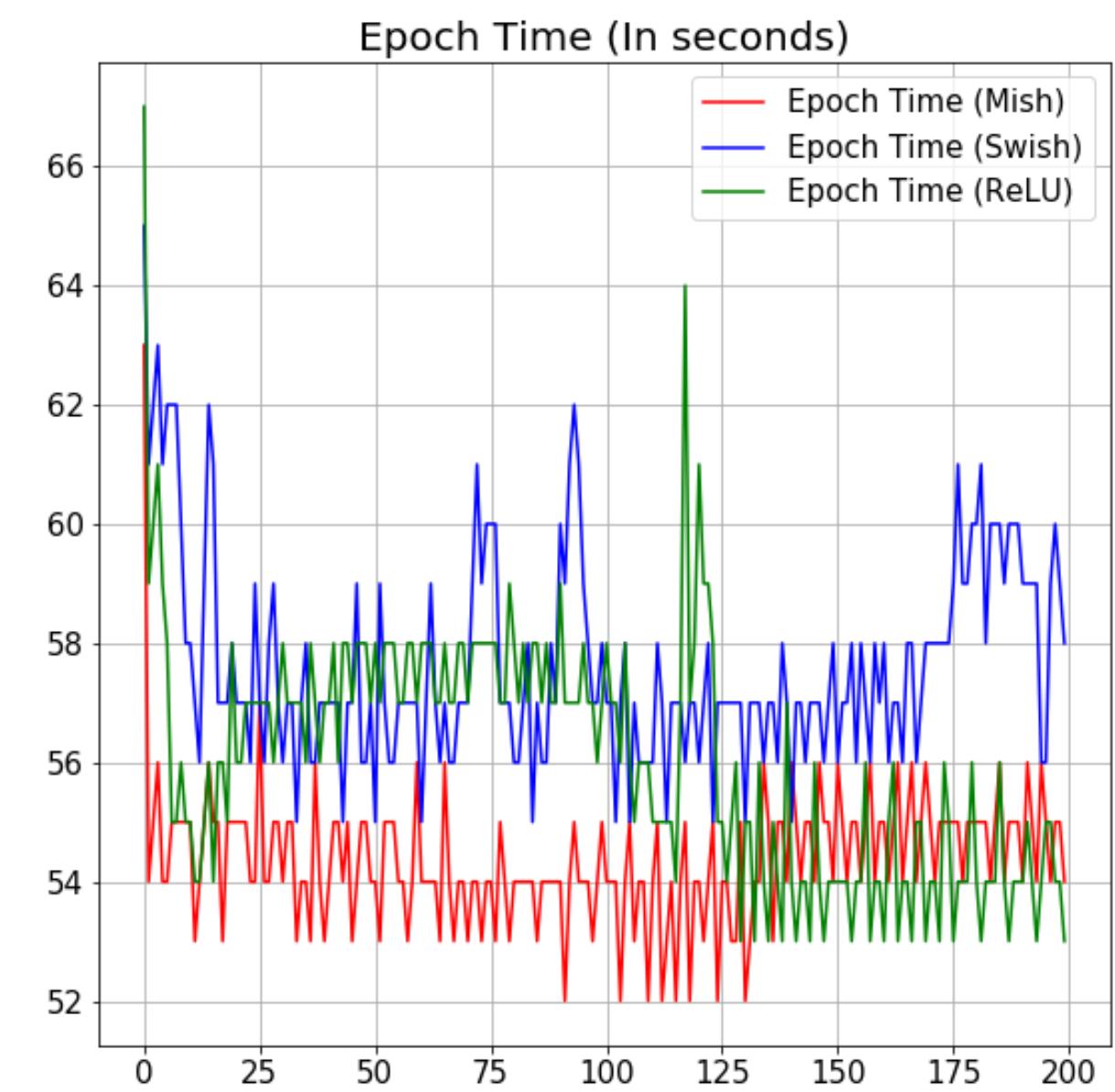
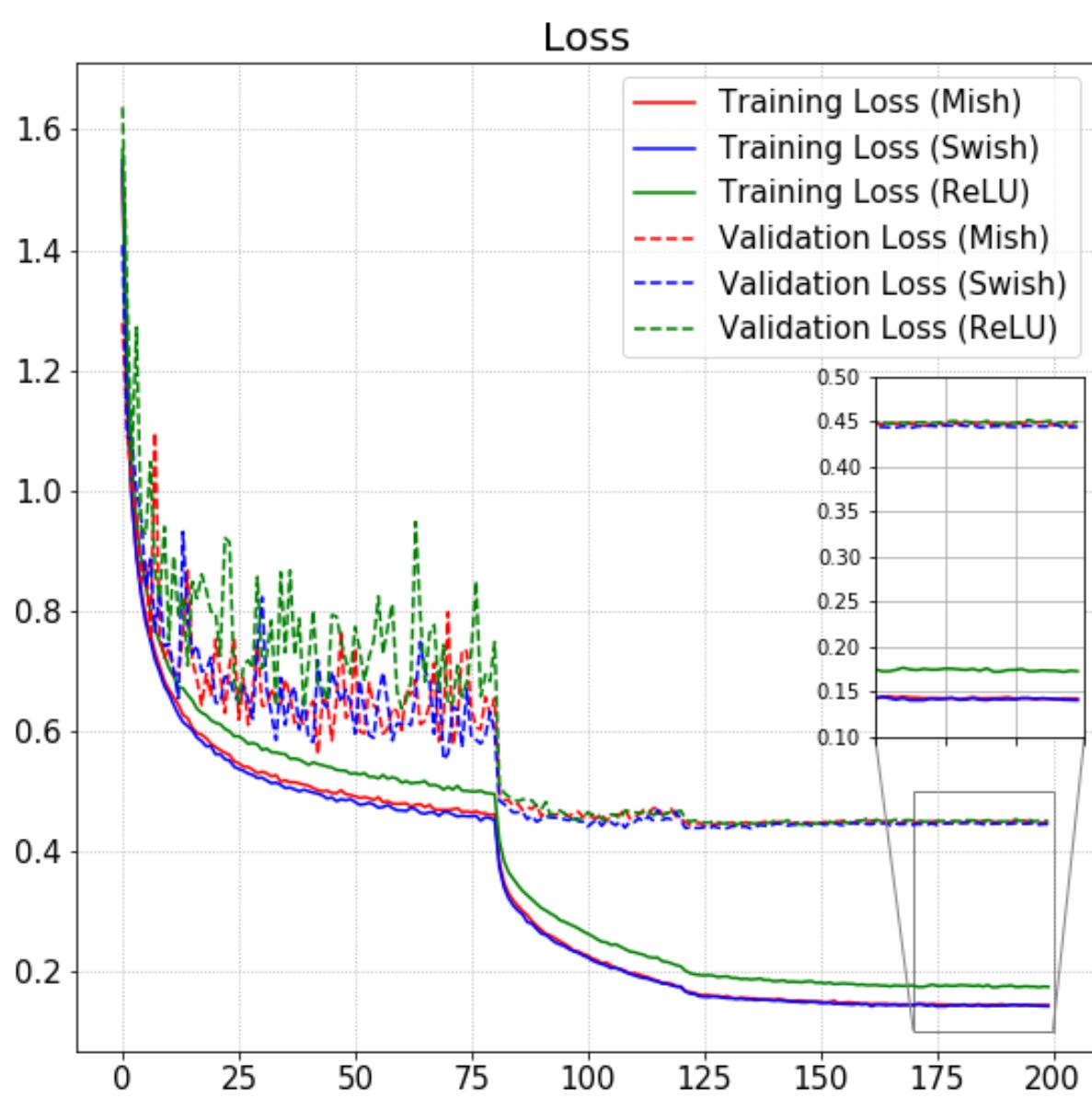
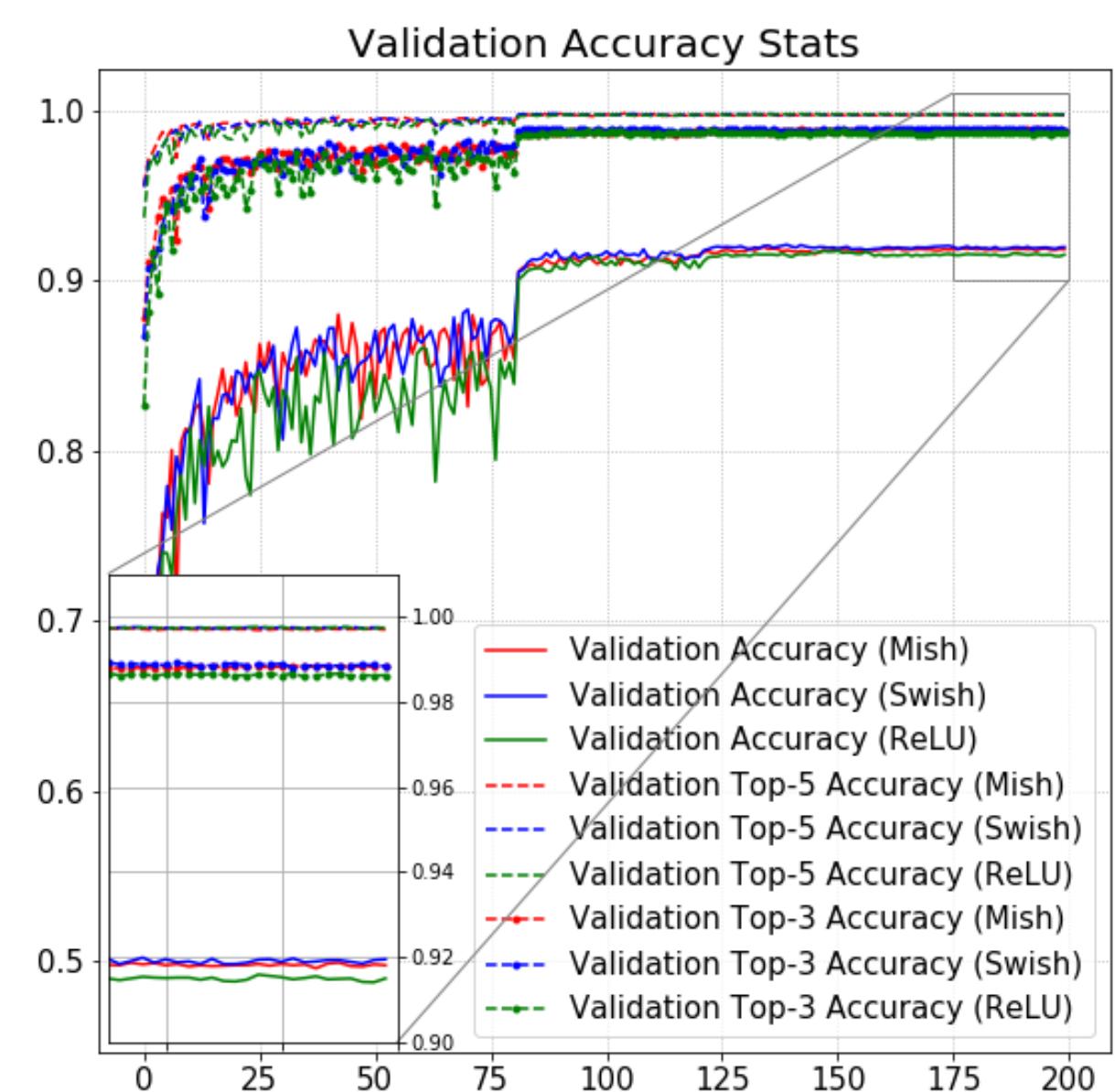
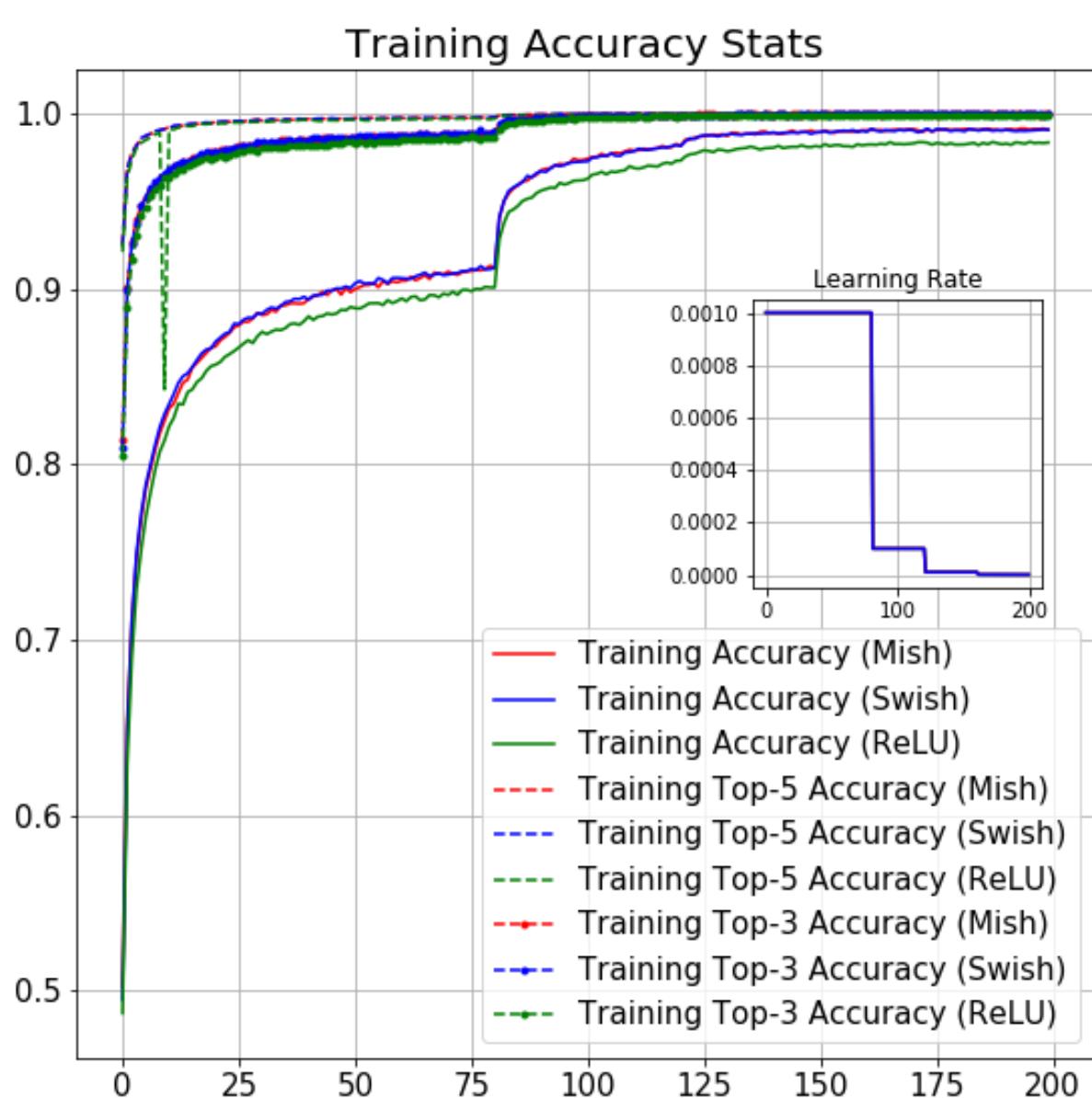
## MOTIVATION

Inspired by Swish  
Activation Function, Mish activation function is a self regularized non-monotonous non-linear adaptive neural activation function. It ranges from approximately -0.31 to infinity. Graphically, it looks nearly the same as Swish, it has a small decay or preserves the gradients in the negative half which results in it having a very sharp global minima.





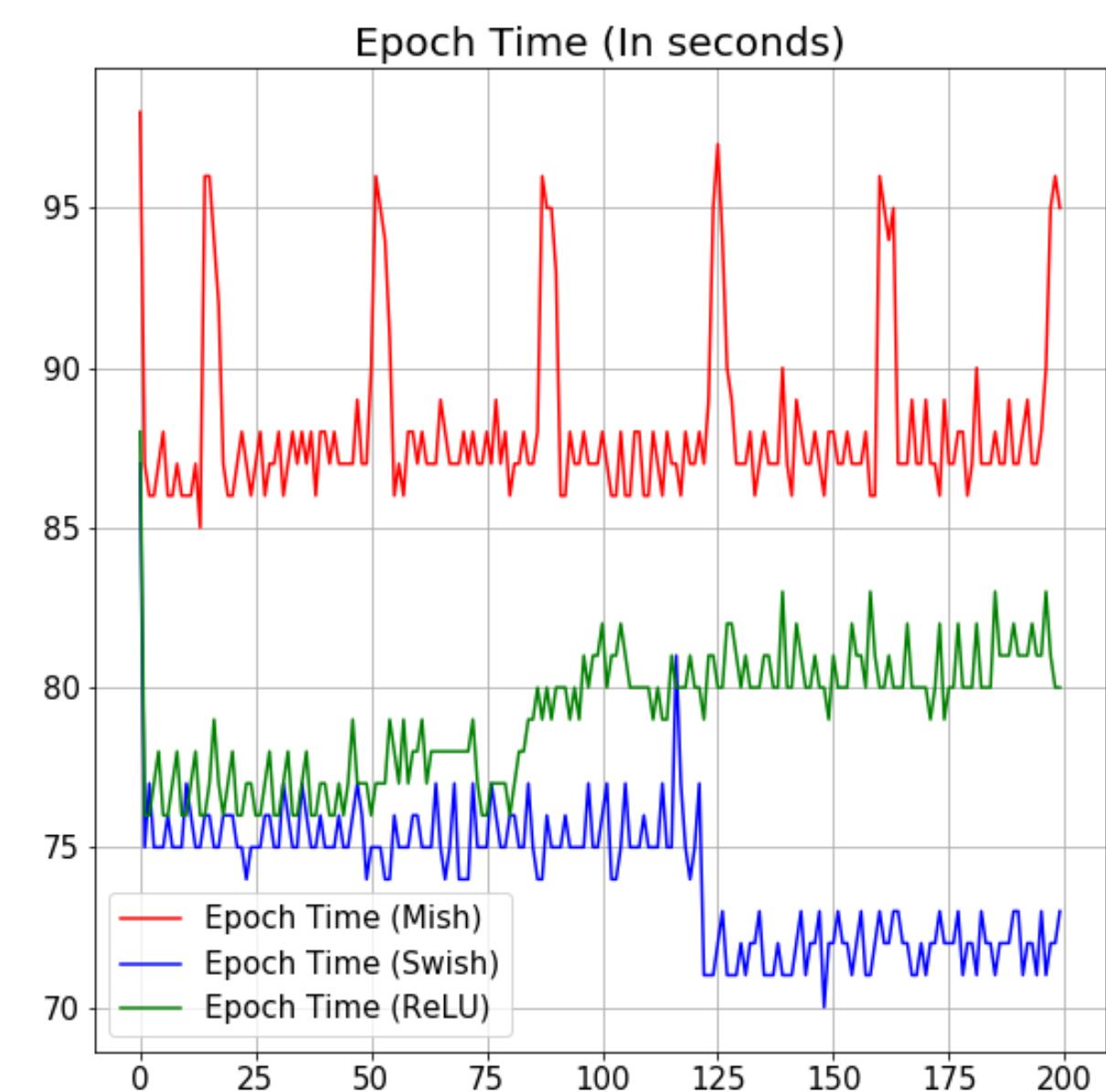
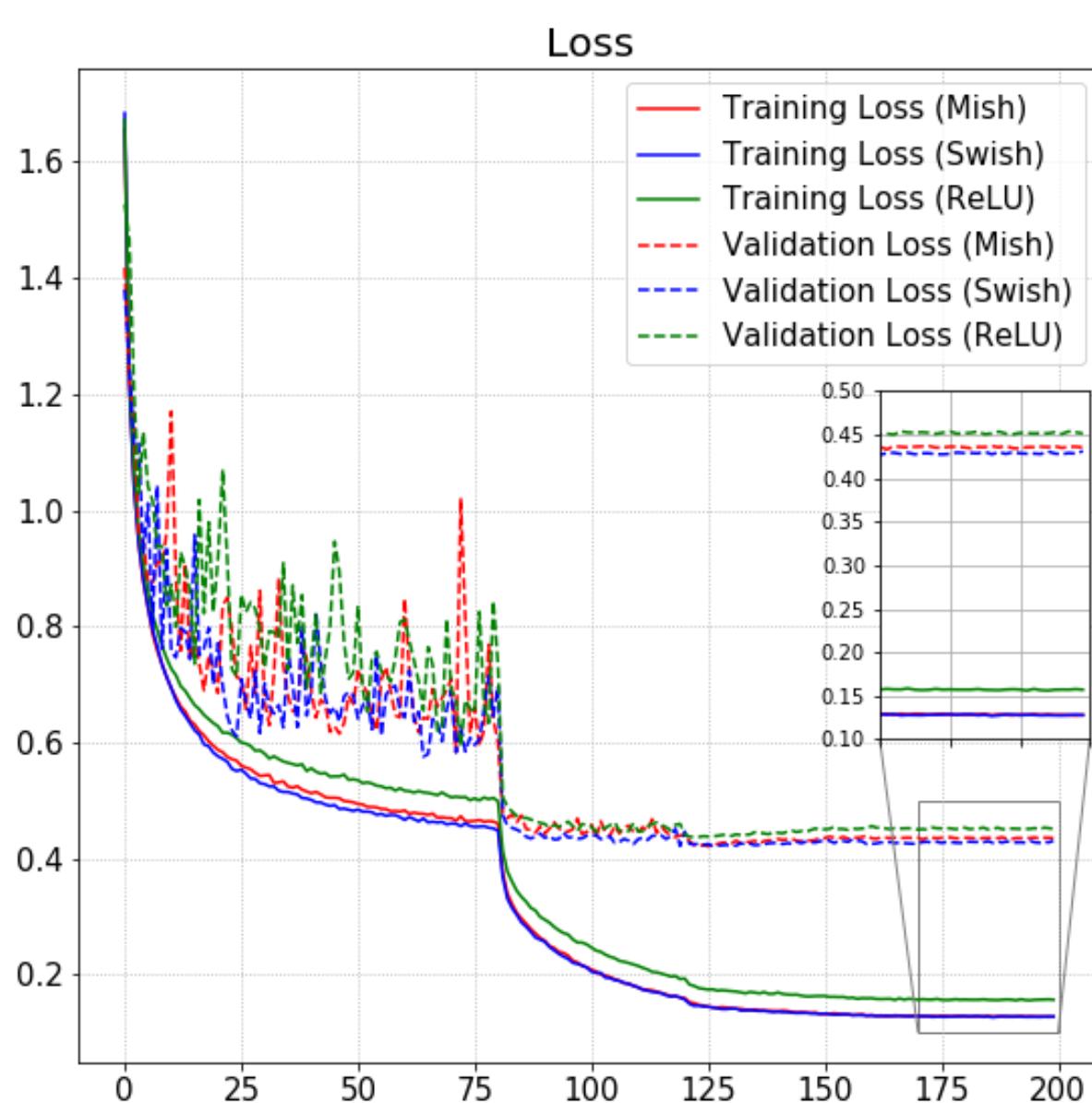
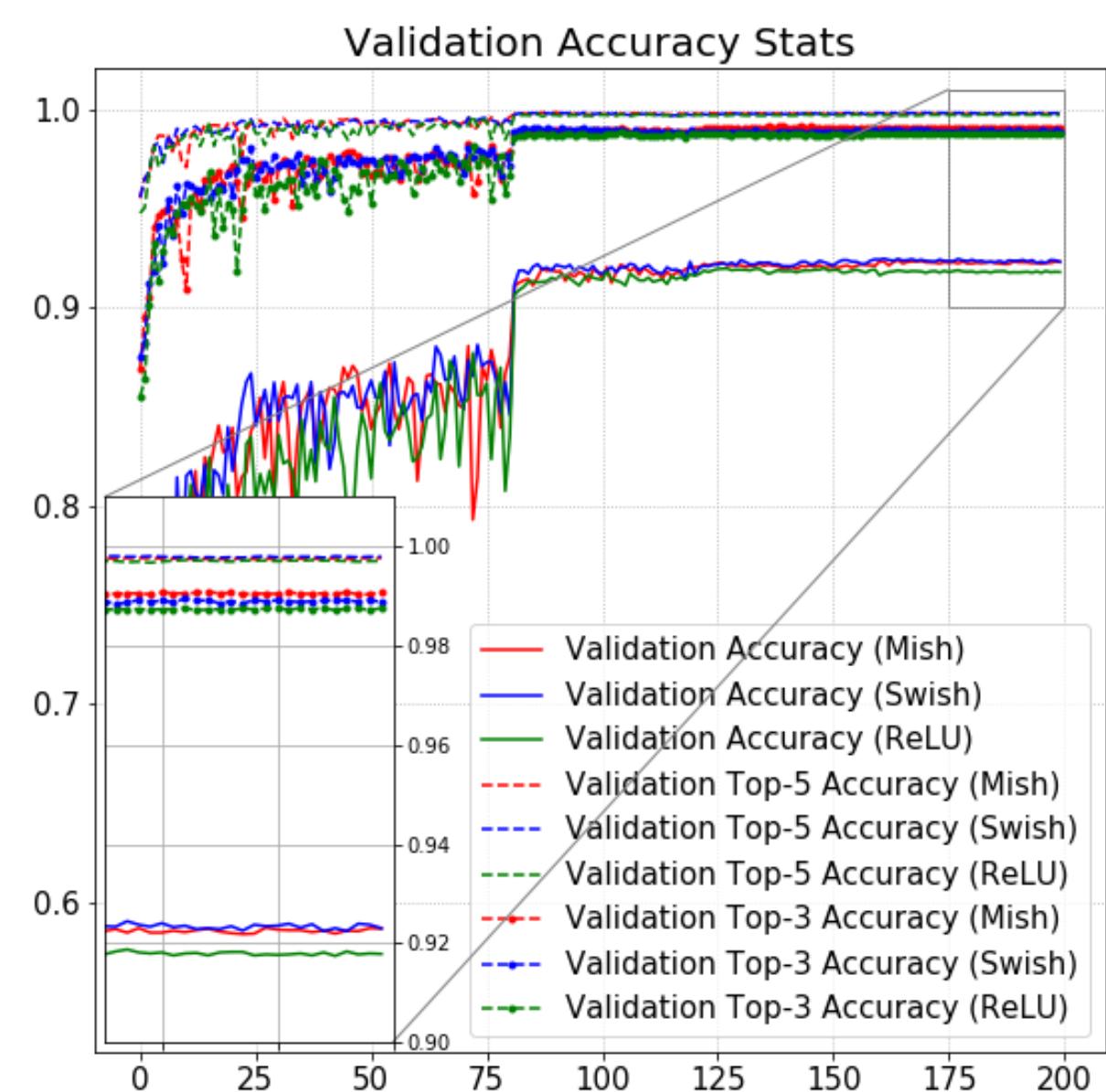
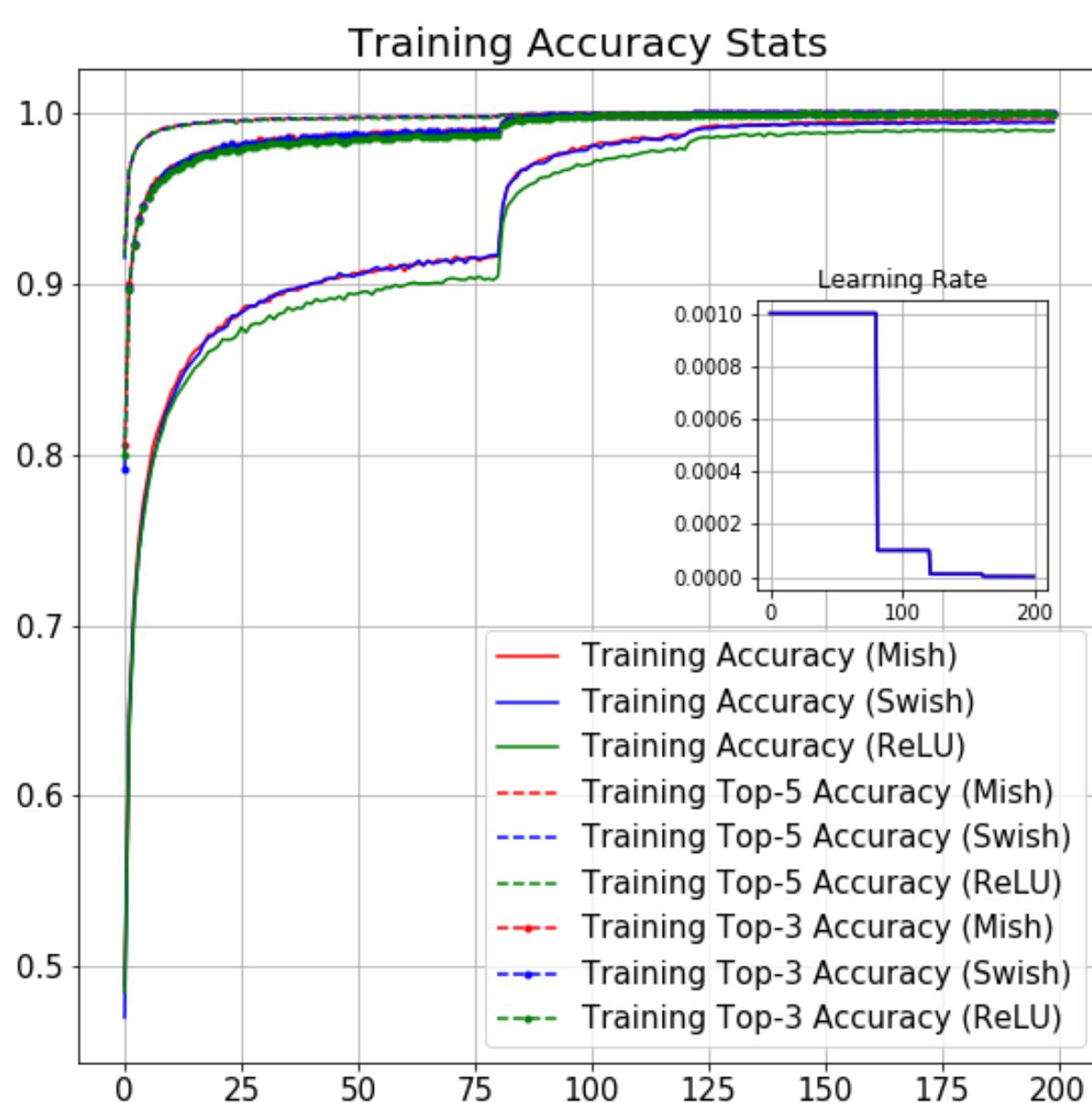
# CIFAR-10



The above result is for ResNet-20 version-1 training on CIFAR-10.



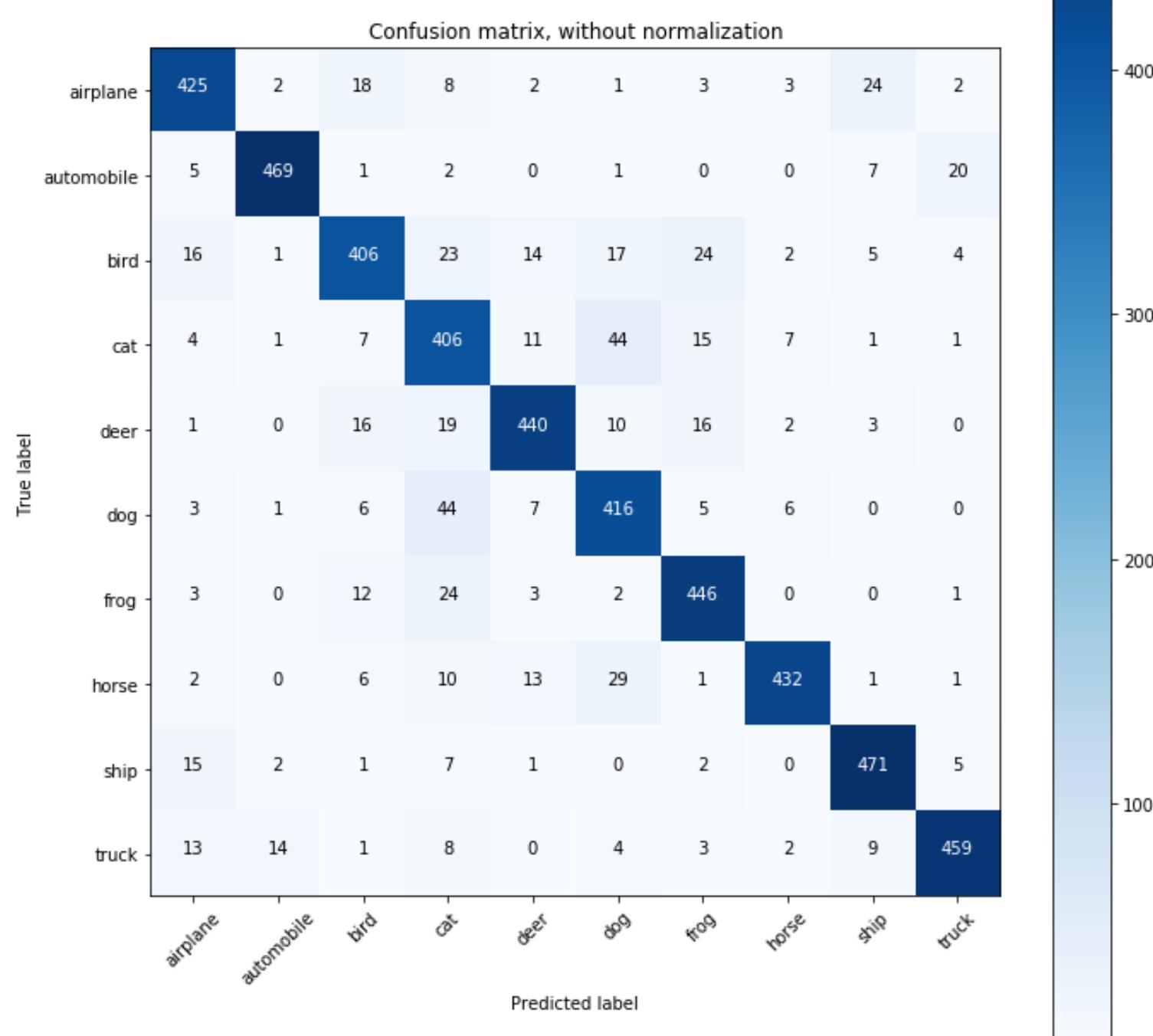
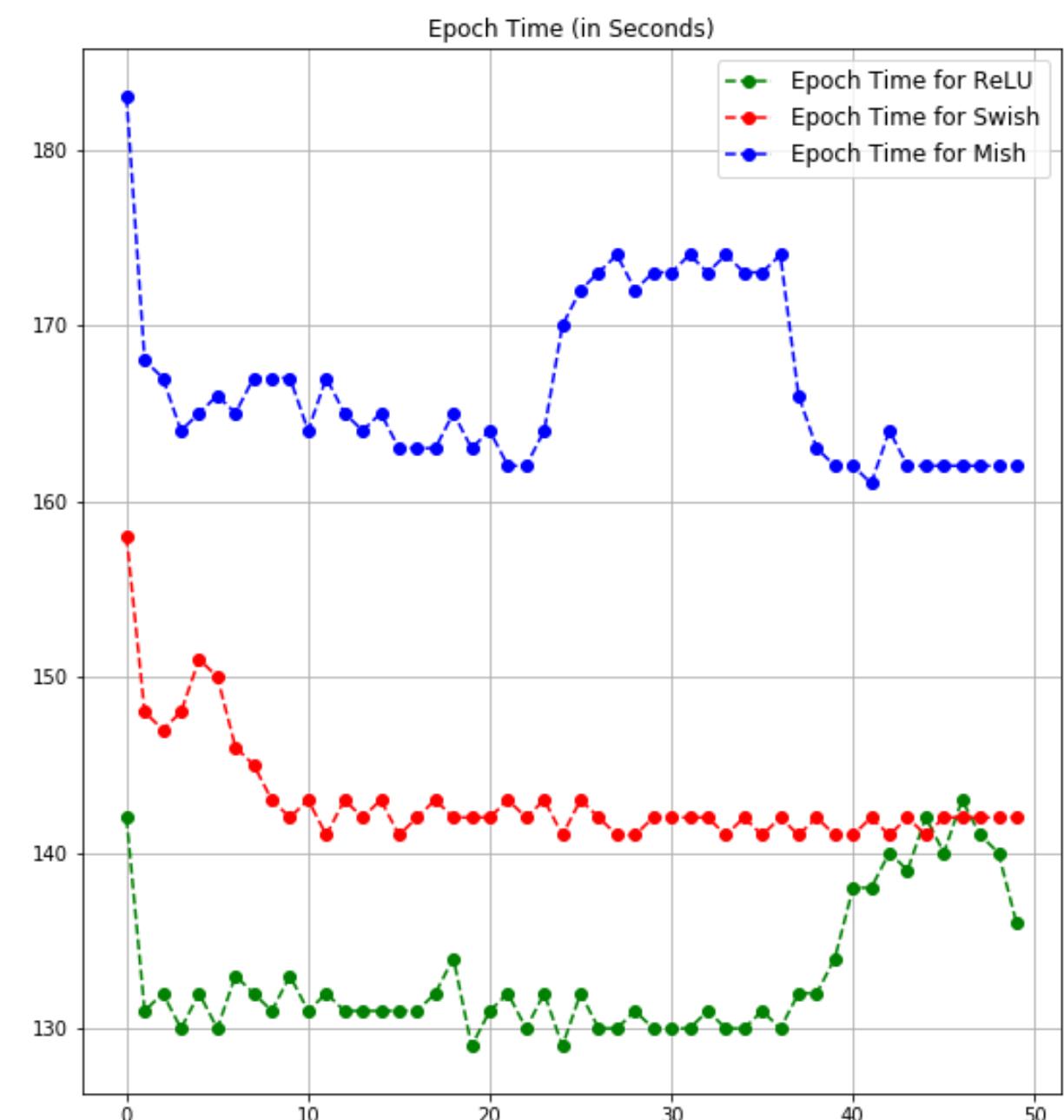
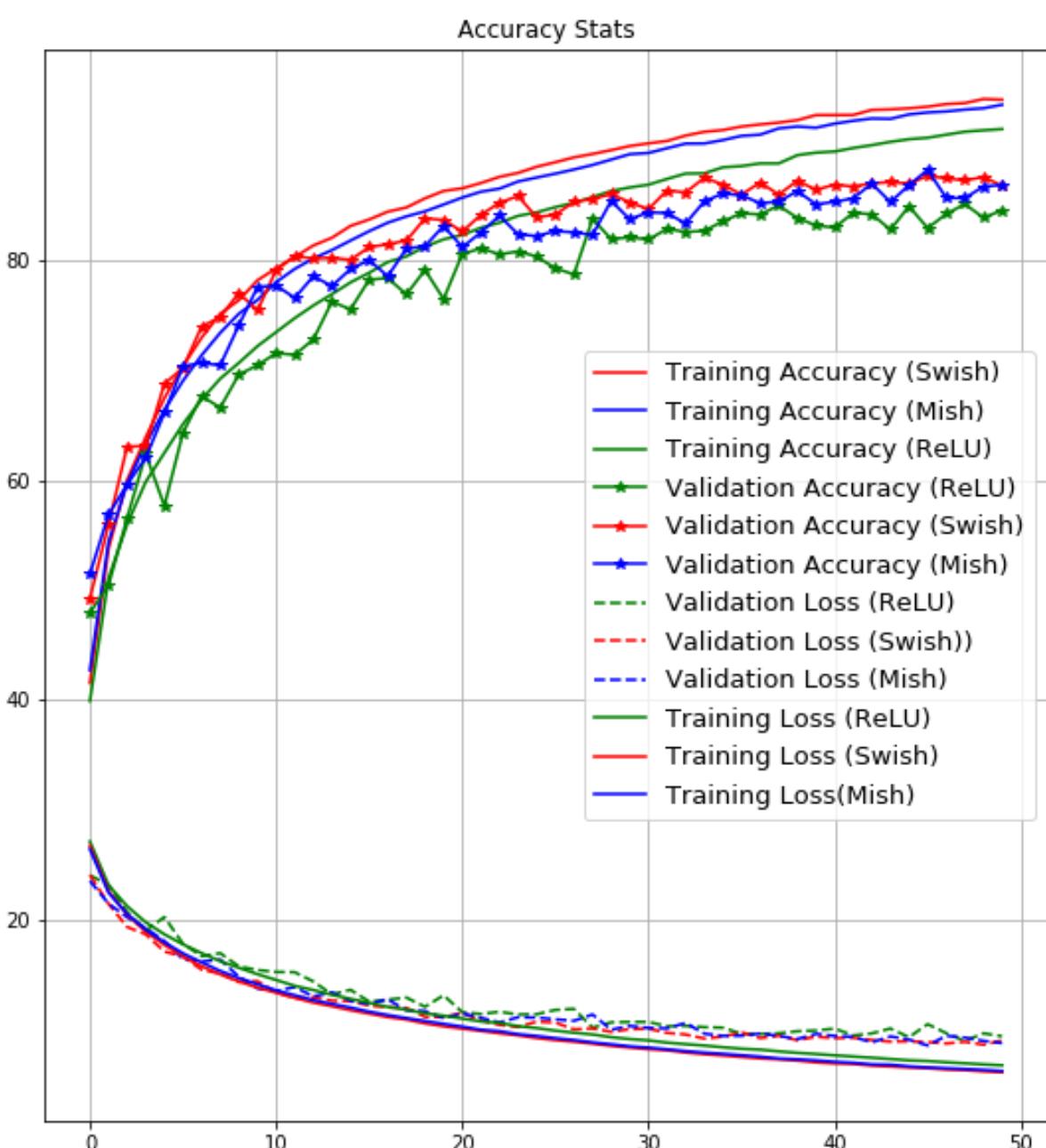
# CIFAR-10



The above result is for ResNet-32 version-1 training on CIFAR-10.



# CIFAR-10



The above result is for ResNet-56 version-2 training on CIFAR-10.



# GANs

Epoch 20

3 4 7 5 3 9 3 5 7 3

Epoch 30

5 4 1 0 3 6 7 0 1 0

Epoch 40

5 3 0 1 9 7 8 3 0 8

Epoch 50

3 3 9 1 1 7 1 3 0 8

Epoch 60

6 1 3 5 2 2 3 1 1 3

Epoch 70

1 1 8 6 4 6 4 1 7 9

Epoch 80

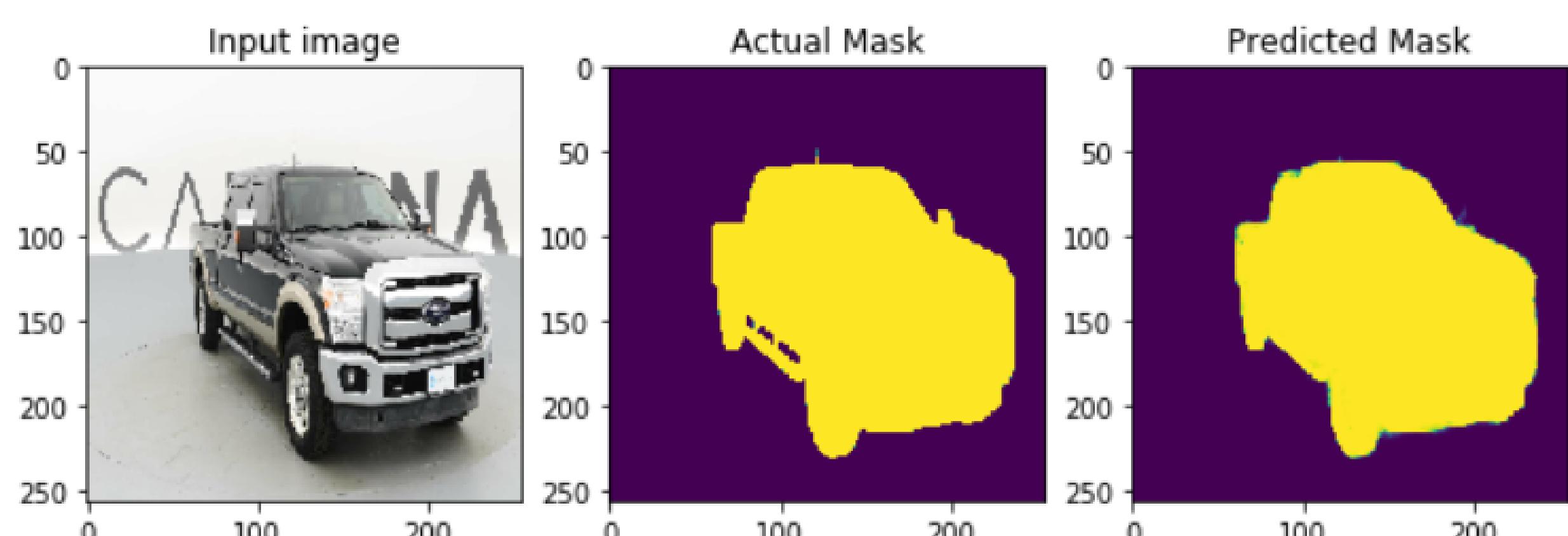
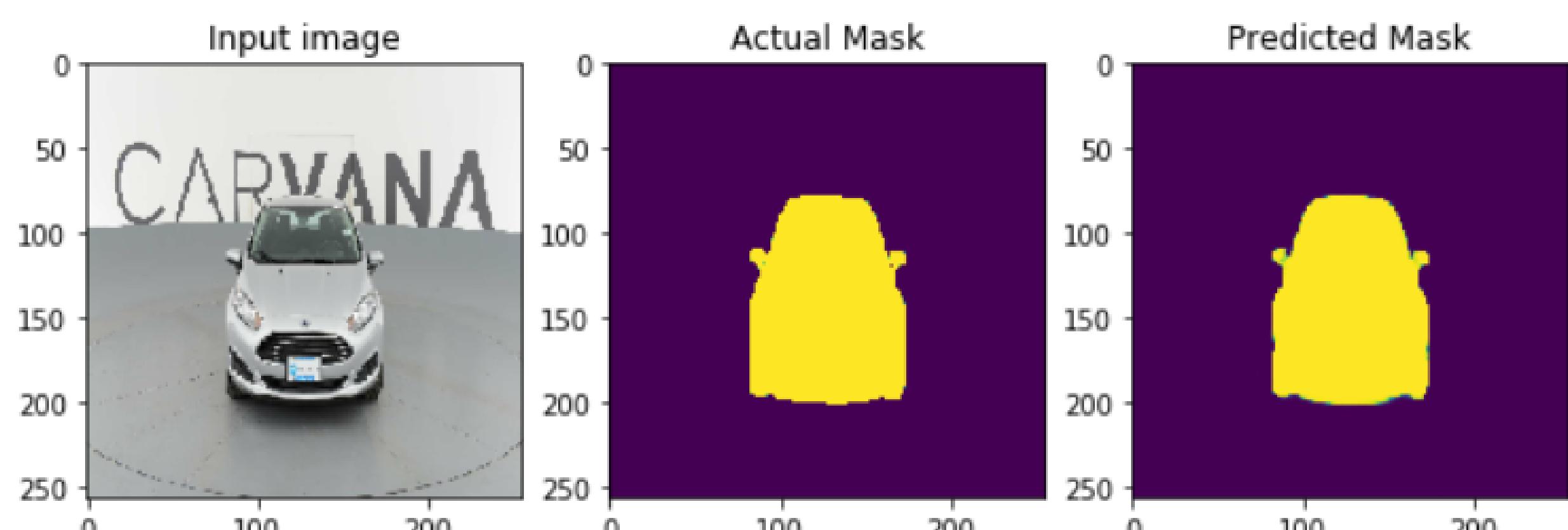
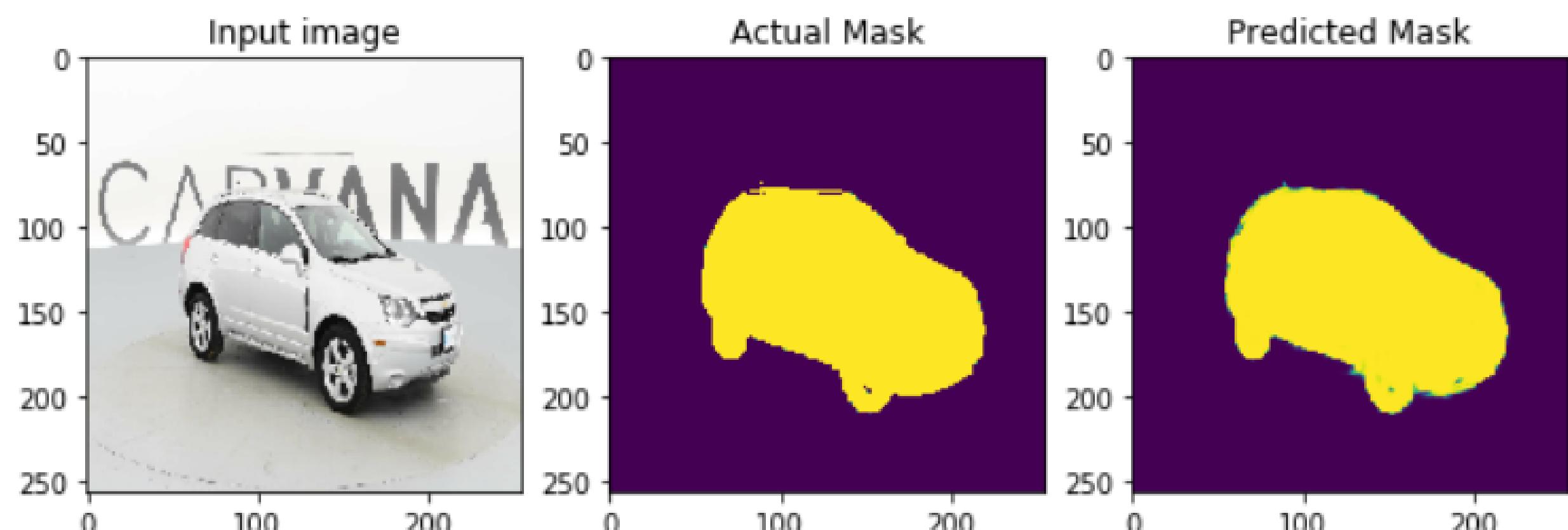
6 1 8 5 0 8 6 7 1 3

Epoch 90

1 1 0 0 8 7 6 7 0 9

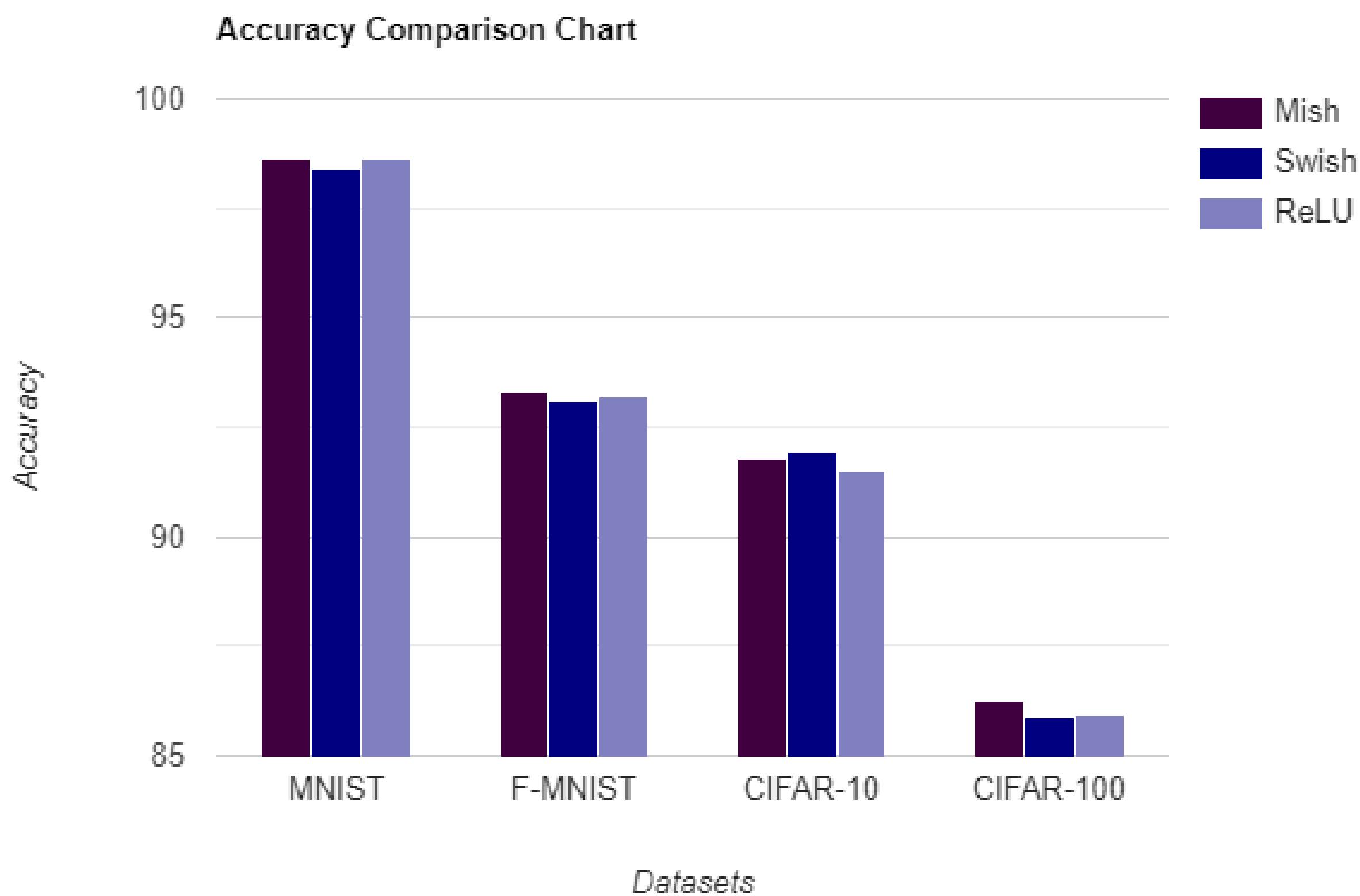


# Mish Object Segmentation





# Results



MNIST was trained on a LeNet-5, F-MNIST was trained on a VGG-Net, CIFAR-10 was trained on ResNet-20 version 1 and CIFAR-100 was trained on ResNet-32 version-1. All the accuracies depicted above are Top-1 Testing Accuracy except for CIFAR-100 which shows the Testing Top-3 Accuracy.



# Thank You!

Additional Details shall be available on [github.com/digantamisra98/Mish](https://github.com/digantamisra98/Mish)