CSCI3230 / ESTR3108 2023-24 First Term Assignment 3

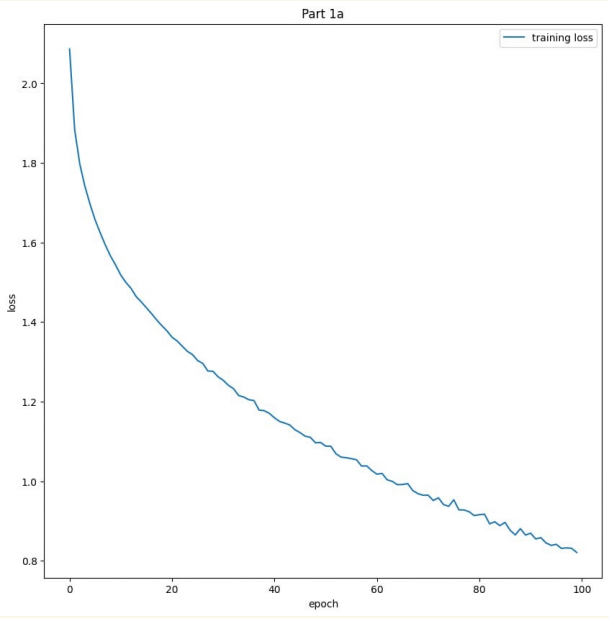
I declare that the assignment here submitted is original except for source material explicitly acknowledged, and that the same or closely related material has not been previously submitted for another course. I also acknowledge that I am aware of University policy and regulations on honesty in academic work, and of the disciplinary guidelines and procedures applicable to breaches of such policy and regulations, as contained in the following websites.

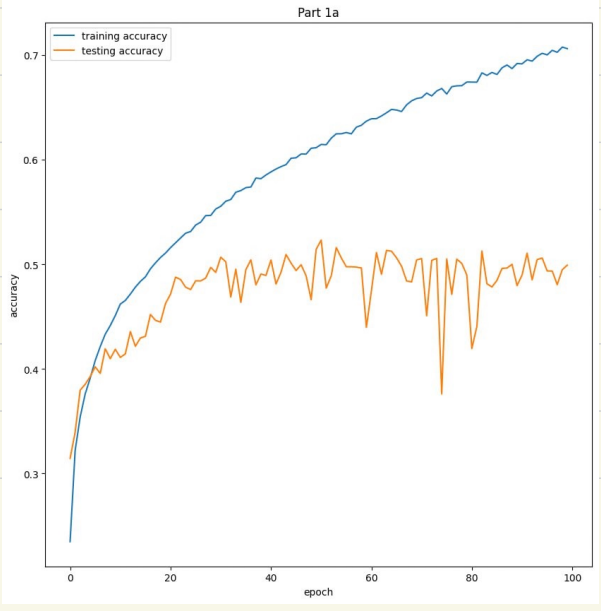
University Guideline on Academic Honesty:
http://www.cuhk.edu.hk/policy/academichonesty/
Faculty of Engineering Guidelines to Academic Honesty:
http://www.erg.cuhk.edu.hk/erg-intra/upload/documents/ENGG_Discipline.pdf

Student Name: Chow Man Fung
Student ID : 1155176920

Part 1

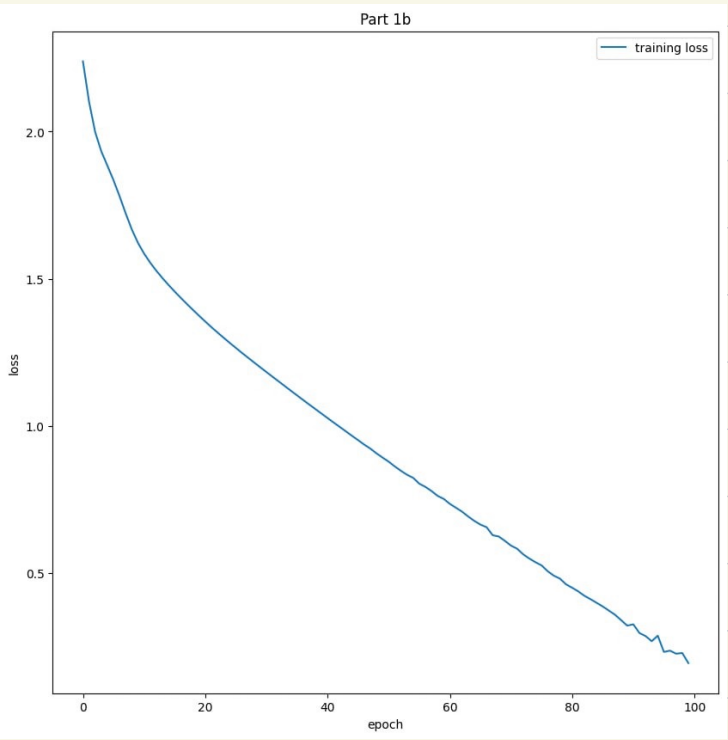
1a) The best test accuracy achieved is 49.9%

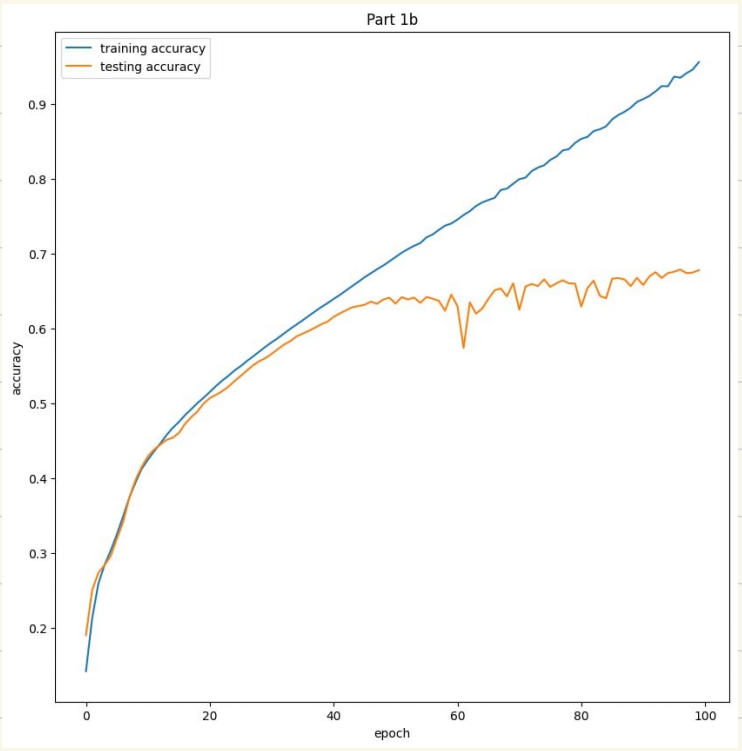





```
100%|████████████| 196/196 [00:08<00:00, 24.33it/s]
Train accuracy: 69.2%, Avg loss: 0.869278
100%|████████████| 40/40 [00:01<00:00, 30.65it/s]
Test accuracy: 49.0%, Avg loss: 1.667663

============== Epoch 92 ==============
100%|████████████| 196/196 [00:07<00:00, 27.19it/s]
Train accuracy: 69.6%, Avg loss: 0.854831
100%|████████████| 40/40 [00:01<00:00, 21.24it/s]
Test accuracy: 51.1%, Avg loss: 1.638719

============== Epoch 93 ==============
100%|████████████| 196/196 [00:06<00:00, 28.23it/s]
Train accuracy: 69.5%, Avg loss: 0.858155
100%|████████████| 40/40 [00:01<00:00, 29.12it/s]
Test accuracy: 48.5%, Avg loss: 1.664658

============== Epoch 94 ==============
100%|████████████| 196/196 [00:08<00:00, 24.42it/s]
Train accuracy: 69.9%, Avg loss: 0.844934
100%|████████████| 40/40 [00:01<00:00, 30.51it/s]
Test accuracy: 50.5%, Avg loss: 1.655292

============== Epoch 95 ==============
100%|████████████| 196/196 [00:06<00:00, 28.04it/s]
Train accuracy: 70.2%, Avg loss: 0.838753
100%|████████████| 40/40 [00:01<00:00, 21.65it/s]
Test accuracy: 50.6%, Avg loss: 1.635126

============== Epoch 96 ==============
100%|████████████| 196/196 [00:07<00:00, 26.68it/s]
Train accuracy: 70.0%, Avg loss: 0.841420
100%|████████████| 40/40 [00:01<00:00, 29.75it/s]
Test accuracy: 49.4%, Avg loss: 1.745748

============== Epoch 97 ==============
100%|████████████| 196/196 [00:07<00:00, 25.22it/s]
Train accuracy: 70.5%, Avg loss: 0.831084
100%|████████████| 40/40 [00:01<00:00, 29.67it/s]
Test accuracy: 49.4%, Avg loss: 1.694757

============== Epoch 98 ==============
100%|████████████| 196/196 [00:06<00:00, 28.93it/s]
Train accuracy: 70.3%, Avg loss: 0.832480
100%|████████████| 40/40 [00:01<00:00, 22.36it/s]
Test accuracy: 48.1%, Avg loss: 1.703168

============== Epoch 99 ==============
100%|████████████| 196/196 [00:07<00:00, 26.14it/s]
Train accuracy: 70.8%, Avg loss: 0.831347
100%|████████████| 40/40 [00:01<00:00, 29.50it/s]
Test accuracy: 49.5%, Avg loss: 1.713408

============== Epoch 100 ==============
100%|████████████| 196/196 [00:07<00:00, 24.90it/s]
Train accuracy: 70.6%, Avg loss: 0.820803
100%|████████████| 40/40 [00:01<00:00, 29.73it/s] Test accuracy: 49.9%, Avg loss: 1.681959
```

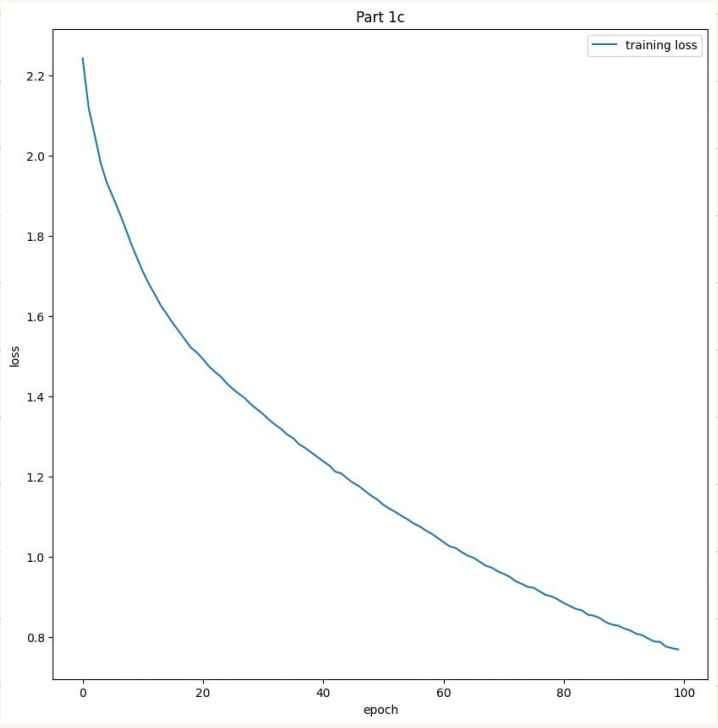
1b) The best test accuracy achieved is 67.8%

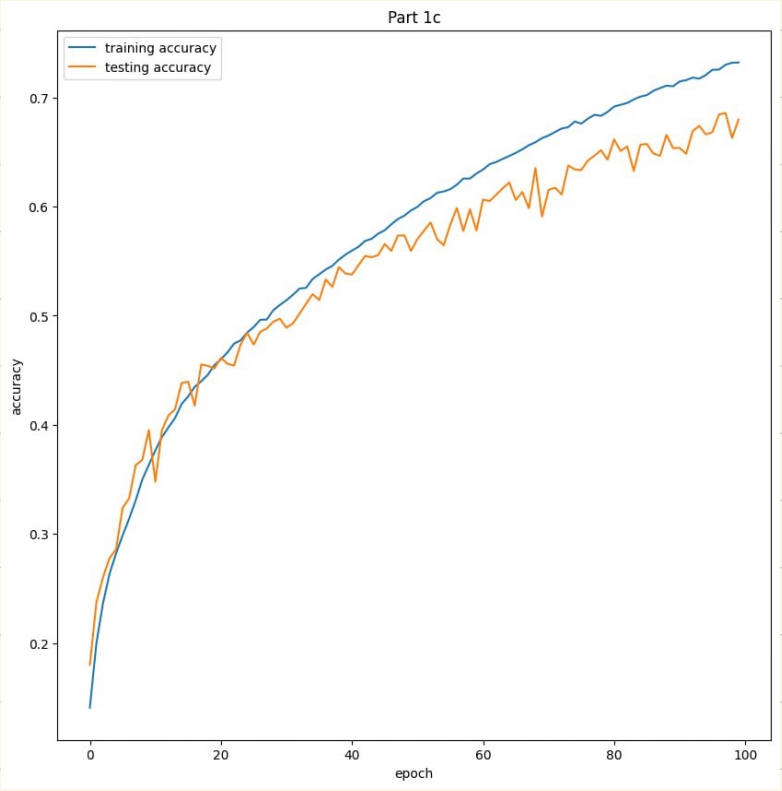





```
100%|████████████████| 196/196 [00:12<00:00, 15.62it/s]
 Train accuracy: 90.7%, Avg loss: 0.324675
100%|████████████████| 40/40 [00:02<00:00, 17.62it/s]
 Test accuracy: 65.9%, Avg loss: 1.191339

 ============== Epoch 92 ==============
100%|████████████████| 196/196 [00:12<00:00, 15.88it/s]
 Train accuracy: 91.1%, Avg loss: 0.295075
100%|████████████████| 40/40 [00:02<00:00, 17.91it/s]
 Test accuracy: 67.0%, Avg loss: 1.139066

 ============== Epoch 93 ==============
100%|████████████████| 196/196 [00:12<00:00, 15.33it/s]
 Train accuracy: 91.7%, Avg loss: 0.284847
100%|████████████████| 40/40 [00:01<00:00, 23.01it/s]
 Test accuracy: 67.6%, Avg loss: 1.141976

 ============== Epoch 94 ==============
100%|████████████████| 196/196 [00:12<00:00, 15.39it/s]
 Train accuracy: 92.4%, Avg loss: 0.267344
100%|████████████████| 40/40 [00:01<00:00, 25.54it/s]
 Test accuracy: 66.8%, Avg loss: 1.170128

 ============== Epoch 95 ==============
100%|████████████████| 196/196 [00:12<00:00, 15.38it/s]
 Train accuracy: 92.4%, Avg loss: 0.286026
100%|████████████████| 40/40 [00:01<00:00, 25.50it/s]
 Test accuracy: 67.5%, Avg loss: 1.142693

 ============== Epoch 96 ==============
100%|████████████████| 196/196 [00:12<00:00, 15.56it/s]
 Train accuracy: 93.7%, Avg loss: 0.230866
100%|████████████████| 40/40 [00:01<00:00, 25.78it/s]
 Test accuracy: 67.6%, Avg loss: 1.161558

 ============== Epoch 97 ==============
100%|████████████████| 196/196 [00:12<00:00, 15.13it/s]
 Train accuracy: 93.6%, Avg loss: 0.235168
100%|████████████████| 40/40 [00:01<00:00, 25.50it/s]
 Test accuracy: 67.9%, Avg loss: 1.173233

 ============== Epoch 98 ==============
100%|████████████████| 196/196 [00:12<00:00, 15.17it/s]
 Train accuracy: 94.2%, Avg loss: 0.224775
100%|████████████████| 40/40 [00:01<00:00, 25.83it/s]
 Test accuracy: 67.4%, Avg loss: 1.201721

 ============== Epoch 99 ==============
100%|████████████████| 196/196 [00:12<00:00, 15.39it/s]
 Train accuracy: 94.6%, Avg loss: 0.227242
100%|████████████████| 40/40 [00:01<00:00, 25.45it/s]
 Test accuracy: 67.5%, Avg loss: 1.191411

 ============== Epoch 100 ==============
100%|████████████████| 196/196 [00:12<00:00, 15.12it/s]
 Train accuracy: 95.7%, Avg loss: 0.192843
100%|████████████████| 40/40 [00:01<00:00, 26.06it/s] Test accuracy: 67.8%, Avg loss: 1.207388
```

1c) The best test accuracy achieved is 68 %






```
100%|████████████| 196/196 [00:21<00:00,  9.14it/s]
 Train accuracy: 71.5%, Avg loss: 0.820984
100%|████████████| 40/40 [00:01<00:00, 22.54it/s]
 Test accuracy: 65.4%, Avg loss: 1.025172

 ============== Epoch 92 ==============
100%|████████████| 196/196 [00:19<00:00,  9.84it/s]
 Train accuracy: 71.6%, Avg loss: 0.815896
100%|████████████| 40/40 [00:02<00:00, 16.41it/s]
 Test accuracy: 64.8%, Avg loss: 1.045268

 ============== Epoch 93 ==============
100%|████████████| 196/196 [00:20<00:00,  9.61it/s]
 Train accuracy: 71.8%, Avg loss: 0.807980
100%|████████████| 40/40 [00:01<00:00, 23.50it/s]
 Test accuracy: 66.9%, Avg loss: 0.990542

 ============== Epoch 94 ==============
100%|████████████| 196/196 [00:20<00:00,  9.41it/s]
 Train accuracy: 71.7%, Avg loss: 0.804190
100%|████████████| 40/40 [00:01<00:00, 23.81it/s]
 Test accuracy: 67.4%, Avg loss: 0.970856

 ============== Epoch 95 ==============
100%|████████████| 196/196 [00:20<00:00,  9.74it/s]
 Train accuracy: 72.1%, Avg loss: 0.795960
100%|████████████| 40/40 [00:01<00:00, 25.36it/s]
 Test accuracy: 66.6%, Avg loss: 0.987702

 ============== Epoch 96 ==============
100%|████████████| 196/196 [00:21<00:00,  9.13it/s]
 Train accuracy: 72.5%, Avg loss: 0.788333
100%|████████████| 40/40 [00:01<00:00, 23.37it/s]
 Test accuracy: 66.8%, Avg loss: 0.995007

 ============== Epoch 97 ==============
100%|████████████| 196/196 [00:21<00:00,  9.27it/s]
 Train accuracy: 72.6%, Avg loss: 0.787328
100%|████████████| 40/40 [00:01<00:00, 21.05it/s]
 Test accuracy: 68.4%, Avg loss: 0.957383

 ============== Epoch 98 ==============
100%|████████████| 196/196 [00:20<00:00,  9.74it/s]
 Train accuracy: 73.0%, Avg loss: 0.775392
100%|████████████| 40/40 [00:01<00:00, 24.15it/s]
 Test accuracy: 68.6%, Avg loss: 0.951911

 ============== Epoch 99 ==============
100%|████████████| 196/196 [00:21<00:00,  9.25it/s]
 Train accuracy: 73.2%, Avg loss: 0.771610
100%|████████████| 40/40 [00:01<00:00, 24.04it/s]
 Test accuracy: 66.3%, Avg loss: 1.034612

 ============== Epoch 100 ==============
100%|████████████| 196/196 [00:20<00:00,  9.43it/s]
 Train accuracy: 73.2%, Avg loss: 0.768457
100%|████████████| 40/40 [00:02<00:00, 17.44it/s] Test accuracy: 68.0%, Avg loss: 0.962249
```

1d) I used the pretrained model directly

```
[58]  1  model = timm.create_model("resnet18_cifar10", pretrained=True)
```

```
[66]  1  model = model.to(device)
      2  print(model)
```

```
      (act2): ReLU(inplace=True)
    )
  )
  (layer3): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(128, 256, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (drop_block): Identity()
      (act1): ReLU(inplace=True)
      (aa): Identity()
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (act2): ReLU(inplace=True)
      (downsample): Sequential(
        (0): Conv2d(128, 256, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (1): BasicBlock(
      (conv1): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (drop_block): Identity()
      (act1): ReLU(inplace=True)
      (aa): Identity()
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (act2): ReLU(inplace=True)
    )
  )
  (layer4): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(256, 512, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (drop_block): Identity()
      (act1): ReLU(inplace=True)
      (aa): Identity()
      (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (act2): ReLU(inplace=True)
      (downsample): Sequential(
        (0): Conv2d(256, 512, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (1): BasicBlock(
      (conv1): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (drop_block): Identity()
      (act1): ReLU(inplace=True)
      (aa): Identity()
      (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (act2): ReLU(inplace=True)
    )
  )
  (global_pool): SelectAdaptivePool2d (pool_type=avg, flatten=Flatten(start_dim=1, end_dim=-1))
  (fc): Linear(in_features=512, out_features=10, bias=True)
)
```

I raised the learning rate to 0.15.

```
[68]  1  # SGD Optimizer
      2  optimizer = torch.optim.SGD(model.parameters(), lr=1.5e-1)
```

I added random cropping to the training data just like the previous question.

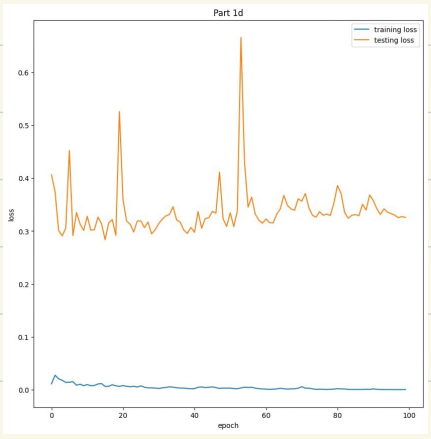
```
[59]  1  # Train transformation
      2  train_transform = transforms.Compose([
      3          transforms.RandomCrop(32, padding=4),
      4          transforms.ToTensor(),
      5  ])
      6  # Test transformation
      7  test_transform = transforms.Compose([
      8          transforms.ToTensor(),
      9  ])
```

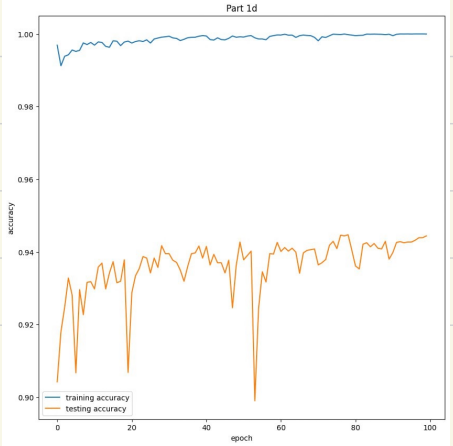
The test function and the code for training are modified so that the testing loss can be recorded

[70]
```python
# Test function
def test(dataloader, model, loss_fn):
    size = len(dataloader.dataset)
    num_batches = len(dataloader)

    # Turn on evalution mode
    model.eval()
    test_loss, correct = 0, 0

    # Turn off gradient descent
    with torch.no_grad():
        for X, y in tqdm(dataloader):
            X, y = X.to(device), y.to(device)
            pred = model(X)

            # record loss
            test_loss += loss_fn(pred, y).item()
            correct += (pred.argmax(1) == y).type(torch.float).sum().item()

    test_loss /= num_batches
    correct /= size

    print(f" Test accuracy: {(100 * correct):>0.1f}%, Avg loss: {test_loss:>8f}")
    return test_loss, correct
```

[71]
```python
# Total training epochs
epochs = 100
training_losses = []
testing_losses = []
training_accuracy = []
testing_accuracy = []
for t in range(epochs):
    print('\n', "=" * 15, "Epoch", t + 1, "=" * 15)
    train_loss, train_accuracy = train(train_dataloader, model, loss_fn, optimizer)
    test_loss, test_accuracy = test(test_dataloader, model, loss_fn)
    training_losses.append(train_loss)
    testing_losses.append(test_loss)
    training_accuracy.append(train_accuracy)
    testing_accuracy.append(test_accuracy)
```

The best test accuracy achieved is 94.4%






```
100%|██████████| 196/196 [00:55<00:00,  3.56it/s]
 Train accuracy: 100.0%, Avg loss: 0.001404
100%|██████████| 40/40 [00:04<00:00,  9.26it/s]
 Test accuracy: 94.0%, Avg loss: 0.357820

 ============= Epoch 92 =============
100%|██████████| 196/196 [00:54<00:00,  3.57it/s]
 Train accuracy: 100.0%, Avg loss: 0.000549
100%|██████████| 40/40 [00:03<00:00, 10.70it/s]
 Test accuracy: 94.3%, Avg loss: 0.342288

 ============= Epoch 93 =============
100%|██████████| 196/196 [00:54<00:00,  3.57it/s]
 Train accuracy: 100.0%, Avg loss: 0.000205
100%|██████████| 40/40 [00:03<00:00, 10.64it/s]
 Test accuracy: 94.3%, Avg loss: 0.331387

 ============= Epoch 94 =============
100%|██████████| 196/196 [00:55<00:00,  3.56it/s]
 Train accuracy: 100.0%, Avg loss: 0.000165
100%|██████████| 40/40 [00:04<00:00,  9.36it/s]
 Test accuracy: 94.2%, Avg loss: 0.342068

 ============= Epoch 95 =============
100%|██████████| 196/196 [00:54<00:00,  3.56it/s]
 Train accuracy: 100.0%, Avg loss: 0.000141
100%|██████████| 40/40 [00:03<00:00, 10.76it/s]
 Test accuracy: 94.3%, Avg loss: 0.335538

 ============= Epoch 96 =============
100%|██████████| 196/196 [00:54<00:00,  3.57it/s]
 Train accuracy: 100.0%, Avg loss: 0.000130
100%|██████████| 40/40 [00:03<00:00, 10.65it/s]
 Test accuracy: 94.3%, Avg loss: 0.332999

 ============= Epoch 97 =============
100%|██████████| 196/196 [00:54<00:00,  3.58it/s]
 Train accuracy: 100.0%, Avg loss: 0.000067
100%|██████████| 40/40 [00:04<00:00,  9.32it/s]
 Test accuracy: 94.3%, Avg loss: 0.330035

 ============= Epoch 98 =============
100%|██████████| 196/196 [00:54<00:00,  3.57it/s]
 Train accuracy: 100.0%, Avg loss: 0.000054
100%|██████████| 40/40 [00:03<00:00, 10.83it/s]
 Test accuracy: 94.4%, Avg loss: 0.325314

 ============= Epoch 99 =============
100%|██████████| 196/196 [00:54<00:00,  3.57it/s]
 Train accuracy: 100.0%, Avg loss: 0.000056
100%|██████████| 40/40 [00:03<00:00, 10.84it/s]
 Test accuracy: 94.4%, Avg loss: 0.327373

 ============= Epoch 100 =============
100%|██████████| 196/196 [00:54<00:00,  3.57it/s]
 Train accuracy: 100.0%, Avg loss: 0.000096
100%|██████████| 40/40 [00:04<00:00,  9.49it/s] Test accuracy: 94.4%, Avg loss: 0.325708
```