Deep Learning in Practice 2nd Coding Assignment

Zhongke Sun
zhongke.sun@student-cs.fr

March 2, 2025

1 Question 1.1 - Train a fully-supervised baseline on CIFAR-100

I use the ResNet-18 on CIFAR-100 for 20 epochs and log both train and test loss and accuracy.

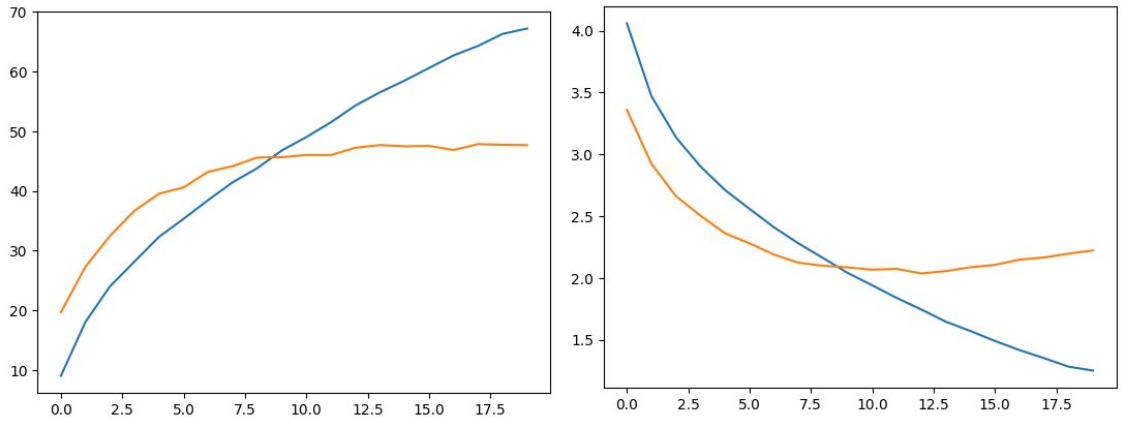


Figure 1: Question 1.1 - Train a fully-supervised baseline on CIFAR-100 Loss & Accuracy Performance

The loss graph shows that both training loss (blue) and test loss (orange) decrease as the number of epochs increases. However, after a certain point, the test loss flattens and even starts to increase slightly, indicating a potential overfitting issue.

The accuracy curve illustrates that both training accuracy (blue) and test accuracy (orange) increase rapidly in the early stages. However, as training progresses (around epoch 8), training accuracy continues to improve quickly, while test accuracy plateaus.

From the training logs, we can observe that over 20 epochs:

- Test loss decreases from 3.359 to 2.036 (minimum).
- Test Acc increases from 19.670% to 47.850% (maximum).

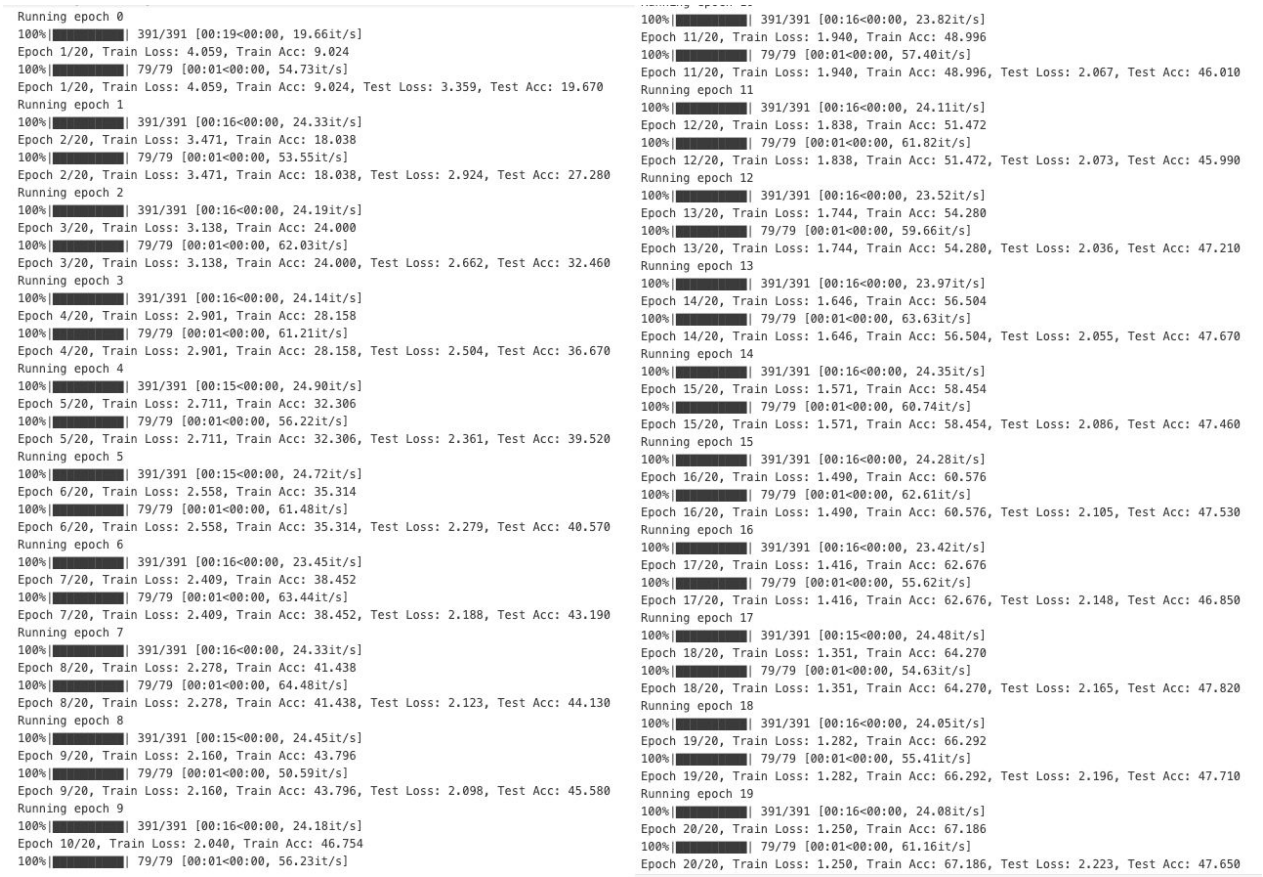


Figure 2: Question 1.1 - Baseline Training Log

2 Question 1.5 - Train the SimCLR model on the CIFAR-100 dataset

Here, I train the SimCLRModel on the CIFAR-100 dataset using the ContrastiveLoss for 35 epochs and log train losses. The model was trained in an unsupervised manner. The key optimization objectives are:

- **Minimize contrastive loss**: Ensure that augmented views of the same image have similar representations while different images have distinct representations.

- **Observe training convergence**: Evaluate whether the model's contrastive loss decreases over time, indicating improved representation learning.

Hyperparameter Setup

- **Loss Function**: Contrastive Loss
- **Optimizer**: Adam with weight decay $1e-5$
- **Batch Size**: 256
- **Learning Rate**: $3e-4$
- **Training Epochs**: 35
- **Temperature for Contrastive Loss**: 0.5

The training logs are as the following:

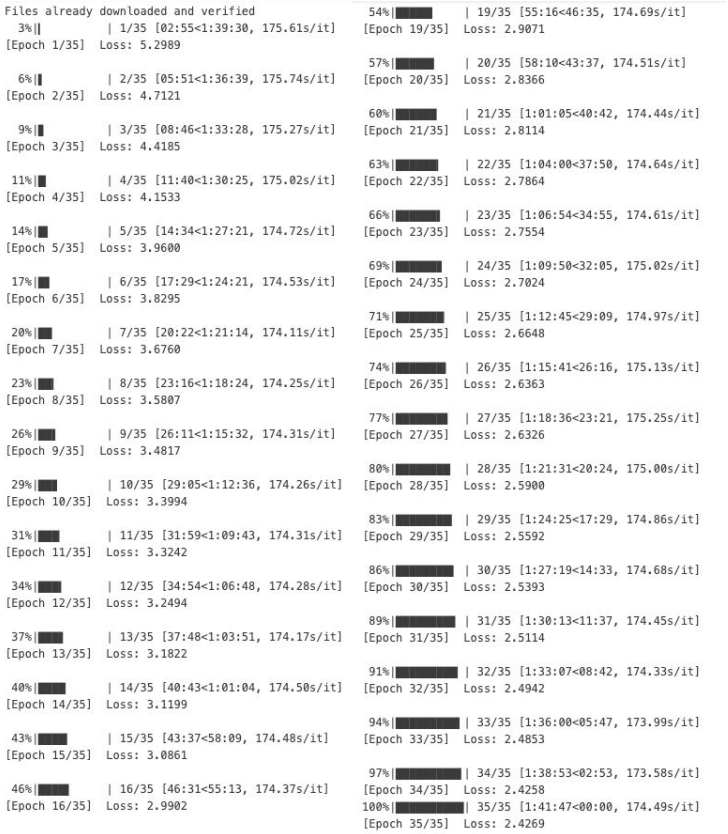


Figure 3: Question 1.5 - Model Training Log

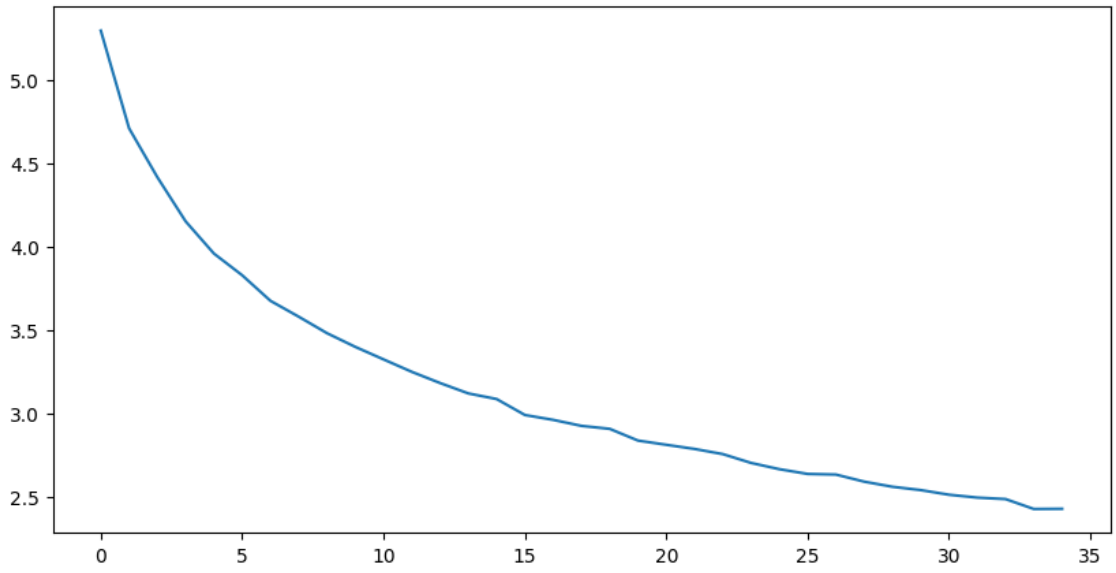


Figure 4: Question 1.5 - Model Loss

The loss starts at 5.2989 in the first epoch and steadily decrease throughout training. By epoch 35, the loss reaches 2.4269, indicating significant learning progress.

The loss curve demonstrates a smooth and consistent decline, settling in to flatten finally, confirming that the model is learning meaningful representations.

3 Question 1.6 - Fine-tune SimCLR model on the downstream task of classification

Using the pre-trained model, I added a linear classifier on top of representation, and trained all on CIFAR-100 for 40 epochs using CrossEntropyLoss and logged both train and test loss and accuracies.

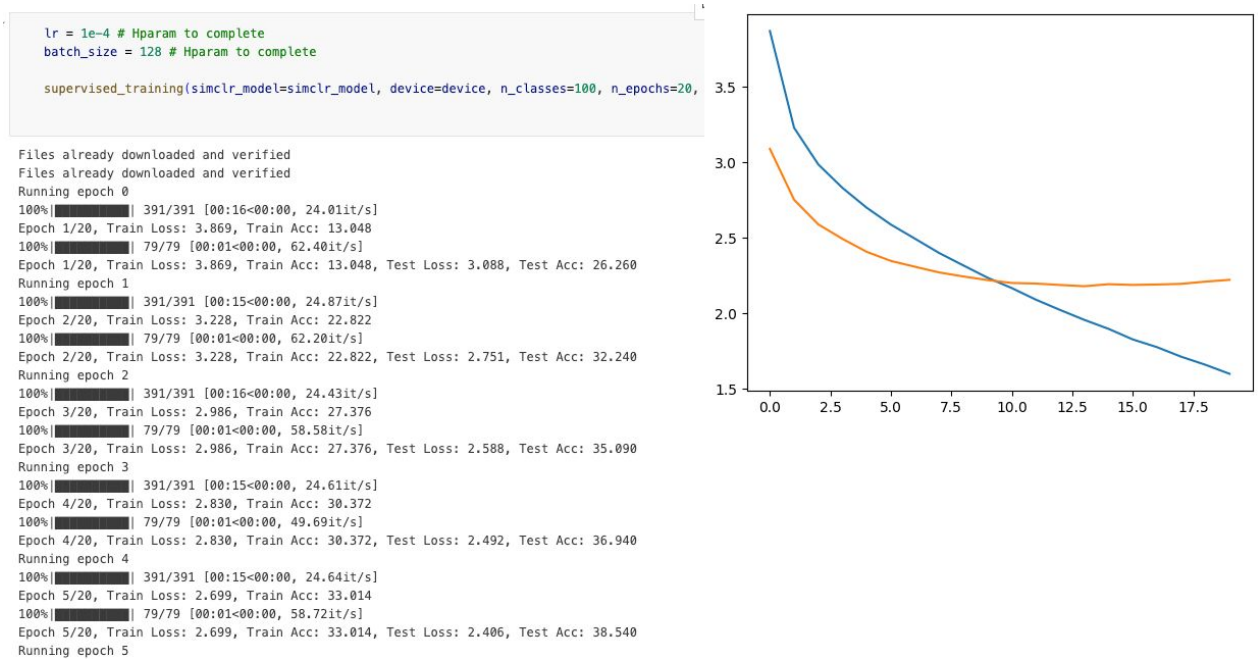
Hyperparameter Setup

- **Loss Function**: CrossEntropyLoss
- **Optimizer**: Adam with weight decay $1e-5$
- **Batch Size**: 128
- **Learning Rate**: $1e-4$
- **Training Epochs**: 20

- **Number of Classes**: 100

The epoch 1 - 20 training outputs are as the followings:

From epoch 1 to epoch 20, the train loss and test loss started at $3.5-4$ (similar to Question 1.1), then decreased rapidly to about 2.1 finally. The test accuracy started from 26.260%, increased rapidly to around 50%.



```
lr = 1e-4 # Hparam to complete
batch_size = 128 # Hparam to complete

supervised_training(simclr_model=simclr_model, device=device, n_classes=100, n_epochs=20,
```

```
Files already downloaded and verified
Files already downloaded and verified
Running epoch 0
100%|██████████| 391/391 [00:16<00:00, 24.01it/s]
Epoch 1/20, Train Loss: 3.869, Train Acc: 13.048
100%|██████████| 79/79 [00:01<00:00, 62.40it/s]
Epoch 1/20, Train Loss: 3.869, Train Acc: 13.048, Test Loss: 3.088, Test Acc: 26.260
Running epoch 1
100%|██████████| 391/391 [00:15<00:00, 24.87it/s]
Epoch 2/20, Train Loss: 3.228, Train Acc: 22.822
100%|██████████| 79/79 [00:01<00:00, 62.20it/s]
Epoch 2/20, Train Loss: 3.228, Train Acc: 22.822, Test Loss: 2.751, Test Acc: 32.240
Running epoch 2
100%|██████████| 391/391 [00:16<00:00, 24.43it/s]
Epoch 3/20, Train Loss: 2.986, Train Acc: 27.376
100%|██████████| 79/79 [00:01<00:00, 58.58it/s]
Epoch 3/20, Train Loss: 2.986, Train Acc: 27.376, Test Loss: 2.588, Test Acc: 35.090
Running epoch 3
100%|██████████| 391/391 [00:15<00:00, 24.61it/s]
Epoch 4/20, Train Loss: 2.830, Train Acc: 30.372
100%|██████████| 79/79 [00:01<00:00, 49.69it/s]
Epoch 4/20, Train Loss: 2.830, Train Acc: 30.372, Test Loss: 2.492, Test Acc: 36.940
Running epoch 4
100%|██████████| 391/391 [00:15<00:00, 24.64it/s]
Epoch 5/20, Train Loss: 2.699, Train Acc: 33.014
100%|██████████| 79/79 [00:01<00:00, 58.72it/s]
Epoch 5/20, Train Loss: 2.699, Train Acc: 33.014, Test Loss: 2.406, Test Acc: 38.540
Running epoch 5
```


Figure 5: Question 1.6 - Model Running Logs & Loss Curve **Epoch 1 - Epoch 20**

(**NOTE: I re-run my classification model after the first 20 epochs for better convergence, so total training epochs is 40 epochs. The graphs below are showing the outputs of 21-40 epochs**).

From epoch 21 to epoch 40, the train loss starts at 2.183 and rapidly decreases to 0.652 by epoch 40, showing effective learning. The test loss initially fluctuates around 2.0000. The train accuracy starts at 56.0% and rises quickly to 83.644% at epoch 40. The test accuracy remains around 51-52%, without significant improvement.

The below model loss and accuracy curve show the performance of the SimCLR classification model:

```
lr = 1e-4 # Hparam to complete
batch_size = 128 # Hparam to complete

supervised_training(simclr_model=simclr_model, device=device, n_classes=100, n_epochs=20,
```

```
Files already downloaded and verified
Files already downloaded and verified
Running epoch 0
100%|██████████| 391/391 [00:15<00:00, 24.64it/s]
Epoch 1/20, Train Loss: 2.183, Train Acc: 56.000
100%|██████████| 79/79 [00:01<00:00, 52.67it/s]
Epoch 1/20, Train Loss: 2.183, Train Acc: 56.000, Test Loss: 1.942, Test Acc: 50.550
Running epoch 1
100%|██████████| 391/391 [00:15<00:00, 24.85it/s]
Epoch 2/20, Train Loss: 1.088, Train Acc: 75.424
100%|██████████| 79/79 [00:01<00:00, 59.54it/s]
Epoch 2/20, Train Loss: 1.088, Train Acc: 75.424, Test Loss: 1.914, Test Acc: 51.320
Running epoch 2
100%|██████████| 391/391 [00:15<00:00, 24.79it/s]
Epoch 3/20, Train Loss: 0.948, Train Acc: 77.356
100%|██████████| 79/79 [00:01<00:00, 53.27it/s]
Epoch 3/20, Train Loss: 0.948, Train Acc: 77.356, Test Loss: 1.964, Test Acc: 51.640
Running epoch 3
100%|██████████| 391/391 [00:15<00:00, 24.92it/s]
Epoch 4/20, Train Loss: 0.868, Train Acc: 78.790
100%|██████████| 79/79 [00:01<00:00, 58.80it/s]
Epoch 4/20, Train Loss: 0.868, Train Acc: 78.790, Test Loss: 1.992, Test Acc: 51.420
Running epoch 4
100%|██████████| 391/391 [00:16<00:00, 24.41it/s]
Epoch 5/20, Train Loss: 0.823, Train Acc: 79.736
100%|██████████| 79/79 [00:01<00:00, 56.44it/s]
Epoch 5/20, Train Loss: 0.823, Train Acc: 79.736, Test Loss: 2.023, Test Acc: 51.670
Running epoch 5
100%|██████████| 391/391 [00:15<00:00, 24.91it/s]
Epoch 6/20, Train Loss: 0.797, Train Acc: 80.302
100%|██████████| 79/79 [00:01<00:00, 56.30it/s]
Epoch 6/20, Train Loss: 0.797, Train Acc: 80.302, Test Loss: 2.045, Test Acc: 51.180
Running epoch 6
100%|██████████| 391/391 [00:15<00:00, 24.51it/s]
Epoch 7/20, Train Loss: 0.776, Train Acc: 80.730
100%|██████████| 79/79 [00:01<00:00, 62.05it/s]
Epoch 7/20, Train Loss: 0.776, Train Acc: 80.730, Test Loss: 2.073, Test Acc: 51.680
Running epoch 7
100%|██████████| 391/391 [00:15<00:00, 24.67it/s]
Epoch 8/20, Train Loss: 0.760, Train Acc: 81.126
100%|██████████| 79/79 [00:01<00:00, 61.46it/s]
Epoch 8/20, Train Loss: 0.760, Train Acc: 81.126, Test Loss: 2.090, Test Acc: 51.910
Epoch 11/20, Train Loss: 0.720, Train Acc: 81.976
100%|██████████| 79/79 [00:01<00:00, 54.99it/s]
Epoch 11/20, Train Loss: 0.720, Train Acc: 81.976, Test Loss: 2.133, Test Acc: 51.310
Running epoch 11
100%|██████████| 391/391 [00:15<00:00, 24.84it/s]
Epoch 12/20, Train Loss: 0.707, Train Acc: 82.328
100%|██████████| 79/79 [00:01<00:00, 57.64it/s]
Epoch 12/20, Train Loss: 0.707, Train Acc: 82.328, Test Loss: 2.146, Test Acc: 51.380
Running epoch 12
100%|██████████| 391/391 [00:15<00:00, 24.86it/s]
Epoch 13/20, Train Loss: 0.691, Train Acc: 82.862
100%|██████████| 79/79 [00:01<00:00, 56.63it/s]
Epoch 13/20, Train Loss: 0.691, Train Acc: 82.862, Test Loss: 2.166, Test Acc: 51.490
Running epoch 13
100%|██████████| 391/391 [00:15<00:00, 24.84it/s]
Epoch 14/20, Train Loss: 0.690, Train Acc: 82.830
100%|██████████| 79/79 [00:01<00:00, 62.19it/s]
Epoch 14/20, Train Loss: 0.690, Train Acc: 82.830, Test Loss: 2.172, Test Acc: 50.860
Running epoch 14
100%|██████████| 391/391 [00:15<00:00, 24.85it/s]
Epoch 15/20, Train Loss: 0.690, Train Acc: 82.764
100%|██████████| 79/79 [00:01<00:00, 59.86it/s]
Epoch 15/20, Train Loss: 0.690, Train Acc: 82.764, Test Loss: 2.177, Test Acc: 51.400
Running epoch 15
100%|██████████| 391/391 [00:15<00:00, 24.88it/s]
Epoch 16/20, Train Loss: 0.677, Train Acc: 83.108
100%|██████████| 79/79 [00:01<00:00, 60.45it/s]
Epoch 16/20, Train Loss: 0.677, Train Acc: 83.108, Test Loss: 2.184, Test Acc: 51.610
Running epoch 16
100%|██████████| 391/391 [00:15<00:00, 24.49it/s]
Epoch 17/20, Train Loss: 0.660, Train Acc: 83.324
100%|██████████| 79/79 [00:01<00:00, 60.97it/s]
Epoch 17/20, Train Loss: 0.660, Train Acc: 83.324, Test Loss: 2.203, Test Acc: 51.660
Running epoch 17
100%|██████████| 391/391 [00:15<00:00, 24.82it/s]
Epoch 18/20, Train Loss: 0.667, Train Acc: 83.350
100%|██████████| 79/79 [00:01<00:00, 55.54it/s]
Epoch 18/20, Train Loss: 0.667, Train Acc: 83.350, Test Loss: 2.194, Test Acc: 51.530
Running epoch 18
100%|██████████| 391/391 [00:16<00:00, 24.25it/s]
Epoch 19/20, Train Loss: 0.653, Train Acc: 83.600
100%|██████████| 79/79 [00:01<00:00, 62.05it/s]
Epoch 19/20, Train Loss: 0.653, Train Acc: 83.600, Test Loss: 2.222, Test Acc: 50.940
Running epoch 19
100%|██████████| 391/391 [00:16<00:00, 24.04it/s]
Epoch 20/20, Train Loss: 0.652, Train Acc: 83.644
100%|██████████| 79/79 [00:01<00:00, 56.07it/s]
Epoch 20/20, Train Loss: 0.652, Train Acc: 83.644, Test Loss: 2.210, Test Acc: 51.220
```

Figure 6: Question 1.6 - Model Training Logs **Epoch 21 - Epoch 40**

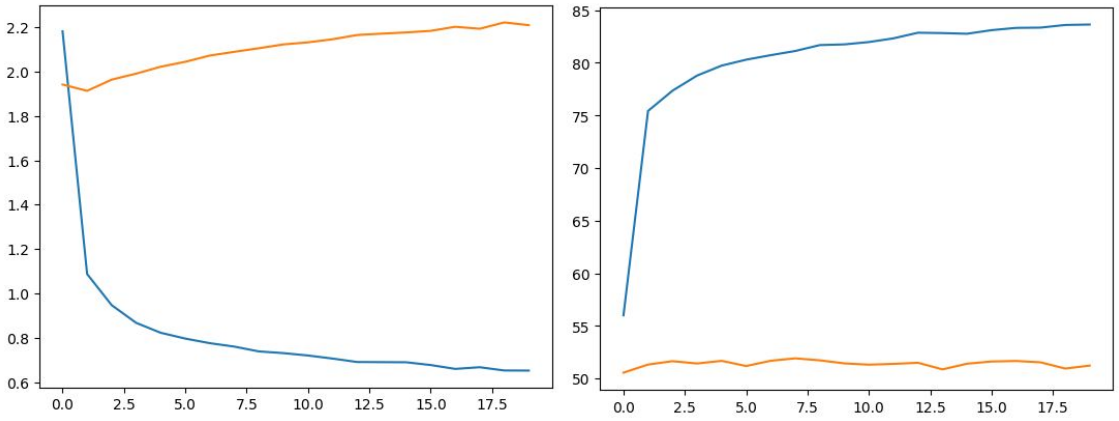


Figure 7: Question 1.6 - Model Loss and Accuracy Curve **Epoch 21 - Epoch 40**

4 Observation by comparing the result from Question 1.1 and Question 1.6

4.1 Faster Convergence in Fine-tuning

In Question 1.1, training ResNet-18 from scratch required a longer time to learn representations, as the model had to learn all features of the dataset. In Question 1.6, pre-trained SimCLR features

allowed the model to converge much faster, as the model already had meaningful representations from contrastive learning.

This highlights the benefit of contrastive pretraining, which captures useful feature embeddings without labeled supervision.

4.2 Better Test Accuracy in Fine-tuning

The final test accuracy in Question 1.6 (51.660%) is **3.84% higher** than Question 1.1 (47.820%) on the same dataset, confirming that pretraining with SimCLR improves generalization. The pre-trained features help the classifier learn more effectively compared to training from scratch, leading to better test performance.