```python
import torch
torch.manual_seed(17)

import numpy as np
from torchsummary import summary
from tqdm import tqdm
import matplotlib.pyplot as plt

from DatasetLoader import DatasetFetcher
from project_model import *

# if torch.backends.mps.is_available():
#     mps_device = torch.device("mps")
#     x = torch.ones(1, device=mps_device)
#     print (x)
# else:
#     print ("MPS device not found.")

device = torch.device("cuda" if torch.cuda.is_available() else "cpu")
print(device)
```

```
cuda
```

```python
# Fetching Dataset
df = DatasetFetcher(dataset = "CIFAR10", batch_size = 128)
df.addHorizontalFlipping()
#df.addVerticalFlipping()
df.addRandomCrop(size = 32, padding = 4)
#df.addAutoAugmentation()
#df.addHistogramEqualization()
df.addNormalizer()
#df.addGaussianNoise()
trainLoader, testLoader = df.getLoaders()
```

```
Initializing fetching CIFAR10 dataset using torchvision
Downloading https://www.cs.toronto.edu/~kriz/cifar-10-python.tar.gz to
./data/cifar-10-python.tar.gz

100%|██████████| 170498071/170498071 [00:03<00:00, 43880869.70it/s]

Extracting ./data/cifar-10-python.tar.gz to ./data
Files already downloaded and verified
Files already downloaded and verified
```

```python
# Get Model
#model = ResNet(BasicBlock, 32, 4, [4, 4, 4, 2], 10, bias=True)
model = project1_model()
model = model.to(device)
print(summary(model, input_size = (3, 32, 32)))
```

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
```

```
================================================================
        Conv2d-1          [-1, 32, 32, 32]             896
   BatchNorm2d-2          [-1, 32, 32, 32]              64
        Conv2d-3          [-1, 32, 32, 32]           9,248
   BatchNorm2d-4          [-1, 32, 32, 32]              64
        Conv2d-5          [-1, 32, 32, 32]           9,248
   BatchNorm2d-6          [-1, 32, 32, 32]              64
    BasicBlock-7          [-1, 32, 32, 32]               0
        Conv2d-8          [-1, 32, 32, 32]           9,248
   BatchNorm2d-9          [-1, 32, 32, 32]              64
       Conv2d-10          [-1, 32, 32, 32]           9,248
  BatchNorm2d-11          [-1, 32, 32, 32]              64
   BasicBlock-12          [-1, 32, 32, 32]               0
       Conv2d-13          [-1, 32, 32, 32]           9,248
  BatchNorm2d-14          [-1, 32, 32, 32]              64
       Conv2d-15          [-1, 32, 32, 32]           9,248
  BatchNorm2d-16          [-1, 32, 32, 32]              64
   BasicBlock-17          [-1, 32, 32, 32]               0
       Conv2d-18          [-1, 32, 32, 32]           9,248
  BatchNorm2d-19          [-1, 32, 32, 32]              64
       Conv2d-20          [-1, 32, 32, 32]           9,248
  BatchNorm2d-21          [-1, 32, 32, 32]              64
   BasicBlock-22          [-1, 32, 32, 32]               0
       Conv2d-23          [-1, 64, 16, 16]          18,496
  BatchNorm2d-24          [-1, 64, 16, 16]             128
       Conv2d-25          [-1, 64, 16, 16]          36,928
  BatchNorm2d-26          [-1, 64, 16, 16]             128
       Conv2d-27          [-1, 64, 16, 16]           2,112
  BatchNorm2d-28          [-1, 64, 16, 16]             128
   BasicBlock-29          [-1, 64, 16, 16]               0
       Conv2d-30          [-1, 64, 16, 16]          36,928
  BatchNorm2d-31          [-1, 64, 16, 16]             128
       Conv2d-32          [-1, 64, 16, 16]          36,928
  BatchNorm2d-33          [-1, 64, 16, 16]             128
   BasicBlock-34          [-1, 64, 16, 16]               0
       Conv2d-35          [-1, 64, 16, 16]          36,928
  BatchNorm2d-36          [-1, 64, 16, 16]             128
       Conv2d-37          [-1, 64, 16, 16]          36,928
  BatchNorm2d-38          [-1, 64, 16, 16]             128
   BasicBlock-39          [-1, 64, 16, 16]               0
       Conv2d-40          [-1, 64, 16, 16]          36,928
  BatchNorm2d-41          [-1, 64, 16, 16]             128
       Conv2d-42          [-1, 64, 16, 16]          36,928
  BatchNorm2d-43          [-1, 64, 16, 16]             128
   BasicBlock-44          [-1, 64, 16, 16]               0
       Conv2d-45           [-1, 128, 8, 8]          73,856
  BatchNorm2d-46           [-1, 128, 8, 8]             256
       Conv2d-47           [-1, 128, 8, 8]         147,584
  BatchNorm2d-48           [-1, 128, 8, 8]             256
       Conv2d-49           [-1, 128, 8, 8]           8,320
```

```
        BatchNorm2d-50          [-1, 128, 8, 8]              256
         BasicBlock-51          [-1, 128, 8, 8]                0
           Conv2d-52            [-1, 128, 8, 8]          147,584
        BatchNorm2d-53          [-1, 128, 8, 8]              256
           Conv2d-54            [-1, 128, 8, 8]          147,584
        BatchNorm2d-55          [-1, 128, 8, 8]              256
         BasicBlock-56          [-1, 128, 8, 8]                0
           Conv2d-57            [-1, 128, 8, 8]          147,584
        BatchNorm2d-58          [-1, 128, 8, 8]              256
           Conv2d-59            [-1, 128, 8, 8]          147,584
        BatchNorm2d-60          [-1, 128, 8, 8]              256
         BasicBlock-61          [-1, 128, 8, 8]                0
           Conv2d-62            [-1, 128, 8, 8]          147,584
        BatchNorm2d-63          [-1, 128, 8, 8]              256
           Conv2d-64            [-1, 128, 8, 8]          147,584
        BatchNorm2d-65          [-1, 128, 8, 8]              256
         BasicBlock-66          [-1, 128, 8, 8]                0
           Conv2d-67            [-1, 256, 4, 4]          295,168
        BatchNorm2d-68          [-1, 256, 4, 4]              512
           Conv2d-69            [-1, 256, 4, 4]          590,080
        BatchNorm2d-70          [-1, 256, 4, 4]              512
           Conv2d-71            [-1, 256, 4, 4]           33,024
        BatchNorm2d-72          [-1, 256, 4, 4]              512
         BasicBlock-73          [-1, 256, 4, 4]                0
           Conv2d-74            [-1, 256, 4, 4]          590,080
        BatchNorm2d-75          [-1, 256, 4, 4]              512
           Conv2d-76            [-1, 256, 4, 4]          590,080
        BatchNorm2d-77          [-1, 256, 4, 4]              512
         BasicBlock-78          [-1, 256, 4, 4]                0
            Linear-79                  [-1, 10]            2,570
================================================================
Total params: 3,576,842
Trainable params: 3,576,842
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.01
Forward/backward pass size (MB): 10.00
Params size (MB): 13.64
Estimated Total Size (MB): 23.66
----------------------------------------------------------------
None

EPOCHS = 300
globalBestAccuracy = 0.0
trainingLoss = []
testingLoss = []
trainingAccuracy = []
testingAccuracy = []

# Defining Loss Function, Learning Rate, Weight Decay, Optimizer)
lossFunction = torch.nn.CrossEntropyLoss(reduction = 'sum')
```

```python
learningRate = 0.1
weightDecay = 0.0001
# optimizer = torch.optim.Adam(model.parameters(), lr=learningRate,
weight_decay=weightDecay)
optimizer = torch.optim.Adagrad(model.parameters(), lr=learningRate,
weight_decay=weightDecay)
# optimizer = torch.optim.Adadelta(model.parameters(), lr =
learningRate, weight_decay = weightDecay)
scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer,
EPOCHS, eta_min = learningRate/10.0)
print(model.eval())
trainable_parameters = sum(p.numel() for p in model.parameters() if
p.requires_grad)
print("Total Trainable Parameters : %s"%(trainable_parameters))
if trainable_parameters > 5 * (10 ** 6):
    raise Exception("Model not under budget!")

ResNet(
  (conv1): Conv2d(3, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
  (bn1): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (layer1): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn1): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (conv2): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential()
    )
    (1): BasicBlock(
      (conv1): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn1): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (conv2): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential()
    )
    (2): BasicBlock(
      (conv1): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn1): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
```

```
      (conv2): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential()
    )
    (3): BasicBlock(
      (conv1): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn1): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (conv2): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential()
    )
  )
  (layer2): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(32, 64, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1))
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential(
        (0): Conv2d(32, 64, kernel_size=(1, 1), stride=(2, 2))
        (1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (1): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential()
    )
    (2): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
```

```
    (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
    (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (shortcut): Sequential()
  )
  (3): BasicBlock(
    (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
    (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
    (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (shortcut): Sequential()
  )
)
(layer3): Sequential(
  (0): BasicBlock(
    (conv1): Conv2d(64, 128, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1))
    (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
    (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (shortcut): Sequential(
      (0): Conv2d(64, 128, kernel_size=(1, 1), stride=(2, 2))
      (1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
  )
  (1): BasicBlock(
    (conv1): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
    (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
    (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (shortcut): Sequential()
  )
  (2): BasicBlock(
    (conv1): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
    (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
```

```
      (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential()
    )
    (3): BasicBlock(
      (conv1): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential()
    )
  )
  (layer4): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(128, 256, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1))
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential(
        (0): Conv2d(128, 256, kernel_size=(1, 1), stride=(2, 2))
        (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))
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1))
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (shortcut): Sequential()
    )
  )
  (linear): Linear(in_features=256, out_features=10, bias=True)
)
Total Trainable Parameters : 3576842
```

```python
# Training
for i in tqdm(range(EPOCHS)):
    for phase in ['train', 'test']:
        if phase == "train":
            loader = trainLoader
            model.train()
            optimizer.zero_grad()
        else:
            loader = testLoader
            model.eval()
        runningLoss = 0.0
        runningCorrects = 0
        for images, labels in loader:
            images = images.to(device)
            labels = labels.to(device)
            output = model(images)
            loss = lossFunction(output, labels)
            predicted_labels = torch.argmax(output, dim=1)
            #runningLoss += loss.item()*images.size(0)
            runningLoss += loss.item()
            runningCorrects += torch.sum(predicted_labels ==
labels).float().item()
            if phase == "train":
                loss.backward()
                optimizer.step()
        epochLoss = runningLoss/len(loader.dataset)
        epochAccuracy = runningCorrects/len(loader.dataset)
        if phase == "train":
            scheduler.step()
            trainingLoss.append(epochLoss)
            trainingAccuracy.append(epochAccuracy)
        else:
            testingLoss.append(epochLoss)
            testingAccuracy.append(epochAccuracy)
            if epochAccuracy > globalBestAccuracy:
                globalBestAccuracy = epochAccuracy
                model.saveToDisk()
    print("Training Loss : %s, Testing Loss : %s, Training Accuracy :
%s, Testing Accuracy : %s"\
          %(trainingLoss[-1], testingLoss[-1], trainingAccuracy[-1],
testingAccuracy[-1]))
```

```
  0%|          | 1/300 [00:13<1:09:05, 13.86s/it]

Training Loss : 2.210271779785156, Testing Loss : 2.0776187704086304,
Training Accuracy : 0.19426, Testing Accuracy : 0.232

  1%|          | 2/300 [00:26<1:05:39, 13.22s/it]

Training Loss : 1.9658442065429687, Testing Loss : 1.8734546237945557,
Training Accuracy : 0.25926, Testing Accuracy : 0.2988
```

```
  1%|              |  3/300 [00:39<1:04:26, 13.02s/it]
```

Training Loss : 1.8579582550048828, Testing Loss : 1.8074705354690552,
Training Accuracy : 0.29862, Testing Accuracy : 0.3291

```
  1%|‖             |  4/300 [00:52<1:04:39, 13.11s/it]
```

Training Loss : 1.786487521057129, Testing Loss : 1.7441342546463012,
Training Accuracy : 0.32944, Testing Accuracy : 0.3488

```
  2%|‖             |  5/300 [01:05<1:04:46, 13.17s/it]
```

Training Loss : 1.7253449432373047, Testing Loss : 1.639097402381897,
Training Accuracy : 0.35218, Testing Accuracy : 0.3764

```
  2%|‖             |  6/300 [01:18<1:03:42, 13.00s/it]
```

Training Loss : 1.6422212664794922, Testing Loss : 1.5863514959335328,
Training Accuracy : 0.38158, Testing Accuracy : 0.4053

```
  2%|‖             |  7/300 [01:30<1:02:26, 12.79s/it]
```

Training Loss : 1.5889783309936523, Testing Loss : 1.5530458061218262,
Training Accuracy : 0.40248, Testing Accuracy : 0.4176

```
  3%|▊             |  8/300 [01:43<1:02:07, 12.77s/it]
```

Training Loss : 1.5566580459594725, Testing Loss : 1.5394046230316163,
Training Accuracy : 0.4197, Testing Accuracy : 0.4331

```
  3%|▊             |  9/300 [01:56<1:01:51, 12.75s/it]
```

Training Loss : 1.5328052172851563, Testing Loss : 1.491718630027771,
Training Accuracy : 0.42946, Testing Accuracy : 0.4422

```
  3%|▊             | 10/300 [02:09<1:02:26, 12.92s/it]
```

Training Loss : 1.5076172845458984, Testing Loss : 1.4853337795257568,
Training Accuracy : 0.43724, Testing Accuracy : 0.4457

```
  4%|▊             | 11/300 [02:22<1:02:24, 12.96s/it]
```

Training Loss : 1.4697139544677735, Testing Loss : 1.3956403562545776,
Training Accuracy : 0.45616, Testing Accuracy : 0.4854

```
  4%|█             | 12/300 [02:35<1:01:37, 12.84s/it]
```

Training Loss : 1.4062362626647948, Testing Loss : 1.3708709671020507,
Training Accuracy : 0.4825, Testing Accuracy : 0.4956

```
  4%|█             | 13/300 [02:47<1:00:51, 12.72s/it]
```

Training Loss : 1.367282255859375, Testing Loss : 1.3027161432266234,
Training Accuracy : 0.5025, Testing Accuracy : 0.5183

```
    5%|█          | 14/300 [03:00<1:00:56, 12.79s/it]
```

Training Loss : 1.3258342610168456, Testing Loss : 1.3046602102279663, Training Accuracy : 0.51758, Testing Accuracy : 0.5197

```
    5%|█          | 15/300 [03:13<1:00:45, 12.79s/it]
```

Training Loss : 1.294127788848877, Testing Loss : 1.2478647279739379, Training Accuracy : 0.53048, Testing Accuracy : 0.5458

```
    5%|█          | 16/300 [03:27<1:01:41, 13.03s/it]
```

Training Loss : 1.2675523260498047, Testing Loss : 1.2463994152069091, Training Accuracy : 0.54084, Testing Accuracy : 0.5391

```
    6%|█          | 17/300 [03:39<1:01:10, 12.97s/it]
```

Training Loss : 1.239879885406494, Testing Loss : 1.299667756462097, Training Accuracy : 0.55234, Testing Accuracy : 0.5237

```
    6%|█          | 18/300 [03:52<1:00:11, 12.81s/it]
```

Training Loss : 1.2260869799804688, Testing Loss : 1.194503775024414, Training Accuracy : 0.55906, Testing Accuracy : 0.5586

```
    6%|█          | 19/300 [04:04<59:12, 12.64s/it]
```

Training Loss : 1.1604465754699707, Testing Loss : 1.116478502368927, Training Accuracy : 0.58238, Testing Accuracy : 0.59

```
    7%|█          | 20/300 [04:17<59:50, 12.82s/it]
```

Training Loss : 1.114789803314209, Testing Loss : 1.110698389148712, Training Accuracy : 0.5996, Testing Accuracy : 0.6001

```
    7%|█          | 21/300 [04:31<1:00:24, 12.99s/it]
```

Training Loss : 1.0812384391784668, Testing Loss : 1.098574313545227, Training Accuracy : 0.61482, Testing Accuracy : 0.6114

```
    7%|█          | 22/300 [04:44<1:00:22, 13.03s/it]
```

Training Loss : 1.0557499084472657, Testing Loss : 1.0852108723640441, Training Accuracy : 0.6225, Testing Accuracy : 0.6167

```
    8%|█          | 23/300 [04:57<1:00:04, 13.01s/it]
```

Training Loss : 1.0249381176757812, Testing Loss : 1.086420977973938, Training Accuracy : 0.63366, Testing Accuracy : 0.6175

```
    8%|█          | 24/300 [05:10<59:39, 12.97s/it]
```

Training Loss : 1.0023556413269044, Testing Loss : 1.0232274389266969, Training Accuracy : 0.63806, Testing Accuracy : 0.6341

```
   8%|█          | 25/300 [05:22<58:59, 12.87s/it]
```

Training Loss : 0.9721787455749512, Testing Loss : 0.9653262029647827, Training Accuracy : 0.65318, Testing Accuracy : 0.6607

```
   9%|█          | 26/300 [05:36<59:14, 12.97s/it]
```

Training Loss : 0.945418916015625, Testing Loss : 0.962262281703949, Training Accuracy : 0.66362, Testing Accuracy : 0.6638

```
   9%|█          | 27/300 [05:49<59:11, 13.01s/it]
```

Training Loss : 0.933024084777832, Testing Loss : 0.9733237300872802, Training Accuracy : 0.6699, Testing Accuracy : 0.6593

```
   9%|█          | 28/300 [06:02<59:01, 13.02s/it]
```

Training Loss : 0.9058255374145507, Testing Loss : 0.9323313039779663, Training Accuracy : 0.67894, Testing Accuracy : 0.6724

```
  10%|█          | 29/300 [06:14<57:59, 12.84s/it]
```

Training Loss : 0.8828874461364746, Testing Loss : 0.9310770510673523, Training Accuracy : 0.68644, Testing Accuracy : 0.6688

```
  10%|█          | 30/300 [06:27<57:22, 12.75s/it]
```

Training Loss : 0.8608403677368164, Testing Loss : 0.8999277117729187, Training Accuracy : 0.69404, Testing Accuracy : 0.6805

```
  10%|█          | 31/300 [06:39<56:52, 12.69s/it]
```

Training Loss : 0.8589099822998046, Testing Loss : 0.8982210741996766, Training Accuracy : 0.6971, Testing Accuracy : 0.6822

```
  11%|█          | 32/300 [06:53<57:32, 12.88s/it]
```

Training Loss : 0.8546044290161133, Testing Loss : 0.8488029915809632, Training Accuracy : 0.69724, Testing Accuracy : 0.7029

```
  11%|█          | 33/300 [07:06<57:49, 12.99s/it]
```

Training Loss : 0.8169798333740235, Testing Loss : 0.841673657989502, Training Accuracy : 0.71218, Testing Accuracy : 0.7059

```
  11%|█          | 34/300 [07:19<57:35, 12.99s/it]
```

Training Loss : 0.7940677331542969, Testing Loss : 0.8311644828796386, Training Accuracy : 0.71956, Testing Accuracy : 0.7108

```
  12%|█          | 35/300 [07:32<57:08, 12.94s/it]
```

Training Loss : 0.7728162466430664, Testing Loss : 0.8486702626228333, Training Accuracy : 0.73022, Testing Accuracy : 0.7016

```
   12%|█             | 36/300 [07:45<57:13, 13.01s/it]
```

Training Loss : 0.766965691986084, Testing Loss : 0.8355301305770874,
Training Accuracy : 0.7293, Testing Accuracy : 0.7085

```
   12%|█             | 37/300 [07:58<57:33, 13.13s/it]
```

Training Loss : 0.7626024369049073, Testing Loss : 0.8133466177940368,
Training Accuracy : 0.7313, Testing Accuracy : 0.7174

```
   13%|█             | 38/300 [08:11<57:28, 13.16s/it]
```

Training Loss : 0.7396993919372559, Testing Loss : 0.784205014038086,
Training Accuracy : 0.74252, Testing Accuracy : 0.7269

```
   13%|█             | 39/300 [08:24<56:59, 13.10s/it]
```

Training Loss : 0.7341398988342285, Testing Loss : 0.808985516166687,
Training Accuracy : 0.74136, Testing Accuracy : 0.7204

```
   13%|█             | 40/300 [08:37<56:09, 12.96s/it]
```

Training Loss : 0.7232949201202392, Testing Loss : 0.7761296573638916,
Training Accuracy : 0.74664, Testing Accuracy : 0.7386

```
   14%|█             | 41/300 [08:49<55:07, 12.77s/it]
```

Training Loss : 0.6975469217681884, Testing Loss : 0.762520223236084,
Training Accuracy : 0.75672, Testing Accuracy : 0.7406

```
   14%|█             | 42/300 [09:02<55:07, 12.82s/it]
```

Training Loss : 0.6812267900848389, Testing Loss : 0.8049013989925384,
Training Accuracy : 0.76332, Testing Accuracy : 0.7233

```
   14%|█             | 43/300 [09:15<55:13, 12.89s/it]
```

Training Loss : 0.6781241802215576, Testing Loss : 0.743831463098526,
Training Accuracy : 0.7624, Testing Accuracy : 0.743

```
   15%|█             | 44/300 [09:29<55:47, 13.08s/it]
```

Training Loss : 0.6651006146240235, Testing Loss : 0.7540731355667114,
Training Accuracy : 0.7672, Testing Accuracy : 0.7439

```
   15%|█             | 45/300 [09:42<55:41, 13.10s/it]
```

Training Loss : 0.6479671980285645, Testing Loss : 0.7309305477142334,
Training Accuracy : 0.77384, Testing Accuracy : 0.7514

```
   15%|█             | 46/300 [09:55<55:24, 13.09s/it]
```

Training Loss : 0.6269497093200683, Testing Loss : 0.7230632201194763,
Training Accuracy : 0.78346, Testing Accuracy : 0.7555

```
  16%|█        | 47/300 [10:08<54:34, 12.94s/it]
```

Training Loss : 0.6253957757568359, Testing Loss : 0.7195323806762696,
Training Accuracy : 0.78212, Testing Accuracy : 0.7578

```
  16%|█        | 48/300 [10:21<54:24, 12.95s/it]
```

Training Loss : 0.6165354170989991, Testing Loss : 0.6963144652366638,
Training Accuracy : 0.78432, Testing Accuracy : 0.7643

```
  16%|█        | 49/300 [10:34<55:03, 13.16s/it]
```

Training Loss : 0.5885531827545166, Testing Loss : 0.6657747723579407,
Training Accuracy : 0.79542, Testing Accuracy : 0.7768

```
  17%|█        | 50/300 [10:48<55:16, 13.27s/it]
```

Training Loss : 0.5844202860260009, Testing Loss : 0.6774873884677887,
Training Accuracy : 0.7981, Testing Accuracy : 0.7695

```
  17%|█        | 51/300 [11:01<55:06, 13.28s/it]
```

Training Loss : 0.5776965895843506, Testing Loss : 0.6879039651870728,
Training Accuracy : 0.7994, Testing Accuracy : 0.7705

```
  17%|█        | 52/300 [11:14<54:05, 13.09s/it]
```

Training Loss : 0.5715209605407715, Testing Loss : 0.6748840007781982,
Training Accuracy : 0.80122, Testing Accuracy : 0.774

```
  18%|█        | 53/300 [11:26<53:14, 12.93s/it]
```

Training Loss : 0.5682792108154296, Testing Loss : 0.7336050190925598,
Training Accuracy : 0.80294, Testing Accuracy : 0.7545

```
  18%|█        | 54/300 [11:39<53:00, 12.93s/it]
```

Training Loss : 0.5711143093872071, Testing Loss : 0.6873980659484863,
Training Accuracy : 0.8012, Testing Accuracy : 0.7652

```
  18%|█        | 55/300 [11:53<53:22, 13.07s/it]
```

Training Loss : 0.5529529873657226, Testing Loss : 0.6576758522987366,
Training Accuracy : 0.8069, Testing Accuracy : 0.7805

```
  19%|█        | 56/300 [12:06<53:24, 13.13s/it]
```

Training Loss : 0.5459928346252442, Testing Loss : 0.6497614676475525,
Training Accuracy : 0.80802, Testing Accuracy : 0.7831

```
  19%|█        | 57/300 [12:19<52:54, 13.06s/it]
```

Training Loss : 0.5285766951751709, Testing Loss : 0.6424392332077027,
Training Accuracy : 0.81556, Testing Accuracy : 0.7837

```
 19%|█            | 58/300 [12:32<52:18, 12.97s/it]
```

Training Loss : 0.5211880172729492, Testing Loss : 0.6576542192459106, Training Accuracy : 0.81902, Testing Accuracy : 0.7818

```
 20%|█            | 59/300 [12:45<52:25, 13.05s/it]
```

Training Loss : 0.5278427574157715, Testing Loss : 0.6580942921638488, Training Accuracy : 0.81632, Testing Accuracy : 0.7802

```
 20%|█            | 60/300 [12:58<52:49, 13.21s/it]
```

Training Loss : 0.5116381945037842, Testing Loss : 0.6357472201347351, Training Accuracy : 0.82292, Testing Accuracy : 0.7874

```
 20%|█            | 61/300 [13:12<52:46, 13.25s/it]
```

Training Loss : 0.4992331523132324, Testing Loss : 0.6186490751266479, Training Accuracy : 0.82512, Testing Accuracy : 0.7912

```
 21%|█            | 62/300 [13:25<52:40, 13.28s/it]
```

Training Loss : 0.4856940021514893, Testing Loss : 0.6260344727993011, Training Accuracy : 0.8313, Testing Accuracy : 0.7842

```
 21%|█            | 63/300 [13:38<52:10, 13.21s/it]
```

Training Loss : 0.48008064018249513, Testing Loss : 0.6177518483161927, Training Accuracy : 0.8325, Testing Accuracy : 0.7939

```
 21%|█            | 64/300 [13:51<51:48, 13.17s/it]
```

Training Loss : 0.4804463861846924, Testing Loss : 0.6057563005447387, Training Accuracy : 0.83374, Testing Accuracy : 0.7963

```
 22%|█            | 65/300 [14:04<51:09, 13.06s/it]
```

Training Loss : 0.4594190396881104, Testing Loss : 0.6019816327095032, Training Accuracy : 0.8391, Testing Accuracy : 0.8009

```
 22%|█            | 66/300 [14:17<51:13, 13.13s/it]
```

Training Loss : 0.462071981048584, Testing Loss : 0.6178116791248321, Training Accuracy : 0.83864, Testing Accuracy : 0.7987

```
 22%|█            | 67/300 [14:30<50:20, 12.96s/it]
```

Training Loss : 0.45585899223327636, Testing Loss : 0.59281337018013, Training Accuracy : 0.84278, Testing Accuracy : 0.8042

```
 23%|█            | 68/300 [14:43<49:57, 12.92s/it]
```

```
Training Loss : 0.44985370002746583, Testing Loss :
0.6026963053703308, Training Accuracy : 0.84282, Testing Accuracy :
0.8047

   23%|██           | 69/300 [14:55<49:18, 12.81s/it]

Training Loss : 0.4413754772567749, Testing Loss : 0.5942634033441544,
Training Accuracy : 0.84594, Testing Accuracy : 0.8083

   23%|██           | 70/300 [15:08<49:04, 12.80s/it]

Training Loss : 0.4301059061431885, Testing Loss : 0.6083448704719543,
Training Accuracy : 0.85042, Testing Accuracy : 0.8045

   24%|██           | 71/300 [15:21<48:38, 12.75s/it]

Training Loss : 0.4243701100540161, Testing Loss : 0.5912619731426239,
Training Accuracy : 0.85198, Testing Accuracy : 0.8071

   24%|██           | 72/300 [15:33<48:04, 12.65s/it]

Training Loss : 0.4251009289932251, Testing Loss : 0.6201574118614197,
Training Accuracy : 0.85086, Testing Accuracy : 0.7995

   24%|██           | 73/300 [15:46<47:39, 12.60s/it]

Training Loss : 0.4162588314819336, Testing Loss : 0.5740034718513489,
Training Accuracy : 0.85536, Testing Accuracy : 0.8089

   25%|██           | 74/300 [15:58<47:46, 12.69s/it]

Training Loss : 0.4089480495071411, Testing Loss : 0.6131972806930542,
Training Accuracy : 0.85662, Testing Accuracy : 0.805

   25%|██           | 75/300 [16:12<48:18, 12.88s/it]

Training Loss : 0.41226461040496826, Testing Loss :
0.5770815107345582, Training Accuracy : 0.85748, Testing Accuracy :
0.8118

   25%|██           | 76/300 [16:24<47:49, 12.81s/it]

Training Loss : 0.4033053939437866, Testing Loss : 0.5805470794200898,
Training Accuracy : 0.86092, Testing Accuracy : 0.8107

   26%|██           | 77/300 [16:37<47:21, 12.74s/it]

Training Loss : 0.3895646762466431, Testing Loss : 0.5838736433029175,
Training Accuracy : 0.86278, Testing Accuracy : 0.8121

   26%|██           | 78/300 [16:50<47:03, 12.72s/it]

Training Loss : 0.4010584759902954, Testing Loss : 0.5904190046310425,
Training Accuracy : 0.8602, Testing Accuracy : 0.8089
```

```
26%|██          | 79/300 [17:02<46:33, 12.64s/it]
```

Training Loss : 0.39817938575744627, Testing Loss : 0.6160044606685638, Training Accuracy : 0.8605, Testing Accuracy : 0.8047

```
27%|██          | 80/300 [17:15<46:19, 12.63s/it]
```

Training Loss : 0.3910089391708374, Testing Loss : 0.5886035094738007, Training Accuracy : 0.86312, Testing Accuracy : 0.8094

```
27%|██          | 81/300 [17:28<46:22, 12.70s/it]
```

Training Loss : 0.38307571311950683, Testing Loss : 0.5906538954257965, Training Accuracy : 0.8659, Testing Accuracy : 0.8091

```
27%|██          | 82/300 [17:40<46:03, 12.67s/it]
```

Training Loss : 0.3845488144302368, Testing Loss : 0.6047785409688949, Training Accuracy : 0.86472, Testing Accuracy : 0.8121

```
28%|██          | 83/300 [17:53<45:58, 12.71s/it]
```

Training Loss : 0.3885998408126831, Testing Loss : 0.5822566366195678, Training Accuracy : 0.86378, Testing Accuracy : 0.8131

```
28%|██          | 84/300 [18:06<45:57, 12.77s/it]
```

Training Loss : 0.3776556831741333, Testing Loss : 0.605935665845871, Training Accuracy : 0.86754, Testing Accuracy : 0.8109

```
28%|██          | 85/300 [18:19<46:09, 12.88s/it]
```

Training Loss : 0.37193075374603274, Testing Loss : 0.5964493837833404, Training Accuracy : 0.8694, Testing Accuracy : 0.8092

```
29%|██          | 86/300 [18:32<46:06, 12.93s/it]
```

Training Loss : 0.3713338168716431, Testing Loss : 0.609518310213089, Training Accuracy : 0.8686, Testing Accuracy : 0.8107

```
29%|██          | 87/300 [18:44<44:52, 12.64s/it]
```

Training Loss : 0.36621388256072995, Testing Loss : 0.5881746784687042, Training Accuracy : 0.87272, Testing Accuracy : 0.8104

```
29%|██          | 88/300 [18:57<45:16, 12.81s/it]
```

Training Loss : 0.36084655464172366, Testing Loss : 0.5697613476991653, Training Accuracy : 0.87364, Testing Accuracy : 0.8232

```
 30%|██          | 89/300 [19:10<44:52, 12.76s/it]
```

Training Loss : 0.34662986919403077, Testing Loss : 0.586603015422821, Training Accuracy : 0.87944, Testing Accuracy : 0.8144

```
 30%|██          | 90/300 [19:22<44:22, 12.68s/it]
```

Training Loss : 0.3438081577682495, Testing Loss : 0.5715156350851059, Training Accuracy : 0.88098, Testing Accuracy : 0.8228

```
 30%|██          | 91/300 [19:35<44:01, 12.64s/it]
```

Training Loss : 0.33658408824920655, Testing Loss : 0.5823370859622955, Training Accuracy : 0.88222, Testing Accuracy : 0.8166

```
 31%|██          | 92/300 [19:47<43:40, 12.60s/it]
```

Training Loss : 0.3381971918487549, Testing Loss : 0.5822079497814179, Training Accuracy : 0.881, Testing Accuracy : 0.8228

```
 31%|██          | 93/300 [20:00<43:16, 12.54s/it]
```

Training Loss : 0.3299005615615845, Testing Loss : 0.5835844218254089, Training Accuracy : 0.88508, Testing Accuracy : 0.8166

```
 31%|██          | 94/300 [20:13<43:35, 12.70s/it]
```

Training Loss : 0.33064772605896, Testing Loss : 0.5860104024410248, Training Accuracy : 0.8834, Testing Accuracy : 0.8257

```
 32%|██          | 95/300 [20:26<43:25, 12.71s/it]
```

Training Loss : 0.3308615256500244, Testing Loss : 0.5942693105697632, Training Accuracy : 0.88248, Testing Accuracy : 0.8204

```
 32%|██          | 96/300 [20:39<43:40, 12.85s/it]
```

Training Loss : 0.3178500291824341, Testing Loss : 0.5535113625049591, Training Accuracy : 0.88932, Testing Accuracy : 0.829

```
 32%|██          | 97/300 [20:51<42:59, 12.71s/it]
```

Training Loss : 0.31047074113845824, Testing Loss : 0.568425732421875, Training Accuracy : 0.89272, Testing Accuracy : 0.8243

```
 33%|██          | 98/300 [21:04<42:55, 12.75s/it]
```

Training Loss : 0.3050571424484253, Testing Loss : 0.5684462862491607, Training Accuracy : 0.89298, Testing Accuracy : 0.826

```
 33%|██          | 99/300 [21:16<42:09, 12.58s/it]
```

Training Loss : 0.3012646688079834, Testing Loss : 0.5659972508907318, Training Accuracy : 0.8944, Testing Accuracy : 0.8225

```
33%|███         | 100/300 [21:29<41:59, 12.60s/it]
```

Training Loss : 0.3063092251205444, Testing Loss : 0.5663371178388595,
Training Accuracy : 0.89228, Testing Accuracy : 0.8229

```
34%|███         | 101/300 [21:42<42:42, 12.88s/it]
```

Training Loss : 0.3026895722961426, Testing Loss : 0.5842990087032318,
Training Accuracy : 0.8937, Testing Accuracy : 0.8206

```
34%|███         | 102/300 [21:57<44:25, 13.46s/it]
```

Training Loss : 0.3005465925598145, Testing Loss : 0.6017999830007553,
Training Accuracy : 0.89286, Testing Accuracy : 0.8212

```
34%|███         | 103/300 [22:12<45:10, 13.76s/it]
```

Training Loss : 0.29762611724853516, Testing Loss :
0.5466107523679733, Training Accuracy : 0.89438, Testing Accuracy :
0.8308

```
35%|███         | 104/300 [22:27<45:58, 14.07s/it]
```

Training Loss : 0.28864576984405516, Testing Loss :
0.5817524383544922, Training Accuracy : 0.89998, Testing Accuracy :
0.8249

```
35%|███         | 105/300 [22:41<46:16, 14.24s/it]
```

Training Loss : 0.28560122356414797, Testing Loss :
0.5564332264423371, Training Accuracy : 0.90016, Testing Accuracy :
0.8314

```
35%|███         | 106/300 [22:56<46:34, 14.40s/it]
```

Training Loss : 0.28163359275817873, Testing Loss :
0.5931995057582855, Training Accuracy : 0.90196, Testing Accuracy :
0.8262

```
36%|███         | 107/300 [23:10<46:13, 14.37s/it]
```

Training Loss : 0.2786657756996155, Testing Loss : 0.5908474447250366,
Training Accuracy : 0.90232, Testing Accuracy : 0.8265

```
36%|███         | 108/300 [23:25<46:13, 14.44s/it]
```

Training Loss : 0.2773500256538391, Testing Loss : 0.5951224136590958,
Training Accuracy : 0.90164, Testing Accuracy : 0.8242

```
36%|███         | 109/300 [23:39<45:59, 14.45s/it]
```

Training Loss : 0.2789917255401611, Testing Loss : 0.5794716274499894,
Training Accuracy : 0.9013, Testing Accuracy : 0.8316

```
37%|███         | 110/300 [23:54<45:39, 14.42s/it]
```

Training Loss : 0.26636116289138795, Testing Loss :
0.5796341747283935, Training Accuracy : 0.90536, Testing Accuracy :
0.8349

37%|███ | 111/300 [24:08<44:59, 14.28s/it]

Training Loss : 0.26045307832717896, Testing Loss :
0.5609304426431656, Training Accuracy : 0.90908, Testing Accuracy :
0.8309

37%|███ | 112/300 [24:22<44:33, 14.22s/it]

Training Loss : 0.2594799780082703, Testing Loss : 0.5983319197416306,
Training Accuracy : 0.9075, Testing Accuracy : 0.8294

38%|███ | 113/300 [24:36<44:31, 14.28s/it]

Training Loss : 0.26896510133743284, Testing Loss :
0.5628561482429505, Training Accuracy : 0.90558, Testing Accuracy :
0.832

38%|███ | 114/300 [24:51<44:41, 14.42s/it]

Training Loss : 0.25245902477264404, Testing Loss :
0.5779407910346985, Training Accuracy : 0.91052, Testing Accuracy :
0.8351

38%|███ | 115/300 [25:06<44:55, 14.57s/it]

Training Loss : 0.24133291776657104, Testing Loss :
0.5599561423778534, Training Accuracy : 0.91628, Testing Accuracy :
0.8371

39%|███ | 116/300 [25:21<45:09, 14.72s/it]

Training Loss : 0.24097785459518434, Testing Loss :
0.5741554753780365, Training Accuracy : 0.91426, Testing Accuracy :
0.8326

39%|███ | 117/300 [25:35<44:44, 14.67s/it]

Training Loss : 0.23633975519180297, Testing Loss :
0.5801919190168381, Training Accuracy : 0.91638, Testing Accuracy :
0.8327

39%|███ | 118/300 [25:50<44:33, 14.69s/it]

Training Loss : 0.23847772430419922, Testing Loss : 0.574350114440918,
Training Accuracy : 0.91584, Testing Accuracy : 0.8314

40%|███ | 119/300 [26:05<44:18, 14.69s/it]

Training Loss : 0.23243807523727417, Testing Loss : 0.579744031071663,
Training Accuracy : 0.91816, Testing Accuracy : 0.8313

```
    40%|████        | 120/300 [26:19<43:49, 14.61s/it]
```

Training Loss : 0.2291406495285034, Testing Loss : 0.5710050269126892, Training Accuracy : 0.91852, Testing Accuracy : 0.8343

```
    40%|████        | 121/300 [26:33<43:14, 14.49s/it]
```

Training Loss : 0.22850960542678833, Testing Loss : 0.5584362402439117, Training Accuracy : 0.91934, Testing Accuracy : 0.8348

```
    41%|████        | 122/300 [26:48<42:57, 14.48s/it]
```

Training Loss : 0.23215229621887207, Testing Loss : 0.5682626835346222, Training Accuracy : 0.9175, Testing Accuracy : 0.8357

```
    41%|████        | 123/300 [27:03<42:52, 14.53s/it]
```

Training Loss : 0.224708242893219, Testing Loss : 0.5857954333543778, Training Accuracy : 0.92058, Testing Accuracy : 0.8368

```
    41%|████        | 124/300 [27:17<42:43, 14.56s/it]
```

Training Loss : 0.22397112775802613, Testing Loss : 0.5685636486053467, Training Accuracy : 0.92092, Testing Accuracy : 0.8401

```
    42%|████        | 125/300 [27:32<42:30, 14.57s/it]
```

Training Loss : 0.22168464318275452, Testing Loss : 0.5722427138805389, Training Accuracy : 0.92232, Testing Accuracy : 0.8389

```
    42%|████        | 126/300 [27:47<42:23, 14.62s/it]
```

Training Loss : 0.21482682754516602, Testing Loss : 0.5541908057689666, Training Accuracy : 0.92406, Testing Accuracy : 0.8419

```
    42%|████        | 127/300 [28:01<42:05, 14.60s/it]
```

Training Loss : 0.21160499462127685, Testing Loss : 0.57845505900383, Training Accuracy : 0.92552, Testing Accuracy : 0.8401

```
    43%|████        | 128/300 [28:16<42:03, 14.67s/it]
```

Training Loss : 0.2088778415298462, Testing Loss : 0.579351811170578, Training Accuracy : 0.92688, Testing Accuracy : 0.8378

```
    43%|████        | 129/300 [28:31<41:51, 14.69s/it]
```

Training Loss : 0.20900318157196046, Testing Loss :
0.5839400049209594, Training Accuracy : 0.92556, Testing Accuracy :
0.8382

  43%|██████ | 130/300 [28:45<41:31, 14.66s/it]

Training Loss : 0.20906648764610292, Testing Loss :
0.5650667197942734, Training Accuracy : 0.92602, Testing Accuracy :
0.8448

  44%|██████ | 131/300 [29:00<41:15, 14.65s/it]

Training Loss : 0.20397900932312013, Testing Loss :
0.6034406856536865, Training Accuracy : 0.92736, Testing Accuracy :
0.8334

  44%|██████ | 132/300 [29:15<41:15, 14.73s/it]

Training Loss : 0.20688124765396118, Testing Loss :
0.5758597497463226, Training Accuracy : 0.92678, Testing Accuracy :
0.8379

  44%|██████ | 133/300 [29:29<40:45, 14.64s/it]

Training Loss : 0.20339282512664794, Testing Loss :
0.5794738677024841, Training Accuracy : 0.9271, Testing Accuracy :
0.8423

  45%|██████ | 134/300 [29:44<40:21, 14.59s/it]

Training Loss : 0.1956898934173584, Testing Loss : 0.5615729581832886,
Training Accuracy : 0.93156, Testing Accuracy : 0.8423

  45%|██████ | 135/300 [29:58<40:16, 14.64s/it]

Training Loss : 0.18894401049613951, Testing Loss :
0.5922660102367401, Training Accuracy : 0.9336, Testing Accuracy :
0.8398

  45%|██████ | 136/300 [30:13<39:58, 14.62s/it]

Training Loss : 0.1869845860862732, Testing Loss : 0.5747343648910522,
Training Accuracy : 0.93376, Testing Accuracy : 0.8461

  46%|██████ | 137/300 [30:28<39:47, 14.65s/it]

Training Loss : 0.18029606172561646, Testing Loss :
0.5935397563934326, Training Accuracy : 0.93602, Testing Accuracy :
0.8433

  46%|██████ | 138/300 [30:42<39:36, 14.67s/it]

Training Loss : 0.184305543384552, Testing Loss : 0.5761169612884521,
Training Accuracy : 0.93472, Testing Accuracy : 0.8455

46%|████ | 139/300 [30:57<39:30, 14.72s/it]

Training Loss : 0.18046229213714599, Testing Loss : 0.598217173576355, Training Accuracy : 0.93606, Testing Accuracy : 0.8392

47%|████ | 140/300 [31:12<39:33, 14.83s/it]

Training Loss : 0.18314362045288085, Testing Loss : 0.5899095553398133, Training Accuracy : 0.93532, Testing Accuracy : 0.843

47%|████ | 141/300 [31:27<39:11, 14.79s/it]

Training Loss : 0.18672931734085083, Testing Loss : 0.6183504626274109, Training Accuracy : 0.93332, Testing Accuracy : 0.8412

47%|████ | 142/300 [31:41<38:37, 14.67s/it]

Training Loss : 0.18730115495681762, Testing Loss : 0.6222307421684266, Training Accuracy : 0.93388, Testing Accuracy : 0.8357

48%|████ | 143/300 [31:56<38:19, 14.65s/it]

Training Loss : 0.18872408826828002, Testing Loss : 0.5988401937961578, Training Accuracy : 0.93276, Testing Accuracy : 0.8417

48%|████ | 144/300 [32:11<38:14, 14.71s/it]

Training Loss : 0.18115683355331422, Testing Loss : 0.6042240878582, Training Accuracy : 0.93582, Testing Accuracy : 0.8397

48%|████ | 145/300 [32:26<38:17, 14.82s/it]

Training Loss : 0.175361667842865, Testing Loss : 0.5852673171520233, Training Accuracy : 0.93724, Testing Accuracy : 0.8444

49%|████ | 146/300 [32:41<37:57, 14.79s/it]

Training Loss : 0.1696713897895813, Testing Loss : 0.5997202711582184, Training Accuracy : 0.94038, Testing Accuracy : 0.8422

49%|████ | 147/300 [32:55<37:38, 14.76s/it]

Training Loss : 0.16404604323387145, Testing Loss : 0.6032806653499603, Training Accuracy : 0.94208, Testing Accuracy : 0.8489

49%|████ | 148/300 [33:10<37:29, 14.80s/it]

```
Training Loss : 0.16689515434265137, Testing Loss :
0.6283687614917756, Training Accuracy : 0.94076, Testing Accuracy :
0.835

  50%|██████        | 149/300 [33:25<36:54, 14.67s/it]

Training Loss : 0.16681827169418334, Testing Loss :
0.6017134065866471, Training Accuracy : 0.94106, Testing Accuracy :
0.8469

  50%|██████        | 150/300 [33:39<36:43, 14.69s/it]

Training Loss : 0.1651969480228424, Testing Loss : 0.6063335008144378,
Training Accuracy : 0.94196, Testing Accuracy : 0.8407

  50%|██████        | 151/300 [33:54<36:21, 14.64s/it]

Training Loss : 0.16143436184883117, Testing Loss :
0.5934721478939057, Training Accuracy : 0.94222, Testing Accuracy :
0.8473

  51%|██████        | 152/300 [34:09<36:23, 14.76s/it]

Training Loss : 0.15341960541725158, Testing Loss :
0.6176096743583679, Training Accuracy : 0.94548, Testing Accuracy :
0.8423

  51%|██████        | 153/300 [34:24<36:12, 14.78s/it]

Training Loss : 0.1586002672958374, Testing Loss : 0.6190509816408157,
Training Accuracy : 0.94454, Testing Accuracy : 0.8419

  51%|██████        | 154/300 [34:38<35:51, 14.74s/it]

Training Loss : 0.15836795718193053, Testing Loss :
0.6166306717395782, Training Accuracy : 0.94248, Testing Accuracy :
0.8431

  52%|██████        | 155/300 [34:53<35:41, 14.77s/it]

Training Loss : 0.15151864468574525, Testing Loss :
0.6300231192231178, Training Accuracy : 0.9465, Testing Accuracy :
0.8388

  52%|██████        | 156/300 [35:08<35:26, 14.77s/it]

Training Loss : 0.15121422521591185, Testing Loss :
0.6122699043273926, Training Accuracy : 0.94576, Testing Accuracy :
0.8446

  52%|██████        | 157/300 [35:23<35:16, 14.80s/it]
```

```
Training Loss : 0.15012972643852235, Testing Loss :
0.6230054361820221, Training Accuracy : 0.94644, Testing Accuracy :
0.8438

 53%|██████        | 158/300 [35:37<34:47, 14.70s/it]

Training Loss : 0.1466237959766388, Testing Loss : 0.6275569225788117,
Training Accuracy : 0.9481, Testing Accuracy : 0.8451

 53%|██████        | 159/300 [35:52<34:41, 14.77s/it]

Training Loss : 0.1460279961204529, Testing Loss : 0.5985301371097564,
Training Accuracy : 0.94712, Testing Accuracy : 0.847

 53%|██████        | 160/300 [36:07<34:10, 14.65s/it]

Training Loss : 0.14367742504119874, Testing Loss :
0.6208131456375122, Training Accuracy : 0.94822, Testing Accuracy :
0.8437

 54%|██████        | 161/300 [36:21<33:51, 14.61s/it]

Training Loss : 0.14019677845001222, Testing Loss :
0.6224750274658203, Training Accuracy : 0.95004, Testing Accuracy :
0.8471

 54%|██████        | 162/300 [36:36<33:32, 14.58s/it]

Training Loss : 0.1381765556049347, Testing Loss : 0.6178542472839356,
Training Accuracy : 0.95032, Testing Accuracy : 0.8464

 54%|██████        | 163/300 [36:50<32:55, 14.42s/it]

Training Loss : 0.142029545917511, Testing Loss : 0.6259154792785645,
Training Accuracy : 0.94986, Testing Accuracy : 0.8462

 55%|██████        | 164/300 [37:05<33:09, 14.63s/it]

Training Loss : 0.13812119091033936, Testing Loss :
0.6279495855808258, Training Accuracy : 0.95098, Testing Accuracy :
0.8436

 55%|██████        | 165/300 [37:20<33:04, 14.70s/it]

Training Loss : 0.1340565928554535, Testing Loss : 0.6284278393745423,
Training Accuracy : 0.95314, Testing Accuracy : 0.8447

 55%|██████        | 166/300 [37:34<32:48, 14.69s/it]

Training Loss : 0.1318954757785797, Testing Loss : 0.6381781487464905,
Training Accuracy : 0.95378, Testing Accuracy : 0.8478

 56%|██████        | 167/300 [37:49<32:30, 14.66s/it]
```

```
Training Loss : 0.12915791215896608, Testing Loss :
0.6275996404647827, Training Accuracy : 0.95354, Testing Accuracy :
0.8464

  56%|██████        | 168/300 [38:04<32:30, 14.78s/it]

Training Loss : 0.12911849113464355, Testing Loss :
0.6436150869846344, Training Accuracy : 0.95432, Testing Accuracy :
0.8449

  56%|██████        | 169/300 [38:19<32:09, 14.73s/it]

Training Loss : 0.13243183733463287, Testing Loss : 0.656114178276062,
Training Accuracy : 0.95314, Testing Accuracy : 0.8398

  57%|██████        | 170/300 [38:33<31:43, 14.64s/it]

Training Loss : 0.1269060293674469, Testing Loss : 0.6362595935344696,
Training Accuracy : 0.95526, Testing Accuracy : 0.8452

  57%|██████        | 171/300 [38:48<31:25, 14.62s/it]

Training Loss : 0.12949284333229064, Testing Loss :
0.6546784000873566, Training Accuracy : 0.9546, Testing Accuracy :
0.8416

  57%|██████        | 172/300 [39:02<31:04, 14.57s/it]

Training Loss : 0.12887051105499267, Testing Loss : 0.634709575510025,
Training Accuracy : 0.95384, Testing Accuracy : 0.8431

  58%|██████        | 173/300 [39:17<31:03, 14.68s/it]

Training Loss : 0.12939018103599548, Testing Loss :
0.6539512180805206, Training Accuracy : 0.95384, Testing Accuracy :
0.8431

  58%|██████        | 174/300 [39:32<30:53, 14.71s/it]

Training Loss : 0.12274213443279267, Testing Loss :
0.6457012056350708, Training Accuracy : 0.95732, Testing Accuracy :
0.8464

  58%|██████        | 175/300 [39:47<30:38, 14.71s/it]

Training Loss : 0.1209101472902298, Testing Loss : 0.6303320509433746,
Training Accuracy : 0.95634, Testing Accuracy : 0.8469

  59%|██████        | 176/300 [40:01<30:07, 14.57s/it]

Training Loss : 0.116859417719841, Testing Loss : 0.6508045325756073,
Training Accuracy : 0.95784, Testing Accuracy : 0.8484

  59%|██████        | 177/300 [40:15<29:52, 14.57s/it]
```

```
Training Loss : 0.11861840815544128, Testing Loss : 0.626621921825409,
Training Accuracy : 0.95796, Testing Accuracy : 0.8456

  59%|██████        | 178/300 [40:30<29:30, 14.51s/it]

Training Loss : 0.11792068127632141, Testing Loss :
0.6504782371520996, Training Accuracy : 0.95748, Testing Accuracy :
0.8438

  60%|██████        | 179/300 [40:44<29:18, 14.53s/it]

Training Loss : 0.11110170256614685, Testing Loss :
0.6193909126281738, Training Accuracy : 0.96004, Testing Accuracy :
0.8497

  60%|██████        | 180/300 [40:59<29:02, 14.52s/it]

Training Loss : 0.11015336357116699, Testing Loss :
0.6521210982322693, Training Accuracy : 0.96114, Testing Accuracy :
0.8513

  60%|██████        | 181/300 [41:13<28:52, 14.56s/it]

Training Loss : 0.10770075759410858, Testing Loss :
0.6434507028579712, Training Accuracy : 0.96124, Testing Accuracy :
0.8523

  61%|██████        | 182/300 [41:28<28:48, 14.65s/it]

Training Loss : 0.11145546368122101, Testing Loss :
0.6404320627689362, Training Accuracy : 0.96014, Testing Accuracy :
0.8473

  61%|██████        | 183/300 [41:43<28:48, 14.77s/it]

Training Loss : 0.10793096237659454, Testing Loss :
0.6528908896446228, Training Accuracy : 0.9624, Testing Accuracy :
0.8485

  61%|██████        | 184/300 [41:58<28:42, 14.85s/it]

Training Loss : 0.1080178373336792, Testing Loss : 0.6480017377853393,
Training Accuracy : 0.96216, Testing Accuracy : 0.8491

  62%|██████        | 185/300 [42:13<28:27, 14.85s/it]

Training Loss : 0.1020466248035431, Testing Loss : 0.6659204786777496,
Training Accuracy : 0.96376, Testing Accuracy : 0.8492

  62%|██████        | 186/300 [42:28<27:55, 14.70s/it]

Training Loss : 0.10308414425373078, Testing Loss :
0.6638622839450836, Training Accuracy : 0.96424, Testing Accuracy :
0.85
```

```
62%|██████     | 187/300 [42:43<27:52, 14.80s/it]
```

Training Loss : 0.10430046090126037, Testing Loss :
0.6492744675159454, Training Accuracy : 0.96304, Testing Accuracy :
0.8508

```
63%|██████     | 188/300 [42:58<27:41, 14.83s/it]
```

Training Loss : 0.10026016318798066, Testing Loss :
0.6591340884447098, Training Accuracy : 0.9653, Testing Accuracy :
0.8498

```
63%|██████     | 189/300 [43:12<27:25, 14.83s/it]
```

Training Loss : 0.09965462335586547, Testing Loss :
0.6498501077651978, Training Accuracy : 0.96388, Testing Accuracy :
0.8538

```
63%|██████     | 190/300 [43:28<27:22, 14.93s/it]
```

Training Loss : 0.09646633209228515, Testing Loss :
0.6592316916465759, Training Accuracy : 0.96556, Testing Accuracy :
0.8535

```
64%|██████     | 191/300 [43:42<27:04, 14.91s/it]
```

Training Loss : 0.09745890415668487, Testing Loss :
0.6619862708091736, Training Accuracy : 0.96542, Testing Accuracy :
0.8507

```
64%|██████     | 192/300 [43:57<26:46, 14.87s/it]
```

Training Loss : 0.09985521060466766, Testing Loss : 0.675312209367752,
Training Accuracy : 0.96474, Testing Accuracy : 0.8478

```
64%|██████     | 193/300 [44:12<26:18, 14.75s/it]
```

Training Loss : 0.10611237815856933, Testing Loss :
0.6730447945594787, Training Accuracy : 0.96232, Testing Accuracy :
0.8503

```
65%|██████     | 194/300 [44:26<25:57, 14.69s/it]
```

Training Loss : 0.09725945237636566, Testing Loss :
0.6723545650482178, Training Accuracy : 0.96544, Testing Accuracy :
0.8497

```
65%|██████     | 195/300 [44:41<25:41, 14.68s/it]
```

Training Loss : 0.09472614323616028, Testing Loss :
0.6759299190998077, Training Accuracy : 0.96662, Testing Accuracy :
0.8536

```
65%|██████     | 196/300 [44:55<25:23, 14.65s/it]
```

Training Loss : 0.09647743122577668, Testing Loss :
0.6757763199806214, Training Accuracy : 0.9653, Testing Accuracy :
0.848

   66%|██████ | 197/300 [45:10<25:20, 14.76s/it]

Training Loss : 0.09628610774993897, Testing Loss :
0.6722831575393676, Training Accuracy : 0.96488, Testing Accuracy :
0.8532

   66%|██████ | 198/300 [45:25<25:10, 14.80s/it]

Training Loss : 0.0910734720659256, Testing Loss : 0.6873094385147095,
Training Accuracy : 0.96842, Testing Accuracy : 0.8518

   66%|██████ | 199/300 [45:40<24:50, 14.75s/it]

Training Loss : 0.08984563920021058, Testing Loss : 0.672694333267212,
Training Accuracy : 0.96862, Testing Accuracy : 0.8537

   67%|██████ | 200/300 [45:55<24:53, 14.93s/it]

Training Loss : 0.09155048265457154, Testing Loss : 0.684890766954422,
Training Accuracy : 0.9679, Testing Accuracy : 0.8517

   67%|██████ | 201/300 [46:10<24:28, 14.83s/it]

Training Loss : 0.09015714128017425, Testing Loss :
0.6661509160041809, Training Accuracy : 0.96784, Testing Accuracy :
0.8555

   67%|██████ | 202/300 [46:25<24:11, 14.81s/it]

Training Loss : 0.08871263599395753, Testing Loss :
0.6854941876411438, Training Accuracy : 0.96882, Testing Accuracy :
0.8524

   68%|██████ | 203/300 [46:39<23:55, 14.79s/it]

Training Loss : 0.089218627409935, Testing Loss : 0.6822199995040894,
Training Accuracy : 0.96814, Testing Accuracy : 0.8549

   68%|██████ | 204/300 [46:54<23:26, 14.65s/it]

Training Loss : 0.087653777718544, Testing Loss : 0.6776058872222901,
Training Accuracy : 0.96842, Testing Accuracy : 0.8544

   68%|██████ | 205/300 [47:09<23:25, 14.79s/it]

Training Loss : 0.08432492959976197, Testing Loss :
0.7001365766525268, Training Accuracy : 0.9691, Testing Accuracy :
0.8485

   69%|██████ | 206/300 [47:24<23:13, 14.82s/it]

```
Training Loss : 0.08294587458133698, Testing Loss :
0.6696866943836212, Training Accuracy : 0.97038, Testing Accuracy :
0.8564

  69%|███████      | 207/300 [47:39<23:02, 14.87s/it]

Training Loss : 0.08043451214551926, Testing Loss :
0.6940341458797454, Training Accuracy : 0.9715, Testing Accuracy :
0.8505

  69%|███████      | 208/300 [47:53<22:44, 14.83s/it]

Training Loss : 0.07937085244894028, Testing Loss : 0.664235668516159,
Training Accuracy : 0.97236, Testing Accuracy : 0.8602

  70%|███████      | 209/300 [48:08<22:27, 14.80s/it]

Training Loss : 0.07656090586662292, Testing Loss :
0.6899644520759582, Training Accuracy : 0.9734, Testing Accuracy :
0.8509

  70%|███████      | 210/300 [48:24<22:26, 14.96s/it]

Training Loss : 0.0757552836894989, Testing Loss : 0.6735709754943847,
Training Accuracy : 0.9724, Testing Accuracy : 0.8578

  70%|███████      | 211/300 [48:38<22:02, 14.86s/it]

Training Loss : 0.07178021090984345, Testing Loss :
0.6874945259094238, Training Accuracy : 0.97512, Testing Accuracy :
0.8542

  71%|███████      | 212/300 [48:53<21:44, 14.82s/it]

Training Loss : 0.07172675694465637, Testing Loss :
0.6835818294048309, Training Accuracy : 0.97432, Testing Accuracy :
0.8586

  71%|███████      | 213/300 [49:08<21:32, 14.86s/it]

Training Loss : 0.06998863555192947, Testing Loss :
0.7012344113349914, Training Accuracy : 0.97534, Testing Accuracy :
0.8536

  71%|███████      | 214/300 [49:23<21:22, 14.92s/it]

Training Loss : 0.07134253517150879, Testing Loss :
0.6963598377227783, Training Accuracy : 0.97502, Testing Accuracy :
0.8559

  72%|███████      | 215/300 [49:38<21:02, 14.86s/it]
```

Training Loss : 0.07182141791224479, Testing Loss :
0.7094413270950317, Training Accuracy : 0.97474, Testing Accuracy :
0.853

 72%|███████ | 216/300 [49:52<20:45, 14.83s/it]

Training Loss : 0.07243164552927017, Testing Loss :
0.6958561646938324, Training Accuracy : 0.97486, Testing Accuracy :
0.8559

 72%|███████ | 217/300 [50:07<20:30, 14.82s/it]

Training Loss : 0.071869901995565887, Testing Loss :
0.7131956358909607, Training Accuracy : 0.97452, Testing Accuracy :
0.8513

 73%|███████ | 218/300 [50:22<20:18, 14.86s/it]

Training Loss : 0.0705226480293274, Testing Loss : 0.6897491882324219,
Training Accuracy : 0.97492, Testing Accuracy : 0.8553

 73%|███████ | 219/300 [50:37<20:03, 14.86s/it]

Training Loss : 0.06888839675188065, Testing Loss : 0.70850656914711,
Training Accuracy : 0.97582, Testing Accuracy : 0.8552

 73%|███████ | 220/300 [50:52<19:45, 14.82s/it]

Training Loss : 0.06832939162492752, Testing Loss :
0.6974722358226776, Training Accuracy : 0.97476, Testing Accuracy :
0.8534

 74%|███████ | 221/300 [51:07<19:33, 14.85s/it]

Training Loss : 0.06817810731172562, Testing Loss :
0.7142514698982239, Training Accuracy : 0.97614, Testing Accuracy :
0.854

 74%|███████ | 222/300 [51:22<19:26, 14.95s/it]

Training Loss : 0.06769857224702835, Testing Loss :
0.6969351161003112, Training Accuracy : 0.97632, Testing Accuracy :
0.859

 74%|███████ | 223/300 [51:36<19:03, 14.86s/it]

Training Loss : 0.06447814516544342, Testing Loss :
0.7155697449207306, Training Accuracy : 0.97704, Testing Accuracy :
0.8539

 75%|███████ | 224/300 [51:52<18:54, 14.93s/it]

Training Loss : 0.06692695227861405, Testing Loss :
0.7060830532550811, Training Accuracy : 0.97612, Testing Accuracy :
0.86

 75%|███████ | 225/300 [52:06<18:32, 14.84s/it]

Training Loss : 0.06582239810466767, Testing Loss :
0.7163696992874146, Training Accuracy : 0.97672, Testing Accuracy :
0.8557

 75%|███████ | 226/300 [52:21<18:22, 14.90s/it]

Training Loss : 0.06308882035017013, Testing Loss :
0.7098920697212219, Training Accuracy : 0.97742, Testing Accuracy :
0.8577

 76%|███████ | 227/300 [52:36<18:08, 14.91s/it]

Training Loss : 0.06362326840758324, Testing Loss :
0.7265035767555237, Training Accuracy : 0.97734, Testing Accuracy :
0.8548

 76%|███████ | 228/300 [52:51<17:42, 14.76s/it]

Training Loss : 0.06106389673233032, Testing Loss :
0.7082245206832886, Training Accuracy : 0.97812, Testing Accuracy :
0.8558

 76%|███████ | 229/300 [53:05<17:26, 14.74s/it]

Training Loss : 0.06142412614107132, Testing Loss :
0.7173233902931213, Training Accuracy : 0.97854, Testing Accuracy :
0.8557

 77%|███████ | 230/300 [53:20<17:10, 14.72s/it]

Training Loss : 0.062168482234478, Testing Loss : 0.7168272431373596,
Training Accuracy : 0.97812, Testing Accuracy : 0.8545

 77%|███████ | 231/300 [53:35<17:00, 14.79s/it]

Training Loss : 0.06259884811639786, Testing Loss :
0.7210296974182129, Training Accuracy : 0.97798, Testing Accuracy :
0.8557

 77%|███████ | 232/300 [53:50<16:49, 14.85s/it]

Training Loss : 0.06115842585325241, Testing Loss :
0.7168644368171692, Training Accuracy : 0.9784, Testing Accuracy :
0.8562

 78%|███████ | 233/300 [54:05<16:36, 14.87s/it]

Training Loss : 0.057756718019843105, Testing Loss :
0.7251059429168701, Training Accuracy : 0.9798, Testing Accuracy :
0.8573

 78%|██████ | 234/300 [54:20<16:25, 14.92s/it]

Training Loss : 0.05908210303783417, Testing Loss :
0.7283456604480744, Training Accuracy : 0.97928, Testing Accuracy :
0.8552

 78%|██████ | 235/300 [54:35<16:11, 14.95s/it]

Training Loss : 0.060831081923246384, Testing Loss :
0.7301415243148803, Training Accuracy : 0.97898, Testing Accuracy :
0.8546

 79%|██████ | 236/300 [54:50<15:50, 14.86s/it]

Training Loss : 0.060958557811975476, Testing Loss :
0.7173619790077209, Training Accuracy : 0.9793, Testing Accuracy :
0.8549

 79%|██████ | 237/300 [55:04<15:37, 14.88s/it]

Training Loss : 0.05770444059848785, Testing Loss :
0.7304026217460632, Training Accuracy : 0.97914, Testing Accuracy :
0.8536

 79%|██████ | 238/300 [55:19<15:24, 14.90s/it]

Training Loss : 0.06011219369769096, Testing Loss :
0.7266506915092468, Training Accuracy : 0.9793, Testing Accuracy :
0.8561

 80%|██████ | 239/300 [55:34<15:09, 14.91s/it]

Training Loss : 0.059105228433609006, Testing Loss :
0.7302264754295349, Training Accuracy : 0.97916, Testing Accuracy :
0.8559

 80%|██████ | 240/300 [55:49<14:51, 14.87s/it]

Training Loss : 0.054349636054039, Testing Loss : 0.7310868762493133,
Training Accuracy : 0.98036, Testing Accuracy : 0.8581

 80%|██████ | 241/300 [56:04<14:35, 14.83s/it]

Training Loss : 0.05675297975540161, Testing Loss :
0.7240343086719513, Training Accuracy : 0.97934, Testing Accuracy :
0.858

 81%|██████ | 242/300 [56:19<14:22, 14.88s/it]

Training Loss : 0.053550688799619675, Testing Loss :
0.7248839141845703, Training Accuracy : 0.98112, Testing Accuracy :
0.855

 81%|██████ | 243/300 [56:34<14:07, 14.87s/it]

Training Loss : 0.05468041951060295, Testing Loss :
0.7166750427722931, Training Accuracy : 0.98162, Testing Accuracy :
0.8604

 81%|██████ | 244/300 [56:49<13:52, 14.87s/it]

Training Loss : 0.05460075744390488, Testing Loss : 0.740034831905365,
Training Accuracy : 0.98112, Testing Accuracy : 0.8538

 82%|██████ | 245/300 [57:03<13:38, 14.89s/it]

Training Loss : 0.052439761881828306, Testing Loss :
0.7283154601097107, Training Accuracy : 0.98164, Testing Accuracy :
0.8568

 82%|██████ | 246/300 [57:18<13:22, 14.86s/it]

Training Loss : 0.054653376942873, Testing Loss : 0.733417223739624,
Training Accuracy : 0.98076, Testing Accuracy : 0.854

 82%|██████ | 247/300 [57:33<13:10, 14.92s/it]

Training Loss : 0.05257310071349144, Testing Loss :
0.7343039061546326, Training Accuracy : 0.98126, Testing Accuracy :
0.8585

 83%|██████ | 248/300 [57:49<13:05, 15.11s/it]

Training Loss : 0.051052484438419345, Testing Loss :
0.7345202083110809, Training Accuracy : 0.98262, Testing Accuracy :
0.8559

 83%|██████ | 249/300 [58:04<12:48, 15.07s/it]

Training Loss : 0.050006743710041046, Testing Loss :
0.7376179597854614, Training Accuracy : 0.98254, Testing Accuracy :
0.8568

 83%|██████ | 250/300 [58:19<12:30, 15.00s/it]

Training Loss : 0.05325740817308426, Testing Loss :
0.7374643845558166, Training Accuracy : 0.98092, Testing Accuracy :
0.8561

 84%|██████ | 251/300 [58:33<12:05, 14.80s/it]

Training Loss : 0.05353519106507301, Testing Loss :
0.7453715013027191, Training Accuracy : 0.98086, Testing Accuracy :
0.8584

84%|███████▋ | 252/300 [58:48<11:52, 14.84s/it]

Training Loss : 0.052183880029916765, Testing Loss :
0.7443199495792389, Training Accuracy : 0.98116, Testing Accuracy :
0.8553

84%|███████▋ | 253/300 [59:03<11:41, 14.92s/it]

Training Loss : 0.05017793232321739, Testing Loss :
0.7465089951038361, Training Accuracy : 0.98222, Testing Accuracy :
0.8592

85%|███████▋ | 254/300 [59:18<11:25, 14.90s/it]

Training Loss : 0.05132534186840058, Testing Loss :
0.7540961122512817, Training Accuracy : 0.98226, Testing Accuracy :
0.8533

85%|███████▋ | 255/300 [59:33<11:09, 14.88s/it]

Training Loss : 0.049834547491073605, Testing Loss :
0.7593983673572541, Training Accuracy : 0.98298, Testing Accuracy :
0.8574

85%|███████▋ | 256/300 [59:48<10:55, 14.89s/it]

Training Loss : 0.048963081922531126, Testing Loss :
0.757156030368805, Training Accuracy : 0.98258, Testing Accuracy :
0.8552

86%|███████▊ | 257/300 [1:00:03<10:42, 14.95s/it]

Training Loss : 0.04868201132774353, Testing Loss :
0.7610020372390747, Training Accuracy : 0.98228, Testing Accuracy :
0.857

86%|███████▊ | 258/300 [1:00:18<10:27, 14.93s/it]

Training Loss : 0.04934921106576919, Testing Loss :
0.7633592974662781, Training Accuracy : 0.98232, Testing Accuracy :
0.8546

86%|███████▊ | 259/300 [1:00:33<10:12, 14.93s/it]

Training Loss : 0.04895237189888954, Testing Loss :
0.7830078925609588, Training Accuracy : 0.98266, Testing Accuracy :
0.8544

87%|███████▊ | 260/300 [1:00:48<10:00, 15.00s/it]

Training Loss : 0.048833699547052384, Testing Loss :
0.7952121075630187, Training Accuracy : 0.98276, Testing Accuracy :
0.8508

 87%|███████| | 261/300 [1:01:03<09:45, 15.02s/it]

Training Loss : 0.051780835280418394, Testing Loss :
0.7743511520385742, Training Accuracy : 0.98124, Testing Accuracy :
0.8528

 87%|███████| | 262/300 [1:01:18<09:32, 15.06s/it]

Training Loss : 0.0513344853591919, Testing Loss : 0.7837949167251587,
Training Accuracy : 0.982, Testing Accuracy : 0.8544

 88%|███████| | 263/300 [1:01:33<09:12, 14.93s/it]

Training Loss : 0.05110806962370872, Testing Loss :
0.7679361395359039, Training Accuracy : 0.98186, Testing Accuracy :
0.8559

 88%|███████| | 264/300 [1:01:48<08:57, 14.93s/it]

Training Loss : 0.047787013564109804, Testing Loss : 0.77046684217453,
Training Accuracy : 0.98286, Testing Accuracy : 0.8546

 88%|███████| | 265/300 [1:02:02<08:40, 14.86s/it]

Training Loss : 0.04946080620408058, Testing Loss :
0.7625497789859772, Training Accuracy : 0.98252, Testing Accuracy :
0.8574

 89%|███████| | 266/300 [1:02:17<08:27, 14.94s/it]

Training Loss : 0.04965348205327988, Testing Loss :
0.7700312458515167, Training Accuracy : 0.9831, Testing Accuracy :
0.856

 89%|███████| | 267/300 [1:02:33<08:15, 15.03s/it]

Training Loss : 0.04616447902202606, Testing Loss :
0.7647084849834442, Training Accuracy : 0.98326, Testing Accuracy :
0.8571

 89%|███████| | 268/300 [1:02:47<07:56, 14.88s/it]

Training Loss : 0.04746733974218369, Testing Loss :
0.7697947658061981, Training Accuracy : 0.9836, Testing Accuracy :
0.8558

 90%|███████| | 269/300 [1:03:02<07:40, 14.86s/it]

Training Loss : 0.0462769946795702, Testing Loss : 0.7652537537574768,
Training Accuracy : 0.98388, Testing Accuracy : 0.8553

```
  90%|██████████        | 270/300 [1:03:17<07:27, 14.91s/it]
```

Training Loss : 0.044707206696271894, Testing Loss :
0.7717072954654693, Training Accuracy : 0.9843, Testing Accuracy :
0.8558

```
  90%|██████████        | 271/300 [1:03:32<07:17, 15.07s/it]
```

Training Loss : 0.04815471057057381, Testing Loss :
0.7652659547328949, Training Accuracy : 0.98292, Testing Accuracy :
0.8587

```
  91%|██████████        | 272/300 [1:03:48<07:05, 15.20s/it]
```

Training Loss : 0.045099790060053825, Testing Loss :
0.7829234031200409, Training Accuracy : 0.98414, Testing Accuracy :
0.854

```
  91%|██████████        | 273/300 [1:04:03<06:53, 15.30s/it]
```

Training Loss : 0.043659545291662216, Testing Loss :
0.7648932756900787, Training Accuracy : 0.98442, Testing Accuracy :
0.8594

```
  91%|██████████        | 274/300 [1:04:18<06:34, 15.18s/it]
```

Training Loss : 0.044680756689310076, Testing Loss :
0.7796842070102692, Training Accuracy : 0.98476, Testing Accuracy :
0.8548

```
  92%|██████████        | 275/300 [1:04:33<06:18, 15.12s/it]
```

Training Loss : 0.04488661799788475, Testing Loss :
0.7668601583480835, Training Accuracy : 0.98386, Testing Accuracy :
0.8582

```
  92%|██████████        | 276/300 [1:04:49<06:04, 15.18s/it]
```

Training Loss : 0.042297340418577194, Testing Loss :
0.7738072890281678, Training Accuracy : 0.98516, Testing Accuracy :
0.8564

```
  92%|██████████        | 277/300 [1:05:04<05:48, 15.14s/it]
```

Training Loss : 0.04433213582992554, Testing Loss :
0.7773398983955383, Training Accuracy : 0.98406, Testing Accuracy :
0.8589

```
  93%|██████████        | 278/300 [1:05:19<05:31, 15.09s/it]
```

Training Loss : 0.042936965987086295, Testing Loss :
0.7777125495910645, Training Accuracy : 0.98498, Testing Accuracy :
0.8552

```
 93%|████████      | 279/300 [1:05:34<05:16, 15.05s/it]
```

Training Loss : 0.04092326947569847, Testing Loss :
0.7706224782943726, Training Accuracy : 0.98576, Testing Accuracy :
0.857

```
 93%|████████      | 280/300 [1:05:49<05:01, 15.08s/it]
```

Training Loss : 0.040335978382229805, Testing Loss :
0.7792717337131501, Training Accuracy : 0.98594, Testing Accuracy :
0.8554

```
 94%|████████      | 281/300 [1:06:03<04:44, 14.97s/it]
```

Training Loss : 0.04096625518500805, Testing Loss :
0.7888147783756256, Training Accuracy : 0.9854, Testing Accuracy :
0.8575

```
 94%|████████      | 282/300 [1:06:18<04:28, 14.93s/it]
```

Training Loss : 0.041747714711427686, Testing Loss :
0.7777100477218628, Training Accuracy : 0.98516, Testing Accuracy :
0.859

```
 94%|████████      | 283/300 [1:06:34<04:15, 15.04s/it]
```

Training Loss : 0.04179693133950233, Testing Loss :
0.7853626253604888, Training Accuracy : 0.98518, Testing Accuracy :
0.8575

```
 95%|████████      | 284/300 [1:06:49<04:01, 15.11s/it]
```

Training Loss : 0.04378843834578991, Testing Loss :
0.7770103446960449, Training Accuracy : 0.98406, Testing Accuracy :
0.8566

```
 95%|████████      | 285/300 [1:07:04<03:45, 15.04s/it]
```

Training Loss : 0.04286963198006153, Testing Loss : 0.783554330778122,
Training Accuracy : 0.98536, Testing Accuracy : 0.8564

```
 95%|████████      | 286/300 [1:07:19<03:31, 15.09s/it]
```

Training Loss : 0.04092264357447624, Testing Loss :
0.7810811515808106, Training Accuracy : 0.98578, Testing Accuracy :
0.8579

```
 96%|████████      | 287/300 [1:07:34<03:17, 15.22s/it]
```

Training Loss : 0.04146460967779159, Testing Loss :
0.7853627054214477, Training Accuracy : 0.98554, Testing Accuracy :
0.8561

```
 96%|████████      | 288/300 [1:07:50<03:02, 15.22s/it]
```

```
Training Loss : 0.043423529076576234, Testing Loss :
0.7850516163825989, Training Accuracy : 0.98456, Testing Accuracy :
0.8545

   96%|████████| 289/300 [1:08:05<02:46, 15.17s/it]

Training Loss : 0.04348712749242783, Testing Loss :
0.7890497318267822, Training Accuracy : 0.98428, Testing Accuracy :
0.8563

   97%|████████| 290/300 [1:08:21<02:34, 15.40s/it]

Training Loss : 0.04204765565752983, Testing Loss :
0.7795873571395874, Training Accuracy : 0.98502, Testing Accuracy :
0.8594

   97%|████████| 291/300 [1:08:36<02:17, 15.26s/it]

Training Loss : 0.044330946266651156, Testing Loss :
0.799743632698059, Training Accuracy : 0.9849, Testing Accuracy :
0.8545

   97%|████████| 292/300 [1:08:51<02:01, 15.15s/it]

Training Loss : 0.043946774841547014, Testing Loss :
0.7882014881134033, Training Accuracy : 0.98448, Testing Accuracy :
0.858

   98%|████████| 293/300 [1:09:06<01:46, 15.19s/it]

Training Loss : 0.04145186347812414, Testing Loss :
0.7854585415840148, Training Accuracy : 0.9856, Testing Accuracy :
0.8583

   98%|████████| 294/300 [1:09:21<01:30, 15.17s/it]

Training Loss : 0.04157210846364498, Testing Loss : 0.784699764919281,
Training Accuracy : 0.98542, Testing Accuracy : 0.8596

   98%|████████| 295/300 [1:09:36<01:15, 15.09s/it]

Training Loss : 0.04232649885416031, Testing Loss :
0.7986132296562195, Training Accuracy : 0.98576, Testing Accuracy :
0.8561

   99%|████████| 296/300 [1:09:51<01:00, 15.12s/it]

Training Loss : 0.042640989879369734, Testing Loss :
0.7905191232681275, Training Accuracy : 0.98514, Testing Accuracy :
0.8593

   99%|████████| 297/300 [1:10:06<00:45, 15.15s/it]
```

Training Loss : 0.041933040608167646, Testing Loss :
0.8047542271614074, Training Accuracy : 0.98528, Testing Accuracy :
0.857

    99%|███████████| 298/300 [1:10:21<00:30, 15.02s/it]

Training Loss : 0.04135710584700108, Testing Loss :
0.7886473711967468, Training Accuracy : 0.98572, Testing Accuracy :
0.8592

   100%|███████████| 299/300 [1:10:36<00:15, 15.05s/it]

Training Loss : 0.040191376990675924, Testing Loss :
0.814676825428009, Training Accuracy : 0.9861, Testing Accuracy :
0.8551

100%|███████████| 300/300 [1:10:51<00:00, 14.17s/it]

Training Loss : 0.039542506706118585, Testing Loss :
0.7996593507766724, Training Accuracy : 0.98638, Testing Accuracy :
0.8574


```python
print("Maximum Testing Accuracy Achieved: %s"%(max(testingAccuracy)))
xmax = np.argmax(testingAccuracy)
ymax = max(testingAccuracy)
```

Maximum Testing Accuracy Achieved: 0.8604

```python
f, (ax1, ax2) = plt.subplots(1, 2, figsize = (20, 10))
n = len(trainingLoss)
ax1.plot(range(n), trainingLoss, '-', linewidth = '3', label = 'Train
Error')
ax1.plot(range(n), testingLoss, '-', linewidth = '3', label = 'Test
Error')
ax2.plot(range(n), trainingAccuracy, '-', linewidth = '3', label =
'Train Accuracy')
ax2.plot(range(n), testingAccuracy, '-', linewidth  = '3', label =
'Test Accuracy')
ax2.annotate('max accuracy = %s'%(ymax), xy = (xmax, ymax), xytext =
(xmax, ymax+0.15), arrowprops = dict(facecolor = 'black', shrink =
0.05))
ax1.grid(True)
ax2.grid(True)
ax1.legend()
ax2.legend()
f.savefig("./trainTestCurve.png")
```

Train Error
Test Error

Train Accuracy
Test Accuracy

max accuracy = 0.8604