Name: Sreeja Guduri
Assignment - 2
Roll number: 2021102007

QUESTION - 1

- The dataset is downloadde from the torchvision library. We then seperate the actual images and labels associsteed with them.
- The SIFT (Scale Invariant Feature Transform) detector and descriptor are implemented using the inbuilt SIFT functions in OpenCV.
- The Bag Of Visual Words (BoVW) approach involves computing SIFT features for the images and then computing a cluster of these features (using kmeans).
- We then represent each image as a histogram of these visual words, which the SVM (Support Vector Machine) uses to train and learn and classify the images.

ACCURACY: 0.401 (for 10 clusters)

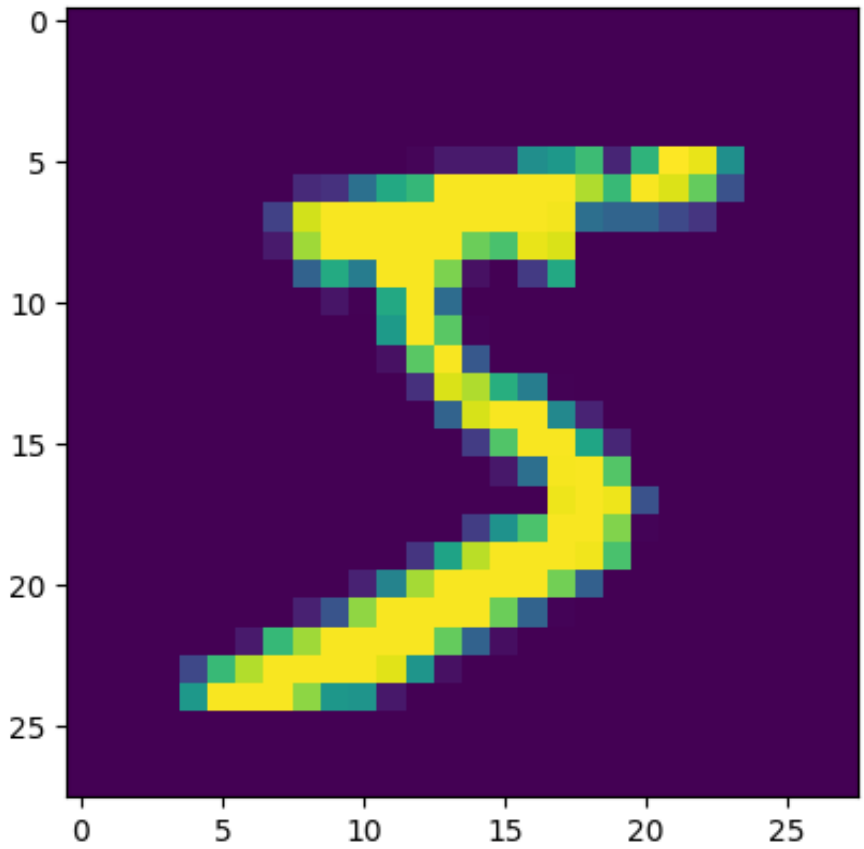
```
import torchvision
import cv2 as cv
import numpy as np
import matplotlib.pyplot as plt
from sklearn.cluster import KMeans
from sklearn.svm import LinearSVC
from sklearn.preprocessing import StandardScaler
from sklearn.pipeline import make_pipeline
from sklearn.model_selection import train_test_split
from sklearn.metrics import accuracy_score
```

```
train_data = torchvision.datasets.MNIST(root='./data', train=True, download=Tru
test_data = torchvision.datasets.MNIST(root='./data', train=False, download=Tru

print("Number of training images:", len(train_data))
print("Number of test images:", len(test_data))
```

```
    Number of training images: 60000
    Number of test images: 10000
```

```
# Separate images and labels from the train dataset
train_images = [image for image, label in train_data]
train_labels = [label for image, label in train_data]

# Separate images and labels from the test dataset
test_images = [image for image, label in test_data]
test_labels = [label for image, label in test_data]


plt.imshow(train_images[0])
print(train_images[0].size)
```

```
(28, 28)
```



```
def sift_detector(image):
    sift = cv.SIFT_create()
    keypoints, descriptors = sift.detectAndCompute(image, None)
    return keypoints, descriptors

# testing the sift function
# image = np.array(train_data[0][0])
# plt.imshow(image)
# keypoints, descriptors = sift_detector(image)
# image_with_keypoints = cv.drawKeypoints(image, keypoints, None)
# plt.imshow(image_with_keypoints)

def BoVW(images, num):
    image_descriptors = []
    for image in images:
        _, descriptors = sift_detector(image)
        if descriptors is not None:
            image_descriptors.extend(descriptors)

    image_descriptors = np.vstack(image_descriptors)
    image_descriptors = image_descriptors.astype(np.float32)

    kmeans = KMeans(n_clusters=num, random_state=42)
    kmeans.fit(image_descriptors)
    return kmeans, kmeans.cluster_centers_ # gives the visual features in the c

def image_histogram(image, visual_words, kmeans_model):
    _, descriptors = sift_detector(image)
    if descriptors is None:
        return np.zeros(len(visual_words))

    histogram = np.zeros(len(visual_words))
    labels = kmeans_model.predict(descriptors)
    for label in labels:
        histogram[label] += 1
    return histogram
```

```
cluster = 10
train_images = np.array(train_images)
test_images = np.array(test_images)
kmeans, visual_words = BoVW(train_images, cluster)

train_hist = np.array([image_histogram(image, visual_words, kmeans) for image i
test_hist = np.array([image_histogram(image, visual_words, kmeans) for image in

# Train
svm = LinearSVC(random_state=42)
svm.fit(train_hist, train_labels)

pred_label = svm.predict(test_hist)
accuracy = accuracy_score(test_labels, pred_label)
print("Accuracy:", accuracy)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
Accuracy: 0.401
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
```

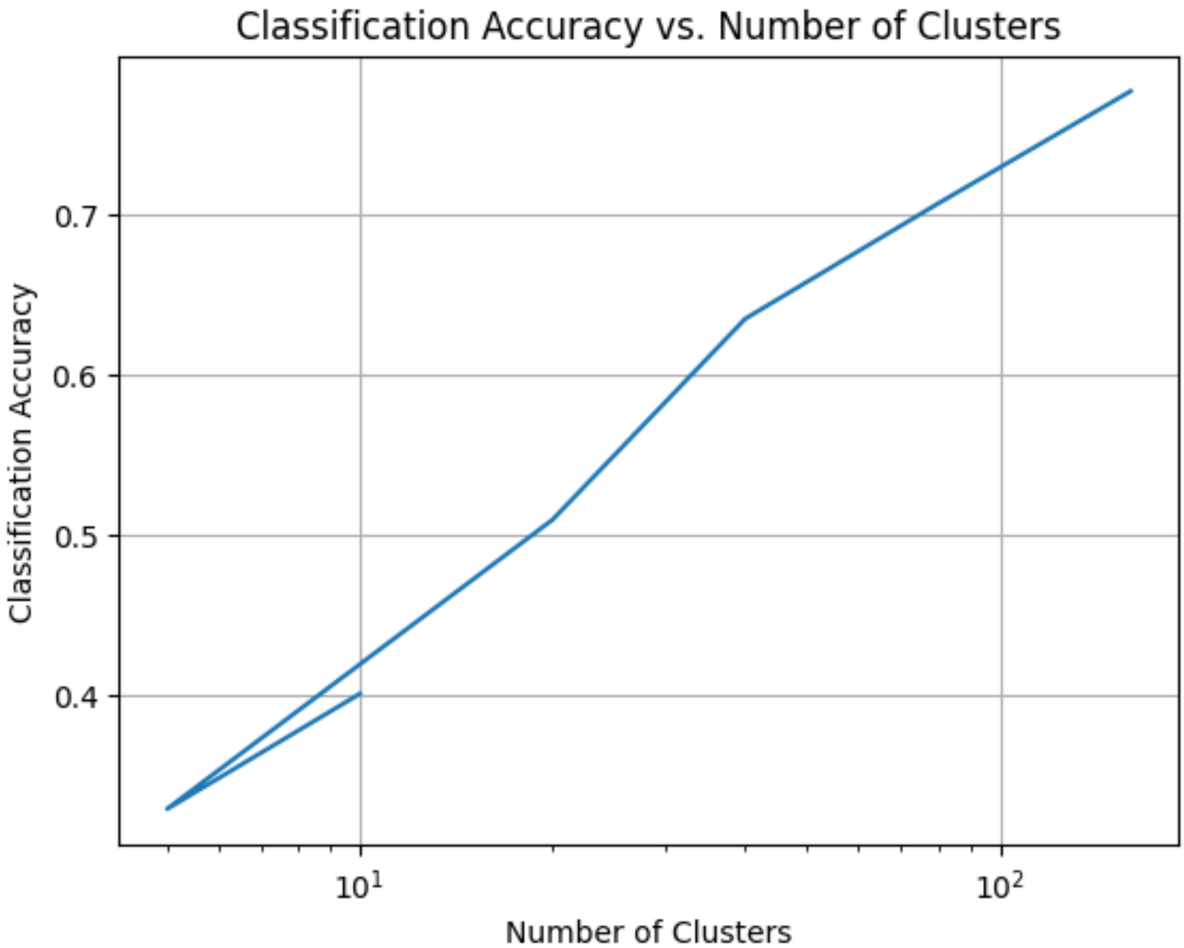
Different cluster values:

- The plot shows us that the classification accuracy increases as the number of clusters increases.
- More clusters means more visual words that are considered, which enables us to capture the finer details of the image.
- The tradeoff to consider here is that if the number of clusters is too high, it will lead to overfitting (when the model is trained too closely to the training data, leading to poor performance on the test data)

```
clusters_values = [10, 5, 20, 40, 80, 160]
accuracies = []
accuracies.append(accuracy)

for clusters in clusters_values[1:]:
    kmeans, visual_words = BoVW(train_images, clusters)

    train_hist = np.array([image_histogram(image, visual_words, kmeans) for ima
    test_hist = np.array([image_histogram(image, visual_words, kmeans) for imag

    svm_model = LinearSVC(random_state=42)
```

```
        svm_model.fit(train_hist, train_labels)

        pred_label = svm_model.predict(test_hist)
        accuracy = accuracy_score(test_labels, pred_label)
        accuracies.append(accuracy)


    plt.figure()
    plt.plot(clusters_values, accuracies)
    plt.title('Classification Accuracy vs. Number of Clusters')
    plt.xlabel('Number of Clusters')
    plt.ylabel('Classification Accuracy')
    plt.xscale('log')  # Set logarithmic scale
    plt.grid(True)
    plt.show()
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
```



Different hyperparameters:

We test with six different hyperparameters:

1. C (Regularization paramter) value of our linear SVM - 10 and 0.1
2. Contrast threshold (determines which keypoints are considered) of our SIFT descriptor - 0.01 and 0.08
3. The sigma value (determines the scale of the scale space) of our SIFT descriptor - 0.5 and 3.5

We notice from the graph that changing the sigma value of our SIFT model had the most effect on accuracy. When the sigma value is aorund 0.5, we get the highest accuracy while a sigma value of 3.5 gave us the worst accuracy. Most other hyperparameters gave us about the same accuracy, with a smaller C value giving us a better result.

```
hyperparameter_settings = [
    {'svm_c': 10},
    {'svm_c': 0.1},
    {'contrast_threshold': 0.01},
    {'contrast_threshold': 0.08},
    {'sift_sigma': 0.5},
    {'sift_sigma': 3.5},
]

accuracies = []
cluster = 10
kmeans, visual_words = BoVW(train_images, cluster)

train_hist = np.array([image_histogram(image, visual_words, kmeans) for image i
test_hist = np.array([image_histogram(image, visual_words, kmeans) for image in

for i in range(2):
    svm_model = LinearSVC(random_state=42)
    svm_model.C = hyperparameter_settings[i]['svm_c']
    svm_model.fit(train_hist, train_labels)

    # Evaluate the model
    y_pred = svm_model.predict(test_hist)
    accuracy = accuracy_score(test_labels, y_pred)
    accuracies.append(accuracy)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
```

```
def sift_detector(image, contrast):
    sift = cv.SIFT_create(contrastThreshold=contrast)
    keypoints, descriptors = sift.detectAndCompute(image, None)
    return keypoints, descriptors


def BoVW(images, num, contrast):
    image_descriptors = []
    for image in images:
        _, descriptors = sift_detector(image, contrast)
        if descriptors is not None:
            image_descriptors.extend(descriptors)

    image_descriptors = np.vstack(image_descriptors)
    image_descriptors = image_descriptors.astype(np.float32)
```

```
    kmeans = KMeans(n_clusters=num, random_state=42)
    kmeans.fit(image_descriptors)
    return kmeans, kmeans.cluster_centers_ # gives the visual features in the c

def image_histogram(image, visual_words, kmeans_model, contrast):
    _, descriptors = sift_detector(image, contrast)
    if descriptors is None:
        return np.zeros(len(visual_words))

    histogram = np.zeros(len(visual_words))
    labels = kmeans_model.predict(descriptors)
    for label in labels:
        histogram[label] += 1
    return histogram

for i in range(2,4):
    kmeans, visual_words = BoVW(train_images, cluster, hyperparameter_settings|

    train_hist = np.array([image_histogram(image, visual_words, kmeans, hyperpa
    test_hist = np.array([image_histogram(image, visual_words, kmeans, hyperpar

    svm_model = LinearSVC(random_state=42)
    svm_model.fit(train_hist, train_labels)

    # Evaluate the model
    y_pred = svm_model.predict(test_hist)
    accuracy = accuracy_score(test_labels, y_pred)
    accuracies.append(accuracy)
```

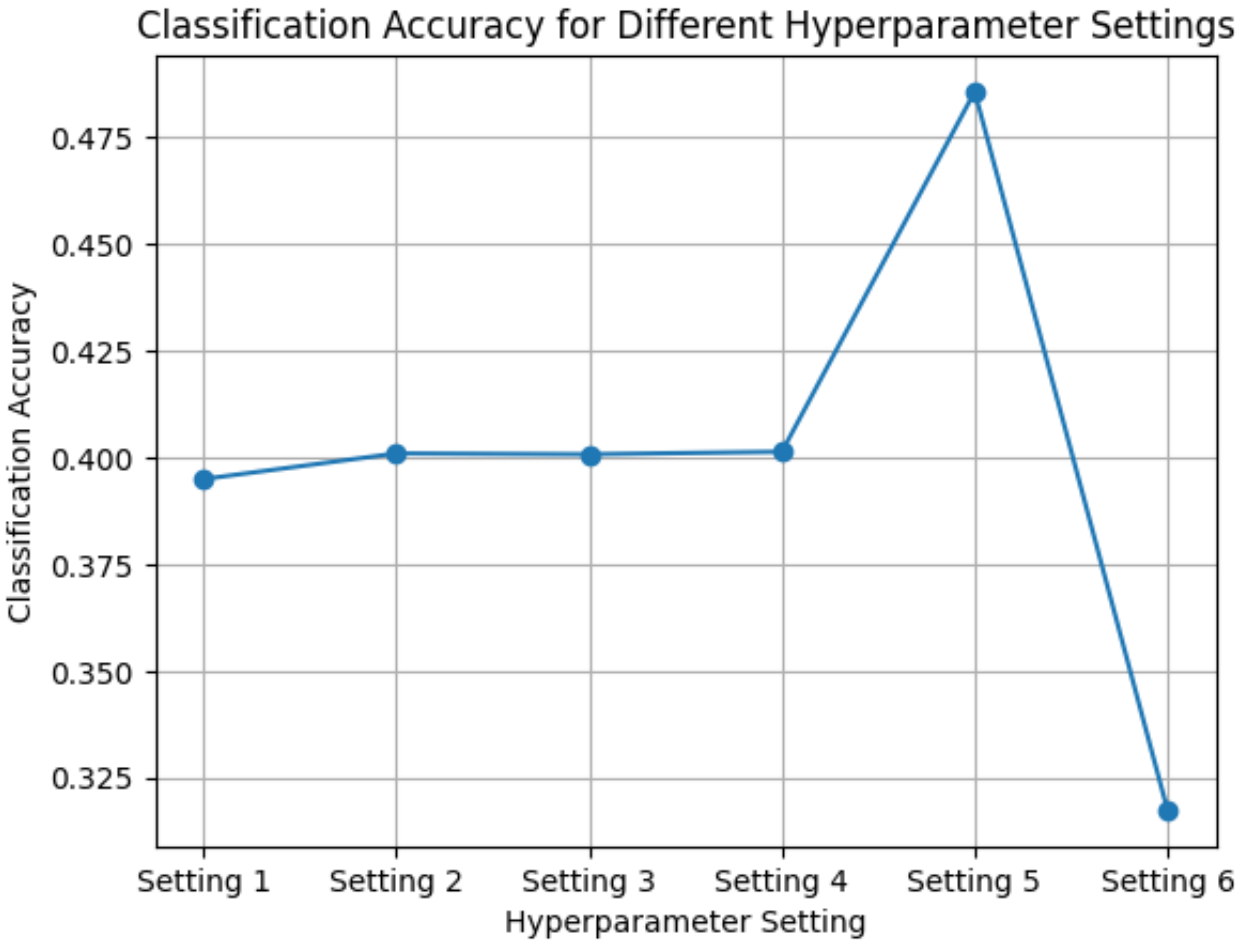
```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
```

```
def sift_detector(image, sigma):
    sift = cv.SIFT_create(sigma=sigma)
    keypoints, descriptors = sift.detectAndCompute(image, None)
    return keypoints, descriptors


def BoVW(images, num, sigma):
    image_descriptors = []
    for image in images:
        _, descriptors = sift_detector(image, sigma)
```

```
            if descriptors is not None:
                image_descriptors.extend(descriptors)

        image_descriptors = np.vstack(image_descriptors)
        image_descriptors = image_descriptors.astype(np.float32)

        kmeans = KMeans(n_clusters=num, random_state=42)
        kmeans.fit(image_descriptors)
        return kmeans, kmeans.cluster_centers_ # gives the visual features in the c

    def image_histogram(image, visual_words, kmeans_model, sigma):
        _, descriptors = sift_detector(image, sigma)
        if descriptors is None:
            return np.zeros(len(visual_words))

        histogram = np.zeros(len(visual_words))
        labels = kmeans_model.predict(descriptors)
        for label in labels:
            histogram[label] += 1
        return histogram

    for i in range(4,6):
        kmeans, visual_words = BoVW(train_images, cluster, hyperparameter_settings|

        train_hist = np.array([image_histogram(image, visual_words, kmeans, hyperpa
        test_hist = np.array([image_histogram(image, visual_words, kmeans, hyperpar

        svm_model = LinearSVC(random_state=42)
        svm_model.fit(train_hist, train_labels)

        # Evaluate the model
        y_pred = svm_model.predict(test_hist)
        accuracy = accuracy_score(test_labels, y_pred)
        accuracies.append(accuracy)
```

```
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/cluster/_kmeans.py:870: Fut
  warnings.warn(
/usr/local/lib/python3.10/dist-packages/sklearn/svm/_base.py:1244: Converge
  warnings.warn(
```

```
# Plot the results
plt.figure()
plt.plot(range(1, len(hyperparameter_settings) + 1), accuracies, marker='o')
plt.title('Classification Accuracy for Different Hyperparameter Settings')
plt.xlabel('Hyperparameter Setting')
plt.ylabel('Classification Accuracy')
plt.xticks(range(1, len(hyperparameter_settings) + 1), [f"Setting {i+1}" for i
plt.grid(True)
plt.show()
```



QUESTION - 2

- Our aim is to now perform the same task, but using a convolutional neural network (CNN) instead.
- We first define the model using the Torchvision Neural Network library.
- LeNet is a CNN model that consists of three main parts - convolution, pooling and non-linear activation function (ReLU)
- The convolution layers extract the features of the image, while pooling is used to subsample. The non-linear function allows the network to learn properly.

- The input data is given to the CNN in batches of 64 images, using the data loader.
- The loss function used is Cross-Entropy loss and the optimiser is the simple stochastic gradient descent.
- We keep a track of the running loss and also the total number of predictions the model got correct to determine accuracy.

ACCURACY: 98.2%

```
!pip install wandb
!pip install torchmetrics
```

```
Requirement already satisfied: wandb in /usr/local/lib/python3.10/dist-pack
Requirement already satisfied: Click!=8.0.0,>=7.1 in /usr/local/lib/python3
Requirement already satisfied: GitPython!=3.1.29,>=1.0.0 in /usr/local/lib/
Requirement already satisfied: requests<3,>=2.0.0 in /usr/local/lib/python3
Requirement already satisfied: psutil>=5.0.0 in /usr/local/lib/python3.10/d
Requirement already satisfied: sentry-sdk>=1.0.0 in /usr/local/lib/python3.
Requirement already satisfied: docker-pycreds>=0.4.0 in /usr/local/lib/pyth
Requirement already satisfied: PyYAML in /usr/local/lib/python3.10/dist-pac
Requirement already satisfied: setproctitle in /usr/local/lib/python3.10/di
Requirement already satisfied: setuptools in /usr/local/lib/python3.10/dist
Requirement already satisfied: appdirs>=1.4.3 in /usr/local/lib/python3.10/
Requirement already satisfied: protobuf!=4.21.0,<5,>=3.19.0 in /usr/local/l
Requirement already satisfied: six>=1.4.0 in /usr/local/lib/python3.10/dist
Requirement already satisfied: gitdb<5,>=4.0.1 in /usr/local/lib/python3.10
Requirement already satisfied: charset-normalizer<4,>=2 in /usr/local/lib/p
Requirement already satisfied: idna<4,>=2.5 in /usr/local/lib/python3.10/di
Requirement already satisfied: urllib3<3,>=1.21.1 in /usr/local/lib/python3
Requirement already satisfied: certifi>=2017.4.17 in /usr/local/lib/python3
Requirement already satisfied: smmap<6,>=3.0.1 in /usr/local/lib/python3.10
```

```
import torch
import torch.nn as nn
import torch.optim as optim
import torchvision
import torchvision.transforms as transforms
import wandb
import numpy as np

# random seed
torch.manual_seed(42)
np.random.seed(42)
wandb.login()
```

```
wandb: Logging into wandb.ai. (Learn how to deploy a W&B server locally: ht
wandb: You can find your API key in your browser here: https://wandb.ai/aut
wandb: Paste an API key from your profile and hit enter, or press ctrl+c to
wandb: Appending key for api.wandb.ai to your netrc file: /root/.netrc
True
```

```
class LeNet(nn.Module):
    def __init__(self):
        super(LeNet, self).__init__()
        self.conv1 = nn.Conv2d(1, 6, 5)
        self.conv2 = nn.Conv2d(6, 16, 5)
        self.pool = nn.MaxPool2d(2, 2)
        self.fc1 = nn.Linear(16*4*4, 120)
        self.fc2 = nn.Linear(120, 84)
        self.fc3 = nn.Linear(84, 10)

    def forward(self, x):
        x = torch.relu(self.conv1(x))
        x = self.pool(x)
        x = torch.relu(self.conv2(x))
        x = self.pool(x)
        x = x.view(-1, 16*4*4)
        x = torch.relu(self.fc1(x))
        x = torch.relu(self.fc2(x))
        x = self.fc3(x)
        return x


transform = transforms.Compose([
    transforms.ToTensor(),
    transforms.Normalize((0.5,), (0.5,))
])

train_dataset = torchvision.datasets.MNIST(root='./data', train=True, download=
test_dataset = torchvision.datasets.MNIST(root='./data', train=False, download=

train_loader = torch.utils.data.DataLoader(train_dataset, batch_size=64, shuffl
test_loader = torch.utils.data.DataLoader(test_dataset, batch_size=64, shuffle=


model = LeNet()
loss_fn = nn.CrossEntropyLoss()
optimizer = optim.SGD(model.parameters(), lr=0.001, momentum=0.9)


def train(model, train_loader, criterion, optimizer, epoch):
  model.train()
  running_loss = 0.0
  total = 0
  correct = 0

  for index, (data, label) in enumerate(train_loader):
      optimizer.zero_grad()
```

```
        output = model(data)
        loss = criterion(output, label)
        loss.backward()

        optimizer.step()

        running_loss += loss.item()
        _, predicted = torch.max(output.data, 1)
        total += label.size(0)
        correct += (predicted == label).sum().item()

        if index % 100 == 0:
            progress = 100. * index / len(train_loader)
            print('Train Epoch: {} [{}/{} ({:.4f}%)]\tLoss: {:.6f}'.format(epoch,

    train_loss = running_loss / len(train_loader)
    accuracy = 100. * correct / total
    return train_loss, accuracy

def test(model, test_loader, criterion):
    model.eval()
    test_loss = 0.0
    total = 0
    correct = 0

    with torch.no_grad():
        for index, (data, label) in enumerate(test_loader):
            output = model(data)
            loss = criterion(output, label)
            test_loss += loss.item()
            _, predicted = torch.max(output.data, 1)
            total += label.size(0)
            correct += (predicted == label).sum().item()

    test_loss /= len(test_loader)
    accuracy = 100. * correct / total
    print('Test set: Average loss: {:.4f}, Accuracy: {}/{} ({:.4f}%)'.format(test
    return test_loss, accuracy


wandb.init(project='LeNet-CNN')

# Training loop
epochs = 10
for epoch in range(1, epochs + 1):
    train_loss, train_accuracy = train(model, train_loader, loss_fn, optimizer,
    test_loss, test_accuracy = test(model, test_loader, loss_fn)
```

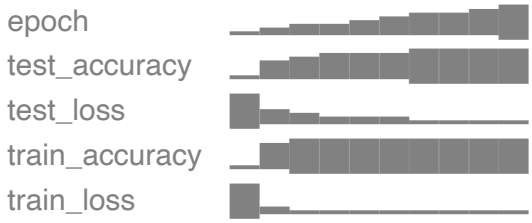
```
    wandb.log({
        "epoch": epoch,
        "train_loss": train_loss,
        "train_accuracy": train_accuracy,
        "test_loss": test_loss,
        "test_accuracy": test_accuracy
    })

wandb.finish()
```

```
wandb: Currently logged in as: hugthebee. Use `wandb login --relogin` to fo
Tracking run with wandb version 0.16.4
Run data is saved locally in /content/wandb/run-20240309_132553-haoemj73
Syncing run neat-snowball-1 to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/LeNet-CNN
View run at https://wandb.ai/hugthebee/LeNet-CNN/runs/haoemj73
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.294955
Train Epoch: 1 [6400/60000 (10.6610%)]  Loss: 2.282586
Train Epoch: 1 [12800/60000 (21.3220%)] Loss: 2.274316
Train Epoch: 1 [19200/60000 (31.9829%)] Loss: 2.238298
Train Epoch: 1 [25600/60000 (42.6439%)] Loss: 2.167442
Train Epoch: 1 [32000/60000 (53.3049%)] Loss: 1.861004
Train Epoch: 1 [38400/60000 (63.9659%)] Loss: 0.958871
Train Epoch: 1 [44800/60000 (74.6269%)] Loss: 0.470287
Train Epoch: 1 [51200/60000 (85.2878%)] Loss: 0.297824
Train Epoch: 1 [57600/60000 (95.9488%)] Loss: 0.263150
Test set: Average loss: 0.3608, Accuracy: 8953/10000 (89.5300%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.457948
Train Epoch: 2 [6400/60000 (10.6610%)]  Loss: 0.392471
Train Epoch: 2 [12800/60000 (21.3220%)] Loss: 0.295093
Train Epoch: 2 [19200/60000 (31.9829%)] Loss: 0.302700
Train Epoch: 2 [25600/60000 (42.6439%)] Loss: 0.357308
Train Epoch: 2 [32000/60000 (53.3049%)] Loss: 0.397990
Train Epoch: 2 [38400/60000 (63.9659%)] Loss: 0.119122
Train Epoch: 2 [44800/60000 (74.6269%)] Loss: 0.119623
Train Epoch: 2 [51200/60000 (85.2878%)] Loss: 0.210150
Train Epoch: 2 [57600/60000 (95.9488%)] Loss: 0.149549
Test set: Average loss: 0.1765, Accuracy: 9450/10000 (94.5000%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.113381
Train Epoch: 3 [6400/60000 (10.6610%)]  Loss: 0.141970
Train Epoch: 3 [12800/60000 (21.3220%)] Loss: 0.173495
Train Epoch: 3 [19200/60000 (31.9829%)] Loss: 0.196615
Train Epoch: 3 [25600/60000 (42.6439%)] Loss: 0.179235
Train Epoch: 3 [32000/60000 (53.3049%)] Loss: 0.138590
Train Epoch: 3 [38400/60000 (63.9659%)] Loss: 0.063260
Train Epoch: 3 [44800/60000 (74.6269%)] Loss: 0.075924
Train Epoch: 3 [51200/60000 (85.2878%)] Loss: 0.213301
Train Epoch: 3 [57600/60000 (95.9488%)] Loss: 0.147457
Test set: Average loss: 0.1291, Accuracy: 9615/10000 (96.1500%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.052232
Train Epoch: 4 [6400/60000 (10.6610%)]  Loss: 0.077141
Train Epoch: 4 [12800/60000 (21.3220%)] Loss: 0.048941
Train Epoch: 4 [19200/60000 (31.9829%)] Loss: 0.110147
```

```
Train Epoch: 4 [25600/60000 (42.6439%)] Loss: 0.087216
Train Epoch: 4 [32000/60000 (53.3049%)] Loss: 0.034290
Train Epoch: 4 [38400/60000 (63.9659%)] Loss: 0.221593
Train Epoch: 4 [44800/60000 (74.6269%)] Loss: 0.108452
Train Epoch: 4 [51200/60000 (85.2878%)] Loss: 0.154750
Train Epoch: 4 [57600/60000 (95.9488%)] Loss: 0.038483
Test set: Average loss: 0.0985, Accuracy: 9677/10000 (96.7700%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.057054
Train Epoch: 5 [6400/60000 (10.6610%)]  Loss: 0.216123
Train Epoch: 5 [12800/60000 (21.3220%)] Loss: 0.066681
Train Epoch: 5 [19200/60000 (31.9829%)] Loss: 0.057051
Train Epoch: 5 [25600/60000 (42.6439%)] Loss: 0.037148
Train Epoch: 5 [32000/60000 (53.3049%)] Loss: 0.073471
Train Epoch: 5 [38400/60000 (63.9659%)] Loss: 0.081579
Train Epoch: 5 [44800/60000 (74.6269%)] Loss: 0.198873
Train Epoch: 5 [51200/60000 (85.2878%)] Loss: 0.103638
Train Epoch: 5 [57600/60000 (95.9488%)] Loss: 0.179211
Test set: Average loss: 0.0775, Accuracy: 9756/10000 (97.5600%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.148687
Train Epoch: 6 [6400/60000 (10.6610%)]  Loss: 0.102897
Train Epoch: 6 [12800/60000 (21.3220%)] Loss: 0.064168
Train Epoch: 6 [19200/60000 (31.9829%)] Loss: 0.011784
Train Epoch: 6 [25600/60000 (42.6439%)] Loss: 0.108049
Train Epoch: 6 [32000/60000 (53.3049%)] Loss: 0.082507
Train Epoch: 6 [38400/60000 (63.9659%)] Loss: 0.020885
Train Epoch: 6 [44800/60000 (74.6269%)] Loss: 0.095316
Train Epoch: 6 [51200/60000 (85.2878%)] Loss: 0.011687
Train Epoch: 6 [57600/60000 (95.9488%)] Loss: 0.032328
Test set: Average loss: 0.0764, Accuracy: 9753/10000 (97.5300%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.102912
Train Epoch: 7 [6400/60000 (10.6610%)]  Loss: 0.103904
Train Epoch: 7 [12800/60000 (21.3220%)] Loss: 0.040773
Train Epoch: 7 [19200/60000 (31.9829%)] Loss: 0.025725
Train Epoch: 7 [25600/60000 (42.6439%)] Loss: 0.022857
Train Epoch: 7 [32000/60000 (53.3049%)] Loss: 0.045091
Train Epoch: 7 [38400/60000 (63.9659%)] Loss: 0.031942
Train Epoch: 7 [44800/60000 (74.6269%)] Loss: 0.052795
Train Epoch: 7 [51200/60000 (85.2878%)] Loss: 0.067705
Train Epoch: 7 [57600/60000 (95.9488%)] Loss: 0.031494
Test set: Average loss: 0.0633, Accuracy: 9796/10000 (97.9600%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.143549
Train Epoch: 8 [6400/60000 (10.6610%)]  Loss: 0.075595
Train Epoch: 8 [12800/60000 (21.3220%)] Loss: 0.015932
Train Epoch: 8 [19200/60000 (31.9829%)] Loss: 0.044214
Train Epoch: 8 [25600/60000 (42.6439%)] Loss: 0.024858
Train Epoch: 8 [32000/60000 (53.3049%)] Loss: 0.064366
Train Epoch: 8 [38400/60000 (63.9659%)] Loss: 0.186551
Train Epoch: 8 [44800/60000 (74.6269%)] Loss: 0.101175
Train Epoch: 8 [51200/60000 (85.2878%)] Loss: 0.023470
Train Epoch: 8 [57600/60000 (95.9488%)] Loss: 0.146370
Test set: Average loss: 0.0607, Accuracy: 9808/10000 (98.0800%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.016068
Train Epoch: 9 [6400/60000 (10.6610%)]  Loss: 0.070499
Train Epoch: 9 [12800/60000 (21.3220%)] Loss: 0.042347
```

```
Train Epoch: 9 [19200/60000 (31.9829%)] Loss: 0.103899
Train Epoch: 9 [25600/60000 (42.6439%)] Loss: 0.107765
Train Epoch: 9 [32000/60000 (53.3049%)] Loss: 0.051333
Train Epoch: 9 [38400/60000 (63.9659%)] Loss: 0.069486
Train Epoch: 9 [44800/60000 (74.6269%)] Loss: 0.220516
Train Epoch: 9 [51200/60000 (85.2878%)] Loss: 0.138529
Train Epoch: 9 [57600/60000 (95.9488%)] Loss: 0.012014
Test set: Average loss: 0.0529, Accuracy: 9824/10000 (98.2400%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.009349
Train Epoch: 10 [6400/60000 (10.6610%)] Loss: 0.015817
Train Epoch: 10 [12800/60000 (21.3220%)]        Loss: 0.037364
Train Epoch: 10 [19200/60000 (31.9829%)]        Loss: 0.048398
Train Epoch: 10 [25600/60000 (42.6439%)]        Loss: 0.186297
Train Epoch: 10 [32000/60000 (53.3049%)]        Loss: 0.091547
Train Epoch: 10 [38400/60000 (63.9659%)]        Loss: 0.075124
Train Epoch: 10 [44800/60000 (74.6269%)]        Loss: 0.086675
Train Epoch: 10 [51200/60000 (85.2878%)]        Loss: 0.115675
Train Epoch: 10 [57600/60000 (95.9488%)]        Loss: 0.018509
Test set: Average loss: 0.0510, Accuracy: 9840/10000 (98.4000%)
```

Run history:



Run summary:

epoch	10
test_accuracy	98.4
test_loss	0.05095
train_accuracy	98.20833
train_loss	0.05886

View run **neat-snowball-1** at: https://wandb.ai/hugthebee/LeNet-CNN/runs/haoemj73
Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: `./wandb/run-20240309_132553-haoemj73/logs`

Different hyperparameters:

We test with six different hyperparameters:

1. Batch size (how many images are considered at a time) - 4 and 128
2. Optimizer (how the new parameters are determined) - Adam and Adagrad
3. Learning rate (step size with which parameters are updated) - 0.01 and 0.1

We notice from the results of the plots and accuracies that:

- For batch size, both models perform equally as good with the batch size of 128 being slightly better performing than that model with a batch size of 4. Having a good batch size is important to prevent overfitting and enable faster convergence.
  Train accuracy (batch size = 4) = 98.07%
  Train accuracy (batch size = 128) = 98.20%

- In the case of using a different optimiser, we notice that Adam performs signifacntly better than Adagrad. The choice of our optimal optimiser depends on the kind of data and model we are using.
  Train accuracy (optimiser = Adam) = 99.42%
  Train accuracy (optmiser = Adagrad) = 94.61%

- For different values of the learning rate, a smaller learning rate vastly outperforms a higher learning rate. Since the learning rate determines the step size of the updated parameters, having a very large learning rate will lead to overshooting and poor accuracy.
  Train accuracy (learning rate = 0.01) = 99.53%
  Train accuracy (learning rate = 0.1) = 10.72%

```
hyperparameters = [
    {"batch_size": 4, "learning_rate": 0.001, "optimizer": optim.SGD},
    {"batch_size": 128, "learning_rate": 0.001, "optimizer": optim.SGD},
    {"batch_size": 64, "learning_rate": 0.001, "optimizer": optim.Adam},
    {"batch_size": 64, "learning_rate": 0.001, "optimizer": optim.Adagrad},
    {"batch_size": 64, "learning_rate": 0.01, "optimizer": optim.SGD},
    {"batch_size": 64, "learning_rate": 0.1, "optimizer": optim.SGD}
]


for i, param in enumerate(hyperparameters):
    print(f"\nHyperparameter {i+1}:")
    print(f"Batch Size: {param['batch_size']}, Learning Rate: {param['learning_

    model = LeNet()
    loss_fn = nn.CrossEntropyLoss()

    if (param['optimizer'] == optim.SGD):
      optimizer = param['optimizer'](model.parameters(), lr=param['learning_rat
    else:
      optimizer = param['optimizer'](model.parameters(), lr=param['learning_rat

    wandb.init(project='Different-Hyperparameters')
    epochs = 10
    for epoch in range(1, epochs + 1):
      train_loss, train_accuracy = train(model, train_loader, loss_fn, optimize
      test_loss, test_accuracy = test(model, test_loader, loss_fn)

      wandb.log({
          f"train_loss (Hyperparameter {i+1})": train_loss,
          f"train_accuracy (Hyperparameter {i+1})": train_accuracy,
          f"test_loss (Hyperparameter {i+1})": test_loss,
          f"test_accuracy (Hyperparameter {i+1})": test_accuracy
      })

    wandb.finish()
```

```
Hyperparameter 1:
Batch Size: 4, Learning Rate: 0.001, Optimizer: SGD
Tracking run with wandb version 0.16.4
Run data is saved locally in /content/wandb/run-20240309_133155-ruw0iv2i
Syncing run graceful-river-1 to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/Different-Hyperparameters
View run at https://wandb.ai/hugthebee/Different-Hyperparameters/runs/ruw0iv2i
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.295034
Train Epoch: 1 [6400/60000 (10.6610%)]  Loss: 2.282092
Train Epoch: 1 [12800/60000 (21.3220%)] Loss: 2.284101
Train Epoch: 1 [19200/60000 (31.9829%)] Loss: 2.281743
```

```
Train Epoch: 1 [25600/60000 (42.6439%)] Loss: 2.281213
Train Epoch: 1 [32000/60000 (53.3049%)] Loss: 2.251649
Train Epoch: 1 [38400/60000 (63.9659%)] Loss: 2.164249
Train Epoch: 1 [44800/60000 (74.6269%)] Loss: 1.925964
Train Epoch: 1 [51200/60000 (85.2878%)] Loss: 1.211648
Train Epoch: 1 [57600/60000 (95.9488%)] Loss: 0.556046
Test set: Average loss: 0.4914, Accuracy: 8537/10000 (85.3700%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.535566
Train Epoch: 2 [6400/60000 (10.6610%)]  Loss: 0.440026
Train Epoch: 2 [12800/60000 (21.3220%)] Loss: 0.397242
Train Epoch: 2 [19200/60000 (31.9829%)] Loss: 0.342157
Train Epoch: 2 [25600/60000 (42.6439%)] Loss: 0.249732
Train Epoch: 2 [32000/60000 (53.3049%)] Loss: 0.287875
Train Epoch: 2 [38400/60000 (63.9659%)] Loss: 0.226770
Train Epoch: 2 [44800/60000 (74.6269%)] Loss: 0.261540
Train Epoch: 2 [51200/60000 (85.2878%)] Loss: 0.290765
Train Epoch: 2 [57600/60000 (95.9488%)] Loss: 0.384077
Test set: Average loss: 0.1915, Accuracy: 9393/10000 (93.9300%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.184161
Train Epoch: 3 [6400/60000 (10.6610%)]  Loss: 0.272161
Train Epoch: 3 [12800/60000 (21.3220%)] Loss: 0.137139
Train Epoch: 3 [19200/60000 (31.9829%)] Loss: 0.198062
Train Epoch: 3 [25600/60000 (42.6439%)] Loss: 0.154313
Train Epoch: 3 [32000/60000 (53.3049%)] Loss: 0.213222
Train Epoch: 3 [38400/60000 (63.9659%)] Loss: 0.202148
Train Epoch: 3 [44800/60000 (74.6269%)] Loss: 0.111720
Train Epoch: 3 [51200/60000 (85.2878%)] Loss: 0.036918
Train Epoch: 3 [57600/60000 (95.9488%)] Loss: 0.129389
Test set: Average loss: 0.1263, Accuracy: 9603/10000 (96.0300%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.109041
Train Epoch: 4 [6400/60000 (10.6610%)]  Loss: 0.305185
Train Epoch: 4 [12800/60000 (21.3220%)] Loss: 0.089468
Train Epoch: 4 [19200/60000 (31.9829%)] Loss: 0.110754
Train Epoch: 4 [25600/60000 (42.6439%)] Loss: 0.074230
Train Epoch: 4 [32000/60000 (53.3049%)] Loss: 0.225371
Train Epoch: 4 [38400/60000 (63.9659%)] Loss: 0.067834
Train Epoch: 4 [44800/60000 (74.6269%)] Loss: 0.098630
Train Epoch: 4 [51200/60000 (85.2878%)] Loss: 0.027691
Train Epoch: 4 [57600/60000 (95.9488%)] Loss: 0.064887
Test set: Average loss: 0.0992, Accuracy: 9682/10000 (96.8200%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.029665
Train Epoch: 5 [6400/60000 (10.6610%)]  Loss: 0.130003
Train Epoch: 5 [12800/60000 (21.3220%)] Loss: 0.072476
Train Epoch: 5 [19200/60000 (31.9829%)] Loss: 0.176538
Train Epoch: 5 [25600/60000 (42.6439%)] Loss: 0.035763
Train Epoch: 5 [32000/60000 (53.3049%)] Loss: 0.103353
Train Epoch: 5 [38400/60000 (63.9659%)] Loss: 0.046803
Train Epoch: 5 [44800/60000 (74.6269%)] Loss: 0.079421
Train Epoch: 5 [51200/60000 (85.2878%)] Loss: 0.019302
Train Epoch: 5 [57600/60000 (95.9488%)] Loss: 0.173577
Test set: Average loss: 0.0822, Accuracy: 9738/10000 (97.3800%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.049539
Train Epoch: 6 [6400/60000 (10.6610%)]  Loss: 0.111063
Train Epoch: 6 [12800/60000 (21.3220%)] Loss: 0.037532
```

```
Train Epoch: 6 [19200/60000 (31.9829%)] Loss: 0.055399
Train Epoch: 6 [25600/60000 (42.6439%)] Loss: 0.048131
Train Epoch: 6 [32000/60000 (53.3049%)] Loss: 0.052821
Train Epoch: 6 [38400/60000 (63.9659%)] Loss: 0.045274
Train Epoch: 6 [44800/60000 (74.6269%)] Loss: 0.080127
Train Epoch: 6 [51200/60000 (85.2878%)] Loss: 0.068770
Train Epoch: 6 [57600/60000 (95.9488%)] Loss: 0.093825
Test set: Average loss: 0.0742, Accuracy: 9768/10000 (97.6800%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.039345
Train Epoch: 7 [6400/60000 (10.6610%)]  Loss: 0.092360
Train Epoch: 7 [12800/60000 (21.3220%)] Loss: 0.048686
Train Epoch: 7 [19200/60000 (31.9829%)] Loss: 0.050975
Train Epoch: 7 [25600/60000 (42.6439%)] Loss: 0.048329
Train Epoch: 7 [32000/60000 (53.3049%)] Loss: 0.041494
Train Epoch: 7 [38400/60000 (63.9659%)] Loss: 0.018882
Train Epoch: 7 [44800/60000 (74.6269%)] Loss: 0.010595
Train Epoch: 7 [51200/60000 (85.2878%)] Loss: 0.040737
Train Epoch: 7 [57600/60000 (95.9488%)] Loss: 0.071342
Test set: Average loss: 0.0701, Accuracy: 9770/10000 (97.7000%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.029308
Train Epoch: 8 [6400/60000 (10.6610%)]  Loss: 0.061637
Train Epoch: 8 [12800/60000 (21.3220%)] Loss: 0.006315
Train Epoch: 8 [19200/60000 (31.9829%)] Loss: 0.148645
Train Epoch: 8 [25600/60000 (42.6439%)] Loss: 0.110318
Train Epoch: 8 [32000/60000 (53.3049%)] Loss: 0.078591
Train Epoch: 8 [38400/60000 (63.9659%)] Loss: 0.050162
Train Epoch: 8 [44800/60000 (74.6269%)] Loss: 0.069302
Train Epoch: 8 [51200/60000 (85.2878%)] Loss: 0.072542
Train Epoch: 8 [57600/60000 (95.9488%)] Loss: 0.222074
Test set: Average loss: 0.0658, Accuracy: 9789/10000 (97.8900%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.059077
Train Epoch: 9 [6400/60000 (10.6610%)]  Loss: 0.034575
Train Epoch: 9 [12800/60000 (21.3220%)] Loss: 0.058566
Train Epoch: 9 [19200/60000 (31.9829%)] Loss: 0.069438
Train Epoch: 9 [25600/60000 (42.6439%)] Loss: 0.042082
Train Epoch: 9 [32000/60000 (53.3049%)] Loss: 0.031233
Train Epoch: 9 [38400/60000 (63.9659%)] Loss: 0.008535
Train Epoch: 9 [44800/60000 (74.6269%)] Loss: 0.036639
Train Epoch: 9 [51200/60000 (85.2878%)] Loss: 0.026515
Train Epoch: 9 [57600/60000 (95.9488%)] Loss: 0.018336
Test set: Average loss: 0.0682, Accuracy: 9769/10000 (97.6900%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.022816
Train Epoch: 10 [6400/60000 (10.6610%)] Loss: 0.028755
Train Epoch: 10 [12800/60000 (21.3220%)]        Loss: 0.068614
Train Epoch: 10 [19200/60000 (31.9829%)]        Loss: 0.062773
Train Epoch: 10 [25600/60000 (42.6439%)]        Loss: 0.031831
Train Epoch: 10 [32000/60000 (53.3049%)]        Loss: 0.015851
Train Epoch: 10 [38400/60000 (63.9659%)]        Loss: 0.012002
Train Epoch: 10 [44800/60000 (74.6269%)]        Loss: 0.033075
Train Epoch: 10 [51200/60000 (85.2878%)]        Loss: 0.090184
Train Epoch: 10 [57600/60000 (95.9488%)]        Loss: 0.059510
Test set: Average loss: 0.0578, Accuracy: 9818/10000 (98.1800%)
```

Run history: **Run summary:**

Run history:

test_accuracy (Hyperparameter 1)	
test_loss (Hyperparameter 1)	
train_accuracy (Hyperparameter 1)	
train_loss (Hyperparameter 1)	

Run summary:

test_accuracy (Hyperparameter 1)	98.18
test_loss (Hyperparameter 1)	0.05785
train_accuracy (Hyperparameter 1)	98.07333
train_loss (Hyperparameter 1)	0.06149

View run **graceful-river-1** at: https://wandb.ai/hugthebee/Different-Hyperparameters/runs/ruw0iv2i
Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: `./wandb/run-20240309_133155-ruw0iv2i/logs`

```
Hyperparameter 2:
Batch Size: 128, Learning Rate: 0.001, Optimizer: SGD
```

Tracking run with wandb version 0.16.4
Run data is saved locally in `/content/wandb/run-20240309_133726-nm3ndoj1`
Syncing run **ancient-armadillo-2** to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/Different-Hyperparameters
View run at https://wandb.ai/hugthebee/Different-Hyperparameters/runs/nm3ndoj1

```
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.327308
Train Epoch: 1 [6400/60000 (10.6610%)]  Loss: 2.295189
Train Epoch: 1 [12800/60000 (21.3220%)] Loss: 2.304505
Train Epoch: 1 [19200/60000 (31.9829%)] Loss: 2.278725
Train Epoch: 1 [25600/60000 (42.6439%)] Loss: 2.253054
Train Epoch: 1 [32000/60000 (53.3049%)] Loss: 2.218751
Train Epoch: 1 [38400/60000 (63.9659%)] Loss: 2.153521
Train Epoch: 1 [44800/60000 (74.6269%)] Loss: 1.776694
Train Epoch: 1 [51200/60000 (85.2878%)] Loss: 1.078003
Train Epoch: 1 [57600/60000 (95.9488%)] Loss: 0.627104
Test set: Average loss: 0.6163, Accuracy: 8089/10000 (80.8900%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.754803
Train Epoch: 2 [6400/60000 (10.6610%)]  Loss: 0.379034
Train Epoch: 2 [12800/60000 (21.3220%)] Loss: 0.436004
Train Epoch: 2 [19200/60000 (31.9829%)] Loss: 0.376382
Train Epoch: 2 [25600/60000 (42.6439%)] Loss: 0.306242
Train Epoch: 2 [32000/60000 (53.3049%)] Loss: 0.259069
Train Epoch: 2 [38400/60000 (63.9659%)] Loss: 0.241895
Train Epoch: 2 [44800/60000 (74.6269%)] Loss: 0.255231
Train Epoch: 2 [51200/60000 (85.2878%)] Loss: 0.185274
Train Epoch: 2 [57600/60000 (95.9488%)] Loss: 0.311747
Test set: Average loss: 0.1877, Accuracy: 9417/10000 (94.1700%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.147741
Train Epoch: 3 [6400/60000 (10.6610%)]  Loss: 0.244986
Train Epoch: 3 [12800/60000 (21.3220%)] Loss: 0.170215
Train Epoch: 3 [19200/60000 (31.9829%)] Loss: 0.244487
Train Epoch: 3 [25600/60000 (42.6439%)] Loss: 0.134432
```

```
Train Epoch: 3 [32000/60000 (53.3049%)] Loss: 0.298649
Train Epoch: 3 [38400/60000 (63.9659%)] Loss: 0.088008
Train Epoch: 3 [44800/60000 (74.6269%)] Loss: 0.114233
Train Epoch: 3 [51200/60000 (85.2878%)] Loss: 0.100170
Train Epoch: 3 [57600/60000 (95.9488%)] Loss: 0.211073
Test set: Average loss: 0.1201, Accuracy: 9629/10000 (96.2900%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.124808
Train Epoch: 4 [6400/60000 (10.6610%)]  Loss: 0.077909
Train Epoch: 4 [12800/60000 (21.3220%)] Loss: 0.284594
Train Epoch: 4 [19200/60000 (31.9829%)] Loss: 0.099775
Train Epoch: 4 [25600/60000 (42.6439%)] Loss: 0.034008
Train Epoch: 4 [32000/60000 (53.3049%)] Loss: 0.050102
Train Epoch: 4 [38400/60000 (63.9659%)] Loss: 0.416476
Train Epoch: 4 [44800/60000 (74.6269%)] Loss: 0.062459
Train Epoch: 4 [51200/60000 (85.2878%)] Loss: 0.191165
Train Epoch: 4 [57600/60000 (95.9488%)] Loss: 0.094514
Test set: Average loss: 0.0915, Accuracy: 9720/10000 (97.2000%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.033603
Train Epoch: 5 [6400/60000 (10.6610%)]  Loss: 0.023601
Train Epoch: 5 [12800/60000 (21.3220%)] Loss: 0.037615
Train Epoch: 5 [19200/60000 (31.9829%)] Loss: 0.109177
Train Epoch: 5 [25600/60000 (42.6439%)] Loss: 0.131975
Train Epoch: 5 [32000/60000 (53.3049%)] Loss: 0.138951
Train Epoch: 5 [38400/60000 (63.9659%)] Loss: 0.208642
Train Epoch: 5 [44800/60000 (74.6269%)] Loss: 0.052715
Train Epoch: 5 [51200/60000 (85.2878%)] Loss: 0.086872
Train Epoch: 5 [57600/60000 (95.9488%)] Loss: 0.027823
Test set: Average loss: 0.0788, Accuracy: 9757/10000 (97.5700%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.061291
Train Epoch: 6 [6400/60000 (10.6610%)]  Loss: 0.124778
Train Epoch: 6 [12800/60000 (21.3220%)] Loss: 0.119125
Train Epoch: 6 [19200/60000 (31.9829%)] Loss: 0.097857
Train Epoch: 6 [25600/60000 (42.6439%)] Loss: 0.104027
Train Epoch: 6 [32000/60000 (53.3049%)] Loss: 0.141308
Train Epoch: 6 [38400/60000 (63.9659%)] Loss: 0.028234
Train Epoch: 6 [44800/60000 (74.6269%)] Loss: 0.033189
Train Epoch: 6 [51200/60000 (85.2878%)] Loss: 0.076854
Train Epoch: 6 [57600/60000 (95.9488%)] Loss: 0.174892
Test set: Average loss: 0.0721, Accuracy: 9765/10000 (97.6500%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.076246
Train Epoch: 7 [6400/60000 (10.6610%)]  Loss: 0.012449
Train Epoch: 7 [12800/60000 (21.3220%)] Loss: 0.051345
Train Epoch: 7 [19200/60000 (31.9829%)] Loss: 0.149276
Train Epoch: 7 [25600/60000 (42.6439%)] Loss: 0.096592
Train Epoch: 7 [32000/60000 (53.3049%)] Loss: 0.021262
Train Epoch: 7 [38400/60000 (63.9659%)] Loss: 0.009197
Train Epoch: 7 [44800/60000 (74.6269%)] Loss: 0.032720
Train Epoch: 7 [51200/60000 (85.2878%)] Loss: 0.051880
Train Epoch: 7 [57600/60000 (95.9488%)] Loss: 0.103685
Test set: Average loss: 0.0620, Accuracy: 9800/10000 (98.0000%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.068470
Train Epoch: 8 [6400/60000 (10.6610%)]  Loss: 0.030622
Train Epoch: 8 [12800/60000 (21.3220%)] Loss: 0.094533
Train Epoch: 8 [19200/60000 (31.9829%)] Loss: 0.046790
```

```
Train Epoch: 8 [25600/60000 (42.6439%)] Loss: 0.129873
Train Epoch: 8 [32000/60000 (53.3049%)] Loss: 0.104056
Train Epoch: 8 [38400/60000 (63.9659%)] Loss: 0.199622
Train Epoch: 8 [44800/60000 (74.6269%)] Loss: 0.057680
Train Epoch: 8 [51200/60000 (85.2878%)] Loss: 0.016930
Train Epoch: 8 [57600/60000 (95.9488%)] Loss: 0.016904
Test set: Average loss: 0.0634, Accuracy: 9787/10000 (97.8700%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.054189
Train Epoch: 9 [6400/60000 (10.6610%)]  Loss: 0.025441
Train Epoch: 9 [12800/60000 (21.3220%)] Loss: 0.104751
Train Epoch: 9 [19200/60000 (31.9829%)] Loss: 0.025936
Train Epoch: 9 [25600/60000 (42.6439%)] Loss: 0.051602
Train Epoch: 9 [32000/60000 (53.3049%)] Loss: 0.086105
Train Epoch: 9 [38400/60000 (63.9659%)] Loss: 0.173813
Train Epoch: 9 [44800/60000 (74.6269%)] Loss: 0.055233
Train Epoch: 9 [51200/60000 (85.2878%)] Loss: 0.016743
Train Epoch: 9 [57600/60000 (95.9488%)] Loss: 0.013323
Test set: Average loss: 0.0524, Accuracy: 9834/10000 (98.3400%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.025501
Train Epoch: 10 [6400/60000 (10.6610%)] Loss: 0.024042
Train Epoch: 10 [12800/60000 (21.3220%)]        Loss: 0.017396
Train Epoch: 10 [19200/60000 (31.9829%)]        Loss: 0.066339
Train Epoch: 10 [25600/60000 (42.6439%)]        Loss: 0.049268
Train Epoch: 10 [32000/60000 (53.3049%)]        Loss: 0.075807
Train Epoch: 10 [38400/60000 (63.9659%)]        Loss: 0.090288
Train Epoch: 10 [44800/60000 (74.6269%)]        Loss: 0.037114
Train Epoch: 10 [51200/60000 (85.2878%)]        Loss: 0.023214
Train Epoch: 10 [57600/60000 (95.9488%)]        Loss: 0.069725
Test set: Average loss: 0.0502, Accuracy: 9837/10000 (98.3700%)
```

Run history:

test_accuracy (Hyperparameter 2)	
test_loss (Hyperparameter 2)	
train_accuracy (Hyperparameter 2)	
train_loss (Hyperparameter 2)	

Run summary:

test_accuracy (Hyperparameter 2)	98.37
test_loss (Hyperparameter 2)	0.05016
train_accuracy (Hyperparameter 2)	98.20667
train_loss (Hyperparameter 2)	0.05731

View run **ancient-armadillo-2** at: https://wandb.ai/hugthebee/Different-Hyperparameters/runs/nm3ndoj1
Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: `./wandb/run-20240309_133726-nm3ndoj1/logs`

```
Hyperparameter 3:
Batch Size: 64, Learning Rate: 0.001, Optimizer: Adam
```

Tracking run with wandb version 0.16.4
Run data is saved locally in /content/wandb/run-20240309_134255-lbfsv0tf
Syncing run **skilled-lion-3** to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/Different-Hyperparameters
View run at https://wandb.ai/hugthebee/Different-Hyperparameters/runs/lbfsv0tf

```
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.303605
Train Epoch: 1 [6400/60000 (10.6610%)]  Loss: 0.418885
Train Epoch: 1 [12800/60000 (21.3220%)] Loss: 0.123931
Train Epoch: 1 [19200/60000 (31.9829%)] Loss: 0.309061
Train Epoch: 1 [25600/60000 (42.6439%)] Loss: 0.160784
Train Epoch: 1 [32000/60000 (53.3049%)] Loss: 0.112147
Train Epoch: 1 [38400/60000 (63.9659%)] Loss: 0.176707
Train Epoch: 1 [44800/60000 (74.6269%)] Loss: 0.048996
Train Epoch: 1 [51200/60000 (85.2878%)] Loss: 0.058572
Train Epoch: 1 [57600/60000 (95.9488%)] Loss: 0.009201
Test set: Average loss: 0.0827, Accuracy: 9737/10000 (97.3700%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.058141
Train Epoch: 2 [6400/60000 (10.6610%)]  Loss: 0.060023
Train Epoch: 2 [12800/60000 (21.3220%)] Loss: 0.065499
Train Epoch: 2 [19200/60000 (31.9829%)] Loss: 0.121236
Train Epoch: 2 [25600/60000 (42.6439%)] Loss: 0.073605
Train Epoch: 2 [32000/60000 (53.3049%)] Loss: 0.031027
Train Epoch: 2 [38400/60000 (63.9659%)] Loss: 0.033881
Train Epoch: 2 [44800/60000 (74.6269%)] Loss: 0.044898
Train Epoch: 2 [51200/60000 (85.2878%)] Loss: 0.015048
Train Epoch: 2 [57600/60000 (95.9488%)] Loss: 0.012319
Test set: Average loss: 0.0547, Accuracy: 9827/10000 (98.2700%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.095367
Train Epoch: 3 [6400/60000 (10.6610%)]  Loss: 0.021339
Train Epoch: 3 [12800/60000 (21.3220%)] Loss: 0.005155
Train Epoch: 3 [19200/60000 (31.9829%)] Loss: 0.079927
Train Epoch: 3 [25600/60000 (42.6439%)] Loss: 0.013934
Train Epoch: 3 [32000/60000 (53.3049%)] Loss: 0.024625
Train Epoch: 3 [38400/60000 (63.9659%)] Loss: 0.026272
Train Epoch: 3 [44800/60000 (74.6269%)] Loss: 0.030956
Train Epoch: 3 [51200/60000 (85.2878%)] Loss: 0.054841
Train Epoch: 3 [57600/60000 (95.9488%)] Loss: 0.073742
Test set: Average loss: 0.0446, Accuracy: 9855/10000 (98.5500%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.050921
Train Epoch: 4 [6400/60000 (10.6610%)]  Loss: 0.048421
Train Epoch: 4 [12800/60000 (21.3220%)] Loss: 0.013065
Train Epoch: 4 [19200/60000 (31.9829%)] Loss: 0.116274
Train Epoch: 4 [25600/60000 (42.6439%)] Loss: 0.018184
Train Epoch: 4 [32000/60000 (53.3049%)] Loss: 0.020620
Train Epoch: 4 [38400/60000 (63.9659%)] Loss: 0.001085
Train Epoch: 4 [44800/60000 (74.6269%)] Loss: 0.005009
Train Epoch: 4 [51200/60000 (85.2878%)] Loss: 0.136345
Train Epoch: 4 [57600/60000 (95.9488%)] Loss: 0.041047
Test set: Average loss: 0.0416, Accuracy: 9861/10000 (98.6100%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.033224
Train Epoch: 5 [6400/60000 (10.6610%)]  Loss: 0.008849
Train Epoch: 5 [12800/60000 (21.3220%)] Loss: 0.027861
Train Epoch: 5 [19200/60000 (31.9829%)] Loss: 0.020453
Train Epoch: 5 [25600/60000 (42.6439%)] Loss: 0.068445
```

```
Train Epoch: 5 [32000/60000 (53.3049%)] Loss: 0.001370
Train Epoch: 5 [38400/60000 (63.9659%)] Loss: 0.003332
Train Epoch: 5 [44800/60000 (74.6269%)] Loss: 0.026455
Train Epoch: 5 [51200/60000 (85.2878%)] Loss: 0.023078
Train Epoch: 5 [57600/60000 (95.9488%)] Loss: 0.027724
Test set: Average loss: 0.0342, Accuracy: 9884/10000 (98.8400%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.010939
Train Epoch: 6 [6400/60000 (10.6610%)]  Loss: 0.011418
Train Epoch: 6 [12800/60000 (21.3220%)] Loss: 0.003338
Train Epoch: 6 [19200/60000 (31.9829%)] Loss: 0.002630
Train Epoch: 6 [25600/60000 (42.6439%)] Loss: 0.029692
Train Epoch: 6 [32000/60000 (53.3049%)] Loss: 0.025801
Train Epoch: 6 [38400/60000 (63.9659%)] Loss: 0.013343
Train Epoch: 6 [44800/60000 (74.6269%)] Loss: 0.002856
Train Epoch: 6 [51200/60000 (85.2878%)] Loss: 0.005015
Train Epoch: 6 [57600/60000 (95.9488%)] Loss: 0.191822
Test set: Average loss: 0.0484, Accuracy: 9850/10000 (98.5000%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.011267
Train Epoch: 7 [6400/60000 (10.6610%)]  Loss: 0.001690
Train Epoch: 7 [12800/60000 (21.3220%)] Loss: 0.000969
Train Epoch: 7 [19200/60000 (31.9829%)] Loss: 0.119073
Train Epoch: 7 [25600/60000 (42.6439%)] Loss: 0.021633
Train Epoch: 7 [32000/60000 (53.3049%)] Loss: 0.040279
Train Epoch: 7 [38400/60000 (63.9659%)] Loss: 0.003616
Train Epoch: 7 [44800/60000 (74.6269%)] Loss: 0.017555
Train Epoch: 7 [51200/60000 (85.2878%)] Loss: 0.003787
Train Epoch: 7 [57600/60000 (95.9488%)] Loss: 0.006645
Test set: Average loss: 0.0489, Accuracy: 9840/10000 (98.4000%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.024093
Train Epoch: 8 [6400/60000 (10.6610%)]  Loss: 0.001612
Train Epoch: 8 [12800/60000 (21.3220%)] Loss: 0.008045
Train Epoch: 8 [19200/60000 (31.9829%)] Loss: 0.183777
Train Epoch: 8 [25600/60000 (42.6439%)] Loss: 0.012900
Train Epoch: 8 [32000/60000 (53.3049%)] Loss: 0.027282
Train Epoch: 8 [38400/60000 (63.9659%)] Loss: 0.046988
Train Epoch: 8 [44800/60000 (74.6269%)] Loss: 0.008598
Train Epoch: 8 [51200/60000 (85.2878%)] Loss: 0.022854
Train Epoch: 8 [57600/60000 (95.9488%)] Loss: 0.000939
Test set: Average loss: 0.0494, Accuracy: 9852/10000 (98.5200%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.002391
Train Epoch: 9 [6400/60000 (10.6610%)]  Loss: 0.057039
Train Epoch: 9 [12800/60000 (21.3220%)] Loss: 0.010792
Train Epoch: 9 [19200/60000 (31.9829%)] Loss: 0.058978
Train Epoch: 9 [25600/60000 (42.6439%)] Loss: 0.005947
Train Epoch: 9 [32000/60000 (53.3049%)] Loss: 0.003551
Train Epoch: 9 [38400/60000 (63.9659%)] Loss: 0.024260
Train Epoch: 9 [44800/60000 (74.6269%)] Loss: 0.025892
Train Epoch: 9 [51200/60000 (85.2878%)] Loss: 0.015797
Train Epoch: 9 [57600/60000 (95.9488%)] Loss: 0.006710
Test set: Average loss: 0.0378, Accuracy: 9889/10000 (98.8900%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.000898
Train Epoch: 10 [6400/60000 (10.6610%)] Loss: 0.001666
Train Epoch: 10 [12800/60000 (21.3220%)]        Loss: 0.001526
Train Epoch: 10 [19200/60000 (31.9829%)]        Loss: 0.003377
```

```
Train Epoch: 10 [25600/60000 (42.6439%)]        Loss: 0.095161
Train Epoch: 10 [32000/60000 (53.3049%)]        Loss: 0.038337
Train Epoch: 10 [38400/60000 (63.9659%)]        Loss: 0.002940
Train Epoch: 10 [44800/60000 (74.6269%)]        Loss: 0.005821
Train Epoch: 10 [51200/60000 (85.2878%)]        Loss: 0.009800
Train Epoch: 10 [57600/60000 (95.9488%)]        Loss: 0.000916
Test set: Average loss: 0.0426, Accuracy: 9882/10000 (98.8200%)
```

Run history:

test_accuracy (Hyperparameter 3)	
test_loss (Hyperparameter 3)	
train_accuracy (Hyperparameter 3)	
train_loss (Hyperparameter 3)	

Run summary:

test_accuracy (Hyperparameter 3)	98.82
test_loss (Hyperparameter 3)	0.04262
train_accuracy (Hyperparameter 3)	99.425
train_loss (Hyperparameter 3)	0.01665

View run **skilled-lion-3** at: https://wandb.ai/hugthebee/Different-Hyperparameters/runs/lbfsv0tf
Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: `./wandb/run-20240309_134255-lbfsv0tf/logs`

```
Hyperparameter 4:
Batch Size: 64, Learning Rate: 0.001, Optimizer: Adagrad
```

Tracking run with wandb version 0.16.4
Run data is saved locally in `/content/wandb/run-20240309_134832-pemq23l3`
Syncing run **lyric-glitter-4** to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/Different-Hyperparameters
View run at https://wandb.ai/hugthebee/Different-Hyperparameters/runs/pemq23l3

```
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.312012
Train Epoch: 1 [6400/60000 (10.6610%)]  Loss: 1.302689
Train Epoch: 1 [12800/60000 (21.3220%)] Loss: 0.887086
Train Epoch: 1 [19200/60000 (31.9829%)] Loss: 0.662107
Train Epoch: 1 [25600/60000 (42.6439%)] Loss: 0.620189
Train Epoch: 1 [32000/60000 (53.3049%)] Loss: 0.522215
Train Epoch: 1 [38400/60000 (63.9659%)] Loss: 0.627649
Train Epoch: 1 [44800/60000 (74.6269%)] Loss: 0.417562
Train Epoch: 1 [51200/60000 (85.2878%)] Loss: 0.494163
Train Epoch: 1 [57600/60000 (95.9488%)] Loss: 0.398178
Test set: Average loss: 0.3932, Accuracy: 8855/10000 (88.5500%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.514870
Train Epoch: 2 [6400/60000 (10.6610%)]  Loss: 0.587752
Train Epoch: 2 [12800/60000 (21.3220%)] Loss: 0.571180
Train Epoch: 2 [19200/60000 (31.9829%)] Loss: 0.430582
Train Epoch: 2 [25600/60000 (42.6439%)] Loss: 0.402665
Train Epoch: 2 [32000/60000 (53.3049%)] Loss: 0.359512
```

```
Train Epoch: 2 [38400/60000 (63.9659%)] Loss: 0.414868
Train Epoch: 2 [44800/60000 (74.6269%)] Loss: 0.290654
Train Epoch: 2 [51200/60000 (85.2878%)] Loss: 0.292704
Train Epoch: 2 [57600/60000 (95.9488%)] Loss: 0.387819
Test set: Average loss: 0.3031, Accuracy: 9106/10000 (91.0600%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.366218
Train Epoch: 3 [6400/60000 (10.6610%)]  Loss: 0.287284
Train Epoch: 3 [12800/60000 (21.3220%)] Loss: 0.239254
Train Epoch: 3 [19200/60000 (31.9829%)] Loss: 0.220089
Train Epoch: 3 [25600/60000 (42.6439%)] Loss: 0.337820
Train Epoch: 3 [32000/60000 (53.3049%)] Loss: 0.322483
Train Epoch: 3 [38400/60000 (63.9659%)] Loss: 0.275993
Train Epoch: 3 [44800/60000 (74.6269%)] Loss: 0.232476
Train Epoch: 3 [51200/60000 (85.2878%)] Loss: 0.176605
Train Epoch: 3 [57600/60000 (95.9488%)] Loss: 0.251088
Test set: Average loss: 0.2605, Accuracy: 9219/10000 (92.1900%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.251382
Train Epoch: 4 [6400/60000 (10.6610%)]  Loss: 0.186055
Train Epoch: 4 [12800/60000 (21.3220%)] Loss: 0.267005
Train Epoch: 4 [19200/60000 (31.9829%)] Loss: 0.118399
Train Epoch: 4 [25600/60000 (42.6439%)] Loss: 0.210474
Train Epoch: 4 [32000/60000 (53.3049%)] Loss: 0.398810
Train Epoch: 4 [38400/60000 (63.9659%)] Loss: 0.208517
Train Epoch: 4 [44800/60000 (74.6269%)] Loss: 0.477640
Train Epoch: 4 [51200/60000 (85.2878%)] Loss: 0.184143
Train Epoch: 4 [57600/60000 (95.9488%)] Loss: 0.351335
Test set: Average loss: 0.2347, Accuracy: 9292/10000 (92.9200%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.191384
Train Epoch: 5 [6400/60000 (10.6610%)]  Loss: 0.090778
Train Epoch: 5 [12800/60000 (21.3220%)] Loss: 0.281273
Train Epoch: 5 [19200/60000 (31.9829%)] Loss: 0.254820
Train Epoch: 5 [25600/60000 (42.6439%)] Loss: 0.371970
Train Epoch: 5 [32000/60000 (53.3049%)] Loss: 0.332188
Train Epoch: 5 [38400/60000 (63.9659%)] Loss: 0.341892
Train Epoch: 5 [44800/60000 (74.6269%)] Loss: 0.064850
Train Epoch: 5 [51200/60000 (85.2878%)] Loss: 0.256321
Train Epoch: 5 [57600/60000 (95.9488%)] Loss: 0.129487
Test set: Average loss: 0.2165, Accuracy: 9359/10000 (93.5900%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.159148
Train Epoch: 6 [6400/60000 (10.6610%)]  Loss: 0.395626
Train Epoch: 6 [12800/60000 (21.3220%)] Loss: 0.373834
Train Epoch: 6 [19200/60000 (31.9829%)] Loss: 0.159612
Train Epoch: 6 [25600/60000 (42.6439%)] Loss: 0.404138
Train Epoch: 6 [32000/60000 (53.3049%)] Loss: 0.187978
Train Epoch: 6 [38400/60000 (63.9659%)] Loss: 0.127546
Train Epoch: 6 [44800/60000 (74.6269%)] Loss: 0.200356
Train Epoch: 6 [51200/60000 (85.2878%)] Loss: 0.291262
Train Epoch: 6 [57600/60000 (95.9488%)] Loss: 0.341603
Test set: Average loss: 0.2010, Accuracy: 9403/10000 (94.0300%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.263962
Train Epoch: 7 [6400/60000 (10.6610%)]  Loss: 0.265997
Train Epoch: 7 [12800/60000 (21.3220%)] Loss: 0.141919
Train Epoch: 7 [19200/60000 (31.9829%)] Loss: 0.258334
Train Epoch: 7 [25600/60000 (42.6439%)] Loss: 0.215106
Train Epoch: 7 [32000/60000 (53.3049%)] Loss: 0.155954
```

```
Train Epoch: 7 [32000/60000 (53.3049%)] Loss: 0.155954
Train Epoch: 7 [38400/60000 (63.9659%)] Loss: 0.250680
Train Epoch: 7 [44800/60000 (74.6269%)] Loss: 0.221973
Train Epoch: 7 [51200/60000 (85.2878%)] Loss: 0.151209
Train Epoch: 7 [57600/60000 (95.9488%)] Loss: 0.142880
Test set: Average loss: 0.1898, Accuracy: 9429/10000 (94.2900%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.228604
Train Epoch: 8 [6400/60000 (10.6610%)]  Loss: 0.155387
Train Epoch: 8 [12800/60000 (21.3220%)] Loss: 0.258969
Train Epoch: 8 [19200/60000 (31.9829%)] Loss: 0.166013
Train Epoch: 8 [25600/60000 (42.6439%)] Loss: 0.163506
Train Epoch: 8 [32000/60000 (53.3049%)] Loss: 0.290063
Train Epoch: 8 [38400/60000 (63.9659%)] Loss: 0.303441
Train Epoch: 8 [44800/60000 (74.6269%)] Loss: 0.219010
Train Epoch: 8 [51200/60000 (85.2878%)] Loss: 0.257990
Train Epoch: 8 [57600/60000 (95.9488%)] Loss: 0.162370
Test set: Average loss: 0.1793, Accuracy: 9474/10000 (94.7400%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.187321
Train Epoch: 9 [6400/60000 (10.6610%)]  Loss: 0.217433
Train Epoch: 9 [12800/60000 (21.3220%)] Loss: 0.322271
Train Epoch: 9 [19200/60000 (31.9829%)] Loss: 0.189322
Train Epoch: 9 [25600/60000 (42.6439%)] Loss: 0.252198
Train Epoch: 9 [32000/60000 (53.3049%)] Loss: 0.459640
Train Epoch: 9 [38400/60000 (63.9659%)] Loss: 0.295391
Train Epoch: 9 [44800/60000 (74.6269%)] Loss: 0.221062
Train Epoch: 9 [51200/60000 (85.2878%)] Loss: 0.136723
Train Epoch: 9 [57600/60000 (95.9488%)] Loss: 0.367230
Test set: Average loss: 0.1716, Accuracy: 9507/10000 (95.0700%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.280567
Train Epoch: 10 [6400/60000 (10.6610%)] Loss: 0.252709
Train Epoch: 10 [12800/60000 (21.3220%)]        Loss: 0.129617
Train Epoch: 10 [19200/60000 (31.9829%)]        Loss: 0.210381
Train Epoch: 10 [25600/60000 (42.6439%)]        Loss: 0.136758
Train Epoch: 10 [32000/60000 (53.3049%)]        Loss: 0.139013
Train Epoch: 10 [38400/60000 (63.9659%)]        Loss: 0.165260
Train Epoch: 10 [44800/60000 (74.6269%)]        Loss: 0.282208
Train Epoch: 10 [51200/60000 (85.2878%)]        Loss: 0.169655
Train Epoch: 10 [57600/60000 (95.9488%)]        Loss: 0.091123
Test set: Average loss: 0.1642, Accuracy: 9527/10000 (95.2700%)
```

Run history:

test_accuracy (Hyperparameter 4)	
test_loss (Hyperparameter 4)	
train_accuracy (Hyperparameter 4)	
train_loss (Hyperparameter	

Run summary:

test_accuracy (Hyperparameter 4)	95.27
test_loss (Hyperparameter 4)	0.16425
train_accuracy (Hyperparameter 4)	94.615
train_loss (Hyperparameter	0.18302

4) 4)

View run **lyric-glitter-4** at: https://wandb.ai/hugthebee/Different-Hyperparameters/runs/pemq23l3
Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: `./wandb/run-20240309_134832-pemq23l3/logs`

```
Hyperparameter 5:
Batch Size: 64, Learning Rate: 0.01, Optimizer: SGD
```
Tracking run with wandb version 0.16.4
Run data is saved locally in `/content/wandb/run-20240309_135406-k80lh6va`
Syncing run **true-feather-5** to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/Different-Hyperparameters
View run at https://wandb.ai/hugthebee/Different-Hyperparameters/runs/k80lh6va
```
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.302463
Train Epoch: 1 [6400/60000 (10.6610%)]  Loss: 2.188815
Train Epoch: 1 [12800/60000 (21.3220%)] Loss: 0.371617
Train Epoch: 1 [19200/60000 (31.9829%)] Loss: 0.214720
Train Epoch: 1 [25600/60000 (42.6439%)] Loss: 0.142778
Train Epoch: 1 [32000/60000 (53.3049%)] Loss: 0.134079
Train Epoch: 1 [38400/60000 (63.9659%)] Loss: 0.172850
Train Epoch: 1 [44800/60000 (74.6269%)] Loss: 0.069229
Train Epoch: 1 [51200/60000 (85.2878%)] Loss: 0.040183
Train Epoch: 1 [57600/60000 (95.9488%)] Loss: 0.144463
Test set: Average loss: 0.1040, Accuracy: 9663/10000 (96.6300%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.104944
Train Epoch: 2 [6400/60000 (10.6610%)]  Loss: 0.049691
Train Epoch: 2 [12800/60000 (21.3220%)] Loss: 0.055483
Train Epoch: 2 [19200/60000 (31.9829%)] Loss: 0.026544
Train Epoch: 2 [25600/60000 (42.6439%)] Loss: 0.024513
Train Epoch: 2 [32000/60000 (53.3049%)] Loss: 0.040119
Train Epoch: 2 [38400/60000 (63.9659%)] Loss: 0.071951
Train Epoch: 2 [44800/60000 (74.6269%)] Loss: 0.199626
Train Epoch: 2 [51200/60000 (85.2878%)] Loss: 0.039430
Train Epoch: 2 [57600/60000 (95.9488%)] Loss: 0.087419
Test set: Average loss: 0.0548, Accuracy: 9791/10000 (97.9100%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.059755
Train Epoch: 3 [6400/60000 (10.6610%)]  Loss: 0.041116
Train Epoch: 3 [12800/60000 (21.3220%)] Loss: 0.020349
Train Epoch: 3 [19200/60000 (31.9829%)] Loss: 0.008571
Train Epoch: 3 [25600/60000 (42.6439%)] Loss: 0.015188
Train Epoch: 3 [32000/60000 (53.3049%)] Loss: 0.027412
Train Epoch: 3 [38400/60000 (63.9659%)] Loss: 0.066923
Train Epoch: 3 [44800/60000 (74.6269%)] Loss: 0.048441
Train Epoch: 3 [51200/60000 (85.2878%)] Loss: 0.045381
Train Epoch: 3 [57600/60000 (95.9488%)] Loss: 0.063744
Test set: Average loss: 0.0441, Accuracy: 9854/10000 (98.5400%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.038748
Train Epoch: 4 [6400/60000 (10.6610%)]  Loss: 0.061719
Train Epoch: 4 [12800/60000 (21.3220%)] Loss: 0.072847
Train Epoch: 4 [19200/60000 (31.9829%)] Loss: 0.032919
Train Epoch: 4 [25600/60000 (42.6439%)] Loss: 0.033705
Train Epoch: 4 [32000/60000 (53.3049%)] Loss: 0.005512
Train Epoch: 4 [38400/60000 (63.9659%)] Loss: 0.040866
```

```
Train Epoch: 4 [44800/60000 (74.6269%)] Loss: 0.009096
Train Epoch: 4 [51200/60000 (85.2878%)] Loss: 0.011461
Train Epoch: 4 [57600/60000 (95.9488%)] Loss: 0.014361
Test set: Average loss: 0.0433, Accuracy: 9858/10000 (98.5800%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.000402
Train Epoch: 5 [6400/60000 (10.6610%)]  Loss: 0.030212
Train Epoch: 5 [12800/60000 (21.3220%)] Loss: 0.006711
Train Epoch: 5 [19200/60000 (31.9829%)] Loss: 0.003461
Train Epoch: 5 [25600/60000 (42.6439%)] Loss: 0.042190
Train Epoch: 5 [32000/60000 (53.3049%)] Loss: 0.033733
Train Epoch: 5 [38400/60000 (63.9659%)] Loss: 0.013134
Train Epoch: 5 [44800/60000 (74.6269%)] Loss: 0.003853
Train Epoch: 5 [51200/60000 (85.2878%)] Loss: 0.017581
Train Epoch: 5 [57600/60000 (95.9488%)] Loss: 0.065962
Test set: Average loss: 0.0349, Accuracy: 9892/10000 (98.9200%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.021198
Train Epoch: 6 [6400/60000 (10.6610%)]  Loss: 0.007474
Train Epoch: 6 [12800/60000 (21.3220%)] Loss: 0.003723
Train Epoch: 6 [19200/60000 (31.9829%)] Loss: 0.025507
Train Epoch: 6 [25600/60000 (42.6439%)] Loss: 0.151857
Train Epoch: 6 [32000/60000 (53.3049%)] Loss: 0.033481
Train Epoch: 6 [38400/60000 (63.9659%)] Loss: 0.005507
Train Epoch: 6 [44800/60000 (74.6269%)] Loss: 0.036237
Train Epoch: 6 [51200/60000 (85.2878%)] Loss: 0.036004
Train Epoch: 6 [57600/60000 (95.9488%)] Loss: 0.001642
Test set: Average loss: 0.0395, Accuracy: 9862/10000 (98.6200%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.027222
Train Epoch: 7 [6400/60000 (10.6610%)]  Loss: 0.007729
Train Epoch: 7 [12800/60000 (21.3220%)] Loss: 0.012973
Train Epoch: 7 [19200/60000 (31.9829%)] Loss: 0.076071
Train Epoch: 7 [25600/60000 (42.6439%)] Loss: 0.017041
Train Epoch: 7 [32000/60000 (53.3049%)] Loss: 0.018818
Train Epoch: 7 [38400/60000 (63.9659%)] Loss: 0.080611
Train Epoch: 7 [44800/60000 (74.6269%)] Loss: 0.003862
Train Epoch: 7 [51200/60000 (85.2878%)] Loss: 0.003937
Train Epoch: 7 [57600/60000 (95.9488%)] Loss: 0.101976
Test set: Average loss: 0.0419, Accuracy: 9868/10000 (98.6800%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.093077
Train Epoch: 8 [6400/60000 (10.6610%)]  Loss: 0.006788
Train Epoch: 8 [12800/60000 (21.3220%)] Loss: 0.000362
Train Epoch: 8 [19200/60000 (31.9829%)] Loss: 0.021969
Train Epoch: 8 [25600/60000 (42.6439%)] Loss: 0.126466
Train Epoch: 8 [32000/60000 (53.3049%)] Loss: 0.004342
Train Epoch: 8 [38400/60000 (63.9659%)] Loss: 0.088554
Train Epoch: 8 [44800/60000 (74.6269%)] Loss: 0.015764
Train Epoch: 8 [51200/60000 (85.2878%)] Loss: 0.009677
Train Epoch: 8 [57600/60000 (95.9488%)] Loss: 0.020693
Test set: Average loss: 0.0351, Accuracy: 9889/10000 (98.8900%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.008792
Train Epoch: 9 [6400/60000 (10.6610%)]  Loss: 0.001077
Train Epoch: 9 [12800/60000 (21.3220%)] Loss: 0.025404
Train Epoch: 9 [19200/60000 (31.9829%)] Loss: 0.000116
Train Epoch: 9 [25600/60000 (42.6439%)] Loss: 0.010433
Train Epoch: 9 [32000/60000 (53.3049%)] Loss: 0.001820
```

```
Train Epoch: 9 [38400/60000 (63.9659%)] Loss: 0.011097
Train Epoch: 9 [44800/60000 (74.6269%)] Loss: 0.002043
Train Epoch: 9 [51200/60000 (85.2878%)] Loss: 0.018089
Train Epoch: 9 [57600/60000 (95.9488%)] Loss: 0.010201
Test set: Average loss: 0.0379, Accuracy: 9878/10000 (98.7800%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.000635
Train Epoch: 10 [6400/60000 (10.6610%)] Loss: 0.002264
Train Epoch: 10 [12800/60000 (21.3220%)]        Loss: 0.005457
Train Epoch: 10 [19200/60000 (31.9829%)]        Loss: 0.015196
Train Epoch: 10 [25600/60000 (42.6439%)]        Loss: 0.005163
Train Epoch: 10 [32000/60000 (53.3049%)]        Loss: 0.000924
Train Epoch: 10 [38400/60000 (63.9659%)]        Loss: 0.004536
Train Epoch: 10 [44800/60000 (74.6269%)]        Loss: 0.009195
Train Epoch: 10 [51200/60000 (85.2878%)]        Loss: 0.000890
Train Epoch: 10 [57600/60000 (95.9488%)]        Loss: 0.065868
Test set: Average loss: 0.0359, Accuracy: 9894/10000 (98.9400%)
```

Run history:

test_accuracy (Hyperparameter 5)	
test_loss (Hyperparameter 5)	
train_accuracy (Hyperparameter 5)	
train_loss (Hyperparameter 5)	

Run summary:

test_accuracy (Hyperparameter 5)	98.94
test_loss (Hyperparameter 5)	0.03589
train_accuracy (Hyperparameter 5)	99.53667
train_loss (Hyperparameter 5)	0.01463

View run **true-feather-5** at: https://wandb.ai/hugthebee/Different-Hyperparameters/runs/k80lh6va
Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: `./wandb/run-20240309_135406-k80lh6va/logs`

```
Hyperparameter 6:
Batch Size: 64, Learning Rate: 0.1, Optimizer: SGD
```

Tracking run with wandb version 0.16.4
Run data is saved locally in `/content/wandb/run-20240309_135936-38ujkdhw`
Syncing run **floral-star-6** to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/Different-Hyperparameters
View run at https://wandb.ai/hugthebee/Different-Hyperparameters/runs/38ujkdhw

```
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.299047
Train Epoch: 1 [6400/60000 (10.6610%)]  Loss: 0.364706
Train Epoch: 1 [12800/60000 (21.3220%)] Loss: 0.428294
Train Epoch: 1 [19200/60000 (31.9829%)] Loss: 0.343822
Train Epoch: 1 [25600/60000 (42.6439%)] Loss: 0.388516
Train Epoch: 1 [32000/60000 (53.3049%)] Loss: 0.214498
Train Epoch: 1 [38400/60000 (63.9659%)] Loss: 0.150879
Train Epoch: 1 [44800/60000 (74.6269%)] Loss: 0.173395
Train Epoch: 1 [51200/60000 (85.2878%)] Loss: 0.208932
```

```
Train Epoch: 1 [57600/60000 (95.9488%)] Loss: 0.165797
Test set: Average loss: 0.1809, Accuracy: 9514/10000 (95.1400%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.117268
Train Epoch: 2 [6400/60000 (10.6610%)]  Loss: 0.098112
Train Epoch: 2 [12800/60000 (21.3220%)] Loss: 0.057890
Train Epoch: 2 [19200/60000 (31.9829%)] Loss: 0.125744
Train Epoch: 2 [25600/60000 (42.6439%)] Loss: 0.059676
Train Epoch: 2 [32000/60000 (53.3049%)] Loss: 0.258069
Train Epoch: 2 [38400/60000 (63.9659%)] Loss: 0.292353
Train Epoch: 2 [44800/60000 (74.6269%)] Loss: 0.309943
Train Epoch: 2 [51200/60000 (85.2878%)] Loss: 0.123867
Train Epoch: 2 [57600/60000 (95.9488%)] Loss: 0.198690
Test set: Average loss: 0.2881, Accuracy: 9510/10000 (95.1000%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.208937
Train Epoch: 3 [6400/60000 (10.6610%)]  Loss: 0.295268
Train Epoch: 3 [12800/60000 (21.3220%)] Loss: 0.161653
Train Epoch: 3 [19200/60000 (31.9829%)] Loss: 0.066979
Train Epoch: 3 [25600/60000 (42.6439%)] Loss: 0.169115
Train Epoch: 3 [32000/60000 (53.3049%)] Loss: 0.613825
Train Epoch: 3 [38400/60000 (63.9659%)] Loss: 2.172187
Train Epoch: 3 [44800/60000 (74.6269%)] Loss: 2.281058
Train Epoch: 3 [51200/60000 (85.2878%)] Loss: 2.316178
Train Epoch: 3 [57600/60000 (95.9488%)] Loss: 2.309067
Test set: Average loss: 2.3036, Accuracy: 1028/10000 (10.2800%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 2.290006
Train Epoch: 4 [6400/60000 (10.6610%)]  Loss: 2.321803
Train Epoch: 4 [12800/60000 (21.3220%)] Loss: 2.323107
Train Epoch: 4 [19200/60000 (31.9829%)] Loss: 2.309690
Train Epoch: 4 [25600/60000 (42.6439%)] Loss: 2.312042
Train Epoch: 4 [32000/60000 (53.3049%)] Loss: 2.291984
Train Epoch: 4 [38400/60000 (63.9659%)] Loss: 2.299426
Train Epoch: 4 [44800/60000 (74.6269%)] Loss: 2.297507
Train Epoch: 4 [51200/60000 (85.2878%)] Loss: 2.304507
Train Epoch: 4 [57600/60000 (95.9488%)] Loss: 2.298981
Test set: Average loss: 2.3028, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 2.299208
Train Epoch: 5 [6400/60000 (10.6610%)]  Loss: 2.301285
Train Epoch: 5 [12800/60000 (21.3220%)] Loss: 2.304439
Train Epoch: 5 [19200/60000 (31.9829%)] Loss: 2.309360
Train Epoch: 5 [25600/60000 (42.6439%)] Loss: 2.295343
Train Epoch: 5 [32000/60000 (53.3049%)] Loss: 2.310885
Train Epoch: 5 [38400/60000 (63.9659%)] Loss: 2.293403
Train Epoch: 5 [44800/60000 (74.6269%)] Loss: 2.312170
Train Epoch: 5 [51200/60000 (85.2878%)] Loss: 2.286355
Train Epoch: 5 [57600/60000 (95.9488%)] Loss: 2.293085
Test set: Average loss: 2.3042, Accuracy: 1028/10000 (10.2800%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 2.308599
Train Epoch: 6 [6400/60000 (10.6610%)]  Loss: 2.312287
Train Epoch: 6 [12800/60000 (21.3220%)] Loss: 2.308493
Train Epoch: 6 [19200/60000 (31.9829%)] Loss: 2.304898
Train Epoch: 6 [25600/60000 (42.6439%)] Loss: 2.301039
Train Epoch: 6 [32000/60000 (53.3049%)] Loss: 2.298284
Train Epoch: 6 [38400/60000 (63.9659%)] Loss: 2.310901
Train Epoch: 6 [44800/60000 (74.6269%)] Loss: 2.287524
```

```
Train Epoch: 6 [51200/60000 (85.2878%)] Loss: 2.298461
Train Epoch: 6 [57600/60000 (95.9488%)] Loss: 2.309213
Test set: Average loss: 2.3039, Accuracy: 1010/10000 (10.1000%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 2.301073
Train Epoch: 7 [6400/60000 (10.6610%)]  Loss: 2.297749
Train Epoch: 7 [12800/60000 (21.3220%)] Loss: 2.317447
Train Epoch: 7 [19200/60000 (31.9829%)] Loss: 2.309944
Train Epoch: 7 [25600/60000 (42.6439%)] Loss: 2.318542
Train Epoch: 7 [32000/60000 (53.3049%)] Loss: 2.310401
Train Epoch: 7 [38400/60000 (63.9659%)] Loss: 2.308295
Train Epoch: 7 [44800/60000 (74.6269%)] Loss: 2.318976
Train Epoch: 7 [51200/60000 (85.2878%)] Loss: 2.321462
Train Epoch: 7 [57600/60000 (95.9488%)] Loss: 2.291306
Test set: Average loss: 2.3027, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 2.304092
Train Epoch: 8 [6400/60000 (10.6610%)]  Loss: 2.311695
Train Epoch: 8 [12800/60000 (21.3220%)] Loss: 2.291932
Train Epoch: 8 [19200/60000 (31.9829%)] Loss: 2.299494
Train Epoch: 8 [25600/60000 (42.6439%)] Loss: 2.287552
Train Epoch: 8 [32000/60000 (53.3049%)] Loss: 2.296731
Train Epoch: 8 [38400/60000 (63.9659%)] Loss: 2.299324
Train Epoch: 8 [44800/60000 (74.6269%)] Loss: 2.304566
Train Epoch: 8 [51200/60000 (85.2878%)] Loss: 2.292848
Train Epoch: 8 [57600/60000 (95.9488%)] Loss: 2.316478
Test set: Average loss: 2.3052, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 2.331720
Train Epoch: 9 [6400/60000 (10.6610%)]  Loss: 2.312703
Train Epoch: 9 [12800/60000 (21.3220%)] Loss: 2.305795
Train Epoch: 9 [19200/60000 (31.9829%)] Loss: 2.291967
Train Epoch: 9 [25600/60000 (42.6439%)] Loss: 2.307033
Train Epoch: 9 [32000/60000 (53.3049%)] Loss: 2.297285
Train Epoch: 9 [38400/60000 (63.9659%)] Loss: 2.284554
Train Epoch: 9 [44800/60000 (74.6269%)] Loss: 2.292360
Train Epoch: 9 [51200/60000 (85.2878%)] Loss: 2.299354
Train Epoch: 9 [57600/60000 (95.9488%)] Loss: 2.301579
Test set: Average loss: 2.3038, Accuracy: 1032/10000 (10.3200%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 2.298898
Train Epoch: 10 [6400/60000 (10.6610%)] Loss: 2.311862
Train Epoch: 10 [12800/60000 (21.3220%)]        Loss: 2.312367
Train Epoch: 10 [19200/60000 (31.9829%)]        Loss: 2.279628
Train Epoch: 10 [25600/60000 (42.6439%)]        Loss: 2.298239
Train Epoch: 10 [32000/60000 (53.3049%)]        Loss: 2.303918
Train Epoch: 10 [38400/60000 (63.9659%)]        Loss: 2.281651
Train Epoch: 10 [44800/60000 (74.6269%)]        Loss: 2.310203
Train Epoch: 10 [51200/60000 (85.2878%)]        Loss: 2.311241
Train Epoch: 10 [57600/60000 (95.9488%)]        Loss: 2.319961
Test set: Average loss: 2.3051, Accuracy: 1135/10000 (11.3500%)
```

Run history:

test_accuracy (Hyperparameter 6) ▁▁▁▁▁▁▁▁▁▁

Run summary:

test_accuracy (Hyperparameter 6) 11.35

test_loss (Hyperparameter 6)		test_loss (Hyperparameter 6) 2.3051
train_accuracy (Hyperparameter 6)		train_accuracy (Hyperparameter 6) 10.72
train_loss (Hyperparameter 6)		train_loss (Hyperparameter 6) 2.30478

View run **floral-star-6** at: https://wandb.ai/hugthebee/Different-Hyperparameters/runs/38ujkdhw
Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: ./wandb/run-20240309_135936-38ujkdhw/logs

Both the LeNet CNN and SIFT-BoVW perform reasonably well for the task of digit recognition. CNN's however are generally more preferred because they capture important image elements faster and more easily, when compared to SIFT.

The problem with the SIFT-BoVW approach is that the performance depends a lot on the cluster size we use, and an increase in cluster size also causes an increase in computation time and efficiency. This is another area where the CNN performs better.

Though the CNN takes some time to train, once trained it performs better and faster than BoVW.

Double # of convolutional layers

Intuitively, doubling the number of convolutional layers should increase the accuracy of a CNN. This is because convolutional layers capture the features of the image, and if there are more of these layers, we will be able to train the model for finer details.

However, there is a tradeoff to consider because it could also lead to overfitting and also significantly increases the computatinal cost. Further, if the dataset isn't too complex (like in out case), having more number of layers is just an overkill.

Previous Accuracy = 98.2%
Accuracy (x2 convolutional layers) = 93.32%

```
class LeNet(nn.Module):
    def __init__(self):
        super(LeNet, self).__init__()
        self.conv1 = nn.Conv2d(1, 6, 5)
        self.conv2 = nn.Conv2d(6, 16, 5)
        self.conv3 = nn.Conv2d(16, 32, 5)
        self.conv4 = nn.Conv2d(32, 64, 5)
        self.pool = nn.MaxPool2d(2, 2)
        self.fc1 = nn.Linear(16*4*4, 120)
        self.fc2 = nn.Linear(120, 84)
        self.fc3 = nn.Linear(84, 10)
        self.pad = nn.ZeroPad2d(2)
        self.pad1 = nn.ZeroPad2d(1)

    def forward(self, x):
        x = torch.relu(self.conv1(x))
        x = self.pool(x)
        # print(x.size())
        x = torch.relu(self.conv2(x))
        x = self.pool(x)
        x = self.pad(x)
        x = torch.relu(self.conv3(x))
        x = self.pool(x)
        x = self.pad(x)
        x = torch.relu(self.conv4(x))
        x = x.view(-1, 16*4*4)
        x = torch.relu(self.fc1(x))
        x = torch.relu(self.fc2(x))
        x = self.fc3(x)
        return x


model = LeNet()
loss_fn = nn.CrossEntropyLoss()
optimizer = optim.SGD(model.parameters(), lr=0.001, momentum=0.9)

wandb.init(project='Double_conv_layers')

# Training loop
epochs = 10
for epoch in range(1, epochs + 1):
    train_loss, train_accuracy = train(model, train_loader, loss_fn, optimizer,
    test_loss, test_accuracy = test(model, test_loader, loss_fn)

    wandb.log({
        "epoch": epoch,
        "train_loss": train_loss,
        "train_accuracy": train_accuracy,
```

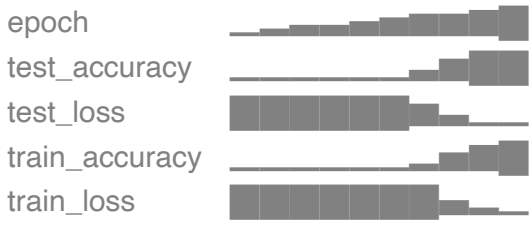
```
        "test_loss": test_loss,
        "test_accuracy": test_accuracy
    })

wandb.finish()
```

```
Tracking run with wandb version 0.16.4
Run data is saved locally in /content/wandb/run-20240309_140935-fpjgk89y
Syncing run rare-frost-1 to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/Double_conv_layers
View run at https://wandb.ai/hugthebee/Double_conv_layers/runs/fpjgk89y
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.300234
Train Epoch: 1 [6400/60000 (10.6610%)]  Loss: 2.295676
Train Epoch: 1 [12800/60000 (21.3220%)] Loss: 2.311294
Train Epoch: 1 [19200/60000 (31.9829%)] Loss: 2.303059
Train Epoch: 1 [25600/60000 (42.6439%)] Loss: 2.295641
Train Epoch: 1 [32000/60000 (53.3049%)] Loss: 2.301938
Train Epoch: 1 [38400/60000 (63.9659%)] Loss: 2.293066
Train Epoch: 1 [44800/60000 (74.6269%)] Loss: 2.298978
Train Epoch: 1 [51200/60000 (85.2878%)] Loss: 2.297978
Train Epoch: 1 [57600/60000 (95.9488%)] Loss: 2.298210
Test set: Average loss: 2.3011, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 2.306732
Train Epoch: 2 [6400/60000 (10.6610%)]  Loss: 2.300900
Train Epoch: 2 [12800/60000 (21.3220%)] Loss: 2.307973
Train Epoch: 2 [19200/60000 (31.9829%)] Loss: 2.316901
Train Epoch: 2 [25600/60000 (42.6439%)] Loss: 2.294209
Train Epoch: 2 [32000/60000 (53.3049%)] Loss: 2.301602
Train Epoch: 2 [38400/60000 (63.9659%)] Loss: 2.295606
Train Epoch: 2 [44800/60000 (74.6269%)] Loss: 2.308000
Train Epoch: 2 [51200/60000 (85.2878%)] Loss: 2.301825
Train Epoch: 2 [57600/60000 (95.9488%)] Loss: 2.298784
Test set: Average loss: 2.3005, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 2.297160
Train Epoch: 3 [6400/60000 (10.6610%)]  Loss: 2.305074
Train Epoch: 3 [12800/60000 (21.3220%)] Loss: 2.311946
Train Epoch: 3 [19200/60000 (31.9829%)] Loss: 2.297770
Train Epoch: 3 [25600/60000 (42.6439%)] Loss: 2.301685
Train Epoch: 3 [32000/60000 (53.3049%)] Loss: 2.281498
Train Epoch: 3 [38400/60000 (63.9659%)] Loss: 2.304216
Train Epoch: 3 [44800/60000 (74.6269%)] Loss: 2.301247
Train Epoch: 3 [51200/60000 (85.2878%)] Loss: 2.295618
Train Epoch: 3 [57600/60000 (95.9488%)] Loss: 2.312664
Test set: Average loss: 2.3002, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 2.304865
Train Epoch: 4 [6400/60000 (10.6610%)]  Loss: 2.302162
Train Epoch: 4 [12800/60000 (21.3220%)] Loss: 2.304982
Train Epoch: 4 [19200/60000 (31.9829%)] Loss: 2.301778
Train Epoch: 4 [25600/60000 (42.6439%)] Loss: 2.298467
Train Epoch: 4 [32000/60000 (53.3049%)] Loss: 2.296018
Train Epoch: 4 [38400/60000 (63.9659%)] Loss: 2.301852
Train Epoch: 4 [44800/60000 (74.6269%)] Loss: 2.293951
Train Epoch: 4 [51200/60000 (85.2878%)] Loss: 2.301214
```

```
Train Epoch: 4 [57600/60000 (95.9488%)] Loss: 2.301291
Test set: Average loss: 2.2995, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 2.294071
Train Epoch: 5 [6400/60000 (10.6610%)]  Loss: 2.295516
Train Epoch: 5 [12800/60000 (21.3220%)] Loss: 2.300949
Train Epoch: 5 [19200/60000 (31.9829%)] Loss: 2.306619
Train Epoch: 5 [25600/60000 (42.6439%)] Loss: 2.298256
Train Epoch: 5 [32000/60000 (53.3049%)] Loss: 2.292068
Train Epoch: 5 [38400/60000 (63.9659%)] Loss: 2.300461
Train Epoch: 5 [44800/60000 (74.6269%)] Loss: 2.301924
Train Epoch: 5 [51200/60000 (85.2878%)] Loss: 2.300368
Train Epoch: 5 [57600/60000 (95.9488%)] Loss: 2.300587
Test set: Average loss: 2.2978, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 2.299306
Train Epoch: 6 [6400/60000 (10.6610%)]  Loss: 2.298756
Train Epoch: 6 [12800/60000 (21.3220%)] Loss: 2.295475
Train Epoch: 6 [19200/60000 (31.9829%)] Loss: 2.301492
Train Epoch: 6 [25600/60000 (42.6439%)] Loss: 2.306556
Train Epoch: 6 [32000/60000 (53.3049%)] Loss: 2.288894
Train Epoch: 6 [38400/60000 (63.9659%)] Loss: 2.287434
Train Epoch: 6 [44800/60000 (74.6269%)] Loss: 2.289429
Train Epoch: 6 [51200/60000 (85.2878%)] Loss: 2.304688
Train Epoch: 6 [57600/60000 (95.9488%)] Loss: 2.280239
Test set: Average loss: 2.2888, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 2.294541
Train Epoch: 7 [6400/60000 (10.6610%)]  Loss: 2.300433
Train Epoch: 7 [12800/60000 (21.3220%)] Loss: 2.289636
Train Epoch: 7 [19200/60000 (31.9829%)] Loss: 2.274356
Train Epoch: 7 [25600/60000 (42.6439%)] Loss: 2.250561
Train Epoch: 7 [32000/60000 (53.3049%)] Loss: 2.248251
Train Epoch: 7 [38400/60000 (63.9659%)] Loss: 2.154402
Train Epoch: 7 [44800/60000 (74.6269%)] Loss: 2.113749
Train Epoch: 7 [51200/60000 (85.2878%)] Loss: 2.036591
Train Epoch: 7 [57600/60000 (95.9488%)] Loss: 1.879172
Test set: Average loss: 1.7801, Accuracy: 3591/10000 (35.9100%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 1.851516
Train Epoch: 8 [6400/60000 (10.6610%)]  Loss: 1.603719
Train Epoch: 8 [12800/60000 (21.3220%)] Loss: 1.352755
Train Epoch: 8 [19200/60000 (31.9829%)] Loss: 1.231416
Train Epoch: 8 [25600/60000 (42.6439%)] Loss: 1.099013
Train Epoch: 8 [32000/60000 (53.3049%)] Loss: 0.893448
Train Epoch: 8 [38400/60000 (63.9659%)] Loss: 1.050586
Train Epoch: 8 [44800/60000 (74.6269%)] Loss: 0.758809
Train Epoch: 8 [51200/60000 (85.2878%)] Loss: 0.847886
Train Epoch: 8 [57600/60000 (95.9488%)] Loss: 0.474293
Test set: Average loss: 0.6969, Accuracy: 7482/10000 (74.8200%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 1.040618
Train Epoch: 9 [6400/60000 (10.6610%)]  Loss: 0.488202
Train Epoch: 9 [12800/60000 (21.3220%)] Loss: 0.358873
Train Epoch: 9 [19200/60000 (31.9829%)] Loss: 0.387202
Train Epoch: 9 [25600/60000 (42.6439%)] Loss: 0.507746
Train Epoch: 9 [32000/60000 (53.3049%)] Loss: 0.445096
Train Epoch: 9 [38400/60000 (63.9659%)] Loss: 0.326384
Train Epoch: 9 [44800/60000 (74.6269%)] Loss: 0.341914
```

```
Train Epoch: 9 [51200/60000 (85.2878%)] Loss: 0.250757
Train Epoch: 9 [57600/60000 (95.9488%)] Loss: 0.354615
Test set: Average loss: 0.2859, Accuracy: 9087/10000 (90.8700%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.236997
Train Epoch: 10 [6400/60000 (10.6610%)] Loss: 0.391291
Train Epoch: 10 [12800/60000 (21.3220%)]        Loss: 0.312363
Train Epoch: 10 [19200/60000 (31.9829%)]        Loss: 0.138591
Train Epoch: 10 [25600/60000 (42.6439%)]        Loss: 0.372969
Train Epoch: 10 [32000/60000 (53.3049%)]        Loss: 0.144708
Train Epoch: 10 [38400/60000 (63.9659%)]        Loss: 0.171020
Train Epoch: 10 [44800/60000 (74.6269%)]        Loss: 0.108838
Train Epoch: 10 [51200/60000 (85.2878%)]        Loss: 0.221497
Train Epoch: 10 [57600/60000 (95.9488%)]        Loss: 0.079562
Test set: Average loss: 0.1582, Accuracy: 9521/10000 (95.2100%)
```

Run history:



Run summary:

epoch	10
test_accuracy	95.21
test_loss	0.15822
train_accuracy	93.32333
train_loss	0.21827

View run **rare-frost-1** at: https://wandb.ai/hugthebee/Double_conv_layers/runs/fpjgk89y
Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: `./wandb/run-20240309_140935-fpjgk89y/logs`

Double sizes of input data

To ensure a fair data input, we make sure that every size of input data has all the classes represented equally.

1. Input size = 600, Accuracy = 18.33%
2. Input size = 1800, Accuracy = 46.32%
3. Input size = 6000, Accuracy = 97.42%
4. Input size = 18000, Accuracy = 98.42%
5. Input size = 60000, Accuracy = 98.95%

Thus, as is clear from the graphs and results obtained, increasing the input data size significantly increases the output.

Having very little input data means the model will not be able to extract all of the important features that would be required for classification. Larger data size provides more diverse samples for the model to learn from, which can lead to better performance on new, unseen data.

```
classes = torch.bincount(train_dataset.targets)

train_sizes = [600, 1800, 6000, 18000, 60000]
model = LeNet()
loss_fn = nn.CrossEntropyLoss()
optimizer = optim.SGD(model.parameters(), lr=0.001, momentum=0.9)

wandb.init(project='Different-traindata-size')
for size in train_sizes:
    print(f"\nTraining with {size} samples:")

    samples_per_class = size // 10  # 10 classes in MNIST dataset
    selected_indices = []

    # Iterate over each class
    for i in range(10):
        class_indices = (train_dataset.targets == i).nonzero(as_tuple=True)[0]

        # randomise
        class_indices = class_indices[torch.randperm(len(class_indices))]
        selected_indices.extend(class_indices[:samples_per_class].tolist())
```

```
    subset_sampler = torch.utils.data.SubsetRandomSampler(selected_indices)
    train_loader = torch.utils.data.DataLoader(train_dataset, batch_size=4, sam

    # Training loop
    epochs = 10
    for epoch in range(1, epochs + 1):
      train_loss, train_accuracy = train(model, train_loader, loss_fn, optimize
      test_loss, test_accuracy = test(model, test_loader, loss_fn)

      wandb.log({
          "epoch": epoch,
          f"train_loss (size:{size})": train_loss,
          f"train_accuracy (size:{size})": train_accuracy,
          f"test_loss (size:{size})": test_loss,
          f"test_accuracy (size:{size})": test_accuracy
      })

wandb.finish()
```

```
Tracking run with wandb version 0.16.4
Run data is saved locally in /content/wandb/run-20240309_141705-sicmbor4
Syncing run sparkling-disco-1 to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/Different-traindata-size
View run at https://wandb.ai/hugthebee/Different-traindata-size/runs/sicmbor4

Training with 600 samples:
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.298886
Train Epoch: 1 [400/60000 (66.6667%)]   Loss: 2.249553
Test set: Average loss: 2.3032, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 2.285604
Train Epoch: 2 [400/60000 (66.6667%)]   Loss: 2.280403
Test set: Average loss: 2.3027, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 2.330031
Train Epoch: 3 [400/60000 (66.6667%)]   Loss: 2.310724
Test set: Average loss: 2.3026, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 2.259864
Train Epoch: 4 [400/60000 (66.6667%)]   Loss: 2.304604
Test set: Average loss: 2.3024, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 2.329255
Train Epoch: 5 [400/60000 (66.6667%)]   Loss: 2.312717
Test set: Average loss: 2.3023, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 2.309626
Train Epoch: 6 [400/60000 (66.6667%)]   Loss: 2.293909
Test set: Average loss: 2.3022, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 2.296906
Train Epoch: 7 [400/60000 (66.6667%)]   Loss: 2.296407
Test set: Average loss: 2.3022, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 2.292102
Train Epoch: 8 [400/60000 (66.6667%)]   Loss: 2.305664
Test set: Average loss: 2.3022, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 2.298430
Train Epoch: 9 [400/60000 (66.6667%)]   Loss: 2.297713
```

```
Test set: Average loss: 2.3020, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 2.300594
Train Epoch: 10 [400/60000 (66.6667%)]  Loss: 2.312999
Test set: Average loss: 2.3021, Accuracy: 1833/10000 (18.3300%)

Training with 1800 samples:
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.303197
Train Epoch: 1 [400/60000 (22.2222%)]   Loss: 2.299713
Train Epoch: 1 [800/60000 (44.4444%)]   Loss: 2.318676
Train Epoch: 1 [1200/60000 (66.6667%)]  Loss: 2.302165
Train Epoch: 1 [1600/60000 (88.8889%)]  Loss: 2.312472
Test set: Average loss: 2.3020, Accuracy: 892/10000 (8.9200%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 2.306256
Train Epoch: 2 [400/60000 (22.2222%)]   Loss: 2.309534
Train Epoch: 2 [800/60000 (44.4444%)]   Loss: 2.308643
Train Epoch: 2 [1200/60000 (66.6667%)]  Loss: 2.297830
Train Epoch: 2 [1600/60000 (88.8889%)]  Loss: 2.301830
Test set: Average loss: 2.3017, Accuracy: 1235/10000 (12.3500%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 2.306054
Train Epoch: 3 [400/60000 (22.2222%)]   Loss: 2.306410
Train Epoch: 3 [800/60000 (44.4444%)]   Loss: 2.305359
Train Epoch: 3 [1200/60000 (66.6667%)]  Loss: 2.307584
Train Epoch: 3 [1600/60000 (88.8889%)]  Loss: 2.305704
Test set: Average loss: 2.3014, Accuracy: 1591/10000 (15.9100%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 2.300904
Train Epoch: 4 [400/60000 (22.2222%)]   Loss: 2.307925
Train Epoch: 4 [800/60000 (44.4444%)]   Loss: 2.291931
Train Epoch: 4 [1200/60000 (66.6667%)]  Loss: 2.307729
Train Epoch: 4 [1600/60000 (88.8889%)]  Loss: 2.293281
Test set: Average loss: 2.3011, Accuracy: 1658/10000 (16.5800%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 2.296307
Train Epoch: 5 [400/60000 (22.2222%)]   Loss: 2.299821
Train Epoch: 5 [800/60000 (44.4444%)]   Loss: 2.297099
Train Epoch: 5 [1200/60000 (66.6667%)]  Loss: 2.302952
Train Epoch: 5 [1600/60000 (88.8889%)]  Loss: 2.301417
Test set: Average loss: 2.3003, Accuracy: 1738/10000 (17.3800%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 2.300718
Train Epoch: 6 [400/60000 (22.2222%)]   Loss: 2.317502
Train Epoch: 6 [800/60000 (44.4444%)]   Loss: 2.306867
Train Epoch: 6 [1200/60000 (66.6667%)]  Loss: 2.284281
Train Epoch: 6 [1600/60000 (88.8889%)]  Loss: 2.297437
Test set: Average loss: 2.2994, Accuracy: 1135/10000 (11.3500%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 2.296647
Train Epoch: 7 [400/60000 (22.2222%)]   Loss: 2.327686
Train Epoch: 7 [800/60000 (44.4444%)]   Loss: 2.295613
Train Epoch: 7 [1200/60000 (66.6667%)]  Loss: 2.296510
Train Epoch: 7 [1600/60000 (88.8889%)]  Loss: 2.304519
Test set: Average loss: 2.2980, Accuracy: 2597/10000 (25.9700%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 2.294032
Train Epoch: 8 [400/60000 (22.2222%)]   Loss: 2.299747
Train Epoch: 8 [800/60000 (44.4444%)]   Loss: 2.288682
Train Epoch: 8 [1200/60000 (66.6667%)]  Loss: 2.293507
Train Epoch: 8 [1600/60000 (88.8889%)]  Loss: 2.313107
Test set: Average loss: 2.2955, Accuracy: 2310/10000 (23.1000%)
```

```
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 2.295470
Train Epoch: 9 [400/60000 (22.2222%)]   Loss: 2.290008
Train Epoch: 9 [800/60000 (44.4444%)]   Loss: 2.288448
Train Epoch: 9 [1200/60000 (66.6667%)]  Loss: 2.298070
Train Epoch: 9 [1600/60000 (88.8889%)]  Loss: 2.286863
Test set: Average loss: 2.2882, Accuracy: 3095/10000 (30.9500%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 2.284304
Train Epoch: 10 [400/60000 (22.2222%)]  Loss: 2.285856
Train Epoch: 10 [800/60000 (44.4444%)]  Loss: 2.290143
Train Epoch: 10 [1200/60000 (66.6667%)] Loss: 2.255402
Train Epoch: 10 [1600/60000 (88.8889%)] Loss: 2.251816
Test set: Average loss: 2.2536, Accuracy: 4632/10000 (46.3200%)

Training with 6000 samples:
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.246469
Train Epoch: 1 [400/60000 (6.6667%)]    Loss: 2.252727
Train Epoch: 1 [800/60000 (13.3333%)]   Loss: 2.165347
Train Epoch: 1 [1200/60000 (20.0000%)]  Loss: 2.231647
Train Epoch: 1 [1600/60000 (26.6667%)]  Loss: 1.840636
Train Epoch: 1 [2000/60000 (33.3333%)]  Loss: 2.707223
Train Epoch: 1 [2400/60000 (40.0000%)]  Loss: 2.402066
Train Epoch: 1 [2800/60000 (46.6667%)]  Loss: 0.670231
Train Epoch: 1 [3200/60000 (53.3333%)]  Loss: 0.581167
Train Epoch: 1 [3600/60000 (60.0000%)]  Loss: 1.240760
Train Epoch: 1 [4000/60000 (66.6667%)]  Loss: 0.360702
Train Epoch: 1 [4400/60000 (73.3333%)]  Loss: 0.664767
Train Epoch: 1 [4800/60000 (80.0000%)]  Loss: 1.610089
Train Epoch: 1 [5200/60000 (86.6667%)]  Loss: 1.546953
Train Epoch: 1 [5600/60000 (93.3333%)]  Loss: 0.408832
Test set: Average loss: 0.4259, Accuracy: 8735/10000 (87.3500%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.458233
Train Epoch: 2 [400/60000 (6.6667%)]    Loss: 1.616388
Train Epoch: 2 [800/60000 (13.3333%)]   Loss: 0.134129
Train Epoch: 2 [1200/60000 (20.0000%)]  Loss: 0.199540
Train Epoch: 2 [1600/60000 (26.6667%)]  Loss: 0.109859
Train Epoch: 2 [2000/60000 (33.3333%)]  Loss: 0.060126
Train Epoch: 2 [2400/60000 (40.0000%)]  Loss: 0.088138
Train Epoch: 2 [2800/60000 (46.6667%)]  Loss: 0.021954
Train Epoch: 2 [3200/60000 (53.3333%)]  Loss: 0.047197
Train Epoch: 2 [3600/60000 (60.0000%)]  Loss: 0.094358
Train Epoch: 2 [4000/60000 (66.6667%)]  Loss: 0.042774
Train Epoch: 2 [4400/60000 (73.3333%)]  Loss: 0.360398
Train Epoch: 2 [4800/60000 (80.0000%)]  Loss: 0.050388
Train Epoch: 2 [5200/60000 (86.6667%)]  Loss: 0.096528
Train Epoch: 2 [5600/60000 (93.3333%)]  Loss: 0.022910
Test set: Average loss: 0.1582, Accuracy: 9523/10000 (95.2300%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.091033
Train Epoch: 3 [400/60000 (6.6667%)]    Loss: 0.011954
Train Epoch: 3 [800/60000 (13.3333%)]   Loss: 0.026503
Train Epoch: 3 [1200/60000 (20.0000%)]  Loss: 0.649825
Train Epoch: 3 [1600/60000 (26.6667%)]  Loss: 0.019135
Train Epoch: 3 [2000/60000 (33.3333%)]  Loss: 0.079188
Train Epoch: 3 [2400/60000 (40.0000%)]  Loss: 0.026631
Train Epoch: 3 [2800/60000 (46.6667%)]  Loss: 0.008562
```

```
Train Epoch: 3 [3200/60000 (53.3333%)]  Loss: 1.417481
Train Epoch: 3 [3600/60000 (60.0000%)]  Loss: 0.021888
Train Epoch: 3 [4000/60000 (66.6667%)]  Loss: 0.065322
Train Epoch: 3 [4400/60000 (73.3333%)]  Loss: 0.023754
Train Epoch: 3 [4800/60000 (80.0000%)]  Loss: 0.007815
Train Epoch: 3 [5200/60000 (86.6667%)]  Loss: 0.010283
Train Epoch: 3 [5600/60000 (93.3333%)]  Loss: 0.087687
Test set: Average loss: 0.1561, Accuracy: 9504/10000 (95.0400%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.300818
Train Epoch: 4 [400/60000 (6.6667%)]    Loss: 0.019732
Train Epoch: 4 [800/60000 (13.3333%)]   Loss: 0.006670
Train Epoch: 4 [1200/60000 (20.0000%)]  Loss: 0.001789
Train Epoch: 4 [1600/60000 (26.6667%)]  Loss: 0.055788
Train Epoch: 4 [2000/60000 (33.3333%)]  Loss: 0.022386
Train Epoch: 4 [2400/60000 (40.0000%)]  Loss: 0.003781
Train Epoch: 4 [2800/60000 (46.6667%)]  Loss: 0.022768
Train Epoch: 4 [3200/60000 (53.3333%)]  Loss: 0.020603
Train Epoch: 4 [3600/60000 (60.0000%)]  Loss: 0.011040
Train Epoch: 4 [4000/60000 (66.6667%)]  Loss: 0.144180
Train Epoch: 4 [4400/60000 (73.3333%)]  Loss: 0.023346
Train Epoch: 4 [4800/60000 (80.0000%)]  Loss: 0.013227
Train Epoch: 4 [5200/60000 (86.6667%)]  Loss: 0.005026
Train Epoch: 4 [5600/60000 (93.3333%)]  Loss: 0.004636
Test set: Average loss: 0.1128, Accuracy: 9641/10000 (96.4100%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.002460
Train Epoch: 5 [400/60000 (6.6667%)]    Loss: 0.020846
Train Epoch: 5 [800/60000 (13.3333%)]   Loss: 0.007422
Train Epoch: 5 [1200/60000 (20.0000%)]  Loss: 0.001236
Train Epoch: 5 [1600/60000 (26.6667%)]  Loss: 0.014601
Train Epoch: 5 [2000/60000 (33.3333%)]  Loss: 0.002287
Train Epoch: 5 [2400/60000 (40.0000%)]  Loss: 0.580309
Train Epoch: 5 [2800/60000 (46.6667%)]  Loss: 0.003417
Train Epoch: 5 [3200/60000 (53.3333%)]  Loss: 0.001653
Train Epoch: 5 [3600/60000 (60.0000%)]  Loss: 0.004670
Train Epoch: 5 [4000/60000 (66.6667%)]  Loss: 0.016215
Train Epoch: 5 [4400/60000 (73.3333%)]  Loss: 0.004621
Train Epoch: 5 [4800/60000 (80.0000%)]  Loss: 0.000608
Train Epoch: 5 [5200/60000 (86.6667%)]  Loss: 0.001063
Train Epoch: 5 [5600/60000 (93.3333%)]  Loss: 0.319097
Test set: Average loss: 0.1072, Accuracy: 9669/10000 (96.6900%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.005374
Train Epoch: 6 [400/60000 (6.6667%)]    Loss: 0.002060
Train Epoch: 6 [800/60000 (13.3333%)]   Loss: 0.001855
Train Epoch: 6 [1200/60000 (20.0000%)]  Loss: 0.004950
Train Epoch: 6 [1600/60000 (26.6667%)]  Loss: 0.022325
Train Epoch: 6 [2000/60000 (33.3333%)]  Loss: 0.009205
Train Epoch: 6 [2400/60000 (40.0000%)]  Loss: 0.089990
Train Epoch: 6 [2800/60000 (46.6667%)]  Loss: 0.003219
Train Epoch: 6 [3200/60000 (53.3333%)]  Loss: 0.002039
Train Epoch: 6 [3600/60000 (60.0000%)]  Loss: 0.134559
Train Epoch: 6 [4000/60000 (66.6667%)]  Loss: 0.016132
Train Epoch: 6 [4400/60000 (73.3333%)]  Loss: 0.235679
Train Epoch: 6 [4800/60000 (80.0000%)]  Loss: 0.000617
```

```
Train Epoch: 6 [5200/60000 (86.6667%)]  Loss: 0.012205
Train Epoch: 6 [5600/60000 (93.3333%)]  Loss: 0.044451
Test set: Average loss: 0.1406, Accuracy: 9574/10000 (95.7400%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.002622
Train Epoch: 7 [400/60000 (6.6667%)]    Loss: 0.020386
Train Epoch: 7 [800/60000 (13.3333%)]   Loss: 0.097227
Train Epoch: 7 [1200/60000 (20.0000%)]  Loss: 0.002135
Train Epoch: 7 [1600/60000 (26.6667%)]  Loss: 0.002416
Train Epoch: 7 [2000/60000 (33.3333%)]  Loss: 0.000308
Train Epoch: 7 [2400/60000 (40.0000%)]  Loss: 0.000405
Train Epoch: 7 [2800/60000 (46.6667%)]  Loss: 0.001106
Train Epoch: 7 [3200/60000 (53.3333%)]  Loss: 0.001357
Train Epoch: 7 [3600/60000 (60.0000%)]  Loss: 0.007110
Train Epoch: 7 [4000/60000 (66.6667%)]  Loss: 0.021195
Train Epoch: 7 [4400/60000 (73.3333%)]  Loss: 0.029762
Train Epoch: 7 [4800/60000 (80.0000%)]  Loss: 0.006004
Train Epoch: 7 [5200/60000 (86.6667%)]  Loss: 0.002088
Train Epoch: 7 [5600/60000 (93.3333%)]  Loss: 0.284898
Test set: Average loss: 0.0906, Accuracy: 9725/10000 (97.2500%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.005987
Train Epoch: 8 [400/60000 (6.6667%)]    Loss: 0.019809
Train Epoch: 8 [800/60000 (13.3333%)]   Loss: 0.000709
Train Epoch: 8 [1200/60000 (20.0000%)]  Loss: 0.000340
Train Epoch: 8 [1600/60000 (26.6667%)]  Loss: 0.028147
Train Epoch: 8 [2000/60000 (33.3333%)]  Loss: 0.000095
Train Epoch: 8 [2400/60000 (40.0000%)]  Loss: 0.009932
Train Epoch: 8 [2800/60000 (46.6667%)]  Loss: 0.000125
Train Epoch: 8 [3200/60000 (53.3333%)]  Loss: 0.000407
Train Epoch: 8 [3600/60000 (60.0000%)]  Loss: 0.318162
Train Epoch: 8 [4000/60000 (66.6667%)]  Loss: 0.003190
Train Epoch: 8 [4400/60000 (73.3333%)]  Loss: 0.066329
Train Epoch: 8 [4800/60000 (80.0000%)]  Loss: 0.000497
Train Epoch: 8 [5200/60000 (86.6667%)]  Loss: 0.006419
Train Epoch: 8 [5600/60000 (93.3333%)]  Loss: 0.133376
Test set: Average loss: 0.1119, Accuracy: 9678/10000 (96.7800%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.000193
Train Epoch: 9 [400/60000 (6.6667%)]    Loss: 0.000371
Train Epoch: 9 [800/60000 (13.3333%)]   Loss: 0.001012
Train Epoch: 9 [1200/60000 (20.0000%)]  Loss: 0.005208
Train Epoch: 9 [1600/60000 (26.6667%)]  Loss: 0.000146
Train Epoch: 9 [2000/60000 (33.3333%)]  Loss: 0.000028
Train Epoch: 9 [2400/60000 (40.0000%)]  Loss: 0.000025
Train Epoch: 9 [2800/60000 (46.6667%)]  Loss: 0.000038
Train Epoch: 9 [3200/60000 (53.3333%)]  Loss: 0.000186
Train Epoch: 9 [3600/60000 (60.0000%)]  Loss: 0.001653
Train Epoch: 9 [4000/60000 (66.6667%)]  Loss: 0.000560
Train Epoch: 9 [4400/60000 (73.3333%)]  Loss: 0.000164
Train Epoch: 9 [4800/60000 (80.0000%)]  Loss: 0.013924
Train Epoch: 9 [5200/60000 (86.6667%)]  Loss: 0.000559
Train Epoch: 9 [5600/60000 (93.3333%)]  Loss: 0.005654
Test set: Average loss: 0.1375, Accuracy: 9634/10000 (96.3400%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.000539
Train Epoch: 10 [400/60000 (6.6667%)]   Loss: 0.001753
Train Epoch: 10 [800/60000 (13.3333%)]  Loss: 0.000261
Train Epoch: 10 [1200/60000 (20.0000%)] Loss: 0.000208
```

```
Train Epoch: 10 [1200/60000 (20.0000%)] Loss: 0.000208
Train Epoch: 10 [1600/60000 (26.6667%)] Loss: 0.000031
Train Epoch: 10 [2000/60000 (33.3333%)] Loss: 0.034552
Train Epoch: 10 [2400/60000 (40.0000%)] Loss: 0.000676
Train Epoch: 10 [2800/60000 (46.6667%)] Loss: 0.000209
Train Epoch: 10 [3200/60000 (53.3333%)] Loss: 0.001619
Train Epoch: 10 [3600/60000 (60.0000%)] Loss: 0.001663
Train Epoch: 10 [4000/60000 (66.6667%)] Loss: 0.000142
Train Epoch: 10 [4400/60000 (73.3333%)] Loss: 0.000039
Train Epoch: 10 [4800/60000 (80.0000%)] Loss: 0.000036
Train Epoch: 10 [5200/60000 (86.6667%)] Loss: 0.024277
Train Epoch: 10 [5600/60000 (93.3333%)] Loss: 0.004822
Test set: Average loss: 0.0917, Accuracy: 9742/10000 (97.4200%)

Training with 18000 samples:
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 0.022086
Train Epoch: 1 [400/60000 (2.2222%)]    Loss: 0.004750
Train Epoch: 1 [800/60000 (4.4444%)]    Loss: 0.003717
Train Epoch: 1 [1200/60000 (6.6667%)]   Loss: 0.563344
Train Epoch: 1 [1600/60000 (8.8889%)]   Loss: 0.074187
Train Epoch: 1 [2000/60000 (11.1111%)]  Loss: 0.002466
Train Epoch: 1 [2400/60000 (13.3333%)]  Loss: 0.004437
Train Epoch: 1 [2800/60000 (15.5556%)]  Loss: 0.034838
Train Epoch: 1 [3200/60000 (17.7778%)]  Loss: 0.016173
Train Epoch: 1 [3600/60000 (20.0000%)]  Loss: 0.020301
Train Epoch: 1 [4000/60000 (22.2222%)]  Loss: 0.016334
Train Epoch: 1 [4400/60000 (24.4444%)]  Loss: 0.006413
Train Epoch: 1 [4800/60000 (26.6667%)]  Loss: 0.003466
Train Epoch: 1 [5200/60000 (28.8889%)]  Loss: 0.003195
Train Epoch: 1 [5600/60000 (31.1111%)]  Loss: 0.002606
Train Epoch: 1 [6000/60000 (33.3333%)]  Loss: 0.005463
Train Epoch: 1 [6400/60000 (35.5556%)]  Loss: 0.009161
Train Epoch: 1 [6800/60000 (37.7778%)]  Loss: 0.011724
Train Epoch: 1 [7200/60000 (40.0000%)]  Loss: 0.001245
Train Epoch: 1 [7600/60000 (42.2222%)]  Loss: 0.168978
Train Epoch: 1 [8000/60000 (44.4444%)]  Loss: 0.000888
Train Epoch: 1 [8400/60000 (46.6667%)]  Loss: 0.022257
Train Epoch: 1 [8800/60000 (48.8889%)]  Loss: 0.003867
Train Epoch: 1 [9200/60000 (51.1111%)]  Loss: 0.307956
Train Epoch: 1 [9600/60000 (53.3333%)]  Loss: 0.000843
Train Epoch: 1 [10000/60000 (55.5556%)] Loss: 0.000170
Train Epoch: 1 [10400/60000 (57.7778%)] Loss: 0.115316
Train Epoch: 1 [10800/60000 (60.0000%)] Loss: 0.023112
Train Epoch: 1 [11200/60000 (62.2222%)] Loss: 0.001760
Train Epoch: 1 [11600/60000 (64.4444%)] Loss: 0.012740
Train Epoch: 1 [12000/60000 (66.6667%)] Loss: 0.014588
Train Epoch: 1 [12400/60000 (68.8889%)] Loss: 0.004485
Train Epoch: 1 [12800/60000 (71.1111%)] Loss: 0.000432
Train Epoch: 1 [13200/60000 (73.3333%)] Loss: 0.003583
Train Epoch: 1 [13600/60000 (75.5556%)] Loss: 0.000162
Train Epoch: 1 [14000/60000 (77.7778%)] Loss: 0.006239
Train Epoch: 1 [14400/60000 (80.0000%)] Loss: 0.223876
Train Epoch: 1 [14800/60000 (82.2222%)] Loss: 0.000331
Train Epoch: 1 [15200/60000 (84.4444%)] Loss: 0.692023
Train Epoch: 1 [15600/60000 (86.6667%)] Loss: 0.000624
```

```
Train Epoch: 1 [15600/60000 (86.6667%)] Loss: 0.000624
Train Epoch: 1 [16000/60000 (88.8889%)] Loss: 0.001317
Train Epoch: 1 [16400/60000 (91.1111%)] Loss: 0.001247
Train Epoch: 1 [16800/60000 (93.3333%)] Loss: 0.001569
Train Epoch: 1 [17200/60000 (95.5556%)] Loss: 0.002338
Train Epoch: 1 [17600/60000 (97.7778%)] Loss: 0.000112
Test set: Average loss: 0.0790, Accuracy: 9760/10000 (97.6000%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.004445
Train Epoch: 2 [400/60000 (2.2222%)]    Loss: 0.006138
Train Epoch: 2 [800/60000 (4.4444%)]    Loss: 0.020329
Train Epoch: 2 [1200/60000 (6.6667%)]   Loss: 0.022212
Train Epoch: 2 [1600/60000 (8.8889%)]   Loss: 0.000521
Train Epoch: 2 [2000/60000 (11.1111%)]  Loss: 0.000385
Train Epoch: 2 [2400/60000 (13.3333%)]  Loss: 0.000108
Train Epoch: 2 [2800/60000 (15.5556%)]  Loss: 0.000968
Train Epoch: 2 [3200/60000 (17.7778%)]  Loss: 0.047226
Train Epoch: 2 [3600/60000 (20.0000%)]  Loss: 0.310431
Train Epoch: 2 [4000/60000 (22.2222%)]  Loss: 0.008661
Train Epoch: 2 [4400/60000 (24.4444%)]  Loss: 0.006574
Train Epoch: 2 [4800/60000 (26.6667%)]  Loss: 0.009093
Train Epoch: 2 [5200/60000 (28.8889%)]  Loss: 0.034469
Train Epoch: 2 [5600/60000 (31.1111%)]  Loss: 0.007972
Train Epoch: 2 [6000/60000 (33.3333%)]  Loss: 0.001752
Train Epoch: 2 [6400/60000 (35.5556%)]  Loss: 0.004807
Train Epoch: 2 [6800/60000 (37.7778%)]  Loss: 0.001005
Train Epoch: 2 [7200/60000 (40.0000%)]  Loss: 0.034209
Train Epoch: 2 [7600/60000 (42.2222%)]  Loss: 0.003529
Train Epoch: 2 [8000/60000 (44.4444%)]  Loss: 0.004216
Train Epoch: 2 [8400/60000 (46.6667%)]  Loss: 0.001889
Train Epoch: 2 [8800/60000 (48.8889%)]  Loss: 0.002820
Train Epoch: 2 [9200/60000 (51.1111%)]  Loss: 0.008430
Train Epoch: 2 [9600/60000 (53.3333%)]  Loss: 0.005940
Train Epoch: 2 [10000/60000 (55.5556%)] Loss: 0.006693
Train Epoch: 2 [10400/60000 (57.7778%)] Loss: 0.028196
Train Epoch: 2 [10800/60000 (60.0000%)] Loss: 0.120949
Train Epoch: 2 [11200/60000 (62.2222%)] Loss: 0.000572
Train Epoch: 2 [11600/60000 (64.4444%)] Loss: 0.047919
Train Epoch: 2 [12000/60000 (66.6667%)] Loss: 0.058097
Train Epoch: 2 [12400/60000 (68.8889%)] Loss: 0.024323
Train Epoch: 2 [12800/60000 (71.1111%)] Loss: 0.000573
Train Epoch: 2 [13200/60000 (73.3333%)] Loss: 0.000018
Train Epoch: 2 [13600/60000 (75.5556%)] Loss: 0.000093
Train Epoch: 2 [14000/60000 (77.7778%)] Loss: 0.417536
Train Epoch: 2 [14400/60000 (80.0000%)] Loss: 0.001273
Train Epoch: 2 [14800/60000 (82.2222%)] Loss: 0.000095
Train Epoch: 2 [15200/60000 (84.4444%)] Loss: 1.374198
Train Epoch: 2 [15600/60000 (86.6667%)] Loss: 0.000904
Train Epoch: 2 [16000/60000 (88.8889%)] Loss: 0.002839
Train Epoch: 2 [16400/60000 (91.1111%)] Loss: 0.737422
Train Epoch: 2 [16800/60000 (93.3333%)] Loss: 0.000208
Train Epoch: 2 [17200/60000 (95.5556%)] Loss: 0.455314
Train Epoch: 2 [17600/60000 (97.7778%)] Loss: 0.004784
Test set: Average loss: 0.0739, Accuracy: 9757/10000 (97.5700%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.000813
```

```
Train Epoch: 3 [400/60000 (2.2222%)]    Loss: 0.000623
Train Epoch: 3 [800/60000 (4.4444%)]    Loss: 0.016873
Train Epoch: 3 [1200/60000 (6.6667%)]   Loss: 0.010937
Train Epoch: 3 [1600/60000 (8.8889%)]   Loss: 0.003177
Train Epoch: 3 [2000/60000 (11.1111%)]  Loss: 0.014159
Train Epoch: 3 [2400/60000 (13.3333%)]  Loss: 0.362292
Train Epoch: 3 [2800/60000 (15.5556%)]  Loss: 0.111401
Train Epoch: 3 [3200/60000 (17.7778%)]  Loss: 0.002491
Train Epoch: 3 [3600/60000 (20.0000%)]  Loss: 0.000995
Train Epoch: 3 [4000/60000 (22.2222%)]  Loss: 1.499174
Train Epoch: 3 [4400/60000 (24.4444%)]  Loss: 0.005713
Train Epoch: 3 [4800/60000 (26.6667%)]  Loss: 0.001689
Train Epoch: 3 [5200/60000 (28.8889%)]  Loss: 0.000116
Train Epoch: 3 [5600/60000 (31.1111%)]  Loss: 0.000016
Train Epoch: 3 [6000/60000 (33.3333%)]  Loss: 0.000059
Train Epoch: 3 [6400/60000 (35.5556%)]  Loss: 0.001608
Train Epoch: 3 [6800/60000 (37.7778%)]  Loss: 0.000129
Train Epoch: 3 [7200/60000 (40.0000%)]  Loss: 0.000487
Train Epoch: 3 [7600/60000 (42.2222%)]  Loss: 0.004080
Train Epoch: 3 [8000/60000 (44.4444%)]  Loss: 0.691130
Train Epoch: 3 [8400/60000 (46.6667%)]  Loss: 0.001860
Train Epoch: 3 [8800/60000 (48.8889%)]  Loss: 0.000023
Train Epoch: 3 [9200/60000 (51.1111%)]  Loss: 0.005746
Train Epoch: 3 [9600/60000 (53.3333%)]  Loss: 0.208176
Train Epoch: 3 [10000/60000 (55.5556%)] Loss: 0.001183
Train Epoch: 3 [10400/60000 (57.7778%)] Loss: 0.000120
Train Epoch: 3 [10800/60000 (60.0000%)] Loss: 0.001332
Train Epoch: 3 [11200/60000 (62.2222%)] Loss: 0.000416
Train Epoch: 3 [11600/60000 (64.4444%)] Loss: 0.048060
Train Epoch: 3 [12000/60000 (66.6667%)] Loss: 0.000883
Train Epoch: 3 [12400/60000 (68.8889%)] Loss: 0.010271
Train Epoch: 3 [12800/60000 (71.1111%)] Loss: 0.000080
Train Epoch: 3 [13200/60000 (73.3333%)] Loss: 1.069287
Train Epoch: 3 [13600/60000 (75.5556%)] Loss: 0.003952
Train Epoch: 3 [14000/60000 (77.7778%)] Loss: 0.003861
Train Epoch: 3 [14400/60000 (80.0000%)] Loss: 0.001380
Train Epoch: 3 [14800/60000 (82.2222%)] Loss: 0.016548
Train Epoch: 3 [15200/60000 (84.4444%)] Loss: 0.198276
Train Epoch: 3 [15600/60000 (86.6667%)] Loss: 0.350101
Train Epoch: 3 [16000/60000 (88.8889%)] Loss: 0.004635
Train Epoch: 3 [16400/60000 (91.1111%)] Loss: 0.002361
Train Epoch: 3 [16800/60000 (93.3333%)] Loss: 0.098100
Train Epoch: 3 [17200/60000 (95.5556%)] Loss: 0.001515
Train Epoch: 3 [17600/60000 (97.7778%)] Loss: 0.005552
Test set: Average loss: 0.0565, Accuracy: 9822/10000 (98.2200%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.057705
Train Epoch: 4 [400/60000 (2.2222%)]    Loss: 0.111628
Train Epoch: 4 [800/60000 (4.4444%)]    Loss: 0.002593
Train Epoch: 4 [1200/60000 (6.6667%)]   Loss: 0.001298
Train Epoch: 4 [1600/60000 (8.8889%)]   Loss: 0.003756
Train Epoch: 4 [2000/60000 (11.1111%)]  Loss: 0.000078
Train Epoch: 4 [2400/60000 (13.3333%)]  Loss: 0.000006
Train Epoch: 4 [2800/60000 (15.5556%)]  Loss: 0.000001
Train Epoch: 4 [3200/60000 (17.7778%)]  Loss: 0.000162
```

```
Train Epoch: 4 [3600/60000 (20.0000%)]  Loss: 0.000589
Train Epoch: 4 [4000/60000 (22.2222%)]  Loss: 0.001057
Train Epoch: 4 [4400/60000 (24.4444%)]  Loss: 0.001245
Train Epoch: 4 [4800/60000 (26.6667%)]  Loss: 0.000284
Train Epoch: 4 [5200/60000 (28.8889%)]  Loss: 0.000084
Train Epoch: 4 [5600/60000 (31.1111%)]  Loss: 0.000293
Train Epoch: 4 [6000/60000 (33.3333%)]  Loss: 0.004347
Train Epoch: 4 [6400/60000 (35.5556%)]  Loss: 0.000104
Train Epoch: 4 [6800/60000 (37.7778%)]  Loss: 0.000020
Train Epoch: 4 [7200/60000 (40.0000%)]  Loss: 0.000608
Train Epoch: 4 [7600/60000 (42.2222%)]  Loss: 0.000010
Train Epoch: 4 [8000/60000 (44.4444%)]  Loss: 0.087479
Train Epoch: 4 [8400/60000 (46.6667%)]  Loss: 0.000046
Train Epoch: 4 [8800/60000 (48.8889%)]  Loss: 0.000466
Train Epoch: 4 [9200/60000 (51.1111%)]  Loss: 0.000212
Train Epoch: 4 [9600/60000 (53.3333%)]  Loss: 0.000146
Train Epoch: 4 [10000/60000 (55.5556%)] Loss: 0.000171
Train Epoch: 4 [10400/60000 (57.7778%)] Loss: 0.002358
Train Epoch: 4 [10800/60000 (60.0000%)] Loss: 0.000562
Train Epoch: 4 [11200/60000 (62.2222%)] Loss: 0.100353
Train Epoch: 4 [11600/60000 (64.4444%)] Loss: 0.000147
Train Epoch: 4 [12000/60000 (66.6667%)] Loss: 0.000028
Train Epoch: 4 [12400/60000 (68.8889%)] Loss: 0.218782
Train Epoch: 4 [12800/60000 (71.1111%)] Loss: 0.000157
Train Epoch: 4 [13200/60000 (73.3333%)] Loss: 0.001112
Train Epoch: 4 [13600/60000 (75.5556%)] Loss: 0.000038
Train Epoch: 4 [14000/60000 (77.7778%)] Loss: 0.000005
Train Epoch: 4 [14400/60000 (80.0000%)] Loss: 0.005452
Train Epoch: 4 [14800/60000 (82.2222%)] Loss: 0.000584
Train Epoch: 4 [15200/60000 (84.4444%)] Loss: 0.000483
Train Epoch: 4 [15600/60000 (86.6667%)] Loss: 0.000840
Train Epoch: 4 [16000/60000 (88.8889%)] Loss: 0.000588
Train Epoch: 4 [16400/60000 (91.1111%)] Loss: 0.000056
Train Epoch: 4 [16800/60000 (93.3333%)] Loss: 0.019119
Train Epoch: 4 [17200/60000 (95.5556%)] Loss: 0.000018
Train Epoch: 4 [17600/60000 (97.7778%)] Loss: 0.006013
Test set: Average loss: 0.0770, Accuracy: 9772/10000 (97.7200%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.006367
Train Epoch: 5 [400/60000 (2.2222%)]    Loss: 0.001062
Train Epoch: 5 [800/60000 (4.4444%)]    Loss: 0.000816
Train Epoch: 5 [1200/60000 (6.6667%)]   Loss: 0.000058
Train Epoch: 5 [1600/60000 (8.8889%)]   Loss: 0.000044
Train Epoch: 5 [2000/60000 (11.1111%)]  Loss: 0.000421
Train Epoch: 5 [2400/60000 (13.3333%)]  Loss: 0.000307
Train Epoch: 5 [2800/60000 (15.5556%)]  Loss: 0.015369
Train Epoch: 5 [3200/60000 (17.7778%)]  Loss: 0.000378
Train Epoch: 5 [3600/60000 (20.0000%)]  Loss: 0.009847
Train Epoch: 5 [4000/60000 (22.2222%)]  Loss: 0.000361
Train Epoch: 5 [4400/60000 (24.4444%)]  Loss: 0.004408
Train Epoch: 5 [4800/60000 (26.6667%)]  Loss: 0.015079
Train Epoch: 5 [5200/60000 (28.8889%)]  Loss: 0.001282
Train Epoch: 5 [5600/60000 (31.1111%)]  Loss: 0.003523
Train Epoch: 5 [6000/60000 (33.3333%)]  Loss: 0.011930
Train Epoch: 5 [6400/60000 (35.5556%)]  Loss: 0.000185
```

```
Train Epoch: 5 [6800/60000 (37.7778%)]  Loss: 0.005633
Train Epoch: 5 [7200/60000 (40.0000%)]  Loss: 0.000688
Train Epoch: 5 [7600/60000 (42.2222%)]  Loss: 0.138818
Train Epoch: 5 [8000/60000 (44.4444%)]  Loss: 0.002863
Train Epoch: 5 [8400/60000 (46.6667%)]  Loss: 0.248372
Train Epoch: 5 [8800/60000 (48.8889%)]  Loss: 0.000623
Train Epoch: 5 [9200/60000 (51.1111%)]  Loss: 0.001879
Train Epoch: 5 [9600/60000 (53.3333%)]  Loss: 0.001452
Train Epoch: 5 [10000/60000 (55.5556%)] Loss: 0.006234
Train Epoch: 5 [10400/60000 (57.7778%)] Loss: 0.003018
Train Epoch: 5 [10800/60000 (60.0000%)] Loss: 0.003691
Train Epoch: 5 [11200/60000 (62.2222%)] Loss: 0.023977
Train Epoch: 5 [11600/60000 (64.4444%)] Loss: 0.000252
Train Epoch: 5 [12000/60000 (66.6667%)] Loss: 0.000646
Train Epoch: 5 [12400/60000 (68.8889%)] Loss: 0.000373
Train Epoch: 5 [12800/60000 (71.1111%)] Loss: 0.000784
Train Epoch: 5 [13200/60000 (73.3333%)] Loss: 0.000916
Train Epoch: 5 [13600/60000 (75.5556%)] Loss: 0.001315
Train Epoch: 5 [14000/60000 (77.7778%)] Loss: 0.000046
Train Epoch: 5 [14400/60000 (80.0000%)] Loss: 0.000166
Train Epoch: 5 [14800/60000 (82.2222%)] Loss: 0.000040
Train Epoch: 5 [15200/60000 (84.4444%)] Loss: 0.022974
Train Epoch: 5 [15600/60000 (86.6667%)] Loss: 0.001740
Train Epoch: 5 [16000/60000 (88.8889%)] Loss: 0.001458
Train Epoch: 5 [16400/60000 (91.1111%)] Loss: 0.000765
Train Epoch: 5 [16800/60000 (93.3333%)] Loss: 0.000150
Train Epoch: 5 [17200/60000 (95.5556%)] Loss: 0.000058
Train Epoch: 5 [17600/60000 (97.7778%)] Loss: 0.000082
Test set: Average loss: 0.0579, Accuracy: 9839/10000 (98.3900%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.003492
Train Epoch: 6 [400/60000 (2.2222%)]    Loss: 0.000039
Train Epoch: 6 [800/60000 (4.4444%)]    Loss: 0.000212
Train Epoch: 6 [1200/60000 (6.6667%)]   Loss: 0.000220
Train Epoch: 6 [1600/60000 (8.8889%)]   Loss: 0.000014
Train Epoch: 6 [2000/60000 (11.1111%)]  Loss: 0.000673
Train Epoch: 6 [2400/60000 (13.3333%)]  Loss: 0.170220
Train Epoch: 6 [2800/60000 (15.5556%)]  Loss: 0.000008
Train Epoch: 6 [3200/60000 (17.7778%)]  Loss: 0.000002
Train Epoch: 6 [3600/60000 (20.0000%)]  Loss: 0.000186
Train Epoch: 6 [4000/60000 (22.2222%)]  Loss: 0.001314
Train Epoch: 6 [4400/60000 (24.4444%)]  Loss: 0.000047
Train Epoch: 6 [4800/60000 (26.6667%)]  Loss: 0.030214
Train Epoch: 6 [5200/60000 (28.8889%)]  Loss: 0.003607
Train Epoch: 6 [5600/60000 (31.1111%)]  Loss: 0.086975
Train Epoch: 6 [6000/60000 (33.3333%)]  Loss: 0.001102
Train Epoch: 6 [6400/60000 (35.5556%)]  Loss: 0.000178
Train Epoch: 6 [6800/60000 (37.7778%)]  Loss: 0.000404
Train Epoch: 6 [7200/60000 (40.0000%)]  Loss: 0.011207
Train Epoch: 6 [7600/60000 (42.2222%)]  Loss: 0.000241
Train Epoch: 6 [8000/60000 (44.4444%)]  Loss: 0.000055
Train Epoch: 6 [8400/60000 (46.6667%)]  Loss: 0.006270
Train Epoch: 6 [8800/60000 (48.8889%)]  Loss: 0.000001
Train Epoch: 6 [9200/60000 (51.1111%)]  Loss: 0.001808
Train Epoch: 6 [9600/60000 (53.3333%)]  Loss: 0.000201
```

```
Train Epoch: 6 [10000/60000 (55.5556%)] Loss: 0.001426
Train Epoch: 6 [10400/60000 (57.7778%)] Loss: 0.000033
Train Epoch: 6 [10800/60000 (60.0000%)] Loss: 0.955136
Train Epoch: 6 [11200/60000 (62.2222%)] Loss: 0.004390
Train Epoch: 6 [11600/60000 (64.4444%)] Loss: 0.000035
Train Epoch: 6 [12000/60000 (66.6667%)] Loss: 0.000011
Train Epoch: 6 [12400/60000 (68.8889%)] Loss: 0.000015
Train Epoch: 6 [12800/60000 (71.1111%)] Loss: 0.011971
Train Epoch: 6 [13200/60000 (73.3333%)] Loss: 0.001685
Train Epoch: 6 [13600/60000 (75.5556%)] Loss: 0.000010
Train Epoch: 6 [14000/60000 (77.7778%)] Loss: 0.000381
Train Epoch: 6 [14400/60000 (80.0000%)] Loss: 0.043450
Train Epoch: 6 [14800/60000 (82.2222%)] Loss: 0.000639
Train Epoch: 6 [15200/60000 (84.4444%)] Loss: 0.000455
Train Epoch: 6 [15600/60000 (86.6667%)] Loss: 0.000156
Train Epoch: 6 [16000/60000 (88.8889%)] Loss: 0.000228
Train Epoch: 6 [16400/60000 (91.1111%)] Loss: 0.000651
Train Epoch: 6 [16800/60000 (93.3333%)] Loss: 0.002945
Train Epoch: 6 [17200/60000 (95.5556%)] Loss: 0.000047
Train Epoch: 6 [17600/60000 (97.7778%)] Loss: 0.000263
Test set: Average loss: 0.0683, Accuracy: 9807/10000 (98.0700%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.009465
Train Epoch: 7 [400/60000 (2.2222%)]    Loss: 0.000061
Train Epoch: 7 [800/60000 (4.4444%)]    Loss: 0.001895
Train Epoch: 7 [1200/60000 (6.6667%)]   Loss: 0.053557
Train Epoch: 7 [1600/60000 (8.8889%)]   Loss: 0.000065
Train Epoch: 7 [2000/60000 (11.1111%)]  Loss: 0.003807
Train Epoch: 7 [2400/60000 (13.3333%)]  Loss: 0.000834
Train Epoch: 7 [2800/60000 (15.5556%)]  Loss: 0.002986
Train Epoch: 7 [3200/60000 (17.7778%)]  Loss: 0.000683
Train Epoch: 7 [3600/60000 (20.0000%)]  Loss: 0.001258
Train Epoch: 7 [4000/60000 (22.2222%)]  Loss: 0.000061
Train Epoch: 7 [4400/60000 (24.4444%)]  Loss: 0.000121
Train Epoch: 7 [4800/60000 (26.6667%)]  Loss: 0.000169
Train Epoch: 7 [5200/60000 (28.8889%)]  Loss: 0.000038
Train Epoch: 7 [5600/60000 (31.1111%)]  Loss: 0.000037
Train Epoch: 7 [6000/60000 (33.3333%)]  Loss: 0.000000
Train Epoch: 7 [6400/60000 (35.5556%)]  Loss: 0.000770
Train Epoch: 7 [6800/60000 (37.7778%)]  Loss: 0.001435
Train Epoch: 7 [7200/60000 (40.0000%)]  Loss: 0.000088
Train Epoch: 7 [7600/60000 (42.2222%)]  Loss: 0.000118
Train Epoch: 7 [8000/60000 (44.4444%)]  Loss: 0.000154
Train Epoch: 7 [8400/60000 (46.6667%)]  Loss: 0.000047
Train Epoch: 7 [8800/60000 (48.8889%)]  Loss: 0.005771
Train Epoch: 7 [9200/60000 (51.1111%)]  Loss: 0.000028
Train Epoch: 7 [9600/60000 (53.3333%)]  Loss: 0.000032
Train Epoch: 7 [10000/60000 (55.5556%)] Loss: 0.000117
Train Epoch: 7 [10400/60000 (57.7778%)] Loss: 0.000890
Train Epoch: 7 [10800/60000 (60.0000%)] Loss: 0.000051
Train Epoch: 7 [11200/60000 (62.2222%)] Loss: 0.000031
Train Epoch: 7 [11600/60000 (64.4444%)] Loss: 0.148176
Train Epoch: 7 [12000/60000 (66.6667%)] Loss: 0.138565
Train Epoch: 7 [12400/60000 (68.8889%)] Loss: 0.000565
Train Epoch: 7 [12800/60000 (71.1111%)] Loss: 0.001471
```

```
Train Epoch: 7 [13200/60000 (73.3333%)] Loss: 0.000171
Train Epoch: 7 [13600/60000 (75.5556%)] Loss: 0.000143
Train Epoch: 7 [14000/60000 (77.7778%)] Loss: 0.000065
Train Epoch: 7 [14400/60000 (80.0000%)] Loss: 0.000028
Train Epoch: 7 [14800/60000 (82.2222%)] Loss: 0.000125
Train Epoch: 7 [15200/60000 (84.4444%)] Loss: 0.000432
Train Epoch: 7 [15600/60000 (86.6667%)] Loss: 0.000652
Train Epoch: 7 [16000/60000 (88.8889%)] Loss: 0.000140
Train Epoch: 7 [16400/60000 (91.1111%)] Loss: 0.000706
Train Epoch: 7 [16800/60000 (93.3333%)] Loss: 0.057807
Train Epoch: 7 [17200/60000 (95.5556%)] Loss: 0.000059
Train Epoch: 7 [17600/60000 (97.7778%)] Loss: 0.266889
Test set: Average loss: 0.0532, Accuracy: 9863/10000 (98.6300%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.002887
Train Epoch: 8 [400/60000 (2.2222%)]    Loss: 0.108906
Train Epoch: 8 [800/60000 (4.4444%)]    Loss: 0.000002
Train Epoch: 8 [1200/60000 (6.6667%)]   Loss: 0.000007
Train Epoch: 8 [1600/60000 (8.8889%)]   Loss: 0.000127
Train Epoch: 8 [2000/60000 (11.1111%)]  Loss: 0.042993
Train Epoch: 8 [2400/60000 (13.3333%)]  Loss: 0.085740
Train Epoch: 8 [2800/60000 (15.5556%)]  Loss: 0.082127
Train Epoch: 8 [3200/60000 (17.7778%)]  Loss: 0.000011
Train Epoch: 8 [3600/60000 (20.0000%)]  Loss: 1.624889
Train Epoch: 8 [4000/60000 (22.2222%)]  Loss: 0.000010
Train Epoch: 8 [4400/60000 (24.4444%)]  Loss: 0.000032
Train Epoch: 8 [4800/60000 (26.6667%)]  Loss: 0.000006
Train Epoch: 8 [5200/60000 (28.8889%)]  Loss: 0.000011
Train Epoch: 8 [5600/60000 (31.1111%)]  Loss: 0.000315
Train Epoch: 8 [6000/60000 (33.3333%)]  Loss: 0.000038
Train Epoch: 8 [6400/60000 (35.5556%)]  Loss: 0.018846
Train Epoch: 8 [6800/60000 (37.7778%)]  Loss: 0.000027
Train Epoch: 8 [7200/60000 (40.0000%)]  Loss: 0.000003
Train Epoch: 8 [7600/60000 (42.2222%)]  Loss: 0.000018
Train Epoch: 8 [8000/60000 (44.4444%)]  Loss: 0.000601
Train Epoch: 8 [8400/60000 (46.6667%)]  Loss: 0.000016
Train Epoch: 8 [8800/60000 (48.8889%)]  Loss: 0.000022
Train Epoch: 8 [9200/60000 (51.1111%)]  Loss: 0.007810
Train Epoch: 8 [9600/60000 (53.3333%)]  Loss: 0.042067
Train Epoch: 8 [10000/60000 (55.5556%)] Loss: 0.000013
Train Epoch: 8 [10400/60000 (57.7778%)] Loss: 0.000149
Train Epoch: 8 [10800/60000 (60.0000%)] Loss: 0.000230
Train Epoch: 8 [11200/60000 (62.2222%)] Loss: 0.001283
Train Epoch: 8 [11600/60000 (64.4444%)] Loss: 0.000019
Train Epoch: 8 [12000/60000 (66.6667%)] Loss: 0.001477
Train Epoch: 8 [12400/60000 (68.8889%)] Loss: 0.006462
Train Epoch: 8 [12800/60000 (71.1111%)] Loss: 0.000625
Train Epoch: 8 [13200/60000 (73.3333%)] Loss: 0.000056
Train Epoch: 8 [13600/60000 (75.5556%)] Loss: 0.042395
Train Epoch: 8 [14000/60000 (77.7778%)] Loss: 0.000065
Train Epoch: 8 [14400/60000 (80.0000%)] Loss: 0.003818
Train Epoch: 8 [14800/60000 (82.2222%)] Loss: 0.017381
Train Epoch: 8 [15200/60000 (84.4444%)] Loss: 0.000729
Train Epoch: 8 [15600/60000 (86.6667%)] Loss: 0.000335
Train Epoch: 8 [16000/60000 (88.8889%)] Loss: 0.000034
```

```
Train Epoch: 8 [16400/60000 (91.1111%)] Loss: 0.000344
Train Epoch: 8 [16800/60000 (93.3333%)] Loss: 0.000975
Train Epoch: 8 [17200/60000 (95.5556%)] Loss: 0.000270
Train Epoch: 8 [17600/60000 (97.7778%)] Loss: 0.003157
Test set: Average loss: 0.0602, Accuracy: 9847/10000 (98.4700%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.000019
Train Epoch: 9 [400/60000 (2.2222%)]    Loss: 0.000711
Train Epoch: 9 [800/60000 (4.4444%)]    Loss: 0.001251
Train Epoch: 9 [1200/60000 (6.6667%)]   Loss: 0.004729
Train Epoch: 9 [1600/60000 (8.8889%)]   Loss: 0.000003
Train Epoch: 9 [2000/60000 (11.1111%)]  Loss: 0.000106
Train Epoch: 9 [2400/60000 (13.3333%)]  Loss: 0.000134
Train Epoch: 9 [2800/60000 (15.5556%)]  Loss: 0.000005
Train Epoch: 9 [3200/60000 (17.7778%)]  Loss: 1.333259
Train Epoch: 9 [3600/60000 (20.0000%)]  Loss: 0.000022
Train Epoch: 9 [4000/60000 (22.2222%)]  Loss: 0.106859
Train Epoch: 9 [4400/60000 (24.4444%)]  Loss: 0.000083
Train Epoch: 9 [4800/60000 (26.6667%)]  Loss: 0.144424
Train Epoch: 9 [5200/60000 (28.8889%)]  Loss: 0.000514
Train Epoch: 9 [5600/60000 (31.1111%)]  Loss: 0.000007
Train Epoch: 9 [6000/60000 (33.3333%)]  Loss: 0.000050
Train Epoch: 9 [6400/60000 (35.5556%)]  Loss: 0.000018
Train Epoch: 9 [6800/60000 (37.7778%)]  Loss: 0.000187
Train Epoch: 9 [7200/60000 (40.0000%)]  Loss: 0.000471
Train Epoch: 9 [7600/60000 (42.2222%)]  Loss: 0.000045
Train Epoch: 9 [8000/60000 (44.4444%)]  Loss: 0.324573
Train Epoch: 9 [8400/60000 (46.6667%)]  Loss: 0.000002
Train Epoch: 9 [8800/60000 (48.8889%)]  Loss: 0.000078
Train Epoch: 9 [9200/60000 (51.1111%)]  Loss: 0.004772
Train Epoch: 9 [9600/60000 (53.3333%)]  Loss: 0.000079
Train Epoch: 9 [10000/60000 (55.5556%)] Loss: 0.000199
Train Epoch: 9 [10400/60000 (57.7778%)] Loss: 0.000014
Train Epoch: 9 [10800/60000 (60.0000%)] Loss: 0.000000
Train Epoch: 9 [11200/60000 (62.2222%)] Loss: 0.000106
Train Epoch: 9 [11600/60000 (64.4444%)] Loss: 0.000002
Train Epoch: 9 [12000/60000 (66.6667%)] Loss: 0.000055
Train Epoch: 9 [12400/60000 (68.8889%)] Loss: 0.038370
Train Epoch: 9 [12800/60000 (71.1111%)] Loss: 0.053901
Train Epoch: 9 [13200/60000 (73.3333%)] Loss: 0.000191
Train Epoch: 9 [13600/60000 (75.5556%)] Loss: 0.000445
Train Epoch: 9 [14000/60000 (77.7778%)] Loss: 0.000005
Train Epoch: 9 [14400/60000 (80.0000%)] Loss: 0.000023
Train Epoch: 9 [14800/60000 (82.2222%)] Loss: 0.309160
Train Epoch: 9 [15200/60000 (84.4444%)] Loss: 0.000061
Train Epoch: 9 [15600/60000 (86.6667%)] Loss: 0.000069
Train Epoch: 9 [16000/60000 (88.8889%)] Loss: 0.000001
Train Epoch: 9 [16400/60000 (91.1111%)] Loss: 0.000121
Train Epoch: 9 [16800/60000 (93.3333%)] Loss: 0.000666
Train Epoch: 9 [17200/60000 (95.5556%)] Loss: 0.139970
Train Epoch: 9 [17600/60000 (97.7778%)] Loss: 0.021234
Test set: Average loss: 0.0575, Accuracy: 9840/10000 (98.4000%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.000865
Train Epoch: 10 [400/60000 (2.2222%)]   Loss: 0.000005
Train Epoch: 10 [800/60000 (4.4444%)]   Loss: 0.000661
```

```
Train Epoch: 10 [1200/60000 (6.6667%)]  Loss: 0.000165
Train Epoch: 10 [1600/60000 (8.8889%)]  Loss: 0.000004
Train Epoch: 10 [2000/60000 (11.1111%)] Loss: 0.015745
Train Epoch: 10 [2400/60000 (13.3333%)] Loss: 0.000001
Train Epoch: 10 [2800/60000 (15.5556%)] Loss: 0.000014
Train Epoch: 10 [3200/60000 (17.7778%)] Loss: 0.000957
Train Epoch: 10 [3600/60000 (20.0000%)] Loss: 0.001022
Train Epoch: 10 [4000/60000 (22.2222%)] Loss: 0.000414
Train Epoch: 10 [4400/60000 (24.4444%)] Loss: 0.000148
Train Epoch: 10 [4800/60000 (26.6667%)] Loss: 0.000000
Train Epoch: 10 [5200/60000 (28.8889%)] Loss: 0.000059
Train Epoch: 10 [5600/60000 (31.1111%)] Loss: 0.000005
Train Epoch: 10 [6000/60000 (33.3333%)] Loss: 0.000002
Train Epoch: 10 [6400/60000 (35.5556%)] Loss: 0.000023
Train Epoch: 10 [6800/60000 (37.7778%)] Loss: 0.000006
Train Epoch: 10 [7200/60000 (40.0000%)] Loss: 0.000009
Train Epoch: 10 [7600/60000 (42.2222%)] Loss: 0.000000
Train Epoch: 10 [8000/60000 (44.4444%)] Loss: 0.000000
Train Epoch: 10 [8400/60000 (46.6667%)] Loss: 0.000061
Train Epoch: 10 [8800/60000 (48.8889%)] Loss: 0.004566
Train Epoch: 10 [9200/60000 (51.1111%)] Loss: 0.000001
Train Epoch: 10 [9600/60000 (53.3333%)] Loss: 0.000007
Train Epoch: 10 [10000/60000 (55.5556%)]        Loss: 0.000007
Train Epoch: 10 [10400/60000 (57.7778%)]        Loss: 0.000010
Train Epoch: 10 [10800/60000 (60.0000%)]        Loss: 0.000018
Train Epoch: 10 [11200/60000 (62.2222%)]        Loss: 0.000001
Train Epoch: 10 [11600/60000 (64.4444%)]        Loss: 0.000008
Train Epoch: 10 [12000/60000 (66.6667%)]        Loss: 0.004052
Train Epoch: 10 [12400/60000 (68.8889%)]        Loss: 0.000002
Train Epoch: 10 [12800/60000 (71.1111%)]        Loss: 0.000001
Train Epoch: 10 [13200/60000 (73.3333%)]        Loss: 0.000001
Train Epoch: 10 [13600/60000 (75.5556%)]        Loss: 0.000001
Train Epoch: 10 [14000/60000 (77.7778%)]        Loss: 0.000000
Train Epoch: 10 [14400/60000 (80.0000%)]        Loss: 0.003528
Train Epoch: 10 [14800/60000 (82.2222%)]        Loss: 0.000001
Train Epoch: 10 [15200/60000 (84.4444%)]        Loss: 0.001597
Train Epoch: 10 [15600/60000 (86.6667%)]        Loss: 0.000002
Train Epoch: 10 [16000/60000 (88.8889%)]        Loss: 0.000014
Train Epoch: 10 [16400/60000 (91.1111%)]        Loss: 0.000152
Train Epoch: 10 [16800/60000 (93.3333%)]        Loss: 0.000154
Train Epoch: 10 [17200/60000 (95.5556%)]        Loss: 0.000019
Train Epoch: 10 [17600/60000 (97.7778%)]        Loss: 0.011042
Test set: Average loss: 0.0685, Accuracy: 9842/10000 (98.4200%)

Training with 60000 samples:
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 0.000034
Train Epoch: 1 [400/60000 (0.6795%)]    Loss: 0.000002
Train Epoch: 1 [800/60000 (1.3591%)]    Loss: 0.146964
Train Epoch: 1 [1200/60000 (2.0386%)]   Loss: 0.000036
Train Epoch: 1 [1600/60000 (2.7181%)]   Loss: 0.000024
Train Epoch: 1 [2000/60000 (3.3977%)]   Loss: 0.002794
Train Epoch: 1 [2400/60000 (4.0772%)]   Loss: 0.000459
Train Epoch: 1 [2800/60000 (4.7567%)]   Loss: 0.003450
Train Epoch: 1 [3200/60000 (5.4363%)]   Loss: 0.009628
```

```
Train Epoch: 1 [3600/60000 (6.1158%)]   Loss: 0.000766
Train Epoch: 1 [4000/60000 (6.7953%)]   Loss: 0.111249
Train Epoch: 1 [4400/60000 (7.4749%)]   Loss: 0.004543
Train Epoch: 1 [4800/60000 (8.1544%)]   Loss: 0.000805
Train Epoch: 1 [5200/60000 (8.8339%)]   Loss: 0.000372
Train Epoch: 1 [5600/60000 (9.5135%)]   Loss: 0.080357
Train Epoch: 1 [6000/60000 (10.1930%)]  Loss: 0.000061
Train Epoch: 1 [6400/60000 (10.8725%)]  Loss: 0.001717
Train Epoch: 1 [6800/60000 (11.5521%)]  Loss: 0.019888
Train Epoch: 1 [7200/60000 (12.2316%)]  Loss: 0.000306
Train Epoch: 1 [7600/60000 (12.9111%)]  Loss: 0.028780
Train Epoch: 1 [8000/60000 (13.5906%)]  Loss: 0.000036
Train Epoch: 1 [8400/60000 (14.2702%)]  Loss: 0.003805
Train Epoch: 1 [8800/60000 (14.9497%)]  Loss: 2.271615
Train Epoch: 1 [9200/60000 (15.6292%)]  Loss: 0.272983
Train Epoch: 1 [9600/60000 (16.3088%)]  Loss: 0.000358
Train Epoch: 1 [10000/60000 (16.9883%)] Loss: 0.001454
Train Epoch: 1 [10400/60000 (17.6678%)] Loss: 0.020625
Train Epoch: 1 [10800/60000 (18.3474%)] Loss: 0.001486
Train Epoch: 1 [11200/60000 (19.0269%)] Loss: 0.000241
Train Epoch: 1 [11600/60000 (19.7064%)] Loss: 0.001104
Train Epoch: 1 [12000/60000 (20.3860%)] Loss: 0.817438
Train Epoch: 1 [12400/60000 (21.0655%)] Loss: 0.001894
Train Epoch: 1 [12800/60000 (21.7450%)] Loss: 0.001632
Train Epoch: 1 [13200/60000 (22.4246%)] Loss: 1.436715
Train Epoch: 1 [13600/60000 (23.1041%)] Loss: 0.000086
Train Epoch: 1 [14000/60000 (23.7836%)] Loss: 0.000095
Train Epoch: 1 [14400/60000 (24.4632%)] Loss: 0.600475
Train Epoch: 1 [14800/60000 (25.1427%)] Loss: 0.004681
Train Epoch: 1 [15200/60000 (25.8222%)] Loss: 0.003050
Train Epoch: 1 [15600/60000 (26.5018%)] Loss: 0.002449
Train Epoch: 1 [16000/60000 (27.1813%)] Loss: 0.945951
Train Epoch: 1 [16400/60000 (27.8608%)] Loss: 0.001431
Train Epoch: 1 [16800/60000 (28.5404%)] Loss: 0.003672
Train Epoch: 1 [17200/60000 (29.2199%)] Loss: 0.003202
Train Epoch: 1 [17600/60000 (29.8994%)] Loss: 0.008841
Train Epoch: 1 [18000/60000 (30.5790%)] Loss: 0.000018
Train Epoch: 1 [18400/60000 (31.2585%)] Loss: 0.002722
Train Epoch: 1 [18800/60000 (31.9380%)] Loss: 0.092470
Train Epoch: 1 [19200/60000 (32.6176%)] Loss: 0.002304
Train Epoch: 1 [19600/60000 (33.2971%)] Loss: 0.000091
Train Epoch: 1 [20000/60000 (33.9766%)] Loss: 0.000009
Train Epoch: 1 [20400/60000 (34.6562%)] Loss: 0.002077
Train Epoch: 1 [20800/60000 (35.3357%)] Loss: 0.727116
Train Epoch: 1 [21200/60000 (36.0152%)] Loss: 0.000091
Train Epoch: 1 [21600/60000 (36.6948%)] Loss: 0.006178
Train Epoch: 1 [22000/60000 (37.3743%)] Loss: 0.091875
Train Epoch: 1 [22400/60000 (38.0538%)] Loss: 0.000222
Train Epoch: 1 [22800/60000 (38.7334%)] Loss: 0.004822
Train Epoch: 1 [23200/60000 (39.4129%)] Loss: 0.000933
Train Epoch: 1 [23600/60000 (40.0924%)] Loss: 0.000048
Train Epoch: 1 [24000/60000 (40.7719%)] Loss: 0.000141
Train Epoch: 1 [24400/60000 (41.4515%)] Loss: 0.042491
Train Epoch: 1 [24800/60000 (42.1310%)] Loss: 0.000072
```

```
Train Epoch: 1 [25200/60000 (42.8105%)] Loss: 0.000343
Train Epoch: 1 [25600/60000 (43.4901%)] Loss: 0.000677
Train Epoch: 1 [26000/60000 (44.1696%)] Loss: 0.000002
Train Epoch: 1 [26400/60000 (44.8491%)] Loss: 0.000620
Train Epoch: 1 [26800/60000 (45.5287%)] Loss: 0.000047
Train Epoch: 1 [27200/60000 (46.2082%)] Loss: 0.002045
Train Epoch: 1 [27600/60000 (46.8877%)] Loss: 0.000400
Train Epoch: 1 [28000/60000 (47.5673%)] Loss: 0.011918
Train Epoch: 1 [28400/60000 (48.2468%)] Loss: 0.000129
Train Epoch: 1 [28800/60000 (48.9263%)] Loss: 0.017652
Train Epoch: 1 [29200/60000 (49.6059%)] Loss: 0.001921
Train Epoch: 1 [29600/60000 (50.2854%)] Loss: 0.000227
Train Epoch: 1 [30000/60000 (50.9649%)] Loss: 0.000804
Train Epoch: 1 [30400/60000 (51.6445%)] Loss: 0.000183
Train Epoch: 1 [30800/60000 (52.3240%)] Loss: 0.000519
Train Epoch: 1 [31200/60000 (53.0035%)] Loss: 0.000073
Train Epoch: 1 [31600/60000 (53.6831%)] Loss: 0.001965
Train Epoch: 1 [32000/60000 (54.3626%)] Loss: 0.026462
Train Epoch: 1 [32400/60000 (55.0421%)] Loss: 0.000365
Train Epoch: 1 [32800/60000 (55.7217%)] Loss: 0.000152
Train Epoch: 1 [33200/60000 (56.4012%)] Loss: 0.006712
Train Epoch: 1 [33600/60000 (57.0807%)] Loss: 0.000154
Train Epoch: 1 [34000/60000 (57.7603%)] Loss: 1.266917
Train Epoch: 1 [34400/60000 (58.4398%)] Loss: 0.300411
Train Epoch: 1 [34800/60000 (59.1193%)] Loss: 0.000587
Train Epoch: 1 [35200/60000 (59.7989%)] Loss: 0.001635
Train Epoch: 1 [35600/60000 (60.4784%)] Loss: 0.012939
Train Epoch: 1 [36000/60000 (61.1579%)] Loss: 0.000126
Train Epoch: 1 [36400/60000 (61.8375%)] Loss: 0.001578
Train Epoch: 1 [36800/60000 (62.5170%)] Loss: 0.000035
Train Epoch: 1 [37200/60000 (63.1965%)] Loss: 0.000110
Train Epoch: 1 [37600/60000 (63.8761%)] Loss: 0.001098
Train Epoch: 1 [38000/60000 (64.5556%)] Loss: 0.000175
Train Epoch: 1 [38400/60000 (65.2351%)] Loss: 0.000407
Train Epoch: 1 [38800/60000 (65.9147%)] Loss: 0.004183
Train Epoch: 1 [39200/60000 (66.5942%)] Loss: 0.000398
Train Epoch: 1 [39600/60000 (67.2737%)] Loss: 0.000404
Train Epoch: 1 [40000/60000 (67.9532%)] Loss: 0.000054
Train Epoch: 1 [40400/60000 (68.6328%)] Loss: 0.000019
Train Epoch: 1 [40800/60000 (69.3123%)] Loss: 0.294717
Train Epoch: 1 [41200/60000 (69.9918%)] Loss: 0.000009
Train Epoch: 1 [41600/60000 (70.6714%)] Loss: 0.016852
Train Epoch: 1 [42000/60000 (71.3509%)] Loss: 0.032568
Train Epoch: 1 [42400/60000 (72.0304%)] Loss: 0.014155
Train Epoch: 1 [42800/60000 (72.7100%)] Loss: 0.000525
Train Epoch: 1 [43200/60000 (73.3895%)] Loss: 0.002751
Train Epoch: 1 [43600/60000 (74.0690%)] Loss: 0.000037
Train Epoch: 1 [44000/60000 (74.7486%)] Loss: 0.003836
Train Epoch: 1 [44400/60000 (75.4281%)] Loss: 0.001116
Train Epoch: 1 [44800/60000 (76.1076%)] Loss: 0.002386
Train Epoch: 1 [45200/60000 (76.7872%)] Loss: 0.000277
Train Epoch: 1 [45600/60000 (77.4667%)] Loss: 0.002744
Train Epoch: 1 [46000/60000 (78.1462%)] Loss: 0.001238
Train Epoch: 1 [46400/60000 (78.8258%)] Loss: 0.004547
```

```
Train Epoch: 1 [46800/60000 (79.5053%)] Loss: 0.035145
Train Epoch: 1 [47200/60000 (80.1848%)] Loss: 0.023019
Train Epoch: 1 [47600/60000 (80.8644%)] Loss: 0.000436
Train Epoch: 1 [48000/60000 (81.5439%)] Loss: 0.000021
Train Epoch: 1 [48400/60000 (82.2234%)] Loss: 0.000990
Train Epoch: 1 [48800/60000 (82.9030%)] Loss: 0.016253
Train Epoch: 1 [49200/60000 (83.5825%)] Loss: 0.001197
Train Epoch: 1 [49600/60000 (84.2620%)] Loss: 0.010523
Train Epoch: 1 [50000/60000 (84.9416%)] Loss: 0.263861
Train Epoch: 1 [50400/60000 (85.6211%)] Loss: 0.000313
Train Epoch: 1 [50800/60000 (86.3006%)] Loss: 0.000409
Train Epoch: 1 [51200/60000 (86.9802%)] Loss: 0.000156
Train Epoch: 1 [51600/60000 (87.6597%)] Loss: 0.227895
Train Epoch: 1 [52000/60000 (88.3392%)] Loss: 0.000180
Train Epoch: 1 [52400/60000 (89.0188%)] Loss: 0.000380
Train Epoch: 1 [52800/60000 (89.6983%)] Loss: 0.007781
Train Epoch: 1 [53200/60000 (90.3778%)] Loss: 0.000520
Train Epoch: 1 [53600/60000 (91.0574%)] Loss: 0.000128
Train Epoch: 1 [54000/60000 (91.7369%)] Loss: 0.497529
Train Epoch: 1 [54400/60000 (92.4164%)] Loss: 0.000171
Train Epoch: 1 [54800/60000 (93.0959%)] Loss: 0.000542
Train Epoch: 1 [55200/60000 (93.7755%)] Loss: 0.004042
Train Epoch: 1 [55600/60000 (94.4550%)] Loss: 0.000133
Train Epoch: 1 [56000/60000 (95.1345%)] Loss: 0.000859
Train Epoch: 1 [56400/60000 (95.8141%)] Loss: 0.000007
Train Epoch: 1 [56800/60000 (96.4936%)] Loss: 0.000052
Train Epoch: 1 [57200/60000 (97.1731%)] Loss: 0.000093
Train Epoch: 1 [57600/60000 (97.8527%)] Loss: 0.014577
Train Epoch: 1 [58000/60000 (98.5322%)] Loss: 0.033881
Train Epoch: 1 [58400/60000 (99.2117%)] Loss: 0.000451
Train Epoch: 1 [58800/60000 (99.8913%)] Loss: 0.000236
Test set: Average loss: 0.0462, Accuracy: 9864/10000 (98.6400%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.000390
Train Epoch: 2 [400/60000 (0.6795%)]    Loss: 0.001652
Train Epoch: 2 [800/60000 (1.3591%)]    Loss: 0.001348
Train Epoch: 2 [1200/60000 (2.0386%)]   Loss: 0.007891
Train Epoch: 2 [1600/60000 (2.7181%)]   Loss: 0.001068
Train Epoch: 2 [2000/60000 (3.3977%)]   Loss: 0.000159
Train Epoch: 2 [2400/60000 (4.0772%)]   Loss: 0.000554
Train Epoch: 2 [2800/60000 (4.7567%)]   Loss: 0.000456
Train Epoch: 2 [3200/60000 (5.4363%)]   Loss: 0.000275
Train Epoch: 2 [3600/60000 (6.1158%)]   Loss: 0.000334
Train Epoch: 2 [4000/60000 (6.7953%)]   Loss: 0.103375
Train Epoch: 2 [4400/60000 (7.4749%)]   Loss: 0.000017
Train Epoch: 2 [4800/60000 (8.1544%)]   Loss: 0.005355
Train Epoch: 2 [5200/60000 (8.8339%)]   Loss: 0.000009
Train Epoch: 2 [5600/60000 (9.5135%)]   Loss: 0.000074
Train Epoch: 2 [6000/60000 (10.1930%)]  Loss: 0.000767
Train Epoch: 2 [6400/60000 (10.8725%)]  Loss: 0.000022
Train Epoch: 2 [6800/60000 (11.5521%)]  Loss: 0.009059
Train Epoch: 2 [7200/60000 (12.2316%)]  Loss: 0.000969
Train Epoch: 2 [7600/60000 (12.9111%)]  Loss: 0.001219
Train Epoch: 2 [8000/60000 (13.5906%)]  Loss: 0.017488
Train Epoch: 2 [8400/60000 (14.2702%)]  Loss: 0.000006
```

```
Train Epoch: 2 [8800/60000 (14.9497%)]  Loss: 0.000112
Train Epoch: 2 [9200/60000 (15.6292%)]  Loss: 0.002181
Train Epoch: 2 [9600/60000 (16.3088%)]  Loss: 0.001368
Train Epoch: 2 [10000/60000 (16.9883%)] Loss: 0.012771
Train Epoch: 2 [10400/60000 (17.6678%)] Loss: 0.000914
Train Epoch: 2 [10800/60000 (18.3474%)] Loss: 0.000040
Train Epoch: 2 [11200/60000 (19.0269%)] Loss: 0.004662
Train Epoch: 2 [11600/60000 (19.7064%)] Loss: 0.033280
Train Epoch: 2 [12000/60000 (20.3860%)] Loss: 0.018126
Train Epoch: 2 [12400/60000 (21.0655%)] Loss: 0.001268
Train Epoch: 2 [12800/60000 (21.7450%)] Loss: 0.000550
Train Epoch: 2 [13200/60000 (22.4246%)] Loss: 0.000224
Train Epoch: 2 [13600/60000 (23.1041%)] Loss: 0.002476
Train Epoch: 2 [14000/60000 (23.7836%)] Loss: 0.000013
Train Epoch: 2 [14400/60000 (24.4632%)] Loss: 0.000176
Train Epoch: 2 [14800/60000 (25.1427%)] Loss: 0.000193
Train Epoch: 2 [15200/60000 (25.8222%)] Loss: 0.001046
Train Epoch: 2 [15600/60000 (26.5018%)] Loss: 0.001185
Train Epoch: 2 [16000/60000 (27.1813%)] Loss: 0.575274
Train Epoch: 2 [16400/60000 (27.8608%)] Loss: 0.000370
Train Epoch: 2 [16800/60000 (28.5404%)] Loss: 0.000098
Train Epoch: 2 [17200/60000 (29.2199%)] Loss: 0.017035
Train Epoch: 2 [17600/60000 (29.8994%)] Loss: 0.000082
Train Epoch: 2 [18000/60000 (30.5790%)] Loss: 0.000035
Train Epoch: 2 [18400/60000 (31.2585%)] Loss: 0.072571
Train Epoch: 2 [18800/60000 (31.9380%)] Loss: 0.000058
Train Epoch: 2 [19200/60000 (32.6176%)] Loss: 0.000006
Train Epoch: 2 [19600/60000 (33.2971%)] Loss: 0.004880
Train Epoch: 2 [20000/60000 (33.9766%)] Loss: 0.000102
Train Epoch: 2 [20400/60000 (34.6562%)] Loss: 0.097889
Train Epoch: 2 [20800/60000 (35.3357%)] Loss: 0.000023
Train Epoch: 2 [21200/60000 (36.0152%)] Loss: 0.004649
Train Epoch: 2 [21600/60000 (36.6948%)] Loss: 0.001177
Train Epoch: 2 [22000/60000 (37.3743%)] Loss: 0.001108
Train Epoch: 2 [22400/60000 (38.0538%)] Loss: 0.128851
Train Epoch: 2 [22800/60000 (38.7334%)] Loss: 0.009019
Train Epoch: 2 [23200/60000 (39.4129%)] Loss: 0.000420
Train Epoch: 2 [23600/60000 (40.0924%)] Loss: 0.002034
Train Epoch: 2 [24000/60000 (40.7719%)] Loss: 0.008262
Train Epoch: 2 [24400/60000 (41.4515%)] Loss: 0.010045
Train Epoch: 2 [24800/60000 (42.1310%)] Loss: 0.000353
Train Epoch: 2 [25200/60000 (42.8105%)] Loss: 0.000572
Train Epoch: 2 [25600/60000 (43.4901%)] Loss: 0.005417
Train Epoch: 2 [26000/60000 (44.1696%)] Loss: 0.374886
Train Epoch: 2 [26400/60000 (44.8491%)] Loss: 0.000076
Train Epoch: 2 [26800/60000 (45.5287%)] Loss: 0.000722
Train Epoch: 2 [27200/60000 (46.2082%)] Loss: 0.003031
Train Epoch: 2 [27600/60000 (46.8877%)] Loss: 0.000257
Train Epoch: 2 [28000/60000 (47.5673%)] Loss: 0.000642
Train Epoch: 2 [28400/60000 (48.2468%)] Loss: 0.000032
Train Epoch: 2 [28800/60000 (48.9263%)] Loss: 0.000134
Train Epoch: 2 [29200/60000 (49.6059%)] Loss: 0.116359
Train Epoch: 2 [29600/60000 (50.2854%)] Loss: 0.569775
Train Epoch: 2 [30000/60000 (50.9649%)] Loss: 0.000480
```

```
Train Epoch: 2 [30400/60000 (51.6445%)] Loss: 0.016431
Train Epoch: 2 [30800/60000 (52.3240%)] Loss: 0.000569
Train Epoch: 2 [31200/60000 (53.0035%)] Loss: 0.000410
Train Epoch: 2 [31600/60000 (53.6831%)] Loss: 0.000036
Train Epoch: 2 [32000/60000 (54.3626%)] Loss: 0.000576
Train Epoch: 2 [32400/60000 (55.0421%)] Loss: 0.020767
Train Epoch: 2 [32800/60000 (55.7217%)] Loss: 0.000366
Train Epoch: 2 [33200/60000 (56.4012%)] Loss: 0.000696
Train Epoch: 2 [33600/60000 (57.0807%)] Loss: 0.049268
Train Epoch: 2 [34000/60000 (57.7603%)] Loss: 0.000025
Train Epoch: 2 [34400/60000 (58.4398%)] Loss: 0.004622
Train Epoch: 2 [34800/60000 (59.1193%)] Loss: 0.443354
Train Epoch: 2 [35200/60000 (59.7989%)] Loss: 0.000020
Train Epoch: 2 [35600/60000 (60.4784%)] Loss: 0.000310
Train Epoch: 2 [36000/60000 (61.1579%)] Loss: 0.000369
Train Epoch: 2 [36400/60000 (61.8375%)] Loss: 0.010738
Train Epoch: 2 [36800/60000 (62.5170%)] Loss: 0.000044
Train Epoch: 2 [37200/60000 (63.1965%)] Loss: 0.000294
Train Epoch: 2 [37600/60000 (63.8761%)] Loss: 0.000028
Train Epoch: 2 [38000/60000 (64.5556%)] Loss: 0.000572
Train Epoch: 2 [38400/60000 (65.2351%)] Loss: 0.000045
Train Epoch: 2 [38800/60000 (65.9147%)] Loss: 0.000079
Train Epoch: 2 [39200/60000 (66.5942%)] Loss: 0.002545
Train Epoch: 2 [39600/60000 (67.2737%)] Loss: 0.000156
Train Epoch: 2 [40000/60000 (67.9532%)] Loss: 0.001242
Train Epoch: 2 [40400/60000 (68.6328%)] Loss: 0.000232
Train Epoch: 2 [40800/60000 (69.3123%)] Loss: 0.000201
Train Epoch: 2 [41200/60000 (69.9918%)] Loss: 0.000743
Train Epoch: 2 [41600/60000 (70.6714%)] Loss: 0.010008
Train Epoch: 2 [42000/60000 (71.3509%)] Loss: 0.008349
Train Epoch: 2 [42400/60000 (72.0304%)] Loss: 0.000112
Train Epoch: 2 [42800/60000 (72.7100%)] Loss: 0.000385
Train Epoch: 2 [43200/60000 (73.3895%)] Loss: 0.014646
Train Epoch: 2 [43600/60000 (74.0690%)] Loss: 0.000503
Train Epoch: 2 [44000/60000 (74.7486%)] Loss: 0.006487
Train Epoch: 2 [44400/60000 (75.4281%)] Loss: 0.000017
Train Epoch: 2 [44800/60000 (76.1076%)] Loss: 0.007845
Train Epoch: 2 [45200/60000 (76.7872%)] Loss: 0.000082
Train Epoch: 2 [45600/60000 (77.4667%)] Loss: 0.000017
Train Epoch: 2 [46000/60000 (78.1462%)] Loss: 0.110349
Train Epoch: 2 [46400/60000 (78.8258%)] Loss: 0.000247
Train Epoch: 2 [46800/60000 (79.5053%)] Loss: 0.073621
Train Epoch: 2 [47200/60000 (80.1848%)] Loss: 0.000259
Train Epoch: 2 [47600/60000 (80.8644%)] Loss: 0.221183
Train Epoch: 2 [48000/60000 (81.5439%)] Loss: 0.034993
Train Epoch: 2 [48400/60000 (82.2234%)] Loss: 0.022858
Train Epoch: 2 [48800/60000 (82.9030%)] Loss: 0.000517
Train Epoch: 2 [49200/60000 (83.5825%)] Loss: 0.000085
Train Epoch: 2 [49600/60000 (84.2620%)] Loss: 0.000050
Train Epoch: 2 [50000/60000 (84.9416%)] Loss: 0.000424
Train Epoch: 2 [50400/60000 (85.6211%)] Loss: 0.000142
Train Epoch: 2 [50800/60000 (86.3006%)] Loss: 0.046690
Train Epoch: 2 [51200/60000 (86.9802%)] Loss: 0.000010
Train Epoch: 2 [51600/60000 (87.6597%)] Loss: 0.000992
```

```
Train Epoch: 2 [52000/60000 (88.3392%)] Loss: 0.000786
Train Epoch: 2 [52400/60000 (89.0188%)] Loss: 0.000055
Train Epoch: 2 [52800/60000 (89.6983%)] Loss: 0.000123
Train Epoch: 2 [53200/60000 (90.3778%)] Loss: 0.001484
Train Epoch: 2 [53600/60000 (91.0574%)] Loss: 0.001843
Train Epoch: 2 [54000/60000 (91.7369%)] Loss: 0.003509
Train Epoch: 2 [54400/60000 (92.4164%)] Loss: 0.000264
Train Epoch: 2 [54800/60000 (93.0959%)] Loss: 0.000272
Train Epoch: 2 [55200/60000 (93.7755%)] Loss: 0.000018
Train Epoch: 2 [55600/60000 (94.4550%)] Loss: 0.000365
Train Epoch: 2 [56000/60000 (95.1345%)] Loss: 0.000104
Train Epoch: 2 [56400/60000 (95.8141%)] Loss: 0.000178
Train Epoch: 2 [56800/60000 (96.4936%)] Loss: 0.000152
Train Epoch: 2 [57200/60000 (97.1731%)] Loss: 0.000438
Train Epoch: 2 [57600/60000 (97.8527%)] Loss: 0.015318
Train Epoch: 2 [58000/60000 (98.5322%)] Loss: 0.013655
Train Epoch: 2 [58400/60000 (99.2117%)] Loss: 0.000294
Train Epoch: 2 [58800/60000 (99.8913%)] Loss: 0.003506
Test set: Average loss: 0.0595, Accuracy: 9821/10000 (98.2100%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.002790
Train Epoch: 3 [400/60000 (0.6795%)]    Loss: 0.019622
Train Epoch: 3 [800/60000 (1.3591%)]    Loss: 0.000023
Train Epoch: 3 [1200/60000 (2.0386%)]   Loss: 0.000008
Train Epoch: 3 [1600/60000 (2.7181%)]   Loss: 0.005307
Train Epoch: 3 [2000/60000 (3.3977%)]   Loss: 0.001329
Train Epoch: 3 [2400/60000 (4.0772%)]   Loss: 0.000006
Train Epoch: 3 [2800/60000 (4.7567%)]   Loss: 0.001876
Train Epoch: 3 [3200/60000 (5.4363%)]   Loss: 0.007237
Train Epoch: 3 [3600/60000 (6.1158%)]   Loss: 0.000120
Train Epoch: 3 [4000/60000 (6.7953%)]   Loss: 0.000072
Train Epoch: 3 [4400/60000 (7.4749%)]   Loss: 0.000025
Train Epoch: 3 [4800/60000 (8.1544%)]   Loss: 0.005244
Train Epoch: 3 [5200/60000 (8.8339%)]   Loss: 0.023808
Train Epoch: 3 [5600/60000 (9.5135%)]   Loss: 0.000575
Train Epoch: 3 [6000/60000 (10.1930%)]  Loss: 0.000063
Train Epoch: 3 [6400/60000 (10.8725%)]  Loss: 0.000140
Train Epoch: 3 [6800/60000 (11.5521%)]  Loss: 0.003346
Train Epoch: 3 [7200/60000 (12.2316%)]  Loss: 0.001633
Train Epoch: 3 [7600/60000 (12.9111%)]  Loss: 0.000049
Train Epoch: 3 [8000/60000 (13.5906%)]  Loss: 0.001208
Train Epoch: 3 [8400/60000 (14.2702%)]  Loss: 0.000053
Train Epoch: 3 [8800/60000 (14.9497%)]  Loss: 0.000099
Train Epoch: 3 [9200/60000 (15.6292%)]  Loss: 0.312926
Train Epoch: 3 [9600/60000 (16.3088%)]  Loss: 0.002243
Train Epoch: 3 [10000/60000 (16.9883%)] Loss: 0.000172
Train Epoch: 3 [10400/60000 (17.6678%)] Loss: 0.000032
Train Epoch: 3 [10800/60000 (18.3474%)] Loss: 0.102946
Train Epoch: 3 [11200/60000 (19.0269%)] Loss: 0.006078
Train Epoch: 3 [11600/60000 (19.7064%)] Loss: 0.000019
Train Epoch: 3 [12000/60000 (20.3860%)] Loss: 0.001144
Train Epoch: 3 [12400/60000 (21.0655%)] Loss: 0.002339
Train Epoch: 3 [12800/60000 (21.7450%)] Loss: 0.001117
Train Epoch: 3 [13200/60000 (22.4246%)] Loss: 0.000411
```

```
Train Epoch: 3 [14000/60000 (23.7836%)] Loss: 0.000116
Train Epoch: 3 [14400/60000 (24.4632%)] Loss: 0.000076
Train Epoch: 3 [14800/60000 (25.1427%)] Loss: 0.000035
Train Epoch: 3 [15200/60000 (25.8222%)] Loss: 0.000308
Train Epoch: 3 [15600/60000 (26.5018%)] Loss: 0.000020
Train Epoch: 3 [16000/60000 (27.1813%)] Loss: 0.012687
Train Epoch: 3 [16400/60000 (27.8608%)] Loss: 0.000670
Train Epoch: 3 [16800/60000 (28.5404%)] Loss: 0.000021
Train Epoch: 3 [17200/60000 (29.2199%)] Loss: 0.000484
Train Epoch: 3 [17600/60000 (29.8994%)] Loss: 0.000138
Train Epoch: 3 [18000/60000 (30.5790%)] Loss: 0.000099
Train Epoch: 3 [18400/60000 (31.2585%)] Loss: 0.046449
Train Epoch: 3 [18800/60000 (31.9380%)] Loss: 0.000068
Train Epoch: 3 [19200/60000 (32.6176%)] Loss: 0.004221
Train Epoch: 3 [19600/60000 (33.2971%)] Loss: 0.001287
Train Epoch: 3 [20000/60000 (33.9766%)] Loss: 0.000134
Train Epoch: 3 [20400/60000 (34.6562%)] Loss: 0.000639
Train Epoch: 3 [20800/60000 (35.3357%)] Loss: 0.000034
Train Epoch: 3 [21200/60000 (36.0152%)] Loss: 0.000150
Train Epoch: 3 [21600/60000 (36.6948%)] Loss: 0.024966
Train Epoch: 3 [22000/60000 (37.3743%)] Loss: 0.000012
Train Epoch: 3 [22400/60000 (38.0538%)] Loss: 0.000102
Train Epoch: 3 [22800/60000 (38.7334%)] Loss: 0.000040
Train Epoch: 3 [23200/60000 (39.4129%)] Loss: 0.000839
Train Epoch: 3 [23600/60000 (40.0924%)] Loss: 0.000704
Train Epoch: 3 [24000/60000 (40.7719%)] Loss: 0.000011
Train Epoch: 3 [24400/60000 (41.4515%)] Loss: 0.000137
Train Epoch: 3 [24800/60000 (42.1310%)] Loss: 0.000118
Train Epoch: 3 [25200/60000 (42.8105%)] Loss: 1.503516
Train Epoch: 3 [25600/60000 (43.4901%)] Loss: 0.000277
Train Epoch: 3 [26000/60000 (44.1696%)] Loss: 0.006512
Train Epoch: 3 [26400/60000 (44.8491%)] Loss: 0.001013
Train Epoch: 3 [26800/60000 (45.5287%)] Loss: 0.000169
Train Epoch: 3 [27200/60000 (46.2082%)] Loss: 0.000505
Train Epoch: 3 [27600/60000 (46.8877%)] Loss: 0.062491
Train Epoch: 3 [28000/60000 (47.5673%)] Loss: 0.000185
Train Epoch: 3 [28400/60000 (48.2468%)] Loss: 0.000041
Train Epoch: 3 [28800/60000 (48.9263%)] Loss: 0.000092
Train Epoch: 3 [29200/60000 (49.6059%)] Loss: 0.000888
Train Epoch: 3 [29600/60000 (50.2854%)] Loss: 0.000038
Train Epoch: 3 [30000/60000 (50.9649%)] Loss: 0.004133
Train Epoch: 3 [30400/60000 (51.6445%)] Loss: 0.001023
Train Epoch: 3 [30800/60000 (52.3240%)] Loss: 0.000001
Train Epoch: 3 [31200/60000 (53.0035%)] Loss: 0.000247
Train Epoch: 3 [31600/60000 (53.6831%)] Loss: 0.012340
Train Epoch: 3 [32000/60000 (54.3626%)] Loss: 0.000433
Train Epoch: 3 [32400/60000 (55.0421%)] Loss: 0.000068
Train Epoch: 3 [32800/60000 (55.7217%)] Loss: 0.029284
Train Epoch: 3 [33200/60000 (56.4012%)] Loss: 0.000321
Train Epoch: 3 [33600/60000 (57.0807%)] Loss: 0.000018
Train Epoch: 3 [34000/60000 (57.7603%)] Loss: 0.918150
Train Epoch: 3 [34400/60000 (58.4398%)] Loss: 0.000073
Train Epoch: 3 [34800/60000 (59.1193%)] Loss: 0.005613
```

```
Train Epoch: 3 [35600/60000 (60.4784%)] Loss: 0.000068
Train Epoch: 3 [36000/60000 (61.1579%)] Loss: 0.005158
Train Epoch: 3 [36400/60000 (61.8375%)] Loss: 0.012210
Train Epoch: 3 [36800/60000 (62.5170%)] Loss: 0.000000
Train Epoch: 3 [37200/60000 (63.1965%)] Loss: 0.009496
Train Epoch: 3 [37600/60000 (63.8761%)] Loss: 0.046893
Train Epoch: 3 [38000/60000 (64.5556%)] Loss: 0.000394
Train Epoch: 3 [38400/60000 (65.2351%)] Loss: 0.000000
Train Epoch: 3 [38800/60000 (65.9147%)] Loss: 0.000160
Train Epoch: 3 [39200/60000 (66.5942%)] Loss: 0.000043
Train Epoch: 3 [39600/60000 (67.2737%)] Loss: 0.000002
Train Epoch: 3 [40000/60000 (67.9532%)] Loss: 0.000038
Train Epoch: 3 [40400/60000 (68.6328%)] Loss: 0.000019
Train Epoch: 3 [40800/60000 (69.3123%)] Loss: 0.000000
Train Epoch: 3 [41200/60000 (69.9918%)] Loss: 0.000936
Train Epoch: 3 [41600/60000 (70.6714%)] Loss: 0.013424
Train Epoch: 3 [42000/60000 (71.3509%)] Loss: 0.000520
Train Epoch: 3 [42400/60000 (72.0304%)] Loss: 0.000610
Train Epoch: 3 [42800/60000 (72.7100%)] Loss: 0.000647
Train Epoch: 3 [43200/60000 (73.3895%)] Loss: 0.001428
Train Epoch: 3 [43600/60000 (74.0690%)] Loss: 0.022148
Train Epoch: 3 [44000/60000 (74.7486%)] Loss: 0.000016
Train Epoch: 3 [44400/60000 (75.4281%)] Loss: 0.512978
Train Epoch: 3 [44800/60000 (76.1076%)] Loss: 0.000020
Train Epoch: 3 [45200/60000 (76.7872%)] Loss: 0.000139
Train Epoch: 3 [45600/60000 (77.4667%)] Loss: 0.002212
Train Epoch: 3 [46000/60000 (78.1462%)] Loss: 0.000243
Train Epoch: 3 [46400/60000 (78.8258%)] Loss: 0.016892
Train Epoch: 3 [46800/60000 (79.5053%)] Loss: 0.000224
Train Epoch: 3 [47200/60000 (80.1848%)] Loss: 0.000110
Train Epoch: 3 [47600/60000 (80.8644%)] Loss: 0.008438
Train Epoch: 3 [48000/60000 (81.5439%)] Loss: 0.000005
Train Epoch: 3 [48400/60000 (82.2234%)] Loss: 0.000013
Train Epoch: 3 [48800/60000 (82.9030%)] Loss: 0.001900
Train Epoch: 3 [49200/60000 (83.5825%)] Loss: 0.002968
Train Epoch: 3 [49600/60000 (84.2620%)] Loss: 0.000100
Train Epoch: 3 [50000/60000 (84.9416%)] Loss: 0.000001
Train Epoch: 3 [50400/60000 (85.6211%)] Loss: 0.009021
Train Epoch: 3 [50800/60000 (86.3006%)] Loss: 0.000030
Train Epoch: 3 [51200/60000 (86.9802%)] Loss: 0.000013
Train Epoch: 3 [51600/60000 (87.6597%)] Loss: 0.009277
Train Epoch: 3 [52000/60000 (88.3392%)] Loss: 0.000055
Train Epoch: 3 [52400/60000 (89.0188%)] Loss: 0.001556
Train Epoch: 3 [52800/60000 (89.6983%)] Loss: 0.000240
Train Epoch: 3 [53200/60000 (90.3778%)] Loss: 0.007882
Train Epoch: 3 [53600/60000 (91.0574%)] Loss: 0.000009
Train Epoch: 3 [54000/60000 (91.7369%)] Loss: 0.033234
Train Epoch: 3 [54400/60000 (92.4164%)] Loss: 0.000105
Train Epoch: 3 [54800/60000 (93.0959%)] Loss: 0.000641
Train Epoch: 3 [55200/60000 (93.7755%)] Loss: 0.029179
Train Epoch: 3 [55600/60000 (94.4550%)] Loss: 0.001602
Train Epoch: 3 [56000/60000 (95.1345%)] Loss: 0.009772
Train Epoch: 3 [56400/60000 (95.8141%)] Loss: 0.001484
```

```
Train Epoch: 3 [56800/60000 (96.4956%)] Loss: 0.000296
Train Epoch: 3 [57200/60000 (97.1731%)] Loss: 0.000163
Train Epoch: 3 [57600/60000 (97.8527%)] Loss: 0.000020
Train Epoch: 3 [58000/60000 (98.5322%)] Loss: 0.000022
Train Epoch: 3 [58400/60000 (99.2117%)] Loss: 0.000030
Train Epoch: 3 [58800/60000 (99.8913%)] Loss: 0.315003
Test set: Average loss: 0.0426, Accuracy: 9881/10000 (98.8100%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.006962
Train Epoch: 4 [400/60000 (0.6795%)]    Loss: 0.000028
Train Epoch: 4 [800/60000 (1.3591%)]    Loss: 0.000411
Train Epoch: 4 [1200/60000 (2.0386%)]   Loss: 0.129438
Train Epoch: 4 [1600/60000 (2.7181%)]   Loss: 0.000376
Train Epoch: 4 [2000/60000 (3.3977%)]   Loss: 0.000001
Train Epoch: 4 [2400/60000 (4.0772%)]   Loss: 0.007022
Train Epoch: 4 [2800/60000 (4.7567%)]   Loss: 0.000006
Train Epoch: 4 [3200/60000 (5.4363%)]   Loss: 0.000833
Train Epoch: 4 [3600/60000 (6.1158%)]   Loss: 0.000372
Train Epoch: 4 [4000/60000 (6.7953%)]   Loss: 0.000689
Train Epoch: 4 [4400/60000 (7.4749%)]   Loss: 0.000010
Train Epoch: 4 [4800/60000 (8.1544%)]   Loss: 0.000001
Train Epoch: 4 [5200/60000 (8.8339%)]   Loss: 0.000202
Train Epoch: 4 [5600/60000 (9.5135%)]   Loss: 0.002634
Train Epoch: 4 [6000/60000 (10.1930%)]  Loss: 0.006649
Train Epoch: 4 [6400/60000 (10.8725%)]  Loss: 0.004506
Train Epoch: 4 [6800/60000 (11.5521%)]  Loss: 0.000349
Train Epoch: 4 [7200/60000 (12.2316%)]  Loss: 0.000103
Train Epoch: 4 [7600/60000 (12.9111%)]  Loss: 0.000016
Train Epoch: 4 [8000/60000 (13.5906%)]  Loss: 0.000065
Train Epoch: 4 [8400/60000 (14.2702%)]  Loss: 0.003256
Train Epoch: 4 [8800/60000 (14.9497%)]  Loss: 0.001329
Train Epoch: 4 [9200/60000 (15.6292%)]  Loss: 0.002650
Train Epoch: 4 [9600/60000 (16.3088%)]  Loss: 0.000083
Train Epoch: 4 [10000/60000 (16.9883%)] Loss: 0.000015
Train Epoch: 4 [10400/60000 (17.6678%)] Loss: 0.001386
Train Epoch: 4 [10800/60000 (18.3474%)] Loss: 0.002455
Train Epoch: 4 [11200/60000 (19.0269%)] Loss: 0.000172
Train Epoch: 4 [11600/60000 (19.7064%)] Loss: 0.000114
Train Epoch: 4 [12000/60000 (20.3860%)] Loss: 0.000586
Train Epoch: 4 [12400/60000 (21.0655%)] Loss: 0.003994
Train Epoch: 4 [12800/60000 (21.7450%)] Loss: 0.000124
Train Epoch: 4 [13200/60000 (22.4246%)] Loss: 0.280086
Train Epoch: 4 [13600/60000 (23.1041%)] Loss: 0.009684
Train Epoch: 4 [14000/60000 (23.7836%)] Loss: 0.000006
Train Epoch: 4 [14400/60000 (24.4632%)] Loss: 0.000016
Train Epoch: 4 [14800/60000 (25.1427%)] Loss: 0.018477
Train Epoch: 4 [15200/60000 (25.8222%)] Loss: 0.000045
Train Epoch: 4 [15600/60000 (26.5018%)] Loss: 0.000045
Train Epoch: 4 [16000/60000 (27.1813%)] Loss: 0.000016
Train Epoch: 4 [16400/60000 (27.8608%)] Loss: 0.000006
Train Epoch: 4 [16800/60000 (28.5404%)] Loss: 0.000203
Train Epoch: 4 [17200/60000 (29.2199%)] Loss: 0.000353
Train Epoch: 4 [17600/60000 (29.8994%)] Loss: 0.000167
Train Epoch: 4 [18000/60000 (30.5790%)] Loss: 0.000010
Train Epoch: 4 [18400/60000 (31.2585%)] Loss: 0.000240
Train Epoch: 4 [18800/60000 (31.9380%)] Loss: 0.000113
```

```
Train Epoch: 4 [18800/60000 (31.9380%)] Loss: 0.000113
Train Epoch: 4 [19200/60000 (32.6176%)] Loss: 0.000017
Train Epoch: 4 [19600/60000 (33.2971%)] Loss: 0.000110
Train Epoch: 4 [20000/60000 (33.9766%)] Loss: 0.000306
Train Epoch: 4 [20400/60000 (34.6562%)] Loss: 0.000002
Train Epoch: 4 [20800/60000 (35.3357%)] Loss: 0.000000
Train Epoch: 4 [21200/60000 (36.0152%)] Loss: 0.000013
Train Epoch: 4 [21600/60000 (36.6948%)] Loss: 0.000103
Train Epoch: 4 [22000/60000 (37.3743%)] Loss: 0.000057
Train Epoch: 4 [22400/60000 (38.0538%)] Loss: 0.000002
Train Epoch: 4 [22800/60000 (38.7334%)] Loss: 0.011849
Train Epoch: 4 [23200/60000 (39.4129%)] Loss: 0.000001
Train Epoch: 4 [23600/60000 (40.0924%)] Loss: 0.000019
Train Epoch: 4 [24000/60000 (40.7719%)] Loss: 0.000015
Train Epoch: 4 [24400/60000 (41.4515%)] Loss: 0.000950
Train Epoch: 4 [24800/60000 (42.1310%)] Loss: 0.001721
Train Epoch: 4 [25200/60000 (42.8105%)] Loss: 0.000017
Train Epoch: 4 [25600/60000 (43.4901%)] Loss: 0.000613
Train Epoch: 4 [26000/60000 (44.1696%)] Loss: 0.000045
Train Epoch: 4 [26400/60000 (44.8491%)] Loss: 0.000002
Train Epoch: 4 [26800/60000 (45.5287%)] Loss: 0.078734
Train Epoch: 4 [27200/60000 (46.2082%)] Loss: 0.000080
Train Epoch: 4 [27600/60000 (46.8877%)] Loss: 0.000683
Train Epoch: 4 [28000/60000 (47.5673%)] Loss: 0.000766
Train Epoch: 4 [28400/60000 (48.2468%)] Loss: 0.000009
Train Epoch: 4 [28800/60000 (48.9263%)] Loss: 0.000729
Train Epoch: 4 [29200/60000 (49.6059%)] Loss: 0.016346
Train Epoch: 4 [29600/60000 (50.2854%)] Loss: 0.000126
Train Epoch: 4 [30000/60000 (50.9649%)] Loss: 0.000943
Train Epoch: 4 [30400/60000 (51.6445%)] Loss: 0.000015
Train Epoch: 4 [30800/60000 (52.3240%)] Loss: 0.000433
Train Epoch: 4 [31200/60000 (53.0035%)] Loss: 0.000002
Train Epoch: 4 [31600/60000 (53.6831%)] Loss: 0.000011
Train Epoch: 4 [32000/60000 (54.3626%)] Loss: 0.000296
Train Epoch: 4 [32400/60000 (55.0421%)] Loss: 0.000158
Train Epoch: 4 [32800/60000 (55.7217%)] Loss: 0.000366
Train Epoch: 4 [33200/60000 (56.4012%)] Loss: 0.000133
Train Epoch: 4 [33600/60000 (57.0807%)] Loss: 0.003022
Train Epoch: 4 [34000/60000 (57.7603%)] Loss: 0.000002
Train Epoch: 4 [34400/60000 (58.4398%)] Loss: 0.000065
Train Epoch: 4 [34800/60000 (59.1193%)] Loss: 0.000078
Train Epoch: 4 [35200/60000 (59.7989%)] Loss: 0.000106
Train Epoch: 4 [35600/60000 (60.4784%)] Loss: 0.000472
Train Epoch: 4 [36000/60000 (61.1579%)] Loss: 0.000937
Train Epoch: 4 [36400/60000 (61.8375%)] Loss: 0.000219
Train Epoch: 4 [36800/60000 (62.5170%)] Loss: 0.009340
Train Epoch: 4 [37200/60000 (63.1965%)] Loss: 0.201161
Train Epoch: 4 [37600/60000 (63.8761%)] Loss: 0.001434
Train Epoch: 4 [38000/60000 (64.5556%)] Loss: 0.000006
Train Epoch: 4 [38400/60000 (65.2351%)] Loss: 0.000262
Train Epoch: 4 [38800/60000 (65.9147%)] Loss: 0.000376
Train Epoch: 4 [39200/60000 (66.5942%)] Loss: 0.000133
Train Epoch: 4 [39600/60000 (67.2737%)] Loss: 0.000175
Train Epoch: 4 [40000/60000 (67.9532%)] Loss: 0.000008
```

```
Train Epoch: 4 [40400/60000 (68.6328%)] Loss: 0.000028
Train Epoch: 4 [40800/60000 (69.3123%)] Loss: 0.000028
Train Epoch: 4 [41200/60000 (69.9918%)] Loss: 0.000318
Train Epoch: 4 [41600/60000 (70.6714%)] Loss: 0.000002
Train Epoch: 4 [42000/60000 (71.3509%)] Loss: 0.003464
Train Epoch: 4 [42400/60000 (72.0304%)] Loss: 0.000059
Train Epoch: 4 [42800/60000 (72.7100%)] Loss: 0.001116
Train Epoch: 4 [43200/60000 (73.3895%)] Loss: 0.000011
Train Epoch: 4 [43600/60000 (74.0690%)] Loss: 0.000002
Train Epoch: 4 [44000/60000 (74.7486%)] Loss: 0.149901
Train Epoch: 4 [44400/60000 (75.4281%)] Loss: 0.009947
Train Epoch: 4 [44800/60000 (76.1076%)] Loss: 0.000047
Train Epoch: 4 [45200/60000 (76.7872%)] Loss: 0.000031
Train Epoch: 4 [45600/60000 (77.4667%)] Loss: 0.004406
Train Epoch: 4 [46000/60000 (78.1462%)] Loss: 0.000328
Train Epoch: 4 [46400/60000 (78.8258%)] Loss: 0.004664
Train Epoch: 4 [46800/60000 (79.5053%)] Loss: 0.000017
Train Epoch: 4 [47200/60000 (80.1848%)] Loss: 0.000001
Train Epoch: 4 [47600/60000 (80.8644%)] Loss: 0.002432
Train Epoch: 4 [48000/60000 (81.5439%)] Loss: 0.000912
Train Epoch: 4 [48400/60000 (82.2234%)] Loss: 0.002026
Train Epoch: 4 [48800/60000 (82.9030%)] Loss: 0.044360
Train Epoch: 4 [49200/60000 (83.5825%)] Loss: 0.000598
Train Epoch: 4 [49600/60000 (84.2620%)] Loss: 0.000040
Train Epoch: 4 [50000/60000 (84.9416%)] Loss: 0.000040
Train Epoch: 4 [50400/60000 (85.6211%)] Loss: 0.002148
Train Epoch: 4 [50800/60000 (86.3006%)] Loss: 0.000036
Train Epoch: 4 [51200/60000 (86.9802%)] Loss: 0.011457
Train Epoch: 4 [51600/60000 (87.6597%)] Loss: 0.000253
Train Epoch: 4 [52000/60000 (88.3392%)] Loss: 0.059662
Train Epoch: 4 [52400/60000 (89.0188%)] Loss: 0.000096
Train Epoch: 4 [52800/60000 (89.6983%)] Loss: 0.000002
Train Epoch: 4 [53200/60000 (90.3778%)] Loss: 0.000071
Train Epoch: 4 [53600/60000 (91.0574%)] Loss: 0.000067
Train Epoch: 4 [54000/60000 (91.7369%)] Loss: 0.005782
Train Epoch: 4 [54400/60000 (92.4164%)] Loss: 0.002722
Train Epoch: 4 [54800/60000 (93.0959%)] Loss: 0.000003
Train Epoch: 4 [55200/60000 (93.7755%)] Loss: 0.018324
Train Epoch: 4 [55600/60000 (94.4550%)] Loss: 0.042418
Train Epoch: 4 [56000/60000 (95.1345%)] Loss: 0.000001
Train Epoch: 4 [56400/60000 (95.8141%)] Loss: 0.059050
Train Epoch: 4 [56800/60000 (96.4936%)] Loss: 0.000150
Train Epoch: 4 [57200/60000 (97.1731%)] Loss: 0.000192
Train Epoch: 4 [57600/60000 (97.8527%)] Loss: 0.004482
Train Epoch: 4 [58000/60000 (98.5322%)] Loss: 0.000120
Train Epoch: 4 [58400/60000 (99.2117%)] Loss: 0.000251
Train Epoch: 4 [58800/60000 (99.8913%)] Loss: 0.000042
Test set: Average loss: 0.0480, Accuracy: 9850/10000 (98.5000%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.000464
Train Epoch: 5 [400/60000 (0.6795%)]    Loss: 0.000089
Train Epoch: 5 [800/60000 (1.3591%)]    Loss: 0.000022
Train Epoch: 5 [1200/60000 (2.0386%)]   Loss: 0.015919
Train Epoch: 5 [1600/60000 (2.7181%)]   Loss: 0.000546
Train Epoch: 5 [2000/60000 (3.3977%)]   Loss: 0.000058
```

```
Train Epoch: 5 [2400/60000 (4.0772%)]   Loss: 0.000108
Train Epoch: 5 [2800/60000 (4.7567%)]   Loss: 0.000065
Train Epoch: 5 [3200/60000 (5.4363%)]   Loss: 0.008936
Train Epoch: 5 [3600/60000 (6.1158%)]   Loss: 0.000018
Train Epoch: 5 [4000/60000 (6.7953%)]   Loss: 0.000514
Train Epoch: 5 [4400/60000 (7.4749%)]   Loss: 0.000012
Train Epoch: 5 [4800/60000 (8.1544%)]   Loss: 0.000006
Train Epoch: 5 [5200/60000 (8.8339%)]   Loss: 0.032446
Train Epoch: 5 [5600/60000 (9.5135%)]   Loss: 0.000246
Train Epoch: 5 [6000/60000 (10.1930%)]  Loss: 0.000482
Train Epoch: 5 [6400/60000 (10.8725%)]  Loss: 0.006483
Train Epoch: 5 [6800/60000 (11.5521%)]  Loss: 0.000003
Train Epoch: 5 [7200/60000 (12.2316%)]  Loss: 0.000014
Train Epoch: 5 [7600/60000 (12.9111%)]  Loss: 0.003689
Train Epoch: 5 [8000/60000 (13.5906%)]  Loss: 0.000484
Train Epoch: 5 [8400/60000 (14.2702%)]  Loss: 0.005960
Train Epoch: 5 [8800/60000 (14.9497%)]  Loss: 0.000007
Train Epoch: 5 [9200/60000 (15.6292%)]  Loss: 0.342428
Train Epoch: 5 [9600/60000 (16.3088%)]  Loss: 0.052121
Train Epoch: 5 [10000/60000 (16.9883%)] Loss: 0.000334
Train Epoch: 5 [10400/60000 (17.6678%)] Loss: 0.000015
Train Epoch: 5 [10800/60000 (18.3474%)] Loss: 0.000002
Train Epoch: 5 [11200/60000 (19.0269%)] Loss: 0.000056
Train Epoch: 5 [11600/60000 (19.7064%)] Loss: 0.000023
Train Epoch: 5 [12000/60000 (20.3860%)] Loss: 0.000000
Train Epoch: 5 [12400/60000 (21.0655%)] Loss: 0.000140
Train Epoch: 5 [12800/60000 (21.7450%)] Loss: 0.000008
Train Epoch: 5 [13200/60000 (22.4246%)] Loss: 0.000095
Train Epoch: 5 [13600/60000 (23.1041%)] Loss: 0.002289
Train Epoch: 5 [14000/60000 (23.7836%)] Loss: 0.000165
Train Epoch: 5 [14400/60000 (24.4632%)] Loss: 0.000003
Train Epoch: 5 [14800/60000 (25.1427%)] Loss: 0.000010
Train Epoch: 5 [15200/60000 (25.8222%)] Loss: 0.000002
Train Epoch: 5 [15600/60000 (26.5018%)] Loss: 0.000079
Train Epoch: 5 [16000/60000 (27.1813%)] Loss: 0.000272
Train Epoch: 5 [16400/60000 (27.8608%)] Loss: 0.052414
Train Epoch: 5 [16800/60000 (28.5404%)] Loss: 0.000144
Train Epoch: 5 [17200/60000 (29.2199%)] Loss: 0.000085
Train Epoch: 5 [17600/60000 (29.8994%)] Loss: 0.000026
Train Epoch: 5 [18000/60000 (30.5790%)] Loss: 0.000011
Train Epoch: 5 [18400/60000 (31.2585%)] Loss: 0.000004
Train Epoch: 5 [18800/60000 (31.9380%)] Loss: 0.000253
Train Epoch: 5 [19200/60000 (32.6176%)] Loss: 0.000001
Train Epoch: 5 [19600/60000 (33.2971%)] Loss: 0.002933
Train Epoch: 5 [20000/60000 (33.9766%)] Loss: 0.000113
Train Epoch: 5 [20400/60000 (34.6562%)] Loss: 0.000039
Train Epoch: 5 [20800/60000 (35.3357%)] Loss: 0.000011
Train Epoch: 5 [21200/60000 (36.0152%)] Loss: 0.000001
Train Epoch: 5 [21600/60000 (36.6948%)] Loss: 0.000002
Train Epoch: 5 [22000/60000 (37.3743%)] Loss: 0.000001
Train Epoch: 5 [22400/60000 (38.0538%)] Loss: 0.006494
Train Epoch: 5 [22800/60000 (38.7334%)] Loss: 0.000011
Train Epoch: 5 [23200/60000 (39.4129%)] Loss: 0.000689
Train Epoch: 5 [23600/60000 (40.0924%)] Loss: 0.000020
```

```
Train Epoch: 5 [24000/60000 (40.7719%)] Loss: 0.000013
Train Epoch: 5 [24400/60000 (41.4515%)] Loss: 0.000006
Train Epoch: 5 [24800/60000 (42.1310%)] Loss: 0.000795
Train Epoch: 5 [25200/60000 (42.8105%)] Loss: 0.000043
Train Epoch: 5 [25600/60000 (43.4901%)] Loss: 0.000020
Train Epoch: 5 [26000/60000 (44.1696%)] Loss: 0.000010
Train Epoch: 5 [26400/60000 (44.8491%)] Loss: 0.000003
Train Epoch: 5 [26800/60000 (45.5287%)] Loss: 0.002571
Train Epoch: 5 [27200/60000 (46.2082%)] Loss: 0.000003
Train Epoch: 5 [27600/60000 (46.8877%)] Loss: 0.000007
Train Epoch: 5 [28000/60000 (47.5673%)] Loss: 0.000044
Train Epoch: 5 [28400/60000 (48.2468%)] Loss: 0.000000
Train Epoch: 5 [28800/60000 (48.9263%)] Loss: 0.000022
Train Epoch: 5 [29200/60000 (49.6059%)] Loss: 0.000001
Train Epoch: 5 [29600/60000 (50.2854%)] Loss: 0.000025
Train Epoch: 5 [30000/60000 (50.9649%)] Loss: 0.011529
Train Epoch: 5 [30400/60000 (51.6445%)] Loss: 0.000479
Train Epoch: 5 [30800/60000 (52.3240%)] Loss: 0.000013
Train Epoch: 5 [31200/60000 (53.0035%)] Loss: 0.000006
Train Epoch: 5 [31600/60000 (53.6831%)] Loss: 0.000060
Train Epoch: 5 [32000/60000 (54.3626%)] Loss: 0.000008
Train Epoch: 5 [32400/60000 (55.0421%)] Loss: 0.009684
Train Epoch: 5 [32800/60000 (55.7217%)] Loss: 0.000015
Train Epoch: 5 [33200/60000 (56.4012%)] Loss: 0.000059
Train Epoch: 5 [33600/60000 (57.0807%)] Loss: 0.000014
Train Epoch: 5 [34000/60000 (57.7603%)] Loss: 0.000038
Train Epoch: 5 [34400/60000 (58.4398%)] Loss: 0.001348
Train Epoch: 5 [34800/60000 (59.1193%)] Loss: 0.000471
Train Epoch: 5 [35200/60000 (59.7989%)] Loss: 0.000003
Train Epoch: 5 [35600/60000 (60.4784%)] Loss: 0.000001
Train Epoch: 5 [36000/60000 (61.1579%)] Loss: 0.001202
Train Epoch: 5 [36400/60000 (61.8375%)] Loss: 0.002920
Train Epoch: 5 [36800/60000 (62.5170%)] Loss: 0.000006
Train Epoch: 5 [37200/60000 (63.1965%)] Loss: 0.002370
Train Epoch: 5 [37600/60000 (63.8761%)] Loss: 0.016069
Train Epoch: 5 [38000/60000 (64.5556%)] Loss: 0.000708
Train Epoch: 5 [38400/60000 (65.2351%)] Loss: 0.000018
Train Epoch: 5 [38800/60000 (65.9147%)] Loss: 0.000004
Train Epoch: 5 [39200/60000 (66.5942%)] Loss: 0.000043
Train Epoch: 5 [39600/60000 (67.2737%)] Loss: 0.000003
Train Epoch: 5 [40000/60000 (67.9532%)] Loss: 0.001596
Train Epoch: 5 [40400/60000 (68.6328%)] Loss: 0.000319
Train Epoch: 5 [40800/60000 (69.3123%)] Loss: 0.000218
Train Epoch: 5 [41200/60000 (69.9918%)] Loss: 0.001607
Train Epoch: 5 [41600/60000 (70.6714%)] Loss: 0.000110
Train Epoch: 5 [42000/60000 (71.3509%)] Loss: 0.000002
Train Epoch: 5 [42400/60000 (72.0304%)] Loss: 0.000080
Train Epoch: 5 [42800/60000 (72.7100%)] Loss: 0.000177
Train Epoch: 5 [43200/60000 (73.3895%)] Loss: 0.000483
Train Epoch: 5 [43600/60000 (74.0690%)] Loss: 0.000006
Train Epoch: 5 [44000/60000 (74.7486%)] Loss: 0.235895
Train Epoch: 5 [44400/60000 (75.4281%)] Loss: 0.000032
Train Epoch: 5 [44800/60000 (76.1076%)] Loss: 0.000006
Train Epoch: 5 [45200/60000 (76.7872%)] Loss: 0.000603
```

```
Train Epoch: 5 [45600/60000 (77.4667%)] Loss: 0.000466
Train Epoch: 5 [46000/60000 (78.1462%)] Loss: 0.000019
Train Epoch: 5 [46400/60000 (78.8258%)] Loss: 0.000171
Train Epoch: 5 [46800/60000 (79.5053%)] Loss: 0.000422
Train Epoch: 5 [47200/60000 (80.1848%)] Loss: 0.000213
Train Epoch: 5 [47600/60000 (80.8644%)] Loss: 0.003529
Train Epoch: 5 [48000/60000 (81.5439%)] Loss: 0.000000
Train Epoch: 5 [48400/60000 (82.2234%)] Loss: 0.000634
Train Epoch: 5 [48800/60000 (82.9030%)] Loss: 0.000033
Train Epoch: 5 [49200/60000 (83.5825%)] Loss: 0.000160
Train Epoch: 5 [49600/60000 (84.2620%)] Loss: 0.000003
Train Epoch: 5 [50000/60000 (84.9416%)] Loss: 0.000034
Train Epoch: 5 [50400/60000 (85.6211%)] Loss: 0.000000
Train Epoch: 5 [50800/60000 (86.3006%)] Loss: 0.000004
Train Epoch: 5 [51200/60000 (86.9802%)] Loss: 0.000005
Train Epoch: 5 [51600/60000 (87.6597%)] Loss: 0.000078
Train Epoch: 5 [52000/60000 (88.3392%)] Loss: 0.000327
Train Epoch: 5 [52400/60000 (89.0188%)] Loss: 0.000006
Train Epoch: 5 [52800/60000 (89.6983%)] Loss: 0.001028
Train Epoch: 5 [53200/60000 (90.3778%)] Loss: 0.001761
Train Epoch: 5 [53600/60000 (91.0574%)] Loss: 0.132693
Train Epoch: 5 [54000/60000 (91.7369%)] Loss: 0.000179
Train Epoch: 5 [54400/60000 (92.4164%)] Loss: 0.000000
Train Epoch: 5 [54800/60000 (93.0959%)] Loss: 0.000154
Train Epoch: 5 [55200/60000 (93.7755%)] Loss: 0.000138
Train Epoch: 5 [55600/60000 (94.4550%)] Loss: 0.050468
Train Epoch: 5 [56000/60000 (95.1345%)] Loss: 0.000014
Train Epoch: 5 [56400/60000 (95.8141%)] Loss: 0.000001
Train Epoch: 5 [56800/60000 (96.4936%)] Loss: 0.000008
Train Epoch: 5 [57200/60000 (97.1731%)] Loss: 0.000579
Train Epoch: 5 [57600/60000 (97.8527%)] Loss: 0.000007
Train Epoch: 5 [58000/60000 (98.5322%)] Loss: 0.000028
Train Epoch: 5 [58400/60000 (99.2117%)] Loss: 0.000007
Train Epoch: 5 [58800/60000 (99.8913%)] Loss: 0.000611
Test set: Average loss: 0.0479, Accuracy: 9880/10000 (98.8000%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.000052
Train Epoch: 6 [400/60000 (0.6795%)]    Loss: 0.000001
Train Epoch: 6 [800/60000 (1.3591%)]    Loss: 0.000009
Train Epoch: 6 [1200/60000 (2.0386%)]   Loss: 0.000000
Train Epoch: 6 [1600/60000 (2.7181%)]   Loss: 0.000088
Train Epoch: 6 [2000/60000 (3.3977%)]   Loss: 0.000072
Train Epoch: 6 [2400/60000 (4.0772%)]   Loss: 0.000011
Train Epoch: 6 [2800/60000 (4.7567%)]   Loss: 0.000022
Train Epoch: 6 [3200/60000 (5.4363%)]   Loss: 0.000152
Train Epoch: 6 [3600/60000 (6.1158%)]   Loss: 0.002245
Train Epoch: 6 [4000/60000 (6.7953%)]   Loss: 0.000081
Train Epoch: 6 [4400/60000 (7.4749%)]   Loss: 0.000963
Train Epoch: 6 [4800/60000 (8.1544%)]   Loss: 0.000010
Train Epoch: 6 [5200/60000 (8.8339%)]   Loss: 0.022805
Train Epoch: 6 [5600/60000 (9.5135%)]   Loss: 0.003749
Train Epoch: 6 [6000/60000 (10.1930%)]  Loss: 0.057309
Train Epoch: 6 [6400/60000 (10.8725%)]  Loss: 0.003789
Train Epoch: 6 [6800/60000 (11.5521%)]  Loss: 0.001278
Train Epoch: 6 [7200/60000 (12.2316%)]  Loss: 0.000006
```

```
Train Epoch: 6 [7600/60000 (12.9111%)]  Loss: 0.001202
Train Epoch: 6 [8000/60000 (13.5906%)]  Loss: 0.000036
Train Epoch: 6 [8400/60000 (14.2702%)]  Loss: 0.000032
Train Epoch: 6 [8800/60000 (14.9497%)]  Loss: 0.000496
Train Epoch: 6 [9200/60000 (15.6292%)]  Loss: 0.000001
Train Epoch: 6 [9600/60000 (16.3088%)]  Loss: 0.000087
Train Epoch: 6 [10000/60000 (16.9883%)] Loss: 0.001362
Train Epoch: 6 [10400/60000 (17.6678%)] Loss: 0.001781
Train Epoch: 6 [10800/60000 (18.3474%)] Loss: 0.000001
Train Epoch: 6 [11200/60000 (19.0269%)] Loss: 0.000003
Train Epoch: 6 [11600/60000 (19.7064%)] Loss: 0.000036
Train Epoch: 6 [12000/60000 (20.3860%)] Loss: 0.004657
Train Epoch: 6 [12400/60000 (21.0655%)] Loss: 0.000017
Train Epoch: 6 [12800/60000 (21.7450%)] Loss: 0.000232
Train Epoch: 6 [13200/60000 (22.4246%)] Loss: 0.000015
Train Epoch: 6 [13600/60000 (23.1041%)] Loss: 0.000012
Train Epoch: 6 [14000/60000 (23.7836%)] Loss: 0.002029
Train Epoch: 6 [14400/60000 (24.4632%)] Loss: 0.000098
Train Epoch: 6 [14800/60000 (25.1427%)] Loss: 0.104395
Train Epoch: 6 [15200/60000 (25.8222%)] Loss: 0.000024
Train Epoch: 6 [15600/60000 (26.5018%)] Loss: 0.000022
Train Epoch: 6 [16000/60000 (27.1813%)] Loss: 0.001340
Train Epoch: 6 [16400/60000 (27.8608%)] Loss: 0.000005
Train Epoch: 6 [16800/60000 (28.5404%)] Loss: 0.000286
Train Epoch: 6 [17200/60000 (29.2199%)] Loss: 0.000145
Train Epoch: 6 [17600/60000 (29.8994%)] Loss: 0.004400
Train Epoch: 6 [18000/60000 (30.5790%)] Loss: 0.000016
Train Epoch: 6 [18400/60000 (31.2585%)] Loss: 0.000017
Train Epoch: 6 [18800/60000 (31.9380%)] Loss: 0.005247
Train Epoch: 6 [19200/60000 (32.6176%)] Loss: 0.000001
Train Epoch: 6 [19600/60000 (33.2971%)] Loss: 0.001542
Train Epoch: 6 [20000/60000 (33.9766%)] Loss: 0.000078
Train Epoch: 6 [20400/60000 (34.6562%)] Loss: 0.461500
Train Epoch: 6 [20800/60000 (35.3357%)] Loss: 0.000015
Train Epoch: 6 [21200/60000 (36.0152%)] Loss: 0.000203
Train Epoch: 6 [21600/60000 (36.6948%)] Loss: 0.000015
Train Epoch: 6 [22000/60000 (37.3743%)] Loss: 0.000004
Train Epoch: 6 [22400/60000 (38.0538%)] Loss: 0.000007
Train Epoch: 6 [22800/60000 (38.7334%)] Loss: 0.000814
Train Epoch: 6 [23200/60000 (39.4129%)] Loss: 0.000024
Train Epoch: 6 [23600/60000 (40.0924%)] Loss: 0.003038
Train Epoch: 6 [24000/60000 (40.7719%)] Loss: 0.000044
Train Epoch: 6 [24400/60000 (41.4515%)] Loss: 0.000230
Train Epoch: 6 [24800/60000 (42.1310%)] Loss: 0.000006
Train Epoch: 6 [25200/60000 (42.8105%)] Loss: 0.000001
Train Epoch: 6 [25600/60000 (43.4901%)] Loss: 0.017339
Train Epoch: 6 [26000/60000 (44.1696%)] Loss: 0.000065
Train Epoch: 6 [26400/60000 (44.8491%)] Loss: 0.000055
Train Epoch: 6 [26800/60000 (45.5287%)] Loss: 0.000010
Train Epoch: 6 [27200/60000 (46.2082%)] Loss: 0.000143
Train Epoch: 6 [27600/60000 (46.8877%)] Loss: 0.000002
Train Epoch: 6 [28000/60000 (47.5673%)] Loss: 0.000001
Train Epoch: 6 [28400/60000 (48.2468%)] Loss: 0.000004
Train Epoch: 6 [28800/60000 (48.9263%)] Loss: 0.000011
```

```
Train Epoch: 6 [29200/60000 (49.6059%)] Loss: 0.000061
Train Epoch: 6 [29600/60000 (50.2854%)] Loss: 0.000010
Train Epoch: 6 [30000/60000 (50.9649%)] Loss: 0.000001
Train Epoch: 6 [30400/60000 (51.6445%)] Loss: 0.000097
Train Epoch: 6 [30800/60000 (52.3240%)] Loss: 0.000031
Train Epoch: 6 [31200/60000 (53.0035%)] Loss: 0.000002
Train Epoch: 6 [31600/60000 (53.6831%)] Loss: 0.000586
Train Epoch: 6 [32000/60000 (54.3626%)] Loss: 0.000184
Train Epoch: 6 [32400/60000 (55.0421%)] Loss: 0.006963
Train Epoch: 6 [32800/60000 (55.7217%)] Loss: 0.000019
Train Epoch: 6 [33200/60000 (56.4012%)] Loss: 0.000000
Train Epoch: 6 [33600/60000 (57.0807%)] Loss: 0.000000
Train Epoch: 6 [34000/60000 (57.7603%)] Loss: 0.000024
Train Epoch: 6 [34400/60000 (58.4398%)] Loss: 0.000027
Train Epoch: 6 [34800/60000 (59.1193%)] Loss: 0.000168
Train Epoch: 6 [35200/60000 (59.7989%)] Loss: 0.000019
Train Epoch: 6 [35600/60000 (60.4784%)] Loss: 0.000020
Train Epoch: 6 [36000/60000 (61.1579%)] Loss: 0.000609
Train Epoch: 6 [36400/60000 (61.8375%)] Loss: 0.004965
Train Epoch: 6 [36800/60000 (62.5170%)] Loss: 0.000105
Train Epoch: 6 [37200/60000 (63.1965%)] Loss: 0.000105
Train Epoch: 6 [37600/60000 (63.8761%)] Loss: 0.000001
Train Epoch: 6 [38000/60000 (64.5556%)] Loss: 0.000406
Train Epoch: 6 [38400/60000 (65.2351%)] Loss: 0.000004
Train Epoch: 6 [38800/60000 (65.9147%)] Loss: 0.000007
Train Epoch: 6 [39200/60000 (66.5942%)] Loss: 0.005360
Train Epoch: 6 [39600/60000 (67.2737%)] Loss: 0.000004
Train Epoch: 6 [40000/60000 (67.9532%)] Loss: 0.000383
Train Epoch: 6 [40400/60000 (68.6328%)] Loss: 0.000040
Train Epoch: 6 [40800/60000 (69.3123%)] Loss: 0.000005
Train Epoch: 6 [41200/60000 (69.9918%)] Loss: 0.674963
Train Epoch: 6 [41600/60000 (70.6714%)] Loss: 0.000071
Train Epoch: 6 [42000/60000 (71.3509%)] Loss: 0.000004
Train Epoch: 6 [42400/60000 (72.0304%)] Loss: 0.011159
Train Epoch: 6 [42800/60000 (72.7100%)] Loss: 0.001887
Train Epoch: 6 [43200/60000 (73.3895%)] Loss: 0.000035
Train Epoch: 6 [43600/60000 (74.0690%)] Loss: 0.000111
Train Epoch: 6 [44000/60000 (74.7486%)] Loss: 0.000044
Train Epoch: 6 [44400/60000 (75.4281%)] Loss: 0.000766
Train Epoch: 6 [44800/60000 (76.1076%)] Loss: 0.009368
Train Epoch: 6 [45200/60000 (76.7872%)] Loss: 0.000003
Train Epoch: 6 [45600/60000 (77.4667%)] Loss: 0.000016
Train Epoch: 6 [46000/60000 (78.1462%)] Loss: 0.000062
Train Epoch: 6 [46400/60000 (78.8258%)] Loss: 0.000344
Train Epoch: 6 [46800/60000 (79.5053%)] Loss: 0.000017
Train Epoch: 6 [47200/60000 (80.1848%)] Loss: 0.000112
Train Epoch: 6 [47600/60000 (80.8644%)] Loss: 0.000000
Train Epoch: 6 [48000/60000 (81.5439%)] Loss: 0.000001
Train Epoch: 6 [48400/60000 (82.2234%)] Loss: 0.135527
Train Epoch: 6 [48800/60000 (82.9030%)] Loss: 0.018530
Train Epoch: 6 [49200/60000 (83.5825%)] Loss: 0.000002
Train Epoch: 6 [49600/60000 (84.2620%)] Loss: 0.020816
Train Epoch: 6 [50000/60000 (84.9416%)] Loss: 0.000010
Train Epoch: 6 [50400/60000 (85.6211%)] Loss: 0.004621
```

```
Train Epoch: 6 [50800/60000 (86.3006%)] Loss: 0.000387
Train Epoch: 6 [51200/60000 (86.9802%)] Loss: 0.000015
Train Epoch: 6 [51600/60000 (87.6597%)] Loss: 0.007036
Train Epoch: 6 [52000/60000 (88.3392%)] Loss: 0.000095
Train Epoch: 6 [52400/60000 (89.0188%)] Loss: 0.000001
Train Epoch: 6 [52800/60000 (89.6983%)] Loss: 0.010590
Train Epoch: 6 [53200/60000 (90.3778%)] Loss: 0.000117
Train Epoch: 6 [53600/60000 (91.0574%)] Loss: 0.000696
Train Epoch: 6 [54000/60000 (91.7369%)] Loss: 0.000006
Train Epoch: 6 [54400/60000 (92.4164%)] Loss: 0.288760
Train Epoch: 6 [54800/60000 (93.0959%)] Loss: 0.000004
Train Epoch: 6 [55200/60000 (93.7755%)] Loss: 0.000007
Train Epoch: 6 [55600/60000 (94.4550%)] Loss: 0.000061
Train Epoch: 6 [56000/60000 (95.1345%)] Loss: 0.000001
Train Epoch: 6 [56400/60000 (95.8141%)] Loss: 0.002445
Train Epoch: 6 [56800/60000 (96.4936%)] Loss: 0.001788
Train Epoch: 6 [57200/60000 (97.1731%)] Loss: 0.000004
Train Epoch: 6 [57600/60000 (97.8527%)] Loss: 0.000095
Train Epoch: 6 [58000/60000 (98.5322%)] Loss: 0.000003
Train Epoch: 6 [58400/60000 (99.2117%)] Loss: 0.001103
Train Epoch: 6 [58800/60000 (99.8913%)] Loss: 0.002628
Test set: Average loss: 0.0334, Accuracy: 9907/10000 (99.0700%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.001200
Train Epoch: 7 [400/60000 (0.6795%)]    Loss: 0.000388
Train Epoch: 7 [800/60000 (1.3591%)]    Loss: 0.000032
Train Epoch: 7 [1200/60000 (2.0386%)]   Loss: 0.000014
Train Epoch: 7 [1600/60000 (2.7181%)]   Loss: 0.000021
Train Epoch: 7 [2000/60000 (3.3977%)]   Loss: 0.003472
Train Epoch: 7 [2400/60000 (4.0772%)]   Loss: 0.000002
Train Epoch: 7 [2800/60000 (4.7567%)]   Loss: 0.000018
Train Epoch: 7 [3200/60000 (5.4363%)]   Loss: 0.000713
Train Epoch: 7 [3600/60000 (6.1158%)]   Loss: 0.000538
Train Epoch: 7 [4000/60000 (6.7953%)]   Loss: 0.000006
Train Epoch: 7 [4400/60000 (7.4749%)]   Loss: 0.000004
Train Epoch: 7 [4800/60000 (8.1544%)]   Loss: 0.000093
Train Epoch: 7 [5200/60000 (8.8339%)]   Loss: 0.000001
Train Epoch: 7 [5600/60000 (9.5135%)]   Loss: 0.000209
Train Epoch: 7 [6000/60000 (10.1930%)]  Loss: 0.000144
Train Epoch: 7 [6400/60000 (10.8725%)]  Loss: 0.000089
Train Epoch: 7 [6800/60000 (11.5521%)]  Loss: 0.000043
Train Epoch: 7 [7200/60000 (12.2316%)]  Loss: 0.000651
Train Epoch: 7 [7600/60000 (12.9111%)]  Loss: 0.006318
Train Epoch: 7 [8000/60000 (13.5906%)]  Loss: 0.000185
Train Epoch: 7 [8400/60000 (14.2702%)]  Loss: 0.000005
Train Epoch: 7 [8800/60000 (14.9497%)]  Loss: 0.000315
Train Epoch: 7 [9200/60000 (15.6292%)]  Loss: 0.000586
Train Epoch: 7 [9600/60000 (16.3088%)]  Loss: 0.006508
Train Epoch: 7 [10000/60000 (16.9883%)] Loss: 0.000000
Train Epoch: 7 [10400/60000 (17.6678%)] Loss: 0.001768
Train Epoch: 7 [10800/60000 (18.3474%)] Loss: 0.000035
Train Epoch: 7 [11200/60000 (19.0269%)] Loss: 0.000171
Train Epoch: 7 [11600/60000 (19.7064%)] Loss: 0.000046
Train Epoch: 7 [12000/60000 (20.3860%)] Loss: 0.000038
Train Epoch: 7 [12400/60000 (21.0655%)] Loss: 0.000013
```

```
Train Epoch: 7 [12800/60000 (21.7450%)] Loss: 0.000004
Train Epoch: 7 [13200/60000 (22.4246%)] Loss: 0.000202
Train Epoch: 7 [13600/60000 (23.1041%)] Loss: 0.000071
Train Epoch: 7 [14000/60000 (23.7836%)] Loss: 0.000015
Train Epoch: 7 [14400/60000 (24.4632%)] Loss: 0.000319
Train Epoch: 7 [14800/60000 (25.1427%)] Loss: 0.000271
Train Epoch: 7 [15200/60000 (25.8222%)] Loss: 0.000203
Train Epoch: 7 [15600/60000 (26.5018%)] Loss: 0.000046
Train Epoch: 7 [16000/60000 (27.1813%)] Loss: 0.000001
Train Epoch: 7 [16400/60000 (27.8608%)] Loss: 0.000055
Train Epoch: 7 [16800/60000 (28.5404%)] Loss: 0.000060
Train Epoch: 7 [17200/60000 (29.2199%)] Loss: 0.000028
Train Epoch: 7 [17600/60000 (29.8994%)] Loss: 0.000021
Train Epoch: 7 [18000/60000 (30.5790%)] Loss: 0.004617
Train Epoch: 7 [18400/60000 (31.2585%)] Loss: 0.004547
Train Epoch: 7 [18800/60000 (31.9380%)] Loss: 0.000467
Train Epoch: 7 [19200/60000 (32.6176%)] Loss: 0.011841
Train Epoch: 7 [19600/60000 (33.2971%)] Loss: 0.000005
Train Epoch: 7 [20000/60000 (33.9766%)] Loss: 0.000264
Train Epoch: 7 [20400/60000 (34.6562%)] Loss: 0.000009
Train Epoch: 7 [20800/60000 (35.3357%)] Loss: 0.000361
Train Epoch: 7 [21200/60000 (36.0152%)] Loss: 0.000062
Train Epoch: 7 [21600/60000 (36.6948%)] Loss: 0.000026
Train Epoch: 7 [22000/60000 (37.3743%)] Loss: 0.000001
Train Epoch: 7 [22400/60000 (38.0538%)] Loss: 0.000001
Train Epoch: 7 [22800/60000 (38.7334%)] Loss: 0.000015
Train Epoch: 7 [23200/60000 (39.4129%)] Loss: 0.000549
Train Epoch: 7 [23600/60000 (40.0924%)] Loss: 0.000203
Train Epoch: 7 [24000/60000 (40.7719%)] Loss: 0.000012
Train Epoch: 7 [24400/60000 (41.4515%)] Loss: 0.003038
Train Epoch: 7 [24800/60000 (42.1310%)] Loss: 0.000817
Train Epoch: 7 [25200/60000 (42.8105%)] Loss: 0.001321
Train Epoch: 7 [25600/60000 (43.4901%)] Loss: 0.000018
Train Epoch: 7 [26000/60000 (44.1696%)] Loss: 0.000003
Train Epoch: 7 [26400/60000 (44.8491%)] Loss: 0.000005
Train Epoch: 7 [26800/60000 (45.5287%)] Loss: 0.000025
Train Epoch: 7 [27200/60000 (46.2082%)] Loss: 0.000451
Train Epoch: 7 [27600/60000 (46.8877%)] Loss: 0.000078
Train Epoch: 7 [28000/60000 (47.5673%)] Loss: 0.000001
Train Epoch: 7 [28400/60000 (48.2468%)] Loss: 0.001643
Train Epoch: 7 [28800/60000 (48.9263%)] Loss: 0.000010
Train Epoch: 7 [29200/60000 (49.6059%)] Loss: 0.001399
Train Epoch: 7 [29600/60000 (50.2854%)] Loss: 0.007360
Train Epoch: 7 [30000/60000 (50.9649%)] Loss: 0.000030
Train Epoch: 7 [30400/60000 (51.6445%)] Loss: 0.001075
Train Epoch: 7 [30800/60000 (52.3240%)] Loss: 0.000012
Train Epoch: 7 [31200/60000 (53.0035%)] Loss: 0.000001
Train Epoch: 7 [31600/60000 (53.6831%)] Loss: 0.000186
Train Epoch: 7 [32000/60000 (54.3626%)] Loss: 0.000006
Train Epoch: 7 [32400/60000 (55.0421%)] Loss: 0.000063
Train Epoch: 7 [32800/60000 (55.7217%)] Loss: 0.000000
Train Epoch: 7 [33200/60000 (56.4012%)] Loss: 0.000001
Train Epoch: 7 [33600/60000 (57.0807%)] Loss: 0.000002
Train Epoch: 7 [34000/60000 (57.7603%)] Loss: 0.007646
```

```
Train Epoch: 7 [34400/60000 (58.4398%)] Loss: 0.000001
Train Epoch: 7 [34800/60000 (59.1193%)] Loss: 0.000000
Train Epoch: 7 [35200/60000 (59.7989%)] Loss: 0.000000
Train Epoch: 7 [35600/60000 (60.4784%)] Loss: 0.000006
Train Epoch: 7 [36000/60000 (61.1579%)] Loss: 0.000008
Train Epoch: 7 [36400/60000 (61.8375%)] Loss: 0.001165
Train Epoch: 7 [36800/60000 (62.5170%)] Loss: 0.000000
Train Epoch: 7 [37200/60000 (63.1965%)] Loss: 0.000224
Train Epoch: 7 [37600/60000 (63.8761%)] Loss: 0.000035
Train Epoch: 7 [38000/60000 (64.5556%)] Loss: 0.000001
Train Epoch: 7 [38400/60000 (65.2351%)] Loss: 0.000585
Train Epoch: 7 [38800/60000 (65.9147%)] Loss: 0.000003
Train Epoch: 7 [39200/60000 (66.5942%)] Loss: 0.000297
Train Epoch: 7 [39600/60000 (67.2737%)] Loss: 0.000038
Train Epoch: 7 [40000/60000 (67.9532%)] Loss: 0.000032
Train Epoch: 7 [40400/60000 (68.6328%)] Loss: 0.000002
Train Epoch: 7 [40800/60000 (69.3123%)] Loss: 0.000093
Train Epoch: 7 [41200/60000 (69.9918%)] Loss: 0.960346
Train Epoch: 7 [41600/60000 (70.6714%)] Loss: 0.010640
Train Epoch: 7 [42000/60000 (71.3509%)] Loss: 0.000034
Train Epoch: 7 [42400/60000 (72.0304%)] Loss: 0.000035
Train Epoch: 7 [42800/60000 (72.7100%)] Loss: 0.000731
Train Epoch: 7 [43200/60000 (73.3895%)] Loss: 0.000047
Train Epoch: 7 [43600/60000 (74.0690%)] Loss: 0.000201
Train Epoch: 7 [44000/60000 (74.7486%)] Loss: 0.003724
Train Epoch: 7 [44400/60000 (75.4281%)] Loss: 0.000004
Train Epoch: 7 [44800/60000 (76.1076%)] Loss: 0.709364
Train Epoch: 7 [45200/60000 (76.7872%)] Loss: 0.000082
Train Epoch: 7 [45600/60000 (77.4667%)] Loss: 0.000025
Train Epoch: 7 [46000/60000 (78.1462%)] Loss: 0.000077
Train Epoch: 7 [46400/60000 (78.8258%)] Loss: 0.000003
Train Epoch: 7 [46800/60000 (79.5053%)] Loss: 0.000000
Train Epoch: 7 [47200/60000 (80.1848%)] Loss: 0.000001
Train Epoch: 7 [47600/60000 (80.8644%)] Loss: 0.002441
Train Epoch: 7 [48000/60000 (81.5439%)] Loss: 0.000013
Train Epoch: 7 [48400/60000 (82.2234%)] Loss: 0.000501
Train Epoch: 7 [48800/60000 (82.9030%)] Loss: 0.000151
Train Epoch: 7 [49200/60000 (83.5825%)] Loss: 0.000055
Train Epoch: 7 [49600/60000 (84.2620%)] Loss: 0.008562
Train Epoch: 7 [50000/60000 (84.9416%)] Loss: 0.000003
Train Epoch: 7 [50400/60000 (85.6211%)] Loss: 0.000103
Train Epoch: 7 [50800/60000 (86.3006%)] Loss: 0.020236
Train Epoch: 7 [51200/60000 (86.9802%)] Loss: 0.069736
Train Epoch: 7 [51600/60000 (87.6597%)] Loss: 0.030806
Train Epoch: 7 [52000/60000 (88.3392%)] Loss: 0.000005
Train Epoch: 7 [52400/60000 (89.0188%)] Loss: 0.001199
Train Epoch: 7 [52800/60000 (89.6983%)] Loss: 0.303504
Train Epoch: 7 [53200/60000 (90.3778%)] Loss: 0.000047
Train Epoch: 7 [53600/60000 (91.0574%)] Loss: 0.000004
Train Epoch: 7 [54000/60000 (91.7369%)] Loss: 0.000227
Train Epoch: 7 [54400/60000 (92.4164%)] Loss: 0.000270
Train Epoch: 7 [54800/60000 (93.0959%)] Loss: 0.000001
Train Epoch: 7 [55200/60000 (93.7755%)] Loss: 0.000192
Train Epoch: 7 [55600/60000 (94.4550%)] Loss: 0.000518
```

```
Train Epoch: 7 [56000/60000 (95.1345%)] Loss: 0.000183
Train Epoch: 7 [56400/60000 (95.8141%)] Loss: 0.000020
Train Epoch: 7 [56800/60000 (96.4936%)] Loss: 0.007246
Train Epoch: 7 [57200/60000 (97.1731%)] Loss: 0.000032
Train Epoch: 7 [57600/60000 (97.8527%)] Loss: 0.000088
Train Epoch: 7 [58000/60000 (98.5322%)] Loss: 0.000642
Train Epoch: 7 [58400/60000 (99.2117%)] Loss: 0.000000
Train Epoch: 7 [58800/60000 (99.8913%)] Loss: 0.000571
Test set: Average loss: 0.0402, Accuracy: 9895/10000 (98.9500%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.000031
Train Epoch: 8 [400/60000 (0.6795%)]    Loss: 0.002361
Train Epoch: 8 [800/60000 (1.3591%)]    Loss: 0.000007
Train Epoch: 8 [1200/60000 (2.0386%)]   Loss: 0.000094
Train Epoch: 8 [1600/60000 (2.7181%)]   Loss: 0.000052
Train Epoch: 8 [2000/60000 (3.3977%)]   Loss: 0.000061
Train Epoch: 8 [2400/60000 (4.0772%)]   Loss: 0.003793
Train Epoch: 8 [2800/60000 (4.7567%)]   Loss: 0.004534
Train Epoch: 8 [3200/60000 (5.4363%)]   Loss: 0.000000
Train Epoch: 8 [3600/60000 (6.1158%)]   Loss: 0.000003
Train Epoch: 8 [4000/60000 (6.7953%)]   Loss: 0.000001
Train Epoch: 8 [4400/60000 (7.4749%)]   Loss: 0.000000
Train Epoch: 8 [4800/60000 (8.1544%)]   Loss: 0.000029
Train Epoch: 8 [5200/60000 (8.8339%)]   Loss: 0.000007
Train Epoch: 8 [5600/60000 (9.5135%)]   Loss: 0.000000
Train Epoch: 8 [6000/60000 (10.1930%)]  Loss: 0.000070
Train Epoch: 8 [6400/60000 (10.8725%)]  Loss: 0.003229
Train Epoch: 8 [6800/60000 (11.5521%)]  Loss: 0.000420
Train Epoch: 8 [7200/60000 (12.2316%)]  Loss: 0.000000
Train Epoch: 8 [7600/60000 (12.9111%)]  Loss: 0.000004
Train Epoch: 8 [8000/60000 (13.5906%)]  Loss: 0.002644
Train Epoch: 8 [8400/60000 (14.2702%)]  Loss: 0.000136
Train Epoch: 8 [8800/60000 (14.9497%)]  Loss: 0.000000
Train Epoch: 8 [9200/60000 (15.6292%)]  Loss: 0.000002
Train Epoch: 8 [9600/60000 (16.3088%)]  Loss: 0.019744
Train Epoch: 8 [10000/60000 (16.9883%)] Loss: 0.000001
Train Epoch: 8 [10400/60000 (17.6678%)] Loss: 0.000000
Train Epoch: 8 [10800/60000 (18.3474%)] Loss: 0.000883
Train Epoch: 8 [11200/60000 (19.0269%)] Loss: 0.000014
Train Epoch: 8 [11600/60000 (19.7064%)] Loss: 0.000075
Train Epoch: 8 [12000/60000 (20.3860%)] Loss: 0.000154
Train Epoch: 8 [12400/60000 (21.0655%)] Loss: 0.000003
Train Epoch: 8 [12800/60000 (21.7450%)] Loss: 0.000030
Train Epoch: 8 [13200/60000 (22.4246%)] Loss: 0.000019
Train Epoch: 8 [13600/60000 (23.1041%)] Loss: 0.000017
Train Epoch: 8 [14000/60000 (23.7836%)] Loss: 0.098498
Train Epoch: 8 [14400/60000 (24.4632%)] Loss: 0.000001
Train Epoch: 8 [14800/60000 (25.1427%)] Loss: 0.001025
Train Epoch: 8 [15200/60000 (25.8222%)] Loss: 0.000004
Train Epoch: 8 [15600/60000 (26.5018%)] Loss: 0.000002
Train Epoch: 8 [16000/60000 (27.1813%)] Loss: 0.000026
Train Epoch: 8 [16400/60000 (27.8608%)] Loss: 0.005559
Train Epoch: 8 [16800/60000 (28.5404%)] Loss: 0.000095
Train Epoch: 8 [17200/60000 (29.2199%)] Loss: 0.000167
Train Epoch: 8 [17600/60000 (29.8994%)] Loss: 0.000001
```

```
Train Epoch: 8 [18000/60000 (30.5790%)] Loss: 0.000012
Train Epoch: 8 [18400/60000 (31.2585%)] Loss: 0.001248
Train Epoch: 8 [18800/60000 (31.9380%)] Loss: 0.000065
Train Epoch: 8 [19200/60000 (32.6176%)] Loss: 0.000000
Train Epoch: 8 [19600/60000 (33.2971%)] Loss: 0.000027
Train Epoch: 8 [20000/60000 (33.9766%)] Loss: 0.002569
Train Epoch: 8 [20400/60000 (34.6562%)] Loss: 0.000044
Train Epoch: 8 [20800/60000 (35.3357%)] Loss: 0.000105
Train Epoch: 8 [21200/60000 (36.0152%)] Loss: 0.000027
Train Epoch: 8 [21600/60000 (36.6948%)] Loss: 0.000035
Train Epoch: 8 [22000/60000 (37.3743%)] Loss: 0.000051
Train Epoch: 8 [22400/60000 (38.0538%)] Loss: 0.001860
Train Epoch: 8 [22800/60000 (38.7334%)] Loss: 0.000185
Train Epoch: 8 [23200/60000 (39.4129%)] Loss: 0.000085
Train Epoch: 8 [23600/60000 (40.0924%)] Loss: 0.000000
Train Epoch: 8 [24000/60000 (40.7719%)] Loss: 0.000003
Train Epoch: 8 [24400/60000 (41.4515%)] Loss: 0.000001
Train Epoch: 8 [24800/60000 (42.1310%)] Loss: 0.000095
Train Epoch: 8 [25200/60000 (42.8105%)] Loss: 0.000716
Train Epoch: 8 [25600/60000 (43.4901%)] Loss: 0.000000
Train Epoch: 8 [26000/60000 (44.1696%)] Loss: 0.000001
Train Epoch: 8 [26400/60000 (44.8491%)] Loss: 0.000076
Train Epoch: 8 [26800/60000 (45.5287%)] Loss: 0.002257
Train Epoch: 8 [27200/60000 (46.2082%)] Loss: 0.000008
Train Epoch: 8 [27600/60000 (46.8877%)] Loss: 0.000002
Train Epoch: 8 [28000/60000 (47.5673%)] Loss: 0.001347
Train Epoch: 8 [28400/60000 (48.2468%)] Loss: 0.000281
Train Epoch: 8 [28800/60000 (48.9263%)] Loss: 0.000003
Train Epoch: 8 [29200/60000 (49.6059%)] Loss: 0.000001
Train Epoch: 8 [29600/60000 (50.2854%)] Loss: 0.000116
Train Epoch: 8 [30000/60000 (50.9649%)] Loss: 0.000154
Train Epoch: 8 [30400/60000 (51.6445%)] Loss: 0.000000
Train Epoch: 8 [30800/60000 (52.3240%)] Loss: 0.000061
Train Epoch: 8 [31200/60000 (53.0035%)] Loss: 0.007338
Train Epoch: 8 [31600/60000 (53.6831%)] Loss: 0.000002
Train Epoch: 8 [32000/60000 (54.3626%)] Loss: 0.010155
Train Epoch: 8 [32400/60000 (55.0421%)] Loss: 0.000063
Train Epoch: 8 [32800/60000 (55.7217%)] Loss: 0.000340
Train Epoch: 8 [33200/60000 (56.4012%)] Loss: 0.000065
Train Epoch: 8 [33600/60000 (57.0807%)] Loss: 0.000000
Train Epoch: 8 [34000/60000 (57.7603%)] Loss: 0.000021
Train Epoch: 8 [34400/60000 (58.4398%)] Loss: 0.000001
Train Epoch: 8 [34800/60000 (59.1193%)] Loss: 0.000008
Train Epoch: 8 [35200/60000 (59.7989%)] Loss: 1.452599
Train Epoch: 8 [35600/60000 (60.4784%)] Loss: 0.000001
Train Epoch: 8 [36000/60000 (61.1579%)] Loss: 0.000034
Train Epoch: 8 [36400/60000 (61.8375%)] Loss: 0.000011
Train Epoch: 8 [36800/60000 (62.5170%)] Loss: 0.002320
Train Epoch: 8 [37200/60000 (63.1965%)] Loss: 0.000002
Train Epoch: 8 [37600/60000 (63.8761%)] Loss: 0.000072
Train Epoch: 8 [38000/60000 (64.5556%)] Loss: 0.000040
Train Epoch: 8 [38400/60000 (65.2351%)] Loss: 0.000017
Train Epoch: 8 [38800/60000 (65.9147%)] Loss: 0.000001
Train Epoch: 8 [39200/60000 (66.5942%)] Loss: 0.000002
```

```
Train Epoch: 8 [39600/60000 (67.2737%)] Loss: 0.197867
Train Epoch: 8 [40000/60000 (67.9532%)] Loss: 0.000072
Train Epoch: 8 [40400/60000 (68.6328%)] Loss: 0.000090
Train Epoch: 8 [40800/60000 (69.3123%)] Loss: 0.000030
Train Epoch: 8 [41200/60000 (69.9918%)] Loss: 0.005300
Train Epoch: 8 [41600/60000 (70.6714%)] Loss: 0.000149
Train Epoch: 8 [42000/60000 (71.3509%)] Loss: 0.000001
Train Epoch: 8 [42400/60000 (72.0304%)] Loss: 0.000345
Train Epoch: 8 [42800/60000 (72.7100%)] Loss: 0.009510
Train Epoch: 8 [43200/60000 (73.3895%)] Loss: 0.002457
Train Epoch: 8 [43600/60000 (74.0690%)] Loss: 0.001047
Train Epoch: 8 [44000/60000 (74.7486%)] Loss: 0.000000
Train Epoch: 8 [44400/60000 (75.4281%)] Loss: 0.000026
Train Epoch: 8 [44800/60000 (76.1076%)] Loss: 0.005689
Train Epoch: 8 [45200/60000 (76.7872%)] Loss: 0.000672
Train Epoch: 8 [45600/60000 (77.4667%)] Loss: 0.000575
Train Epoch: 8 [46000/60000 (78.1462%)] Loss: 0.000025
Train Epoch: 8 [46400/60000 (78.8258%)] Loss: 0.000098
Train Epoch: 8 [46800/60000 (79.5053%)] Loss: 0.020023
Train Epoch: 8 [47200/60000 (80.1848%)] Loss: 0.001028
Train Epoch: 8 [47600/60000 (80.8644%)] Loss: 0.000021
Train Epoch: 8 [48000/60000 (81.5439%)] Loss: 0.018492
Train Epoch: 8 [48400/60000 (82.2234%)] Loss: 0.000008
Train Epoch: 8 [48800/60000 (82.9030%)] Loss: 0.000025
Train Epoch: 8 [49200/60000 (83.5825%)] Loss: 0.033447
Train Epoch: 8 [49600/60000 (84.2620%)] Loss: 0.000160
Train Epoch: 8 [50000/60000 (84.9416%)] Loss: 0.000009
Train Epoch: 8 [50400/60000 (85.6211%)] Loss: 0.000050
Train Epoch: 8 [50800/60000 (86.3006%)] Loss: 0.000012
Train Epoch: 8 [51200/60000 (86.9802%)] Loss: 0.001864
Train Epoch: 8 [51600/60000 (87.6597%)] Loss: 0.000337
Train Epoch: 8 [52000/60000 (88.3392%)] Loss: 0.000010
Train Epoch: 8 [52400/60000 (89.0188%)] Loss: 0.000558
Train Epoch: 8 [52800/60000 (89.6983%)] Loss: 0.000106
Train Epoch: 8 [53200/60000 (90.3778%)] Loss: 0.000006
Train Epoch: 8 [53600/60000 (91.0574%)] Loss: 0.000287
Train Epoch: 8 [54000/60000 (91.7369%)] Loss: 0.000079
Train Epoch: 8 [54400/60000 (92.4164%)] Loss: 0.000046
Train Epoch: 8 [54800/60000 (93.0959%)] Loss: 0.000000
Train Epoch: 8 [55200/60000 (93.7755%)] Loss: 0.000095
Train Epoch: 8 [55600/60000 (94.4550%)] Loss: 0.000001
Train Epoch: 8 [56000/60000 (95.1345%)] Loss: 0.000003
Train Epoch: 8 [56400/60000 (95.8141%)] Loss: 0.000071
Train Epoch: 8 [56800/60000 (96.4936%)] Loss: 0.000013
Train Epoch: 8 [57200/60000 (97.1731%)] Loss: 0.000015
Train Epoch: 8 [57600/60000 (97.8527%)] Loss: 0.000002
Train Epoch: 8 [58000/60000 (98.5322%)] Loss: 0.000000
Train Epoch: 8 [58400/60000 (99.2117%)] Loss: 0.000918
Train Epoch: 8 [58800/60000 (99.8913%)] Loss: 0.000409
Test set: Average loss: 0.0425, Accuracy: 9886/10000 (98.8600%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.002156
Train Epoch: 9 [400/60000 (0.6795%)]    Loss: 0.000013
Train Epoch: 9 [800/60000 (1.3591%)]    Loss: 0.000010
Train Epoch: 9 [1200/60000 (2.0386%)]   Loss: 0.000001
```

```
Train Epoch: 9 [1600/60000 (2.7181%)]   Loss: 0.015411
Train Epoch: 9 [2000/60000 (3.3977%)]   Loss: 0.000000
Train Epoch: 9 [2400/60000 (4.0772%)]   Loss: 0.000001
Train Epoch: 9 [2800/60000 (4.7567%)]   Loss: 0.000000
Train Epoch: 9 [3200/60000 (5.4363%)]   Loss: 0.000058
Train Epoch: 9 [3600/60000 (6.1158%)]   Loss: 0.000002
Train Epoch: 9 [4000/60000 (6.7953%)]   Loss: 0.000021
Train Epoch: 9 [4400/60000 (7.4749%)]   Loss: 0.000000
Train Epoch: 9 [4800/60000 (8.1544%)]   Loss: 0.000152
Train Epoch: 9 [5200/60000 (8.8339%)]   Loss: 0.017992
Train Epoch: 9 [5600/60000 (9.5135%)]   Loss: 0.000002
Train Epoch: 9 [6000/60000 (10.1930%)]  Loss: 0.000000
Train Epoch: 9 [6400/60000 (10.8725%)]  Loss: 0.000021
Train Epoch: 9 [6800/60000 (11.5521%)]  Loss: 0.000023
Train Epoch: 9 [7200/60000 (12.2316%)]  Loss: 0.000603
Train Epoch: 9 [7600/60000 (12.9111%)]  Loss: 0.000067
Train Epoch: 9 [8000/60000 (13.5906%)]  Loss: 0.000012
Train Epoch: 9 [8400/60000 (14.2702%)]  Loss: 0.000034
Train Epoch: 9 [8800/60000 (14.9497%)]  Loss: 0.000036
Train Epoch: 9 [9200/60000 (15.6292%)]  Loss: 0.000000
Train Epoch: 9 [9600/60000 (16.3088%)]  Loss: 0.001197
Train Epoch: 9 [10000/60000 (16.9883%)] Loss: 0.000363
Train Epoch: 9 [10400/60000 (17.6678%)] Loss: 0.000002
Train Epoch: 9 [10800/60000 (18.3474%)] Loss: 0.000001
Train Epoch: 9 [11200/60000 (19.0269%)] Loss: 0.000003
Train Epoch: 9 [11600/60000 (19.7064%)] Loss: 0.001056
Train Epoch: 9 [12000/60000 (20.3860%)] Loss: 0.000000
Train Epoch: 9 [12400/60000 (21.0655%)] Loss: 0.000001
Train Epoch: 9 [12800/60000 (21.7450%)] Loss: 0.000039
Train Epoch: 9 [13200/60000 (22.4246%)] Loss: 0.000000
Train Epoch: 9 [13600/60000 (23.1041%)] Loss: 0.000035
Train Epoch: 9 [14000/60000 (23.7836%)] Loss: 0.000000
Train Epoch: 9 [14400/60000 (24.4632%)] Loss: 0.000069
Train Epoch: 9 [14800/60000 (25.1427%)] Loss: 0.000002
Train Epoch: 9 [15200/60000 (25.8222%)] Loss: 0.000000
Train Epoch: 9 [15600/60000 (26.5018%)] Loss: 0.000000
Train Epoch: 9 [16000/60000 (27.1813%)] Loss: 0.000217
Train Epoch: 9 [16400/60000 (27.8608%)] Loss: 0.000037
Train Epoch: 9 [16800/60000 (28.5404%)] Loss: 0.000000
Train Epoch: 9 [17200/60000 (29.2199%)] Loss: 0.000002
Train Epoch: 9 [17600/60000 (29.8994%)] Loss: 0.000004
Train Epoch: 9 [18000/60000 (30.5790%)] Loss: 0.000003
Train Epoch: 9 [18400/60000 (31.2585%)] Loss: 0.000073
Train Epoch: 9 [18800/60000 (31.9380%)] Loss: 0.001605
Train Epoch: 9 [19200/60000 (32.6176%)] Loss: 0.000000
Train Epoch: 9 [19600/60000 (33.2971%)] Loss: 0.000001
Train Epoch: 9 [20000/60000 (33.9766%)] Loss: 0.000010
Train Epoch: 9 [20400/60000 (34.6562%)] Loss: 0.000000
Train Epoch: 9 [20800/60000 (35.3357%)] Loss: 0.000006
Train Epoch: 9 [21200/60000 (36.0152%)] Loss: 0.000000
Train Epoch: 9 [21600/60000 (36.6948%)] Loss: 0.000000
Train Epoch: 9 [22000/60000 (37.3743%)] Loss: 0.000026
Train Epoch: 9 [22400/60000 (38.0538%)] Loss: 0.000026
```

```
Train Epoch: 9 [23200/60000 (39.4129%)] Loss: 0.000030
Train Epoch: 9 [23600/60000 (40.0924%)] Loss: 0.000551
Train Epoch: 9 [24000/60000 (40.7719%)] Loss: 0.000000
Train Epoch: 9 [24400/60000 (41.4515%)] Loss: 0.000001
Train Epoch: 9 [24800/60000 (42.1310%)] Loss: 0.000160
Train Epoch: 9 [25200/60000 (42.8105%)] Loss: 0.875517
Train Epoch: 9 [25600/60000 (43.4901%)] Loss: 0.000015
Train Epoch: 9 [26000/60000 (44.1696%)] Loss: 0.003322
Train Epoch: 9 [26400/60000 (44.8491%)] Loss: 0.000001
Train Epoch: 9 [26800/60000 (45.5287%)] Loss: 0.000008
Train Epoch: 9 [27200/60000 (46.2082%)] Loss: 0.000002
Train Epoch: 9 [27600/60000 (46.8877%)] Loss: 0.000000
Train Epoch: 9 [28000/60000 (47.5673%)] Loss: 0.000273
Train Epoch: 9 [28400/60000 (48.2468%)] Loss: 0.034198
Train Epoch: 9 [28800/60000 (48.9263%)] Loss: 0.002315
Train Epoch: 9 [29200/60000 (49.6059%)] Loss: 0.000071
Train Epoch: 9 [29600/60000 (50.2854%)] Loss: 0.000017
Train Epoch: 9 [30000/60000 (50.9649%)] Loss: 0.000000
Train Epoch: 9 [30400/60000 (51.6445%)] Loss: 0.000093
Train Epoch: 9 [30800/60000 (52.3240%)] Loss: 0.000153
Train Epoch: 9 [31200/60000 (53.0035%)] Loss: 0.000000
Train Epoch: 9 [31600/60000 (53.6831%)] Loss: 0.000010
Train Epoch: 9 [32000/60000 (54.3626%)] Loss: 0.000001
Train Epoch: 9 [32400/60000 (55.0421%)] Loss: 0.215318
Train Epoch: 9 [32800/60000 (55.7217%)] Loss: 0.000006
Train Epoch: 9 [33200/60000 (56.4012%)] Loss: 0.000008
Train Epoch: 9 [33600/60000 (57.0807%)] Loss: 0.009395
Train Epoch: 9 [34000/60000 (57.7603%)] Loss: 0.000000
Train Epoch: 9 [34400/60000 (58.4398%)] Loss: 0.000009
Train Epoch: 9 [34800/60000 (59.1193%)] Loss: 0.000075
Train Epoch: 9 [35200/60000 (59.7989%)] Loss: 0.000000
Train Epoch: 9 [35600/60000 (60.4784%)] Loss: 0.008038
Train Epoch: 9 [36000/60000 (61.1579%)] Loss: 0.000000
Train Epoch: 9 [36400/60000 (61.8375%)] Loss: 0.000000
Train Epoch: 9 [36800/60000 (62.5170%)] Loss: 0.000000
Train Epoch: 9 [37200/60000 (63.1965%)] Loss: 0.000035
Train Epoch: 9 [37600/60000 (63.8761%)] Loss: 0.000011
Train Epoch: 9 [38000/60000 (64.5556%)] Loss: 0.000092
Train Epoch: 9 [38400/60000 (65.2351%)] Loss: 0.000504
Train Epoch: 9 [38800/60000 (65.9147%)] Loss: 0.000007
Train Epoch: 9 [39200/60000 (66.5942%)] Loss: 0.000003
Train Epoch: 9 [39600/60000 (67.2737%)] Loss: 0.000083
Train Epoch: 9 [40000/60000 (67.9532%)] Loss: 0.000253
Train Epoch: 9 [40400/60000 (68.6328%)] Loss: 0.002519
Train Epoch: 9 [40800/60000 (69.3123%)] Loss: 0.000467
Train Epoch: 9 [41200/60000 (69.9918%)] Loss: 0.000005
Train Epoch: 9 [41600/60000 (70.6714%)] Loss: 0.000330
Train Epoch: 9 [42000/60000 (71.3509%)] Loss: 0.000419
Train Epoch: 9 [42400/60000 (72.0304%)] Loss: 0.000672
Train Epoch: 9 [42800/60000 (72.7100%)] Loss: 0.000003
Train Epoch: 9 [43200/60000 (73.3895%)] Loss: 0.000133
Train Epoch: 9 [43600/60000 (74.0690%)] Loss: 0.000001
Train Epoch: 9 [44000/60000 (74.7486%)] Loss: 0.010710
```

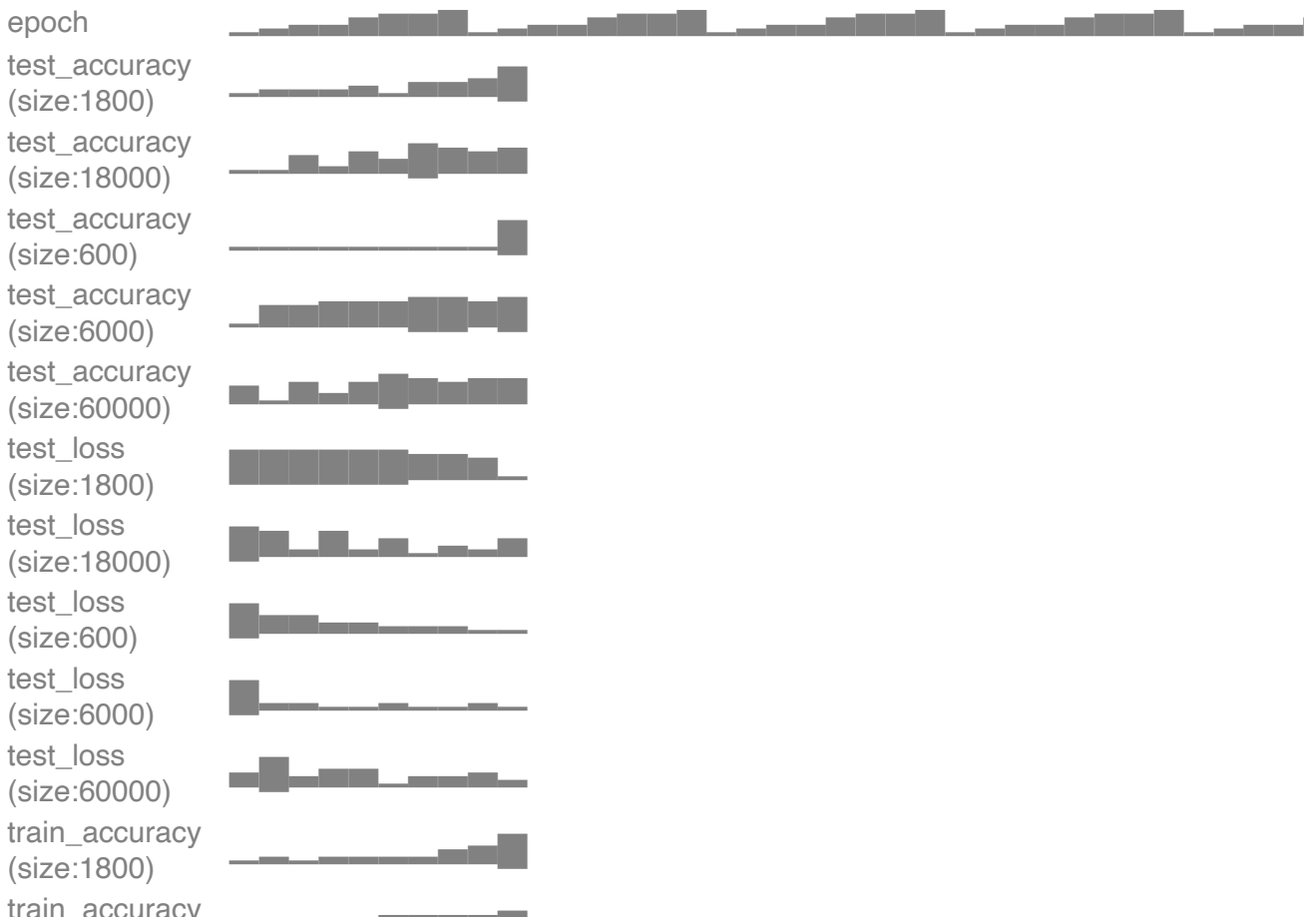
```
Train Epoch: 9 [44400/60000 (74.0068%)] Loss: 0.000009
Train Epoch: 9 [44800/60000 (76.1076%)] Loss: 0.000018
Train Epoch: 9 [45200/60000 (76.7872%)] Loss: 0.000029
Train Epoch: 9 [45600/60000 (77.4667%)] Loss: 0.000006
Train Epoch: 9 [46000/60000 (78.1462%)] Loss: 0.000000
Train Epoch: 9 [46400/60000 (78.8258%)] Loss: 0.000070
Train Epoch: 9 [46800/60000 (79.5053%)] Loss: 0.000012
Train Epoch: 9 [47200/60000 (80.1848%)] Loss: 0.000001
Train Epoch: 9 [47600/60000 (80.8644%)] Loss: 0.000021
Train Epoch: 9 [48000/60000 (81.5439%)] Loss: 0.000039
Train Epoch: 9 [48400/60000 (82.2234%)] Loss: 0.000000
Train Epoch: 9 [48800/60000 (82.9030%)] Loss: 0.000000
Train Epoch: 9 [49200/60000 (83.5825%)] Loss: 0.000001
Train Epoch: 9 [49600/60000 (84.2620%)] Loss: 0.000001
Train Epoch: 9 [50000/60000 (84.9416%)] Loss: 0.000010
Train Epoch: 9 [50400/60000 (85.6211%)] Loss: 0.000041
Train Epoch: 9 [50800/60000 (86.3006%)] Loss: 0.000585
Train Epoch: 9 [51200/60000 (86.9802%)] Loss: 0.000117
Train Epoch: 9 [51600/60000 (87.6597%)] Loss: 0.000064
Train Epoch: 9 [52000/60000 (88.3392%)] Loss: 0.000009
Train Epoch: 9 [52400/60000 (89.0188%)] Loss: 0.000139
Train Epoch: 9 [52800/60000 (89.6983%)] Loss: 0.000042
Train Epoch: 9 [53200/60000 (90.3778%)] Loss: 0.019108
Train Epoch: 9 [53600/60000 (91.0574%)] Loss: 0.000006
Train Epoch: 9 [54000/60000 (91.7369%)] Loss: 0.000001
Train Epoch: 9 [54400/60000 (92.4164%)] Loss: 0.000050
Train Epoch: 9 [54800/60000 (93.0959%)] Loss: 0.000006
Train Epoch: 9 [55200/60000 (93.7755%)] Loss: 0.000000
Train Epoch: 9 [55600/60000 (94.4550%)] Loss: 0.000000
Train Epoch: 9 [56000/60000 (95.1345%)] Loss: 0.006472
Train Epoch: 9 [56400/60000 (95.8141%)] Loss: 0.001087
Train Epoch: 9 [56800/60000 (96.4936%)] Loss: 0.001707
Train Epoch: 9 [57200/60000 (97.1731%)] Loss: 0.000008
Train Epoch: 9 [57600/60000 (97.8527%)] Loss: 0.000040
Train Epoch: 9 [58000/60000 (98.5322%)] Loss: 0.000015
Train Epoch: 9 [58400/60000 (99.2117%)] Loss: 0.000002
Train Epoch: 9 [58800/60000 (99.8913%)] Loss: 0.109261

Test set: Average loss: 0.0435, Accuracy: 9899/10000 (98.9900%)

Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.000004
Train Epoch: 10 [400/60000 (0.6795%)]   Loss: 0.030792
Train Epoch: 10 [800/60000 (1.3591%)]   Loss: 0.000493
Train Epoch: 10 [1200/60000 (2.0386%)]  Loss: 0.000002
Train Epoch: 10 [1600/60000 (2.7181%)]  Loss: 0.000001
Train Epoch: 10 [2000/60000 (3.3977%)]  Loss: 0.000000
Train Epoch: 10 [2400/60000 (4.0772%)]  Loss: 0.000675
Train Epoch: 10 [2800/60000 (4.7567%)]  Loss: 0.000126
Train Epoch: 10 [3200/60000 (5.4363%)]  Loss: 0.000004
Train Epoch: 10 [3600/60000 (6.1158%)]  Loss: 0.000042
Train Epoch: 10 [4000/60000 (6.7953%)]  Loss: 0.000330
Train Epoch: 10 [4400/60000 (7.4749%)]  Loss: 0.000115
Train Epoch: 10 [4800/60000 (8.1544%)]  Loss: 0.000000
Train Epoch: 10 [5200/60000 (8.8339%)]  Loss: 0.000026
Train Epoch: 10 [5600/60000 (9.5135%)]  Loss: 0.000166
Train Epoch: 10 [6000/60000 (10.1930%)] Loss: 0.000024
Train Epoch: 10 [6400/60000 (10.8725%)] Loss: 0.000000
```

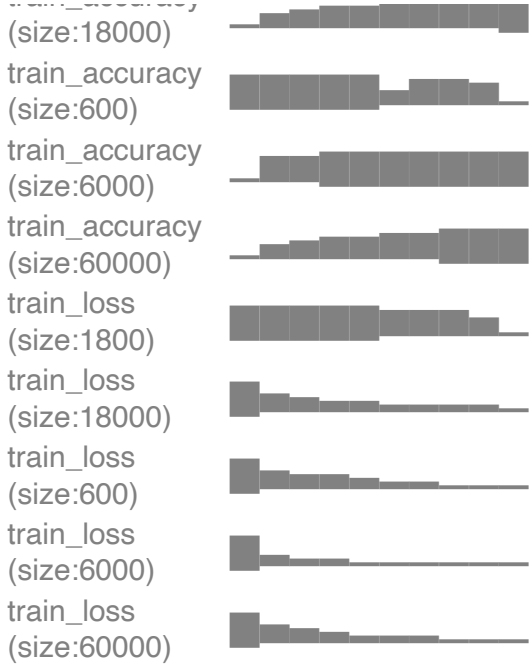
```
Train Epoch: 10 [6400/60000 (10.8723%)] Loss: 0.000000
Train Epoch: 10 [6800/60000 (11.5521%)] Loss: 0.000005
Train Epoch: 10 [7200/60000 (12.2316%)] Loss: 0.000015
Train Epoch: 10 [7600/60000 (12.9111%)] Loss: 0.000359
Train Epoch: 10 [8000/60000 (13.5906%)] Loss: 0.000106
Train Epoch: 10 [8400/60000 (14.2702%)] Loss: 0.000039
Train Epoch: 10 [8800/60000 (14.9497%)] Loss: 0.000001
Train Epoch: 10 [9200/60000 (15.6292%)] Loss: 0.000009
Train Epoch: 10 [9600/60000 (16.3088%)] Loss: 0.000103
Train Epoch: 10 [10000/60000 (16.9883%)]        Loss: 0.000100
Train Epoch: 10 [10400/60000 (17.6678%)]        Loss: 0.000000
Train Epoch: 10 [10800/60000 (18.3474%)]        Loss: 0.000051
Train Epoch: 10 [11200/60000 (19.0269%)]        Loss: 0.000004
Train Epoch: 10 [11600/60000 (19.7064%)]        Loss: 0.004548
Train Epoch: 10 [12000/60000 (20.3860%)]        Loss: 0.000002
Train Epoch: 10 [12400/60000 (21.0655%)]        Loss: 0.000006
Train Epoch: 10 [12800/60000 (21.7450%)]        Loss: 0.000885
Train Epoch: 10 [13200/60000 (22.4246%)]        Loss: 0.000024
Train Epoch: 10 [13600/60000 (23.1041%)]        Loss: 0.000007
Train Epoch: 10 [14000/60000 (23.7836%)]        Loss: 0.000013
Train Epoch: 10 [14400/60000 (24.4632%)]        Loss: 0.073676
Train Epoch: 10 [14800/60000 (25.1427%)]        Loss: 0.000002
Train Epoch: 10 [15200/60000 (25.8222%)]        Loss: 0.000000
Train Epoch: 10 [15600/60000 (26.5018%)]        Loss: 0.000061
Train Epoch: 10 [16000/60000 (27.1813%)]        Loss: 0.000010
Train Epoch: 10 [16400/60000 (27.8608%)]        Loss: 0.000547
Train Epoch: 10 [16800/60000 (28.5404%)]        Loss: 0.000144
Train Epoch: 10 [17200/60000 (29.2199%)]        Loss: 0.000001
Train Epoch: 10 [17600/60000 (29.8994%)]        Loss: 0.000008
Train Epoch: 10 [18000/60000 (30.5790%)]        Loss: 0.000329
Train Epoch: 10 [18400/60000 (31.2585%)]        Loss: 0.000002
Train Epoch: 10 [18800/60000 (31.9380%)]        Loss: 0.000048
Train Epoch: 10 [19200/60000 (32.6176%)]        Loss: 0.000053
Train Epoch: 10 [19600/60000 (33.2971%)]        Loss: 0.000000
Train Epoch: 10 [20000/60000 (33.9766%)]        Loss: 0.000001
Train Epoch: 10 [20400/60000 (34.6562%)]        Loss: 0.000006
Train Epoch: 10 [20800/60000 (35.3357%)]        Loss: 0.000002
Train Epoch: 10 [21200/60000 (36.0152%)]        Loss: 0.107438
Train Epoch: 10 [21600/60000 (36.6948%)]        Loss: 0.000047
Train Epoch: 10 [22000/60000 (37.3743%)]        Loss: 0.000008
Train Epoch: 10 [22400/60000 (38.0538%)]        Loss: 0.001471
Train Epoch: 10 [22800/60000 (38.7334%)]        Loss: 0.000026
Train Epoch: 10 [23200/60000 (39.4129%)]        Loss: 0.001185
Train Epoch: 10 [23600/60000 (40.0924%)]        Loss: 0.000001
Train Epoch: 10 [24000/60000 (40.7719%)]        Loss: 0.000002
Train Epoch: 10 [24400/60000 (41.4515%)]        Loss: 0.000008
Train Epoch: 10 [24800/60000 (42.1310%)]        Loss: 0.000025
Train Epoch: 10 [25200/60000 (42.8105%)]        Loss: 0.000001
Train Epoch: 10 [25600/60000 (43.4901%)]        Loss: 0.000000
Train Epoch: 10 [26000/60000 (44.1696%)]        Loss: 0.097826
Train Epoch: 10 [26400/60000 (44.8491%)]        Loss: 0.000004
Train Epoch: 10 [26800/60000 (45.5287%)]        Loss: 0.000002
Train Epoch: 10 [27200/60000 (46.2082%)]        Loss: 0.000260
Train Epoch: 10 [27600/60000 (46.8877%)]        Loss: 0.000523
Train Epoch: 10 [28000/60000 (47.5673%)]        Loss: 0.000000
```

```
Train Epoch: 10 [28000/60000 (47.5673%)]        Loss: 0.000000
Train Epoch: 10 [28400/60000 (48.2468%)]        Loss: 0.000095
Train Epoch: 10 [28800/60000 (48.9263%)]        Loss: 0.000000
Train Epoch: 10 [29200/60000 (49.6059%)]        Loss: 0.000001
Train Epoch: 10 [29600/60000 (50.2854%)]        Loss: 0.000000
Train Epoch: 10 [30000/60000 (50.9649%)]        Loss: 0.000000
Train Epoch: 10 [30400/60000 (51.6445%)]        Loss: 0.000008
Train Epoch: 10 [30800/60000 (52.3240%)]        Loss: 0.000003
Train Epoch: 10 [31200/60000 (53.0035%)]        Loss: 0.000010
Train Epoch: 10 [31600/60000 (53.6831%)]        Loss: 0.000020
Train Epoch: 10 [32000/60000 (54.3626%)]        Loss: 0.000225
Train Epoch: 10 [32400/60000 (55.0421%)]        Loss: 0.000002
Train Epoch: 10 [32800/60000 (55.7217%)]        Loss: 0.000171
Train Epoch: 10 [33200/60000 (56.4012%)]        Loss: 0.000007
Train Epoch: 10 [33600/60000 (57.0807%)]        Loss: 0.000007
Train Epoch: 10 [34000/60000 (57.7603%)]        Loss: 0.000004
Train Epoch: 10 [34400/60000 (58.4398%)]        Loss: 0.000001
Train Epoch: 10 [34800/60000 (59.1193%)]        Loss: 0.000478
Train Epoch: 10 [35200/60000 (59.7989%)]        Loss: 0.001318
Train Epoch: 10 [35600/60000 (60.4784%)]        Loss: 0.085143
Train Epoch: 10 [36000/60000 (61.1579%)]        Loss: 0.000425
Train Epoch: 10 [36400/60000 (61.8375%)]        Loss: 0.001219
Train Epoch: 10 [36800/60000 (62.5170%)]        Loss: 0.000001
Train Epoch: 10 [37200/60000 (63.1965%)]        Loss: 0.000287
Train Epoch: 10 [37600/60000 (63.8761%)]        Loss: 0.000015
Train Epoch: 10 [38000/60000 (64.5556%)]        Loss: 0.000004
Train Epoch: 10 [38400/60000 (65.2351%)]        Loss: 0.017086
Train Epoch: 10 [38800/60000 (65.9147%)]        Loss: 0.000430
Train Epoch: 10 [39200/60000 (66.5942%)]        Loss: 0.000387
Train Epoch: 10 [39600/60000 (67.2737%)]        Loss: 0.000029
Train Epoch: 10 [40000/60000 (67.9532%)]        Loss: 0.000008
Train Epoch: 10 [40400/60000 (68.6328%)]        Loss: 0.000096
Train Epoch: 10 [40800/60000 (69.3123%)]        Loss: 0.000004
Train Epoch: 10 [41200/60000 (69.9918%)]        Loss: 0.000003
Train Epoch: 10 [41600/60000 (70.6714%)]        Loss: 0.008958
Train Epoch: 10 [42000/60000 (71.3509%)]        Loss: 0.000000
Train Epoch: 10 [42400/60000 (72.0304%)]        Loss: 0.000237
Train Epoch: 10 [42800/60000 (72.7100%)]        Loss: 0.000000
Train Epoch: 10 [43200/60000 (73.3895%)]        Loss: 0.000009
Train Epoch: 10 [43600/60000 (74.0690%)]        Loss: 0.000001
Train Epoch: 10 [44000/60000 (74.7486%)]        Loss: 0.000000
Train Epoch: 10 [44400/60000 (75.4281%)]        Loss: 0.000004
Train Epoch: 10 [44800/60000 (76.1076%)]        Loss: 0.000106
Train Epoch: 10 [45200/60000 (76.7872%)]        Loss: 0.000007
Train Epoch: 10 [45600/60000 (77.4667%)]        Loss: 0.000104
Train Epoch: 10 [46000/60000 (78.1462%)]        Loss: 0.000001
Train Epoch: 10 [46400/60000 (78.8258%)]        Loss: 0.000011
Train Epoch: 10 [46800/60000 (79.5053%)]        Loss: 0.000189
Train Epoch: 10 [47200/60000 (80.1848%)]        Loss: 0.000019
Train Epoch: 10 [47600/60000 (80.8644%)]        Loss: 0.000208
Train Epoch: 10 [48000/60000 (81.5439%)]        Loss: 0.000008
Train Epoch: 10 [48400/60000 (82.2234%)]        Loss: 0.000000
Train Epoch: 10 [48800/60000 (82.9030%)]        Loss: 0.000168
Train Epoch: 10 [49200/60000 (83.5825%)]        Loss: 0.000772
```

```
Train Epoch: 10 [49600/60000 (84.2620%)]        Loss: 0.000011
Train Epoch: 10 [50000/60000 (84.9416%)]        Loss: 0.000202
Train Epoch: 10 [50400/60000 (85.6211%)]        Loss: 0.000001
Train Epoch: 10 [50800/60000 (86.3006%)]        Loss: 0.000016
Train Epoch: 10 [51200/60000 (86.9802%)]        Loss: 0.001011
Train Epoch: 10 [51600/60000 (87.6597%)]        Loss: 0.000005
Train Epoch: 10 [52000/60000 (88.3392%)]        Loss: 0.000027
Train Epoch: 10 [52400/60000 (89.0188%)]        Loss: 0.000000
Train Epoch: 10 [52800/60000 (89.6983%)]        Loss: 0.003228
Train Epoch: 10 [53200/60000 (90.3778%)]        Loss: 0.000033
Train Epoch: 10 [53600/60000 (91.0574%)]        Loss: 0.000009
Train Epoch: 10 [54000/60000 (91.7369%)]        Loss: 0.000000
Train Epoch: 10 [54400/60000 (92.4164%)]        Loss: 0.000034
Train Epoch: 10 [54800/60000 (93.0959%)]        Loss: 0.001717
Train Epoch: 10 [55200/60000 (93.7755%)]        Loss: 0.000203
Train Epoch: 10 [55600/60000 (94.4550%)]        Loss: 0.000000
Train Epoch: 10 [56000/60000 (95.1345%)]        Loss: 0.000373
Train Epoch: 10 [56400/60000 (95.8141%)]        Loss: 0.001831
Train Epoch: 10 [56800/60000 (96.4936%)]        Loss: 0.000009
Train Epoch: 10 [57200/60000 (97.1731%)]        Loss: 0.000044
Train Epoch: 10 [57600/60000 (97.8527%)]        Loss: 0.000000
Train Epoch: 10 [58000/60000 (98.5322%)]        Loss: 0.001779
Train Epoch: 10 [58400/60000 (99.2117%)]        Loss: 0.000000
Train Epoch: 10 [58800/60000 (99.8913%)]        Loss: 0.001852
Test set: Average loss: 0.0387, Accuracy: 9895/10000 (98.9500%)
```

Run history:





Run summary:

Transformer VS CNN

- The ViT style transformer encoder model consists of self-attention mechanisms, linear and convolutional layers, stacked in two layers.
- We input the image into the transformer (with no positinal embeddings) to the transformer with 4 heads and 256 hidden dimensions

Accuracy (for 6k images) = 92.95%
Accuracy (for 60k images) = 97.88%

Both the transformer and CNN perform very well for the task of digit recognition. The LeNet CNN architecture has been specially designed for this task and so it is not surprising that it works so well.
The transformer takes a little bit more time to train but in general, performs lot better than CNN's because of its attention mechanism and ability to look ahead. However, in the context of digit recognition, they are very similar.

```
class TransformerModel(nn.Module):
    def __init__(self, input_dim, hidden_dim, output_dim, num_layers, num_heads
        super(TransformerModel, self).__init__()
        self.embedding = nn.Linear(input_dim, hidden_dim)
        self.inputdim = input_dim
        encoder_layer = nn.TransformerEncoderLayer(d_model=hidden_dim, nhead=nu
        self.transformer_encoder = nn.TransformerEncoder(encoder_layer, num_lay
        self.fc = nn.Linear(hidden_dim, output_dim)

    def forward(self, x):
        x = x.view(-1, self.inputdim)
        x = self.embedding(x)
        x = x.unsqueeze(1)
        x = self.transformer_encoder(x)
        x = x.squeeze(1)  # original shape
        x = self.fc(x)
        return x


train_sizes = [6000, 60000]

# Training parameters
batch_size = 64
learning_rate = 0.001
epochs = 10

# Transformer parameters
input_dim = 28 * 28  # MNIST image size
hidden_dim = 256
output_dim = 10  # Number of labels
num_layers = 2
num_heads = 4

transform_model = TransformerModel(input_dim, hidden_dim, output_dim, num_layer
loss_fn = nn.CrossEntropyLoss()
optimizer = optim.SGD(transform_model.parameters(), lr=0.001, momentum=0.9)

wandb.init(project='transformer-vs-cnn')
for size in train_sizes:
    print(f"\nTraining with {size} samples:")

    samples_per_class = size // 10  # 10 labels
    selected_indices = []

    # Iterate over each class
    for i in range(10):
        class_indices = (train_dataset.targets == i).nonzero(as_tuple=True)[0]
```

```
        # randomise
        class_indices = class_indices[torch.randperm(len(class_indices))]
        selected_indices.extend(class_indices[:samples_per_class].tolist())

    subset_sampler = torch.utils.data.SubsetRandomSampler(selected_indices)
    train_loader = torch.utils.data.DataLoader(train_dataset, batch_size, sampl

    # Training loop
    for epoch in range(1, epochs + 1):
      train_loss, train_accuracy = train(transform_model, train_loader, loss_fr
      test_loss, test_accuracy = test(transform_model, test_loader, loss_fn)

      wandb.log({
          "epoch": epoch,
          f"train_loss (size:{size})": train_loss,
          f"train_accuracy (size:{size})": train_accuracy,
          f"test_loss (size:{size})": test_loss,
          f"test_accuracy (size:{size})": test_accuracy
      })

wandb.finish()
```

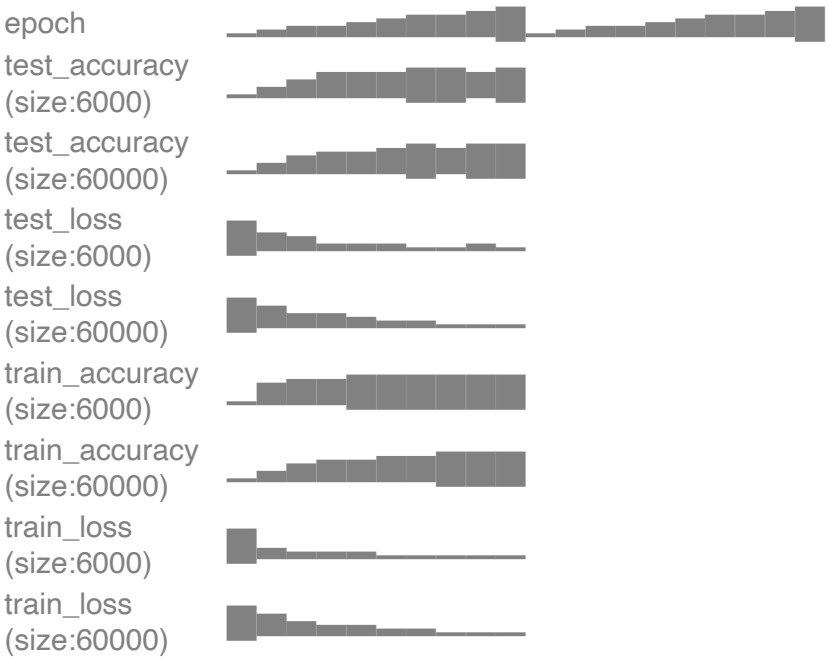
```
/usr/local/lib/python3.10/dist-packages/torch/nn/modules/transformer.py:282
  warnings.warn(f"enable_nested_tensor is True, but self.use_nested_tensor
Tracking run with wandb version 0.16.4
Run data is saved locally in /content/wandb/run-20240309_145338-dz97htcj
Syncing run frosty-glitter-1 to Weights & Biases (docs)
View project at https://wandb.ai/hugthebee/transformer-vs-cnn
View run at https://wandb.ai/hugthebee/transformer-vs-cnn/runs/dz97htcj

Training with 6000 samples:
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 2.391948
Test set: Average loss: 0.5161, Accuracy: 8606/10000 (86.0600%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.473207
Test set: Average loss: 0.4210, Accuracy: 8785/10000 (87.8500%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.356965
Test set: Average loss: 0.3767, Accuracy: 8914/10000 (89.1400%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.560345
Test set: Average loss: 0.3306, Accuracy: 9067/10000 (90.6700%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.379106
Test set: Average loss: 0.3227, Accuracy: 9050/10000 (90.5000%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.585279
Test set: Average loss: 0.3076, Accuracy: 9097/10000 (90.9700%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.358232
Test set: Average loss: 0.2999, Accuracy: 9118/10000 (91.1800%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.205135
Test set: Average loss: 0.2876, Accuracy: 9133/10000 (91.3300%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.276899
Test set: Average loss: 0.3022, Accuracy: 9096/10000 (90.9600%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.219399
Test set: Average loss: 0.2838, Accuracy: 9135/10000 (91.3500%)
```

```
Training with 60000 samples:
Train Epoch: 1 [0/60000 (0.0000%)]      Loss: 0.324269
Train Epoch: 1 [6400/60000 (10.8696%)]  Loss: 0.332779
Train Epoch: 1 [12800/60000 (21.7391%)] Loss: 0.305146
Train Epoch: 1 [19200/60000 (32.6087%)] Loss: 0.240253
Train Epoch: 1 [25600/60000 (43.4783%)] Loss: 0.117149
Train Epoch: 1 [32000/60000 (54.3478%)] Loss: 0.312782
Train Epoch: 1 [38400/60000 (65.2174%)] Loss: 0.156097
Train Epoch: 1 [44800/60000 (76.0870%)] Loss: 0.170965
Train Epoch: 1 [51200/60000 (86.9565%)] Loss: 0.223555
Train Epoch: 1 [57600/60000 (97.8261%)] Loss: 0.193927
Test set: Average loss: 0.2076, Accuracy: 9385/10000 (93.8500%)
Train Epoch: 2 [0/60000 (0.0000%)]      Loss: 0.231250
Train Epoch: 2 [6400/60000 (10.8696%)]  Loss: 0.127795
Train Epoch: 2 [12800/60000 (21.7391%)] Loss: 0.258064
Train Epoch: 2 [19200/60000 (32.6087%)] Loss: 0.184969
Train Epoch: 2 [25600/60000 (43.4783%)] Loss: 0.255712
Train Epoch: 2 [32000/60000 (54.3478%)] Loss: 0.141783
Train Epoch: 2 [38400/60000 (65.2174%)] Loss: 0.177224
Train Epoch: 2 [44800/60000 (76.0870%)] Loss: 0.140045
Train Epoch: 2 [51200/60000 (86.9565%)] Loss: 0.169421
Train Epoch: 2 [57600/60000 (97.8261%)] Loss: 0.252566
Test set: Average loss: 0.1661, Accuracy: 9488/10000 (94.8800%)
Train Epoch: 3 [0/60000 (0.0000%)]      Loss: 0.161267
Train Epoch: 3 [6400/60000 (10.8696%)]  Loss: 0.216562
Train Epoch: 3 [12800/60000 (21.7391%)] Loss: 0.286527
Train Epoch: 3 [19200/60000 (32.6087%)] Loss: 0.106958
Train Epoch: 3 [25600/60000 (43.4783%)] Loss: 0.177589
Train Epoch: 3 [32000/60000 (54.3478%)] Loss: 0.083619
Train Epoch: 3 [38400/60000 (65.2174%)] Loss: 0.159843
Train Epoch: 3 [44800/60000 (76.0870%)] Loss: 0.153432
Train Epoch: 3 [51200/60000 (86.9565%)] Loss: 0.058433
Train Epoch: 3 [57600/60000 (97.8261%)] Loss: 0.191613
Test set: Average loss: 0.1442, Accuracy: 9566/10000 (95.6600%)
Train Epoch: 4 [0/60000 (0.0000%)]      Loss: 0.087090
Train Epoch: 4 [6400/60000 (10.8696%)]  Loss: 0.073048
Train Epoch: 4 [12800/60000 (21.7391%)] Loss: 0.080556
Train Epoch: 4 [19200/60000 (32.6087%)] Loss: 0.080767
Train Epoch: 4 [25600/60000 (43.4783%)] Loss: 0.162454
Train Epoch: 4 [32000/60000 (54.3478%)] Loss: 0.036066
Train Epoch: 4 [38400/60000 (65.2174%)] Loss: 0.049830
Train Epoch: 4 [44800/60000 (76.0870%)] Loss: 0.154208
Train Epoch: 4 [51200/60000 (86.9565%)] Loss: 0.016236
Train Epoch: 4 [57600/60000 (97.8261%)] Loss: 0.149174
Test set: Average loss: 0.1286, Accuracy: 9615/10000 (96.1500%)
Train Epoch: 5 [0/60000 (0.0000%)]      Loss: 0.090738
Train Epoch: 5 [6400/60000 (10.8696%)]  Loss: 0.121973
Train Epoch: 5 [12800/60000 (21.7391%)] Loss: 0.162467
Train Epoch: 5 [19200/60000 (32.6087%)] Loss: 0.077015
Train Epoch: 5 [25600/60000 (43.4783%)] Loss: 0.169393
Train Epoch: 5 [32000/60000 (54.3478%)] Loss: 0.068552
Train Epoch: 5 [38400/60000 (65.2174%)] Loss: 0.121042
Train Epoch: 5 [44800/60000 (76.0870%)] Loss: 0.054864
```

```
Train Epoch: 5 [51200/60000 (86.9565%)] Loss: 0.290322
Train Epoch: 5 [57600/60000 (97.8261%)] Loss: 0.184788
Test set: Average loss: 0.1197, Accuracy: 9636/10000 (96.3600%)
Train Epoch: 6 [0/60000 (0.0000%)]      Loss: 0.173149
Train Epoch: 6 [6400/60000 (10.8696%)]  Loss: 0.081023
Train Epoch: 6 [12800/60000 (21.7391%)] Loss: 0.097416
Train Epoch: 6 [19200/60000 (32.6087%)] Loss: 0.103403
Train Epoch: 6 [25600/60000 (43.4783%)] Loss: 0.064080
Train Epoch: 6 [32000/60000 (54.3478%)] Loss: 0.032877
Train Epoch: 6 [38400/60000 (65.2174%)] Loss: 0.083407
Train Epoch: 6 [44800/60000 (76.0870%)] Loss: 0.039341
Train Epoch: 6 [51200/60000 (86.9565%)] Loss: 0.056034
Train Epoch: 6 [57600/60000 (97.8261%)] Loss: 0.035900
Test set: Average loss: 0.1036, Accuracy: 9687/10000 (96.8700%)
Train Epoch: 7 [0/60000 (0.0000%)]      Loss: 0.086543
Train Epoch: 7 [6400/60000 (10.8696%)]  Loss: 0.104190
Train Epoch: 7 [12800/60000 (21.7391%)] Loss: 0.062211
Train Epoch: 7 [19200/60000 (32.6087%)] Loss: 0.104554
Train Epoch: 7 [25600/60000 (43.4783%)] Loss: 0.102022
Train Epoch: 7 [32000/60000 (54.3478%)] Loss: 0.044491
Train Epoch: 7 [38400/60000 (65.2174%)] Loss: 0.197589
Train Epoch: 7 [44800/60000 (76.0870%)] Loss: 0.061403
Train Epoch: 7 [51200/60000 (86.9565%)] Loss: 0.153194
Train Epoch: 7 [57600/60000 (97.8261%)] Loss: 0.192313
Test set: Average loss: 0.0940, Accuracy: 9714/10000 (97.1400%)
Train Epoch: 8 [0/60000 (0.0000%)]      Loss: 0.069220
Train Epoch: 8 [6400/60000 (10.8696%)]  Loss: 0.075456
Train Epoch: 8 [12800/60000 (21.7391%)] Loss: 0.012933
Train Epoch: 8 [19200/60000 (32.6087%)] Loss: 0.106142
Train Epoch: 8 [25600/60000 (43.4783%)] Loss: 0.011423
Train Epoch: 8 [32000/60000 (54.3478%)] Loss: 0.070977
Train Epoch: 8 [38400/60000 (65.2174%)] Loss: 0.129796
Train Epoch: 8 [44800/60000 (76.0870%)] Loss: 0.023499
Train Epoch: 8 [51200/60000 (86.9565%)] Loss: 0.173263
Train Epoch: 8 [57600/60000 (97.8261%)] Loss: 0.056025
Test set: Average loss: 0.0921, Accuracy: 9704/10000 (97.0400%)
Train Epoch: 9 [0/60000 (0.0000%)]      Loss: 0.052518
Train Epoch: 9 [6400/60000 (10.8696%)]  Loss: 0.089016
Train Epoch: 9 [12800/60000 (21.7391%)] Loss: 0.020131
Train Epoch: 9 [19200/60000 (32.6087%)] Loss: 0.097440
Train Epoch: 9 [25600/60000 (43.4783%)] Loss: 0.047324
Train Epoch: 9 [32000/60000 (54.3478%)] Loss: 0.047848
Train Epoch: 9 [38400/60000 (65.2174%)] Loss: 0.078224
Train Epoch: 9 [44800/60000 (76.0870%)] Loss: 0.088431
Train Epoch: 9 [51200/60000 (86.9565%)] Loss: 0.020215
Train Epoch: 9 [57600/60000 (97.8261%)] Loss: 0.054466
Test set: Average loss: 0.0904, Accuracy: 9725/10000 (97.2500%)
Train Epoch: 10 [0/60000 (0.0000%)]     Loss: 0.041494
Train Epoch: 10 [6400/60000 (10.8696%)] Loss: 0.119638
Train Epoch: 10 [12800/60000 (21.7391%)]        Loss: 0.041422
Train Epoch: 10 [19200/60000 (32.6087%)]        Loss: 0.037405
Train Epoch: 10 [25600/60000 (43.4783%)]        Loss: 0.083136
Train Epoch: 10 [32000/60000 (54.3478%)]        Loss: 0.015838
Train Epoch: 10 [38400/60000 (65.2174%)]        Loss: 0.046603
```

```
Train Epoch: 10 [44800/60000 (76.0870%)]        Loss: 0.048688
Train Epoch: 10 [51200/60000 (86.9565%)]        Loss: 0.128194
Train Epoch: 10 [57600/60000 (97.8261%)]        Loss: 0.029055
Test set: Average loss: 0.0847, Accuracy: 9739/10000 (97.3900%)
```

Run history:



Run summary:

epoch	10
test_accuracy (size:6000)	91.35
test_accuracy (size:60000)	97.39
test_loss (size:6000)	0.28379
test_loss (size:60000)	0.08474
train_accuracy (size:6000)	92.95
train_accuracy (size:60000)	97.88828
train_loss (size:6000)	0.23422
train_loss (size:60000)	0.06866

View run **frosty-glitter-1** at: https://wandb.ai/hugthebee/transformer-vs-cnn/runs/dz97htcj
Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Find logs at: `./wandb/run-20240309_145338-dz97htcj/logs`