A1_4_1

March 27, 2024

0.1 Part I

[ ]:
```
import torch
import torchvision
import torchvision.transforms as transforms
import torch.nn as nn
import torch.optim as optim

# Device configuration
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')

# Hyperparameters
num_epochs = 10
batch_size = 64
learning_rate = 0.001

# CIFAR-10 dataset
transform = transforms.Compose([
    transforms.ToTensor(),
    transforms.Normalize((0.5, 0.5, 0.5), (0.5, 0.5, 0.5))
])

train_dataset = torchvision.datasets.CIFAR10(root='./data', train=True,␣
 ↪download=True, transform=transform)
train_loader = torch.utils.data.DataLoader(train_dataset,␣
 ↪batch_size=batch_size, shuffle=True)

test_dataset = torchvision.datasets.CIFAR10(root='./data', train=False,␣
 ↪download=True, transform=transform)
test_loader = torch.utils.data.DataLoader(test_dataset, batch_size=batch_size,␣
 ↪shuffle=False)

# Convolutional neural network (3 Convolutional layers + 1 Fully connected␣
 ↪layer)
class CNN(nn.Module):
    def __init__(self):
        super(CNN, self).__init__()
        self.conv1 = nn.Conv2d(3, 16, kernel_size=3, padding=1)
```

```
        self.conv2 = nn.Conv2d(16, 32, kernel_size=3, padding=1)
        self.conv3 = nn.Conv2d(32, 64, kernel_size=3, padding=1)
        self.pool = nn.MaxPool2d(2, 2)
        self.fc1 = nn.Linear(64 * 4 * 4, 256)
        self.fc2 = nn.Linear(256, 10)

    def forward(self, x):
        x = self.pool(torch.relu(self.conv1(x)))
        x = self.pool(torch.relu(self.conv2(x)))
        x = self.pool(torch.relu(self.conv3(x)))
        x = x.view(-1, 64 * 4 * 4)
        x = torch.relu(self.fc1(x))
        x = self.fc2(x)
        return x

model = CNN().to(device)

# Loss and optimizer
criterion = nn.CrossEntropyLoss()
optimizer = optim.Adam(model.parameters(), lr=learning_rate)

# Train the model
total_step = len(train_loader)
for epoch in range(num_epochs):
    for i, (images, labels) in enumerate(train_loader):
        images = images.to(device)
        labels = labels.to(device)

        # Forward pass
        outputs = model(images)
        loss = criterion(outputs, labels)

        # Backward and optimize
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()

        if (i+1) % 100 == 0:
            print ('Epoch [{}/{}], Step [{}/{}], Loss: {:.4f}'.format(epoch+1,␣
 ↪num_epochs, i+1, total_step, loss.item()))

# Test the model
model.eval()
with torch.no_grad():
    correct = 0
    total = 0
    for images, labels in test_loader:
```

```
        images = images.to(device)
        labels = labels.to(device)
        outputs = model(images)
        _, predicted = torch.max(outputs.data, 1)
        total += labels.size(0)
        correct += (predicted == labels).sum().item()

    print('Test Accuracy of the model on the 10000 test images: {} %'.
↪format(100 * correct / total))
```

```
Downloading https://www.cs.toronto.edu/~kriz/cifar-10-python.tar.gz to
./data/cifar-10-python.tar.gz
```

```
100%|      | 170498071/170498071 [00:13<00:00, 12715387.02it/s]
```

```
Extracting ./data/cifar-10-python.tar.gz to ./data
Files already downloaded and verified
Epoch [1/10], Step [100/782], Loss: 1.6014
Epoch [1/10], Step [200/782], Loss: 1.7285
Epoch [1/10], Step [300/782], Loss: 1.2765
Epoch [1/10], Step [400/782], Loss: 1.2570
Epoch [1/10], Step [500/782], Loss: 1.2392
Epoch [1/10], Step [600/782], Loss: 0.9949
Epoch [1/10], Step [700/782], Loss: 1.2892
Epoch [2/10], Step [100/782], Loss: 1.2479
Epoch [2/10], Step [200/782], Loss: 1.0150
Epoch [2/10], Step [300/782], Loss: 1.0945
Epoch [2/10], Step [400/782], Loss: 0.8884
Epoch [2/10], Step [500/782], Loss: 0.7958
Epoch [2/10], Step [600/782], Loss: 1.0781
Epoch [2/10], Step [700/782], Loss: 0.9117
Epoch [3/10], Step [100/782], Loss: 1.1361
Epoch [3/10], Step [200/782], Loss: 0.9273
Epoch [3/10], Step [300/782], Loss: 0.8810
Epoch [3/10], Step [400/782], Loss: 0.9124
Epoch [3/10], Step [500/782], Loss: 0.7568
Epoch [3/10], Step [600/782], Loss: 0.9125
Epoch [3/10], Step [700/782], Loss: 0.8653
Epoch [4/10], Step [100/782], Loss: 0.7774
Epoch [4/10], Step [200/782], Loss: 0.7719
Epoch [4/10], Step [300/782], Loss: 0.7761
Epoch [4/10], Step [400/782], Loss: 0.5034
Epoch [4/10], Step [500/782], Loss: 0.6923
Epoch [4/10], Step [600/782], Loss: 0.9134
Epoch [4/10], Step [700/782], Loss: 0.7542
Epoch [5/10], Step [100/782], Loss: 0.6674
Epoch [5/10], Step [200/782], Loss: 0.8065
Epoch [5/10], Step [300/782], Loss: 0.6523
```

```
Epoch [5/10], Step [400/782], Loss: 0.5765
Epoch [5/10], Step [500/782], Loss: 0.7250
Epoch [5/10], Step [600/782], Loss: 0.6852
Epoch [5/10], Step [700/782], Loss: 0.5991
Epoch [6/10], Step [100/782], Loss: 0.6234
Epoch [6/10], Step [200/782], Loss: 0.6914
Epoch [6/10], Step [300/782], Loss: 0.6816
Epoch [6/10], Step [400/782], Loss: 0.5768
Epoch [6/10], Step [500/782], Loss: 0.6364
Epoch [6/10], Step [600/782], Loss: 0.4679
Epoch [6/10], Step [700/782], Loss: 0.5848
Epoch [7/10], Step [100/782], Loss: 0.4693
Epoch [7/10], Step [200/782], Loss: 0.5736
Epoch [7/10], Step [300/782], Loss: 0.4834
Epoch [7/10], Step [400/782], Loss: 0.5990
Epoch [7/10], Step [500/782], Loss: 0.4021
Epoch [7/10], Step [600/782], Loss: 0.5518
Epoch [7/10], Step [700/782], Loss: 0.3942
Epoch [8/10], Step [100/782], Loss: 0.4133
Epoch [8/10], Step [200/782], Loss: 0.3664
Epoch [8/10], Step [300/782], Loss: 0.2451
Epoch [8/10], Step [400/782], Loss: 0.4930
Epoch [8/10], Step [500/782], Loss: 0.3893
Epoch [8/10], Step [600/782], Loss: 0.5398
Epoch [8/10], Step [700/782], Loss: 0.5041
Epoch [9/10], Step [100/782], Loss: 0.4044
Epoch [9/10], Step [200/782], Loss: 0.3716
Epoch [9/10], Step [300/782], Loss: 0.4640
Epoch [9/10], Step [400/782], Loss: 0.4240
Epoch [9/10], Step [500/782], Loss: 0.4071
Epoch [9/10], Step [600/782], Loss: 0.6575
Epoch [9/10], Step [700/782], Loss: 0.5266
Epoch [10/10], Step [100/782], Loss: 0.2383
Epoch [10/10], Step [200/782], Loss: 0.2672
Epoch [10/10], Step [300/782], Loss: 0.6874
Epoch [10/10], Step [400/782], Loss: 0.2830
Epoch [10/10], Step [500/782], Loss: 0.3002
Epoch [10/10], Step [600/782], Loss: 0.5157
Epoch [10/10], Step [700/782], Loss: 0.3185
Test Accuracy of the model on the 10000 test images: 73.94 %
```

[ ]:
```python
import torch
import torchvision
import torchvision.transforms as transforms
import torch.nn as nn
import torch.optim as optim
```

```
# Device configuration
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')

# Hyperparameters
num_epochs = 100
batch_size = 64
learning_rate = 0.001
momentum = 0.9
weight_decay = 1e-4
dropout_rate = 0.3

# CIFAR-10 dataset with data augmentation
transform_train = transforms.Compose([
    transforms.RandomHorizontalFlip(),
    transforms.RandomCrop(32, padding=4),
    transforms.ToTensor(),
    transforms.Normalize((0.5, 0.5, 0.5), (0.5, 0.5, 0.5)),
])

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

train_dataset = torchvision.datasets.CIFAR10(root='./data', train=True,␣
 ↪download=True, transform=transform_train)
train_loader = torch.utils.data.DataLoader(train_dataset,␣
 ↪batch_size=batch_size, shuffle=True)

test_dataset = torchvision.datasets.CIFAR10(root='./data', train=False,␣
 ↪download=True, transform=transform_test)
test_loader = torch.utils.data.DataLoader(test_dataset, batch_size=batch_size,␣
 ↪shuffle=False)

# Define the model
class CNN(nn.Module):
    def __init__(self):
        super(CNN, self).__init__()
        self.conv1 = nn.Conv2d(3, 32, kernel_size=3, padding=1)
        self.bn1 = nn.BatchNorm2d(32)
        self.conv2 = nn.Conv2d(32, 64, kernel_size=3, padding=1)
        self.bn2 = nn.BatchNorm2d(64)
        self.conv3 = nn.Conv2d(64, 128, kernel_size=3, padding=1)
        self.bn3 = nn.BatchNorm2d(128)
        self.pool = nn.MaxPool2d(2, 2)
        self.fc1 = nn.Linear(128 * 4 * 4, 512)
        self.fc2 = nn.Linear(512, 10)
```

```
        self.dropout = nn.Dropout(dropout_rate)

    def forward(self, x):
        x = self.pool(torch.relu(self.bn1(self.conv1(x))))
        x = self.pool(torch.relu(self.bn2(self.conv2(x))))
        x = self.pool(torch.relu(self.bn3(self.conv3(x))))
        x = x.view(-1, 128 * 4 * 4)
        x = self.dropout(torch.relu(self.fc1(x)))
        x = self.fc2(x)
        return x

model = CNN().to(device)

# Loss and optimizer
criterion = nn.CrossEntropyLoss()
optimizer = optim.SGD(model.parameters(), lr=learning_rate, momentum=momentum,␣
 ↪weight_decay=weight_decay)

# Learning rate scheduler
scheduler = optim.lr_scheduler.MultiStepLR(optimizer, milestones=[10, 15],␣
 ↪gamma=0.1)

# Train the model
total_step = len(train_loader)
for epoch in range(num_epochs):
    model.train()
    for i, (images, labels) in enumerate(train_loader):
        images = images.to(device)
        labels = labels.to(device)

        # Forward pass
        outputs = model(images)
        loss = criterion(outputs, labels)

        # Backward and optimize
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()

        if (i+1) % 100 == 0:
            print ('Epoch [{}/{}], Step [{}/{}], Loss: {:.4f}'.format(epoch+1,␣
 ↪num_epochs, i+1, total_step, loss.item()))

    # Decay learning rate
    scheduler.step()

# Test the model
```

```
model.eval()
with torch.no_grad():
    correct = 0
    total = 0
    for images, labels in test_loader:
        images = images.to(device)
        labels = labels.to(device)
        outputs = model(images)
        _, predicted = torch.max(outputs.data, 1)
        total += labels.size(0)
        correct += (predicted == labels).sum().item()

    print('Test Accuracy of the model on the 10000 test images: {} %'.
 ↪format(100 * correct / total))
```

```
Files already downloaded and verified
Files already downloaded and verified
Epoch [1/100], Step [100/782], Loss: 1.9820
Epoch [1/100], Step [200/782], Loss: 1.7135
Epoch [1/100], Step [300/782], Loss: 1.6176
Epoch [1/100], Step [400/782], Loss: 1.4844
Epoch [1/100], Step [500/782], Loss: 1.6111
Epoch [1/100], Step [600/782], Loss: 1.6146
Epoch [1/100], Step [700/782], Loss: 1.3368
Epoch [2/100], Step [100/782], Loss: 1.5575
Epoch [2/100], Step [200/782], Loss: 1.3339
Epoch [2/100], Step [300/782], Loss: 1.4902
Epoch [2/100], Step [400/782], Loss: 1.2101
Epoch [2/100], Step [500/782], Loss: 1.1109
Epoch [2/100], Step [600/782], Loss: 0.9728
Epoch [2/100], Step [700/782], Loss: 1.0572
Epoch [3/100], Step [100/782], Loss: 1.3173
Epoch [3/100], Step [200/782], Loss: 1.2741
Epoch [3/100], Step [300/782], Loss: 1.4302
Epoch [3/100], Step [400/782], Loss: 1.2798
Epoch [3/100], Step [500/782], Loss: 1.4131
Epoch [3/100], Step [600/782], Loss: 1.2670
Epoch [3/100], Step [700/782], Loss: 1.1009
Epoch [4/100], Step [100/782], Loss: 1.1808
Epoch [4/100], Step [200/782], Loss: 1.0415
Epoch [4/100], Step [300/782], Loss: 1.2919
Epoch [4/100], Step [400/782], Loss: 1.1713
Epoch [4/100], Step [500/782], Loss: 1.0211
Epoch [4/100], Step [600/782], Loss: 0.9457
Epoch [4/100], Step [700/782], Loss: 1.1088
Epoch [5/100], Step [100/782], Loss: 1.0276
Epoch [5/100], Step [200/782], Loss: 1.0364
```

```
Epoch [5/100], Step [300/782], Loss: 1.1126
Epoch [5/100], Step [400/782], Loss: 1.0716
Epoch [5/100], Step [500/782], Loss: 1.0382
Epoch [5/100], Step [600/782], Loss: 1.0881
Epoch [5/100], Step [700/782], Loss: 1.0051
Epoch [6/100], Step [100/782], Loss: 0.8950
Epoch [6/100], Step [200/782], Loss: 0.9201
Epoch [6/100], Step [300/782], Loss: 0.9283
Epoch [6/100], Step [400/782], Loss: 0.8832
Epoch [6/100], Step [500/782], Loss: 0.8411
Epoch [6/100], Step [600/782], Loss: 0.9008
Epoch [6/100], Step [700/782], Loss: 1.0167
Epoch [7/100], Step [100/782], Loss: 0.8144
Epoch [7/100], Step [200/782], Loss: 0.8912
Epoch [7/100], Step [300/782], Loss: 1.0213
Epoch [7/100], Step [400/782], Loss: 0.8483
Epoch [7/100], Step [500/782], Loss: 0.7289
Epoch [7/100], Step [600/782], Loss: 0.9009
Epoch [7/100], Step [700/782], Loss: 0.7400
Epoch [8/100], Step [100/782], Loss: 0.7165
Epoch [8/100], Step [200/782], Loss: 0.9253
Epoch [8/100], Step [300/782], Loss: 0.8996
Epoch [8/100], Step [400/782], Loss: 0.8354
Epoch [8/100], Step [500/782], Loss: 0.7923
Epoch [8/100], Step [600/782], Loss: 0.9129
Epoch [8/100], Step [700/782], Loss: 1.0218
Epoch [9/100], Step [100/782], Loss: 0.9692
Epoch [9/100], Step [200/782], Loss: 0.8451
Epoch [9/100], Step [300/782], Loss: 0.7376
Epoch [9/100], Step [400/782], Loss: 0.7111
Epoch [9/100], Step [500/782], Loss: 0.6798
Epoch [9/100], Step [600/782], Loss: 0.9408
Epoch [9/100], Step [700/782], Loss: 0.8922
Epoch [10/100], Step [100/782], Loss: 0.7035
Epoch [10/100], Step [200/782], Loss: 0.6602
Epoch [10/100], Step [300/782], Loss: 0.8417
Epoch [10/100], Step [400/782], Loss: 0.9840
Epoch [10/100], Step [500/782], Loss: 0.5727
Epoch [10/100], Step [600/782], Loss: 0.7210
Epoch [10/100], Step [700/782], Loss: 0.6983
Epoch [11/100], Step [100/782], Loss: 0.7510
Epoch [11/100], Step [200/782], Loss: 0.8897
Epoch [11/100], Step [300/782], Loss: 0.8210
Epoch [11/100], Step [400/782], Loss: 0.7187
Epoch [11/100], Step [500/782], Loss: 0.8820
Epoch [11/100], Step [600/782], Loss: 0.7166
Epoch [11/100], Step [700/782], Loss: 0.7315
Epoch [12/100], Step [100/782], Loss: 0.7706
```

```
Epoch [12/100], Step [200/782], Loss: 0.9119
Epoch [12/100], Step [300/782], Loss: 0.8220
Epoch [12/100], Step [400/782], Loss: 0.8740
Epoch [12/100], Step [500/782], Loss: 0.8691
Epoch [12/100], Step [600/782], Loss: 0.8021
Epoch [12/100], Step [700/782], Loss: 0.7710
Epoch [13/100], Step [100/782], Loss: 0.7688
Epoch [13/100], Step [200/782], Loss: 0.8747
Epoch [13/100], Step [300/782], Loss: 0.7091
Epoch [13/100], Step [400/782], Loss: 0.9241
Epoch [13/100], Step [500/782], Loss: 0.6018
Epoch [13/100], Step [600/782], Loss: 0.7249
Epoch [13/100], Step [700/782], Loss: 0.7430
Epoch [14/100], Step [100/782], Loss: 0.7585
Epoch [14/100], Step [200/782], Loss: 0.5901
Epoch [14/100], Step [300/782], Loss: 0.6974
Epoch [14/100], Step [400/782], Loss: 0.7836
Epoch [14/100], Step [500/782], Loss: 0.8946
Epoch [14/100], Step [600/782], Loss: 0.8931
Epoch [14/100], Step [700/782], Loss: 0.9654
Epoch [15/100], Step [100/782], Loss: 0.9905
Epoch [15/100], Step [200/782], Loss: 0.8753
Epoch [15/100], Step [300/782], Loss: 0.6686
Epoch [15/100], Step [400/782], Loss: 0.7310
Epoch [15/100], Step [500/782], Loss: 0.6135
Epoch [15/100], Step [600/782], Loss: 0.5091
Epoch [15/100], Step [700/782], Loss: 0.6115
Epoch [16/100], Step [100/782], Loss: 0.6715
Epoch [16/100], Step [200/782], Loss: 0.9117
Epoch [16/100], Step [300/782], Loss: 0.7698
Epoch [16/100], Step [400/782], Loss: 0.6805
Epoch [16/100], Step [500/782], Loss: 0.5869
Epoch [16/100], Step [600/782], Loss: 0.5877
Epoch [16/100], Step [700/782], Loss: 0.6925
Epoch [17/100], Step [100/782], Loss: 0.5774
Epoch [17/100], Step [200/782], Loss: 0.9783
Epoch [17/100], Step [300/782], Loss: 0.6630
Epoch [17/100], Step [400/782], Loss: 0.7256
Epoch [17/100], Step [500/782], Loss: 0.7685
Epoch [17/100], Step [600/782], Loss: 0.7111
Epoch [17/100], Step [700/782], Loss: 0.8790
Epoch [18/100], Step [100/782], Loss: 0.6174
Epoch [18/100], Step [200/782], Loss: 0.7509
Epoch [18/100], Step [300/782], Loss: 0.8179
Epoch [18/100], Step [400/782], Loss: 0.5431
Epoch [18/100], Step [500/782], Loss: 0.6102
Epoch [18/100], Step [600/782], Loss: 0.9106
Epoch [18/100], Step [700/782], Loss: 0.7306
```

```
Epoch [19/100], Step [100/782], Loss: 0.6570
Epoch [19/100], Step [200/782], Loss: 0.6658
Epoch [19/100], Step [300/782], Loss: 0.6812
Epoch [19/100], Step [400/782], Loss: 0.7050
Epoch [19/100], Step [500/782], Loss: 0.5241
Epoch [19/100], Step [600/782], Loss: 0.7561
Epoch [19/100], Step [700/782], Loss: 0.7893
Epoch [20/100], Step [100/782], Loss: 0.5788
Epoch [20/100], Step [200/782], Loss: 0.7439
Epoch [20/100], Step [300/782], Loss: 0.6820
Epoch [20/100], Step [400/782], Loss: 0.3870
Epoch [20/100], Step [500/782], Loss: 0.8762
Epoch [20/100], Step [600/782], Loss: 0.6887
Epoch [20/100], Step [700/782], Loss: 0.9187
Epoch [21/100], Step [100/782], Loss: 0.6467
Epoch [21/100], Step [200/782], Loss: 0.7200
Epoch [21/100], Step [300/782], Loss: 0.7508
Epoch [21/100], Step [400/782], Loss: 0.7453
Epoch [21/100], Step [500/782], Loss: 0.6002
Epoch [21/100], Step [600/782], Loss: 0.7976
Epoch [21/100], Step [700/782], Loss: 1.0220
Epoch [22/100], Step [100/782], Loss: 0.6428
Epoch [22/100], Step [200/782], Loss: 0.8207
Epoch [22/100], Step [300/782], Loss: 0.6540
Epoch [22/100], Step [400/782], Loss: 0.6402
Epoch [22/100], Step [500/782], Loss: 0.8542
Epoch [22/100], Step [600/782], Loss: 0.8238
Epoch [22/100], Step [700/782], Loss: 0.7370
Epoch [23/100], Step [100/782], Loss: 0.8082
Epoch [23/100], Step [200/782], Loss: 0.5412
Epoch [23/100], Step [300/782], Loss: 0.5892
Epoch [23/100], Step [400/782], Loss: 0.5808
Epoch [23/100], Step [500/782], Loss: 0.6002
Epoch [23/100], Step [600/782], Loss: 0.6784
Epoch [23/100], Step [700/782], Loss: 0.6276
Epoch [24/100], Step [100/782], Loss: 0.7268
Epoch [24/100], Step [200/782], Loss: 0.7844
Epoch [24/100], Step [300/782], Loss: 0.7731
Epoch [24/100], Step [400/782], Loss: 0.8101
Epoch [24/100], Step [500/782], Loss: 0.6456
Epoch [24/100], Step [600/782], Loss: 0.7214
Epoch [24/100], Step [700/782], Loss: 0.6912
Epoch [25/100], Step [100/782], Loss: 0.5662
Epoch [25/100], Step [200/782], Loss: 0.8222
Epoch [25/100], Step [300/782], Loss: 0.5861
Epoch [25/100], Step [400/782], Loss: 0.7419
Epoch [25/100], Step [500/782], Loss: 0.7012
Epoch [25/100], Step [600/782], Loss: 0.8191
```

```
Epoch [25/100], Step [700/782], Loss: 0.6794
Epoch [26/100], Step [100/782], Loss: 0.8585
Epoch [26/100], Step [200/782], Loss: 0.8973
Epoch [26/100], Step [300/782], Loss: 0.4485
Epoch [26/100], Step [400/782], Loss: 0.6479
Epoch [26/100], Step [500/782], Loss: 0.7548
Epoch [26/100], Step [600/782], Loss: 0.6226
Epoch [26/100], Step [700/782], Loss: 0.6081
Epoch [27/100], Step [100/782], Loss: 0.6797
Epoch [27/100], Step [200/782], Loss: 0.6290
Epoch [27/100], Step [300/782], Loss: 0.6587
Epoch [27/100], Step [400/782], Loss: 0.6464
Epoch [27/100], Step [500/782], Loss: 0.7747
Epoch [27/100], Step [600/782], Loss: 0.7834
Epoch [27/100], Step [700/782], Loss: 0.7201
Epoch [28/100], Step [100/782], Loss: 0.6301
Epoch [28/100], Step [200/782], Loss: 0.6179
Epoch [28/100], Step [300/782], Loss: 0.8839
Epoch [28/100], Step [400/782], Loss: 0.8127
Epoch [28/100], Step [500/782], Loss: 0.6308
Epoch [28/100], Step [600/782], Loss: 0.3961
Epoch [28/100], Step [700/782], Loss: 0.7373
Epoch [29/100], Step [100/782], Loss: 0.7339
Epoch [29/100], Step [200/782], Loss: 0.6314
Epoch [29/100], Step [300/782], Loss: 0.6450
Epoch [29/100], Step [400/782], Loss: 0.5997
Epoch [29/100], Step [500/782], Loss: 0.5186
Epoch [29/100], Step [600/782], Loss: 0.9372
Epoch [29/100], Step [700/782], Loss: 0.8041
Epoch [30/100], Step [100/782], Loss: 0.7216
Epoch [30/100], Step [200/782], Loss: 0.7531
Epoch [30/100], Step [300/782], Loss: 0.7273
Epoch [30/100], Step [400/782], Loss: 1.0005
Epoch [30/100], Step [500/782], Loss: 0.6355
Epoch [30/100], Step [600/782], Loss: 0.6850
Epoch [30/100], Step [700/782], Loss: 0.7882
Epoch [31/100], Step [100/782], Loss: 0.7397
Epoch [31/100], Step [200/782], Loss: 0.7874
Epoch [31/100], Step [300/782], Loss: 0.6906
Epoch [31/100], Step [400/782], Loss: 0.6872
Epoch [31/100], Step [500/782], Loss: 0.6110
Epoch [31/100], Step [600/782], Loss: 0.7973
Epoch [31/100], Step [700/782], Loss: 0.7149
Epoch [32/100], Step [100/782], Loss: 0.9296
Epoch [32/100], Step [200/782], Loss: 0.6635
Epoch [32/100], Step [300/782], Loss: 0.7520
Epoch [32/100], Step [400/782], Loss: 0.8178
Epoch [32/100], Step [500/782], Loss: 0.8447
```

```
Epoch [32/100], Step [600/782], Loss: 0.6493
Epoch [32/100], Step [700/782], Loss: 0.6317
Epoch [33/100], Step [100/782], Loss: 1.0104
Epoch [33/100], Step [200/782], Loss: 0.7245
Epoch [33/100], Step [300/782], Loss: 0.5414
Epoch [33/100], Step [400/782], Loss: 0.6480
Epoch [33/100], Step [500/782], Loss: 0.7324
Epoch [33/100], Step [600/782], Loss: 0.6621
Epoch [33/100], Step [700/782], Loss: 0.6647
Epoch [34/100], Step [100/782], Loss: 0.8217
Epoch [34/100], Step [200/782], Loss: 0.9496
Epoch [34/100], Step [300/782], Loss: 0.6686
Epoch [34/100], Step [400/782], Loss: 0.7667
Epoch [34/100], Step [500/782], Loss: 0.5906
Epoch [34/100], Step [600/782], Loss: 0.8597
Epoch [34/100], Step [700/782], Loss: 0.5902
Epoch [35/100], Step [100/782], Loss: 0.8363
Epoch [35/100], Step [200/782], Loss: 0.6627
Epoch [35/100], Step [300/782], Loss: 0.9780
Epoch [35/100], Step [400/782], Loss: 0.6763
Epoch [35/100], Step [500/782], Loss: 0.6249
Epoch [35/100], Step [600/782], Loss: 0.6695
Epoch [35/100], Step [700/782], Loss: 0.5557
Epoch [36/100], Step [100/782], Loss: 0.7523
Epoch [36/100], Step [200/782], Loss: 0.5224
Epoch [36/100], Step [300/782], Loss: 0.6955
Epoch [36/100], Step [400/782], Loss: 0.5853
Epoch [36/100], Step [500/782], Loss: 0.7270
Epoch [36/100], Step [600/782], Loss: 0.8534
Epoch [36/100], Step [700/782], Loss: 0.8012
Epoch [37/100], Step [100/782], Loss: 0.5348
Epoch [37/100], Step [200/782], Loss: 0.5921
Epoch [37/100], Step [300/782], Loss: 0.6342
Epoch [37/100], Step [400/782], Loss: 0.7201
Epoch [37/100], Step [500/782], Loss: 0.6090
Epoch [37/100], Step [600/782], Loss: 0.7557
Epoch [37/100], Step [700/782], Loss: 0.6363
Epoch [38/100], Step [100/782], Loss: 0.6023
Epoch [38/100], Step [200/782], Loss: 0.6247
Epoch [38/100], Step [300/782], Loss: 0.7205
Epoch [38/100], Step [400/782], Loss: 0.6516
Epoch [38/100], Step [500/782], Loss: 0.6508
Epoch [38/100], Step [600/782], Loss: 0.6014
Epoch [38/100], Step [700/782], Loss: 0.7184
Epoch [39/100], Step [100/782], Loss: 0.5838
Epoch [39/100], Step [200/782], Loss: 0.9050
Epoch [39/100], Step [300/782], Loss: 0.7816
Epoch [39/100], Step [400/782], Loss: 0.6280
```

```
Epoch [39/100], Step [500/782], Loss: 0.8978
Epoch [39/100], Step [600/782], Loss: 0.7733
Epoch [39/100], Step [700/782], Loss: 0.6808
Epoch [40/100], Step [100/782], Loss: 0.6878
Epoch [40/100], Step [200/782], Loss: 0.6209
Epoch [40/100], Step [300/782], Loss: 0.7579
Epoch [40/100], Step [400/782], Loss: 0.5953
Epoch [40/100], Step [500/782], Loss: 0.6998
Epoch [40/100], Step [600/782], Loss: 0.5485
Epoch [40/100], Step [700/782], Loss: 0.7560
Epoch [41/100], Step [100/782], Loss: 0.8624
Epoch [41/100], Step [200/782], Loss: 0.6839
Epoch [41/100], Step [300/782], Loss: 0.7464
Epoch [41/100], Step [400/782], Loss: 0.8471
Epoch [41/100], Step [500/782], Loss: 0.7898
Epoch [41/100], Step [600/782], Loss: 0.6198
Epoch [41/100], Step [700/782], Loss: 0.7713
Epoch [42/100], Step [100/782], Loss: 0.5291
Epoch [42/100], Step [200/782], Loss: 0.7188
Epoch [42/100], Step [300/782], Loss: 0.7319
Epoch [42/100], Step [400/782], Loss: 0.6948
Epoch [42/100], Step [500/782], Loss: 0.8097
Epoch [42/100], Step [600/782], Loss: 0.5195
Epoch [42/100], Step [700/782], Loss: 0.4359
Epoch [43/100], Step [100/782], Loss: 0.6198
Epoch [43/100], Step [200/782], Loss: 0.6132
Epoch [43/100], Step [300/782], Loss: 0.5496
Epoch [43/100], Step [400/782], Loss: 0.5275
Epoch [43/100], Step [500/782], Loss: 0.6631
Epoch [43/100], Step [600/782], Loss: 0.5292
Epoch [43/100], Step [700/782], Loss: 0.7128
Epoch [44/100], Step [100/782], Loss: 0.7638
Epoch [44/100], Step [200/782], Loss: 0.7714
Epoch [44/100], Step [300/782], Loss: 0.6601
Epoch [44/100], Step [400/782], Loss: 0.7589
Epoch [44/100], Step [500/782], Loss: 0.6952
Epoch [44/100], Step [600/782], Loss: 0.6559
Epoch [44/100], Step [700/782], Loss: 0.8069
Epoch [45/100], Step [100/782], Loss: 0.7784
Epoch [45/100], Step [200/782], Loss: 0.7501
Epoch [45/100], Step [300/782], Loss: 0.6351
Epoch [45/100], Step [400/782], Loss: 0.5427
Epoch [45/100], Step [500/782], Loss: 0.8474
Epoch [45/100], Step [600/782], Loss: 0.7819
Epoch [45/100], Step [700/782], Loss: 0.5453
Epoch [46/100], Step [100/782], Loss: 0.5981
Epoch [46/100], Step [200/782], Loss: 0.5549
Epoch [46/100], Step [300/782], Loss: 0.5439
```

```
Epoch [46/100], Step [400/782], Loss: 0.5922
Epoch [46/100], Step [500/782], Loss: 0.8949
Epoch [46/100], Step [600/782], Loss: 0.6951
Epoch [46/100], Step [700/782], Loss: 0.4772
Epoch [47/100], Step [100/782], Loss: 0.8136
Epoch [47/100], Step [200/782], Loss: 0.6699
Epoch [47/100], Step [300/782], Loss: 0.9779
Epoch [47/100], Step [400/782], Loss: 0.6316
Epoch [47/100], Step [500/782], Loss: 0.9225
Epoch [47/100], Step [600/782], Loss: 0.6839
Epoch [47/100], Step [700/782], Loss: 0.5875
Epoch [48/100], Step [100/782], Loss: 0.5970
Epoch [48/100], Step [200/782], Loss: 0.4605
Epoch [48/100], Step [300/782], Loss: 0.8678
Epoch [48/100], Step [400/782], Loss: 0.9416
Epoch [48/100], Step [500/782], Loss: 0.7134
Epoch [48/100], Step [600/782], Loss: 0.7772
Epoch [48/100], Step [700/782], Loss: 0.5826
Epoch [49/100], Step [100/782], Loss: 0.5097
Epoch [49/100], Step [200/782], Loss: 0.5178
Epoch [49/100], Step [300/782], Loss: 0.7850
Epoch [49/100], Step [400/782], Loss: 0.6628
Epoch [49/100], Step [500/782], Loss: 0.6585
Epoch [49/100], Step [600/782], Loss: 0.8162
Epoch [49/100], Step [700/782], Loss: 0.6703
Epoch [50/100], Step [100/782], Loss: 0.6586
Epoch [50/100], Step [200/782], Loss: 0.8128
Epoch [50/100], Step [300/782], Loss: 0.6814
Epoch [50/100], Step [400/782], Loss: 0.9737
Epoch [50/100], Step [500/782], Loss: 0.7482
Epoch [50/100], Step [600/782], Loss: 0.5651
Epoch [50/100], Step [700/782], Loss: 0.6651
Epoch [51/100], Step [100/782], Loss: 0.8002
Epoch [51/100], Step [200/782], Loss: 0.7325
Epoch [51/100], Step [300/782], Loss: 0.6981
Epoch [51/100], Step [400/782], Loss: 0.7196
Epoch [51/100], Step [500/782], Loss: 0.5447
Epoch [51/100], Step [600/782], Loss: 0.5403
Epoch [51/100], Step [700/782], Loss: 0.7805
Epoch [52/100], Step [100/782], Loss: 0.5982
Epoch [52/100], Step [200/782], Loss: 0.6110
Epoch [52/100], Step [300/782], Loss: 0.7679
Epoch [52/100], Step [400/782], Loss: 0.7374
Epoch [52/100], Step [500/782], Loss: 0.7461
Epoch [52/100], Step [600/782], Loss: 0.7129
Epoch [52/100], Step [700/782], Loss: 0.6940
Epoch [53/100], Step [100/782], Loss: 0.8388
Epoch [53/100], Step [200/782], Loss: 0.8599
```

```
Epoch [53/100], Step [300/782], Loss: 0.7810
Epoch [53/100], Step [400/782], Loss: 0.8303
Epoch [53/100], Step [500/782], Loss: 0.6715
Epoch [53/100], Step [600/782], Loss: 0.7879
Epoch [53/100], Step [700/782], Loss: 0.8618
Epoch [54/100], Step [100/782], Loss: 0.7146
Epoch [54/100], Step [200/782], Loss: 0.5342
Epoch [54/100], Step [300/782], Loss: 0.5933
Epoch [54/100], Step [400/782], Loss: 0.6434
Epoch [54/100], Step [500/782], Loss: 0.7525
Epoch [54/100], Step [600/782], Loss: 0.7254
Epoch [54/100], Step [700/782], Loss: 0.6469
Epoch [55/100], Step [100/782], Loss: 0.7596
Epoch [55/100], Step [200/782], Loss: 0.6346
Epoch [55/100], Step [300/782], Loss: 0.5286
Epoch [55/100], Step [400/782], Loss: 0.4646
Epoch [55/100], Step [500/782], Loss: 0.7893
Epoch [55/100], Step [600/782], Loss: 0.6061
Epoch [55/100], Step [700/782], Loss: 0.6218
Epoch [56/100], Step [100/782], Loss: 0.5808
Epoch [56/100], Step [200/782], Loss: 0.7875
Epoch [56/100], Step [300/782], Loss: 0.7584
Epoch [56/100], Step [400/782], Loss: 0.6002
Epoch [56/100], Step [500/782], Loss: 0.5200
Epoch [56/100], Step [600/782], Loss: 0.7815
Epoch [56/100], Step [700/782], Loss: 0.6327
Epoch [57/100], Step [100/782], Loss: 0.6097
Epoch [57/100], Step [200/782], Loss: 0.6594
Epoch [57/100], Step [300/782], Loss: 0.5616
Epoch [57/100], Step [400/782], Loss: 0.5087
Epoch [57/100], Step [500/782], Loss: 0.7336
Epoch [57/100], Step [600/782], Loss: 0.5514
Epoch [57/100], Step [700/782], Loss: 0.7322
Epoch [58/100], Step [100/782], Loss: 0.7248
Epoch [58/100], Step [200/782], Loss: 0.7301
Epoch [58/100], Step [300/782], Loss: 0.6522
Epoch [58/100], Step [400/782], Loss: 0.6865
Epoch [58/100], Step [500/782], Loss: 0.9371
Epoch [58/100], Step [600/782], Loss: 0.6808
Epoch [58/100], Step [700/782], Loss: 0.7382
Epoch [59/100], Step [100/782], Loss: 0.7629
Epoch [59/100], Step [200/782], Loss: 0.6137
Epoch [59/100], Step [300/782], Loss: 0.6063
Epoch [59/100], Step [400/782], Loss: 0.4453
Epoch [59/100], Step [500/782], Loss: 0.7861
Epoch [59/100], Step [600/782], Loss: 0.5280
Epoch [59/100], Step [700/782], Loss: 0.6736
Epoch [60/100], Step [100/782], Loss: 0.4538
```

```
Epoch [60/100], Step [200/782], Loss: 0.5059
Epoch [60/100], Step [300/782], Loss: 0.6484
Epoch [60/100], Step [400/782], Loss: 0.4791
Epoch [60/100], Step [500/782], Loss: 0.6212
Epoch [60/100], Step [600/782], Loss: 0.8908
Epoch [60/100], Step [700/782], Loss: 0.7653
Epoch [61/100], Step [100/782], Loss: 0.6174
Epoch [61/100], Step [200/782], Loss: 0.7115
Epoch [61/100], Step [300/782], Loss: 0.9935
Epoch [61/100], Step [400/782], Loss: 0.5105
Epoch [61/100], Step [500/782], Loss: 0.5201
Epoch [61/100], Step [600/782], Loss: 0.6177
Epoch [61/100], Step [700/782], Loss: 0.6812
Epoch [62/100], Step [100/782], Loss: 0.7063
Epoch [62/100], Step [200/782], Loss: 0.8583
Epoch [62/100], Step [300/782], Loss: 0.6031
Epoch [62/100], Step [400/782], Loss: 0.7304
Epoch [62/100], Step [500/782], Loss: 0.7782
Epoch [62/100], Step [600/782], Loss: 0.6374
Epoch [62/100], Step [700/782], Loss: 0.7471
Epoch [63/100], Step [100/782], Loss: 0.5081
Epoch [63/100], Step [200/782], Loss: 0.8500
Epoch [63/100], Step [300/782], Loss: 0.5751
Epoch [63/100], Step [400/782], Loss: 0.9753
Epoch [63/100], Step [500/782], Loss: 0.4967
Epoch [63/100], Step [600/782], Loss: 0.9156
Epoch [63/100], Step [700/782], Loss: 0.7887
Epoch [64/100], Step [100/782], Loss: 0.5703
Epoch [64/100], Step [200/782], Loss: 0.6977
Epoch [64/100], Step [300/782], Loss: 0.7946
Epoch [64/100], Step [400/782], Loss: 0.7009
Epoch [64/100], Step [500/782], Loss: 0.7128
Epoch [64/100], Step [600/782], Loss: 0.5997
Epoch [64/100], Step [700/782], Loss: 0.8016
Epoch [65/100], Step [100/782], Loss: 0.6239
Epoch [65/100], Step [200/782], Loss: 0.5178
Epoch [65/100], Step [300/782], Loss: 0.5647
Epoch [65/100], Step [400/782], Loss: 0.5971
Epoch [65/100], Step [500/782], Loss: 0.5184
Epoch [65/100], Step [600/782], Loss: 0.6280
Epoch [65/100], Step [700/782], Loss: 0.7877
Epoch [66/100], Step [100/782], Loss: 0.7614
Epoch [66/100], Step [200/782], Loss: 0.5468
Epoch [66/100], Step [300/782], Loss: 0.4451
Epoch [66/100], Step [400/782], Loss: 0.8108
Epoch [66/100], Step [500/782], Loss: 0.5768
Epoch [66/100], Step [600/782], Loss: 0.6571
Epoch [66/100], Step [700/782], Loss: 0.6447
```

```
Epoch [67/100], Step [100/782], Loss: 0.6080
Epoch [67/100], Step [200/782], Loss: 0.5299
Epoch [67/100], Step [300/782], Loss: 0.7193
Epoch [67/100], Step [400/782], Loss: 0.7268
Epoch [67/100], Step [500/782], Loss: 0.4594
Epoch [67/100], Step [600/782], Loss: 0.6885
Epoch [67/100], Step [700/782], Loss: 0.7571
Epoch [68/100], Step [100/782], Loss: 0.5218
Epoch [68/100], Step [200/782], Loss: 0.6341
Epoch [68/100], Step [300/782], Loss: 0.5510
Epoch [68/100], Step [400/782], Loss: 0.6610
Epoch [68/100], Step [500/782], Loss: 0.6588
Epoch [68/100], Step [600/782], Loss: 0.6373
Epoch [68/100], Step [700/782], Loss: 0.6422
Epoch [69/100], Step [100/782], Loss: 0.7623
Epoch [69/100], Step [200/782], Loss: 0.6366
Epoch [69/100], Step [300/782], Loss: 0.7590
Epoch [69/100], Step [400/782], Loss: 0.6655
Epoch [69/100], Step [500/782], Loss: 0.4684
Epoch [69/100], Step [600/782], Loss: 0.5007
Epoch [69/100], Step [700/782], Loss: 0.8668
Epoch [70/100], Step [100/782], Loss: 0.8790
Epoch [70/100], Step [200/782], Loss: 0.6010
Epoch [70/100], Step [300/782], Loss: 0.7793
Epoch [70/100], Step [400/782], Loss: 0.7826
Epoch [70/100], Step [500/782], Loss: 0.9122
Epoch [70/100], Step [600/782], Loss: 0.6939
Epoch [70/100], Step [700/782], Loss: 0.6332
Epoch [71/100], Step [100/782], Loss: 0.6544
Epoch [71/100], Step [200/782], Loss: 0.7257
Epoch [71/100], Step [300/782], Loss: 0.5254
Epoch [71/100], Step [400/782], Loss: 0.5246
Epoch [71/100], Step [500/782], Loss: 0.4195
Epoch [71/100], Step [600/782], Loss: 0.6916
Epoch [71/100], Step [700/782], Loss: 0.7273
Epoch [72/100], Step [100/782], Loss: 0.5332
Epoch [72/100], Step [200/782], Loss: 0.9022
Epoch [72/100], Step [300/782], Loss: 0.6739
Epoch [72/100], Step [400/782], Loss: 0.7886
Epoch [72/100], Step [500/782], Loss: 0.6878
Epoch [72/100], Step [600/782], Loss: 0.8681
Epoch [72/100], Step [700/782], Loss: 0.6693
Epoch [73/100], Step [100/782], Loss: 0.7369
Epoch [73/100], Step [200/782], Loss: 0.5926
Epoch [73/100], Step [300/782], Loss: 0.8365
Epoch [73/100], Step [400/782], Loss: 0.6850
Epoch [73/100], Step [500/782], Loss: 0.5959
Epoch [73/100], Step [600/782], Loss: 0.6616
```

```
Epoch [73/100], Step [700/782], Loss: 0.6409
Epoch [74/100], Step [100/782], Loss: 0.6133
Epoch [74/100], Step [200/782], Loss: 0.5947
Epoch [74/100], Step [300/782], Loss: 0.4983
Epoch [74/100], Step [400/782], Loss: 0.7389
Epoch [74/100], Step [500/782], Loss: 0.7144
Epoch [74/100], Step [600/782], Loss: 0.4918
Epoch [74/100], Step [700/782], Loss: 0.6280
Epoch [75/100], Step [100/782], Loss: 0.4732
Epoch [75/100], Step [200/782], Loss: 0.6381
Epoch [75/100], Step [300/782], Loss: 0.5559
Epoch [75/100], Step [400/782], Loss: 0.5289
Epoch [75/100], Step [500/782], Loss: 0.6764
Epoch [75/100], Step [600/782], Loss: 0.6512
Epoch [75/100], Step [700/782], Loss: 0.5740
Epoch [76/100], Step [100/782], Loss: 0.6987
Epoch [76/100], Step [200/782], Loss: 0.3859
Epoch [76/100], Step [300/782], Loss: 0.6115
Epoch [76/100], Step [400/782], Loss: 0.6583
Epoch [76/100], Step [500/782], Loss: 0.8896
Epoch [76/100], Step [600/782], Loss: 0.8272
Epoch [76/100], Step [700/782], Loss: 0.6173
Epoch [77/100], Step [100/782], Loss: 0.5532
Epoch [77/100], Step [200/782], Loss: 0.6093
Epoch [77/100], Step [300/782], Loss: 0.5124
Epoch [77/100], Step [400/782], Loss: 0.8285
Epoch [77/100], Step [500/782], Loss: 0.8566
Epoch [77/100], Step [600/782], Loss: 0.5787
Epoch [77/100], Step [700/782], Loss: 0.4632
Epoch [78/100], Step [100/782], Loss: 0.8462
Epoch [78/100], Step [200/782], Loss: 0.7328
Epoch [78/100], Step [300/782], Loss: 0.8211
Epoch [78/100], Step [400/782], Loss: 0.6712
Epoch [78/100], Step [500/782], Loss: 0.7419
Epoch [78/100], Step [600/782], Loss: 0.6560
Epoch [78/100], Step [700/782], Loss: 0.7381
Epoch [79/100], Step [100/782], Loss: 0.7446
Epoch [79/100], Step [200/782], Loss: 0.5920
Epoch [79/100], Step [300/782], Loss: 0.5942
Epoch [79/100], Step [400/782], Loss: 0.9005
Epoch [79/100], Step [500/782], Loss: 0.7685
Epoch [79/100], Step [600/782], Loss: 0.6471
Epoch [79/100], Step [700/782], Loss: 0.6521
Epoch [80/100], Step [100/782], Loss: 0.7615
Epoch [80/100], Step [200/782], Loss: 0.6150
Epoch [80/100], Step [300/782], Loss: 0.6119
Epoch [80/100], Step [400/782], Loss: 0.6786
Epoch [80/100], Step [500/782], Loss: 0.4427
```

```
Epoch [80/100], Step [600/782], Loss: 0.5628
Epoch [80/100], Step [700/782], Loss: 0.7376
Epoch [81/100], Step [100/782], Loss: 0.5389
Epoch [81/100], Step [200/782], Loss: 0.6187
Epoch [81/100], Step [300/782], Loss: 0.8163
Epoch [81/100], Step [400/782], Loss: 0.4648
Epoch [81/100], Step [500/782], Loss: 0.7858
Epoch [81/100], Step [600/782], Loss: 0.6593
Epoch [81/100], Step [700/782], Loss: 0.8053
Epoch [82/100], Step [100/782], Loss: 0.5959
Epoch [82/100], Step [200/782], Loss: 0.7460
Epoch [82/100], Step [300/782], Loss: 0.7810
Epoch [82/100], Step [400/782], Loss: 0.8607
Epoch [82/100], Step [500/782], Loss: 0.6502
Epoch [82/100], Step [600/782], Loss: 0.5859
Epoch [82/100], Step [700/782], Loss: 0.5828
Epoch [83/100], Step [100/782], Loss: 0.5525
Epoch [83/100], Step [200/782], Loss: 0.7138
Epoch [83/100], Step [300/782], Loss: 0.4692
Epoch [83/100], Step [400/782], Loss: 0.5964
Epoch [83/100], Step [500/782], Loss: 0.5347
Epoch [83/100], Step [600/782], Loss: 0.8008
Epoch [83/100], Step [700/782], Loss: 0.6739
Epoch [84/100], Step [100/782], Loss: 0.7159
Epoch [84/100], Step [200/782], Loss: 0.7615
Epoch [84/100], Step [300/782], Loss: 0.5924
Epoch [84/100], Step [400/782], Loss: 0.5796
Epoch [84/100], Step [500/782], Loss: 0.7654
Epoch [84/100], Step [600/782], Loss: 0.7748
Epoch [84/100], Step [700/782], Loss: 0.6170
Epoch [85/100], Step [100/782], Loss: 0.7176
Epoch [85/100], Step [200/782], Loss: 0.6833
Epoch [85/100], Step [300/782], Loss: 0.7439
Epoch [85/100], Step [400/782], Loss: 0.7260
Epoch [85/100], Step [500/782], Loss: 0.8211
Epoch [85/100], Step [600/782], Loss: 0.8631
Epoch [85/100], Step [700/782], Loss: 0.6111
Epoch [86/100], Step [100/782], Loss: 0.6078
Epoch [86/100], Step [200/782], Loss: 0.5902
Epoch [86/100], Step [300/782], Loss: 0.5138
Epoch [86/100], Step [400/782], Loss: 0.5140
Epoch [86/100], Step [500/782], Loss: 0.6015
Epoch [86/100], Step [600/782], Loss: 0.7531
Epoch [86/100], Step [700/782], Loss: 0.6518
Epoch [87/100], Step [100/782], Loss: 0.7211
Epoch [87/100], Step [200/782], Loss: 0.6084
Epoch [87/100], Step [300/782], Loss: 0.9635
Epoch [87/100], Step [400/782], Loss: 0.6445
```

```
Epoch [87/100], Step [500/782], Loss: 0.8039
Epoch [87/100], Step [600/782], Loss: 0.7605
Epoch [87/100], Step [700/782], Loss: 0.6058
Epoch [88/100], Step [100/782], Loss: 0.7547
Epoch [88/100], Step [200/782], Loss: 0.7143
Epoch [88/100], Step [300/782], Loss: 0.7254
Epoch [88/100], Step [400/782], Loss: 0.5630
Epoch [88/100], Step [500/782], Loss: 0.6646
Epoch [88/100], Step [600/782], Loss: 0.6555
Epoch [88/100], Step [700/782], Loss: 0.8808
Epoch [89/100], Step [100/782], Loss: 0.7886
Epoch [89/100], Step [200/782], Loss: 0.8396
Epoch [89/100], Step [300/782], Loss: 0.4566
Epoch [89/100], Step [400/782], Loss: 0.7518
Epoch [89/100], Step [500/782], Loss: 0.7255
Epoch [89/100], Step [600/782], Loss: 0.6899
Epoch [89/100], Step [700/782], Loss: 0.8338
Epoch [90/100], Step [100/782], Loss: 0.5846
Epoch [90/100], Step [200/782], Loss: 0.5606
Epoch [90/100], Step [300/782], Loss: 0.9518
Epoch [90/100], Step [400/782], Loss: 0.7362
Epoch [90/100], Step [500/782], Loss: 0.6753
Epoch [90/100], Step [600/782], Loss: 0.6203
Epoch [90/100], Step [700/782], Loss: 0.9134
Epoch [91/100], Step [100/782], Loss: 0.6348
Epoch [91/100], Step [200/782], Loss: 0.6748
Epoch [91/100], Step [300/782], Loss: 0.6487
Epoch [91/100], Step [400/782], Loss: 0.9350
Epoch [91/100], Step [500/782], Loss: 0.8651
Epoch [91/100], Step [600/782], Loss: 0.7020
Epoch [91/100], Step [700/782], Loss: 0.6102
Epoch [92/100], Step [100/782], Loss: 0.6615
Epoch [92/100], Step [200/782], Loss: 0.5860
Epoch [92/100], Step [300/782], Loss: 0.6759
Epoch [92/100], Step [400/782], Loss: 0.4601
Epoch [92/100], Step [500/782], Loss: 0.6747
Epoch [92/100], Step [600/782], Loss: 0.4767
Epoch [92/100], Step [700/782], Loss: 1.0071
Epoch [93/100], Step [100/782], Loss: 0.7373
Epoch [93/100], Step [200/782], Loss: 0.8844
Epoch [93/100], Step [300/782], Loss: 0.8642
Epoch [93/100], Step [400/782], Loss: 0.7499
Epoch [93/100], Step [500/782], Loss: 0.5850
Epoch [93/100], Step [600/782], Loss: 0.6006
Epoch [93/100], Step [700/782], Loss: 0.7103
Epoch [94/100], Step [100/782], Loss: 0.7391
Epoch [94/100], Step [200/782], Loss: 0.7005
Epoch [94/100], Step [300/782], Loss: 0.4771
```

```
Epoch [94/100], Step [400/782], Loss: 0.5669
Epoch [94/100], Step [500/782], Loss: 1.0971
Epoch [94/100], Step [600/782], Loss: 0.5993
Epoch [94/100], Step [700/782], Loss: 0.7744
Epoch [95/100], Step [100/782], Loss: 0.8813
Epoch [95/100], Step [200/782], Loss: 0.6949
Epoch [95/100], Step [300/782], Loss: 0.7090
Epoch [95/100], Step [400/782], Loss: 0.6860
Epoch [95/100], Step [500/782], Loss: 0.6912
Epoch [95/100], Step [600/782], Loss: 0.5147
Epoch [95/100], Step [700/782], Loss: 0.6380
Epoch [96/100], Step [100/782], Loss: 0.7942
Epoch [96/100], Step [200/782], Loss: 0.5552
Epoch [96/100], Step [300/782], Loss: 0.7891
Epoch [96/100], Step [400/782], Loss: 0.7404
Epoch [96/100], Step [500/782], Loss: 0.7508
Epoch [96/100], Step [600/782], Loss: 0.5721
Epoch [96/100], Step [700/782], Loss: 0.6192
Epoch [97/100], Step [100/782], Loss: 0.4631
Epoch [97/100], Step [200/782], Loss: 0.7615
Epoch [97/100], Step [300/782], Loss: 0.6935
Epoch [97/100], Step [400/782], Loss: 0.6953
Epoch [97/100], Step [500/782], Loss: 0.6140
Epoch [97/100], Step [600/782], Loss: 0.7212
Epoch [97/100], Step [700/782], Loss: 0.7630
Epoch [98/100], Step [100/782], Loss: 0.5644
Epoch [98/100], Step [200/782], Loss: 0.6990
Epoch [98/100], Step [300/782], Loss: 0.8157
Epoch [98/100], Step [400/782], Loss: 0.5613
Epoch [98/100], Step [500/782], Loss: 0.7205
Epoch [98/100], Step [600/782], Loss: 0.6162
Epoch [98/100], Step [700/782], Loss: 0.7545
Epoch [99/100], Step [100/782], Loss: 0.7957
Epoch [99/100], Step [200/782], Loss: 0.5297
Epoch [99/100], Step [300/782], Loss: 0.6302
Epoch [99/100], Step [400/782], Loss: 0.5602
Epoch [99/100], Step [500/782], Loss: 0.7909
Epoch [99/100], Step [600/782], Loss: 0.8265
Epoch [99/100], Step [700/782], Loss: 0.6460
Epoch [100/100], Step [100/782], Loss: 0.7618
Epoch [100/100], Step [200/782], Loss: 0.7106
Epoch [100/100], Step [300/782], Loss: 0.7013
Epoch [100/100], Step [400/782], Loss: 0.6239
Epoch [100/100], Step [500/782], Loss: 0.6513
Epoch [100/100], Step [600/782], Loss: 0.8748
Epoch [100/100], Step [700/782], Loss: 0.7796
Test Accuracy of the model on the 10000 test images: 78.04 %
```