

Report CARN Homework3

Introduction

This report presents my experiments and results for Homework 3 on image classification using CIFAR-100 dataset. All experiments were conducted on custom training pipelines implemented in PyTorch, with additional utilities such as a batch-size scheduler and test-time augmentation (TTA).

Models

Several architectures were tested: ResNet18, ResNet50 and ResNet26D (the largest model and the main focus of my experiments). All models were modified by changing the final layer to output 100 classes, and a dropout layer was added.

Optimization

Optimizers: consisted of SGD, ADAM, ADAMW and SAM. Across experiments SGD provided the best results more consistent but SAM achieved a slight improvement however with more training time.

Scheduler: StepLR or ReduceLROnPlateau. Both were effective.

For **batch size scheduler** I used linearBS from [Git](#). The size ranged between 32-128.

SAM optimizer was imported from [Git](#).

Data Augmentation: RandomCrop, RandomHorizontalFlip, RandAugment, Normalization.

As for efficiency I used AMP training, the batchsize scheduler and a number of 4 data load workers.

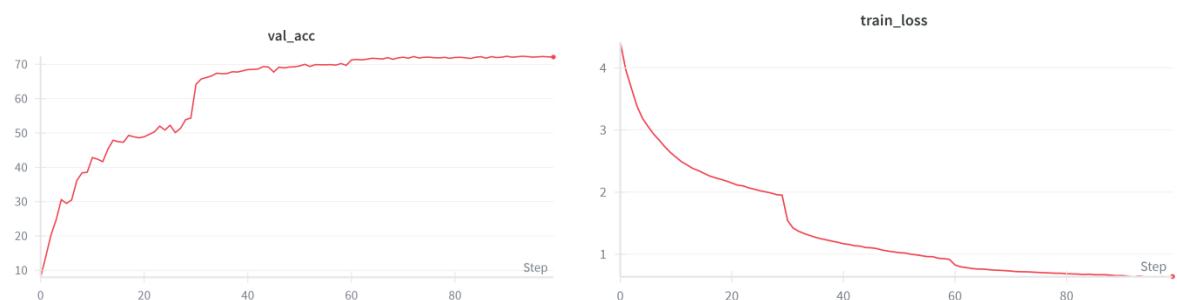
Experimental Results

The configurations that obtained >70. All of them were run on 100 epochs with early stopping if 20 epochs were not improving. Additionally TTA from the previous homework was added.

#	Optimizer	scheduler	Momentum	Lr	Dropout	Time	Acc
1.	SGD	StepLR	0.9	0.01	0	46 min	71.2
2.	SGD	StepLR	0.9	0.01	0.2	40 min	71.33
3.	SGD	StepLR	0.9	0.01	0	50 min	71.54
4.	Sam	StepLR	0.8	0.1	0.1	1h 30min	73.1

5.	SGD	Plateau	0.95	0.1	0.1	57min	73.17
6.	SGD	Plateau	0.8	0.1	0	1h 4 min	73.24
7.	SGD	stepLr	0.9	0.1	0	35 min	73.35
8.	Sam	steplr	0.8	0.1	0	1h 30min	73.44

Best accuracy I've obtained using AdamW has been 67.91 in 39 minutes. Using a lr of 0.01 and ReduceLROnPlateau. Below are attached graphs from the best model and an overall train loss.



For the set up just uploading the notebook and running all of them should be sufficient. Despite my login into wandb the code shall be able to run correctly.

As for an estimation for the grade, I'd assume between (15-18)/25 since I did not achieve a high score with or without a pretrained model and did not improve the efficiency in a very significant matter.