MNIST_RNN_GRU_LSTM

October 25, 2023

1 MNIST_RNN

- ELEC 576 HW 2
- Robert Heeter
- 25 October 2023

1.1 Structure:

- Adjustments:
  - Number of nodes in the hidden layer
  - Learning rate
  - Number of iterations
  - Optimizer

1) Set PyTorch metadata
   - Seed
   - TensorFlow output
   - Whether to transfer to gpu (cuda)
2) Import data
   - Download data
   - Create data loaders with batchsize, transforms, scaling
3) Define model architecture, loss, and optimizer
4) Define test and training loops
   - Train:
     - Get next batch
     - Forward pass through model-
     - Calculate loss
     - Backward pass from loss (calculates the gradient for each parameter)
     - Optimizer: performs weight updates
5) Perform training over multiple epochs
   - Each epoch:
     - Call train loop
     - Call test loop

1.2 Acknowledgements:

- Worked with Arielle Sanford

```python
[1]: import torch
     import torch.nn as nn
     import torch.nn.functional as F
     import torch.optim as optim
     from torchvision import datasets, transforms
     from torch.autograd import Variable
     from torch.utils.data import DataLoader
     import numpy as np

     from torch.utils.tensorboard import SummaryWriter
     from datetime import datetime
     import os

     import matplotlib.pyplot as plt

     %load_ext tensorboard
```

2023-10-25 02:53:45.929462: I tensorflow/core/platform/cpu_feature_guard.cc:193]
This TensorFlow binary is optimized with oneAPI Deep Neural Network Library
(oneDNN) to use the following CPU instructions in performance-critical
operations: SSE4.1 SSE4.2 AVX AVX2 FMA
To enable them in other operations, rebuild TensorFlow with the appropriate
compiler flags.

```python
[58]: # 1. Set PyTorch metadata

      batch_size = 64
      test_batch_size = 1000
      epochs = 10
      lr = 0.001
      try_cuda = True
      seed = 1000
      logging_interval = 10 # how many batches to wait before logging
      logging_dir = None

      # setting up the logging
      log_dir = os.path.join(os.getcwd(),'log/MNIST', datetime.now().
       ↪strftime('%b%d_%H-%M-%S'))
      writer = SummaryWriter(log_dir=log_dir)

      # deciding whether to send to the cpu or not if available
      if torch.cuda.is_available() and try_cuda:
          cuda = True
          torch.cuda.mnaual_seed(seed)
      else:
          cuda = False
          torch.manual_seed(seed)
```

```python
# 2. Import data

transform = transforms.Compose([transforms.ToTensor(), transforms.Normalize((0.
 ↪1307,), (0.3081,))])

train_dataset = datasets.MNIST('data', train=True, transform=transform,
 ↪download=True)
test_dataset = datasets.MNIST('data', train=False, transform=transform,
 ↪download=True)

train_loader = DataLoader(dataset=train_dataset, batch_size=batch_size,
 ↪shuffle=True)
test_loader = DataLoader(dataset=test_dataset, batch_size=test_batch_size,
 ↪shuffle=False)

# example of data
print(train_dataset.train_data.size())
print(train_dataset.train_labels.size())

plt.imshow(train_dataset.train_data[0].numpy(), cmap='gray')
plt.title('%i' % train_dataset.train_labels[0])
plt.show()
```
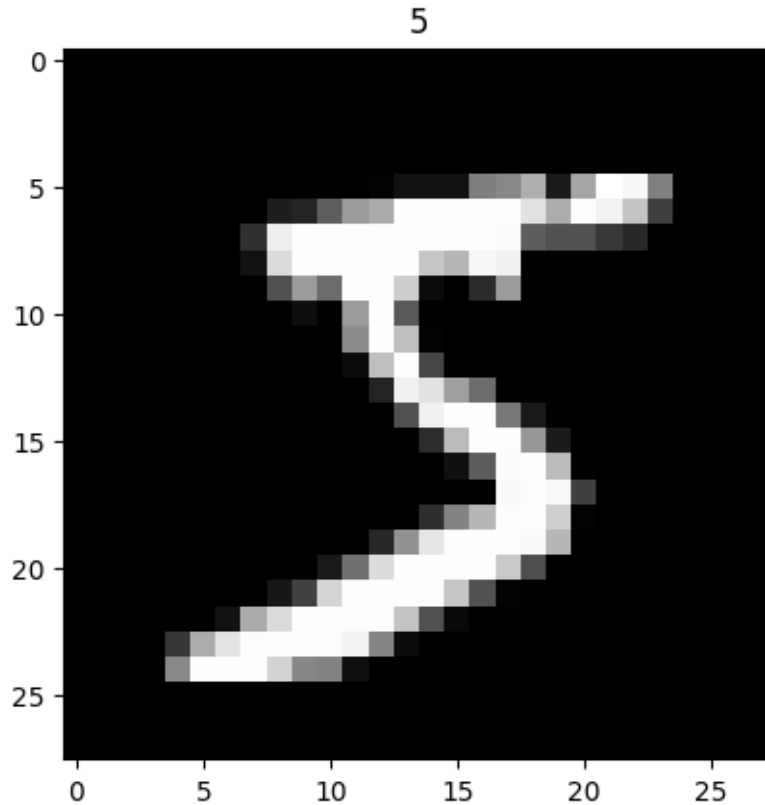
```
torch.Size([60000, 28, 28])
torch.Size([60000])
```

5

```python
# 3. Defining model architecture, loss, and optimizer

class Net(nn.Module):
    def __init__(self, model_type, hidden_size):
        super(Net, self).__init__()

        self.model_type = model_type
        self.hidden_size = hidden_size

        if self.model_type == 'RNN':
            # define RNN layer
            self.network = nn.RNN(
                input_size=28, # input size (number of features in each time
 ↪step)
                hidden_size=self.hidden_size, # number of hidden units
                num_layers=2, # number of RNN layers
                batch_first=True # input and output tensors are (batch,
 ↪time_step, input_size)
            )

        elif self.model_type == 'GRU':
```

```python
            # define GRU layer
            self.network = nn.GRU(
                input_size=28, # input size (number of features in each time⊔
→step)

                hidden_size=self.hidden_size, # number of hidden units
                num_layers=2, # number of RNN layers
                batch_first=True # input and output tensors are (batch,⊔
→time_step, input_size)
            )

        elif self.model_type == 'LSTM':
            # define LSTM layer
            self.network = nn.LSTM(
                input_size=28, # input size (number of features in each time⊔
→step)

                hidden_size=self.hidden_size, # number of hidden units
                num_layers=2, # number of RNN layers
                batch_first=True # input and output tensors are (batch,⊔
→time_step, input_size)
            )

        # define output layer
        self.out = nn.Linear(self.hidden_size, 10) # output size = 10

    def forward(self, x):
        n_out, hidden = self.network(x, None)
        return self.out(n_out[:, -1, :])

model = Net(model_type='LSTM', hidden_size=30)
print(model)

optimizer = optim.Adam(model.parameters(), lr=lr)
# optimizer = optim.SGD(model.parameters(), lr=lr)
print(optimizer)
```

```
Net(
  (network): LSTM(28, 30, num_layers=2, batch_first=True)
  (out): Linear(in_features=30, out_features=10, bias=True)
)
Adam (
Parameter Group 0
    amsgrad: False
    betas: (0.9, 0.999)
    capturable: False
    differentiable: False
    eps: 1e-08
    foreach: None
```

```
     fused: None
     lr: 0.001
     maximize: False
     weight_decay: 0
)
```

[61]:
```python
# 4. Define test and training loops

def train(epoch):
    model.train()

    criterion = nn.CrossEntropyLoss()
    total_loss = 0
    correct = 0

    for batch_idx, (data, target) in enumerate(train_loader):
        if cuda:
            data, target = data.cuda(), target.cuda()

        data = data.view(-1, 28, 28)

        optimizer.zero_grad()
        output = model(data) # forward pass
        loss = criterion(output, target)
        loss.backward() # backward pass
        optimizer.step()
        total_loss += loss.item()

        # calculate training accuracy
        _, pred = output.max(1)
        correct += pred.eq(target).sum().item()

        if batch_idx % logging_interval == 0:
            print(f'Train Epoch: {epoch} [{batch_idx * len(data)}/
 ↪{len(train_loader.dataset)} ({100. * batch_idx / len(train_loader):.
 ↪0f}%)]\tLoss: {loss.item()}')

    # calculate and log training metrics
    train_loss = total_loss / len(train_loader.dataset)
    train_accuracy = 100. * correct / len(train_loader.dataset)
    print(f"Epoch {epoch}: Training Loss: {train_loss:.4f}, Accuracy:␣
 ↪{train_accuracy:.2f}%")

    # add to TensorBoard
    writer.add_scalar('Loss/Train', train_loss, epoch)
    writer.add_scalar('Accuracy/Train', train_accuracy, epoch)
```

```python
def test(epoch):
    model.eval()
    test_loss = 0
    correct = 0

    criterion = nn.CrossEntropyLoss()

    with torch.no_grad():
        for data, target in test_loader:
            if cuda:
                data, target = data.cuda(), target.cuda()

            data = data.view(-1, 28, 28)
            output = model(data)

            test_loss += criterion(output, target,).item() # sum up batch loss␣
↪(later, averaged over all test samples)
            pred = output.data.max(1, keepdim=True)[1] # get index of max
            correct += pred.eq(target.data.view_as(pred)).cpu().sum()

    # calculate and log testing metrics
    test_loss /= len(test_loader.dataset)
    test_accuracy = 100. * correct / len(test_loader.dataset)
    print(f"Test Loss: {test_loss:.4f}, Accuracy: {test_accuracy:.2f}%")

    # add to TensorBoard
    writer.add_scalar('Accuracy/Test', test_accuracy, epoch)
```

```python
[62]: # 5. Perform training over multiple epochs

# start training
for epoch in range(1, epochs + 1):
    train(epoch)
    test(epoch)

writer.close()
```

```
Train Epoch: 1 [0/60000 (0%)]    Loss: 2.2945289611816406
Train Epoch: 1 [640/60000 (1%)] Loss: 2.310185432434082
Train Epoch: 1 [1280/60000 (2%)]       Loss: 2.2863149642944336
Train Epoch: 1 [1920/60000 (3%)]       Loss: 2.2622506618499756
Train Epoch: 1 [2560/60000 (4%)]       Loss: 2.1761507987976074
Train Epoch: 1 [3200/60000 (5%)]       Loss: 2.106879472732544
Train Epoch: 1 [3840/60000 (6%)]       Loss: 1.9779083728790283
Train Epoch: 1 [4480/60000 (7%)]       Loss: 1.9446831941604614
Train Epoch: 1 [5120/60000 (9%)]       Loss: 1.7644163370132446
```

```
Train Epoch: 1 [5760/60000 (10%)]      Loss: 1.6619151830673218
Train Epoch: 1 [6400/60000 (11%)]      Loss: 1.658410906791687
Train Epoch: 1 [7040/60000 (12%)]      Loss: 1.5282047986984253
Train Epoch: 1 [7680/60000 (13%)]      Loss: 1.394400954246521
Train Epoch: 1 [8320/60000 (14%)]      Loss: 1.3041701316833496
Train Epoch: 1 [8960/60000 (15%)]      Loss: 1.2269949913024902
Train Epoch: 1 [9600/60000 (16%)]      Loss: 1.3227688074111938
Train Epoch: 1 [10240/60000 (17%)]     Loss: 1.1708877086639404
Train Epoch: 1 [10880/60000 (18%)]     Loss: 1.0302032232284546
Train Epoch: 1 [11520/60000 (19%)]     Loss: 1.1463431119918823
Train Epoch: 1 [12160/60000 (20%)]     Loss: 0.8685747385025024
Train Epoch: 1 [12800/60000 (21%)]     Loss: 0.8419879674911499
Train Epoch: 1 [13440/60000 (22%)]     Loss: 0.9438325762748718
Train Epoch: 1 [14080/60000 (23%)]     Loss: 0.8639280796051025
Train Epoch: 1 [14720/60000 (25%)]     Loss: 0.7056488990783691
Train Epoch: 1 [15360/60000 (26%)]     Loss: 0.7904483675956726
Train Epoch: 1 [16000/60000 (27%)]     Loss: 0.9090250134468079
Train Epoch: 1 [16640/60000 (28%)]     Loss: 0.6252530813217163
Train Epoch: 1 [17280/60000 (29%)]     Loss: 0.71385657787323
Train Epoch: 1 [17920/60000 (30%)]     Loss: 0.6772246360778809
Train Epoch: 1 [18560/60000 (31%)]     Loss: 0.5937366485595703
Train Epoch: 1 [19200/60000 (32%)]     Loss: 0.5778682827949524
Train Epoch: 1 [19840/60000 (33%)]     Loss: 0.4683002233505249
Train Epoch: 1 [20480/60000 (34%)]     Loss: 0.5364527702331543
Train Epoch: 1 [21120/60000 (35%)]     Loss: 0.4667869508266449
Train Epoch: 1 [21760/60000 (36%)]     Loss: 0.6446887850761414
Train Epoch: 1 [22400/60000 (37%)]     Loss: 0.6362211108207703
Train Epoch: 1 [23040/60000 (38%)]     Loss: 0.5622038245201111
Train Epoch: 1 [23680/60000 (39%)]     Loss: 0.4937388002872467
Train Epoch: 1 [24320/60000 (41%)]     Loss: 0.5408675670623779
Train Epoch: 1 [24960/60000 (42%)]     Loss: 0.3813870847225189
Train Epoch: 1 [25600/60000 (43%)]     Loss: 0.3680822551250458
Train Epoch: 1 [26240/60000 (44%)]     Loss: 0.4595952332019806
Train Epoch: 1 [26880/60000 (45%)]     Loss: 0.35872092843055725
Train Epoch: 1 [27520/60000 (46%)]     Loss: 0.38749030232429504
Train Epoch: 1 [28160/60000 (47%)]     Loss: 0.3334735929965973
Train Epoch: 1 [28800/60000 (48%)]     Loss: 0.3708879351615906
Train Epoch: 1 [29440/60000 (49%)]     Loss: 0.3144991993904114
Train Epoch: 1 [30080/60000 (50%)]     Loss: 0.4568016529083252
Train Epoch: 1 [30720/60000 (51%)]     Loss: 0.39923760294914246
Train Epoch: 1 [31360/60000 (52%)]     Loss: 0.29439374804496765
Train Epoch: 1 [32000/60000 (53%)]     Loss: 0.2720756530761719
Train Epoch: 1 [32640/60000 (54%)]     Loss: 0.4548777937889099
Train Epoch: 1 [33280/60000 (55%)]     Loss: 0.31229597330093384
Train Epoch: 1 [33920/60000 (57%)]     Loss: 0.4307449460029602
Train Epoch: 1 [34560/60000 (58%)]     Loss: 0.3777798116207123
Train Epoch: 1 [35200/60000 (59%)]     Loss: 0.20632891356945038
Train Epoch: 1 [35840/60000 (60%)]     Loss: 0.31650620698928833
```

```
Train Epoch: 1 [36480/60000 (61%)]        Loss: 0.348665714263916
Train Epoch: 1 [37120/60000 (62%)]        Loss: 0.4276980757713318
Train Epoch: 1 [37760/60000 (63%)]        Loss: 0.3171851634979248
Train Epoch: 1 [38400/60000 (64%)]        Loss: 0.3350239098072052
Train Epoch: 1 [39040/60000 (65%)]        Loss: 0.20009705424308777
Train Epoch: 1 [39680/60000 (66%)]        Loss: 0.33397796750068665
Train Epoch: 1 [40320/60000 (67%)]        Loss: 0.3542984426021576
Train Epoch: 1 [40960/60000 (68%)]        Loss: 0.39755144715309143
Train Epoch: 1 [41600/60000 (69%)]        Loss: 0.20689691603183746
Train Epoch: 1 [42240/60000 (70%)]        Loss: 0.3371772766113281
Train Epoch: 1 [42880/60000 (71%)]        Loss: 0.26389288902282715
Train Epoch: 1 [43520/60000 (72%)]        Loss: 0.36912408471107483
Train Epoch: 1 [44160/60000 (74%)]        Loss: 0.29359832406044006
Train Epoch: 1 [44800/60000 (75%)]        Loss: 0.36473971605300903
Train Epoch: 1 [45440/60000 (76%)]        Loss: 0.25217390060424805
Train Epoch: 1 [46080/60000 (77%)]        Loss: 0.13130666315555573
Train Epoch: 1 [46720/60000 (78%)]        Loss: 0.16144095361232758
Train Epoch: 1 [47360/60000 (79%)]        Loss: 0.38100963830947876
Train Epoch: 1 [48000/60000 (80%)]        Loss: 0.3219181299209595
Train Epoch: 1 [48640/60000 (81%)]        Loss: 0.303625226020813
Train Epoch: 1 [49280/60000 (82%)]        Loss: 0.20752952992916107
Train Epoch: 1 [49920/60000 (83%)]        Loss: 0.28156477212905884
Train Epoch: 1 [50560/60000 (84%)]        Loss: 0.18706314265727997
Train Epoch: 1 [51200/60000 (85%)]        Loss: 0.21575868129730225
Train Epoch: 1 [51840/60000 (86%)]        Loss: 0.18342538177967072
Train Epoch: 1 [52480/60000 (87%)]        Loss: 0.23085516691207886
Train Epoch: 1 [53120/60000 (88%)]        Loss: 0.4470926821231842
Train Epoch: 1 [53760/60000 (90%)]        Loss: 0.1286468803882599
Train Epoch: 1 [54400/60000 (91%)]        Loss: 0.2753753960132599
Train Epoch: 1 [55040/60000 (92%)]        Loss: 0.3555644750595093
Train Epoch: 1 [55680/60000 (93%)]        Loss: 0.22616390883922577
Train Epoch: 1 [56320/60000 (94%)]        Loss: 0.21610042452812195
Train Epoch: 1 [56960/60000 (95%)]        Loss: 0.1126556470990181
Train Epoch: 1 [57600/60000 (96%)]        Loss: 0.13718624413013458
Train Epoch: 1 [58240/60000 (97%)]        Loss: 0.15010002255439758
Train Epoch: 1 [58880/60000 (98%)]        Loss: 0.21739205718040466
Train Epoch: 1 [59520/60000 (99%)]        Loss: 0.054770808666944504
Epoch 1: Training Loss: 0.0104, Accuracy: 78.67%
Test Loss: 0.0002, Accuracy: 94.16%
Train Epoch: 2 [0/60000 (0%)]   Loss: 0.17198677361011505
Train Epoch: 2 [640/60000 (1%)] Loss: 0.3094375431537628
Train Epoch: 2 [1280/60000 (2%)]        Loss: 0.37173157930374146
Train Epoch: 2 [1920/60000 (3%)]        Loss: 0.37021562457084656
Train Epoch: 2 [2560/60000 (4%)]        Loss: 0.2392449676990509
Train Epoch: 2 [3200/60000 (5%)]        Loss: 0.5128071904182434
Train Epoch: 2 [3840/60000 (6%)]        Loss: 0.19455893337726593
Train Epoch: 2 [4480/60000 (7%)]        Loss: 0.20626495778560638
Train Epoch: 2 [5120/60000 (9%)]        Loss: 0.12003334611654282
```

```
Train Epoch: 2 [5760/60000 (10%)]      Loss: 0.30628490447998047
Train Epoch: 2 [6400/60000 (11%)]      Loss: 0.05124559998512268
Train Epoch: 2 [7040/60000 (12%)]      Loss: 0.247537299990654
Train Epoch: 2 [7680/60000 (13%)]      Loss: 0.13727982342243195
Train Epoch: 2 [8320/60000 (14%)]      Loss: 0.24304601550102234
Train Epoch: 2 [8960/60000 (15%)]      Loss: 0.0846015214920044
Train Epoch: 2 [9600/60000 (16%)]      Loss: 0.2253648191690445
Train Epoch: 2 [10240/60000 (17%)]     Loss: 0.17394313216209412
Train Epoch: 2 [10880/60000 (18%)]     Loss: 0.23230072855949402
Train Epoch: 2 [11520/60000 (19%)]     Loss: 0.2246512472629547
Train Epoch: 2 [12160/60000 (20%)]     Loss: 0.3746359944343567
Train Epoch: 2 [12800/60000 (21%)]     Loss: 0.25469622015953064
Train Epoch: 2 [13440/60000 (22%)]     Loss: 0.2611026167869568
Train Epoch: 2 [14080/60000 (23%)]     Loss: 0.1535918265581131
Train Epoch: 2 [14720/60000 (25%)]     Loss: 0.34096282720565796
Train Epoch: 2 [15360/60000 (26%)]     Loss: 0.36082029342651367
Train Epoch: 2 [16000/60000 (27%)]     Loss: 0.2596152126789093
Train Epoch: 2 [16640/60000 (28%)]     Loss: 0.09666206687688828
Train Epoch: 2 [17280/60000 (29%)]     Loss: 0.17576637864112854
Train Epoch: 2 [17920/60000 (30%)]     Loss: 0.23635660111904144
Train Epoch: 2 [18560/60000 (31%)]     Loss: 0.1287546306848526
Train Epoch: 2 [19200/60000 (32%)]     Loss: 0.1502864509820938
Train Epoch: 2 [19840/60000 (33%)]     Loss: 0.09590309858322144
Train Epoch: 2 [20480/60000 (34%)]     Loss: 0.18120957911014557
Train Epoch: 2 [21120/60000 (35%)]     Loss: 0.2322518229484558
Train Epoch: 2 [21760/60000 (36%)]     Loss: 0.24659237265586853
Train Epoch: 2 [22400/60000 (37%)]     Loss: 0.12524837255477905
Train Epoch: 2 [23040/60000 (38%)]     Loss: 0.09233067184686661
Train Epoch: 2 [23680/60000 (39%)]     Loss: 0.06369709223508835
Train Epoch: 2 [24320/60000 (41%)]     Loss: 0.09215261042118073
Train Epoch: 2 [24960/60000 (42%)]     Loss: 0.09633141756057739
Train Epoch: 2 [25600/60000 (43%)]     Loss: 0.12268120795488358
Train Epoch: 2 [26240/60000 (44%)]     Loss: 0.11106028407812119
Train Epoch: 2 [26880/60000 (45%)]     Loss: 0.22345808148384094
Train Epoch: 2 [27520/60000 (46%)]     Loss: 0.13616792857646942
Train Epoch: 2 [28160/60000 (47%)]     Loss: 0.04196320101618767
Train Epoch: 2 [28800/60000 (48%)]     Loss: 0.11789335310459137
Train Epoch: 2 [29440/60000 (49%)]     Loss: 0.3281185030937195
Train Epoch: 2 [30080/60000 (50%)]     Loss: 0.15844587981700897
Train Epoch: 2 [30720/60000 (51%)]     Loss: 0.3856131136417389
Train Epoch: 2 [31360/60000 (52%)]     Loss: 0.09819116443395615
Train Epoch: 2 [32000/60000 (53%)]     Loss: 0.1261643022298813
Train Epoch: 2 [32640/60000 (54%)]     Loss: 0.13869211077690125
Train Epoch: 2 [33280/60000 (55%)]     Loss: 0.05392887443304062
Train Epoch: 2 [33920/60000 (57%)]     Loss: 0.3070884943008423
Train Epoch: 2 [34560/60000 (58%)]     Loss: 0.2714058756828308
Train Epoch: 2 [35200/60000 (59%)]     Loss: 0.0865621417760849
Train Epoch: 2 [35840/60000 (60%)]     Loss: 0.1491464227437973
```

```
Train Epoch: 2 [36480/60000 (61%)]      Loss: 0.10783379524946213
Train Epoch: 2 [37120/60000 (62%)]      Loss: 0.1586132049560547
Train Epoch: 2 [37760/60000 (63%)]      Loss: 0.10976088047027588
Train Epoch: 2 [38400/60000 (64%)]      Loss: 0.10003693401813507
Train Epoch: 2 [39040/60000 (65%)]      Loss: 0.056936223059892654
Train Epoch: 2 [39680/60000 (66%)]      Loss: 0.1046886146068573
Train Epoch: 2 [40320/60000 (67%)]      Loss: 0.12938499450683594
Train Epoch: 2 [40960/60000 (68%)]      Loss: 0.2423151284456253
Train Epoch: 2 [41600/60000 (69%)]      Loss: 0.1270882487297058
Train Epoch: 2 [42240/60000 (70%)]      Loss: 0.18772421777248383
Train Epoch: 2 [42880/60000 (71%)]      Loss: 0.1252443492412567
Train Epoch: 2 [43520/60000 (72%)]      Loss: 0.12931403517723083
Train Epoch: 2 [44160/60000 (74%)]      Loss: 0.18881607055664062
Train Epoch: 2 [44800/60000 (75%)]      Loss: 0.0976172685623169
Train Epoch: 2 [45440/60000 (76%)]      Loss: 0.1074165552854538
Train Epoch: 2 [46080/60000 (77%)]      Loss: 0.13705863058567047
Train Epoch: 2 [46720/60000 (78%)]      Loss: 0.25016841292381287
Train Epoch: 2 [47360/60000 (79%)]      Loss: 0.18035393953323364
Train Epoch: 2 [48000/60000 (80%)]      Loss: 0.300092488527298
Train Epoch: 2 [48640/60000 (81%)]      Loss: 0.15651419758796692
Train Epoch: 2 [49280/60000 (82%)]      Loss: 0.22690904140472412
Train Epoch: 2 [49920/60000 (83%)]      Loss: 0.1605144441127777
Train Epoch: 2 [50560/60000 (84%)]      Loss: 0.1600484699010849
Train Epoch: 2 [51200/60000 (85%)]      Loss: 0.06499714404344559
Train Epoch: 2 [51840/60000 (86%)]      Loss: 0.05809737741947174
Train Epoch: 2 [52480/60000 (87%)]      Loss: 0.1119866594672203
Train Epoch: 2 [53120/60000 (88%)]      Loss: 0.15728387236595154
Train Epoch: 2 [53760/60000 (90%)]      Loss: 0.06590581685304642
Train Epoch: 2 [54400/60000 (91%)]      Loss: 0.3351259231567383
Train Epoch: 2 [55040/60000 (92%)]      Loss: 0.22556564211845398
Train Epoch: 2 [55680/60000 (93%)]      Loss: 0.35566338896751404
Train Epoch: 2 [56320/60000 (94%)]      Loss: 0.2130226492881775
Train Epoch: 2 [56960/60000 (95%)]      Loss: 0.08642593026161194
Train Epoch: 2 [57600/60000 (96%)]      Loss: 0.055187977850437164
Train Epoch: 2 [58240/60000 (97%)]      Loss: 0.11076291650533676
Train Epoch: 2 [58880/60000 (98%)]      Loss: 0.25481611490249634
Train Epoch: 2 [59520/60000 (99%)]      Loss: 0.18853631615638733
Epoch 2: Training Loss: 0.0027, Accuracy: 95.01%
Test Loss: 0.0001, Accuracy: 95.80%
Train Epoch: 3 [0/60000 (0%)]   Loss: 0.15060633420944214
Train Epoch: 3 [640/60000 (1%)] Loss: 0.09614292532205582
Train Epoch: 3 [1280/60000 (2%)]        Loss: 0.07992998510599136
Train Epoch: 3 [1920/60000 (3%)]        Loss: 0.07647403329610825
Train Epoch: 3 [2560/60000 (4%)]        Loss: 0.092201537072658539
Train Epoch: 3 [3200/60000 (5%)]        Loss: 0.2806614637374878
Train Epoch: 3 [3840/60000 (6%)]        Loss: 0.06280981749296188
Train Epoch: 3 [4480/60000 (7%)]        Loss: 0.10447729378938675
Train Epoch: 3 [5120/60000 (9%)]        Loss: 0.2270536571741104
```

```
Train Epoch: 3 [5760/60000 (10%)]      Loss: 0.05435444042086601
Train Epoch: 3 [6400/60000 (11%)]      Loss: 0.12420186400413513
Train Epoch: 3 [7040/60000 (12%)]      Loss: 0.1809394657611847
Train Epoch: 3 [7680/60000 (13%)]      Loss: 0.11447129398584366
Train Epoch: 3 [8320/60000 (14%)]      Loss: 0.16642069816589355
Train Epoch: 3 [8960/60000 (15%)]      Loss: 0.1060926765203476
Train Epoch: 3 [9600/60000 (16%)]      Loss: 0.11041676998138428
Train Epoch: 3 [10240/60000 (17%)]     Loss: 0.20007653534412384
Train Epoch: 3 [10880/60000 (18%)]     Loss: 0.1485455334186554
Train Epoch: 3 [11520/60000 (19%)]     Loss: 0.12410756945610046
Train Epoch: 3 [12160/60000 (20%)]     Loss: 0.06727205961942673
Train Epoch: 3 [12800/60000 (21%)]     Loss: 0.2564881145954132
Train Epoch: 3 [13440/60000 (22%)]     Loss: 0.049380719661712646
Train Epoch: 3 [14080/60000 (23%)]     Loss: 0.2152499407529831
Train Epoch: 3 [14720/60000 (25%)]     Loss: 0.16822238266468048
Train Epoch: 3 [15360/60000 (26%)]     Loss: 0.15397536754608154
Train Epoch: 3 [16000/60000 (27%)]     Loss: 0.1881684958934784
Train Epoch: 3 [16640/60000 (28%)]     Loss: 0.22968095541000366
Train Epoch: 3 [17280/60000 (29%)]     Loss: 0.07307745516300201
Train Epoch: 3 [17920/60000 (30%)]     Loss: 0.12444467842578888
Train Epoch: 3 [18560/60000 (31%)]     Loss: 0.02417365461587906
Train Epoch: 3 [19200/60000 (32%)]     Loss: 0.10296367108821869
Train Epoch: 3 [19840/60000 (33%)]     Loss: 0.04851500317454338
Train Epoch: 3 [20480/60000 (34%)]     Loss: 0.1246711015701294
Train Epoch: 3 [21120/60000 (35%)]     Loss: 0.14681239426136017
Train Epoch: 3 [21760/60000 (36%)]     Loss: 0.10299243777990341
Train Epoch: 3 [22400/60000 (37%)]     Loss: 0.13993243873119354
Train Epoch: 3 [23040/60000 (38%)]     Loss: 0.1167730912566185
Train Epoch: 3 [23680/60000 (39%)]     Loss: 0.1692352145910263
Train Epoch: 3 [24320/60000 (41%)]     Loss: 0.15797856450080872
Train Epoch: 3 [24960/60000 (42%)]     Loss: 0.05923561379313469
Train Epoch: 3 [25600/60000 (43%)]     Loss: 0.11033082753419876
Train Epoch: 3 [26240/60000 (44%)]     Loss: 0.2330516129732132
Train Epoch: 3 [26880/60000 (45%)]     Loss: 0.06316038221120834
Train Epoch: 3 [27520/60000 (46%)]     Loss: 0.12287494540214539
Train Epoch: 3 [28160/60000 (47%)]     Loss: 0.12058781832456589
Train Epoch: 3 [28800/60000 (48%)]     Loss: 0.04870736971497536
Train Epoch: 3 [29440/60000 (49%)]     Loss: 0.2565833032131195
Train Epoch: 3 [30080/60000 (50%)]     Loss: 0.010767338797450066
Train Epoch: 3 [30720/60000 (51%)]     Loss: 0.191471129655838
Train Epoch: 3 [31360/60000 (52%)]     Loss: 0.041955333203077316
Train Epoch: 3 [32000/60000 (53%)]     Loss: 0.2200736403465271
Train Epoch: 3 [32640/60000 (54%)]     Loss: 0.05864259600639343
Train Epoch: 3 [33280/60000 (55%)]     Loss: 0.2812963128089905
Train Epoch: 3 [33920/60000 (57%)]     Loss: 0.084490105509758
Train Epoch: 3 [34560/60000 (58%)]     Loss: 0.05661746487021446
Train Epoch: 3 [35200/60000 (59%)]     Loss: 0.1367335170507431
Train Epoch: 3 [35840/60000 (60%)]     Loss: 0.11972302943468094
```

```
Train Epoch: 3 [36480/60000 (61%)]      Loss: 0.04546009749174118
Train Epoch: 3 [37120/60000 (62%)]      Loss: 0.14182248711585999
Train Epoch: 3 [37760/60000 (63%)]      Loss: 0.03713281452655792
Train Epoch: 3 [38400/60000 (64%)]      Loss: 0.02412562631070614
Train Epoch: 3 [39040/60000 (65%)]      Loss: 0.07307452708482742
Train Epoch: 3 [39680/60000 (66%)]      Loss: 0.1917893886566162
Train Epoch: 3 [40320/60000 (67%)]      Loss: 0.17093464732170105
Train Epoch: 3 [40960/60000 (68%)]      Loss: 0.03985047712922096
Train Epoch: 3 [41600/60000 (69%)]      Loss: 0.10221493244171143
Train Epoch: 3 [42240/60000 (70%)]      Loss: 0.1426333785057068
Train Epoch: 3 [42880/60000 (71%)]      Loss: 0.1420499086380005
Train Epoch: 3 [43520/60000 (72%)]      Loss: 0.1376029998064041
Train Epoch: 3 [44160/60000 (74%)]      Loss: 0.0512627474963665
Train Epoch: 3 [44800/60000 (75%)]      Loss: 0.14062896370887756
Train Epoch: 3 [45440/60000 (76%)]      Loss: 0.061785969883203506
Train Epoch: 3 [46080/60000 (77%)]      Loss: 0.08868647366762161
Train Epoch: 3 [46720/60000 (78%)]      Loss: 0.13078659772872925
Train Epoch: 3 [47360/60000 (79%)]      Loss: 0.1261131465435028
Train Epoch: 3 [48000/60000 (80%)]      Loss: 0.012385232374072075
Train Epoch: 3 [48640/60000 (81%)]      Loss: 0.168213576078411492
Train Epoch: 3 [49280/60000 (82%)]      Loss: 0.20297667384147644
Train Epoch: 3 [49920/60000 (83%)]      Loss: 0.098629161171550751
Train Epoch: 3 [50560/60000 (84%)]      Loss: 0.1471257358789444
Train Epoch: 3 [51200/60000 (85%)]      Loss: 0.17215946316719055
Train Epoch: 3 [51840/60000 (86%)]      Loss: 0.09612949937582016
Train Epoch: 3 [52480/60000 (87%)]      Loss: 0.24351347982883453
Train Epoch: 3 [53120/60000 (88%)]      Loss: 0.2068229615688324
Train Epoch: 3 [53760/60000 (90%)]      Loss: 0.1578577160835266
Train Epoch: 3 [54400/60000 (91%)]      Loss: 0.07384596019983292
Train Epoch: 3 [55040/60000 (92%)]      Loss: 0.2359781116247177
Train Epoch: 3 [55680/60000 (93%)]      Loss: 0.11157127469778061
Train Epoch: 3 [56320/60000 (94%)]      Loss: 0.060908567160367966
Train Epoch: 3 [56960/60000 (95%)]      Loss: 0.09585653245449066
Train Epoch: 3 [57600/60000 (96%)]      Loss: 0.08047093451023102
Train Epoch: 3 [58240/60000 (97%)]      Loss: 0.072468183393468857
Train Epoch: 3 [58880/60000 (98%)]      Loss: 0.06716294586658478
Train Epoch: 3 [59520/60000 (99%)]      Loss: 0.018191123381257057
Epoch 3: Training Loss: 0.0019, Accuracy: 96.52%
Test Loss: 0.0001, Accuracy: 96.22%
Train Epoch: 4 [0/60000 (0%)]   Loss: 0.15940910577774048
Train Epoch: 4 [640/60000 (1%)] Loss: 0.20875060558319092
Train Epoch: 4 [1280/60000 (2%)]        Loss: 0.12492070347070694
Train Epoch: 4 [1920/60000 (3%)]        Loss: 0.08392620831727982
Train Epoch: 4 [2560/60000 (4%)]        Loss: 0.09796911478042603
Train Epoch: 4 [3200/60000 (5%)]        Loss: 0.11695822328329086
Train Epoch: 4 [3840/60000 (6%)]        Loss: 0.19431817531585693
Train Epoch: 4 [4480/60000 (7%)]        Loss: 0.21295081079006195
Train Epoch: 4 [5120/60000 (9%)]        Loss: 0.09320583939552307
```

```
Train Epoch: 4 [5760/60000 (10%)]       Loss: 0.18956762552261353
Train Epoch: 4 [6400/60000 (11%)]       Loss: 0.07055515050888062
Train Epoch: 4 [7040/60000 (12%)]       Loss: 0.1048620194196701
Train Epoch: 4 [7680/60000 (13%)]       Loss: 0.08666314929723740
Train Epoch: 4 [8320/60000 (14%)]       Loss: 0.3239706754684448
Train Epoch: 4 [8960/60000 (15%)]       Loss: 0.06611952930688858
Train Epoch: 4 [9600/60000 (16%)]       Loss: 0.08727412670850754
Train Epoch: 4 [10240/60000 (17%)]      Loss: 0.05782540515065193
Train Epoch: 4 [10880/60000 (18%)]      Loss: 0.1134781613945961
Train Epoch: 4 [11520/60000 (19%)]      Loss: 0.04172533005475998
Train Epoch: 4 [12160/60000 (20%)]      Loss: 0.01716267690062523
Train Epoch: 4 [12800/60000 (21%)]      Loss: 0.07109201699495316
Train Epoch: 4 [13440/60000 (22%)]      Loss: 0.16506539285182953
Train Epoch: 4 [14080/60000 (23%)]      Loss: 0.08235379308462143
Train Epoch: 4 [14720/60000 (25%)]      Loss: 0.028616663068532944
Train Epoch: 4 [15360/60000 (26%)]      Loss: 0.19303584098815918
Train Epoch: 4 [16000/60000 (27%)]      Loss: 0.08153948187828064
Train Epoch: 4 [16640/60000 (28%)]      Loss: 0.11752806603908539
Train Epoch: 4 [17280/60000 (29%)]      Loss: 0.1499924212694168
Train Epoch: 4 [17920/60000 (30%)]      Loss: 0.12204084545373917
Train Epoch: 4 [18560/60000 (31%)]      Loss: 0.026202861219644547
Train Epoch: 4 [19200/60000 (32%)]      Loss: 0.13494209945201874
Train Epoch: 4 [19840/60000 (33%)]      Loss: 0.09143088012933731
Train Epoch: 4 [20480/60000 (34%)]      Loss: 0.09133370965719223
Train Epoch: 4 [21120/60000 (35%)]      Loss: 0.07926234602928162
Train Epoch: 4 [21760/60000 (36%)]      Loss: 0.06576032936573029
Train Epoch: 4 [22400/60000 (37%)]      Loss: 0.29091182351112366
Train Epoch: 4 [23040/60000 (38%)]      Loss: 0.188930481672287
Train Epoch: 4 [23680/60000 (39%)]      Loss: 0.06918118894100189
Train Epoch: 4 [24320/60000 (41%)]      Loss: 0.08810659497976303
Train Epoch: 4 [24960/60000 (42%)]      Loss: 0.23312894999980927
Train Epoch: 4 [25600/60000 (43%)]      Loss: 0.040479306131601334
Train Epoch: 4 [26240/60000 (44%)]      Loss: 0.1252056360244751
Train Epoch: 4 [26880/60000 (45%)]      Loss: 0.026117239147424698
Train Epoch: 4 [27520/60000 (46%)]      Loss: 0.09939195215702057
Train Epoch: 4 [28160/60000 (47%)]      Loss: 0.03251553326845169
Train Epoch: 4 [28800/60000 (48%)]      Loss: 0.1276405304670334
Train Epoch: 4 [29440/60000 (49%)]      Loss: 0.12429890781641006
Train Epoch: 4 [30080/60000 (50%)]      Loss: 0.059666816145181656
Train Epoch: 4 [30720/60000 (51%)]      Loss: 0.04693400487303734
Train Epoch: 4 [31360/60000 (52%)]      Loss: 0.07163196802139282
Train Epoch: 4 [32000/60000 (53%)]      Loss: 0.027555545791983604
Train Epoch: 4 [32640/60000 (54%)]      Loss: 0.0571870990097522274
Train Epoch: 4 [33280/60000 (55%)]      Loss: 0.1420605182647705
Train Epoch: 4 [33920/60000 (57%)]      Loss: 0.038578033447265625
Train Epoch: 4 [34560/60000 (58%)]      Loss: 0.061884570866823196
Train Epoch: 4 [35200/60000 (59%)]      Loss: 0.16838733851909637
Train Epoch: 4 [35840/60000 (60%)]      Loss: 0.1754157394170761
```

```
Train Epoch: 4 [36480/60000 (61%)]      Loss: 0.031095417216420174
Train Epoch: 4 [37120/60000 (62%)]      Loss: 0.022793611511588097
Train Epoch: 4 [37760/60000 (63%)]      Loss: 0.06630464643239975
Train Epoch: 4 [38400/60000 (64%)]      Loss: 0.06667856872081757
Train Epoch: 4 [39040/60000 (65%)]      Loss: 0.13633711636066437
Train Epoch: 4 [39680/60000 (66%)]      Loss: 0.08615348480092422
Train Epoch: 4 [40320/60000 (67%)]      Loss: 0.01977170631289482
Train Epoch: 4 [40960/60000 (68%)]      Loss: 0.011665122583508492
Train Epoch: 4 [41600/60000 (69%)]      Loss: 0.15371040999889374
Train Epoch: 4 [42240/60000 (70%)]      Loss: 0.10337015241384506
Train Epoch: 4 [42880/60000 (71%)]      Loss: 0.27321961522102356
Train Epoch: 4 [43520/60000 (72%)]      Loss: 0.15841463208198547
Train Epoch: 4 [44160/60000 (74%)]      Loss: 0.0826229602098465
Train Epoch: 4 [44800/60000 (75%)]      Loss: 0.026829149574041367
Train Epoch: 4 [45440/60000 (76%)]      Loss: 0.05982248857617378
Train Epoch: 4 [46080/60000 (77%)]      Loss: 0.1340295672416687
Train Epoch: 4 [46720/60000 (78%)]      Loss: 0.03182338923215866
Train Epoch: 4 [47360/60000 (79%)]      Loss: 0.052005160599946976
Train Epoch: 4 [48000/60000 (80%)]      Loss: 0.058387283235788345
Train Epoch: 4 [48640/60000 (81%)]      Loss: 0.070803326974391937
Train Epoch: 4 [49280/60000 (82%)]      Loss: 0.1187104806303978
Train Epoch: 4 [49920/60000 (83%)]      Loss: 0.15529263019561768
Train Epoch: 4 [50560/60000 (84%)]      Loss: 0.0242131520062685
Train Epoch: 4 [51200/60000 (85%)]      Loss: 0.026580406352877617
Train Epoch: 4 [51840/60000 (86%)]      Loss: 0.1182308942079544
Train Epoch: 4 [52480/60000 (87%)]      Loss: 0.12455020099878311
Train Epoch: 4 [53120/60000 (88%)]      Loss: 0.061582136899232864
Train Epoch: 4 [53760/60000 (90%)]      Loss: 0.13784000277519226
Train Epoch: 4 [54400/60000 (91%)]      Loss: 0.24187862873077393
Train Epoch: 4 [55040/60000 (92%)]      Loss: 0.20562392473220825
Train Epoch: 4 [55680/60000 (93%)]      Loss: 0.2608974277973175
Train Epoch: 4 [56320/60000 (94%)]      Loss: 0.13122990727424622
Train Epoch: 4 [56960/60000 (95%)]      Loss: 0.08322027325630188
Train Epoch: 4 [57600/60000 (96%)]      Loss: 0.23287004232406616
Train Epoch: 4 [58240/60000 (97%)]      Loss: 0.034886129200045853
Train Epoch: 4 [58880/60000 (98%)]      Loss: 0.10401330143213272
Train Epoch: 4 [59520/60000 (99%)]      Loss: 0.09012478590011597
Epoch 4: Training Loss: 0.0016, Accuracy: 96.94%
Test Loss: 0.0001, Accuracy: 97.19%
Train Epoch: 5 [0/60000 (0%)]   Loss: 0.03844011202454567
Train Epoch: 5 [640/60000 (1%)] Loss: 0.11656370759010315
Train Epoch: 5 [1280/60000 (2%)]        Loss: 0.044077228754758835
Train Epoch: 5 [1920/60000 (3%)]        Loss: 0.2592853605747223
Train Epoch: 5 [2560/60000 (4%)]        Loss: 0.09250041842460632
Train Epoch: 5 [3200/60000 (5%)]        Loss: 0.09256596863269806
Train Epoch: 5 [3840/60000 (6%)]        Loss: 0.0328744575381279
Train Epoch: 5 [4480/60000 (7%)]        Loss: 0.1081237718462944
Train Epoch: 5 [5120/60000 (9%)]        Loss: 0.06504357606172562
```

```
Train Epoch: 5 [5760/60000 (10%)]      Loss: 0.018293146044015884
Train Epoch: 5 [6400/60000 (11%)]      Loss: 0.1773754507303238
Train Epoch: 5 [7040/60000 (12%)]      Loss: 0.353697806596756
Train Epoch: 5 [7680/60000 (13%)]      Loss: 0.17453594505786896
Train Epoch: 5 [8320/60000 (14%)]      Loss: 0.0538177490234375
Train Epoch: 5 [8960/60000 (15%)]      Loss: 0.20071837306022644
Train Epoch: 5 [9600/60000 (16%)]      Loss: 0.10169593244791031
Train Epoch: 5 [10240/60000 (17%)]     Loss: 0.017047367990016937
Train Epoch: 5 [10880/60000 (18%)]     Loss: 0.1374366283416748
Train Epoch: 5 [11520/60000 (19%)]     Loss: 0.11400455236434937
Train Epoch: 5 [12160/60000 (20%)]     Loss: 0.09394153952598572
Train Epoch: 5 [12800/60000 (21%)]     Loss: 0.051858749240636826
Train Epoch: 5 [13440/60000 (22%)]     Loss: 0.11549912393093109
Train Epoch: 5 [14080/60000 (23%)]     Loss: 0.13234218955039978
Train Epoch: 5 [14720/60000 (25%)]     Loss: 0.02681279368698597
Train Epoch: 5 [15360/60000 (26%)]     Loss: 0.06917355954647064
Train Epoch: 5 [16000/60000 (27%)]     Loss: 0.042503040283918388
Train Epoch: 5 [16640/60000 (28%)]     Loss: 0.21100379526615143
Train Epoch: 5 [17280/60000 (29%)]     Loss: 0.060908064246177767
Train Epoch: 5 [17920/60000 (30%)]     Loss: 0.21013018488883972
Train Epoch: 5 [18560/60000 (31%)]     Loss: 0.011505628004670143
Train Epoch: 5 [19200/60000 (32%)]     Loss: 0.01587432064116001
Train Epoch: 5 [19840/60000 (33%)]     Loss: 0.22426021099090576
Train Epoch: 5 [20480/60000 (34%)]     Loss: 0.1414818912744522
Train Epoch: 5 [21120/60000 (35%)]     Loss: 0.20659922063350677
Train Epoch: 5 [21760/60000 (36%)]     Loss: 0.133823126554448914
Train Epoch: 5 [22400/60000 (37%)]     Loss: 0.10896188020706177
Train Epoch: 5 [23040/60000 (38%)]     Loss: 0.053796801716089254
Train Epoch: 5 [23680/60000 (39%)]     Loss: 0.18020252883434296
Train Epoch: 5 [24320/60000 (41%)]     Loss: 0.1536257117986679
Train Epoch: 5 [24960/60000 (42%)]     Loss: 0.14093174040317535
Train Epoch: 5 [25600/60000 (43%)]     Loss: 0.127618670463562
Train Epoch: 5 [26240/60000 (44%)]     Loss: 0.15204845368862152
Train Epoch: 5 [26880/60000 (45%)]     Loss: 0.07341477274894714
Train Epoch: 5 [27520/60000 (46%)]     Loss: 0.015673860907554626
Train Epoch: 5 [28160/60000 (47%)]     Loss: 0.026264343410730362
Train Epoch: 5 [28800/60000 (48%)]     Loss: 0.050386544317007065
Train Epoch: 5 [29440/60000 (49%)]     Loss: 0.14973245561122894
Train Epoch: 5 [30080/60000 (50%)]     Loss: 0.13079029321670532
Train Epoch: 5 [30720/60000 (51%)]     Loss: 0.029139099642634392
Train Epoch: 5 [31360/60000 (52%)]     Loss: 0.10761843621730804
Train Epoch: 5 [32000/60000 (53%)]     Loss: 0.14813056588172913
Train Epoch: 5 [32640/60000 (54%)]     Loss: 0.17832359671592712
Train Epoch: 5 [33280/60000 (55%)]     Loss: 0.1352306753396988
Train Epoch: 5 [33920/60000 (57%)]     Loss: 0.02706296555697918
Train Epoch: 5 [34560/60000 (58%)]     Loss: 0.03620027378201485
Train Epoch: 5 [35200/60000 (59%)]     Loss: 0.028743652626872063
Train Epoch: 5 [35840/60000 (60%)]     Loss: 0.14176984131336212
```

```
Train Epoch: 5 [36480/60000 (61%)]        Loss: 0.0458180271089077
Train Epoch: 5 [37120/60000 (62%)]        Loss: 0.009511301293969154
Train Epoch: 5 [37760/60000 (63%)]        Loss: 0.07681429386138916
Train Epoch: 5 [38400/60000 (64%)]        Loss: 0.031057871878147125
Train Epoch: 5 [39040/60000 (65%)]        Loss: 0.12588468194007874
Train Epoch: 5 [39680/60000 (66%)]        Loss: 0.13662128150463104
Train Epoch: 5 [40320/60000 (67%)]        Loss: 0.1555720865726471
Train Epoch: 5 [40960/60000 (68%)]        Loss: 0.03709964081645012
Train Epoch: 5 [41600/60000 (69%)]        Loss: 0.22294816374778748
Train Epoch: 5 [42240/60000 (70%)]        Loss: 0.017313607037067413
Train Epoch: 5 [42880/60000 (71%)]        Loss: 0.028628036379814148
Train Epoch: 5 [43520/60000 (72%)]        Loss: 0.22270551161670685
Train Epoch: 5 [44160/60000 (74%)]        Loss: 0.10587574541568756
Train Epoch: 5 [44800/60000 (75%)]        Loss: 0.018538184463977814
Train Epoch: 5 [45440/60000 (76%)]        Loss: 0.08913295716047287
Train Epoch: 5 [46080/60000 (77%)]        Loss: 0.028935495764017105
Train Epoch: 5 [46720/60000 (78%)]        Loss: 0.12731686234474182
Train Epoch: 5 [47360/60000 (79%)]        Loss: 0.2096991240978241
Train Epoch: 5 [48000/60000 (80%)]        Loss: 0.12974141538143158
Train Epoch: 5 [48640/60000 (81%)]        Loss: 0.169158935546875
Train Epoch: 5 [49280/60000 (82%)]        Loss: 0.06459783017635345
Train Epoch: 5 [49920/60000 (83%)]        Loss: 0.07833244651556015
Train Epoch: 5 [50560/60000 (84%)]        Loss: 0.07146286219358444
Train Epoch: 5 [51200/60000 (85%)]        Loss: 0.11621267348527908
Train Epoch: 5 [51840/60000 (86%)]        Loss: 0.39875128865242004
Train Epoch: 5 [52480/60000 (87%)]        Loss: 0.06378942728042603
Train Epoch: 5 [53120/60000 (88%)]        Loss: 0.05711818113923073
Train Epoch: 5 [53760/60000 (90%)]        Loss: 0.128086194396011898
Train Epoch: 5 [54400/60000 (91%)]        Loss: 0.029803194105625153
Train Epoch: 5 [55040/60000 (92%)]        Loss: 0.12160080671310425
Train Epoch: 5 [55680/60000 (93%)]        Loss: 0.021601945161819458
Train Epoch: 5 [56320/60000 (94%)]        Loss: 0.02646557055413723
Train Epoch: 5 [56960/60000 (95%)]        Loss: 0.09651637822389603
Train Epoch: 5 [57600/60000 (96%)]        Loss: 0.10775717347860336
Train Epoch: 5 [58240/60000 (97%)]        Loss: 0.033375276178121567
Train Epoch: 5 [58880/60000 (98%)]        Loss: 0.25623631477355957
Train Epoch: 5 [59520/60000 (99%)]        Loss: 0.019595788791775703
Epoch 5: Training Loss: 0.0014, Accuracy: 97.49%
Test Loss: 0.0001, Accuracy: 97.68%
Train Epoch: 6 [0/60000 (0%)]   Loss: 0.08110550791025162
Train Epoch: 6 [640/60000 (1%)] Loss: 0.05644409731030464
Train Epoch: 6 [1280/60000 (2%)]          Loss: 0.21336232125759125
Train Epoch: 6 [1920/60000 (3%)]          Loss: 0.0703141838312149
Train Epoch: 6 [2560/60000 (4%)]          Loss: 0.06429192423820496
Train Epoch: 6 [3200/60000 (5%)]          Loss: 0.10854922980070114
Train Epoch: 6 [3840/60000 (6%)]          Loss: 0.04286909103393555
Train Epoch: 6 [4480/60000 (7%)]          Loss: 0.1439521461725235
Train Epoch: 6 [5120/60000 (9%)]          Loss: 0.03076135367155075
```

```
Train Epoch: 6 [5760/60000 (10%)]      Loss: 0.05288704112172127
Train Epoch: 6 [6400/60000 (11%)]      Loss: 0.02547217532992363
Train Epoch: 6 [7040/60000 (12%)]      Loss: 0.19784657657146454
Train Epoch: 6 [7680/60000 (13%)]      Loss: 0.14449946582317352
Train Epoch: 6 [8320/60000 (14%)]      Loss: 0.03535604849457741
Train Epoch: 6 [8960/60000 (15%)]      Loss: 0.0267721582204I0347
Train Epoch: 6 [9600/60000 (16%)]      Loss: 0.03513962775468826
Train Epoch: 6 [10240/60000 (17%)]     Loss: 0.010584989562630653
Train Epoch: 6 [10880/60000 (18%)]     Loss: 0.03970927372574806
Train Epoch: 6 [11520/60000 (19%)]     Loss: 0.18951945006847382
Train Epoch: 6 [12160/60000 (20%)]     Loss: 0.033911824226379395
Train Epoch: 6 [12800/60000 (21%)]     Loss: 0.05906685069203377
Train Epoch: 6 [13440/60000 (22%)]     Loss: 0.08158185333013535
Train Epoch: 6 [14080/60000 (23%)]     Loss: 0.013848042115569115
Train Epoch: 6 [14720/60000 (25%)]     Loss: 0.10341477394104004
Train Epoch: 6 [15360/60000 (26%)]     Loss: 0.051148440688848495
Train Epoch: 6 [16000/60000 (27%)]     Loss: 0.0799131840467453
Train Epoch: 6 [16640/60000 (28%)]     Loss: 0.06362628936767578
Train Epoch: 6 [17280/60000 (29%)]     Loss: 0.0826692059636116
Train Epoch: 6 [17920/60000 (30%)]     Loss: 0.02185036987066269
Train Epoch: 6 [18560/60000 (31%)]     Loss: 0.054500989615917206
Train Epoch: 6 [19200/60000 (32%)]     Loss: 0.023152055218815804
Train Epoch: 6 [19840/60000 (33%)]     Loss: 0.022281957790255547
Train Epoch: 6 [20480/60000 (34%)]     Loss: 0.015451537445187569
Train Epoch: 6 [21120/60000 (35%)]     Loss: 0.043618559709024429
Train Epoch: 6 [21760/60000 (36%)]     Loss: 0.1393440216779709
Train Epoch: 6 [22400/60000 (37%)]     Loss: 0.06865524500608444
Train Epoch: 6 [23040/60000 (38%)]     Loss: 0.10468790680170059
Train Epoch: 6 [23680/60000 (39%)]     Loss: 0.06388480961322784
Train Epoch: 6 [24320/60000 (41%)]     Loss: 0.024439040571451187
Train Epoch: 6 [24960/60000 (42%)]     Loss: 0.14256174862384796
Train Epoch: 6 [25600/60000 (43%)]     Loss: 0.07145658135414124
Train Epoch: 6 [26240/60000 (44%)]     Loss: 0.073672704439863205
Train Epoch: 6 [26880/60000 (45%)]     Loss: 0.12189199775457382
Train Epoch: 6 [27520/60000 (46%)]     Loss: 0.09417702257633209
Train Epoch: 6 [28160/60000 (47%)]     Loss: 0.13630183041095734
Train Epoch: 6 [28800/60000 (48%)]     Loss: 0.20102398097515106
Train Epoch: 6 [29440/60000 (49%)]     Loss: 0.027291733771562576
Train Epoch: 6 [30080/60000 (50%)]     Loss: 0.024890342727303505
Train Epoch: 6 [30720/60000 (51%)]     Loss: 0.09140019118785858
Train Epoch: 6 [31360/60000 (52%)]     Loss: 0.04339708015322685
Train Epoch: 6 [32000/60000 (53%)]     Loss: 0.021879184991121292
Train Epoch: 6 [32640/60000 (54%)]     Loss: 0.04578009247779846
Train Epoch: 6 [33280/60000 (55%)]     Loss: 0.155564695596669495
Train Epoch: 6 [33920/60000 (57%)]     Loss: 0.03382530063390732
Train Epoch: 6 [34560/60000 (58%)]     Loss: 0.025933004915714264
Train Epoch: 6 [35200/60000 (59%)]     Loss: 0.21203766763210297
Train Epoch: 6 [35840/60000 (60%)]     Loss: 0.009173872880637646
```

```
Train Epoch: 6 [36480/60000 (61%)]      Loss: 0.06199938431382179
Train Epoch: 6 [37120/60000 (62%)]      Loss: 0.010630035772919655
Train Epoch: 6 [37760/60000 (63%)]      Loss: 0.13756370544433594
Train Epoch: 6 [38400/60000 (64%)]      Loss: 0.06943994015455246
Train Epoch: 6 [39040/60000 (65%)]      Loss: 0.04328933358192444
Train Epoch: 6 [39680/60000 (66%)]      Loss: 0.0036972740199416876
Train Epoch: 6 [40320/60000 (67%)]      Loss: 0.07474483549594879
Train Epoch: 6 [40960/60000 (68%)]      Loss: 0.04942787438631058
Train Epoch: 6 [41600/60000 (69%)]      Loss: 0.08565808832645416
Train Epoch: 6 [42240/60000 (70%)]      Loss: 0.06754539161920547
Train Epoch: 6 [42880/60000 (71%)]      Loss: 0.05337926372885704
Train Epoch: 6 [43520/60000 (72%)]      Loss: 0.11770277551460266
Train Epoch: 6 [44160/60000 (74%)]      Loss: 0.04751996695995331
Train Epoch: 6 [44800/60000 (75%)]      Loss: 0.04656316339969635
Train Epoch: 6 [45440/60000 (76%)]      Loss: 0.03080292046070099
Train Epoch: 6 [46080/60000 (77%)]      Loss: 0.08063174784183502
Train Epoch: 6 [46720/60000 (78%)]      Loss: 0.01784062199294567
Train Epoch: 6 [47360/60000 (79%)]      Loss: 0.023326469585299492
Train Epoch: 6 [48000/60000 (80%)]      Loss: 0.07939612865447998
Train Epoch: 6 [48640/60000 (81%)]      Loss: 0.1470966786146164
Train Epoch: 6 [49280/60000 (82%)]      Loss: 0.004473178181797266
Train Epoch: 6 [49920/60000 (83%)]      Loss: 0.04988798871636391
Train Epoch: 6 [50560/60000 (84%)]      Loss: 0.022232089191675186
Train Epoch: 6 [51200/60000 (85%)]      Loss: 0.04211222007870674
Train Epoch: 6 [51840/60000 (86%)]      Loss: 0.051815714687108994
Train Epoch: 6 [52480/60000 (87%)]      Loss: 0.12004239857196808
Train Epoch: 6 [53120/60000 (88%)]      Loss: 0.08650272339582443
Train Epoch: 6 [53760/60000 (90%)]      Loss: 0.04723498597741127
Train Epoch: 6 [54400/60000 (91%)]      Loss: 0.05790979042649269
Train Epoch: 6 [55040/60000 (92%)]      Loss: 0.05573449656367302
Train Epoch: 6 [55680/60000 (93%)]      Loss: 0.04619918763637543
Train Epoch: 6 [56320/60000 (94%)]      Loss: 0.12291458249092102
Train Epoch: 6 [56960/60000 (95%)]      Loss: 0.03934129327535629
Train Epoch: 6 [57600/60000 (96%)]      Loss: 0.20550855994224548
Train Epoch: 6 [58240/60000 (97%)]      Loss: 0.03523800149559975
Train Epoch: 6 [58880/60000 (98%)]      Loss: 0.059994008392095566
Train Epoch: 6 [59520/60000 (99%)]      Loss: 0.04739918187260628
Epoch 6: Training Loss: 0.0011, Accuracy: 97.81%
Test Loss: 0.0001, Accuracy: 97.91%
Train Epoch: 7 [0/60000 (0%)]   Loss: 0.04529151692986488
Train Epoch: 7 [640/60000 (1%)] Loss: 0.06657196581363678
Train Epoch: 7 [1280/60000 (2%)]        Loss: 0.069626221212005615
Train Epoch: 7 [1920/60000 (3%)]        Loss: 0.01349425595253706
Train Epoch: 7 [2560/60000 (4%)]        Loss: 0.020827101543545723
Train Epoch: 7 [3200/60000 (5%)]        Loss: 0.11666480451822281
Train Epoch: 7 [3840/60000 (6%)]        Loss: 0.0699005052447319
Train Epoch: 7 [4480/60000 (7%)]        Loss: 0.10725575685501099
Train Epoch: 7 [5120/60000 (9%)]        Loss: 0.017986241728067398
```

```
Train Epoch: 7 [5760/60000 (10%)]      Loss: 0.07745257765054703
Train Epoch: 7 [6400/60000 (11%)]      Loss: 0.05646171048283577
Train Epoch: 7 [7040/60000 (12%)]      Loss: 0.06435856968164444
Train Epoch: 7 [7680/60000 (13%)]      Loss: 0.08597834408283234
Train Epoch: 7 [8320/60000 (14%)]      Loss: 0.04943186789751053
Train Epoch: 7 [8960/60000 (15%)]      Loss: 0.019918298348784447
Train Epoch: 7 [9600/60000 (16%)]      Loss: 0.011581255123019218
Train Epoch: 7 [10240/60000 (17%)]     Loss: 0.040635351091623306
Train Epoch: 7 [10880/60000 (18%)]     Loss: 0.04006193205714226
Train Epoch: 7 [11520/60000 (19%)]     Loss: 0.04435594752430916
Train Epoch: 7 [12160/60000 (20%)]     Loss: 0.11191114783287048
Train Epoch: 7 [12800/60000 (21%)]     Loss: 0.015062726102769375
Train Epoch: 7 [13440/60000 (22%)]     Loss: 0.06409517675638199
Train Epoch: 7 [14080/60000 (23%)]     Loss: 0.009210055693984032
Train Epoch: 7 [14720/60000 (25%)]     Loss: 0.04854174703359604
Train Epoch: 7 [15360/60000 (26%)]     Loss: 0.07900889217853546
Train Epoch: 7 [16000/60000 (27%)]     Loss: 0.03316906467080116
Train Epoch: 7 [16640/60000 (28%)]     Loss: 0.08649293333292007
Train Epoch: 7 [17280/60000 (29%)]     Loss: 0.06768598407506943
Train Epoch: 7 [17920/60000 (30%)]     Loss: 0.07991968095302582
Train Epoch: 7 [18560/60000 (31%)]     Loss: 0.07122315466403961
Train Epoch: 7 [19200/60000 (32%)]     Loss: 0.009212230332195759
Train Epoch: 7 [19840/60000 (33%)]     Loss: 0.08597870171070099
Train Epoch: 7 [20480/60000 (34%)]     Loss: 0.0328020416200161
Train Epoch: 7 [21120/60000 (35%)]     Loss: 0.03923725709319115
Train Epoch: 7 [21760/60000 (36%)]     Loss: 0.0636308416724205
Train Epoch: 7 [22400/60000 (37%)]     Loss: 0.094424430131912231
Train Epoch: 7 [23040/60000 (38%)]     Loss: 0.027539795264601707
Train Epoch: 7 [23680/60000 (39%)]     Loss: 0.016800200566649437
Train Epoch: 7 [24320/60000 (41%)]     Loss: 0.059770990163087845
Train Epoch: 7 [24960/60000 (42%)]     Loss: 0.10158231854438782
Train Epoch: 7 [25600/60000 (43%)]     Loss: 0.017706183716654778
Train Epoch: 7 [26240/60000 (44%)]     Loss: 0.03422287479043007
Train Epoch: 7 [26880/60000 (45%)]     Loss: 0.12156163156032562
Train Epoch: 7 [27520/60000 (46%)]     Loss: 0.009127852506935596
Train Epoch: 7 [28160/60000 (47%)]     Loss: 0.050060927867889404
Train Epoch: 7 [28800/60000 (48%)]     Loss: 0.011505097150802612
Train Epoch: 7 [29440/60000 (49%)]     Loss: 0.16173304617404938
Train Epoch: 7 [30080/60000 (50%)]     Loss: 0.136664309680461884
Train Epoch: 7 [30720/60000 (51%)]     Loss: 0.16488119959831238
Train Epoch: 7 [31360/60000 (52%)]     Loss: 0.09648597240447998
Train Epoch: 7 [32000/60000 (53%)]     Loss: 0.12633337080478668
Train Epoch: 7 [32640/60000 (54%)]     Loss: 0.12374678254127502
Train Epoch: 7 [33280/60000 (55%)]     Loss: 0.1991404891014099
Train Epoch: 7 [33920/60000 (57%)]     Loss: 0.013657265342772007
Train Epoch: 7 [34560/60000 (58%)]     Loss: 0.08978036046028137
Train Epoch: 7 [35200/60000 (59%)]     Loss: 0.14059025049209595
Train Epoch: 7 [35840/60000 (60%)]     Loss: 0.012640341185033321
```

```
Train Epoch: 7 [36480/60000 (61%)]      Loss: 0.06457734853029251
Train Epoch: 7 [37120/60000 (62%)]      Loss: 0.055932000279426575
Train Epoch: 7 [37760/60000 (63%)]      Loss: 0.0661948099732399
Train Epoch: 7 [38400/60000 (64%)]      Loss: 0.062034182250499725
Train Epoch: 7 [39040/60000 (65%)]      Loss: 0.2039961963891983
Train Epoch: 7 [39680/60000 (66%)]      Loss: 0.046258457005023956
Train Epoch: 7 [40320/60000 (67%)]      Loss: 0.020134827122092247
Train Epoch: 7 [40960/60000 (68%)]      Loss: 0.03532683104276657
Train Epoch: 7 [41600/60000 (69%)]      Loss: 0.023614061996340755
Train Epoch: 7 [42240/60000 (70%)]      Loss: 0.0338163822889328
Train Epoch: 7 [42880/60000 (71%)]      Loss: 0.016044620424509905
Train Epoch: 7 [43520/60000 (72%)]      Loss: 0.10040933638811111
Train Epoch: 7 [44160/60000 (74%)]      Loss: 0.020955977961421013
Train Epoch: 7 [44800/60000 (75%)]      Loss: 0.018079787492752075
Train Epoch: 7 [45440/60000 (76%)]      Loss: 0.04111455753445625
Train Epoch: 7 [46080/60000 (77%)]      Loss: 0.07496371120214462
Train Epoch: 7 [46720/60000 (78%)]      Loss: 0.010469984263181686
Train Epoch: 7 [47360/60000 (79%)]      Loss: 0.04622308909893036
Train Epoch: 7 [48000/60000 (80%)]      Loss: 0.02537374012172222
Train Epoch: 7 [48640/60000 (81%)]      Loss: 0.06661836057901382
Train Epoch: 7 [49280/60000 (82%)]      Loss: 0.1602323055267334
Train Epoch: 7 [49920/60000 (83%)]      Loss: 0.036163538694381714
Train Epoch: 7 [50560/60000 (84%)]      Loss: 0.034479524940252304
Train Epoch: 7 [51200/60000 (85%)]      Loss: 0.11892833560705185
Train Epoch: 7 [51840/60000 (86%)]      Loss: 0.009526753798127174
Train Epoch: 7 [52480/60000 (87%)]      Loss: 0.07374266535043716
Train Epoch: 7 [53120/60000 (88%)]      Loss: 0.008960599079728127
Train Epoch: 7 [53760/60000 (90%)]      Loss: 0.06948988884687424
Train Epoch: 7 [54400/60000 (91%)]      Loss: 0.102542139589786533
Train Epoch: 7 [55040/60000 (92%)]      Loss: 0.0675894021987915
Train Epoch: 7 [55680/60000 (93%)]      Loss: 0.042182400822639465
Train Epoch: 7 [56320/60000 (94%)]      Loss: 0.19745109975337982
Train Epoch: 7 [56960/60000 (95%)]      Loss: 0.08934978395700455
Train Epoch: 7 [57600/60000 (96%)]      Loss: 0.02609509415924549
Train Epoch: 7 [58240/60000 (97%)]      Loss: 0.13944967091083527
Train Epoch: 7 [58880/60000 (98%)]      Loss: 0.18065191805362701
Train Epoch: 7 [59520/60000 (99%)]      Loss: 0.035018544644117355
Epoch 7: Training Loss: 0.0011, Accuracy: 97.97%
Test Loss: 0.0001, Accuracy: 97.96%
Train Epoch: 8 [0/60000 (0%)]   Loss: 0.1378110945224762
Train Epoch: 8 [640/60000 (1%)] Loss: 0.133957177400589
Train Epoch: 8 [1280/60000 (2%)]        Loss: 0.052692919969558716
Train Epoch: 8 [1920/60000 (3%)]        Loss: 0.14114892482757568
Train Epoch: 8 [2560/60000 (4%)]        Loss: 0.018604757264256477
Train Epoch: 8 [3200/60000 (5%)]        Loss: 0.04838918149471283
Train Epoch: 8 [3840/60000 (6%)]        Loss: 0.012605438940227032
Train Epoch: 8 [4480/60000 (7%)]        Loss: 0.04217258095741272
Train Epoch: 8 [5120/60000 (9%)]        Loss: 0.015103700570762157
```

```
Train Epoch: 8 [5760/60000 (10%)]      Loss: 0.047960441559553146
Train Epoch: 8 [6400/60000 (11%)]      Loss: 0.012910488992929459
Train Epoch: 8 [7040/60000 (12%)]      Loss: 0.007801611442118883
Train Epoch: 8 [7680/60000 (13%)]      Loss: 0.02034585177898407
Train Epoch: 8 [8320/60000 (14%)]      Loss: 0.006644136272370815
Train Epoch: 8 [8960/60000 (15%)]      Loss: 0.032004132866859436
Train Epoch: 8 [9600/60000 (16%)]      Loss: 0.05469847470521927
Train Epoch: 8 [10240/60000 (17%)]     Loss: 0.019867610186338425
Train Epoch: 8 [10880/60000 (18%)]     Loss: 0.048746682703495026
Train Epoch: 8 [11520/60000 (19%)]     Loss: 0.03739789128303528
Train Epoch: 8 [12160/60000 (20%)]     Loss: 0.039289899917039871
Train Epoch: 8 [12800/60000 (21%)]     Loss: 0.098953381180276324
Train Epoch: 8 [13440/60000 (22%)]     Loss: 0.056421294808387756
Train Epoch: 8 [14080/60000 (23%)]     Loss: 0.017256207764148712
Train Epoch: 8 [14720/60000 (25%)]     Loss: 0.09985906630754471
Train Epoch: 8 [15360/60000 (26%)]     Loss: 0.14870703220367432
Train Epoch: 8 [16000/60000 (27%)]     Loss: 0.1144748404622078
Train Epoch: 8 [16640/60000 (28%)]     Loss: 0.0482235848903656
Train Epoch: 8 [17280/60000 (29%)]     Loss: 0.03673377260565758
Train Epoch: 8 [17920/60000 (30%)]     Loss: 0.034084171056747437
Train Epoch: 8 [18560/60000 (31%)]     Loss: 0.07135553658008575
Train Epoch: 8 [19200/60000 (32%)]     Loss: 0.02307356894016266
Train Epoch: 8 [19840/60000 (33%)]     Loss: 0.024053622037172318
Train Epoch: 8 [20480/60000 (34%)]     Loss: 0.010054020211100578
Train Epoch: 8 [21120/60000 (35%)]     Loss: 0.026833955198526382
Train Epoch: 8 [21760/60000 (36%)]     Loss: 0.1232292652130127
Train Epoch: 8 [22400/60000 (37%)]     Loss: 0.12462225556373596
Train Epoch: 8 [23040/60000 (38%)]     Loss: 0.10730334371328354
Train Epoch: 8 [23680/60000 (39%)]     Loss: 0.08302547037601471
Train Epoch: 8 [24320/60000 (41%)]     Loss: 0.14715883135795593
Train Epoch: 8 [24960/60000 (42%)]     Loss: 0.07252812385559082
Train Epoch: 8 [25600/60000 (43%)]     Loss: 0.027926068753004074
Train Epoch: 8 [26240/60000 (44%)]     Loss: 0.007221829611808062
Train Epoch: 8 [26880/60000 (45%)]     Loss: 0.06509488821029663
Train Epoch: 8 [27520/60000 (46%)]     Loss: 0.059282831847667694
Train Epoch: 8 [28160/60000 (47%)]     Loss: 0.020346185192465782
Train Epoch: 8 [28800/60000 (48%)]     Loss: 0.010517265647649765
Train Epoch: 8 [29440/60000 (49%)]     Loss: 0.008129780180752277
Train Epoch: 8 [30080/60000 (50%)]     Loss: 0.025214316323399544
Train Epoch: 8 [30720/60000 (51%)]     Loss: 0.09780651330947876
Train Epoch: 8 [31360/60000 (52%)]     Loss: 0.04416594281792641
Train Epoch: 8 [32000/60000 (53%)]     Loss: 0.004045569337904453
Train Epoch: 8 [32640/60000 (54%)]     Loss: 0.009298824705183506
Train Epoch: 8 [33280/60000 (55%)]     Loss: 0.044708479195833206
Train Epoch: 8 [33920/60000 (57%)]     Loss: 0.06771881878376007
Train Epoch: 8 [34560/60000 (58%)]     Loss: 0.13755857944488525
Train Epoch: 8 [35200/60000 (59%)]     Loss: 0.051338110119104385
Train Epoch: 8 [35840/60000 (60%)]     Loss: 0.012001583352684975
```

```
Train Epoch: 8 [36480/60000 (61%)]      Loss: 0.04991784319281578
Train Epoch: 8 [37120/60000 (62%)]      Loss: 0.021118206903338432
Train Epoch: 8 [37760/60000 (63%)]      Loss: 0.06738118082284927
Train Epoch: 8 [38400/60000 (64%)]      Loss: 0.017810065299272537
Train Epoch: 8 [39040/60000 (65%)]      Loss: 0.07127736508846283
Train Epoch: 8 [39680/60000 (66%)]      Loss: 0.04562351480126381
Train Epoch: 8 [40320/60000 (67%)]      Loss: 0.06823047995567322
Train Epoch: 8 [40960/60000 (68%)]      Loss: 0.18480288982391357
Train Epoch: 8 [41600/60000 (69%)]      Loss: 0.012502663768827915
Train Epoch: 8 [42240/60000 (70%)]      Loss: 0.09585174918174744
Train Epoch: 8 [42880/60000 (71%)]      Loss: 0.015594473108649254
Train Epoch: 8 [43520/60000 (72%)]      Loss: 0.016399767249822617
Train Epoch: 8 [44160/60000 (74%)]      Loss: 0.025542661547660828
Train Epoch: 8 [44800/60000 (75%)]      Loss: 0.019002342596650124
Train Epoch: 8 [45440/60000 (76%)]      Loss: 0.0646715834736824
Train Epoch: 8 [46080/60000 (77%)]      Loss: 0.008967321366071701
Train Epoch: 8 [46720/60000 (78%)]      Loss: 0.017806220799684525
Train Epoch: 8 [47360/60000 (79%)]      Loss: 0.007620032411068678
Train Epoch: 8 [48000/60000 (80%)]      Loss: 0.07621297985315323
Train Epoch: 8 [48640/60000 (81%)]      Loss: 0.11989504098892212
Train Epoch: 8 [49280/60000 (82%)]      Loss: 0.06177949160337448
Train Epoch: 8 [49920/60000 (83%)]      Loss: 0.05189407989382744
Train Epoch: 8 [50560/60000 (84%)]      Loss: 0.06424332410097122
Train Epoch: 8 [51200/60000 (85%)]      Loss: 0.02106437459588051
Train Epoch: 8 [51840/60000 (86%)]      Loss: 0.08736684918403625
Train Epoch: 8 [52480/60000 (87%)]      Loss: 0.10233617573976517
Train Epoch: 8 [53120/60000 (88%)]      Loss: 0.10714094340801239
Train Epoch: 8 [53760/60000 (90%)]      Loss: 0.0860336497426033
Train Epoch: 8 [54400/60000 (91%)]      Loss: 0.012421077117323875
Train Epoch: 8 [55040/60000 (92%)]      Loss: 0.011852190829981348
Train Epoch: 8 [55680/60000 (93%)]      Loss: 0.08246977627277374
Train Epoch: 8 [56320/60000 (94%)]      Loss: 0.03275807574391365
Train Epoch: 8 [56960/60000 (95%)]      Loss: 0.06002595275640488
Train Epoch: 8 [57600/60000 (96%)]      Loss: 0.06598295271396637
Train Epoch: 8 [58240/60000 (97%)]      Loss: 0.11051517724990845
Train Epoch: 8 [58880/60000 (98%)]      Loss: 0.029587052762508392
Train Epoch: 8 [59520/60000 (99%)]      Loss: 0.021879995241761208
Epoch 8: Training Loss: 0.0009, Accuracy: 98.24%
Test Loss: 0.0001, Accuracy: 97.70%
Train Epoch: 9 [0/60000 (0%)]   Loss: 0.045921001583337784
Train Epoch: 9 [640/60000 (1%)] Loss: 0.04571857675909996
Train Epoch: 9 [1280/60000 (2%)]        Loss: 0.014250364154577255
Train Epoch: 9 [1920/60000 (3%)]        Loss: 0.01157520804554224
Train Epoch: 9 [2560/60000 (4%)]        Loss: 0.08977872878313065
Train Epoch: 9 [3200/60000 (5%)]        Loss: 0.004707299172878265
Train Epoch: 9 [3840/60000 (6%)]        Loss: 0.013025548309087753
Train Epoch: 9 [4480/60000 (7%)]        Loss: 0.016232194378972054
Train Epoch: 9 [5120/60000 (9%)]        Loss: 0.03017127886414528
```

```
Train Epoch: 9 [5760/60000 (10%)]      Loss: 0.008844444528222084
Train Epoch: 9 [6400/60000 (11%)]      Loss: 0.034103672951459885
Train Epoch: 9 [7040/60000 (12%)]      Loss: 0.07133488357067108
Train Epoch: 9 [7680/60000 (13%)]      Loss: 0.018489832058548927
Train Epoch: 9 [8320/60000 (14%)]      Loss: 0.023041192442178726
Train Epoch: 9 [8960/60000 (15%)]      Loss: 0.0727444589138031
Train Epoch: 9 [9600/60000 (16%)]      Loss: 0.07389278709888458
Train Epoch: 9 [10240/60000 (17%)]     Loss: 0.04998977109789848
Train Epoch: 9 [10880/60000 (18%)]     Loss: 0.07489044964313507
Train Epoch: 9 [11520/60000 (19%)]     Loss: 0.016679849475622177
Train Epoch: 9 [12160/60000 (20%)]     Loss: 0.102732814848423
Train Epoch: 9 [12800/60000 (21%)]     Loss: 0.08203672617673874
Train Epoch: 9 [13440/60000 (22%)]     Loss: 0.08506347239017487
Train Epoch: 9 [14080/60000 (23%)]     Loss: 0.07567037642002106
Train Epoch: 9 [14720/60000 (25%)]     Loss: 0.027544083073735237
Train Epoch: 9 [15360/60000 (26%)]     Loss: 0.010614322498440742
Train Epoch: 9 [16000/60000 (27%)]     Loss: 0.006932755000889301
Train Epoch: 9 [16640/60000 (28%)]     Loss: 0.0061975750140845776
Train Epoch: 9 [17280/60000 (29%)]     Loss: 0.08121331036090851
Train Epoch: 9 [17920/60000 (30%)]     Loss: 0.061783380806446075
Train Epoch: 9 [18560/60000 (31%)]     Loss: 0.03355195373296738
Train Epoch: 9 [19200/60000 (32%)]     Loss: 0.07539252936840057
Train Epoch: 9 [19840/60000 (33%)]     Loss: 0.024222202005982399
Train Epoch: 9 [20480/60000 (34%)]     Loss: 0.014822708442807198
Train Epoch: 9 [21120/60000 (35%)]     Loss: 0.07844281941652298
Train Epoch: 9 [21760/60000 (36%)]     Loss: 0.008511999621987343
Train Epoch: 9 [22400/60000 (37%)]     Loss: 0.07756936550140381
Train Epoch: 9 [23040/60000 (38%)]     Loss: 0.04835008457303047
Train Epoch: 9 [23680/60000 (39%)]     Loss: 0.09636754542589188
Train Epoch: 9 [24320/60000 (41%)]     Loss: 0.02704034186899662
Train Epoch: 9 [24960/60000 (42%)]     Loss: 0.0262555293738842
Train Epoch: 9 [25600/60000 (43%)]     Loss: 0.017109761014580727
Train Epoch: 9 [26240/60000 (44%)]     Loss: 0.007722093723714352
Train Epoch: 9 [26880/60000 (45%)]     Loss: 0.013746676966547966
Train Epoch: 9 [27520/60000 (46%)]     Loss: 0.004909774754196405
Train Epoch: 9 [28160/60000 (47%)]     Loss: 0.180734783411026
Train Epoch: 9 [28800/60000 (48%)]     Loss: 0.08108732849359512
Train Epoch: 9 [29440/60000 (49%)]     Loss: 0.05177519842982292
Train Epoch: 9 [30080/60000 (50%)]     Loss: 0.01980036124587059
Train Epoch: 9 [30720/60000 (51%)]     Loss: 0.05374939739704132
Train Epoch: 9 [31360/60000 (52%)]     Loss: 0.051787205040454865
Train Epoch: 9 [32000/60000 (53%)]     Loss: 0.11122187227010727
Train Epoch: 9 [32640/60000 (54%)]     Loss: 0.1666269302368164
Train Epoch: 9 [33280/60000 (55%)]     Loss: 0.0615818165242672
Train Epoch: 9 [33920/60000 (57%)]     Loss: 0.03494195640087128
Train Epoch: 9 [34560/60000 (58%)]     Loss: 0.11392521858215332
Train Epoch: 9 [35200/60000 (59%)]     Loss: 0.02733844704926014
Train Epoch: 9 [35840/60000 (60%)]     Loss: 0.07499045133590698
```

```
Train Epoch: 9 [36480/60000 (61%)]      Loss: 0.10038653016090393
Train Epoch: 9 [37120/60000 (62%)]      Loss: 0.09322988241910934
Train Epoch: 9 [37760/60000 (63%)]      Loss: 0.019437314942479134
Train Epoch: 9 [38400/60000 (64%)]      Loss: 0.1238601803779602
Train Epoch: 9 [39040/60000 (65%)]      Loss: 0.06339089572429657
Train Epoch: 9 [39680/60000 (66%)]      Loss: 0.0064401524141430855
Train Epoch: 9 [40320/60000 (67%)]      Loss: 0.055768780410289764
Train Epoch: 9 [40960/60000 (68%)]      Loss: 0.061200036914944649
Train Epoch: 9 [41600/60000 (69%)]      Loss: 0.099591806653095245
Train Epoch: 9 [42240/60000 (70%)]      Loss: 0.09789781272411346
Train Epoch: 9 [42880/60000 (71%)]      Loss: 0.023550594225525856
Train Epoch: 9 [43520/60000 (72%)]      Loss: 0.020438548177480698
Train Epoch: 9 [44160/60000 (74%)]      Loss: 0.04732874780893326
Train Epoch: 9 [44800/60000 (75%)]      Loss: 0.09088119864463806
Train Epoch: 9 [45440/60000 (76%)]      Loss: 0.02080119401216507
Train Epoch: 9 [46080/60000 (77%)]      Loss: 0.003930837847292423
Train Epoch: 9 [46720/60000 (78%)]      Loss: 0.03757483884692192
Train Epoch: 9 [47360/60000 (79%)]      Loss: 0.002659385558217764
Train Epoch: 9 [48000/60000 (80%)]      Loss: 0.05238849297165871
Train Epoch: 9 [48640/60000 (81%)]      Loss: 0.03695380687713623
Train Epoch: 9 [49280/60000 (82%)]      Loss: 0.011301306076347828
Train Epoch: 9 [49920/60000 (83%)]      Loss: 0.04770320653915405
Train Epoch: 9 [50560/60000 (84%)]      Loss: 0.02235259860754013
Train Epoch: 9 [51200/60000 (85%)]      Loss: 0.04540349170565605
Train Epoch: 9 [51840/60000 (86%)]      Loss: 0.030338196083903313
Train Epoch: 9 [52480/60000 (87%)]      Loss: 0.07903792709112167
Train Epoch: 9 [53120/60000 (88%)]      Loss: 0.03426114469766617
Train Epoch: 9 [53760/60000 (90%)]      Loss: 0.10044210404157639
Train Epoch: 9 [54400/60000 (91%)]      Loss: 0.006996702402830124
Train Epoch: 9 [55040/60000 (92%)]      Loss: 0.05549793317914009
Train Epoch: 9 [55680/60000 (93%)]      Loss: 0.03937613591551781
Train Epoch: 9 [56320/60000 (94%)]      Loss: 0.006070930510759354
Train Epoch: 9 [56960/60000 (95%)]      Loss: 0.014551742933690548
Train Epoch: 9 [57600/60000 (96%)]      Loss: 0.011197556741535664
Train Epoch: 9 [58240/60000 (97%)]      Loss: 0.020401807501912117
Train Epoch: 9 [58880/60000 (98%)]      Loss: 0.004899901803582907
Train Epoch: 9 [59520/60000 (99%)]      Loss: 0.008335938677191734
Epoch 9: Training Loss: 0.0008, Accuracy: 98.39%
Test Loss: 0.0001, Accuracy: 98.30%
Train Epoch: 10 [0/60000 (0%)]  Loss: 0.05117800831794739
Train Epoch: 10 [640/60000 (1%)]        Loss: 0.01719316467642784
Train Epoch: 10 [1280/60000 (2%)]       Loss: 0.025528663769364357
Train Epoch: 10 [1920/60000 (3%)]       Loss: 0.07885055243968964
Train Epoch: 10 [2560/60000 (4%)]       Loss: 0.006762672681361437
Train Epoch: 10 [3200/60000 (5%)]       Loss: 0.08690792322158813
Train Epoch: 10 [3840/60000 (6%)]       Loss: 0.09442201256752014
Train Epoch: 10 [4480/60000 (7%)]       Loss: 0.04073110595345497
Train Epoch: 10 [5120/60000 (9%)]       Loss: 0.0036063797306269407
```

```
Train Epoch: 10 [5760/60000 (10%)]      Loss: 0.016529452055692673
Train Epoch: 10 [6400/60000 (11%)]      Loss: 0.049213819205760956
Train Epoch: 10 [7040/60000 (12%)]      Loss: 0.007067019119858742
Train Epoch: 10 [7680/60000 (13%)]      Loss: 0.03129727393388748
Train Epoch: 10 [8320/60000 (14%)]      Loss: 0.026720426976680756
Train Epoch: 10 [8960/60000 (15%)]      Loss: 0.07753036916255951
Train Epoch: 10 [9600/60000 (16%)]      Loss: 0.07568122446537018
Train Epoch: 10 [10240/60000 (17%)]     Loss: 0.008806700818240643
Train Epoch: 10 [10880/60000 (18%)]     Loss: 0.006534470245242119
Train Epoch: 10 [11520/60000 (19%)]     Loss: 0.009755044244229794
Train Epoch: 10 [12160/60000 (20%)]     Loss: 0.012893276289105415
Train Epoch: 10 [12800/60000 (21%)]     Loss: 0.04909821227192879
Train Epoch: 10 [13440/60000 (22%)]     Loss: 0.11827389895915985
Train Epoch: 10 [14080/60000 (23%)]     Loss: 0.00865648128092289
Train Epoch: 10 [14720/60000 (25%)]     Loss: 0.004822662565857172
Train Epoch: 10 [15360/60000 (26%)]     Loss: 0.017141181975603104
Train Epoch: 10 [16000/60000 (27%)]     Loss: 0.06516236066818237
Train Epoch: 10 [16640/60000 (28%)]     Loss: 0.06092591956825782
Train Epoch: 10 [17280/60000 (29%)]     Loss: 0.039382826536893845
Train Epoch: 10 [17920/60000 (30%)]     Loss: 0.006253070197999477
Train Epoch: 10 [18560/60000 (31%)]     Loss: 0.025980567559599876
Train Epoch: 10 [19200/60000 (32%)]     Loss: 0.007095531094819307
Train Epoch: 10 [19840/60000 (33%)]     Loss: 0.040840234607458115
Train Epoch: 10 [20480/60000 (34%)]     Loss: 0.023627012968063354
Train Epoch: 10 [21120/60000 (35%)]     Loss: 0.012849946506632143
Train Epoch: 10 [21760/60000 (36%)]     Loss: 0.05117126181721687
Train Epoch: 10 [22400/60000 (37%)]     Loss: 0.09611494839191437
Train Epoch: 10 [23040/60000 (38%)]     Loss: 0.030051004141569138
Train Epoch: 10 [23680/60000 (39%)]     Loss: 0.19895704090595245
Train Epoch: 10 [24320/60000 (41%)]     Loss: 0.05680454894900322
Train Epoch: 10 [24960/60000 (42%)]     Loss: 0.0382823646068573
Train Epoch: 10 [25600/60000 (43%)]     Loss: 0.060245268046855927
Train Epoch: 10 [26240/60000 (44%)]     Loss: 0.10107982903718948
Train Epoch: 10 [26880/60000 (45%)]     Loss: 0.02182743139564991
Train Epoch: 10 [27520/60000 (46%)]     Loss: 0.09794303774833679
Train Epoch: 10 [28160/60000 (47%)]     Loss: 0.10113205760717392
Train Epoch: 10 [28800/60000 (48%)]     Loss: 0.066297590732574446
Train Epoch: 10 [29440/60000 (49%)]     Loss: 0.004804034251719713
Train Epoch: 10 [30080/60000 (50%)]     Loss: 0.039265763014554983
Train Epoch: 10 [30720/60000 (51%)]     Loss: 0.047132015228271484
Train Epoch: 10 [31360/60000 (52%)]     Loss: 0.011537712067365646
Train Epoch: 10 [32000/60000 (53%)]     Loss: 0.09373736381530762
Train Epoch: 10 [32640/60000 (54%)]     Loss: 0.02799180895090103
Train Epoch: 10 [33280/60000 (55%)]     Loss: 0.012990034185349941
Train Epoch: 10 [33920/60000 (57%)]     Loss: 0.021683447062969208
Train Epoch: 10 [34560/60000 (58%)]     Loss: 0.019709544256329536
Train Epoch: 10 [35200/60000 (59%)]     Loss: 0.07398165017366409
Train Epoch: 10 [35840/60000 (60%)]     Loss: 0.09893017262220383
```

```
Train Epoch: 10 [36480/60000 (61%)]     Loss: 0.027068767696619034
Train Epoch: 10 [37120/60000 (62%)]     Loss: 0.03278297930955887
Train Epoch: 10 [37760/60000 (63%)]     Loss: 0.09515000134706497
Train Epoch: 10 [38400/60000 (64%)]     Loss: 0.05884203687310219
Train Epoch: 10 [39040/60000 (65%)]     Loss: 0.022077258676290512
Train Epoch: 10 [39680/60000 (66%)]     Loss: 0.015806578099972763
Train Epoch: 10 [40320/60000 (67%)]     Loss: 0.0066222501918673515
Train Epoch: 10 [40960/60000 (68%)]     Loss: 0.013557937927544117
Train Epoch: 10 [41600/60000 (69%)]     Loss: 0.031584933400154114
Train Epoch: 10 [42240/60000 (70%)]     Loss: 0.14985103905200958
Train Epoch: 10 [42880/60000 (71%)]     Loss: 0.11202509701251984
Train Epoch: 10 [43520/60000 (72%)]     Loss: 0.02185123600065708
Train Epoch: 10 [44160/60000 (74%)]     Loss: 0.10420534759759903
Train Epoch: 10 [44800/60000 (75%)]     Loss: 0.061281967908144
Train Epoch: 10 [45440/60000 (76%)]     Loss: 0.11793708056211472
Train Epoch: 10 [46080/60000 (77%)]     Loss: 0.1319528967142105
Train Epoch: 10 [46720/60000 (78%)]     Loss: 0.0661466047167778
Train Epoch: 10 [47360/60000 (79%)]     Loss: 0.00936101097613573
Train Epoch: 10 [48000/60000 (80%)]     Loss: 0.0068985121132465839
Train Epoch: 10 [48640/60000 (81%)]     Loss: 0.10856509208679199
Train Epoch: 10 [49280/60000 (82%)]     Loss: 0.06621689349412918
Train Epoch: 10 [49920/60000 (83%)]     Loss: 0.013828255236148834
Train Epoch: 10 [50560/60000 (84%)]     Loss: 0.004793965723365545
Train Epoch: 10 [51200/60000 (85%)]     Loss: 0.03448067978024483
Train Epoch: 10 [51840/60000 (86%)]     Loss: 0.014548074454069138
Train Epoch: 10 [52480/60000 (87%)]     Loss: 0.018600089475512505
Train Epoch: 10 [53120/60000 (88%)]     Loss: 0.015785392373800278
Train Epoch: 10 [53760/60000 (90%)]     Loss: 0.07669944316148758
Train Epoch: 10 [54400/60000 (91%)]     Loss: 0.003952422644942999
Train Epoch: 10 [55040/60000 (92%)]     Loss: 0.049044094979976303
Train Epoch: 10 [55680/60000 (93%)]     Loss: 0.019222848117351532
Train Epoch: 10 [56320/60000 (94%)]     Loss: 0.023120002821087837
Train Epoch: 10 [56960/60000 (95%)]     Loss: 0.015323969535529613
Train Epoch: 10 [57600/60000 (96%)]     Loss: 0.15106616914272308
Train Epoch: 10 [58240/60000 (97%)]     Loss: 0.02203264646232128
Train Epoch: 10 [58880/60000 (98%)]     Loss: 0.058005269616684227
Train Epoch: 10 [59520/60000 (99%)]     Loss: 0.04264214634895325
Epoch 10: Training Loss: 0.0008, Accuracy: 98.56%
Test Loss: 0.0001, Accuracy: 97.84%
```

[7]: ```
%tensorboard --logdir log/MNIST --port=8008
```

```
<IPython.core.display.HTML object>
```

[ ]: