

mnist_convnet

October 6, 2023

1 mnist_convnet

- ELEC 576 HW 1
- Robert Heeter
- 4 October 2023

1.1 Structure:

- 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:

- <https://colab.research.google.com/drive/1i9KpbQyFU4zfq8zLLns8a2Kd8PRMGsaZ>
- https://github.com/motokimura/pytorch_tensorboard/blob/master/main.py
- https://colab.research.google.com/github/tensorflow/tensorboard/blob/master/docs/tensorboard_in_notebooks.ipynb
- <https://github.com/christianversloot/machine-learning-articles/blob/main/how-to-use-tensorboard-with-pytorch.md>

```
[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
import numpy as np

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

%load_ext tensorboard

```

2023-10-06 01:46:54.245715: 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.

```

[2]: # 1. Set PyTorch metadata

batch_size = 64
test_batch_size = 1000
epochs = 10
lr = 0.01
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.manual_seed(seed)
else:
    cuda = False
    torch.manual_seed(seed)

```

```

[3]: # 2. Import data

transform = transforms.Compose([transforms.ToTensor(), transforms.Normalize((0.
    ↳ 01307,), (0.3081,))])

```

```

train_loader = torch.utils.data.DataLoader(datasets.MNIST('data', train=True,
↳download=True, transform=transform),
                                     batch_size=batch_size,
                                     shuffle=True)
test_loader = torch.utils.data.DataLoader(datasets.MNIST('data', train=False,
↳download=True, transform=transform),
                                     batch_size=test_batch_size,
                                     shuffle=True)

```

[4]: # 3. Defining model architecture, loss, and optimizer

```

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

        self.layers = nn.Sequential(nn.Conv2d(1, 10, kernel_size=5),
                                    nn.Tanh(),
                                    nn.MaxPool2d(2),
                                    nn.Conv2d(10, 20, kernel_size=5),
                                    nn.Tanh(),
                                    nn.MaxPool2d(2),
                                    nn.Flatten(),
                                    nn.Linear(320, 50),
                                    nn.Tanh(),
                                    nn.Dropout2d(0.5),
                                    nn.Linear(50, 10),
                                    nn.Softmax(dim=1)
                                )

    def forward(self, x):
        '''Forward pass'''
        return self.layers(x)

#         self.conv1 = nn.Conv2d(1, 10, kernel_size=5)
#         self.conv2 = nn.Conv2d(10, 20, kernel_size=5)
#         self.conv2_drop = nn.Dropout2d()
#         self.fc1 = nn.Linear(320, 50)
#         self.fc2 = nn.Linear(50, 10)

#     def forward(self, x):
#         # original network architecture
#         # x = F.relu(F.max_pool2d(self.conv1(x), 2))
#         # x = F.relu(F.max_pool2d(self.conv2_drop(self.conv2(x)), 2))
#         # x = x.view(-1, 320) # (batch_size, units)
#         # x = F.relu(self.fc1(x))

```

```

#         # x = F.dropout(x, training=self.training)
#         # x = self.fc2(x)
#         # x = F.softmax(x, dim=1)

#         # new network architecture
#         x = self.conv1(x)
#         x = F.relu(x)
#         x = F.max_pool2d(x, 2)
#         x = self.conv2(x)
#         x = F.relu(x)
#         x = F.max_pool2d(x, 2)
#         x = x.view(-1, 320)
#         x = self.fc1(x)
#         x = F.relu(x)
#         x = F.dropout(x, p=0.5)
#         x = self.fc2(x)
#         x = F.softmax(x, dim=1)

#         return x

model = Net()
print(model)

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

```

```

Net(
  (layers): Sequential(
    (0): Conv2d(1, 10, kernel_size=(5, 5), stride=(1, 1))
    (1): Tanh()
    (2): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceiling_mode=False)
    (3): Conv2d(10, 20, kernel_size=(5, 5), stride=(1, 1))
    (4): Tanh()
    (5): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceiling_mode=False)
    (6): Flatten(start_dim=1, end_dim=-1)
    (7): Linear(in_features=320, out_features=50, bias=True)
    (8): Tanh()
    (9): Dropout2d(p=0.5, inplace=False)
    (10): Linear(in_features=50, out_features=10, bias=True)
    (11): Softmax(dim=1)
  )
)
SGD (
Parameter Group 0
  dampening: 0

```

```

differentiable: False
foreach: None
lr: 0.01
maximize: False
momentum: 0
nesterov: False
weight_decay: 0
)

```

```

[5]: def weight_histograms_conv2d(writer, n_iter, weights, biases, layer_number):
    weights_shape = weights.shape
    num_kernels = weights_shape[0]

    for k in range(num_kernels):

        flattened_weights = weights[k].flatten()
        tag = f"layer_{layer_number}/kernel_{k}_weight_histogram"
        writer.add_histogram(tag, flattened_weights, n_iter, bins='tensorflow')

        tag = f"layer_{layer_number}/kernel_{k}_weight_min"
        writer.add_scalar(tag, torch.min(flattened_weights), n_iter)
        tag = f"layer_{layer_number}/kernel_{k}_weight_max"
        writer.add_scalar(tag, torch.max(flattened_weights), n_iter)
        tag = f"layer_{layer_number}/kernel_{k}_weight_mean"
        writer.add_scalar(tag, torch.mean(flattened_weights), n_iter)
        tag = f"layer_{layer_number}/kernel_{k}_weight_stddev"
        writer.add_scalar(tag, torch.std(flattened_weights), n_iter)

        flattened_biases = biases[k].flatten()
        tag = f"layer_{layer_number}/kernel_{k}_biases_histogram"
        writer.add_histogram(tag, flattened_biases, n_iter, bins='tensorflow')

        tag = f"layer_{layer_number}/kernel_{k}_biases_min"
        writer.add_scalar(tag, torch.min(flattened_biases), n_iter)
        tag = f"layer_{layer_number}/kernel_{k}_biases_max"
        writer.add_scalar(tag, torch.max(flattened_biases), n_iter)
        tag = f"layer_{layer_number}/kernel_{k}_biases_mean"
        writer.add_scalar(tag, torch.mean(flattened_biases), n_iter)
        tag = f"layer_{layer_number}/kernel_{k}_biases_stddev"
        writer.add_scalar(tag, torch.std(flattened_biases), n_iter)

    break

def weight_histograms_linear(writer, n_iter, weights, biases, layer_number):
    flattened_weights = weights.flatten()
    tag = f"layer_{layer_number}_weight_histogram"
    writer.add_histogram(tag, flattened_weights, n_iter, bins='tensorflow')

```

```

tag = f"layer_{layer_number}_weight_min"
writer.add_scalar(tag, torch.min(flattened_weights), n_iter)
tag = f"layer_{layer_number}_weight_max"
writer.add_scalar(tag, torch.max(flattened_weights), n_iter)
tag = f"layer_{layer_number}_weight_mean"
writer.add_scalar(tag, torch.mean(flattened_weights), n_iter)
tag = f"layer_{layer_number}_weight_stddev"
writer.add_scalar(tag, torch.std(flattened_weights), n_iter)

flattened_biases = biases.flatten()
tag = f"layer_{layer_number}_biases_histogram"
writer.add_histogram(tag, flattened_biases, n_iter, bins='tensorflow')

tag = f"layer_{layer_number}_biases_min"
writer.add_scalar(tag, torch.min(flattened_biases), n_iter)
tag = f"layer_{layer_number}_biases_max"
writer.add_scalar(tag, torch.max(flattened_biases), n_iter)
tag = f"layer_{layer_number}_biases_mean"
writer.add_scalar(tag, torch.mean(flattened_biases), n_iter)
tag = f"layer_{layer_number}_biases_stddev"
writer.add_scalar(tag, torch.std(flattened_biases), n_iter)

def weight_histograms(writer, n_iter, model):
    # iterate over all model layers
    for layer_number in range(len(model.layers)):

        # get layer
        layer = model.layers[layer_number]

        # compute weight histograms for appropriate layer
        if isinstance(layer, nn.Conv2d):
            weights = layer.weight
            biases = layer.bias
            weight_histograms_conv2d(writer, n_iter, weights, biases,
→layer_number)

            elif isinstance(layer, nn.Linear):
                weights = layer.weight
                biases = layer.bias
                weight_histograms_linear(writer, n_iter, weights, biases,
→layer_number)

```

[7]: *# 4. Define test and training loops*

```

eps = 1e-13

```

```

def train(epoch):
    model.train()

    # criterion = nn.CrossEntropyLoss()
    criterion = nn.NLLLoss(size_average=False)

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

        optimizer.zero_grad()
        output = model(data) # forward pass
        loss = criterion(torch.log(output+eps), target) # = sum_k(-t_k *
→ log(y_k))
        loss.backward() # backward pass
        optimizer.step()

        if batch_idx % logging_interval == 0:
            print('Train Epoch: {} [{}/{} ({:.0f}%)]\tLoss: {:.6f}'.format(
                epoch, batch_idx * len(data), len(train_loader.dataset),
                100. * batch_idx / len(train_loader), loss.data)
            )

            # log train/loss to TensorBoard at every iteration
            n_iter = (epoch - 1) * len(train_loader) + batch_idx + 1
            writer.add_scalar('train/loss', loss.data, n_iter)

    # visualize layer weights and biases
    weight_histograms(writer, n_iter, model)

    # log model parameters to TensorBoard at every epoch
    for name, param in model.named_parameters():
        layer, attr = os.path.splitext(name)
        attr = attr[1:]
        writer.add_histogram('{}/{}'.format(layer, attr), param.clone().cpu().
→ data.numpy(), n_iter)

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

    # criterion = nn.CrossEntropyLoss()
    # criterion = nn.CrossEntropyLoss(size_average = False)
    criterion = nn.NLLLoss(size_average = False)

    for data, target in test_loader:

```

```

    if cuda:
        data, target = data.cuda(), target.cuda()

    output = model(data)

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

    test_loss /= len(test_loader.dataset)
    test_accuracy = 100. * correct / len(test_loader.dataset)
    print('\nTest set: Average loss: {:.4f}, Accuracy: {}/{} ({:.2f}%)\n'.
    ↪ format(
        test_loss, correct, len(test_loader.dataset), test_accuracy)
    )

    # log test/loss and test/accuracy to TensorBoard at every epoch
    n_iter = epoch * len(train_loader)
    writer.add_scalar('test/loss', test_loss, n_iter)
    writer.add_scalar('test/accuracy', test_accuracy, n_iter)

```

[8]: # 5. Perform training over multiple epochs

```

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

writer.close()

```

/Users/rch/opt/anaconda3/lib/python3.9/site-packages/torch/nn/_reduction.py:42: UserWarning: size_average and reduce args will be deprecated, please use reduction='sum' instead.

```
warnings.warn(warning.format(ret))
```

/Users/rch/opt/anaconda3/lib/python3.9/site-packages/torch/nn/functional.py:1345: UserWarning: dropout2d: Received a 2-D input to dropout2d, which is deprecated and will result in an error in a future release. To retain the behavior and silence this warning, please use dropout instead. Note that dropout2d exists to provide channel-wise dropout on inputs with 2 spatial dimensions, a channel dimension, and an optional batch dimension (i.e. 3D or 4D inputs).

```
warnings.warn(warn_msg)
```

Train Epoch: 1 [0/60000 (0%)] Loss: 148.350174

Train Epoch: 1 [640/60000 (1%)] Loss: 65.574791

Train Epoch: 1	[1280/60000 (2%)]	Loss: 50.805614
Train Epoch: 1	[1920/60000 (3%)]	Loss: 27.808430
Train Epoch: 1	[2560/60000 (4%)]	Loss: 29.381617
Train Epoch: 1	[3200/60000 (5%)]	Loss: 47.149063
Train Epoch: 1	[3840/60000 (6%)]	Loss: 27.568836
Train Epoch: 1	[4480/60000 (7%)]	Loss: 25.695684
Train Epoch: 1	[5120/60000 (9%)]	Loss: 32.483009
Train Epoch: 1	[5760/60000 (10%)]	Loss: 28.119480
Train Epoch: 1	[6400/60000 (11%)]	Loss: 16.325533
Train Epoch: 1	[7040/60000 (12%)]	Loss: 21.478676
Train Epoch: 1	[7680/60000 (13%)]	Loss: 23.415968
Train Epoch: 1	[8320/60000 (14%)]	Loss: 21.872169
Train Epoch: 1	[8960/60000 (15%)]	Loss: 13.121871
Train Epoch: 1	[9600/60000 (16%)]	Loss: 13.346699
Train Epoch: 1	[10240/60000 (17%)]	Loss: 30.787872
Train Epoch: 1	[10880/60000 (18%)]	Loss: 22.634539
Train Epoch: 1	[11520/60000 (19%)]	Loss: 25.395401
Train Epoch: 1	[12160/60000 (20%)]	Loss: 14.447589
Train Epoch: 1	[12800/60000 (21%)]	Loss: 10.730078
Train Epoch: 1	[13440/60000 (22%)]	Loss: 12.101591
Train Epoch: 1	[14080/60000 (23%)]	Loss: 55.785133
Train Epoch: 1	[14720/60000 (25%)]	Loss: 10.913371
Train Epoch: 1	[15360/60000 (26%)]	Loss: 15.996431
Train Epoch: 1	[16000/60000 (27%)]	Loss: 12.564799
Train Epoch: 1	[16640/60000 (28%)]	Loss: 10.983263
Train Epoch: 1	[17280/60000 (29%)]	Loss: 8.856709
Train Epoch: 1	[17920/60000 (30%)]	Loss: 8.925025
Train Epoch: 1	[18560/60000 (31%)]	Loss: 13.580734
Train Epoch: 1	[19200/60000 (32%)]	Loss: 23.908947
Train Epoch: 1	[19840/60000 (33%)]	Loss: 9.361331
Train Epoch: 1	[20480/60000 (34%)]	Loss: 26.058790
Train Epoch: 1	[21120/60000 (35%)]	Loss: 20.889919
Train Epoch: 1	[21760/60000 (36%)]	Loss: 16.410866
Train Epoch: 1	[22400/60000 (37%)]	Loss: 20.695486
Train Epoch: 1	[23040/60000 (38%)]	Loss: 8.645380
Train Epoch: 1	[23680/60000 (39%)]	Loss: 2.316973
Train Epoch: 1	[24320/60000 (41%)]	Loss: 12.669093
Train Epoch: 1	[24960/60000 (42%)]	Loss: 12.652076
Train Epoch: 1	[25600/60000 (43%)]	Loss: 6.577795
Train Epoch: 1	[26240/60000 (44%)]	Loss: 15.498992
Train Epoch: 1	[26880/60000 (45%)]	Loss: 10.603088
Train Epoch: 1	[27520/60000 (46%)]	Loss: 6.678514
Train Epoch: 1	[28160/60000 (47%)]	Loss: 4.343204
Train Epoch: 1	[28800/60000 (48%)]	Loss: 26.278395
Train Epoch: 1	[29440/60000 (49%)]	Loss: 9.498826
Train Epoch: 1	[30080/60000 (50%)]	Loss: 11.930044
Train Epoch: 1	[30720/60000 (51%)]	Loss: 16.838755
Train Epoch: 1	[31360/60000 (52%)]	Loss: 7.070224

Train Epoch: 1	[32000/60000 (53%)]	Loss: 18.772337
Train Epoch: 1	[32640/60000 (54%)]	Loss: 17.717880
Train Epoch: 1	[33280/60000 (55%)]	Loss: 6.297252
Train Epoch: 1	[33920/60000 (57%)]	Loss: 15.736999
Train Epoch: 1	[34560/60000 (58%)]	Loss: 9.797343
Train Epoch: 1	[35200/60000 (59%)]	Loss: 6.874809
Train Epoch: 1	[35840/60000 (60%)]	Loss: 16.641665
Train Epoch: 1	[36480/60000 (61%)]	Loss: 18.873581
Train Epoch: 1	[37120/60000 (62%)]	Loss: 10.327193
Train Epoch: 1	[37760/60000 (63%)]	Loss: 10.541352
Train Epoch: 1	[38400/60000 (64%)]	Loss: 20.968702
Train Epoch: 1	[39040/60000 (65%)]	Loss: 14.714314
Train Epoch: 1	[39680/60000 (66%)]	Loss: 8.235843
Train Epoch: 1	[40320/60000 (67%)]	Loss: 8.279308
Train Epoch: 1	[40960/60000 (68%)]	Loss: 12.288799
Train Epoch: 1	[41600/60000 (69%)]	Loss: 2.113179
Train Epoch: 1	[42240/60000 (70%)]	Loss: 12.241662
Train Epoch: 1	[42880/60000 (71%)]	Loss: 9.805692
Train Epoch: 1	[43520/60000 (72%)]	Loss: 18.253258
Train Epoch: 1	[44160/60000 (74%)]	Loss: 10.824991
Train Epoch: 1	[44800/60000 (75%)]	Loss: 5.566197
Train Epoch: 1	[45440/60000 (76%)]	Loss: 14.457938
Train Epoch: 1	[46080/60000 (77%)]	Loss: 10.472195
Train Epoch: 1	[46720/60000 (78%)]	Loss: 1.622838
Train Epoch: 1	[47360/60000 (79%)]	Loss: 32.133198
Train Epoch: 1	[48000/60000 (80%)]	Loss: 22.061737
Train Epoch: 1	[48640/60000 (81%)]	Loss: 6.789975
Train Epoch: 1	[49280/60000 (82%)]	Loss: 23.259237
Train Epoch: 1	[49920/60000 (83%)]	Loss: 9.818250
Train Epoch: 1	[50560/60000 (84%)]	Loss: 8.120830
Train Epoch: 1	[51200/60000 (85%)]	Loss: 19.550013
Train Epoch: 1	[51840/60000 (86%)]	Loss: 7.249803
Train Epoch: 1	[52480/60000 (87%)]	Loss: 11.707809
Train Epoch: 1	[53120/60000 (88%)]	Loss: 10.630896
Train Epoch: 1	[53760/60000 (90%)]	Loss: 7.630034
Train Epoch: 1	[54400/60000 (91%)]	Loss: 9.601753
Train Epoch: 1	[55040/60000 (92%)]	Loss: 5.875074
Train Epoch: 1	[55680/60000 (93%)]	Loss: 16.653101
Train Epoch: 1	[56320/60000 (94%)]	Loss: 7.744590
Train Epoch: 1	[56960/60000 (95%)]	Loss: 10.115010
Train Epoch: 1	[57600/60000 (96%)]	Loss: 8.761250
Train Epoch: 1	[58240/60000 (97%)]	Loss: 2.970207
Train Epoch: 1	[58880/60000 (98%)]	Loss: 4.338736
Train Epoch: 1	[59520/60000 (99%)]	Loss: 10.248787

Test set: Average loss: 0.0816, Accuracy: 9742/10000 (97.42%)

Train Epoch: 2 [0/60000 (0%)] Loss: 5.794899

Train Epoch: 2	[640/60000 (1%)]	Loss: 7.820867
Train Epoch: 2	[1280/60000 (2%)]	Loss: 9.429663
Train Epoch: 2	[1920/60000 (3%)]	Loss: 3.547286
Train Epoch: 2	[2560/60000 (4%)]	Loss: 8.738937
Train Epoch: 2	[3200/60000 (5%)]	Loss: 12.709581
Train Epoch: 2	[3840/60000 (6%)]	Loss: 6.283051
Train Epoch: 2	[4480/60000 (7%)]	Loss: 10.825681
Train Epoch: 2	[5120/60000 (9%)]	Loss: 1.381034
Train Epoch: 2	[5760/60000 (10%)]	Loss: 9.333070
Train Epoch: 2	[6400/60000 (11%)]	Loss: 6.135649
Train Epoch: 2	[7040/60000 (12%)]	Loss: 7.319004
Train Epoch: 2	[7680/60000 (13%)]	Loss: 3.200693
Train Epoch: 2	[8320/60000 (14%)]	Loss: 3.356311
Train Epoch: 2	[8960/60000 (15%)]	Loss: 18.460999
Train Epoch: 2	[9600/60000 (16%)]	Loss: 12.877852
Train Epoch: 2	[10240/60000 (17%)]	Loss: 4.580849
Train Epoch: 2	[10880/60000 (18%)]	Loss: 7.309263
Train Epoch: 2	[11520/60000 (19%)]	Loss: 9.147858
Train Epoch: 2	[12160/60000 (20%)]	Loss: 8.414282
Train Epoch: 2	[12800/60000 (21%)]	Loss: 2.468552
Train Epoch: 2	[13440/60000 (22%)]	Loss: 11.544351
Train Epoch: 2	[14080/60000 (23%)]	Loss: 5.105357
Train Epoch: 2	[14720/60000 (25%)]	Loss: 2.410307
Train Epoch: 2	[15360/60000 (26%)]	Loss: 1.240903
Train Epoch: 2	[16000/60000 (27%)]	Loss: 3.651239
Train Epoch: 2	[16640/60000 (28%)]	Loss: 14.097479
Train Epoch: 2	[17280/60000 (29%)]	Loss: 11.660074
Train Epoch: 2	[17920/60000 (30%)]	Loss: 9.857907
Train Epoch: 2	[18560/60000 (31%)]	Loss: 11.347275
Train Epoch: 2	[19200/60000 (32%)]	Loss: 15.710032
Train Epoch: 2	[19840/60000 (33%)]	Loss: 7.412633
Train Epoch: 2	[20480/60000 (34%)]	Loss: 9.905286
Train Epoch: 2	[21120/60000 (35%)]	Loss: 5.549774
Train Epoch: 2	[21760/60000 (36%)]	Loss: 16.673695
Train Epoch: 2	[22400/60000 (37%)]	Loss: 15.923005
Train Epoch: 2	[23040/60000 (38%)]	Loss: 10.912141
Train Epoch: 2	[23680/60000 (39%)]	Loss: 9.906033
Train Epoch: 2	[24320/60000 (41%)]	Loss: 4.116310
Train Epoch: 2	[24960/60000 (42%)]	Loss: 1.313582
Train Epoch: 2	[25600/60000 (43%)]	Loss: 6.027876
Train Epoch: 2	[26240/60000 (44%)]	Loss: 11.249135
Train Epoch: 2	[26880/60000 (45%)]	Loss: 10.030438
Train Epoch: 2	[27520/60000 (46%)]	Loss: 8.641503
Train Epoch: 2	[28160/60000 (47%)]	Loss: 5.242270
Train Epoch: 2	[28800/60000 (48%)]	Loss: 3.222556
Train Epoch: 2	[29440/60000 (49%)]	Loss: 4.121618
Train Epoch: 2	[30080/60000 (50%)]	Loss: 5.472337
Train Epoch: 2	[30720/60000 (51%)]	Loss: 9.213414

Train Epoch: 2	[31360/60000 (52%)]	Loss: 3.990453
Train Epoch: 2	[32000/60000 (53%)]	Loss: 4.348527
Train Epoch: 2	[32640/60000 (54%)]	Loss: 4.675620
Train Epoch: 2	[33280/60000 (55%)]	Loss: 21.696838
Train Epoch: 2	[33920/60000 (57%)]	Loss: 0.893426
Train Epoch: 2	[34560/60000 (58%)]	Loss: 5.837167
Train Epoch: 2	[35200/60000 (59%)]	Loss: 5.856362
Train Epoch: 2	[35840/60000 (60%)]	Loss: 6.947505
Train Epoch: 2	[36480/60000 (61%)]	Loss: 9.290040
Train Epoch: 2	[37120/60000 (62%)]	Loss: 11.697110
Train Epoch: 2	[37760/60000 (63%)]	Loss: 12.097234
Train Epoch: 2	[38400/60000 (64%)]	Loss: 6.305795
Train Epoch: 2	[39040/60000 (65%)]	Loss: 2.917878
Train Epoch: 2	[39680/60000 (66%)]	Loss: 7.600556
Train Epoch: 2	[40320/60000 (67%)]	Loss: 4.543109
Train Epoch: 2	[40960/60000 (68%)]	Loss: 6.848329
Train Epoch: 2	[41600/60000 (69%)]	Loss: 11.867298
Train Epoch: 2	[42240/60000 (70%)]	Loss: 6.556248
Train Epoch: 2	[42880/60000 (71%)]	Loss: 10.470507
Train Epoch: 2	[43520/60000 (72%)]	Loss: 2.900258
Train Epoch: 2	[44160/60000 (74%)]	Loss: 10.441376
Train Epoch: 2	[44800/60000 (75%)]	Loss: 7.591688
Train Epoch: 2	[45440/60000 (76%)]	Loss: 16.178593
Train Epoch: 2	[46080/60000 (77%)]	Loss: 6.151915
Train Epoch: 2	[46720/60000 (78%)]	Loss: 16.359280
Train Epoch: 2	[47360/60000 (79%)]	Loss: 16.866432
Train Epoch: 2	[48000/60000 (80%)]	Loss: 5.174510
Train Epoch: 2	[48640/60000 (81%)]	Loss: 7.196651
Train Epoch: 2	[49280/60000 (82%)]	Loss: 6.357285
Train Epoch: 2	[49920/60000 (83%)]	Loss: 16.031305
Train Epoch: 2	[50560/60000 (84%)]	Loss: 5.842550
Train Epoch: 2	[51200/60000 (85%)]	Loss: 20.573612
Train Epoch: 2	[51840/60000 (86%)]	Loss: 5.943273
Train Epoch: 2	[52480/60000 (87%)]	Loss: 2.580393
Train Epoch: 2	[53120/60000 (88%)]	Loss: 7.100457
Train Epoch: 2	[53760/60000 (90%)]	Loss: 26.377951
Train Epoch: 2	[54400/60000 (91%)]	Loss: 7.567432
Train Epoch: 2	[55040/60000 (92%)]	Loss: 3.035346
Train Epoch: 2	[55680/60000 (93%)]	Loss: 2.375730
Train Epoch: 2	[56320/60000 (94%)]	Loss: 12.264088
Train Epoch: 2	[56960/60000 (95%)]	Loss: 8.824215
Train Epoch: 2	[57600/60000 (96%)]	Loss: 5.415017
Train Epoch: 2	[58240/60000 (97%)]	Loss: 3.947428
Train Epoch: 2	[58880/60000 (98%)]	Loss: 5.410202
Train Epoch: 2	[59520/60000 (99%)]	Loss: 4.062138

Test set: Average loss: 0.0692, Accuracy: 9783/10000 (97.83%)

Train Epoch: 3	[0/60000 (0%)]	Loss: 2.392176
Train Epoch: 3	[640/60000 (1%)]	Loss: 23.679876
Train Epoch: 3	[1280/60000 (2%)]	Loss: 5.030544
Train Epoch: 3	[1920/60000 (3%)]	Loss: 19.759233
Train Epoch: 3	[2560/60000 (4%)]	Loss: 8.832271
Train Epoch: 3	[3200/60000 (5%)]	Loss: 13.441246
Train Epoch: 3	[3840/60000 (6%)]	Loss: 13.123025
Train Epoch: 3	[4480/60000 (7%)]	Loss: 7.412638
Train Epoch: 3	[5120/60000 (9%)]	Loss: 4.251581
Train Epoch: 3	[5760/60000 (10%)]	Loss: 5.962806
Train Epoch: 3	[6400/60000 (11%)]	Loss: 6.662619
Train Epoch: 3	[7040/60000 (12%)]	Loss: 2.393607
Train Epoch: 3	[7680/60000 (13%)]	Loss: 2.496614
Train Epoch: 3	[8320/60000 (14%)]	Loss: 1.160033
Train Epoch: 3	[8960/60000 (15%)]	Loss: 7.224338
Train Epoch: 3	[9600/60000 (16%)]	Loss: 6.754421
Train Epoch: 3	[10240/60000 (17%)]	Loss: 2.993656
Train Epoch: 3	[10880/60000 (18%)]	Loss: 9.229931
Train Epoch: 3	[11520/60000 (19%)]	Loss: 2.856031
Train Epoch: 3	[12160/60000 (20%)]	Loss: 12.549516
Train Epoch: 3	[12800/60000 (21%)]	Loss: 8.845797
Train Epoch: 3	[13440/60000 (22%)]	Loss: 1.208502
Train Epoch: 3	[14080/60000 (23%)]	Loss: 1.130532
Train Epoch: 3	[14720/60000 (25%)]	Loss: 3.245269
Train Epoch: 3	[15360/60000 (26%)]	Loss: 10.745577
Train Epoch: 3	[16000/60000 (27%)]	Loss: 13.248148
Train Epoch: 3	[16640/60000 (28%)]	Loss: 5.076397
Train Epoch: 3	[17280/60000 (29%)]	Loss: 6.256505
Train Epoch: 3	[17920/60000 (30%)]	Loss: 3.376264
Train Epoch: 3	[18560/60000 (31%)]	Loss: 4.911217
Train Epoch: 3	[19200/60000 (32%)]	Loss: 4.161876
Train Epoch: 3	[19840/60000 (33%)]	Loss: 5.511187
Train Epoch: 3	[20480/60000 (34%)]	Loss: 9.990060
Train Epoch: 3	[21120/60000 (35%)]	Loss: 6.755213
Train Epoch: 3	[21760/60000 (36%)]	Loss: 6.218166
Train Epoch: 3	[22400/60000 (37%)]	Loss: 9.202353
Train Epoch: 3	[23040/60000 (38%)]	Loss: 4.193927
Train Epoch: 3	[23680/60000 (39%)]	Loss: 3.067010
Train Epoch: 3	[24320/60000 (41%)]	Loss: 2.352905
Train Epoch: 3	[24960/60000 (42%)]	Loss: 0.969219
Train Epoch: 3	[25600/60000 (43%)]	Loss: 4.735764
Train Epoch: 3	[26240/60000 (44%)]	Loss: 8.341851
Train Epoch: 3	[26880/60000 (45%)]	Loss: 2.404269
Train Epoch: 3	[27520/60000 (46%)]	Loss: 0.940102
Train Epoch: 3	[28160/60000 (47%)]	Loss: 3.918036
Train Epoch: 3	[28800/60000 (48%)]	Loss: 12.873190
Train Epoch: 3	[29440/60000 (49%)]	Loss: 10.763255
Train Epoch: 3	[30080/60000 (50%)]	Loss: 4.989667

Train Epoch: 3	[30720/60000 (51%)]	Loss: 8.735826
Train Epoch: 3	[31360/60000 (52%)]	Loss: 1.024777
Train Epoch: 3	[32000/60000 (53%)]	Loss: 3.098973
Train Epoch: 3	[32640/60000 (54%)]	Loss: 12.592302
Train Epoch: 3	[33280/60000 (55%)]	Loss: 9.946776
Train Epoch: 3	[33920/60000 (57%)]	Loss: 4.094009
Train Epoch: 3	[34560/60000 (58%)]	Loss: 3.818836
Train Epoch: 3	[35200/60000 (59%)]	Loss: 1.338328
Train Epoch: 3	[35840/60000 (60%)]	Loss: 9.876770
Train Epoch: 3	[36480/60000 (61%)]	Loss: 3.634490
Train Epoch: 3	[37120/60000 (62%)]	Loss: 0.821941
Train Epoch: 3	[37760/60000 (63%)]	Loss: 5.335407
Train Epoch: 3	[38400/60000 (64%)]	Loss: 1.914721
Train Epoch: 3	[39040/60000 (65%)]	Loss: 9.632752
Train Epoch: 3	[39680/60000 (66%)]	Loss: 6.033620
Train Epoch: 3	[40320/60000 (67%)]	Loss: 6.842088
Train Epoch: 3	[40960/60000 (68%)]	Loss: 1.610606
Train Epoch: 3	[41600/60000 (69%)]	Loss: 1.251294
Train Epoch: 3	[42240/60000 (70%)]	Loss: 20.121061
Train Epoch: 3	[42880/60000 (71%)]	Loss: 7.533456
Train Epoch: 3	[43520/60000 (72%)]	Loss: 8.079223
Train Epoch: 3	[44160/60000 (74%)]	Loss: 2.490916
Train Epoch: 3	[44800/60000 (75%)]	Loss: 13.690059
Train Epoch: 3	[45440/60000 (76%)]	Loss: 17.511292
Train Epoch: 3	[46080/60000 (77%)]	Loss: 14.385363
Train Epoch: 3	[46720/60000 (78%)]	Loss: 6.217794
Train Epoch: 3	[47360/60000 (79%)]	Loss: 18.025816
Train Epoch: 3	[48000/60000 (80%)]	Loss: 4.146681
Train Epoch: 3	[48640/60000 (81%)]	Loss: 7.283103
Train Epoch: 3	[49280/60000 (82%)]	Loss: 8.716898
Train Epoch: 3	[49920/60000 (83%)]	Loss: 12.746178
Train Epoch: 3	[50560/60000 (84%)]	Loss: 10.398217
Train Epoch: 3	[51200/60000 (85%)]	Loss: 11.219976
Train Epoch: 3	[51840/60000 (86%)]	Loss: 8.326018
Train Epoch: 3	[52480/60000 (87%)]	Loss: 7.930647
Train Epoch: 3	[53120/60000 (88%)]	Loss: 4.383054
Train Epoch: 3	[53760/60000 (90%)]	Loss: 6.676928
Train Epoch: 3	[54400/60000 (91%)]	Loss: 5.456255
Train Epoch: 3	[55040/60000 (92%)]	Loss: 6.304915
Train Epoch: 3	[55680/60000 (93%)]	Loss: 6.002474
Train Epoch: 3	[56320/60000 (94%)]	Loss: 1.274454
Train Epoch: 3	[56960/60000 (95%)]	Loss: 0.718004
Train Epoch: 3	[57600/60000 (96%)]	Loss: 7.351045
Train Epoch: 3	[58240/60000 (97%)]	Loss: 1.383693
Train Epoch: 3	[58880/60000 (98%)]	Loss: 1.011876
Train Epoch: 3	[59520/60000 (99%)]	Loss: 19.323925

Test set: Average loss: 0.0581, Accuracy: 9817/10000 (98.17%)

Train Epoch: 4	[0/60000 (0%)]	Loss: 1.760554
Train Epoch: 4	[640/60000 (1%)]	Loss: 10.024610
Train Epoch: 4	[1280/60000 (2%)]	Loss: 2.663712
Train Epoch: 4	[1920/60000 (3%)]	Loss: 6.489213
Train Epoch: 4	[2560/60000 (4%)]	Loss: 7.033193
Train Epoch: 4	[3200/60000 (5%)]	Loss: 2.532507
Train Epoch: 4	[3840/60000 (6%)]	Loss: 11.275025
Train Epoch: 4	[4480/60000 (7%)]	Loss: 11.003782
Train Epoch: 4	[5120/60000 (9%)]	Loss: 6.053628
Train Epoch: 4	[5760/60000 (10%)]	Loss: 3.514347
Train Epoch: 4	[6400/60000 (11%)]	Loss: 3.162746
Train Epoch: 4	[7040/60000 (12%)]	Loss: 1.572698
Train Epoch: 4	[7680/60000 (13%)]	Loss: 8.379597
Train Epoch: 4	[8320/60000 (14%)]	Loss: 1.317598
Train Epoch: 4	[8960/60000 (15%)]	Loss: 3.214137
Train Epoch: 4	[9600/60000 (16%)]	Loss: 12.529469
Train Epoch: 4	[10240/60000 (17%)]	Loss: 4.253539
Train Epoch: 4	[10880/60000 (18%)]	Loss: 5.843931
Train Epoch: 4	[11520/60000 (19%)]	Loss: 8.946067
Train Epoch: 4	[12160/60000 (20%)]	Loss: 8.974072
Train Epoch: 4	[12800/60000 (21%)]	Loss: 6.495608
Train Epoch: 4	[13440/60000 (22%)]	Loss: 2.123494
Train Epoch: 4	[14080/60000 (23%)]	Loss: 1.127161
Train Epoch: 4	[14720/60000 (25%)]	Loss: 4.231235
Train Epoch: 4	[15360/60000 (26%)]	Loss: 16.105898
Train Epoch: 4	[16000/60000 (27%)]	Loss: 4.430163
Train Epoch: 4	[16640/60000 (28%)]	Loss: 1.264707
Train Epoch: 4	[17280/60000 (29%)]	Loss: 1.233963
Train Epoch: 4	[17920/60000 (30%)]	Loss: 3.708989
Train Epoch: 4	[18560/60000 (31%)]	Loss: 15.077209
Train Epoch: 4	[19200/60000 (32%)]	Loss: 18.212618
Train Epoch: 4	[19840/60000 (33%)]	Loss: 2.751913
Train Epoch: 4	[20480/60000 (34%)]	Loss: 2.928551
Train Epoch: 4	[21120/60000 (35%)]	Loss: 7.999966
Train Epoch: 4	[21760/60000 (36%)]	Loss: 2.869562
Train Epoch: 4	[22400/60000 (37%)]	Loss: 9.410314
Train Epoch: 4	[23040/60000 (38%)]	Loss: 14.839697
Train Epoch: 4	[23680/60000 (39%)]	Loss: 5.865270
Train Epoch: 4	[24320/60000 (41%)]	Loss: 6.881284
Train Epoch: 4	[24960/60000 (42%)]	Loss: 12.897303
Train Epoch: 4	[25600/60000 (43%)]	Loss: 6.068412
Train Epoch: 4	[26240/60000 (44%)]	Loss: 9.807303
Train Epoch: 4	[26880/60000 (45%)]	Loss: 8.685884
Train Epoch: 4	[27520/60000 (46%)]	Loss: 2.663198
Train Epoch: 4	[28160/60000 (47%)]	Loss: 14.199425
Train Epoch: 4	[28800/60000 (48%)]	Loss: 11.116791
Train Epoch: 4	[29440/60000 (49%)]	Loss: 2.402918

Train Epoch: 4	[30080/60000 (50%)]	Loss: 4.401239
Train Epoch: 4	[30720/60000 (51%)]	Loss: 4.655196
Train Epoch: 4	[31360/60000 (52%)]	Loss: 4.891407
Train Epoch: 4	[32000/60000 (53%)]	Loss: 2.961356
Train Epoch: 4	[32640/60000 (54%)]	Loss: 10.901689
Train Epoch: 4	[33280/60000 (55%)]	Loss: 0.554322
Train Epoch: 4	[33920/60000 (57%)]	Loss: 9.157825
Train Epoch: 4	[34560/60000 (58%)]	Loss: 6.211045
Train Epoch: 4	[35200/60000 (59%)]	Loss: 3.544980
Train Epoch: 4	[35840/60000 (60%)]	Loss: 5.057472
Train Epoch: 4	[36480/60000 (61%)]	Loss: 3.552627
Train Epoch: 4	[37120/60000 (62%)]	Loss: 2.418842
Train Epoch: 4	[37760/60000 (63%)]	Loss: 12.234315
Train Epoch: 4	[38400/60000 (64%)]	Loss: 2.502735
Train Epoch: 4	[39040/60000 (65%)]	Loss: 5.953590
Train Epoch: 4	[39680/60000 (66%)]	Loss: 8.780099
Train Epoch: 4	[40320/60000 (67%)]	Loss: 1.174813
Train Epoch: 4	[40960/60000 (68%)]	Loss: 4.357615
Train Epoch: 4	[41600/60000 (69%)]	Loss: 6.912323
Train Epoch: 4	[42240/60000 (70%)]	Loss: 6.688109
Train Epoch: 4	[42880/60000 (71%)]	Loss: 5.775367
Train Epoch: 4	[43520/60000 (72%)]	Loss: 2.361507
Train Epoch: 4	[44160/60000 (74%)]	Loss: 3.172145
Train Epoch: 4	[44800/60000 (75%)]	Loss: 5.093514
Train Epoch: 4	[45440/60000 (76%)]	Loss: 5.098560
Train Epoch: 4	[46080/60000 (77%)]	Loss: 4.200459
Train Epoch: 4	[46720/60000 (78%)]	Loss: 19.048227
Train Epoch: 4	[47360/60000 (79%)]	Loss: 5.187266
Train Epoch: 4	[48000/60000 (80%)]	Loss: 4.273799
Train Epoch: 4	[48640/60000 (81%)]	Loss: 3.563021
Train Epoch: 4	[49280/60000 (82%)]	Loss: 2.276229
Train Epoch: 4	[49920/60000 (83%)]	Loss: 4.273859
Train Epoch: 4	[50560/60000 (84%)]	Loss: 8.572746
Train Epoch: 4	[51200/60000 (85%)]	Loss: 1.500271
Train Epoch: 4	[51840/60000 (86%)]	Loss: 9.243051
Train Epoch: 4	[52480/60000 (87%)]	Loss: 2.367111
Train Epoch: 4	[53120/60000 (88%)]	Loss: 4.907752
Train Epoch: 4	[53760/60000 (90%)]	Loss: 0.582087
Train Epoch: 4	[54400/60000 (91%)]	Loss: 11.228354
Train Epoch: 4	[55040/60000 (92%)]	Loss: 11.071354
Train Epoch: 4	[55680/60000 (93%)]	Loss: 2.255403
Train Epoch: 4	[56320/60000 (94%)]	Loss: 4.642889
Train Epoch: 4	[56960/60000 (95%)]	Loss: 13.126541
Train Epoch: 4	[57600/60000 (96%)]	Loss: 5.112108
Train Epoch: 4	[58240/60000 (97%)]	Loss: 0.478988
Train Epoch: 4	[58880/60000 (98%)]	Loss: 4.067495
Train Epoch: 4	[59520/60000 (99%)]	Loss: 7.656174

Test set: Average loss: 0.0544, Accuracy: 9806/10000 (98.06%)

Train Epoch: 5	[0/60000 (0%)]	Loss: 5.114225
Train Epoch: 5	[640/60000 (1%)]	Loss: 15.170694
Train Epoch: 5	[1280/60000 (2%)]	Loss: 6.286937
Train Epoch: 5	[1920/60000 (3%)]	Loss: 9.043308
Train Epoch: 5	[2560/60000 (4%)]	Loss: 3.879147
Train Epoch: 5	[3200/60000 (5%)]	Loss: 8.280998
Train Epoch: 5	[3840/60000 (6%)]	Loss: 7.126341
Train Epoch: 5	[4480/60000 (7%)]	Loss: 2.286620
Train Epoch: 5	[5120/60000 (9%)]	Loss: 8.981131
Train Epoch: 5	[5760/60000 (10%)]	Loss: 4.134356
Train Epoch: 5	[6400/60000 (11%)]	Loss: 8.072508
Train Epoch: 5	[7040/60000 (12%)]	Loss: 1.396859
Train Epoch: 5	[7680/60000 (13%)]	Loss: 1.746360
Train Epoch: 5	[8320/60000 (14%)]	Loss: 0.627814
Train Epoch: 5	[8960/60000 (15%)]	Loss: 3.069103
Train Epoch: 5	[9600/60000 (16%)]	Loss: 3.112091
Train Epoch: 5	[10240/60000 (17%)]	Loss: 9.434679
Train Epoch: 5	[10880/60000 (18%)]	Loss: 3.272834
Train Epoch: 5	[11520/60000 (19%)]	Loss: 1.694075
Train Epoch: 5	[12160/60000 (20%)]	Loss: 4.576614
Train Epoch: 5	[12800/60000 (21%)]	Loss: 7.254579
Train Epoch: 5	[13440/60000 (22%)]	Loss: 11.116033
Train Epoch: 5	[14080/60000 (23%)]	Loss: 19.152424
Train Epoch: 5	[14720/60000 (25%)]	Loss: 1.422675
Train Epoch: 5	[15360/60000 (26%)]	Loss: 4.007272
Train Epoch: 5	[16000/60000 (27%)]	Loss: 0.750134
Train Epoch: 5	[16640/60000 (28%)]	Loss: 3.925351
Train Epoch: 5	[17280/60000 (29%)]	Loss: 21.118919
Train Epoch: 5	[17920/60000 (30%)]	Loss: 12.813771
Train Epoch: 5	[18560/60000 (31%)]	Loss: 3.368471
Train Epoch: 5	[19200/60000 (32%)]	Loss: 2.883131
Train Epoch: 5	[19840/60000 (33%)]	Loss: 4.932166
Train Epoch: 5	[20480/60000 (34%)]	Loss: 1.483761
Train Epoch: 5	[21120/60000 (35%)]	Loss: 2.958440
Train Epoch: 5	[21760/60000 (36%)]	Loss: 9.304618
Train Epoch: 5	[22400/60000 (37%)]	Loss: 2.724267
Train Epoch: 5	[23040/60000 (38%)]	Loss: 0.943876
Train Epoch: 5	[23680/60000 (39%)]	Loss: 6.344690
Train Epoch: 5	[24320/60000 (41%)]	Loss: 8.505790
Train Epoch: 5	[24960/60000 (42%)]	Loss: 12.012214
Train Epoch: 5	[25600/60000 (43%)]	Loss: 7.215436
Train Epoch: 5	[26240/60000 (44%)]	Loss: 13.054042
Train Epoch: 5	[26880/60000 (45%)]	Loss: 3.121171
Train Epoch: 5	[27520/60000 (46%)]	Loss: 6.279863
Train Epoch: 5	[28160/60000 (47%)]	Loss: 3.578970
Train Epoch: 5	[28800/60000 (48%)]	Loss: 18.609404

Train Epoch: 5	[29440/60000 (49%)]	Loss: 4.110584
Train Epoch: 5	[30080/60000 (50%)]	Loss: 11.201428
Train Epoch: 5	[30720/60000 (51%)]	Loss: 0.463449
Train Epoch: 5	[31360/60000 (52%)]	Loss: 1.479442
Train Epoch: 5	[32000/60000 (53%)]	Loss: 3.813271
Train Epoch: 5	[32640/60000 (54%)]	Loss: 11.557169
Train Epoch: 5	[33280/60000 (55%)]	Loss: 10.667287
Train Epoch: 5	[33920/60000 (57%)]	Loss: 9.466030
Train Epoch: 5	[34560/60000 (58%)]	Loss: 3.419337
Train Epoch: 5	[35200/60000 (59%)]	Loss: 12.729682
Train Epoch: 5	[35840/60000 (60%)]	Loss: 5.988071
Train Epoch: 5	[36480/60000 (61%)]	Loss: 3.726662
Train Epoch: 5	[37120/60000 (62%)]	Loss: 4.299050
Train Epoch: 5	[37760/60000 (63%)]	Loss: 4.394792
Train Epoch: 5	[38400/60000 (64%)]	Loss: 14.899055
Train Epoch: 5	[39040/60000 (65%)]	Loss: 2.244035
Train Epoch: 5	[39680/60000 (66%)]	Loss: 6.475419
Train Epoch: 5	[40320/60000 (67%)]	Loss: 1.610528
Train Epoch: 5	[40960/60000 (68%)]	Loss: 17.250280
Train Epoch: 5	[41600/60000 (69%)]	Loss: 8.912785
Train Epoch: 5	[42240/60000 (70%)]	Loss: 2.690377
Train Epoch: 5	[42880/60000 (71%)]	Loss: 19.946255
Train Epoch: 5	[43520/60000 (72%)]	Loss: 0.436165
Train Epoch: 5	[44160/60000 (74%)]	Loss: 9.205739
Train Epoch: 5	[44800/60000 (75%)]	Loss: 5.441026
Train Epoch: 5	[45440/60000 (76%)]	Loss: 4.476599
Train Epoch: 5	[46080/60000 (77%)]	Loss: 6.965637
Train Epoch: 5	[46720/60000 (78%)]	Loss: 4.882964
Train Epoch: 5	[47360/60000 (79%)]	Loss: 5.793066
Train Epoch: 5	[48000/60000 (80%)]	Loss: 6.843711
Train Epoch: 5	[48640/60000 (81%)]	Loss: 0.467783
Train Epoch: 5	[49280/60000 (82%)]	Loss: 1.068542
Train Epoch: 5	[49920/60000 (83%)]	Loss: 8.641273
Train Epoch: 5	[50560/60000 (84%)]	Loss: 7.366841
Train Epoch: 5	[51200/60000 (85%)]	Loss: 6.694343
Train Epoch: 5	[51840/60000 (86%)]	Loss: 7.029500
Train Epoch: 5	[52480/60000 (87%)]	Loss: 13.054803
Train Epoch: 5	[53120/60000 (88%)]	Loss: 8.345400
Train Epoch: 5	[53760/60000 (90%)]	Loss: 14.149642
Train Epoch: 5	[54400/60000 (91%)]	Loss: 2.524837
Train Epoch: 5	[55040/60000 (92%)]	Loss: 2.386958
Train Epoch: 5	[55680/60000 (93%)]	Loss: 1.673861
Train Epoch: 5	[56320/60000 (94%)]	Loss: 2.785835
Train Epoch: 5	[56960/60000 (95%)]	Loss: 1.299375
Train Epoch: 5	[57600/60000 (96%)]	Loss: 16.489639
Train Epoch: 5	[58240/60000 (97%)]	Loss: 9.679877
Train Epoch: 5	[58880/60000 (98%)]	Loss: 8.848861
Train Epoch: 5	[59520/60000 (99%)]	Loss: 1.724942

Test set: Average loss: 0.0534, Accuracy: 9835/10000 (98.35%)

Train Epoch: 6	[0/60000 (0%)]	Loss: 6.630224
Train Epoch: 6	[640/60000 (1%)]	Loss: 20.787951
Train Epoch: 6	[1280/60000 (2%)]	Loss: 8.005960
Train Epoch: 6	[1920/60000 (3%)]	Loss: 2.239980
Train Epoch: 6	[2560/60000 (4%)]	Loss: 6.578001
Train Epoch: 6	[3200/60000 (5%)]	Loss: 3.619935
Train Epoch: 6	[3840/60000 (6%)]	Loss: 8.038226
Train Epoch: 6	[4480/60000 (7%)]	Loss: 1.407748
Train Epoch: 6	[5120/60000 (9%)]	Loss: 6.318404
Train Epoch: 6	[5760/60000 (10%)]	Loss: 3.642601
Train Epoch: 6	[6400/60000 (11%)]	Loss: 14.758779
Train Epoch: 6	[7040/60000 (12%)]	Loss: 7.775433
Train Epoch: 6	[7680/60000 (13%)]	Loss: 1.635665
Train Epoch: 6	[8320/60000 (14%)]	Loss: 5.907955
Train Epoch: 6	[8960/60000 (15%)]	Loss: 6.641166
Train Epoch: 6	[9600/60000 (16%)]	Loss: 8.952413
Train Epoch: 6	[10240/60000 (17%)]	Loss: 2.097774
Train Epoch: 6	[10880/60000 (18%)]	Loss: 1.934465
Train Epoch: 6	[11520/60000 (19%)]	Loss: 7.630402
Train Epoch: 6	[12160/60000 (20%)]	Loss: 10.302482
Train Epoch: 6	[12800/60000 (21%)]	Loss: 6.445203
Train Epoch: 6	[13440/60000 (22%)]	Loss: 1.047648
Train Epoch: 6	[14080/60000 (23%)]	Loss: 13.805943
Train Epoch: 6	[14720/60000 (25%)]	Loss: 6.265696
Train Epoch: 6	[15360/60000 (26%)]	Loss: 4.532309
Train Epoch: 6	[16000/60000 (27%)]	Loss: 9.401494
Train Epoch: 6	[16640/60000 (28%)]	Loss: 3.560211
Train Epoch: 6	[17280/60000 (29%)]	Loss: 6.512724
Train Epoch: 6	[17920/60000 (30%)]	Loss: 6.915971
Train Epoch: 6	[18560/60000 (31%)]	Loss: 6.477233
Train Epoch: 6	[19200/60000 (32%)]	Loss: 1.410539
Train Epoch: 6	[19840/60000 (33%)]	Loss: 13.338208
Train Epoch: 6	[20480/60000 (34%)]	Loss: 3.021915
Train Epoch: 6	[21120/60000 (35%)]	Loss: 13.437016
Train Epoch: 6	[21760/60000 (36%)]	Loss: 2.353795
Train Epoch: 6	[22400/60000 (37%)]	Loss: 1.897857
Train Epoch: 6	[23040/60000 (38%)]	Loss: 7.407558
Train Epoch: 6	[23680/60000 (39%)]	Loss: 9.312606
Train Epoch: 6	[24320/60000 (41%)]	Loss: 7.692962
Train Epoch: 6	[24960/60000 (42%)]	Loss: 1.286190
Train Epoch: 6	[25600/60000 (43%)]	Loss: 0.632134
Train Epoch: 6	[26240/60000 (44%)]	Loss: 1.353311
Train Epoch: 6	[26880/60000 (45%)]	Loss: 8.307919
Train Epoch: 6	[27520/60000 (46%)]	Loss: 5.581457
Train Epoch: 6	[28160/60000 (47%)]	Loss: 8.840015

Train Epoch: 6	[28800/60000 (48%)]	Loss: 6.795627
Train Epoch: 6	[29440/60000 (49%)]	Loss: 3.726153
Train Epoch: 6	[30080/60000 (50%)]	Loss: 1.732226
Train Epoch: 6	[30720/60000 (51%)]	Loss: 11.781883
Train Epoch: 6	[31360/60000 (52%)]	Loss: 4.923038
Train Epoch: 6	[32000/60000 (53%)]	Loss: 6.519748
Train Epoch: 6	[32640/60000 (54%)]	Loss: 5.364420
Train Epoch: 6	[33280/60000 (55%)]	Loss: 7.112742
Train Epoch: 6	[33920/60000 (57%)]	Loss: 2.949376
Train Epoch: 6	[34560/60000 (58%)]	Loss: 3.513264
Train Epoch: 6	[35200/60000 (59%)]	Loss: 1.145744
Train Epoch: 6	[35840/60000 (60%)]	Loss: 1.166405
Train Epoch: 6	[36480/60000 (61%)]	Loss: 1.269870
Train Epoch: 6	[37120/60000 (62%)]	Loss: 2.981927
Train Epoch: 6	[37760/60000 (63%)]	Loss: 2.170358
Train Epoch: 6	[38400/60000 (64%)]	Loss: 9.312818
Train Epoch: 6	[39040/60000 (65%)]	Loss: 15.214147
Train Epoch: 6	[39680/60000 (66%)]	Loss: 11.719201
Train Epoch: 6	[40320/60000 (67%)]	Loss: 13.651773
Train Epoch: 6	[40960/60000 (68%)]	Loss: 0.814875
Train Epoch: 6	[41600/60000 (69%)]	Loss: 11.746740
Train Epoch: 6	[42240/60000 (70%)]	Loss: 2.090991
Train Epoch: 6	[42880/60000 (71%)]	Loss: 3.587035
Train Epoch: 6	[43520/60000 (72%)]	Loss: 10.558620
Train Epoch: 6	[44160/60000 (74%)]	Loss: 2.969190
Train Epoch: 6	[44800/60000 (75%)]	Loss: 3.508309
Train Epoch: 6	[45440/60000 (76%)]	Loss: 5.605505
Train Epoch: 6	[46080/60000 (77%)]	Loss: 1.874039
Train Epoch: 6	[46720/60000 (78%)]	Loss: 0.807255
Train Epoch: 6	[47360/60000 (79%)]	Loss: 7.933768
Train Epoch: 6	[48000/60000 (80%)]	Loss: 2.012141
Train Epoch: 6	[48640/60000 (81%)]	Loss: 8.081161
Train Epoch: 6	[49280/60000 (82%)]	Loss: 6.256367
Train Epoch: 6	[49920/60000 (83%)]	Loss: 3.062191
Train Epoch: 6	[50560/60000 (84%)]	Loss: 2.545498
Train Epoch: 6	[51200/60000 (85%)]	Loss: 2.002692
Train Epoch: 6	[51840/60000 (86%)]	Loss: 9.981467
Train Epoch: 6	[52480/60000 (87%)]	Loss: 5.514098
Train Epoch: 6	[53120/60000 (88%)]	Loss: 15.328537
Train Epoch: 6	[53760/60000 (90%)]	Loss: 4.123604
Train Epoch: 6	[54400/60000 (91%)]	Loss: 6.357366
Train Epoch: 6	[55040/60000 (92%)]	Loss: 1.006755
Train Epoch: 6	[55680/60000 (93%)]	Loss: 6.220472
Train Epoch: 6	[56320/60000 (94%)]	Loss: 8.803873
Train Epoch: 6	[56960/60000 (95%)]	Loss: 3.007529
Train Epoch: 6	[57600/60000 (96%)]	Loss: 24.665640
Train Epoch: 6	[58240/60000 (97%)]	Loss: 7.719577
Train Epoch: 6	[58880/60000 (98%)]	Loss: 4.896190

Train Epoch: 6 [59520/60000 (99%)] Loss: 2.905648

Test set: Average loss: 0.0500, Accuracy: 9841/10000 (98.41%)

Train Epoch: 7 [0/60000 (0%)] Loss: 4.332672
Train Epoch: 7 [640/60000 (1%)] Loss: 3.403158
Train Epoch: 7 [1280/60000 (2%)] Loss: 6.480197
Train Epoch: 7 [1920/60000 (3%)] Loss: 5.810470
Train Epoch: 7 [2560/60000 (4%)] Loss: 1.707820
Train Epoch: 7 [3200/60000 (5%)] Loss: 17.646034
Train Epoch: 7 [3840/60000 (6%)] Loss: 1.011726
Train Epoch: 7 [4480/60000 (7%)] Loss: 3.523808
Train Epoch: 7 [5120/60000 (9%)] Loss: 1.006069
Train Epoch: 7 [5760/60000 (10%)] Loss: 9.496636
Train Epoch: 7 [6400/60000 (11%)] Loss: 1.595050
Train Epoch: 7 [7040/60000 (12%)] Loss: 1.538800
Train Epoch: 7 [7680/60000 (13%)] Loss: 2.765428
Train Epoch: 7 [8320/60000 (14%)] Loss: 2.008750
Train Epoch: 7 [8960/60000 (15%)] Loss: 4.298728
Train Epoch: 7 [9600/60000 (16%)] Loss: 1.099337
Train Epoch: 7 [10240/60000 (17%)] Loss: 18.223530
Train Epoch: 7 [10880/60000 (18%)] Loss: 13.177669
Train Epoch: 7 [11520/60000 (19%)] Loss: 11.055270
Train Epoch: 7 [12160/60000 (20%)] Loss: 6.045821
Train Epoch: 7 [12800/60000 (21%)] Loss: 0.705234
Train Epoch: 7 [13440/60000 (22%)] Loss: 3.253911
Train Epoch: 7 [14080/60000 (23%)] Loss: 1.505289
Train Epoch: 7 [14720/60000 (25%)] Loss: 6.582406
Train Epoch: 7 [15360/60000 (26%)] Loss: 2.989957
Train Epoch: 7 [16000/60000 (27%)] Loss: 1.111599
Train Epoch: 7 [16640/60000 (28%)] Loss: 9.114334
Train Epoch: 7 [17280/60000 (29%)] Loss: 2.861264
Train Epoch: 7 [17920/60000 (30%)] Loss: 4.227240
Train Epoch: 7 [18560/60000 (31%)] Loss: 3.375951
Train Epoch: 7 [19200/60000 (32%)] Loss: 0.490650
Train Epoch: 7 [19840/60000 (33%)] Loss: 5.171482
Train Epoch: 7 [20480/60000 (34%)] Loss: 0.694190
Train Epoch: 7 [21120/60000 (35%)] Loss: 11.296841
Train Epoch: 7 [21760/60000 (36%)] Loss: 3.236701
Train Epoch: 7 [22400/60000 (37%)] Loss: 3.254429
Train Epoch: 7 [23040/60000 (38%)] Loss: 9.260544
Train Epoch: 7 [23680/60000 (39%)] Loss: 4.044302
Train Epoch: 7 [24320/60000 (41%)] Loss: 17.980017
Train Epoch: 7 [24960/60000 (42%)] Loss: 9.320857
Train Epoch: 7 [25600/60000 (43%)] Loss: 8.903996
Train Epoch: 7 [26240/60000 (44%)] Loss: 1.442104
Train Epoch: 7 [26880/60000 (45%)] Loss: 6.671692
Train Epoch: 7 [27520/60000 (46%)] Loss: 5.635300

Train Epoch: 7	[28160/60000 (47%)]	Loss: 9.034647
Train Epoch: 7	[28800/60000 (48%)]	Loss: 2.659579
Train Epoch: 7	[29440/60000 (49%)]	Loss: 1.092082
Train Epoch: 7	[30080/60000 (50%)]	Loss: 4.280841
Train Epoch: 7	[30720/60000 (51%)]	Loss: 5.199413
Train Epoch: 7	[31360/60000 (52%)]	Loss: 5.857262
Train Epoch: 7	[32000/60000 (53%)]	Loss: 0.349206
Train Epoch: 7	[32640/60000 (54%)]	Loss: 3.080471
Train Epoch: 7	[33280/60000 (55%)]	Loss: 9.217436
Train Epoch: 7	[33920/60000 (57%)]	Loss: 8.408026
Train Epoch: 7	[34560/60000 (58%)]	Loss: 15.142101
Train Epoch: 7	[35200/60000 (59%)]	Loss: 3.616246
Train Epoch: 7	[35840/60000 (60%)]	Loss: 12.495605
Train Epoch: 7	[36480/60000 (61%)]	Loss: 6.047259
Train Epoch: 7	[37120/60000 (62%)]	Loss: 4.497419
Train Epoch: 7	[37760/60000 (63%)]	Loss: 7.680035
Train Epoch: 7	[38400/60000 (64%)]	Loss: 1.039654
Train Epoch: 7	[39040/60000 (65%)]	Loss: 5.710527
Train Epoch: 7	[39680/60000 (66%)]	Loss: 7.057245
Train Epoch: 7	[40320/60000 (67%)]	Loss: 4.609823
Train Epoch: 7	[40960/60000 (68%)]	Loss: 1.362619
Train Epoch: 7	[41600/60000 (69%)]	Loss: 0.454375
Train Epoch: 7	[42240/60000 (70%)]	Loss: 7.436452
Train Epoch: 7	[42880/60000 (71%)]	Loss: 6.220904
Train Epoch: 7	[43520/60000 (72%)]	Loss: 0.789020
Train Epoch: 7	[44160/60000 (74%)]	Loss: 5.097136
Train Epoch: 7	[44800/60000 (75%)]	Loss: 4.041225
Train Epoch: 7	[45440/60000 (76%)]	Loss: 2.677005
Train Epoch: 7	[46080/60000 (77%)]	Loss: 11.184466
Train Epoch: 7	[46720/60000 (78%)]	Loss: 6.737823
Train Epoch: 7	[47360/60000 (79%)]	Loss: 1.553954
Train Epoch: 7	[48000/60000 (80%)]	Loss: 8.604844
Train Epoch: 7	[48640/60000 (81%)]	Loss: 4.676332
Train Epoch: 7	[49280/60000 (82%)]	Loss: 1.990419
Train Epoch: 7	[49920/60000 (83%)]	Loss: 6.985807
Train Epoch: 7	[50560/60000 (84%)]	Loss: 2.872166
Train Epoch: 7	[51200/60000 (85%)]	Loss: 6.876919
Train Epoch: 7	[51840/60000 (86%)]	Loss: 4.401384
Train Epoch: 7	[52480/60000 (87%)]	Loss: 15.030044
Train Epoch: 7	[53120/60000 (88%)]	Loss: 0.664423
Train Epoch: 7	[53760/60000 (90%)]	Loss: 6.801167
Train Epoch: 7	[54400/60000 (91%)]	Loss: 0.430540
Train Epoch: 7	[55040/60000 (92%)]	Loss: 5.951643
Train Epoch: 7	[55680/60000 (93%)]	Loss: 15.166791
Train Epoch: 7	[56320/60000 (94%)]	Loss: 2.576188
Train Epoch: 7	[56960/60000 (95%)]	Loss: 8.735505
Train Epoch: 7	[57600/60000 (96%)]	Loss: 2.060519
Train Epoch: 7	[58240/60000 (97%)]	Loss: 4.416695

Train Epoch: 7 [58880/60000 (98%)] Loss: 0.822052
Train Epoch: 7 [59520/60000 (99%)] Loss: 1.440630

Test set: Average loss: 0.0552, Accuracy: 9831/10000 (98.31%)

Train Epoch: 8 [0/60000 (0%)] Loss: 3.563952
Train Epoch: 8 [640/60000 (1%)] Loss: 5.155003
Train Epoch: 8 [1280/60000 (2%)] Loss: 2.238763
Train Epoch: 8 [1920/60000 (3%)] Loss: 7.062019
Train Epoch: 8 [2560/60000 (4%)] Loss: 3.042124
Train Epoch: 8 [3200/60000 (5%)] Loss: 2.825757
Train Epoch: 8 [3840/60000 (6%)] Loss: 6.655331
Train Epoch: 8 [4480/60000 (7%)] Loss: 9.738765
Train Epoch: 8 [5120/60000 (9%)] Loss: 0.335297
Train Epoch: 8 [5760/60000 (10%)] Loss: 1.882492
Train Epoch: 8 [6400/60000 (11%)] Loss: 1.940750
Train Epoch: 8 [7040/60000 (12%)] Loss: 4.579886
Train Epoch: 8 [7680/60000 (13%)] Loss: 11.259303
Train Epoch: 8 [8320/60000 (14%)] Loss: 2.946391
Train Epoch: 8 [8960/60000 (15%)] Loss: 3.968591
Train Epoch: 8 [9600/60000 (16%)] Loss: 2.923450
Train Epoch: 8 [10240/60000 (17%)] Loss: 1.563694
Train Epoch: 8 [10880/60000 (18%)] Loss: 9.737351
Train Epoch: 8 [11520/60000 (19%)] Loss: 4.747862
Train Epoch: 8 [12160/60000 (20%)] Loss: 6.045306
Train Epoch: 8 [12800/60000 (21%)] Loss: 2.094947
Train Epoch: 8 [13440/60000 (22%)] Loss: 9.522662
Train Epoch: 8 [14080/60000 (23%)] Loss: 8.029079
Train Epoch: 8 [14720/60000 (25%)] Loss: 8.021338
Train Epoch: 8 [15360/60000 (26%)] Loss: 5.506800
Train Epoch: 8 [16000/60000 (27%)] Loss: 7.262497
Train Epoch: 8 [16640/60000 (28%)] Loss: 7.336144
Train Epoch: 8 [17280/60000 (29%)] Loss: 12.966516
Train Epoch: 8 [17920/60000 (30%)] Loss: 7.176060
Train Epoch: 8 [18560/60000 (31%)] Loss: 2.643374
Train Epoch: 8 [19200/60000 (32%)] Loss: 3.288464
Train Epoch: 8 [19840/60000 (33%)] Loss: 8.673247
Train Epoch: 8 [20480/60000 (34%)] Loss: 8.329132
Train Epoch: 8 [21120/60000 (35%)] Loss: 4.090757
Train Epoch: 8 [21760/60000 (36%)] Loss: 4.469299
Train Epoch: 8 [22400/60000 (37%)] Loss: 7.758936
Train Epoch: 8 [23040/60000 (38%)] Loss: 1.494558
Train Epoch: 8 [23680/60000 (39%)] Loss: 2.950405
Train Epoch: 8 [24320/60000 (41%)] Loss: 0.509877
Train Epoch: 8 [24960/60000 (42%)] Loss: 2.728011
Train Epoch: 8 [25600/60000 (43%)] Loss: 6.796323
Train Epoch: 8 [26240/60000 (44%)] Loss: 1.376414
Train Epoch: 8 [26880/60000 (45%)] Loss: 0.942263

Train Epoch: 8	[27520/60000 (46%)]	Loss: 9.138129
Train Epoch: 8	[28160/60000 (47%)]	Loss: 3.684197
Train Epoch: 8	[28800/60000 (48%)]	Loss: 1.185676
Train Epoch: 8	[29440/60000 (49%)]	Loss: 12.500170
Train Epoch: 8	[30080/60000 (50%)]	Loss: 0.918692
Train Epoch: 8	[30720/60000 (51%)]	Loss: 8.275303
Train Epoch: 8	[31360/60000 (52%)]	Loss: 4.233202
Train Epoch: 8	[32000/60000 (53%)]	Loss: 7.400211
Train Epoch: 8	[32640/60000 (54%)]	Loss: 1.218750
Train Epoch: 8	[33280/60000 (55%)]	Loss: 5.398952
Train Epoch: 8	[33920/60000 (57%)]	Loss: 2.643663
Train Epoch: 8	[34560/60000 (58%)]	Loss: 8.843253
Train Epoch: 8	[35200/60000 (59%)]	Loss: 0.853465
Train Epoch: 8	[35840/60000 (60%)]	Loss: 5.443512
Train Epoch: 8	[36480/60000 (61%)]	Loss: 8.286023
Train Epoch: 8	[37120/60000 (62%)]	Loss: 1.682471
Train Epoch: 8	[37760/60000 (63%)]	Loss: 4.802897
Train Epoch: 8	[38400/60000 (64%)]	Loss: 1.301162
Train Epoch: 8	[39040/60000 (65%)]	Loss: 2.255265
Train Epoch: 8	[39680/60000 (66%)]	Loss: 3.549284
Train Epoch: 8	[40320/60000 (67%)]	Loss: 6.622640
Train Epoch: 8	[40960/60000 (68%)]	Loss: 0.747360
Train Epoch: 8	[41600/60000 (69%)]	Loss: 5.643166
Train Epoch: 8	[42240/60000 (70%)]	Loss: 2.000731
Train Epoch: 8	[42880/60000 (71%)]	Loss: 9.112757
Train Epoch: 8	[43520/60000 (72%)]	Loss: 2.740674
Train Epoch: 8	[44160/60000 (74%)]	Loss: 3.915157
Train Epoch: 8	[44800/60000 (75%)]	Loss: 9.139977
Train Epoch: 8	[45440/60000 (76%)]	Loss: 3.218711
Train Epoch: 8	[46080/60000 (77%)]	Loss: 4.212547
Train Epoch: 8	[46720/60000 (78%)]	Loss: 1.705429
Train Epoch: 8	[47360/60000 (79%)]	Loss: 13.333056
Train Epoch: 8	[48000/60000 (80%)]	Loss: 17.042568
Train Epoch: 8	[48640/60000 (81%)]	Loss: 3.645360
Train Epoch: 8	[49280/60000 (82%)]	Loss: 11.443022
Train Epoch: 8	[49920/60000 (83%)]	Loss: 3.888028
Train Epoch: 8	[50560/60000 (84%)]	Loss: 3.182363
Train Epoch: 8	[51200/60000 (85%)]	Loss: 12.624931
Train Epoch: 8	[51840/60000 (86%)]	Loss: 3.650775
Train Epoch: 8	[52480/60000 (87%)]	Loss: 6.783047
Train Epoch: 8	[53120/60000 (88%)]	Loss: 4.391937
Train Epoch: 8	[53760/60000 (90%)]	Loss: 5.423202
Train Epoch: 8	[54400/60000 (91%)]	Loss: 3.166685
Train Epoch: 8	[55040/60000 (92%)]	Loss: 1.964956
Train Epoch: 8	[55680/60000 (93%)]	Loss: 8.715889
Train Epoch: 8	[56320/60000 (94%)]	Loss: 4.536964
Train Epoch: 8	[56960/60000 (95%)]	Loss: 5.453135
Train Epoch: 8	[57600/60000 (96%)]	Loss: 2.534173

Train Epoch: 8 [58240/60000 (97%)]	Loss: 0.342883
Train Epoch: 8 [58880/60000 (98%)]	Loss: 0.812616
Train Epoch: 8 [59520/60000 (99%)]	Loss: 7.596770

Test set: Average loss: 0.0528, Accuracy: 9842/10000 (98.42%)

Train Epoch: 9 [0/60000 (0%)]	Loss: 6.299816
Train Epoch: 9 [640/60000 (1%)]	Loss: 6.907352
Train Epoch: 9 [1280/60000 (2%)]	Loss: 3.600256
Train Epoch: 9 [1920/60000 (3%)]	Loss: 6.660454
Train Epoch: 9 [2560/60000 (4%)]	Loss: 11.351507
Train Epoch: 9 [3200/60000 (5%)]	Loss: 2.773911
Train Epoch: 9 [3840/60000 (6%)]	Loss: 4.367436
Train Epoch: 9 [4480/60000 (7%)]	Loss: 2.032342
Train Epoch: 9 [5120/60000 (9%)]	Loss: 3.777251
Train Epoch: 9 [5760/60000 (10%)]	Loss: 17.067390
Train Epoch: 9 [6400/60000 (11%)]	Loss: 7.372547
Train Epoch: 9 [7040/60000 (12%)]	Loss: 16.462072
Train Epoch: 9 [7680/60000 (13%)]	Loss: 3.625682
Train Epoch: 9 [8320/60000 (14%)]	Loss: 5.310660
Train Epoch: 9 [8960/60000 (15%)]	Loss: 0.962364
Train Epoch: 9 [9600/60000 (16%)]	Loss: 0.368191
Train Epoch: 9 [10240/60000 (17%)]	Loss: 10.755780
Train Epoch: 9 [10880/60000 (18%)]	Loss: 1.793166
Train Epoch: 9 [11520/60000 (19%)]	Loss: 2.456767
Train Epoch: 9 [12160/60000 (20%)]	Loss: 5.134060
Train Epoch: 9 [12800/60000 (21%)]	Loss: 1.750126
Train Epoch: 9 [13440/60000 (22%)]	Loss: 5.367651
Train Epoch: 9 [14080/60000 (23%)]	Loss: 1.163476
Train Epoch: 9 [14720/60000 (25%)]	Loss: 4.598181
Train Epoch: 9 [15360/60000 (26%)]	Loss: 8.274391
Train Epoch: 9 [16000/60000 (27%)]	Loss: 8.173906
Train Epoch: 9 [16640/60000 (28%)]	Loss: 5.058841
Train Epoch: 9 [17280/60000 (29%)]	Loss: 4.430741
Train Epoch: 9 [17920/60000 (30%)]	Loss: 5.342674
Train Epoch: 9 [18560/60000 (31%)]	Loss: 4.361444
Train Epoch: 9 [19200/60000 (32%)]	Loss: 4.571357
Train Epoch: 9 [19840/60000 (33%)]	Loss: 0.736085
Train Epoch: 9 [20480/60000 (34%)]	Loss: 4.031384
Train Epoch: 9 [21120/60000 (35%)]	Loss: 3.569153
Train Epoch: 9 [21760/60000 (36%)]	Loss: 5.078662
Train Epoch: 9 [22400/60000 (37%)]	Loss: 7.778805
Train Epoch: 9 [23040/60000 (38%)]	Loss: 2.583452
Train Epoch: 9 [23680/60000 (39%)]	Loss: 6.220582
Train Epoch: 9 [24320/60000 (41%)]	Loss: 4.795364
Train Epoch: 9 [24960/60000 (42%)]	Loss: 12.972521
Train Epoch: 9 [25600/60000 (43%)]	Loss: 3.887052
Train Epoch: 9 [26240/60000 (44%)]	Loss: 2.735393

Train Epoch: 9	[26880/60000 (45%)]	Loss: 9.341623
Train Epoch: 9	[27520/60000 (46%)]	Loss: 3.290767
Train Epoch: 9	[28160/60000 (47%)]	Loss: 5.803006
Train Epoch: 9	[28800/60000 (48%)]	Loss: 2.812087
Train Epoch: 9	[29440/60000 (49%)]	Loss: 10.696334
Train Epoch: 9	[30080/60000 (50%)]	Loss: 2.395273
Train Epoch: 9	[30720/60000 (51%)]	Loss: 0.924807
Train Epoch: 9	[31360/60000 (52%)]	Loss: 16.396051
Train Epoch: 9	[32000/60000 (53%)]	Loss: 16.035578
Train Epoch: 9	[32640/60000 (54%)]	Loss: 2.964968
Train Epoch: 9	[33280/60000 (55%)]	Loss: 0.567809
Train Epoch: 9	[33920/60000 (57%)]	Loss: 12.047561
Train Epoch: 9	[34560/60000 (58%)]	Loss: 0.363724
Train Epoch: 9	[35200/60000 (59%)]	Loss: 0.670693
Train Epoch: 9	[35840/60000 (60%)]	Loss: 9.722827
Train Epoch: 9	[36480/60000 (61%)]	Loss: 1.352764
Train Epoch: 9	[37120/60000 (62%)]	Loss: 5.063897
Train Epoch: 9	[37760/60000 (63%)]	Loss: 18.174213
Train Epoch: 9	[38400/60000 (64%)]	Loss: 7.493132
Train Epoch: 9	[39040/60000 (65%)]	Loss: 0.802953
Train Epoch: 9	[39680/60000 (66%)]	Loss: 0.523510
Train Epoch: 9	[40320/60000 (67%)]	Loss: 3.561802
Train Epoch: 9	[40960/60000 (68%)]	Loss: 10.903907
Train Epoch: 9	[41600/60000 (69%)]	Loss: 6.284098
Train Epoch: 9	[42240/60000 (70%)]	Loss: 7.998821
Train Epoch: 9	[42880/60000 (71%)]	Loss: 14.846068
Train Epoch: 9	[43520/60000 (72%)]	Loss: 5.223794
Train Epoch: 9	[44160/60000 (74%)]	Loss: 8.336000
Train Epoch: 9	[44800/60000 (75%)]	Loss: 3.820257
Train Epoch: 9	[45440/60000 (76%)]	Loss: 12.239676
Train Epoch: 9	[46080/60000 (77%)]	Loss: 8.983435
Train Epoch: 9	[46720/60000 (78%)]	Loss: 1.897945
Train Epoch: 9	[47360/60000 (79%)]	Loss: 1.106027
Train Epoch: 9	[48000/60000 (80%)]	Loss: 11.650100
Train Epoch: 9	[48640/60000 (81%)]	Loss: 2.005645
Train Epoch: 9	[49280/60000 (82%)]	Loss: 2.290253
Train Epoch: 9	[49920/60000 (83%)]	Loss: 3.807944
Train Epoch: 9	[50560/60000 (84%)]	Loss: 9.133227
Train Epoch: 9	[51200/60000 (85%)]	Loss: 2.091491
Train Epoch: 9	[51840/60000 (86%)]	Loss: 1.488215
Train Epoch: 9	[52480/60000 (87%)]	Loss: 10.162395
Train Epoch: 9	[53120/60000 (88%)]	Loss: 1.763453
Train Epoch: 9	[53760/60000 (90%)]	Loss: 7.418035
Train Epoch: 9	[54400/60000 (91%)]	Loss: 1.418293
Train Epoch: 9	[55040/60000 (92%)]	Loss: 3.518251
Train Epoch: 9	[55680/60000 (93%)]	Loss: 5.399552
Train Epoch: 9	[56320/60000 (94%)]	Loss: 1.408760
Train Epoch: 9	[56960/60000 (95%)]	Loss: 1.845590

Train Epoch: 9	[57600/60000 (96%)]	Loss: 4.457830
Train Epoch: 9	[58240/60000 (97%)]	Loss: 5.572523
Train Epoch: 9	[58880/60000 (98%)]	Loss: 3.224164
Train Epoch: 9	[59520/60000 (99%)]	Loss: 1.067660

Test set: Average loss: 0.0476, Accuracy: 9858/10000 (98.58%)

Train Epoch: 10	[0/60000 (0%)]	Loss: 3.189883
Train Epoch: 10	[640/60000 (1%)]	Loss: 0.836223
Train Epoch: 10	[1280/60000 (2%)]	Loss: 9.152547
Train Epoch: 10	[1920/60000 (3%)]	Loss: 3.182679
Train Epoch: 10	[2560/60000 (4%)]	Loss: 1.494692
Train Epoch: 10	[3200/60000 (5%)]	Loss: 2.478202
Train Epoch: 10	[3840/60000 (6%)]	Loss: 6.070901
Train Epoch: 10	[4480/60000 (7%)]	Loss: 3.426918
Train Epoch: 10	[5120/60000 (9%)]	Loss: 5.933118
Train Epoch: 10	[5760/60000 (10%)]	Loss: 10.106821
Train Epoch: 10	[6400/60000 (11%)]	Loss: 10.747767
Train Epoch: 10	[7040/60000 (12%)]	Loss: 4.511617
Train Epoch: 10	[7680/60000 (13%)]	Loss: 11.486852
Train Epoch: 10	[8320/60000 (14%)]	Loss: 1.078712
Train Epoch: 10	[8960/60000 (15%)]	Loss: 11.409203
Train Epoch: 10	[9600/60000 (16%)]	Loss: 3.687876
Train Epoch: 10	[10240/60000 (17%)]	Loss: 0.529134
Train Epoch: 10	[10880/60000 (18%)]	Loss: 4.930624
Train Epoch: 10	[11520/60000 (19%)]	Loss: 1.219167
Train Epoch: 10	[12160/60000 (20%)]	Loss: 9.000279
Train Epoch: 10	[12800/60000 (21%)]	Loss: 4.446573
Train Epoch: 10	[13440/60000 (22%)]	Loss: 1.125186
Train Epoch: 10	[14080/60000 (23%)]	Loss: 4.197731
Train Epoch: 10	[14720/60000 (25%)]	Loss: 4.674078
Train Epoch: 10	[15360/60000 (26%)]	Loss: 14.364170
Train Epoch: 10	[16000/60000 (27%)]	Loss: 4.401692
Train Epoch: 10	[16640/60000 (28%)]	Loss: 2.568971
Train Epoch: 10	[17280/60000 (29%)]	Loss: 1.709392
Train Epoch: 10	[17920/60000 (30%)]	Loss: 1.695273
Train Epoch: 10	[18560/60000 (31%)]	Loss: 4.276515
Train Epoch: 10	[19200/60000 (32%)]	Loss: 8.406143
Train Epoch: 10	[19840/60000 (33%)]	Loss: 1.393481
Train Epoch: 10	[20480/60000 (34%)]	Loss: 2.328926
Train Epoch: 10	[21120/60000 (35%)]	Loss: 1.179146
Train Epoch: 10	[21760/60000 (36%)]	Loss: 18.213463
Train Epoch: 10	[22400/60000 (37%)]	Loss: 28.134809
Train Epoch: 10	[23040/60000 (38%)]	Loss: 5.610700
Train Epoch: 10	[23680/60000 (39%)]	Loss: 9.046723
Train Epoch: 10	[24320/60000 (41%)]	Loss: 5.086613
Train Epoch: 10	[24960/60000 (42%)]	Loss: 10.278698
Train Epoch: 10	[25600/60000 (43%)]	Loss: 6.327031

Train Epoch: 10	[26240/60000 (44%)]	Loss: 4.921795
Train Epoch: 10	[26880/60000 (45%)]	Loss: 2.608008
Train Epoch: 10	[27520/60000 (46%)]	Loss: 9.652815
Train Epoch: 10	[28160/60000 (47%)]	Loss: 4.122248
Train Epoch: 10	[28800/60000 (48%)]	Loss: 10.243618
Train Epoch: 10	[29440/60000 (49%)]	Loss: 0.773587
Train Epoch: 10	[30080/60000 (50%)]	Loss: 1.965021
Train Epoch: 10	[30720/60000 (51%)]	Loss: 4.634542
Train Epoch: 10	[31360/60000 (52%)]	Loss: 12.785232
Train Epoch: 10	[32000/60000 (53%)]	Loss: 7.444949
Train Epoch: 10	[32640/60000 (54%)]	Loss: 2.399635
Train Epoch: 10	[33280/60000 (55%)]	Loss: 1.768976
Train Epoch: 10	[33920/60000 (57%)]	Loss: 11.278794
Train Epoch: 10	[34560/60000 (58%)]	Loss: 5.509012
Train Epoch: 10	[35200/60000 (59%)]	Loss: 1.690901
Train Epoch: 10	[35840/60000 (60%)]	Loss: 4.509453
Train Epoch: 10	[36480/60000 (61%)]	Loss: 3.220570
Train Epoch: 10	[37120/60000 (62%)]	Loss: 4.345402
Train Epoch: 10	[37760/60000 (63%)]	Loss: 3.304995
Train Epoch: 10	[38400/60000 (64%)]	Loss: 0.640988
Train Epoch: 10	[39040/60000 (65%)]	Loss: 0.882594
Train Epoch: 10	[39680/60000 (66%)]	Loss: 3.994906
Train Epoch: 10	[40320/60000 (67%)]	Loss: 1.970464
Train Epoch: 10	[40960/60000 (68%)]	Loss: 12.330709
Train Epoch: 10	[41600/60000 (69%)]	Loss: 8.882204
Train Epoch: 10	[42240/60000 (70%)]	Loss: 2.998013
Train Epoch: 10	[42880/60000 (71%)]	Loss: 4.760514
Train Epoch: 10	[43520/60000 (72%)]	Loss: 1.368909
Train Epoch: 10	[44160/60000 (74%)]	Loss: 5.985575
Train Epoch: 10	[44800/60000 (75%)]	Loss: 2.713306
Train Epoch: 10	[45440/60000 (76%)]	Loss: 5.131996
Train Epoch: 10	[46080/60000 (77%)]	Loss: 8.018957
Train Epoch: 10	[46720/60000 (78%)]	Loss: 17.563559
Train Epoch: 10	[47360/60000 (79%)]	Loss: 4.790299
Train Epoch: 10	[48000/60000 (80%)]	Loss: 1.066709
Train Epoch: 10	[48640/60000 (81%)]	Loss: 0.915334
Train Epoch: 10	[49280/60000 (82%)]	Loss: 5.555498
Train Epoch: 10	[49920/60000 (83%)]	Loss: 19.835142
Train Epoch: 10	[50560/60000 (84%)]	Loss: 5.899791
Train Epoch: 10	[51200/60000 (85%)]	Loss: 3.401639
Train Epoch: 10	[51840/60000 (86%)]	Loss: 3.151226
Train Epoch: 10	[52480/60000 (87%)]	Loss: 2.130831
Train Epoch: 10	[53120/60000 (88%)]	Loss: 7.961141
Train Epoch: 10	[53760/60000 (90%)]	Loss: 10.690460
Train Epoch: 10	[54400/60000 (91%)]	Loss: 5.678731
Train Epoch: 10	[55040/60000 (92%)]	Loss: 14.535790
Train Epoch: 10	[55680/60000 (93%)]	Loss: 0.634219
Train Epoch: 10	[56320/60000 (94%)]	Loss: 0.568256

```
Train Epoch: 10 [56960/60000 (95%)]    Loss: 3.415380
Train Epoch: 10 [57600/60000 (96%)]    Loss: 5.356873
Train Epoch: 10 [58240/60000 (97%)]    Loss: 5.134482
Train Epoch: 10 [58880/60000 (98%)]    Loss: 4.017395
Train Epoch: 10 [59520/60000 (99%)]    Loss: 1.607268
```

```
Test set: Average loss: 0.0473, Accuracy: 9865/10000 (98.65%)
```

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

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

```
[ ]:
```

```
[ ]:
```