Load Libraries

In [1]:
```
import pandas as pd
from sklearn.model_selection import train_test_split
from torch.utils.data import Dataset, DataLoader
from torchvision import transforms, datasets
from PIL import Image
import torch
import torch.nn as nn
import torch.nn.functional as F
import torch.optim as optim
import matplotlib.pyplot as plt
import numpy as np
from sklearn.metrics import accuracy_score
from tqdm import tqdm
```

Load Data

In [2]:
```
# Load the dataset
train_data = pd.read_csv('train.csv')
train_data.head(5)
```

Out[2]:

	label	pixel0	pixel1	pixel2	pixel3	pixel4	pixel5	pixel6	pixel7	pixel8	...	pixel774	pixel775	pixel776	pix
0	1	0	0	0	0	0	0	0	0	0	...	0	0	0	
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
2	1	0	0	0	0	0	0	0	0	0	...	0	0	0	
3	4	0	0	0	0	0	0	0	0	0	...	0	0	0	
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	

5 rows × 785 columns

We noticed that label is in the first column!!

In [3]:
```
# Move labels to the last column in training data
train_data = train_data[[col for col in train_data.columns if col != 'label'] + ['label'

# Separate features (pixels) and labels
X_train_values, y_train = train_data.iloc[:, :-1].values.reshape(-1, 28, 28), train_data
train_data.head(5)
```

Out[3]:

	pixel0	pixel1	pixel2	pixel3	pixel4	pixel5	pixel6	pixel7	pixel8	pixel9	...	pixel775	pixel776	pixel777	pi
0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	

5 rows × 785 columns

Define a custom dataset class that inherits from PyTorch's Dataset class.

In [4]:
```
class CustomDataset(Dataset): # Define a new class named CustomDataset that inherits fro
    def __init__(self, images, labels, transform):
        self.images = images
        self.labels = labels
        self.transform = transform

    def __len__(self):
        return len(self.images)

    def __getitem__(self, idx):
        image, label = self.images[idx], self.labels[idx]
        if self.transform:
            # Add a batch dimension to the image
            image = self.transform(image.unsqueeze(0))
        return image, label
```

Apply transformations to the images

In [5]:
```
# Convert the training data from Pandas DataFrames to PyTorch tensors
X_train = torch.Tensor(X_train_values)
y_train = torch.LongTensor(y_train.values)

# Define a transformation using PyTorch's Compose class.
transform = transforms.Compose([transforms.Normalize((0.5,), (0.5,))])
# tuple 1 set mean to 0.5
# tuple 2 set std to 0.5
# (Original - mean) / std --> Range (0. 1)

#transform = transforms.Compose([
#    RandomRotation(degrees=15),          # Randomly rotate the image by up to 15 degree
#    RandomHorizontalFlip(p=0.5),         # Randomly flip the image horizontally with a
#    RandomVerticalFlip(p=0.5),           # Randomly flip the image vertically with a pr
#    transforms.Normalize((0.5,), (0.5,))
#])

# Create instances of the CustomDataset class created for the training and test datasets
train_dataset = CustomDataset(X_train, y_train, transform=transform)
# image, label

# Split the training data into training and validation sets using train_test_split() fro
train_set, val_set = train_test_split(train_dataset, test_size=0.2, random_state=42, str
```

Create data loaders using PyTorch's DataLoader class

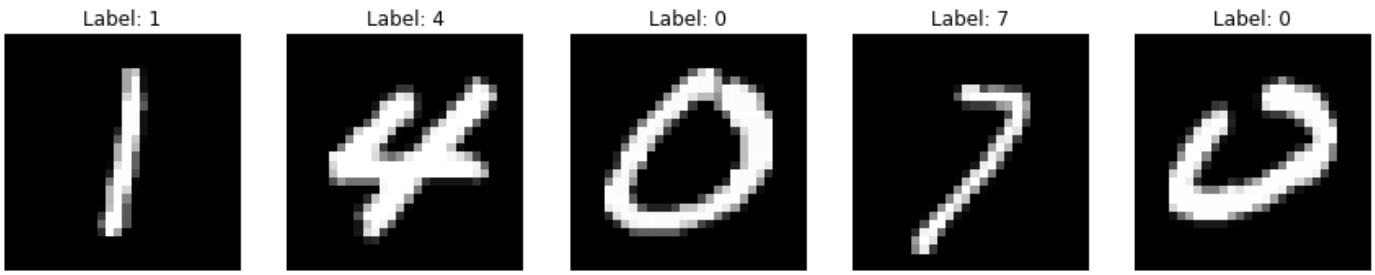
In [6]:
```
# These loaders allow for iterating over batches of data during training, validation, an
# As Batch size decrease to a certain limit accuracy increases
batch_size = 64
# The shuffle parameter is set to True for the training loader, which shuffles the data
train_loader = DataLoader(train_set, batch_size=batch_size, shuffle=True)
# The validation and test loaders have shuffle set to False since ordering doesn't matte
val_loader = DataLoader(val_set, batch_size=batch_size, shuffle=False)
```

Visualize a few samples from the training set (dataset)

In [7]:

```python
# Visualize a few samples from the dataset
def visualize_samples(data_loader, num_samples=5):
    for images, labels in data_loader:
        fig, axes = plt.subplots(1, num_samples, figsize=(15, 3))

        for i in range(num_samples):
            img = images[i].numpy().squeeze()
            label = labels[i].item()

            axes[i].imshow(img, cmap='gray')
            axes[i].set_title(f"Label: {label}")
            axes[i].axis('off')

        plt.show()
        break  # Break to only visualize one batch

# Print shapes of the datasets:
# Print shape of training data
print(f"Training set shapes - Images: {X_train.shape}, Labels: {y_train.shape}")

# Access the images and labels from the tuple in val_set
val_images, val_labels = zip(*val_set)
val_images = torch.stack(val_images)
val_labels = torch.stack(val_labels)

# Print shape of validation_set data
print(f"Validation set shapes - Images: {val_images.shape}, Labels: {val_labels.shape}")

# Note: Removed code related to y_test

# Explore class distribution in the training set
class_distribution = {i: 0 for i in range(10)}
for label in y_train:
    class_distribution[label.item()] += 1

print()
print("--------------------------------------------------")
print()

print("Class Distribution in Training Set:")
for digit, count in class_distribution.items():
    print(f"Digit {digit}: {count} samples")

print()
print("--------------------------------------------------")
print()

# Visualize a few samples from the training set
visualize_samples(train_loader)
```

```
Training set shapes - Images: torch.Size([42000, 28, 28]), Labels: torch.Size([42000])
Validation set shapes - Images: torch.Size([8400, 1, 28, 28]), Labels: torch.Size([840
0])

-------------------------------------------------

Class Distribution in Training Set:
Digit 0: 4132 samples
Digit 1: 4684 samples
Digit 2: 4177 samples
Digit 3: 4351 samples
Digit 4: 4072 samples
Digit 5: 3795 samples
Digit 6: 4137 samples
Digit 7: 4401 samples
Digit 8: 4063 samples
Digit 9: 4188 samples

-------------------------------------------------
```



Define the neural network architecture

In [68]:

```python
# Create a Model Class that inherits nn.Module (Pytorch)
class Model(nn.Module):

    # Define (Constructor) Layers Of the Neural network (using pytorch's 'nn' module)
    def __init__(self, in_features, h1, h2, out_features, dropout_prob1=0.15, dropout_pr
        # construct the nn model
        super(Model, self).__init__()

        # Use He initialization for the weights
        he_init = torch.nn.init.kaiming_normal_
        self.flatten = nn.Flatten() # unroll
        self.fc1 = nn.Linear(in_features, h1) # Defining a linear layer with input size
        self.fc2 = nn.Linear(h1, h2)
        self.out = nn.Linear(h2, out_features)

        # Dropout layers for regularization (helps prevent overfitting)
        self.dropout1 = nn.Dropout(dropout_prob1)
        self.dropout2 = nn.Dropout(dropout_prob2)

        # Layer normalization (helps stabilize the training process)
        self.layer_norm1 = nn.LayerNorm(h1)
        self.layer_norm2 = nn.LayerNorm(h2)

        # Loss function and Optimizer:

        # Measure measure the error
        self.loss = nn.CrossEntropyLoss()

        # Choose Stochastic Gradient Descent (SGD) as the optimizer (simple & Memory Eff
        # learning rate -> if error doesn't decrease after a bunch of iterations (epochs
```

```python
        # weight_decay (the regularization term) adds L2 Regularization to the optimizer
        #self.optimizer = optim.Adam(model.parameters(), lr = 0.001, weight_decay = 1e-5
        self.optimizer = optim.SGD(self.parameters(), lr = 0.01, weight_decay = 1e-5)


    # Implementation of the forward path using the reLU activation function
    def forward(self, x):
        x = self.flatten(x) # Unroll: flatten the input tensor (1D) before passing it th
        #x = x.view(-1, in_features)
        x = F.relu(self.layer_norm1(self.fc1(x)))
        x = self.dropout1(x)
        x = F.relu(self.layer_norm2(self.fc2(x)))
        x = self.dropout2(x)
        x = F.softmax(self.out(x), dim = 1)

        # X_train_softmax = F.softmax(X_train, dim=1)
        # instead of:
        # softmax_layer = nn.Softmax(dim=1)
        # X_train_softmax = softmax_layer(X_train)
        return x


    def fit(self, X, y):
        self. optimizer.zero_grad()  # To avoid accumulating gradients from previous bat
        y_predict = self.forward(X)
        loss = self.loss(y_predict, y)
        loss.backward()  # To compute the gradients of the loss with respect to the mode
        self.optimizer.step() # Update Weights
        return loss.item()

    def predict(self, X):
        with torch.no_grad():
            return torch.argmax(self.forward(X), axis=1)

# Set the input size, hidden layer sizes, and output size
in_features = 28 * 28
h1 = 128
h2 = 64
out_features = 10

# Create an object of the model -> Instantiate the model
model = Model(in_features, h1, h2, out_features)

# Pick a manual seed for randomization
torch.manual_seed(42)
```

Out[68]:

```
<torch._C.Generator at 0x12933ebe070>
```

In [69]:

```python
# Test to see how layers are going to be
model.parameters
```

Out[69]:

```
<bound method Module.parameters of Model(
  (flatten): Flatten(start_dim=1, end_dim=-1)
  (fc1): Linear(in_features=784, out_features=128, bias=True)
  (fc2): Linear(in_features=128, out_features=64, bias=True)
  (out): Linear(in_features=64, out_features=10, bias=True)
  (dropout1): Dropout(p=0.15, inplace=False)
  (dropout2): Dropout(p=0.2, inplace=False)
  (layer_norm1): LayerNorm((128,), eps=1e-05, elementwise_affine=True)
  (layer_norm2): LayerNorm((64,), eps=1e-05, elementwise_affine=True)
  (loss): CrossEntropyLoss()
)>
```

Function For Training

In [70]:
```
# Function for training
def train(model, train_loader):
    model.train()  # Set the model to training mode
    total_loss = 0.0
    correct_train = 0

    for batch_images, batch_labels in tqdm(train_loader, desc='Training'):
        # Forward pass and update weights
        loss = model.fit(batch_images, batch_labels)
        total_loss += loss
        correct_train += (model.predict(batch_images) == batch_labels).sum().item()

    # Calculate average training loss and accuracy
    average_loss = total_loss / len(train_loader)
    train_accuracy = 100 * correct_train / len(train_set)

    return average_loss, train_accuracy
```

Function For Validation

In [71]:
```
# Function for validation
def validate(model, val_loader):
    model.eval()  # Set the model to evaluation mode
    total_val_loss = 0.0
    correct_val = 0

    with torch.no_grad():
        for val_images, val_labels in val_loader:
            val_loss = model.loss(model.forward(val_images), val_labels)
            total_val_loss += val_loss.item()
            correct_val += (model.predict(val_images) == val_labels).sum().item()

    # Calculate average validation loss and accuracy
    average_val_loss = total_val_loss / len(val_loader)
    val_accuracy = 100 * correct_val / len(val_set)

    return average_val_loss, val_accuracy
```

Training Loop

In [72]:
```
# Set the number of epochs
num_epochs = 20

# Lists to store metrics for plotting
train_loss_history = []
val_loss_history = []
train_acc_history = []
val_acc_history = []

# Training loop
for epoch in range(num_epochs):
    # Training
    average_loss, train_accuracy = train(model, train_loader)
    train_loss_history.append(average_loss)
    train_acc_history.append(train_accuracy)
```

```python
    # Validation
    average_val_loss, val_accuracy = validate(model, val_loader)
    val_loss_history.append(average_val_loss)
    val_acc_history.append(val_accuracy)

    # Print and visualize metrics
    print(f'Epoch [{epoch + 1}/{num_epochs}], '
          f'Training Loss: {average_loss:.4f}, Training Accuracy: {train_accuracy:.4f}%,
          f'Validation Loss: {average_val_loss:.4f}, Validation Accuracy: {val_accuracy:
```

```
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 246.92it/s]
Epoch [1/20], Training Loss: 2.2323, Training Accuracy: 33.6161%, Validation Loss: 2.118
6, Validation Accuracy: 51.7143%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 273.49it/s]
Epoch [2/20], Training Loss: 2.0613, Training Accuracy: 54.5685%, Validation Loss: 1.950
0, Validation Accuracy: 61.3095%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 252.35it/s]
Epoch [3/20], Training Loss: 1.9321, Training Accuracy: 65.7649%, Validation Loss: 1.833
8, Validation Accuracy: 74.3452%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 262.44it/s]
Epoch [4/20], Training Loss: 1.8275, Training Accuracy: 76.9077%, Validation Loss: 1.736
6, Validation Accuracy: 81.4048%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 269.43it/s]
Epoch [5/20], Training Loss: 1.7539, Training Accuracy: 80.8155%, Validation Loss: 1.686
4, Validation Accuracy: 82.5476%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 264.26it/s]
Epoch [6/20], Training Loss: 1.7124, Training Accuracy: 82.1756%, Validation Loss: 1.660
3, Validation Accuracy: 83.4286%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 258.33it/s]
Epoch [7/20], Training Loss: 1.6869, Training Accuracy: 83.0744%, Validation Loss: 1.642
9, Validation Accuracy: 84.4881%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 273.60it/s]
Epoch [8/20], Training Loss: 1.6669, Training Accuracy: 84.4107%, Validation Loss: 1.621
8, Validation Accuracy: 87.0833%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 271.76it/s]
Epoch [9/20], Training Loss: 1.6393, Training Accuracy: 88.2500%, Validation Loss: 1.593
3, Validation Accuracy: 90.2024%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 259.76it/s]
Epoch [10/20], Training Loss: 1.6166, Training Accuracy: 90.1250%, Validation Loss: 1.57
68, Validation Accuracy: 91.2143%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 270.67it/s]
Epoch [11/20], Training Loss: 1.6021, Training Accuracy: 90.9018%, Validation Loss: 1.56
70, Validation Accuracy: 91.7619%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 275.83it/s]
Epoch [12/20], Training Loss: 1.5916, Training Accuracy: 91.5119%, Validation Loss: 1.56
00, Validation Accuracy: 92.0952%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 268.52it/s]
```

```
Epoch [13/20], Training Loss: 1.5810, Training Accuracy: 91.7976%, Validation Loss: 1.55
38, Validation Accuracy: 92.3929%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 272.70it/s]
Epoch [14/20], Training Loss: 1.5757, Training Accuracy: 92.0804%, Validation Loss: 1.54
92, Validation Accuracy: 92.5952%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 263.55it/s]
Epoch [15/20], Training Loss: 1.5705, Training Accuracy: 92.2976%, Validation Loss: 1.54
65, Validation Accuracy: 92.8571%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 262.97it/s]
Epoch [16/20], Training Loss: 1.5648, Training Accuracy: 92.7708%, Validation Loss: 1.54
32, Validation Accuracy: 92.9405%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 273.99it/s]
Epoch [17/20], Training Loss: 1.5618, Training Accuracy: 92.9464%, Validation Loss: 1.53
99, Validation Accuracy: 93.2262%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 257.96it/s]
Epoch [18/20], Training Loss: 1.5577, Training Accuracy: 93.1756%, Validation Loss: 1.53
76, Validation Accuracy: 93.3929%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 277.66it/s]
Epoch [19/20], Training Loss: 1.5536, Training Accuracy: 93.3065%, Validation Loss: 1.53
59, Validation Accuracy: 93.5238%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 268.61it/s]
Epoch [20/20], Training Loss: 1.5511, Training Accuracy: 93.5744%, Validation Loss: 1.53
43, Validation Accuracy: 93.6429%
```

In [73]:

```python
# Evaluate the model on the test set
test_loss, test_accuracy = evaluate(model, test_loader)

# Print the test results
print(f"Test Loss: {test_loss:.4f}, Test Accuracy: {test_accuracy:.4f}%")
```

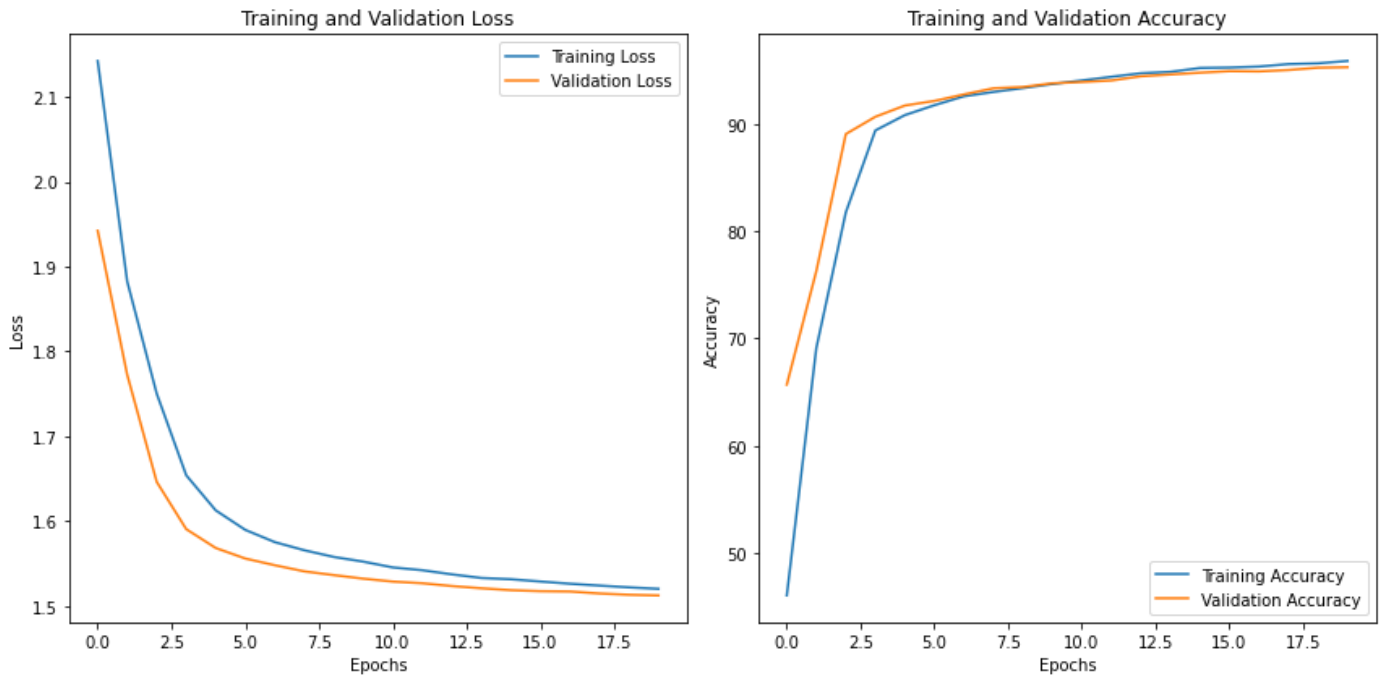
```
Test Loss: 1.7931, Test Accuracy: 69.2400%
```

Visualize Loss and Accuracy through Epochs

In [13]:

```python
# Plot training and validation loss
plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)
plt.plot(train_loss_history, label='Training Loss')
plt.plot(val_loss_history, label='Validation Loss')
plt.title('Training and Validation Loss')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()

# Plot training and validation accuracy
plt.subplot(1, 2, 2)
plt.plot(train_acc_history, label='Training Accuracy')
plt.plot(val_acc_history, label='Validation Accuracy')
plt.title('Training and Validation Accuracy')
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.legend()
```

```
plt.tight_layout()  # Adjust layout for better spacing
plt.show()
```



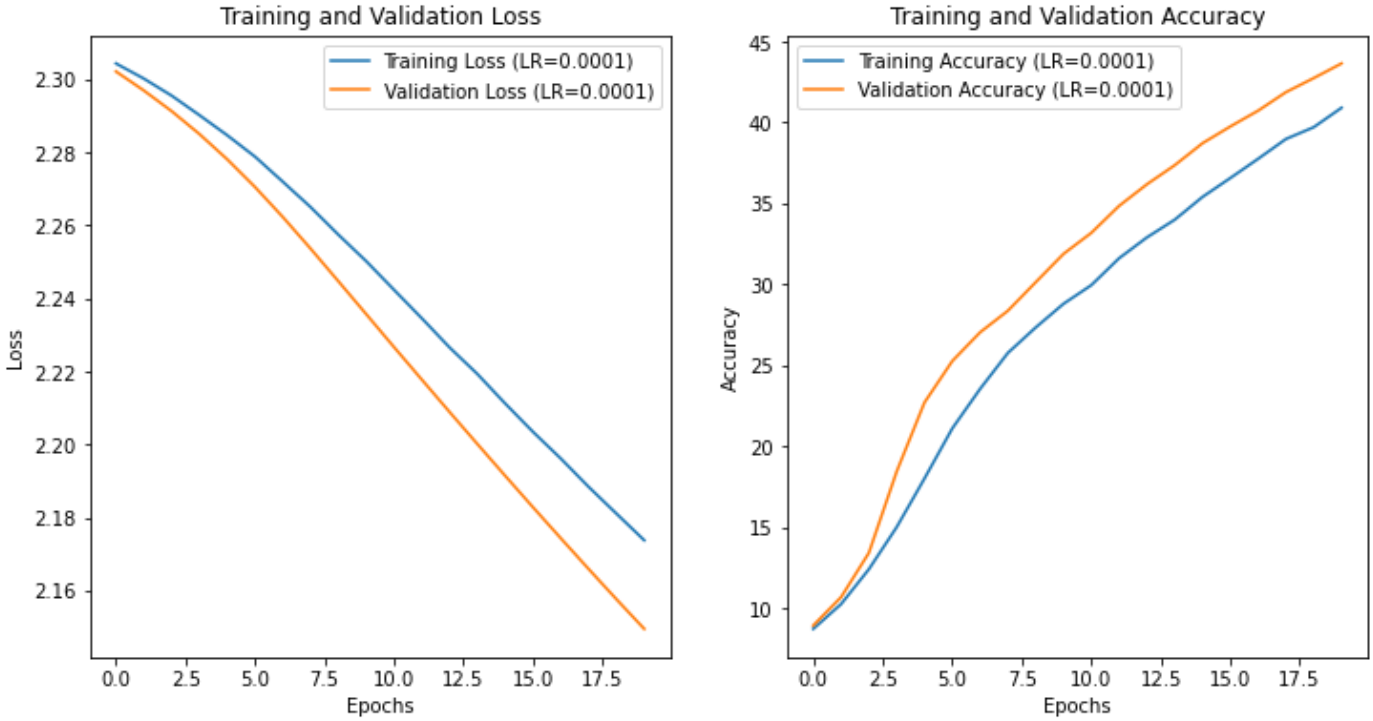
Customized Training Loop to try some different Learning rates

In [30]:
```
learning_rates = [0.0001, 0.001, 0.003, 0.01, 0.03, 0.05, 0.1, 0.3, 0.5]

# Initialize variables to keep track of the best model
best_model = None
best_val_accuracy = 0.0
best_hyperparameters = None

# Lists to store final metrics for each learning rate
final_train_losses = []
final_val_losses = []
final_train_accuracies = []
final_val_accuracies = []

for lr in learning_rates:
    # Instantiate the model
    model = Model(in_features, h1, h2, out_features)

    # Set a new learning rate
    model.optimizer = optim.SGD(model.parameters(), lr=lr, weight_decay=1e-5)

    # Lists to store metrics for plotting
    train_loss_history = []
    val_loss_history = []
    train_acc_history = []
    val_acc_history = []

    # Training loop
    for epoch in range(num_epochs):
        # Training
        average_loss, train_accuracy = train(model, train_loader)
        train_loss_history.append(average_loss)
        train_acc_history.append(train_accuracy)

        # Validation
        average_val_loss, val_accuracy = validate(model, val_loader)
        val_loss_history.append(average_val_loss)
```

```
        val_acc_history.append(val_accuracy)

        # Print and visualize metrics
        print(f'Epoch [{epoch + 1}/{num_epochs}], '
              f'Training Loss: {average_loss:.4f}, Training Accuracy: {train_accuracy:.4
              f'Validation Loss: {average_val_loss:.4f}, Validation Accuracy: {val_accur

    # Plot training and validation loss
    plt.figure(figsize=(12, 6))

    plt.subplot(1, 2, 1)
    plt.plot(train_loss_history, label=f'Training Loss (LR={lr})')
    plt.plot(val_loss_history, label=f'Validation Loss (LR={lr})')
    plt.title('Training and Validation Loss')
    plt.xlabel('Epochs')
    plt.ylabel('Loss')
    plt.legend()

    # Plot training and validation accuracy
    plt.subplot(1, 2, 2)
    plt.plot(train_acc_history, label=f'Training Accuracy (LR={lr})')
    plt.plot(val_acc_history, label=f'Validation Accuracy (LR={lr})')
    plt.title('Training and Validation Accuracy')
    plt.xlabel('Epochs')
    plt.ylabel('Accuracy')
    plt.legend()

    plt.show()

    # Store final metrics for each learning rate
    final_train_losses.append(train_loss_history[-1])
    final_val_losses.append(val_loss_history[-1])
    final_train_accuracies.append(train_acc_history[-1])
    final_val_accuracies.append(val_acc_history[-1])

    # Check if current model has the best validation accuracy
    if val_accuracy > best_val_accuracy:
        best_val_accuracy = val_accuracy
        best_model = model.state_dict().copy()
        best_hyperparameters = {'lr': lr}
```

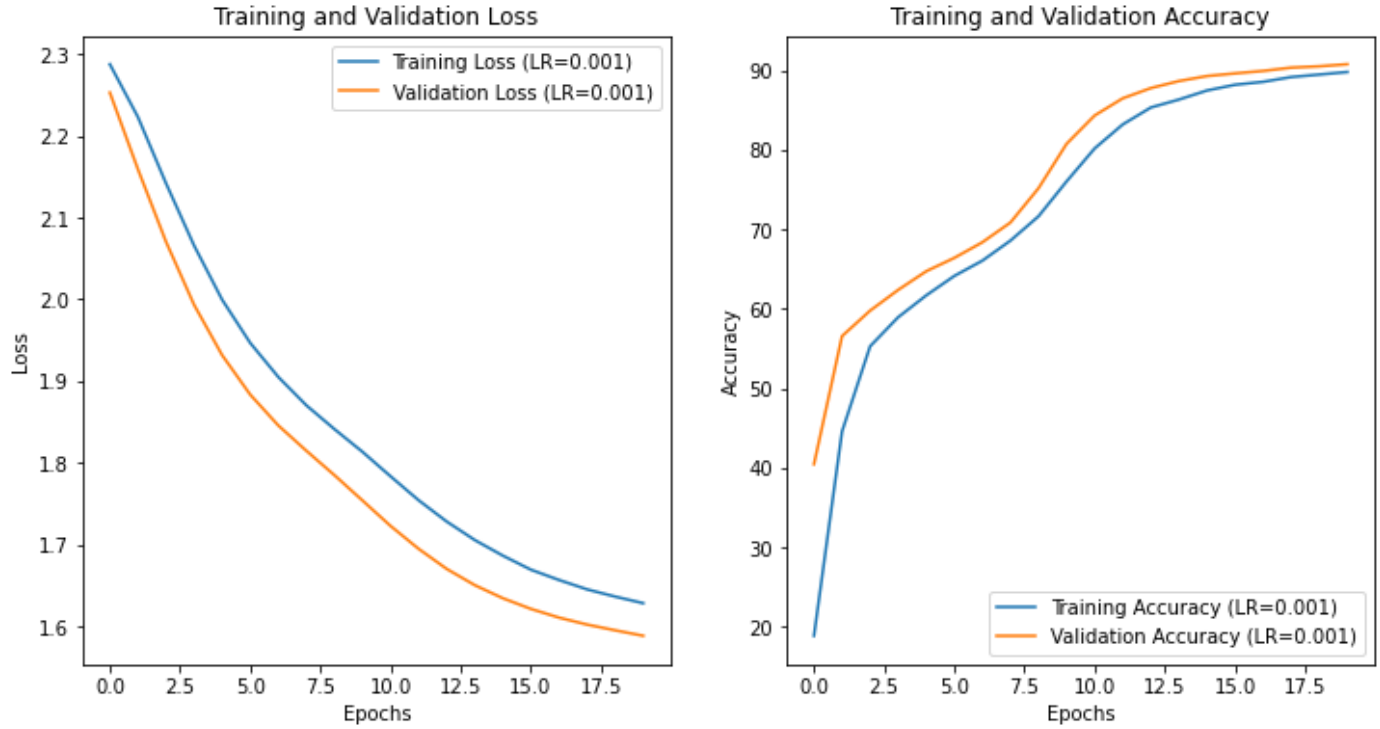
```
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 415.63it/s]
Epoch [1/20], Training Loss: 2.3044, Training Accuracy: 8.7202%, Validation Loss: 2.302
2, Validation Accuracy: 8.9286%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 420.28it/s]
Epoch [2/20], Training Loss: 2.3002, Training Accuracy: 10.2500%, Validation Loss: 2.297
0, Validation Accuracy: 10.6786%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 417.19it/s]
Epoch [3/20], Training Loss: 2.2956, Training Accuracy: 12.4048%, Validation Loss: 2.291
3, Validation Accuracy: 13.4167%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 382.14it/s]
Epoch [4/20], Training Loss: 2.2902, Training Accuracy: 15.0060%, Validation Loss: 2.285
0, Validation Accuracy: 18.4405%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 413.68it/s]
Epoch [5/20], Training Loss: 2.2847, Training Accuracy: 18.0089%, Validation Loss: 2.278
1, Validation Accuracy: 22.7024%
Training: 100%|██████████████████████████████████████████████████████████| 105
<00:00, 402.16it/s]
```

```
Epoch [6/20], Training Loss: 2.2789, Training Accuracy: 21.1012%, Validation Loss: 2.270
5, Validation Accuracy: 25.2619%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 400.38it/s]
Epoch [7/20], Training Loss: 2.2720, Training Accuracy: 23.5417%, Validation Loss: 2.262
3, Validation Accuracy: 27.0357%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 388.78it/s]
Epoch [8/20], Training Loss: 2.2650, Training Accuracy: 25.7649%, Validation Loss: 2.253
7, Validation Accuracy: 28.3690%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.54it/s]
Epoch [9/20], Training Loss: 2.2575, Training Accuracy: 27.3363%, Validation Loss: 2.244
7, Validation Accuracy: 30.1310%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 426.47it/s]
Epoch [10/20], Training Loss: 2.2502, Training Accuracy: 28.7976%, Validation Loss: 2.23
57, Validation Accuracy: 31.8810%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.43it/s]
Epoch [11/20], Training Loss: 2.2424, Training Accuracy: 29.9494%, Validation Loss: 2.22
67, Validation Accuracy: 33.1786%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.16it/s]
Epoch [12/20], Training Loss: 2.2346, Training Accuracy: 31.6131%, Validation Loss: 2.21
78, Validation Accuracy: 34.8452%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.54it/s]
Epoch [13/20], Training Loss: 2.2266, Training Accuracy: 32.8988%, Validation Loss: 2.20
89, Validation Accuracy: 36.1786%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.53it/s]
Epoch [14/20], Training Loss: 2.2193, Training Accuracy: 33.9911%, Validation Loss: 2.20
01, Validation Accuracy: 37.3571%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.73it/s]
Epoch [15/20], Training Loss: 2.2112, Training Accuracy: 35.3988%, Validation Loss: 2.19
14, Validation Accuracy: 38.7143%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.78it/s]
Epoch [16/20], Training Loss: 2.2034, Training Accuracy: 36.5655%, Validation Loss: 2.18
28, Validation Accuracy: 39.7500%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 437.94it/s]
Epoch [17/20], Training Loss: 2.1962, Training Accuracy: 37.7530%, Validation Loss: 2.17
43, Validation Accuracy: 40.7262%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.10it/s]
Epoch [18/20], Training Loss: 2.1884, Training Accuracy: 38.9762%, Validation Loss: 2.16
59, Validation Accuracy: 41.8690%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.38it/s]
Epoch [19/20], Training Loss: 2.1811, Training Accuracy: 39.7054%, Validation Loss: 2.15
76, Validation Accuracy: 42.7381%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.62it/s]
Epoch [20/20], Training Loss: 2.1738, Training Accuracy: 40.9048%, Validation Loss: 2.14
95, Validation Accuracy: 43.6429%
```



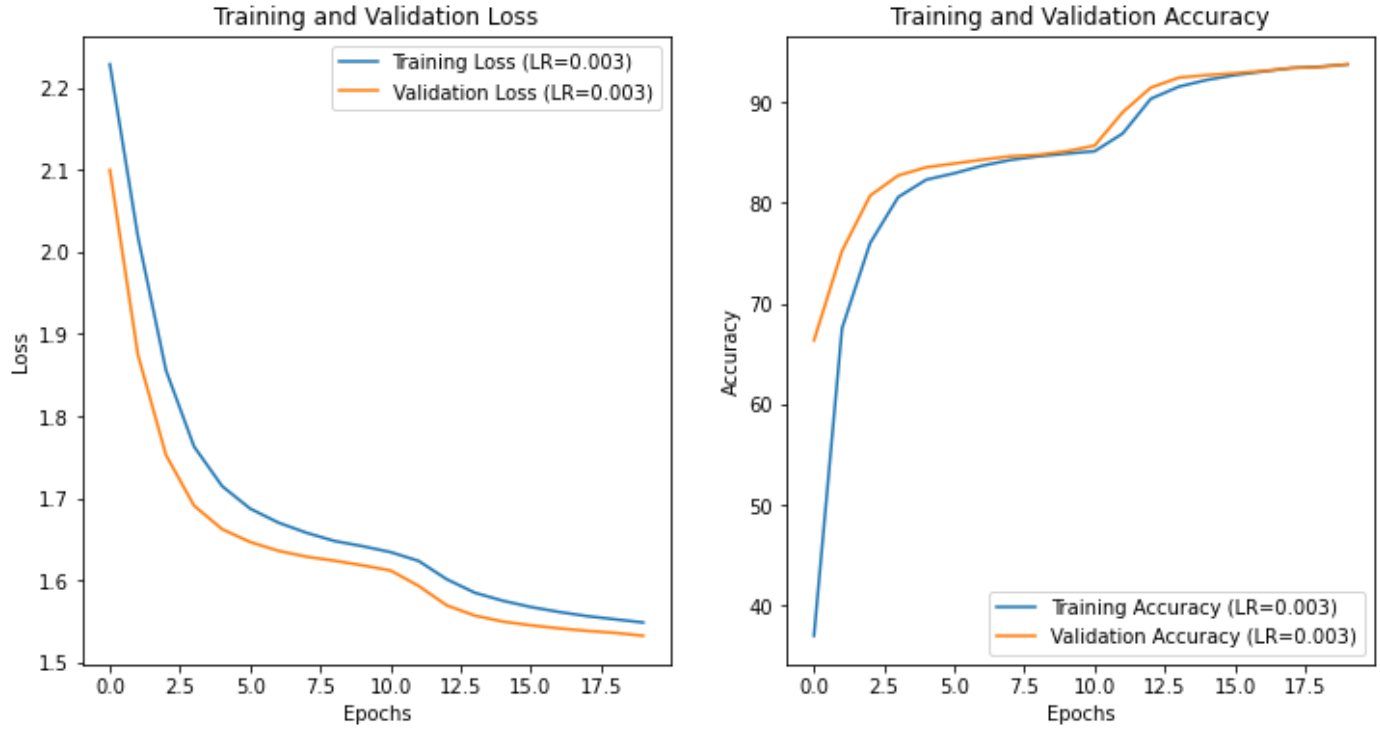

```
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.10it/s]
Epoch [1/20], Training Loss: 2.2876, Training Accuracy: 18.8899%, Validation Loss: 2.252
9, Validation Accuracy: 40.4643%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.07it/s]
Epoch [2/20], Training Loss: 2.2230, Training Accuracy: 44.5744%, Validation Loss: 2.159
1, Validation Accuracy: 56.5595%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.08it/s]
Epoch [3/20], Training Loss: 2.1419, Training Accuracy: 55.2827%, Validation Loss: 2.070
5, Validation Accuracy: 59.7619%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 431.86it/s]
Epoch [4/20], Training Loss: 2.0658, Training Accuracy: 58.9345%, Validation Loss: 1.993
3, Validation Accuracy: 62.3690%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 435.76it/s]
Epoch [5/20], Training Loss: 1.9995, Training Accuracy: 61.6815%, Validation Loss: 1.931
7, Validation Accuracy: 64.7143%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.48it/s]
Epoch [6/20], Training Loss: 1.9467, Training Accuracy: 64.1250%, Validation Loss: 1.883
4, Validation Accuracy: 66.3929%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 451.27it/s]
Epoch [7/20], Training Loss: 1.9050, Training Accuracy: 66.0446%, Validation Loss: 1.845
7, Validation Accuracy: 68.3571%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 434.69it/s]
Epoch [8/20], Training Loss: 1.8703, Training Accuracy: 68.5685%, Validation Loss: 1.814
8, Validation Accuracy: 70.8452%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.09it/s]
Epoch [9/20], Training Loss: 1.8412, Training Accuracy: 71.6250%, Validation Loss: 1.784
9, Validation Accuracy: 75.1548%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 433.23it/s]
```

```
Epoch [10/20], Training Loss: 1.8134, Training Accuracy: 76.0149%, Validation Loss: 1.75
39, Validation Accuracy: 80.7381%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.00it/s]
Epoch [11/20], Training Loss: 1.7835, Training Accuracy: 80.1607%, Validation Loss: 1.72
28, Validation Accuracy: 84.2976%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 433.49it/s]
Epoch [12/20], Training Loss: 1.7542, Training Accuracy: 83.1696%, Validation Loss: 1.69
48, Validation Accuracy: 86.4524%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 450.50it/s]
Epoch [13/20], Training Loss: 1.7282, Training Accuracy: 85.2976%, Validation Loss: 1.67
03, Validation Accuracy: 87.7381%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 440.35it/s]
Epoch [14/20], Training Loss: 1.7056, Training Accuracy: 86.2946%, Validation Loss: 1.65
04, Validation Accuracy: 88.6190%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.19it/s]
Epoch [15/20], Training Loss: 1.6865, Training Accuracy: 87.4375%, Validation Loss: 1.63
44, Validation Accuracy: 89.2500%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 438.95it/s]
Epoch [16/20], Training Loss: 1.6693, Training Accuracy: 88.1548%, Validation Loss: 1.62
13, Validation Accuracy: 89.5833%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.81it/s]
Epoch [17/20], Training Loss: 1.6567, Training Accuracy: 88.5238%, Validation Loss: 1.61
07, Validation Accuracy: 89.8929%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.95it/s]
Epoch [18/20], Training Loss: 1.6452, Training Accuracy: 89.1250%, Validation Loss: 1.60
22, Validation Accuracy: 90.3095%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.03it/s]
Epoch [19/20], Training Loss: 1.6364, Training Accuracy: 89.4375%, Validation Loss: 1.59
52, Validation Accuracy: 90.4762%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.51it/s]
Epoch [20/20], Training Loss: 1.6283, Training Accuracy: 89.7589%, Validation Loss: 1.58
85, Validation Accuracy: 90.7381%
```



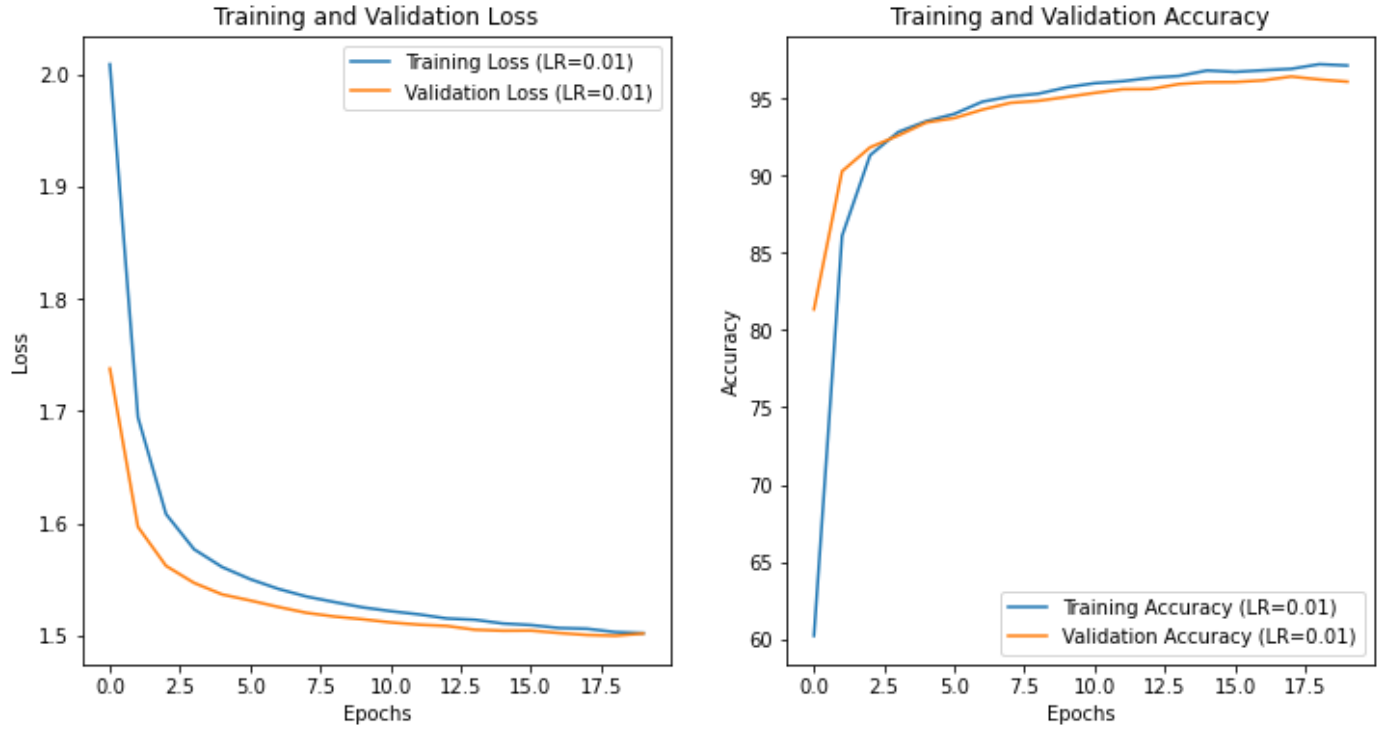

```
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.86it/s]
Epoch [1/20], Training Loss: 2.2288, Training Accuracy: 36.9702%, Validation Loss: 2.100
1, Validation Accuracy: 66.3214%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 452.29it/s]
Epoch [2/20], Training Loss: 2.0186, Training Accuracy: 67.5268%, Validation Loss: 1.874
9, Validation Accuracy: 75.2143%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.66it/s]
Epoch [3/20], Training Loss: 1.8558, Training Accuracy: 75.9911%, Validation Loss: 1.752
6, Validation Accuracy: 80.7024%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.60it/s]
Epoch [4/20], Training Loss: 1.7630, Training Accuracy: 80.5565%, Validation Loss: 1.691
1, Validation Accuracy: 82.7024%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.53it/s]
Epoch [5/20], Training Loss: 1.7148, Training Accuracy: 82.2887%, Validation Loss: 1.662
4, Validation Accuracy: 83.5238%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.38it/s]
Epoch [6/20], Training Loss: 1.6874, Training Accuracy: 82.9256%, Validation Loss: 1.646
8, Validation Accuracy: 83.8929%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.92it/s]
Epoch [7/20], Training Loss: 1.6705, Training Accuracy: 83.6815%, Validation Loss: 1.636
1, Validation Accuracy: 84.2738%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.96it/s]
Epoch [8/20], Training Loss: 1.6581, Training Accuracy: 84.2500%, Validation Loss: 1.628
9, Validation Accuracy: 84.6190%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.18it/s]
Epoch [9/20], Training Loss: 1.6481, Training Accuracy: 84.6131%, Validation Loss: 1.624
0, Validation Accuracy: 84.7619%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.86it/s]
```

```
Epoch [10/20], Training Loss: 1.6417, Training Accuracy: 84.8780%, Validation Loss: 1.61
84, Validation Accuracy: 85.1190%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.86it/s]
Epoch [11/20], Training Loss: 1.6345, Training Accuracy: 85.1161%, Validation Loss: 1.61
21, Validation Accuracy: 85.6786%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.85it/s]
Epoch [12/20], Training Loss: 1.6237, Training Accuracy: 86.8839%, Validation Loss: 1.59
33, Validation Accuracy: 89.0000%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.89it/s]
Epoch [13/20], Training Loss: 1.6014, Training Accuracy: 90.3304%, Validation Loss: 1.56
97, Validation Accuracy: 91.4405%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.35it/s]
Epoch [14/20], Training Loss: 1.5850, Training Accuracy: 91.5357%, Validation Loss: 1.55
72, Validation Accuracy: 92.4286%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.73it/s]
Epoch [15/20], Training Loss: 1.5753, Training Accuracy: 92.1994%, Validation Loss: 1.54
99, Validation Accuracy: 92.6786%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 438.85it/s]
Epoch [16/20], Training Loss: 1.5676, Training Accuracy: 92.6845%, Validation Loss: 1.54
53, Validation Accuracy: 92.8690%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.97it/s]
Epoch [17/20], Training Loss: 1.5616, Training Accuracy: 93.0357%, Validation Loss: 1.54
16, Validation Accuracy: 93.0952%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 427.10it/s]
Epoch [18/20], Training Loss: 1.5565, Training Accuracy: 93.3810%, Validation Loss: 1.53
86, Validation Accuracy: 93.3810%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.41it/s]
Epoch [19/20], Training Loss: 1.5526, Training Accuracy: 93.5179%, Validation Loss: 1.53
62, Validation Accuracy: 93.4762%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.11it/s]
Epoch [20/20], Training Loss: 1.5488, Training Accuracy: 93.7411%, Validation Loss: 1.53
26, Validation Accuracy: 93.7619%
```



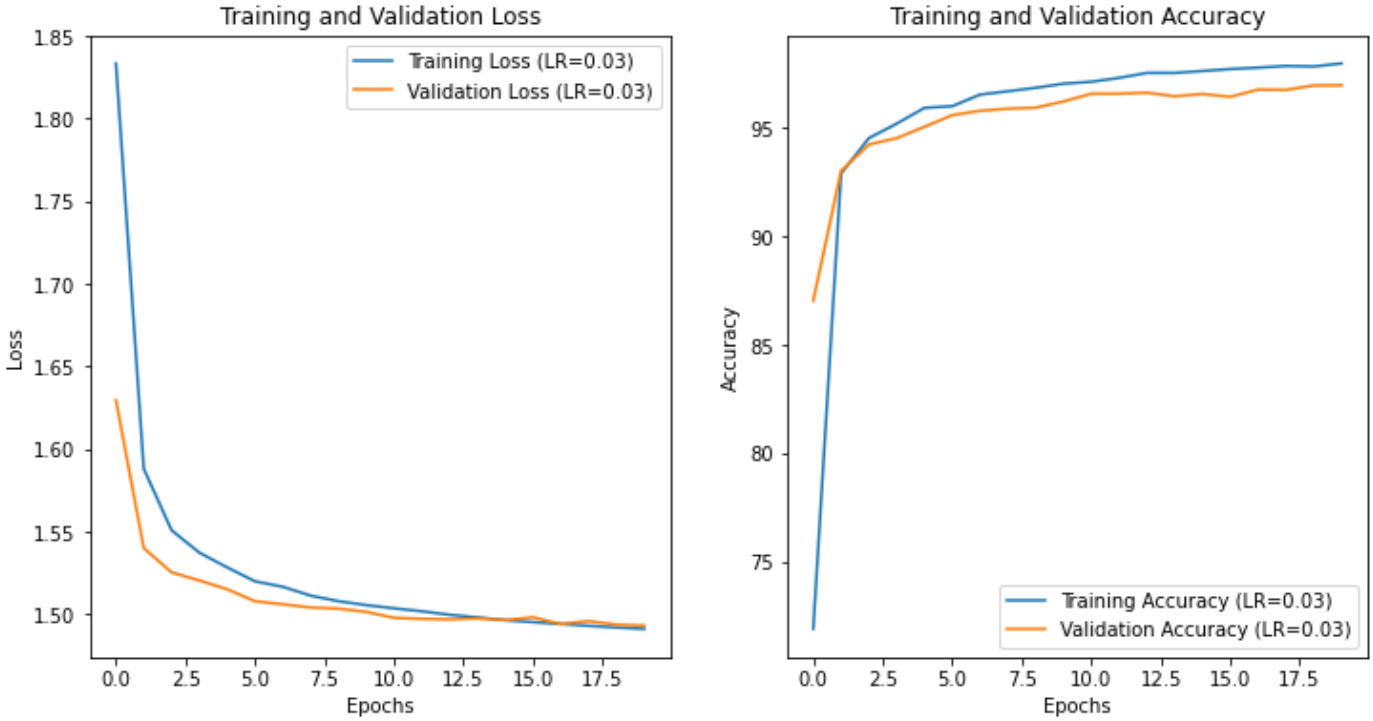

```
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.06it/s]
Epoch [1/20], Training Loss: 2.0089, Training Accuracy: 60.2113%, Validation Loss: 1.737
8, Validation Accuracy: 81.3452%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.30it/s]
Epoch [2/20], Training Loss: 1.6944, Training Accuracy: 86.0982%, Validation Loss: 1.596
8, Validation Accuracy: 90.2857%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.04it/s]
Epoch [3/20], Training Loss: 1.6082, Training Accuracy: 91.3244%, Validation Loss: 1.562
1, Validation Accuracy: 91.8333%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 437.99it/s]
Epoch [4/20], Training Loss: 1.5769, Training Accuracy: 92.8214%, Validation Loss: 1.546
9, Validation Accuracy: 92.5952%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 387.71it/s]
Epoch [5/20], Training Loss: 1.5610, Training Accuracy: 93.5268%, Validation Loss: 1.536
6, Validation Accuracy: 93.4405%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 402.98it/s]
Epoch [6/20], Training Loss: 1.5502, Training Accuracy: 93.9970%, Validation Loss: 1.531
2, Validation Accuracy: 93.7262%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 381.66it/s]
Epoch [7/20], Training Loss: 1.5416, Training Accuracy: 94.7887%, Validation Loss: 1.525
5, Validation Accuracy: 94.2738%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 423.67it/s]
Epoch [8/20], Training Loss: 1.5348, Training Accuracy: 95.1280%, Validation Loss: 1.520
3, Validation Accuracy: 94.7143%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 429.79it/s]
Epoch [9/20], Training Loss: 1.5298, Training Accuracy: 95.3065%, Validation Loss: 1.516
9, Validation Accuracy: 94.8452%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 435.52it/s]
```

```
Epoch [10/20], Training Loss: 1.5251, Training Accuracy: 95.7202%, Validation Loss: 1.51
45, Validation Accuracy: 95.0952%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.99it/s]
Epoch [11/20], Training Loss: 1.5217, Training Accuracy: 95.9851%, Validation Loss: 1.51
17, Validation Accuracy: 95.3571%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 430.50it/s]
Epoch [12/20], Training Loss: 1.5189, Training Accuracy: 96.1131%, Validation Loss: 1.50
97, Validation Accuracy: 95.5952%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 426.07it/s]
Epoch [13/20], Training Loss: 1.5152, Training Accuracy: 96.3333%, Validation Loss: 1.50
86, Validation Accuracy: 95.6071%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 414.65it/s]
Epoch [14/20], Training Loss: 1.5141, Training Accuracy: 96.4524%, Validation Loss: 1.50
52, Validation Accuracy: 95.9286%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 418.61it/s]
Epoch [15/20], Training Loss: 1.5108, Training Accuracy: 96.8065%, Validation Loss: 1.50
43, Validation Accuracy: 96.0476%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.78it/s]
Epoch [16/20], Training Loss: 1.5094, Training Accuracy: 96.7173%, Validation Loss: 1.50
45, Validation Accuracy: 96.0595%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.16it/s]
Epoch [17/20], Training Loss: 1.5067, Training Accuracy: 96.8214%, Validation Loss: 1.50
23, Validation Accuracy: 96.1667%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.04it/s]
Epoch [18/20], Training Loss: 1.5061, Training Accuracy: 96.9137%, Validation Loss: 1.50
05, Validation Accuracy: 96.4167%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.97it/s]
Epoch [19/20], Training Loss: 1.5031, Training Accuracy: 97.2143%, Validation Loss: 1.49
98, Validation Accuracy: 96.2262%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 453.22it/s]
Epoch [20/20], Training Loss: 1.5022, Training Accuracy: 97.1310%, Validation Loss: 1.50
18, Validation Accuracy: 96.0833%
```



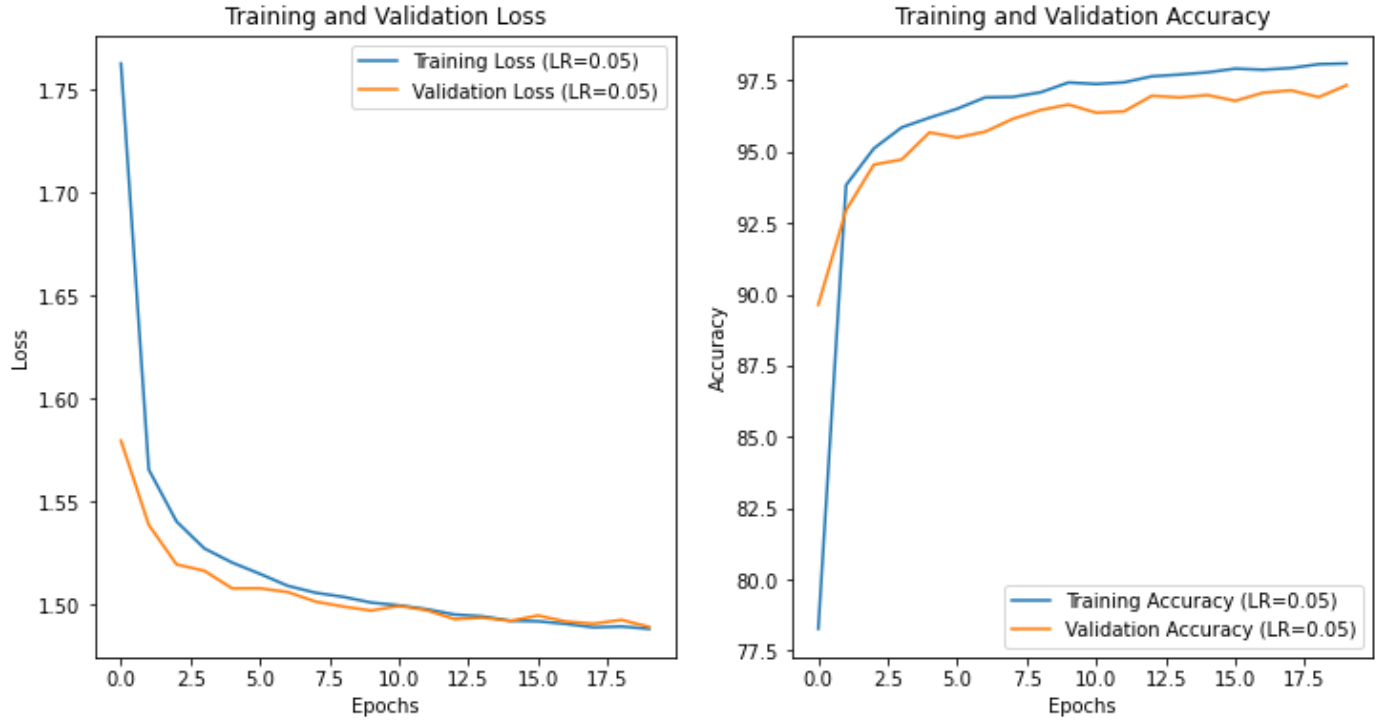

```
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 439.04it/s]
Epoch [1/20], Training Loss: 1.8334, Training Accuracy: 71.8988%, Validation Loss: 1.629
5, Validation Accuracy: 87.0357%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.81it/s]
Epoch [2/20], Training Loss: 1.5880, Training Accuracy: 92.8869%, Validation Loss: 1.540
0, Validation Accuracy: 93.0119%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.67it/s]
Epoch [3/20], Training Loss: 1.5508, Training Accuracy: 94.5238%, Validation Loss: 1.525
4, Validation Accuracy: 94.2262%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.97it/s]
Epoch [4/20], Training Loss: 1.5372, Training Accuracy: 95.1905%, Validation Loss: 1.520
4, Validation Accuracy: 94.5238%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.98it/s]
Epoch [5/20], Training Loss: 1.5283, Training Accuracy: 95.9196%, Validation Loss: 1.515
1, Validation Accuracy: 95.0476%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.84it/s]
Epoch [6/20], Training Loss: 1.5199, Training Accuracy: 95.9970%, Validation Loss: 1.507
9, Validation Accuracy: 95.5833%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.02it/s]
Epoch [7/20], Training Loss: 1.5166, Training Accuracy: 96.5298%, Validation Loss: 1.506
0, Validation Accuracy: 95.7857%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.91it/s]
Epoch [8/20], Training Loss: 1.5112, Training Accuracy: 96.6845%, Validation Loss: 1.504
1, Validation Accuracy: 95.8810%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.36it/s]
Epoch [9/20], Training Loss: 1.5079, Training Accuracy: 96.8452%, Validation Loss: 1.503
4, Validation Accuracy: 95.9286%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.91it/s]
```

```
Epoch [10/20], Training Loss: 1.5055, Training Accuracy: 97.0327%, Validation Loss: 1.50
14, Validation Accuracy: 96.2143%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.87it/s]
Epoch [11/20], Training Loss: 1.5035, Training Accuracy: 97.1220%, Validation Loss: 1.49
78, Validation Accuracy: 96.5714%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.29it/s]
Epoch [12/20], Training Loss: 1.5018, Training Accuracy: 97.3006%, Validation Loss: 1.49
72, Validation Accuracy: 96.5714%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.47it/s]
Epoch [13/20], Training Loss: 1.4996, Training Accuracy: 97.5298%, Validation Loss: 1.49
69, Validation Accuracy: 96.6190%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.79it/s]
Epoch [14/20], Training Loss: 1.4980, Training Accuracy: 97.5298%, Validation Loss: 1.49
75, Validation Accuracy: 96.4524%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.83it/s]
Epoch [15/20], Training Loss: 1.4963, Training Accuracy: 97.6190%, Validation Loss: 1.49
65, Validation Accuracy: 96.5595%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 429.51it/s]
Epoch [16/20], Training Loss: 1.4952, Training Accuracy: 97.7083%, Validation Loss: 1.49
81, Validation Accuracy: 96.4286%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 373.46it/s]
Epoch [17/20], Training Loss: 1.4941, Training Accuracy: 97.7738%, Validation Loss: 1.49
41, Validation Accuracy: 96.7619%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.29it/s]
Epoch [18/20], Training Loss: 1.4930, Training Accuracy: 97.8512%, Validation Loss: 1.49
58, Validation Accuracy: 96.7500%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.08it/s]
Epoch [19/20], Training Loss: 1.4919, Training Accuracy: 97.8274%, Validation Loss: 1.49
37, Validation Accuracy: 96.9524%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.91it/s]
Epoch [20/20], Training Loss: 1.4911, Training Accuracy: 97.9673%, Validation Loss: 1.49
30, Validation Accuracy: 96.9643%
```



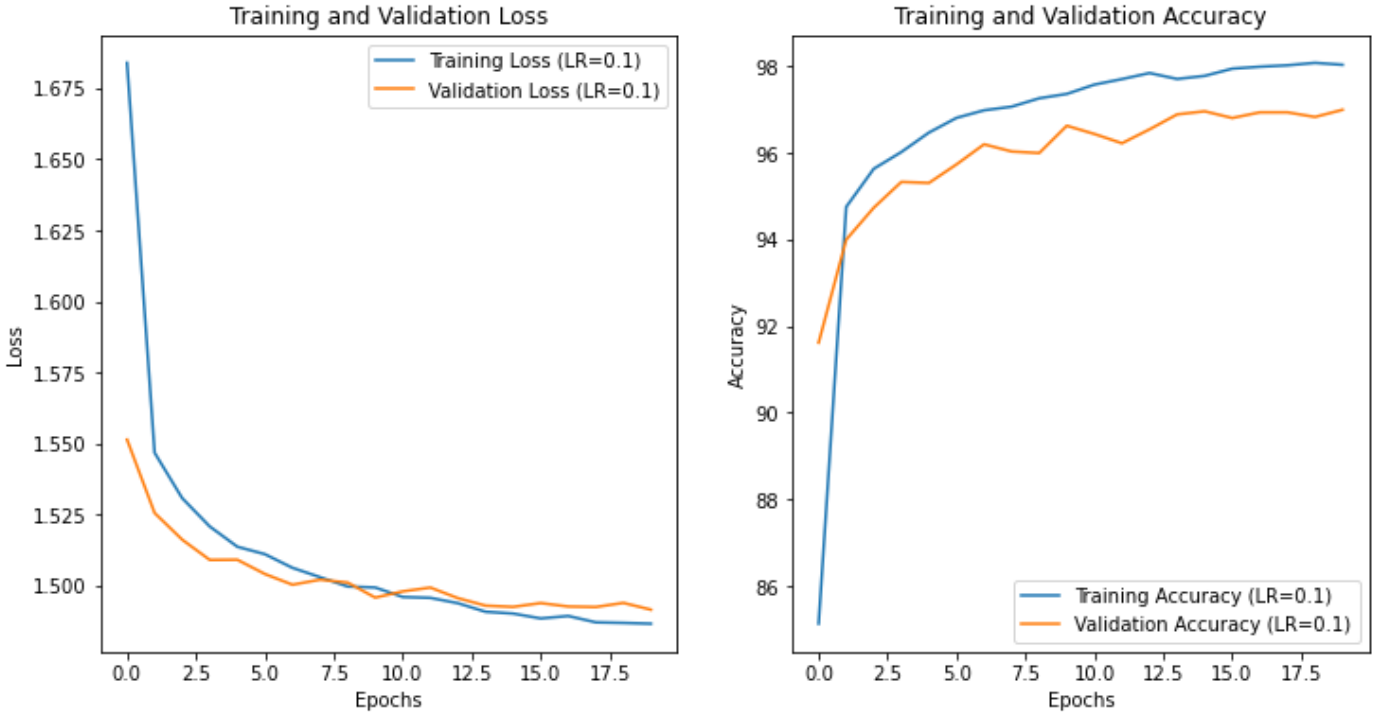
```
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 440.80it/s]
Epoch [1/20], Training Loss: 1.7625, Training Accuracy: 78.2619%, Validation Loss: 1.579
5, Validation Accuracy: 89.6310%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.36it/s]
Epoch [2/20], Training Loss: 1.5655, Training Accuracy: 93.8304%, Validation Loss: 1.538
6, Validation Accuracy: 92.9524%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 451.00it/s]
Epoch [3/20], Training Loss: 1.5402, Training Accuracy: 95.1161%, Validation Loss: 1.519
4, Validation Accuracy: 94.5476%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 436.85it/s]
Epoch [4/20], Training Loss: 1.5272, Training Accuracy: 95.8542%, Validation Loss: 1.516
3, Validation Accuracy: 94.7262%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.86it/s]
Epoch [5/20], Training Loss: 1.5203, Training Accuracy: 96.1935%, Validation Loss: 1.507
8, Validation Accuracy: 95.6786%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 450.69it/s]
Epoch [6/20], Training Loss: 1.5148, Training Accuracy: 96.5089%, Validation Loss: 1.507
8, Validation Accuracy: 95.5000%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 452.54it/s]
Epoch [7/20], Training Loss: 1.5090, Training Accuracy: 96.9048%, Validation Loss: 1.506
0, Validation Accuracy: 95.7024%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.43it/s]
Epoch [8/20], Training Loss: 1.5057, Training Accuracy: 96.9226%, Validation Loss: 1.501
4, Validation Accuracy: 96.1548%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 439.16it/s]
Epoch [9/20], Training Loss: 1.5037, Training Accuracy: 97.0863%, Validation Loss: 1.498
9, Validation Accuracy: 96.4643%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.34it/s]
```

```
Epoch [10/20], Training Loss: 1.5009, Training Accuracy: 97.4286%, Validation Loss: 1.49
70, Validation Accuracy: 96.6548%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.41it/s]
Epoch [11/20], Training Loss: 1.4996, Training Accuracy: 97.3780%, Validation Loss: 1.49
93, Validation Accuracy: 96.3690%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.85it/s]
Epoch [12/20], Training Loss: 1.4977, Training Accuracy: 97.4345%, Validation Loss: 1.49
71, Validation Accuracy: 96.4167%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.22it/s]
Epoch [13/20], Training Loss: 1.4951, Training Accuracy: 97.6458%, Validation Loss: 1.49
29, Validation Accuracy: 96.9643%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.53it/s]
Epoch [14/20], Training Loss: 1.4942, Training Accuracy: 97.7083%, Validation Loss: 1.49
36, Validation Accuracy: 96.9048%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.41it/s]
Epoch [15/20], Training Loss: 1.4921, Training Accuracy: 97.7857%, Validation Loss: 1.49
20, Validation Accuracy: 96.9881%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 411.26it/s]
Epoch [16/20], Training Loss: 1.4918, Training Accuracy: 97.9167%, Validation Loss: 1.49
46, Validation Accuracy: 96.7857%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 422.38it/s]
Epoch [17/20], Training Loss: 1.4906, Training Accuracy: 97.8780%, Validation Loss: 1.49
17, Validation Accuracy: 97.0714%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 450.69it/s]
Epoch [18/20], Training Loss: 1.4889, Training Accuracy: 97.9375%, Validation Loss: 1.49
07, Validation Accuracy: 97.1548%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.69it/s]
Epoch [19/20], Training Loss: 1.4893, Training Accuracy: 98.0714%, Validation Loss: 1.49
25, Validation Accuracy: 96.9167%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.60it/s]
Epoch [20/20], Training Loss: 1.4881, Training Accuracy: 98.0982%, Validation Loss: 1.48
91, Validation Accuracy: 97.3333%
```



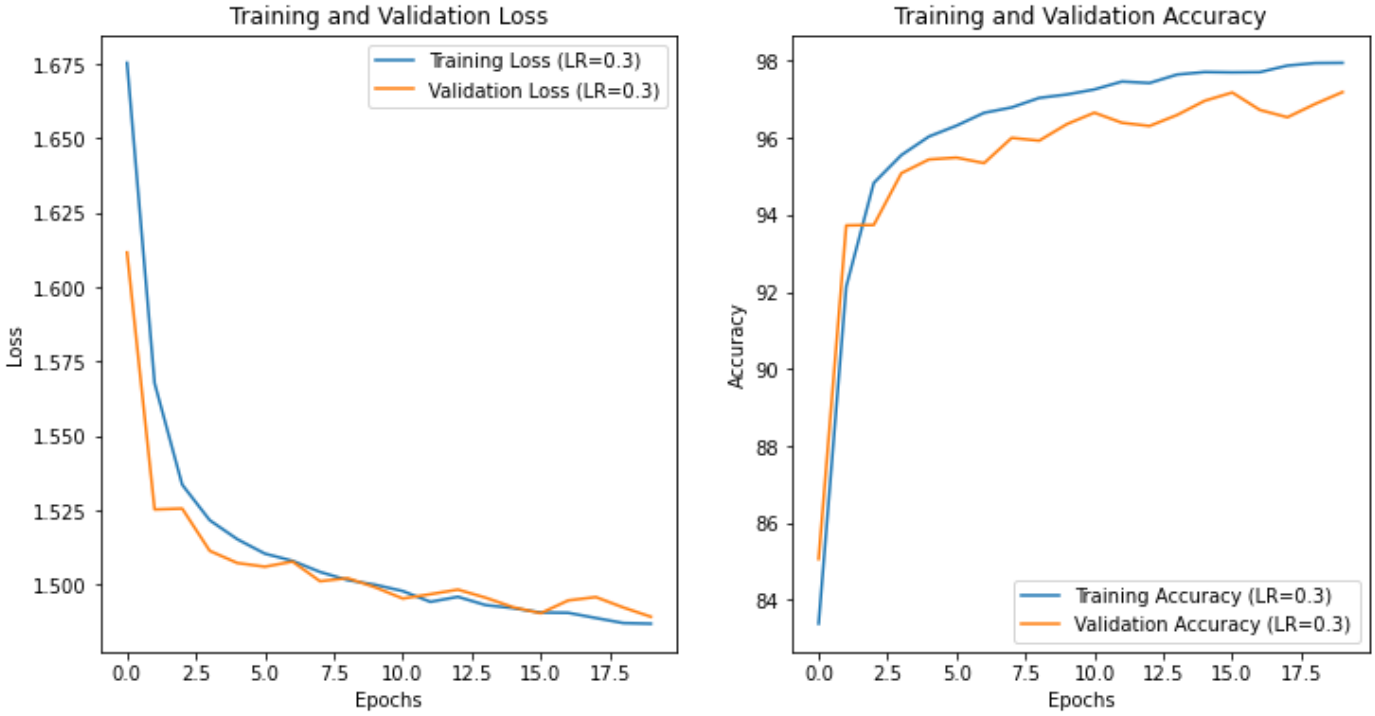
```
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.34it/s]
Epoch [1/20], Training Loss: 1.6839, Training Accuracy: 85.1280%, Validation Loss: 1.551
4, Validation Accuracy: 91.6190%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 386.74it/s]
Epoch [2/20], Training Loss: 1.5468, Training Accuracy: 94.7560%, Validation Loss: 1.525
6, Validation Accuracy: 94.0000%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 392.55it/s]
Epoch [3/20], Training Loss: 1.5308, Training Accuracy: 95.6369%, Validation Loss: 1.516
2, Validation Accuracy: 94.7381%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.10it/s]
Epoch [4/20], Training Loss: 1.5208, Training Accuracy: 96.0238%, Validation Loss: 1.509
1, Validation Accuracy: 95.3333%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.73it/s]
Epoch [5/20], Training Loss: 1.5137, Training Accuracy: 96.4732%, Validation Loss: 1.509
1, Validation Accuracy: 95.3095%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 402.96it/s]
Epoch [6/20], Training Loss: 1.5111, Training Accuracy: 96.8125%, Validation Loss: 1.504
1, Validation Accuracy: 95.7381%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 439.22it/s]
Epoch [7/20], Training Loss: 1.5063, Training Accuracy: 96.9881%, Validation Loss: 1.500
3, Validation Accuracy: 96.2024%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 388.10it/s]
Epoch [8/20], Training Loss: 1.5030, Training Accuracy: 97.0714%, Validation Loss: 1.502
1, Validation Accuracy: 96.0357%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.80it/s]
Epoch [9/20], Training Loss: 1.4997, Training Accuracy: 97.2649%, Validation Loss: 1.501
1, Validation Accuracy: 96.0000%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.86it/s]
```

```
Epoch [10/20], Training Loss: 1.4993, Training Accuracy: 97.3661%, Validation Loss: 1.49
58, Validation Accuracy: 96.6310%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 440.59it/s]
Epoch [11/20], Training Loss: 1.4960, Training Accuracy: 97.5804%, Validation Loss: 1.49
80, Validation Accuracy: 96.4405%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.42it/s]
Epoch [12/20], Training Loss: 1.4957, Training Accuracy: 97.7083%, Validation Loss: 1.49
93, Validation Accuracy: 96.2262%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.67it/s]
Epoch [13/20], Training Loss: 1.4938, Training Accuracy: 97.8482%, Validation Loss: 1.49
56, Validation Accuracy: 96.5476%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.73it/s]
Epoch [14/20], Training Loss: 1.4908, Training Accuracy: 97.7113%, Validation Loss: 1.49
30, Validation Accuracy: 96.8929%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.80it/s]
Epoch [15/20], Training Loss: 1.4902, Training Accuracy: 97.7827%, Validation Loss: 1.49
25, Validation Accuracy: 96.9643%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.81it/s]
Epoch [16/20], Training Loss: 1.4885, Training Accuracy: 97.9494%, Validation Loss: 1.49
39, Validation Accuracy: 96.8095%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.05it/s]
Epoch [17/20], Training Loss: 1.4893, Training Accuracy: 97.9940%, Validation Loss: 1.49
27, Validation Accuracy: 96.9405%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.05it/s]
Epoch [18/20], Training Loss: 1.4871, Training Accuracy: 98.0268%, Validation Loss: 1.49
25, Validation Accuracy: 96.9405%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.52it/s]
Epoch [19/20], Training Loss: 1.4869, Training Accuracy: 98.0833%, Validation Loss: 1.49
39, Validation Accuracy: 96.8333%
Training: 100%|██████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 452.72it/s]
Epoch [20/20], Training Loss: 1.4866, Training Accuracy: 98.0387%, Validation Loss: 1.49
16, Validation Accuracy: 97.0000%
```



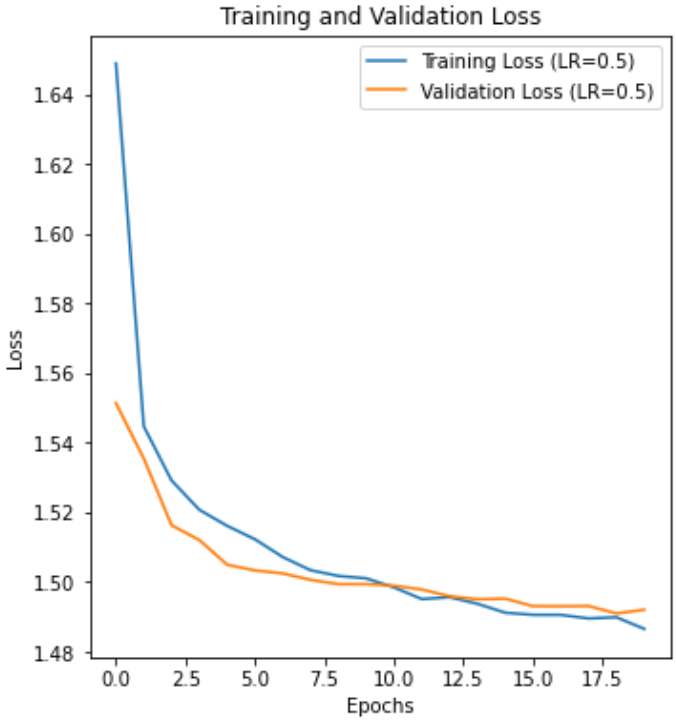

```
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.96it/s]
Epoch [1/20], Training Loss: 1.6753, Training Accuracy: 83.3810%, Validation Loss: 1.611
6, Validation Accuracy: 85.0595%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 408.20it/s]
Epoch [2/20], Training Loss: 1.5677, Training Accuracy: 92.1339%, Validation Loss: 1.525
3, Validation Accuracy: 93.7262%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.99it/s]
Epoch [3/20], Training Loss: 1.5336, Training Accuracy: 94.8304%, Validation Loss: 1.525
6, Validation Accuracy: 93.7381%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 423.92it/s]
Epoch [4/20], Training Loss: 1.5216, Training Accuracy: 95.5536%, Validation Loss: 1.511
3, Validation Accuracy: 95.0833%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 386.21it/s]
Epoch [5/20], Training Loss: 1.5153, Training Accuracy: 96.0357%, Validation Loss: 1.507
3, Validation Accuracy: 95.4405%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 364.04it/s]
Epoch [6/20], Training Loss: 1.5104, Training Accuracy: 96.3185%, Validation Loss: 1.506
0, Validation Accuracy: 95.4881%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 389.35it/s]
Epoch [7/20], Training Loss: 1.5080, Training Accuracy: 96.6518%, Validation Loss: 1.507
7, Validation Accuracy: 95.3452%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 374.27it/s]
Epoch [8/20], Training Loss: 1.5042, Training Accuracy: 96.7917%, Validation Loss: 1.501
1, Validation Accuracy: 96.0000%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 372.54it/s]
Epoch [9/20], Training Loss: 1.5015, Training Accuracy: 97.0387%, Validation Loss: 1.502
2, Validation Accuracy: 95.9286%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 398.64it/s]
```

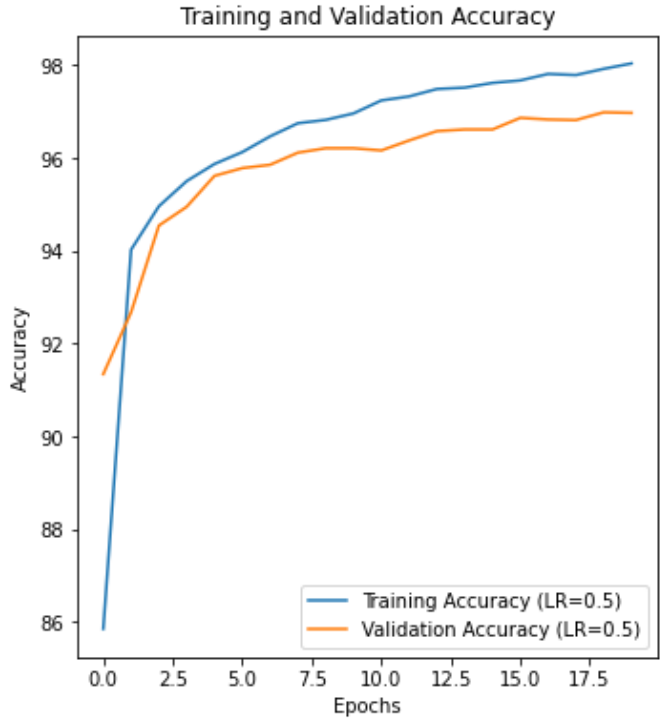
```
Epoch [10/20], Training Loss: 1.5000, Training Accuracy: 97.1280%, Validation Loss: 1.49
91, Validation Accuracy: 96.3571%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 398.49it/s]
Epoch [11/20], Training Loss: 1.4978, Training Accuracy: 97.2589%, Validation Loss: 1.49
53, Validation Accuracy: 96.6548%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 369.02it/s]
Epoch [12/20], Training Loss: 1.4942, Training Accuracy: 97.4643%, Validation Loss: 1.49
68, Validation Accuracy: 96.3929%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 368.12it/s]
Epoch [13/20], Training Loss: 1.4959, Training Accuracy: 97.4256%, Validation Loss: 1.49
83, Validation Accuracy: 96.3095%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 406.80it/s]
Epoch [14/20], Training Loss: 1.4931, Training Accuracy: 97.6458%, Validation Loss: 1.49
56, Validation Accuracy: 96.5952%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 401.53it/s]
Epoch [15/20], Training Loss: 1.4921, Training Accuracy: 97.7113%, Validation Loss: 1.49
24, Validation Accuracy: 96.9643%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 400.60it/s]
Epoch [16/20], Training Loss: 1.4906, Training Accuracy: 97.7024%, Validation Loss: 1.49
04, Validation Accuracy: 97.1786%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 378.30it/s]
Epoch [17/20], Training Loss: 1.4905, Training Accuracy: 97.7083%, Validation Loss: 1.49
47, Validation Accuracy: 96.7262%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 437.58it/s]
Epoch [18/20], Training Loss: 1.4888, Training Accuracy: 97.8780%, Validation Loss: 1.49
58, Validation Accuracy: 96.5357%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 436.49it/s]
Epoch [19/20], Training Loss: 1.4871, Training Accuracy: 97.9435%, Validation Loss: 1.49
23, Validation Accuracy: 96.8810%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 437.76it/s]
Epoch [20/20], Training Loss: 1.4869, Training Accuracy: 97.9494%, Validation Loss: 1.48
92, Validation Accuracy: 97.1905%
```



```
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 370.45it/s]
Epoch [1/20], Training Loss: 1.6488, Training Accuracy: 85.8542%, Validation Loss: 1.551
4, Validation Accuracy: 91.3452%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 374.53it/s]
Epoch [2/20], Training Loss: 1.5447, Training Accuracy: 94.0268%, Validation Loss: 1.535
5, Validation Accuracy: 92.6786%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:03<00:00, 347.51it/s]
Epoch [3/20], Training Loss: 1.5292, Training Accuracy: 94.9643%, Validation Loss: 1.516
3, Validation Accuracy: 94.5476%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 362.41it/s]
Epoch [4/20], Training Loss: 1.5208, Training Accuracy: 95.5060%, Validation Loss: 1.512
1, Validation Accuracy: 94.9524%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 408.22it/s]
Epoch [5/20], Training Loss: 1.5161, Training Accuracy: 95.8750%, Validation Loss: 1.505
0, Validation Accuracy: 95.6190%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:03<00:00, 324.69it/s]
Epoch [6/20], Training Loss: 1.5123, Training Accuracy: 96.1310%, Validation Loss: 1.503
4, Validation Accuracy: 95.7857%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:03<00:00, 344.90it/s]
Epoch [7/20], Training Loss: 1.5072, Training Accuracy: 96.4702%, Validation Loss: 1.502
5, Validation Accuracy: 95.8571%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:03<00:00, 342.04it/s]
Epoch [8/20], Training Loss: 1.5034, Training Accuracy: 96.7530%, Validation Loss: 1.500
7, Validation Accuracy: 96.1190%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 365.69it/s]
Epoch [9/20], Training Loss: 1.5018, Training Accuracy: 96.8244%, Validation Loss: 1.499
4, Validation Accuracy: 96.2143%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 365.94it/s]
```

```
Epoch [10/20], Training Loss: 1.5011, Training Accuracy: 96.9643%, Validation Loss: 1.49
95, Validation Accuracy: 96.2143%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 362.79it/s]
Epoch [11/20], Training Loss: 1.4985, Training Accuracy: 97.2440%, Validation Loss: 1.49
90, Validation Accuracy: 96.1667%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 404.10it/s]
Epoch [12/20], Training Loss: 1.4951, Training Accuracy: 97.3304%, Validation Loss: 1.49
79, Validation Accuracy: 96.3810%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 397.93it/s]
Epoch [13/20], Training Loss: 1.4957, Training Accuracy: 97.4911%, Validation Loss: 1.49
60, Validation Accuracy: 96.5833%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 416.13it/s]
Epoch [14/20], Training Loss: 1.4938, Training Accuracy: 97.5208%, Validation Loss: 1.49
51, Validation Accuracy: 96.6190%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 378.16it/s]
Epoch [15/20], Training Loss: 1.4912, Training Accuracy: 97.6220%, Validation Loss: 1.49
53, Validation Accuracy: 96.6190%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 388.00it/s]
Epoch [16/20], Training Loss: 1.4906, Training Accuracy: 97.6756%, Validation Loss: 1.49
31, Validation Accuracy: 96.8690%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 404.62it/s]
Epoch [17/20], Training Loss: 1.4906, Training Accuracy: 97.8155%, Validation Loss: 1.49
31, Validation Accuracy: 96.8333%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 395.50it/s]
Epoch [18/20], Training Loss: 1.4896, Training Accuracy: 97.7917%, Validation Loss: 1.49
31, Validation Accuracy: 96.8214%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 411.41it/s]
Epoch [19/20], Training Loss: 1.4899, Training Accuracy: 97.9256%, Validation Loss: 1.49
10, Validation Accuracy: 96.9881%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 410.61it/s]
Epoch [20/20], Training Loss: 1.4866, Training Accuracy: 98.0417%, Validation Loss: 1.49
21, Validation Accuracy: 96.9762%
```





Show best model according to Learning rate

In [31]:
```
# Print the best hyperparameters
print("Best Hyperparameter:")
print(best_hyperparameters)

print("-----------------------")

# Print the best model
print("Best Model:")
print(best_model)
```

```
Best Hyperparameter:
{'lr': 0.05}
-----------------------
Best Model:
OrderedDict([('fc1.weight', tensor([[-0.0183,  0.0130, -0.0306,  ..., -0.0177, -0.0006,
0.0177],
        [ 0.0260,  0.0151, -0.0085,  ...,  0.0034,  0.0052, -0.0088],
        [-0.0296,  0.0152,  0.0012,  ...,  0.0128, -0.0221,  0.0280],
        ...,
        [-0.0132,  0.0269,  0.0037,  ...,  0.0250,  0.0064,  0.0296],
        [-0.0047, -0.0240,  0.0139,  ..., -0.0164,  0.0055,  0.0268],
        [-0.0238,  0.0274, -0.0251,  ..., -0.0168,  0.0109, -0.0216]])), ('fc1.bias', te
nsor([-0.0033,  0.0108, -0.0166,  0.0014,  0.0177, -0.0183, -0.0006,  0.0279,
         0.0169,  0.0174, -0.0016, -0.0148,  0.0198,  0.0310, -0.0136, -0.0140,
        -0.0228, -0.0154, -0.0253, -0.0300, -0.0277, -0.0180, -0.0252,  0.0127,
        -0.0176,  0.0176,  0.0004,  0.0162, -0.0239, -0.0159, -0.0212,  0.0305,
        -0.0008, -0.0011,  0.0070,  0.0245,  0.0036,  0.0336,  0.0351,  0.0312,
         0.0265, -0.0150,  0.0325, -0.0320,  0.0009,  0.0098, -0.0190, -0.0202,
         0.0028,  0.0108, -0.0306, -0.0005,  0.0107,  0.0169,  0.0336,  0.0345,
        -0.0080, -0.0290,  0.0150, -0.0084, -0.0185,  0.0086,  0.0062, -0.0236,
        -0.0192, -0.0231,  0.0022, -0.0106, -0.0253, -0.0114, -0.0050, -0.0088,
        -0.0324, -0.0128,  0.0067, -0.0138, -0.0094,  0.0142,  0.0014,  0.0135,
         0.0321, -0.0283,  0.0009,  0.0088,  0.0320, -0.0258, -0.0006, -0.0053,
        -0.0041, -0.0162, -0.0059,  0.0025,  0.0048,  0.0274,  0.0294,  0.0211,
         0.0160, -0.0346, -0.0326, -0.0212, -0.0183,  0.0036,  0.0125,  0.0339,
         0.0053,  0.0220,  0.0017, -0.0013,  0.0259, -0.0019, -0.0128,  0.0293,
         0.0171,  0.0178,  0.0069,  0.0093,  0.0157, -0.0148,  0.0109,  0.0297,
         0.0100,  0.0027,  0.0102, -0.0240,  0.0246,  0.0321, -0.0235,  0.0289])), ('fc
2.weight', tensor([[-0.0056, -0.0508,  0.0126,  ...,  0.0006, -0.0395, -0.0308],
        [ 0.0783, -0.0222, -0.0130,  ...,  0.0386,  0.0424,  0.0108],
        [ 0.0828,  0.1119,  0.0849,  ...,  0.0591,  0.0724, -0.0500],
        ...,
        [-0.1124, -0.0265, -0.0538,  ...,  0.0083,  0.0446,  0.0302],
        [-0.0116,  0.0129,  0.0031,  ..., -0.0344, -0.0740,  0.0558],
        [-0.0416, -0.0415, -0.1014,  ..., -0.0200,  0.0461,  0.0441]])), ('fc2.bias', te
nsor([-0.0465, -0.0223, -0.0186, -0.0011, -0.0255, -0.0809,  0.0331,  0.0278,
         0.0877, -0.0022,  0.0032, -0.1343,  0.0471, -0.0687, -0.0808,  0.0792,
        -0.0433, -0.0455, -0.0398, -0.0218,  0.0161, -0.0064,  0.0480, -0.0154,
        -0.0430, -0.0148,  0.0294,  0.0268,  0.0531,  0.0448, -0.0427, -0.0346,
        -0.0089, -0.0724,  0.0296, -0.0671, -0.0283,  0.0092,  0.0166, -0.0413,
        -0.0834,  0.0354, -0.0140, -0.0603, -0.0525,  0.0573,  0.0422, -0.0421,
         0.0295, -0.0036,  0.0604,  0.0381,  0.0812,  0.0135,  0.0374,  0.0665,
        -0.0354, -0.0701, -0.0469,  0.0477, -0.0244, -0.0249,  0.0157, -0.0768])), ('ou
t.weight', tensor([[ 1.6741e-01, -3.3687e-02,  5.3203e-01, -2.2305e-01,  2.0547e-01,
          1.3052e-01,  3.7336e-02, -2.0561e-01,  6.2072e-01,  4.6805e-02,
          5.9115e-02, -1.9892e-01, -2.0446e-01, -1.5212e-02, -1.4096e-01,
         -1.5331e-01, -1.3690e-01,  8.5651e-03, -1.0630e-01, -1.6841e-01,
         -2.0014e-01, -2.1580e-01,  1.1171e-02, -1.6045e-01, -2.8240e-02,
         -1.5983e-01, -2.0147e-02,  5.0889e-01, -1.8978e-01, -2.6665e-01,
         -2.7600e-01, -2.8285e-01, -1.4226e-01, -3.3910e-02, -3.0180e-01,
         -2.0099e-01,  1.5180e-03,  7.1429e-01, -1.2443e-01,  5.4439e-01,
         -1.8571e-01, -1.2389e-01, -8.4450e-02, -2.1084e-01,  4.5489e-01,
         -1.4298e-01, -4.4524e-02,  3.2120e-01,  1.9875e-03,  6.2675e-01,
         -2.0198e-01,  6.0808e-01,  2.8775e-01, -6.6497e-02, -2.2161e-01,
         -2.3375e-01, -9.8044e-03, -1.9381e-01, -1.2591e-01, -1.6340e-01,
         -1.3106e-01, -2.0854e-01, -2.6904e-01,  7.9987e-02],
        [ 5.3435e-02, -7.4077e-02, -2.9730e-01, -2.6292e-01, -2.5049e-01,
         -2.8502e-02,  5.0817e-01, -2.4522e-01,  4.1598e-01, -2.1779e-01,
         -1.6812e-01, -3.5708e-01, -1.6410e-01,  7.9581e-02, -2.2834e-01,
         -1.9076e-01, -1.0414e-01, -2.3085e-01, -2.6493e-01, -6.1501e-02,
         -2.3592e-01,  5.2446e-01, -2.4033e-01,  5.3684e-02, -3.8542e-02,
         -1.1234e-01, -3.3892e-02, -3.1846e-01,  4.4900e-01,  3.6429e-01,
         -2.9729e-01, -1.5306e-01, -1.4417e-01, -1.1167e-01, -3.1667e-01,
         -1.7606e-01,  5.4064e-01, -8.1544e-02,  6.1665e-01, -1.5755e-01,
```

```
 -2.0886e-01, -1.4138e-01,  1.6004e-01, -1.0178e-01, -1.9846e-01,
  3.8045e-01, -3.4942e-03, -2.2564e-01,  8.2375e-02, -1.7267e-01,
 -1.7805e-01,  1.3573e-03,  2.1089e-01,  8.6629e-02, -1.7327e-01,
  3.4080e-01, -1.0845e-01, -1.3051e-01, -9.6103e-02,  7.0983e-01,
  2.4651e-01,  5.8643e-01, -1.5766e-01,  1.5145e-01],
[-1.1355e-01, -3.8292e-02, -1.6748e-01, -1.6114e-01, -1.2828e-01,
  8.4156e-02,  5.2322e-01, -2.1072e-01, -1.8652e-01,  6.6115e-02,
  2.2690e-03,  5.0460e-01, -1.3290e-01, -2.3328e-01, -1.1978e-01,
  2.0593e-01, -2.1655e-01,  2.8285e-01, -1.0454e-01,  2.7676e-01,
 -8.9812e-02, -2.1141e-01,  8.3059e-01, -3.1638e-01, -2.1100e-01,
 -7.0678e-02, -7.4958e-02, -1.2666e-01, -3.2723e-01, -2.4451e-01,
  4.5884e-01, -9.6885e-02, -7.7725e-02,  6.1827e-01, -2.2411e-01,
  5.5573e-01, -8.9722e-02, -6.6301e-02, -2.1061e-01, -1.9838e-01,
 -1.5981e-01, -1.7913e-01, -9.7909e-02,  3.2219e-01, -1.1353e-01,
  5.6246e-01, -2.6992e-01, -2.1034e-01,  5.7124e-01,  1.5204e-01,
 -3.2988e-01, -4.2946e-02,  2.9681e-02,  4.8509e-01,  2.4309e-01,
 -2.7705e-01,  4.3392e-02, -2.1864e-01, -5.0463e-03, -5.0605e-02,
  3.7725e-02,  2.4598e-02, -1.6629e-01,  1.1505e-01],
[ 2.7621e-02, -9.4849e-03, -1.9100e-01,  5.3471e-01, -2.7462e-01,
 -1.2869e-01, -3.0836e-01, -1.0965e-01, -2.6379e-01,  1.8289e-02,
 -1.2284e-01, -8.0573e-02, -7.5221e-02, -2.2809e-01, -2.2387e-01,
 -2.7718e-01,  5.3835e-01,  4.4287e-01, -1.4033e-01,  4.2038e-01,
  4.8192e-01, -2.1744e-01, -1.2358e-01,  4.7180e-01, -1.9051e-01,
 -1.9792e-01, -9.1147e-02,  4.6815e-02, -3.4370e-01,  2.6027e-01,
 -2.6177e-01,  4.3646e-02, -7.4946e-04, -2.5612e-01, -2.5949e-01,
  5.9694e-01, -1.6151e-01, -1.9397e-01, -1.0729e-01,  4.2228e-01,
  7.0497e-02, -6.7933e-02, -7.3254e-02, -1.7511e-01, -2.2457e-01,
 -4.5784e-02,  5.8771e-01, -2.2572e-01,  5.7559e-01, -1.5365e-01,
  3.8066e-01, -2.9389e-01, -1.7693e-01,  1.2913e-01, -2.7783e-01,
 -3.7110e-01, -4.0641e-02,  7.4525e-01, -1.8735e-01, -4.4691e-02,
  8.7044e-02, -1.5850e-01, -2.6500e-01, -1.1282e-01],
[-1.1960e-01,  5.5760e-01,  5.0903e-01, -3.5603e-01,  4.5789e-01,
  2.7365e-02, -9.3775e-02,  3.4624e-01, -1.7884e-01,  6.5455e-01,
 -1.4321e-02, -2.3499e-01,  6.3356e-01, -1.1554e-01, -2.4097e-01,
 -2.6339e-02, -1.7456e-01, -1.6763e-01,  3.5692e-01, -2.6570e-01,
 -2.4321e-01, -1.1056e-01, -1.2021e-01, -2.5339e-01,  5.3806e-02,
  5.6678e-01,  6.1841e-01, -2.5977e-01,  7.8768e-02, -2.3215e-01,
 -2.7792e-01, -1.3885e-01, -7.8873e-02, -2.0646e-01,  4.5490e-01,
 -9.7420e-02, -5.6249e-02, -7.4550e-02, -1.4422e-01, -3.1844e-01,
 -7.0677e-02, -5.1544e-02,  1.8386e-01,  2.4717e-02, -1.0011e-01,
 -2.9713e-02, -1.7657e-01, -3.1952e-01, -1.5582e-01, -1.3019e-01,
 -3.2601e-01, -2.7474e-01,  8.6203e-02, -9.6692e-02, -2.8376e-01,
  4.2943e-01, -8.1544e-03, -1.6719e-01,  5.3682e-01,  1.3654e-01,
  1.5104e-01,  5.5464e-04, -4.1333e-01, -1.1720e-01],
[ 4.8222e-02,  8.3960e-02, -3.4768e-01, -2.2027e-01, -1.1102e-01,
 -4.4827e-02, -3.0342e-01, -1.0937e-01, -1.0377e-01, -9.5381e-02,
  2.0668e-01, -2.9173e-01, -1.2242e-01,  7.5131e-01, -6.5985e-02,
 -1.4628e-01,  5.9886e-01, -1.7552e-01, -2.0283e-01, -2.1475e-01,
  3.1692e-02, -2.0684e-01, -1.5122e-01, -3.6292e-02, -2.4596e-01,
  1.2543e-01, -1.4824e-01, -2.8864e-01,  9.9498e-02,  4.7676e-01,
 -1.8042e-01,  5.0002e-01, -1.4147e-01, -1.0416e-01,  4.8216e-01,
 -2.0089e-01,  1.0975e-01, -2.4850e-01, -2.4702e-01,  2.7639e-01,
  5.3529e-01,  6.2240e-01, -1.2969e-01, -1.9592e-01, -1.3849e-01,
 -2.2906e-02, -6.3199e-02,  4.2246e-01, -2.2027e-01, -5.9851e-02,
 -3.8496e-05,  5.6389e-01, -2.1176e-01, -1.6829e-01,  4.0246e-01,
 -3.1232e-01, -9.2388e-02,  1.5549e-01, -2.9422e-03, -8.5124e-02,
  9.0910e-02, -2.1431e-01,  4.3678e-01,  4.5887e-02],
[-3.9461e-02, -1.1651e-01,  3.4760e-01, -2.3235e-01, -1.9142e-01,
 -1.4627e-01,  4.9046e-01, -3.1900e-02, -1.3294e-01, -3.5884e-02,
 -1.3543e-01, -2.5796e-02, -2.4549e-01,  9.5849e-02,  5.9899e-01,
  6.7081e-01, -1.4093e-01, -2.2210e-01,  5.8243e-01, -2.5731e-01,
  5.8737e-01,  5.6082e-01, -1.8575e-01, -2.2379e-01, -1.2060e-01,
 -1.2584e-01, -2.5180e-01, -9.4527e-02, -2.9623e-01, -1.9565e-01,
 -1.9887e-01, -1.3212e-01, -5.9170e-02,  2.3803e-01, -3.1068e-01,
```

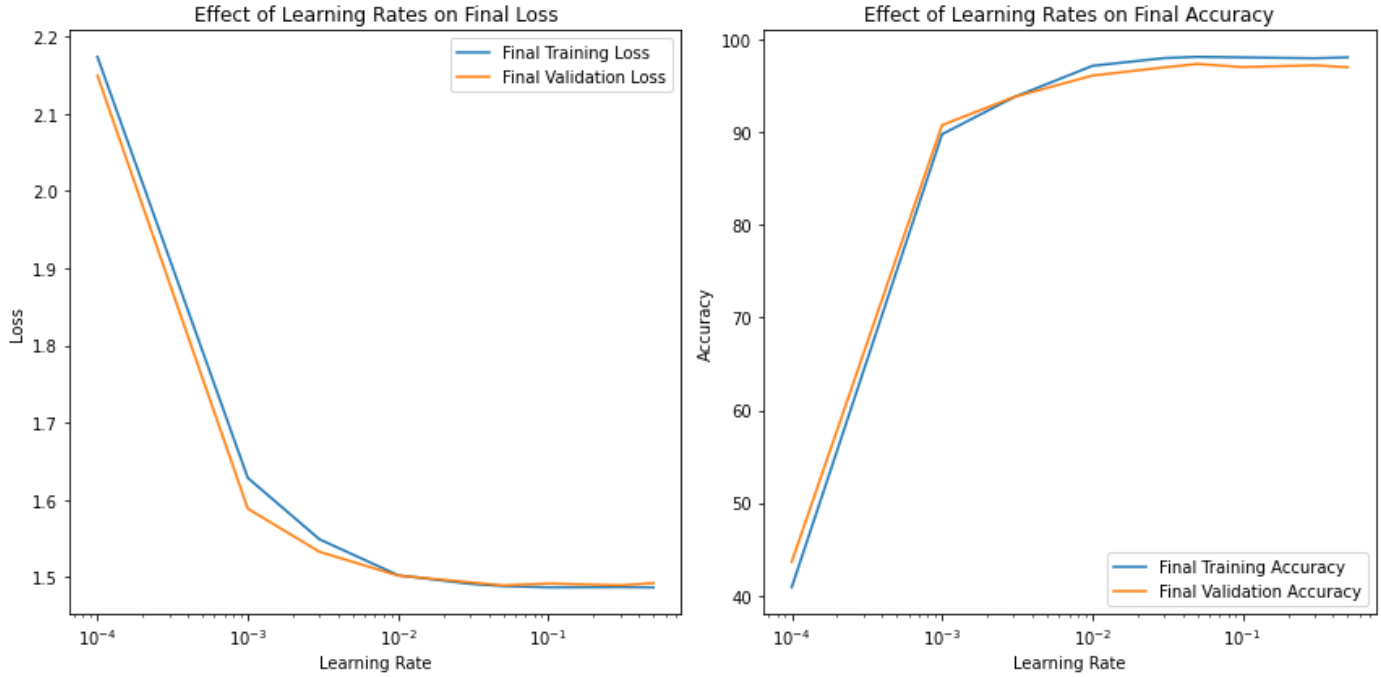
```
         -2.6424e-01, -1.0070e-01,  2.7500e-01, -1.7461e-01,  1.4941e-02,
         -1.0849e-01,  6.0956e-01, -3.1205e-02, -1.5632e-01, -1.7263e-01,
         -2.6627e-01, -8.2569e-02, -1.3350e-01, -2.6238e-01, -1.0036e-01,
         -1.4228e-01, -1.4042e-01,  5.7943e-02, -2.1137e-01, -9.1733e-02,
          5.1613e-01, -4.2016e-03,  2.6548e-01, -8.0144e-02, -1.5443e-01,
         -6.4330e-02, -8.2052e-02,  3.3648e-01, -2.1373e-01],
        [-6.8810e-03, -6.7355e-02, -2.0349e-01, -2.5920e-01, -6.7966e-02,
         -2.6448e-01, -2.1856e-01, -2.5354e-01, -1.0521e-01,  1.6335e-02,
         -1.3553e-01, -3.1520e-01, -1.0102e-01, -7.3933e-02, -1.0879e-03,
         -6.9544e-02, -2.3639e-01, -1.3579e-01, -1.3413e-01,  4.9564e-01,
         -1.3719e-01, -1.0713e-01, -1.0911e-01, -2.1085e-01,  5.9465e-01,
         -1.5989e-01, -2.4014e-01, -1.6475e-01, -2.3299e-01, -2.5729e-01,
          5.1140e-01, -1.8064e-01, -4.7928e-02, -1.9341e-01,  2.5393e-01,
         -1.7497e-01, -9.4135e-02,  1.0781e-02,  5.8430e-01, -7.7013e-02,
         -8.8949e-02, -1.7137e-01,  1.3798e-01,  6.0579e-01,  6.4608e-01,
         -2.0061e-01,  5.8115e-01, -2.3535e-01, -1.8603e-01, -1.0156e-01,
          5.1255e-01, -1.9125e-01, -1.0562e-01,  2.2298e-01,  5.2193e-01,
         -7.6300e-02,  6.0058e-02, -1.8174e-01, -8.0031e-02, -1.8394e-02,
          1.2465e-01, -1.0915e-01,  2.3677e-01, -1.3086e-01],
        [ 6.3967e-02, -1.6093e-01, -2.6488e-01,  3.0229e-01, -1.0825e-01,
         -9.5352e-03, -2.9515e-01, -3.4055e-01, -3.6807e-01, -1.5258e-01,
         -1.2469e-01,  4.8560e-01,  9.5177e-02, -4.8675e-02,  5.9432e-01,
         -2.4891e-01, -2.2655e-01, -5.6028e-02, -5.1616e-02,  4.7697e-01,
         -1.1842e-01, -2.0680e-01, -9.3466e-02,  5.3426e-01, -2.9112e-02,
          3.9967e-01, -1.1331e-01,  3.4200e-01,  4.6439e-01, -1.6805e-01,
         -1.7815e-01,  6.4937e-01, -6.2100e-02,  3.8000e-01, -4.8389e-01,
         -2.7193e-01, -1.6887e-01, -1.8357e-01, -1.4647e-01, -4.2465e-01,
          3.0352e-01, -1.2067e-01, -1.8427e-01, -9.7375e-02, -2.9220e-01,
         -2.9687e-01, -1.6895e-01,  4.7953e-01, -1.1351e-01, -1.1846e-01,
         -1.2191e-01, -4.1362e-02, -1.4692e-01, -2.5829e-01, -2.6590e-01,
          3.6686e-01,  2.9138e-02, -2.2405e-01, -2.5052e-01, -4.5402e-02,
          1.0354e-01,  4.7034e-01, -3.4422e-01,  5.3692e-01],
        [-4.3847e-02, -2.5867e-01, -1.5168e-01,  5.1063e-01,  4.1841e-01,
          6.7257e-01, -2.7309e-01,  6.4156e-01, -1.9547e-01, -2.0839e-01,
         -7.5183e-02,  3.9662e-01,  2.0305e-02, -1.8451e-01, -2.5019e-01,
          6.9968e-03,  6.6337e-03, -1.8527e-01, -2.5687e-01, -4.3358e-01,
         -2.1685e-01, -2.0303e-01, -1.9796e-01,  2.5332e-01,  3.9204e-01,
         -3.2339e-01,  6.1075e-01,  4.8700e-01,  4.1969e-01, -3.0390e-02,
          4.9265e-01, -8.1012e-02,  3.6255e-01, -3.2656e-01,  3.6499e-01,
         -2.0854e-01, -1.9182e-02,  2.3104e-02,  3.2932e-02, -2.8385e-01,
         -2.4705e-01, -1.7394e-01, -2.4428e-01, -3.3666e-01, -2.3103e-01,
         -4.3303e-02, -1.9722e-01, -2.2069e-01, -2.2628e-01, -1.6506e-01,
          5.0034e-01, -1.0725e-01, -2.4429e-02, -2.2896e-01, -3.0304e-01,
         -3.7559e-01, -8.3538e-02, -8.0662e-02,  2.3366e-01, -2.1741e-01,
          2.4271e-02, -2.5419e-01,  3.7163e-01, -2.6468e-01]])), ('out.bias', tensor([0.
0403, 0.0868, 0.0618, 0.0482, 0.1096, 0.1205, 0.0855, 0.0964, 0.0163,
        0.0067])), ('layer_norm1.weight', tensor([0.9744, 1.0071, 0.9599, 0.8581, 0.969
8, 0.9758, 1.1256, 0.9752, 0.9941,
        0.9163, 1.0724, 0.7754, 1.0311, 0.8985, 1.0313, 1.1314, 0.7389, 0.9061,
        1.0381, 0.9957, 0.8898, 1.0783, 0.8848, 0.9441, 1.0593, 0.9320, 1.0665,
        0.9594, 0.9395, 0.9871, 0.9838, 1.0371, 0.7374, 0.7454, 0.9392, 0.8200,
        0.8178, 1.0089, 1.0268, 0.9924, 0.9073, 0.9181, 0.9361, 1.0772, 0.7582,
        0.9043, 0.8629, 0.9442, 0.9551, 1.0298, 1.0654, 1.0234, 1.0535, 1.0716,
        0.8264, 1.0994, 1.0379, 0.7727, 1.1069, 0.8849, 0.9338, 0.9776, 0.9953,
        0.9861, 1.0374, 1.0968, 0.9778, 1.0700, 1.0355, 1.0029, 1.0150, 0.9907,
        1.0783, 0.9820, 0.9465, 0.9777, 1.1190, 1.0236, 0.9468, 0.9944, 0.9497,
        0.9183, 0.9938, 1.0606, 0.7290, 0.9213, 1.0582, 1.0758, 1.0887, 0.7911,
        0.7799, 0.8932, 0.9762, 1.0516, 0.9922, 1.0493, 1.0932, 1.0299, 1.1118,
        0.9790, 1.0524, 1.0291, 0.8574, 0.9306, 1.0800, 1.0793, 0.9581, 0.9607,
        1.0360, 1.0070, 1.0318, 0.9631, 0.9017, 0.9963, 0.9599, 1.0523, 0.9234,
        0.8906, 0.9884, 0.9937, 1.1268, 0.8896, 1.0068, 0.9663, 0.9447, 0.9364,
        1.0072, 0.8855])), ('layer_norm1.bias', tensor([-0.1449, -0.1248, -0.1600, -0.24
72, -0.1517, -0.0973, -0.0873, -0.1791,
        -0.1703, -0.1842, -0.1279, -0.1808, -0.1481, -0.1964, -0.1006, -0.0992,
```

```
        -0.2053, -0.1738, -0.1696, -0.1661, -0.2127, -0.1292, -0.1698, -0.2117,
        -0.1056, -0.1259, -0.1533, -0.1522, -0.2200, -0.1538, -0.1458, -0.1143,
        -0.2203, -0.1880, -0.1419, -0.2014, -0.2023, -0.1052, -0.1471, -0.1634,
        -0.1848, -0.2398, -0.1720, -0.1054, -0.2009, -0.2151, -0.2224, -0.1411,
        -0.1757, -0.1304, -0.1004, -0.1365, -0.1602, -0.1376, -0.1959, -0.1321,
        -0.1771, -0.2152, -0.1112, -0.1845, -0.1973, -0.1525, -0.1503, -0.1028,
        -0.1526, -0.1177, -0.1444, -0.1453, -0.1367, -0.1631, -0.1218, -0.2355,
        -0.1038, -0.1857, -0.1481, -0.1870, -0.0837, -0.1217, -0.1346, -0.1377,
        -0.1664, -0.1795, -0.1894, -0.0830, -0.2092, -0.1436, -0.1321, -0.1624,
        -0.1388, -0.1947, -0.1851, -0.2326, -0.1369, -0.1244, -0.1287, -0.0586,
        -0.1516, -0.1340, -0.1618, -0.1990, -0.1110, -0.1377, -0.2139, -0.2165,
        -0.1464, -0.1327, -0.1445, -0.0898, -0.1042, -0.1259, -0.1411, -0.1923,
        -0.1944, -0.1298, -0.1658, -0.1071, -0.1976, -0.1934, -0.1895, -0.1373,
        -0.1137, -0.1700, -0.1510, -0.2063, -0.2258, -0.1611, -0.1583, -0.1733])), ('lay
er_norm2.weight', tensor([1.0088, 1.1795, 1.4034, 1.4036, 1.2549, 1.2246, 1.4406, 1.337
7, 1.3466,
        1.2249, 1.0316, 1.4019, 1.2372, 1.2911, 1.3759, 1.2789, 1.3523, 1.1876,
        1.2773, 1.4231, 1.3214, 1.3419, 1.3469, 1.3339, 1.2604, 1.2856, 1.3525,
        1.3489, 1.3772, 1.2867, 1.4221, 1.3348, 1.0704, 1.3385, 1.4697, 1.3864,
        1.1527, 1.2950, 1.3618, 1.3710, 1.2211, 1.3575, 1.0700, 1.2920, 1.3538,
        1.2695, 1.3416, 1.3409, 1.3592, 1.2259, 1.3771, 1.3501, 1.0768, 1.1950,
        1.3422, 1.4444, 0.9976, 1.3377, 1.1830, 1.2492, 1.0187, 1.3002, 1.3764,
        1.1849])), ('layer_norm2.bias', tensor([ 0.0281,  0.0147,  0.1021,  0.0693,  0.1
087,  0.0149,  0.1160,  0.0361,
         0.0625,  0.0154,  0.0318,  0.1181,  0.0344,  0.0193,  0.0590,  0.0421,
         0.0492,  0.0829,  0.0528,  0.1390,  0.0781,  0.0432, -0.0166,  0.1057,
         0.0595,  0.0695,  0.0575,  0.1199,  0.1411,  0.0970,  0.0791,  0.0587,
         0.0254,  0.0946,  0.1356,  0.0473,  0.0288,  0.0353,  0.0501,  0.1328,
         0.0805,  0.0656,  0.0571,  0.0485,  0.0391,  0.0577,  0.0323,  0.1022,
         0.0787,  0.0399,  0.0961,  0.0791,  0.0670,  0.0648,  0.0815,  0.1378,
         0.0119,  0.0263,  0.0615,  0.0296,  0.0353,  0.0805,  0.1223,  0.0591]))])
```

Plot Effect of changing learning rates on the accuracy and loss

In [32]:

```python
# Plot the effect of changing learning rates on final accuracy and loss
plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)
plt.plot(learning_rates, final_train_losses, label='Final Training Loss')
plt.plot(learning_rates, final_val_losses, label='Final Validation Loss')
plt.title('Effect of Learning Rates on Final Loss')
plt.xlabel('Learning Rate')
plt.ylabel('Loss')
plt.xscale('log')  # Use log scale for better visualization
plt.legend()

plt.subplot(1, 2, 2)
plt.plot(learning_rates, final_train_accuracies, label='Final Training Accuracy')
plt.plot(learning_rates, final_val_accuracies, label='Final Validation Accuracy')
plt.title('Effect of Learning Rates on Final Accuracy')
plt.xlabel('Learning Rate')
plt.ylabel('Accuracy')
plt.xscale('log')  # Use log scale for better visualization
plt.legend()

plt.tight_layout()
plt.show()
```



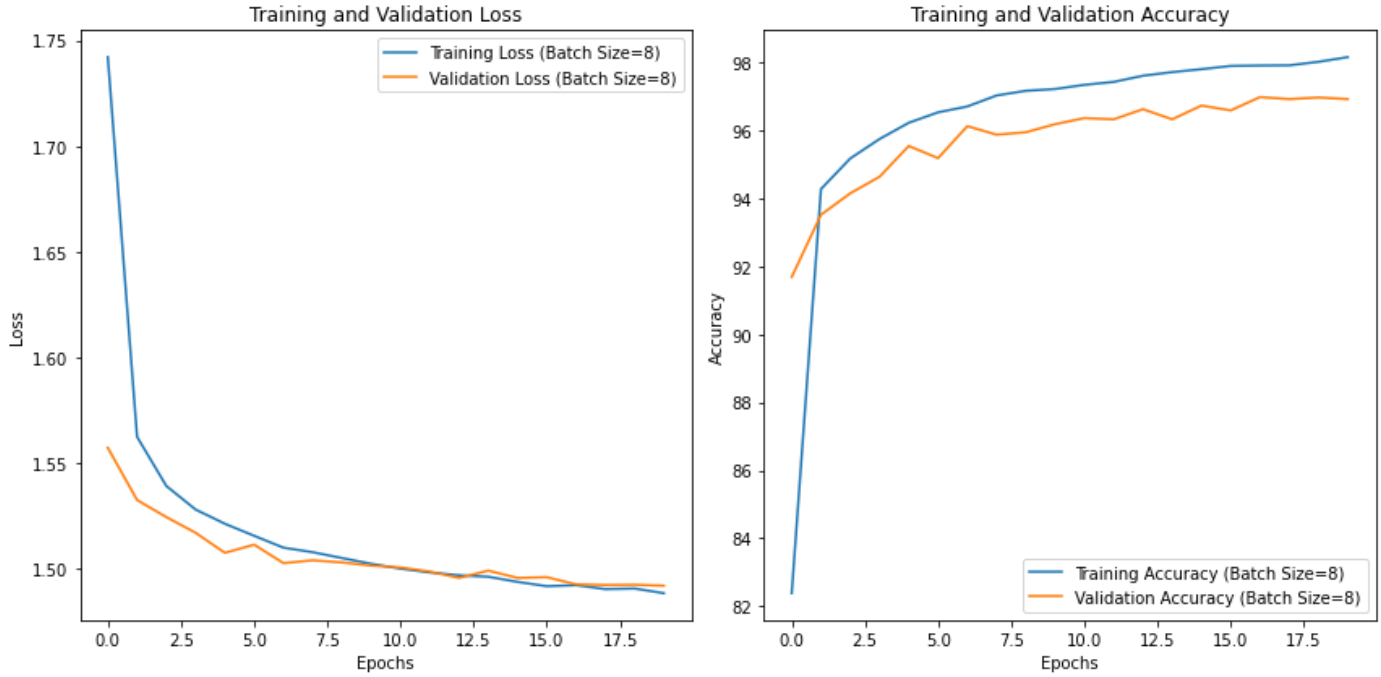
Customized Training Loop to try some different Batch sizes

In [33]:
```
batch_sizes = [8, 16, 32, 64, 128, 256, 1024]

# Lists to store final metrics for each batch size
final_train_losses = []
final_val_losses = []
final_train_accuracies = []
final_val_accuracies = []

# Initialize variables to keep track of the best model
best_model = None
best_val_accuracy = 0.0
best_hyperparameters = None

for batch_size in batch_sizes:
    # Create data loaders with the new batch size
    train_loader = DataLoader(train_set, batch_size=batch_size, shuffle=True)
    val_loader = DataLoader(val_set, batch_size=batch_size, shuffle=False)

    # Instantiate the model
    model = Model(in_features, h1, h2, out_features)

    # Set a new learning rate
    optimizer = optim.SGD(model.parameters(), lr=0.01, weight_decay=1e-5)

    # Lists to store metrics for plotting
    train_loss_history = []
    val_loss_history = []
    train_acc_history = []
    val_acc_history = []

    # Training loop
    for epoch in range(num_epochs):
        # Training
        average_loss, train_accuracy = train(model, train_loader)
        train_loss_history.append(average_loss)
        train_acc_history.append(train_accuracy)

        # Validation
```

```
        average_val_loss, val_accuracy = validate(model, val_loader)
        val_loss_history.append(average_val_loss)
        val_acc_history.append(val_accuracy)

        # Print and visualize metrics
        print(f'Epoch [{epoch + 1}/{num_epochs}], '
              f'Training Loss: {average_loss:.4f}, Training Accuracy: {train_accuracy:.4
              f'Validation Loss: {average_val_loss:.4f}, Validation Accuracy: {val_accur

    # Plot training and validation loss
    plt.figure(figsize=(12, 6))
    plt.subplot(1, 2, 1)
    plt.plot(train_loss_history, label=f'Training Loss (Batch Size={batch_size})')
    plt.plot(val_loss_history, label=f'Validation Loss (Batch Size={batch_size})')
    plt.title('Training and Validation Loss')
    plt.xlabel('Epochs')
    plt.ylabel('Loss')
    plt.legend()

    # Plot training and validation accuracy
    plt.subplot(1, 2, 2)
    plt.plot(train_acc_history, label=f'Training Accuracy (Batch Size={batch_size})')
    plt.plot(val_acc_history, label=f'Validation Accuracy (Batch Size={batch_size})')
    plt.title('Training and Validation Accuracy')
    plt.xlabel('Epochs')
    plt.ylabel('Accuracy')
    plt.legend()

    plt.tight_layout()
    plt.show()

    # Store final metrics for each batch size
    final_train_losses.append(train_loss_history[-1])
    final_val_losses.append(val_loss_history[-1])
    final_train_accuracies.append(train_acc_history[-1])
    final_val_accuracies.append(val_acc_history[-1])

    # Check if current model has the best validation accuracy
    if val_accuracy > best_val_accuracy:
        best_val_accuracy = val_accuracy
        best_model = model.state_dict().copy()
        best_hyperparameters = {'batch_size': batch_size}
```

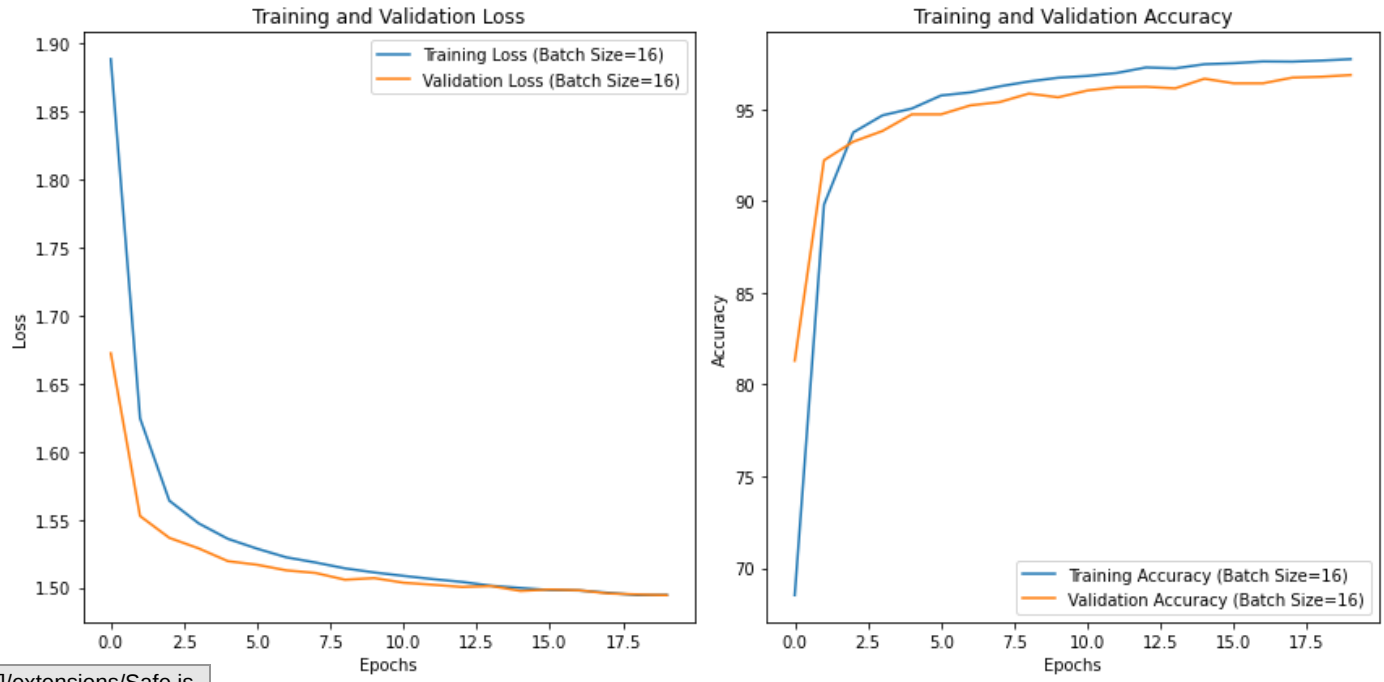
```
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 432.64it/s]
Epoch [1/20], Training Loss: 1.7421, Training Accuracy: 82.3810%, Validation Loss: 1.557
2, Validation Accuracy: 91.6905%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 423.21it/s]
Epoch [2/20], Training Loss: 1.5624, Training Accuracy: 94.2798%, Validation Loss: 1.532
4, Validation Accuracy: 93.5238%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 456.33it/s]
Epoch [3/20], Training Loss: 1.5391, Training Accuracy: 95.1815%, Validation Loss: 1.524
4, Validation Accuracy: 94.1548%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 440.70it/s]
Epoch [4/20], Training Loss: 1.5279, Training Accuracy: 95.7530%, Validation Loss: 1.516
9, Validation Accuracy: 94.6429%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 439.22it/s]
Epoch [5/20], Training Loss: 1.5212, Training Accuracy: 96.2321%, Validation Loss: 1.507
4, Validation Accuracy: 95.5476%
```

```
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 474.37it/s]
Epoch [6/20], Training Loss: 1.5155, Training Accuracy: 96.5387%, Validation Loss: 1.511
3, Validation Accuracy: 95.1905%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 464.72it/s]
Epoch [7/20], Training Loss: 1.5099, Training Accuracy: 96.7113%, Validation Loss: 1.502
6, Validation Accuracy: 96.1310%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 432.46it/s]
Epoch [8/20], Training Loss: 1.5077, Training Accuracy: 97.0357%, Validation Loss: 1.503
9, Validation Accuracy: 95.8810%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 445.28it/s]
Epoch [9/20], Training Loss: 1.5050, Training Accuracy: 97.1726%, Validation Loss: 1.502
9, Validation Accuracy: 95.9524%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 447.40it/s]
Epoch [10/20], Training Loss: 1.5022, Training Accuracy: 97.2262%, Validation Loss: 1.50
14, Validation Accuracy: 96.1905%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 481.59it/s]
Epoch [11/20], Training Loss: 1.4999, Training Accuracy: 97.3482%, Validation Loss: 1.50
05, Validation Accuracy: 96.3690%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 489.03it/s]
Epoch [12/20], Training Loss: 1.4981, Training Accuracy: 97.4345%, Validation Loss: 1.49
85, Validation Accuracy: 96.3333%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 472.27it/s]
Epoch [13/20], Training Loss: 1.4968, Training Accuracy: 97.6161%, Validation Loss: 1.49
55, Validation Accuracy: 96.6310%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 469.68it/s]
Epoch [14/20], Training Loss: 1.4961, Training Accuracy: 97.7232%, Validation Loss: 1.49
90, Validation Accuracy: 96.3333%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 464.10it/s]
Epoch [15/20], Training Loss: 1.4937, Training Accuracy: 97.8095%, Validation Loss: 1.49
56, Validation Accuracy: 96.7381%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 489.52it/s]
Epoch [16/20], Training Loss: 1.4916, Training Accuracy: 97.9048%, Validation Loss: 1.49
59, Validation Accuracy: 96.5952%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.25it/s]
Epoch [17/20], Training Loss: 1.4921, Training Accuracy: 97.9167%, Validation Loss: 1.49
25, Validation Accuracy: 96.9881%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 485.28it/s]
Epoch [18/20], Training Loss: 1.4903, Training Accuracy: 97.9226%, Validation Loss: 1.49
22, Validation Accuracy: 96.9286%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 479.30it/s]
Epoch [19/20], Training Loss: 1.4905, Training Accuracy: 98.0268%, Validation Loss: 1.49
23, Validation Accuracy: 96.9762%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 483.71it/s]
Epoch [20/20], Training Loss: 1.4883, Training Accuracy: 98.1637%, Validation Loss: 1.49
n Accuracy: 96.9286%
```



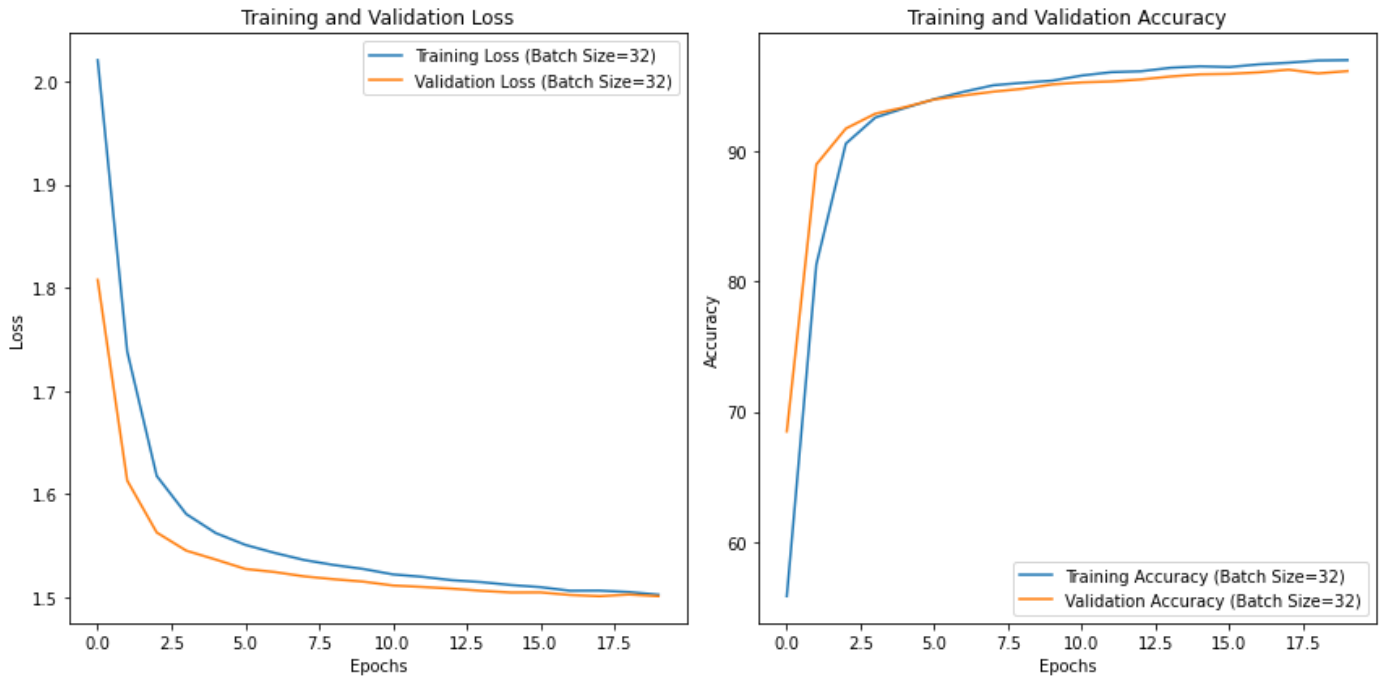

```
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 465.81it/s]
Epoch [1/20], Training Loss: 1.8884, Training Accuracy: 68.5298%, Validation Loss: 1.672
3, Validation Accuracy: 81.2976%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 464.20it/s]
Epoch [2/20], Training Loss: 1.6245, Training Accuracy: 89.7976%, Validation Loss: 1.552
7, Validation Accuracy: 92.2143%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 471.86it/s]
Epoch [3/20], Training Loss: 1.5640, Training Accuracy: 93.7292%, Validation Loss: 1.536
7, Validation Accuracy: 93.2262%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 469.91it/s]
Epoch [4/20], Training Loss: 1.5474, Training Accuracy: 94.6607%, Validation Loss: 1.529
0, Validation Accuracy: 93.8095%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 469.48it/s]
Epoch [5/20], Training Loss: 1.5360, Training Accuracy: 95.0298%, Validation Loss: 1.519
6, Validation Accuracy: 94.7143%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 468.19it/s]
Epoch [6/20], Training Loss: 1.5287, Training Accuracy: 95.7411%, Validation Loss: 1.516
9, Validation Accuracy: 94.7143%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 465.79it/s]
Epoch [7/20], Training Loss: 1.5223, Training Accuracy: 95.9107%, Validation Loss: 1.512
8, Validation Accuracy: 95.2024%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 468.41it/s]
Epoch [8/20], Training Loss: 1.5186, Training Accuracy: 96.2381%, Validation Loss: 1.510
9, Validation Accuracy: 95.3810%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 470.21it/s]
Epoch [9/20], Training Loss: 1.5143, Training Accuracy: 96.5030%, Validation Loss: 1.505
9, Validation Accuracy: 95.8452%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 467.35it/s]
Epoch [10/20], Training Loss: 1.5112, Training Accuracy: 96.7083%, Validation Loss: 1.50
n Accuracy: 95.6429%
```

```
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 467.63it/s]
Epoch [11/20], Training Loss: 1.5087, Training Accuracy: 96.8065%, Validation Loss: 1.50
37, Validation Accuracy: 96.0119%
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 467.91it/s]
Epoch [12/20], Training Loss: 1.5064, Training Accuracy: 96.9673%, Validation Loss: 1.50
22, Validation Accuracy: 96.1905%
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 466.02it/s]
Epoch [13/20], Training Loss: 1.5043, Training Accuracy: 97.2679%, Validation Loss: 1.50
06, Validation Accuracy: 96.2143%
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 461.45it/s]
Epoch [14/20], Training Loss: 1.5015, Training Accuracy: 97.2202%, Validation Loss: 1.50
13, Validation Accuracy: 96.1310%
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 466.33it/s]
Epoch [15/20], Training Loss: 1.4998, Training Accuracy: 97.4435%, Validation Loss: 1.49
77, Validation Accuracy: 96.6548%
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 467.65it/s]
Epoch [16/20], Training Loss: 1.4984, Training Accuracy: 97.5000%, Validation Loss: 1.49
87, Validation Accuracy: 96.4048%
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 462.77it/s]
Epoch [17/20], Training Loss: 1.4980, Training Accuracy: 97.5952%, Validation Loss: 1.49
82, Validation Accuracy: 96.4048%
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 467.26it/s]
Epoch [18/20], Training Loss: 1.4963, Training Accuracy: 97.5833%, Validation Loss: 1.49
58, Validation Accuracy: 96.7143%
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 465.03it/s]
Epoch [19/20], Training Loss: 1.4946, Training Accuracy: 97.6429%, Validation Loss: 1.49
51, Validation Accuracy: 96.7619%
Training: 100%|██████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 465.77it/s]
Epoch [20/20], Training Loss: 1.4948, Training Accuracy: 97.7202%, Validation Loss: 1.49
45, Validation Accuracy: 96.8571%
```



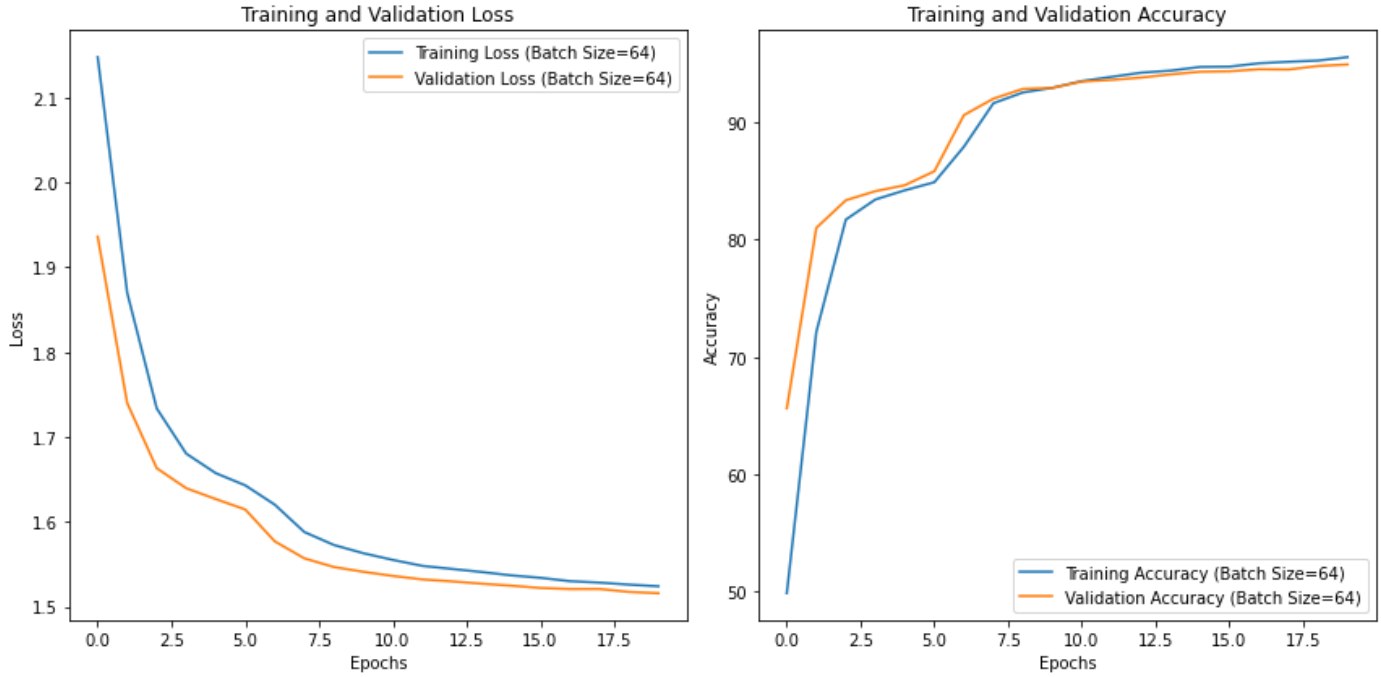
```
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 448.00it/s]
Epoch [1/20], Training Loss: 2.0204, Training Accuracy: 55.8304%, Validation Loss: 1.8077, Validation Accuracy: 68.4881%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 447.55it/s]
Epoch [2/20], Training Loss: 1.7382, Training Accuracy: 81.3095%, Validation Loss: 1.6132, Validation Accuracy: 89.0000%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 446.86it/s]
Epoch [3/20], Training Loss: 1.6177, Training Accuracy: 90.6042%, Validation Loss: 1.5629, Validation Accuracy: 91.7619%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 451.01it/s]
Epoch [4/20], Training Loss: 1.5807, Training Accuracy: 92.6101%, Validation Loss: 1.5453, Validation Accuracy: 92.9048%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 438.52it/s]
Epoch [5/20], Training Loss: 1.5623, Training Accuracy: 93.3393%, Validation Loss: 1.5366, Validation Accuracy: 93.4048%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 450.03it/s]
Epoch [6/20], Training Loss: 1.5509, Training Accuracy: 94.0238%, Validation Loss: 1.5276, Validation Accuracy: 94.0000%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 448.39it/s]
Epoch [7/20], Training Loss: 1.5431, Training Accuracy: 94.5923%, Validation Loss: 1.5245, Validation Accuracy: 94.3095%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 445.47it/s]
Epoch [8/20], Training Loss: 1.5362, Training Accuracy: 95.0893%, Validation Loss: 1.5204, Validation Accuracy: 94.5952%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 443.04it/s]
Epoch [9/20], Training Loss: 1.5314, Training Accuracy: 95.2798%, Validation Loss: 1.5176, Validation Accuracy: 94.8214%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 444.35it/s]
Epoch [10/20], Training Loss: 1.5275, Training Accuracy: 95.4435%, Validation Loss: 1.5153, Validation Accuracy: 95.1548%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 448.51it/s]
Epoch [11/20], Training Loss: 1.5223, Training Accuracy: 95.8363%, Validation Loss: 1.5115, Validation Accuracy: 95.3095%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 446.20it/s]
Epoch [12/20], Training Loss: 1.5199, Training Accuracy: 96.0952%, Validation Loss: 1.5101, Validation Accuracy: 95.3810%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 437.76it/s]
Epoch [13/20], Training Loss: 1.5167, Training Accuracy: 96.1637%, Validation Loss: 1.5085, Validation Accuracy: 95.5357%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 448.77it/s]
Epoch [14/20], Training Loss: 1.5148, Training Accuracy: 96.4315%, Validation Loss: 1.5064, Validation Accuracy: 95.7619%
Training: 100%|██████████████████████████████████████████████████████████| 1050/1050 [00:02<00:00, 442.66it/s]
Epoch [15/20], Training Loss: 1.5120, Training Accuracy: 96.5446%, Validation Loss: 1.50
n Accuracy: 95.9286%
```

```
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.61it/s]
Epoch [16/20], Training Loss: 1.5099, Training Accuracy: 96.4881%, Validation Loss: 1.50
49, Validation Accuracy: 95.9643%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.16it/s]
Epoch [17/20], Training Loss: 1.5065, Training Accuracy: 96.6964%, Validation Loss: 1.50
24, Validation Accuracy: 96.0833%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.92it/s]
Epoch [18/20], Training Loss: 1.5066, Training Accuracy: 96.8214%, Validation Loss: 1.50
12, Validation Accuracy: 96.2857%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.44it/s]
Epoch [19/20], Training Loss: 1.5052, Training Accuracy: 96.9881%, Validation Loss: 1.50
28, Validation Accuracy: 96.0000%
Training: 100%|██████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.97it/s]
Epoch [20/20], Training Loss: 1.5026, Training Accuracy: 97.0179%, Validation Loss: 1.50
12, Validation Accuracy: 96.1786%
```



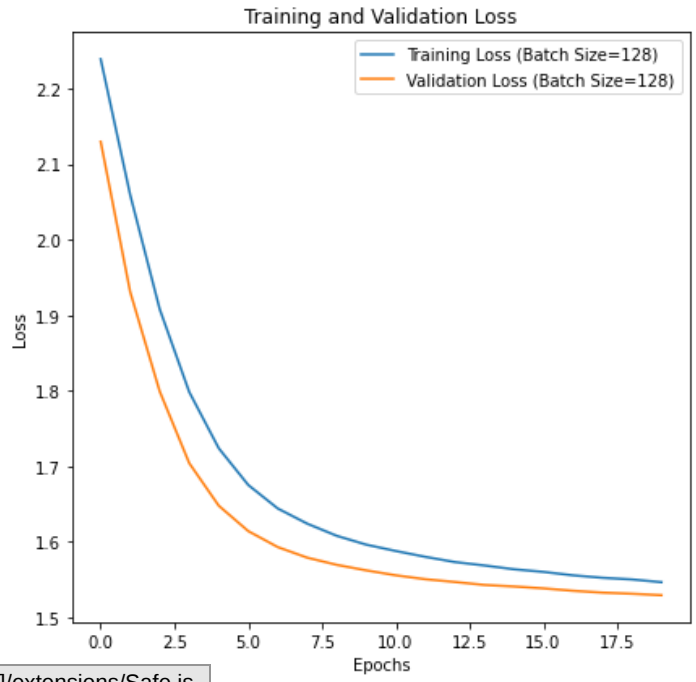

```
Training: 100%|██████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 367.86it/s]
Epoch [1/20], Training Loss: 2.1476, Training Accuracy: 49.8512%, Validation Loss: 1.936
1, Validation Accuracy: 65.6190%
Training: 100%|██████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 372.81it/s]
Epoch [2/20], Training Loss: 1.8698, Training Accuracy: 72.1399%, Validation Loss: 1.739
9, Validation Accuracy: 81.0119%
Training: 100%|██████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 369.93it/s]
Epoch [3/20], Training Loss: 1.7337, Training Accuracy: 81.7202%, Validation Loss: 1.663
2, Validation Accuracy: 83.3690%
Training: 100%|██████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 369.66it/s]
Epoch [4/20], Training Loss: 1.6804, Training Accuracy: 83.4345%, Validation Loss: 1.639
6, Validation Accuracy: 84.1429%
Training: 100%|██████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 366.32it/s]
Epoch [5/20], Training Loss: 1.6574, Training Accuracy: 84.2202%, Validation Loss: 1.626
[illegible] Accuracy: 84.6548%
```

```
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 362.39it/s]
Epoch [6/20], Training Loss: 1.6430, Training Accuracy: 84.9077%, Validation Loss: 1.6146, Validation Accuracy: 85.8452%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 370.71it/s]
Epoch [7/20], Training Loss: 1.6203, Training Accuracy: 87.9345%, Validation Loss: 1.5770, Validation Accuracy: 90.6310%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 369.59it/s]
Epoch [8/20], Training Loss: 1.5880, Training Accuracy: 91.6458%, Validation Loss: 1.5570, Validation Accuracy: 92.0357%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 354.72it/s]
Epoch [9/20], Training Loss: 1.5728, Training Accuracy: 92.5655%, Validation Loss: 1.5468, Validation Accuracy: 92.8690%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 356.40it/s]
Epoch [10/20], Training Loss: 1.5630, Training Accuracy: 92.9613%, Validation Loss: 1.5411, Validation Accuracy: 92.9524%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 360.06it/s]
Epoch [11/20], Training Loss: 1.5552, Training Accuracy: 93.5387%, Validation Loss: 1.5363, Validation Accuracy: 93.5000%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 359.05it/s]
Epoch [12/20], Training Loss: 1.5481, Training Accuracy: 93.8958%, Validation Loss: 1.5322, Validation Accuracy: 93.6310%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 366.07it/s]
Epoch [13/20], Training Loss: 1.5445, Training Accuracy: 94.2500%, Validation Loss: 1.5299, Validation Accuracy: 93.8333%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 353.68it/s]
Epoch [14/20], Training Loss: 1.5409, Training Accuracy: 94.4286%, Validation Loss: 1.5271, Validation Accuracy: 94.1071%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 369.41it/s]
Epoch [15/20], Training Loss: 1.5370, Training Accuracy: 94.7411%, Validation Loss: 1.5248, Validation Accuracy: 94.3333%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 364.04it/s]
Epoch [16/20], Training Loss: 1.5341, Training Accuracy: 94.7679%, Validation Loss: 1.5221, Validation Accuracy: 94.3690%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 364.29it/s]
Epoch [17/20], Training Loss: 1.5301, Training Accuracy: 95.0536%, Validation Loss: 1.5208, Validation Accuracy: 94.5476%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 364.67it/s]
Epoch [18/20], Training Loss: 1.5284, Training Accuracy: 95.1845%, Validation Loss: 1.5209, Validation Accuracy: 94.5238%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 365.30it/s]
Epoch [19/20], Training Loss: 1.5260, Training Accuracy: 95.2857%, Validation Loss: 1.5174, Validation Accuracy: 94.8214%
Training: 100%|██████████████████████████████████████████████████████████| 525/525 [00:01<00:00, 368.11it/s]
Epoch [20/20], Training Loss: 1.5241, Training Accuracy: 95.5774%, Validation Loss: 1.51
n Accuracy: 94.9524%
```



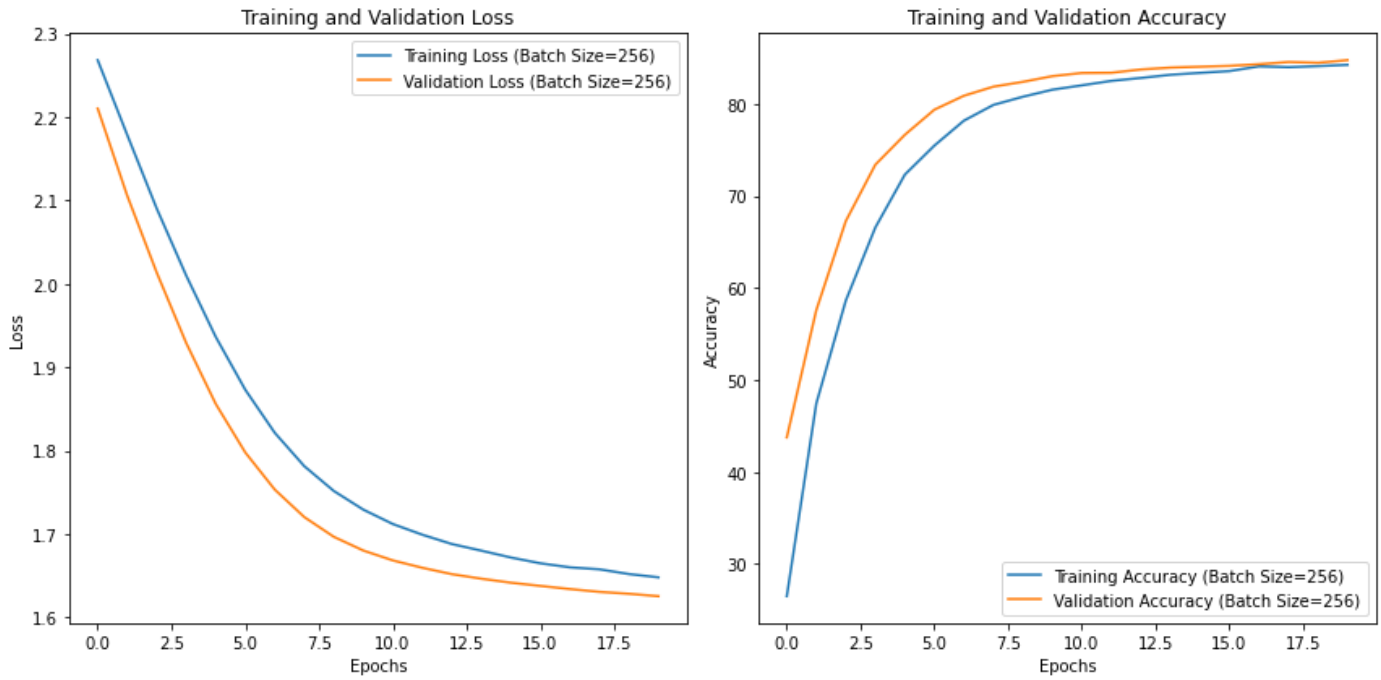

```
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 281.94it/s]
Epoch [1/20], Training Loss: 2.2389, Training Accuracy: 35.2262%, Validation Loss: 2.129
7, Validation Accuracy: 60.0238%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 275.55it/s]
Epoch [2/20], Training Loss: 2.0591, Training Accuracy: 61.7381%, Validation Loss: 1.930
8, Validation Accuracy: 68.0476%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 275.55it/s]
Epoch [3/20], Training Loss: 1.9080, Training Accuracy: 72.0030%, Validation Loss: 1.799
2, Validation Accuracy: 80.9881%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 279.35it/s]
Epoch [4/20], Training Loss: 1.7985, Training Accuracy: 81.8155%, Validation Loss: 1.704
3, Validation Accuracy: 86.4048%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.83it/s]
Epoch [5/20], Training Loss: 1.7240, Training Accuracy: 85.7738%, Validation Loss: 1.648
0, Validation Accuracy: 88.5000%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 277.58it/s]
Epoch [6/20], Training Loss: 1.6754, Training Accuracy: 87.9851%, Validation Loss: 1.614
4, Validation Accuracy: 89.8810%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 276.42it/s]
Epoch [7/20], Training Loss: 1.6443, Training Accuracy: 89.3036%, Validation Loss: 1.593
2, Validation Accuracy: 90.5595%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.54it/s]
Epoch [8/20], Training Loss: 1.6242, Training Accuracy: 90.0655%, Validation Loss: 1.579
2, Validation Accuracy: 91.2143%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.83it/s]
Epoch [9/20], Training Loss: 1.6081, Training Accuracy: 90.7440%, Validation Loss: 1.569
8, Validation Accuracy: 91.5476%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 278.66it/s]
Epoch [10/20], Training Loss: 1.5965, Training Accuracy: 91.2649%, Validation Loss: 1.56
n Accuracy: 92.0357%
```

```
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 281.13it/s]
Epoch [11/20], Training Loss: 1.5881, Training Accuracy: 91.5804%, Validation Loss: 1.55
59, Validation Accuracy: 92.2381%
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 281.13it/s]
Epoch [12/20], Training Loss: 1.5804, Training Accuracy: 91.9345%, Validation Loss: 1.55
07, Validation Accuracy: 92.5952%
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 278.17it/s]
Epoch [13/20], Training Loss: 1.5735, Training Accuracy: 92.3869%, Validation Loss: 1.54
71, Validation Accuracy: 92.9048%
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 277.49it/s]
Epoch [14/20], Training Loss: 1.5690, Training Accuracy: 92.6607%, Validation Loss: 1.54
32, Validation Accuracy: 93.0000%
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.53it/s]
Epoch [15/20], Training Loss: 1.5639, Training Accuracy: 92.8482%, Validation Loss: 1.54
11, Validation Accuracy: 93.1071%
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 279.35it/s]
Epoch [16/20], Training Loss: 1.5605, Training Accuracy: 93.1905%, Validation Loss: 1.53
86, Validation Accuracy: 93.2262%
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 271.58it/s]
Epoch [17/20], Training Loss: 1.5559, Training Accuracy: 93.3304%, Validation Loss: 1.53
55, Validation Accuracy: 93.5595%
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 282.94it/s]
Epoch [18/20], Training Loss: 1.5527, Training Accuracy: 93.6101%, Validation Loss: 1.53
30, Validation Accuracy: 93.7024%
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 278.76it/s]
Epoch [19/20], Training Loss: 1.5505, Training Accuracy: 93.6488%, Validation Loss: 1.53
15, Validation Accuracy: 93.8095%
Training: 100%|██████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.24it/s]
Epoch [20/20], Training Loss: 1.5468, Training Accuracy: 93.8274%, Validation Loss: 1.52
96, Validation Accuracy: 93.9643%
```



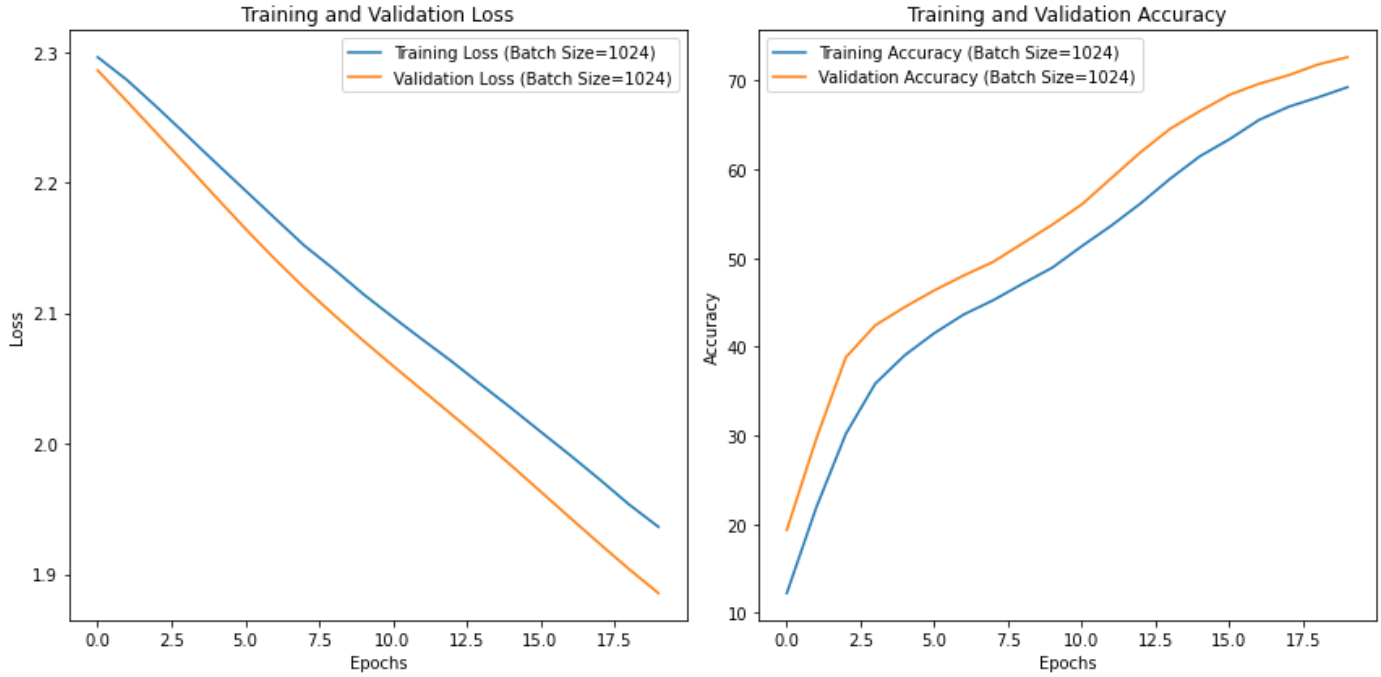
```
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 212.10it/s]
Epoch [1/20], Training Loss: 2.2682, Training Accuracy: 26.5119%, Validation Loss: 2.210
3, Validation Accuracy: 43.7619%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 210.08it/s]
Epoch [2/20], Training Loss: 2.1786, Training Accuracy: 47.4583%, Validation Loss: 2.105
8, Validation Accuracy: 57.5952%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 211.09it/s]
Epoch [3/20], Training Loss: 2.0903, Training Accuracy: 58.6369%, Validation Loss: 2.013
5, Validation Accuracy: 67.2738%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 206.48it/s]
Epoch [4/20], Training Loss: 2.0099, Training Accuracy: 66.5685%, Validation Loss: 1.929
0, Validation Accuracy: 73.3929%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 206.16it/s]
Epoch [5/20], Training Loss: 1.9364, Training Accuracy: 72.3155%, Validation Loss: 1.856
0, Validation Accuracy: 76.6429%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 212.47it/s]
Epoch [6/20], Training Loss: 1.8731, Training Accuracy: 75.4673%, Validation Loss: 1.797
9, Validation Accuracy: 79.3690%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 210.75it/s]
Epoch [7/20], Training Loss: 1.8214, Training Accuracy: 78.1756%, Validation Loss: 1.753
3, Validation Accuracy: 80.8810%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 210.75it/s]
Epoch [8/20], Training Loss: 1.7814, Training Accuracy: 79.8869%, Validation Loss: 1.720
4, Validation Accuracy: 81.8690%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 209.42it/s]
Epoch [9/20], Training Loss: 1.7517, Training Accuracy: 80.7738%, Validation Loss: 1.696
7, Validation Accuracy: 82.3929%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 207.45it/s]
Epoch [10/20], Training Loss: 1.7293, Training Accuracy: 81.5417%, Validation Loss: 1.68
03, Validation Accuracy: 83.0119%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 209.09it/s]
Epoch [11/20], Training Loss: 1.7121, Training Accuracy: 82.0149%, Validation Loss: 1.66
84, Validation Accuracy: 83.3571%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 210.75it/s]
Epoch [12/20], Training Loss: 1.6992, Training Accuracy: 82.5000%, Validation Loss: 1.65
96, Validation Accuracy: 83.3810%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 209.95it/s]
Epoch [13/20], Training Loss: 1.6881, Training Accuracy: 82.8155%, Validation Loss: 1.65
20, Validation Accuracy: 83.7381%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 212.79it/s]
Epoch [14/20], Training Loss: 1.6801, Training Accuracy: 83.1607%, Validation Loss: 1.64
65, Validation Accuracy: 83.9405%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 210.08it/s]
Epoch [15/20], Training Loss: 1.6719, Training Accuracy: 83.3690%, Validation Loss: 1.64
n Accuracy: 84.0357%
```

```
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 211.09it/s]
Epoch [16/20], Training Loss: 1.6650, Training Accuracy: 83.5625%, Validation Loss: 1.63
79, Validation Accuracy: 84.1310%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 208.76it/s]
Epoch [17/20], Training Loss: 1.6601, Training Accuracy: 84.0744%, Validation Loss: 1.63
41, Validation Accuracy: 84.3095%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 207.13it/s]
Epoch [18/20], Training Loss: 1.6578, Training Accuracy: 83.9851%, Validation Loss: 1.63
07, Validation Accuracy: 84.5476%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 205.84it/s]
Epoch [19/20], Training Loss: 1.6521, Training Accuracy: 84.1012%, Validation Loss: 1.62
84, Validation Accuracy: 84.4643%
Training: 100%|██████████████████████████████████████████████████████████| 13
2/132 [00:00<00:00, 212.33it/s]
Epoch [20/20], Training Loss: 1.6482, Training Accuracy: 84.2411%, Validation Loss: 1.62
55, Validation Accuracy: 84.7500%
```




```
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 92.43it/s]
Epoch [1/20], Training Loss: 2.2960, Training Accuracy: 12.2321%, Validation Loss: 2.285
9, Validation Accuracy: 19.3571%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 99.66it/s]
Epoch [2/20], Training Loss: 2.2784, Training Accuracy: 21.9018%, Validation Loss: 2.262
0, Validation Accuracy: 29.6429%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 98.18it/s]
Epoch [3/20], Training Loss: 2.2578, Training Accuracy: 30.1845%, Validation Loss: 2.237
8, Validation Accuracy: 38.8214%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 93.21it/s]
Epoch [4/20], Training Loss: 2.2366, Training Accuracy: 35.8661%, Validation Loss: 2.213
6, Validation Accuracy: 42.4286%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 91.15it/s]
Epoch [5/20], Training Loss: 2.2151, Training Accuracy: 39.0595%, Validation Loss: 2.189
[illegible] Accuracy: 44.4881%
```

```
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 95.63it/s]
Epoch [6/20], Training Loss: 2.1940, Training Accuracy: 41.5417%, Validation Loss: 2.164
7, Validation Accuracy: 46.3690%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 93.47it/s]
Epoch [7/20], Training Loss: 2.1728, Training Accuracy: 43.6518%, Validation Loss: 2.141
5, Validation Accuracy: 48.0476%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 99.36it/s]
Epoch [8/20], Training Loss: 2.1517, Training Accuracy: 45.2589%, Validation Loss: 2.119
4, Validation Accuracy: 49.5833%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 99.96it/s]
Epoch [9/20], Training Loss: 2.1337, Training Accuracy: 47.1280%, Validation Loss: 2.098
7, Validation Accuracy: 51.6786%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 98.48it/s]
Epoch [10/20], Training Loss: 2.1146, Training Accuracy: 48.9286%, Validation Loss: 2.07
89, Validation Accuracy: 53.7738%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 93.47it/s]
Epoch [11/20], Training Loss: 2.0970, Training Accuracy: 51.3512%, Validation Loss: 2.05
98, Validation Accuracy: 56.0595%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 97.03it/s]
Epoch [12/20], Training Loss: 2.0798, Training Accuracy: 53.6637%, Validation Loss: 2.04
10, Validation Accuracy: 59.0238%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 98.48it/s]
Epoch [13/20], Training Loss: 2.0630, Training Accuracy: 56.1935%, Validation Loss: 2.02
22, Validation Accuracy: 61.9643%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 92.17it/s]
Epoch [14/20], Training Loss: 2.0451, Training Accuracy: 58.9702%, Validation Loss: 2.00
31, Validation Accuracy: 64.5952%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 91.66it/s]
Epoch [15/20], Training Loss: 2.0275, Training Accuracy: 61.4821%, Validation Loss: 1.98
36, Validation Accuracy: 66.5714%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 97.32it/s]
Epoch [16/20], Training Loss: 2.0093, Training Accuracy: 63.3958%, Validation Loss: 1.96
36, Validation Accuracy: 68.3929%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 96.19it/s]
Epoch [17/20], Training Loss: 1.9914, Training Accuracy: 65.5655%, Validation Loss: 1.94
35, Validation Accuracy: 69.6429%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 61.16it/s]
Epoch [18/20], Training Loss: 1.9728, Training Accuracy: 67.0476%, Validation Loss: 1.92
35, Validation Accuracy: 70.6071%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 97.61it/s]
Epoch [19/20], Training Loss: 1.9536, Training Accuracy: 68.1131%, Validation Loss: 1.90
40, Validation Accuracy: 71.8214%
Training: 100%|██████████████████████████████████████████████████████████|
33/33 [00:00<00:00, 94.00it/s]
Epoch [20/20], Training Loss: 1.9363, Training Accuracy: 69.2589%, Validation Loss: 1.88
n Accuracy: 72.6429%
```



Show best model according to Batch size

In [34]:
```
# Print the best hyperparameters
print("Best Hyperparameters:")
print(best_hyperparameters)

# Print the best model
print("Best Model:")
print(best_model)
```

```
Best Hyperparameters:
{'batch_size': 8}
Best Model:
OrderedDict([('fc1.weight', tensor([[ 0.0235,  0.0055,  0.0338,  ...,  0.0154,  0.0117,
0.0190],
        [ 0.0020,  0.0048,  0.0262,  ..., -0.0151,  0.0286, -0.0057],
        [-0.0312,  0.0095, -0.0091,  ...,  0.0035,  0.0339,  0.0215],
        ...,
        [ 0.0156, -0.0300,  0.0063,  ...,  0.0328,  0.0350, -0.0097],
        [ 0.0182,  0.0123,  0.0135,  ...,  0.0283, -0.0331, -0.0211],
        [-0.0223, -0.0011,  0.0023,  ..., -0.0332,  0.0161, -0.0238]])), ('fc1.bias', te
nsor([-2.0023e-02,  1.5080e-03, -8.4792e-03,  3.4937e-02, -2.6961e-02,
         5.3109e-03,  8.3756e-03,  9.9554e-03,  4.1020e-03, -8.6610e-03,
         4.3383e-03, -3.3636e-02,  2.5909e-02, -2.1693e-02, -6.4318e-03,
        -1.8942e-02,  1.0250e-02, -1.8895e-02, -2.4110e-02, -2.2429e-02,
        -1.1774e-02, -2.1996e-02,  6.4330e-03, -1.4313e-03,  2.2289e-02,
        -3.4671e-03, -1.7575e-02, -1.7216e-02,  7.7502e-03,  8.6911e-04,
         7.7260e-03,  1.5217e-02,  1.5207e-02,  5.2379e-03,  2.0623e-03,
         5.9827e-03, -2.1279e-02,  4.2001e-05,  1.8920e-02,  2.2499e-03,
        -1.6407e-02,  3.5255e-02, -3.3990e-02, -2.4278e-02,  2.6848e-02,
         2.6398e-02, -1.9597e-02, -2.8663e-02, -1.7327e-02,  1.3988e-02,
         1.7688e-02, -2.9523e-03, -1.4425e-02, -4.1077e-04, -7.6218e-04,
         4.3685e-03, -1.5745e-02, -1.2535e-02, -3.4078e-02, -3.1853e-02,
         9.8184e-03,  2.6832e-02,  3.5274e-04, -4.4149e-04, -3.2909e-02,
         1.4545e-02,  2.9754e-02,  2.5891e-02,  1.9768e-02,  1.2942e-02,
         2.1767e-02, -1.5165e-02, -1.4294e-02, -7.2299e-03,  6.5536e-03,
        -1.3587e-02,  3.8459e-03,  2.0464e-02,  2.4080e-02, -2.6557e-02,
        -2.0445e-02, -2.9648e-02, -2.9986e-02,  2.8535e-02,  1.3282e-02,
         2.5025e-02,  2.2559e-02, -2.8757e-03,  3.2206e-02, -2.0460e-02,
        -2.7725e-02,  1.5940e-03,  3.0447e-02, -3.9297e-04,  2.2707e-02,
        -2.8698e-02, -2.9951e-03,  2.3468e-02,  5.6544e-03,  6.4088e-03,
        -2.0456e-03, -8.4375e-03,  8.5499e-03,  3.5196e-02, -2.3055e-02,
        -3.1740e-02, -1.5520e-02, -3.3599e-02,  1.2532e-02, -4.4821e-03,
        -3.3067e-02,  1.8563e-03, -2.5683e-02, -3.1211e-02,  1.6525e-02,
         3.4516e-02, -1.1236e-02,  1.6088e-02, -2.9311e-02,  3.4599e-02,
         1.4965e-02, -7.7352e-03, -2.1018e-02, -1.1759e-02,  1.0510e-02,
         3.4461e-02,  3.2491e-02, -3.4565e-02])), ('fc2.weight', tensor([[ 0.0246, -0.12
29,  0.0019,  ..., -0.1358,  0.1536, -0.0177],
        [-0.0259,  0.0780, -0.0202,  ...,  0.0272,  0.0660,  0.0091],
        [ 0.0586, -0.0343,  0.0627,  ...,  0.0291,  0.0064, -0.0233],
        ...,
        [-0.0360,  0.0492, -0.0235,  ..., -0.0069, -0.0154, -0.0131],
        [ 0.0046, -0.1598,  0.0262,  ..., -0.0382,  0.1857, -0.0070],
        [ 0.0099,  0.0129,  0.0350,  ..., -0.0456, -0.0215,  0.0316]])), ('fc2.bias', te
nsor([-0.0543, -0.0733, -0.0264, -0.0021,  0.0333, -0.0164, -0.0757, -0.0647,
        -0.0405, -0.0036, -0.0786,  0.0618, -0.0418,  0.0681, -0.0266,  0.0308,
        -0.0181, -0.0452, -0.0330, -0.0425,  0.0064, -0.0129,  0.0302, -0.0535,
        -0.0014, -0.0190,  0.0633, -0.0615,  0.0130, -0.0493, -0.0213,  0.0594,
         0.0534,  0.0014, -0.0489, -0.0549,  0.0065,  0.0303, -0.0171, -0.0058,
        -0.0478,  0.0564,  0.0028,  0.0220, -0.0574,  0.0122,  0.0135, -0.0307,
        -0.0029, -0.0106,  0.0598,  0.0065, -0.0271, -0.0514,  0.0203, -0.0576,
         0.0089,  0.0680,  0.0373, -0.0739,  0.0094, -0.0623, -0.0500, -0.0184])), ('ou
t.weight', tensor([[-2.0702e-01, -1.1471e-01, -1.6205e-01, -2.7605e-01, -6.8306e-02,
         -1.6631e-01, -1.3290e-01, -7.9074e-02,  3.7028e-01,  2.1337e-01,
         -2.2147e-02, -2.2369e-01, -1.8377e-01,  1.6728e-02, -1.5864e-01,
          4.9669e-01, -1.4584e-01, -1.2173e-01, -1.2422e-01, -3.0220e-01,
         -2.7133e-01,  6.0911e-01,  3.4132e-01, -1.5720e-01, -1.8261e-01,
         -2.0312e-01, -1.1237e-01, -7.9520e-02, -1.9766e-01,  1.0538e-02,
         -2.1985e-01, -2.0637e-01,  1.7561e-01, -3.0349e-01, -9.3602e-02,
         -2.5470e-01, -2.0942e-02, -2.5140e-01,  2.2267e-01, -2.2441e-01,
          5.7628e-01,  2.8278e-01, -8.8823e-02, -2.2012e-01, -6.3990e-02,
          7.2073e-01,  5.0816e-04, -1.6979e-01, -1.2868e-02,  1.0178e-01,
         -4.2518e-02, -1.8575e-01,  5.6240e-01, -3.1833e-01, -2.2931e-01,
          2.3878e-02,  3.3601e-01, -2.8208e-01, -1.0928e-02,  5.4677e-01,
```

```
 -9.0979e-02,  2.7445e-01,  5.1708e-01, -9.4300e-02],
[-1.4270e-01,  3.4592e-03, -1.3920e-01, -3.2410e-01, -4.9432e-02,
 -1.5849e-01, -6.6379e-02, -2.3659e-02, -2.3890e-01, -3.4502e-01,
 -1.7091e-01,  6.0506e-01, -1.0794e-01, -1.5826e-01, -3.1093e-01,
 -6.3902e-02, -4.7376e-02,  2.3746e-01,  9.0817e-02,  2.9283e-01,
  4.4175e-01, -2.1122e-01, -1.2317e-02, -1.5427e-01,  6.5303e-01,
 -2.0621e-01,  5.3344e-01, -2.8101e-01, -4.1508e-02, -3.7116e-02,
  4.6932e-01, -1.7906e-01, -4.4919e-02,  1.6738e-01, -2.2544e-01,
 -2.8705e-01, -2.0885e-02, -3.1941e-01, -1.1811e-01, -1.2120e-02,
  3.6480e-01, -2.7513e-01, -9.6517e-02, -1.9832e-01,  3.7538e-01,
 -8.8334e-02,  4.1435e-01, -4.7815e-02,  4.3981e-01,  2.2082e-01,
  5.6294e-01, -3.5066e-01,  2.8995e-02, -8.8287e-02,  3.8274e-01,
 -2.0439e-01, -2.6135e-01, -1.4409e-01, -8.2350e-03, -3.0387e-01,
 -2.3747e-02,  7.3654e-03,  1.5910e-01, -2.6510e-01],
[-2.3958e-01,  1.2080e-01,  2.7108e-01, -2.9164e-02,  1.9005e-01,
 -6.4674e-02, -1.9918e-01, -4.4422e-02, -2.1740e-01, -2.5346e-01,
 -2.0876e-02, -1.2492e-01, -6.5035e-02,  3.1687e-03, -1.8638e-01,
 -7.6838e-02, -1.6163e-01, -9.2983e-02,  6.0365e-01, -1.7455e-01,
 -1.6609e-01,  6.5964e-01, -7.4222e-02, -1.4284e-01, -1.4714e-01,
  5.2325e-01,  8.2848e-02, -2.6164e-01, -2.8923e-01,  4.6258e-01,
  6.9149e-02,  4.9200e-01, -3.5201e-02,  5.1567e-01,  6.3739e-01,
 -2.3107e-01, -1.5354e-01,  5.7315e-02, -1.8378e-01,  6.6064e-01,
 -7.3922e-02,  2.3261e-01,  7.1313e-01, -2.4291e-01, -5.9623e-02,
 -1.2865e-01,  5.0634e-02,  1.3037e-01, -1.1034e-02, -6.7288e-02,
 -2.2042e-01, -3.1583e-01, -4.6271e-02, -1.1279e-01, -2.1911e-01,
  7.7690e-02, -7.2953e-02, -2.8393e-01, -7.6894e-02, -2.5185e-01,
 -1.2925e-01, -1.1978e-01, -2.6490e-01, -5.4117e-02],
[-2.7434e-01,  1.3083e-02, -2.0358e-01,  4.5991e-01, -1.0085e-01,
  6.7857e-01, -1.1188e-01, -3.7486e-02, -1.6001e-01,  4.0718e-01,
 -1.9012e-01, -1.5308e-01,  4.5834e-01,  5.4207e-02,  4.7419e-01,
 -1.4098e-01,  5.6004e-01, -8.4456e-02, -9.2216e-02, -2.4794e-01,
 -3.2578e-02, -1.5617e-01, -1.0299e-01, -1.7890e-01, -2.0765e-01,
  5.1966e-01,  5.8648e-01,  6.6931e-01, -7.7624e-02, -2.0679e-02,
  2.3923e-01,  4.2634e-01, -2.2954e-01, -1.7547e-01, -1.3603e-01,
 -2.8514e-01, -9.5313e-02, -2.3686e-01, -5.4244e-02, -1.4029e-01,
 -3.0773e-01,  3.6503e-01, -9.6459e-02, -1.6615e-01, -1.9452e-01,
  6.2008e-02, -1.7196e-01, -2.6004e-01, -1.1403e-01, -9.4603e-02,
  4.8623e-01, -3.4601e-01, -9.6950e-02,  2.6046e-01, -1.8059e-01,
 -1.0403e-01, -1.0344e-01,  8.0230e-02,  6.8206e-02,  4.8589e-02,
 -1.7261e-01,  2.0396e-01, -3.3028e-01, -1.8138e-01],
[ 1.3735e-01,  6.2468e-01,  4.0201e-02, -1.0571e-01,  6.3137e-01,
 -5.9383e-02, -1.6052e-01, -2.0919e-02, -2.3541e-01, -2.6962e-01,
 -5.6668e-02, -2.2106e-01,  4.9913e-02,  4.2883e-02,  1.8694e-01,
 -5.0949e-02, -3.0345e-01,  5.3001e-01, -1.5064e-01,  1.0507e-01,
 -2.9251e-01, -2.4199e-01, -1.0166e-01, -2.3991e-01, -5.3122e-02,
 -2.0165e-01, -2.2767e-01, -1.8303e-01,  4.1401e-01, -3.3971e-02,
 -1.3842e-01, -2.2396e-01,  3.5261e-01, -2.4115e-01,  2.6574e-01,
  4.4883e-01, -1.8745e-01, -1.5625e-01, -8.3281e-02, -2.0169e-01,
 -3.0806e-01, -2.1903e-01, -8.6014e-02, -2.2135e-01,  5.0446e-02,
  9.4967e-04,  3.5808e-01,  3.3377e-01, -1.4715e-01,  1.1670e-01,
  1.1007e-01,  5.6337e-01, -1.0323e-01, -1.7378e-01,  4.7173e-01,
 -9.3719e-02, -1.3231e-01, -1.5468e-01, -2.9389e-02, -9.2038e-02,
  7.7110e-01,  2.5900e-02, -2.8625e-01, -2.4793e-01],
[ 6.9028e-01,  9.2523e-04, -5.6170e-03,  7.3784e-02, -1.9093e-01,
  1.5601e-01,  1.2299e-01, -5.9588e-02,  3.1943e-01,  7.0625e-01,
  9.8413e-02,  1.1377e-01,  2.4662e-01, -1.2369e-01, -3.3689e-01,
 -2.2724e-01, -2.4544e-02,  3.2761e-01, -2.4037e-01,  1.1161e-01,
 -1.0761e-01, -3.2827e-01, -6.7868e-02, -1.2270e-01, -2.9416e-02,
 -1.5166e-01, -2.7649e-01, -1.8252e-01, -2.6551e-01, -5.9102e-02,
 -1.0089e-01, -1.2666e-01, -1.7980e-01, -1.0485e-01, -9.7222e-02,
 -3.1077e-01,  4.4311e-01, -2.0994e-01, -1.7105e-01, -1.5427e-03,
 -1.0308e-01,  2.3852e-01,  6.3333e-02,  2.7013e-01, -3.2915e-01,
 -1.7497e-01, -6.7547e-02,  6.0398e-01, -3.2497e-02,  1.1121e-01,
 -2.8821e-01,  4.5953e-02, -2.5657e-01,  6.6359e-01,  9.2274e-02,
```

```
         -1.5235e-01,  5.8047e-01,  3.9200e-01,  1.9794e-02, -4.7674e-02,
         -1.7082e-01, -1.5583e-01,  4.8187e-01, -1.8513e-01],
        [-1.8795e-01, -1.7681e-01,  6.1738e-01,  1.2760e-02, -1.4384e-01,
         -1.0973e-01,  5.4950e-01, -1.1150e-01,  1.5479e-01, -1.4839e-01,
         -7.5977e-02, -1.8990e-01, -2.1045e-01, -3.1078e-02,  3.6460e-01,
          5.3661e-01, -2.0608e-01,  3.1622e-02, -1.8536e-01,  4.4768e-01,
         -1.1935e-01, -2.5000e-01, -1.0553e-01, -8.4709e-02,  5.5529e-01,
         -3.1415e-01, -2.0073e-01, -2.4423e-01, -2.6294e-01, -7.8744e-02,
         -2.3419e-01, -8.6368e-02, -2.1923e-01, -2.3655e-01,  4.1530e-01,
          1.5973e-01, -1.9790e-01,  4.4732e-01,  4.5359e-01,  4.8032e-01,
         -2.6682e-01, -2.3211e-01, -8.1600e-02,  4.5468e-01, -7.5257e-02,
         -1.7313e-01, -1.9884e-01, -1.7515e-01, -6.9835e-02,  3.0555e-02,
         -1.2991e-01, -3.2120e-01, -1.4624e-02, -1.6679e-01,  9.8486e-03,
         -1.0839e-01, -5.8157e-02, -9.9437e-02,  3.4003e-01, -1.7973e-01,
          9.3241e-02,  4.5762e-02,  4.6138e-01, -1.8001e-01],
        [-2.5169e-01, -3.6043e-02, -1.7598e-01, -1.9267e-01, -1.4164e-01,
         -6.7233e-02, -1.1834e-01,  4.4604e-03, -1.1303e-01, -2.3838e-01,
         -1.7445e-01, -1.5520e-01, -8.7288e-02, -2.1002e-01, -1.5655e-01,
         -1.8986e-01,  3.6689e-01, -2.3094e-01, -2.4329e-01, -2.9516e-01,
          6.6867e-01,  6.0952e-02,  3.6322e-02, -1.5354e-01, -1.2743e-01,
          5.1619e-01, -8.2559e-02, -2.2001e-01,  6.0981e-01, -4.1823e-04,
          4.0571e-01, -2.1905e-01, -1.4405e-01,  2.3641e-01, -2.4076e-01,
          4.6418e-01, -9.1992e-02,  4.1568e-01, -1.6088e-01, -2.5775e-02,
          6.5048e-01, -9.9760e-02, -1.2448e-01,  2.1074e-01, -1.6368e-01,
         -1.6222e-01, -1.5439e-01, -1.7828e-01, -6.6940e-02,  1.4840e-01,
         -2.1933e-01,  3.8606e-01,  5.7892e-02, -1.8897e-01, -1.8289e-01,
          1.2529e-01,  3.0542e-01, -2.8817e-01,  1.4114e-01, -2.1229e-01,
         -1.4997e-01, -1.5062e-01, -2.5203e-01,  6.1738e-01],
        [-2.9794e-01, -8.0818e-02, -2.0599e-01,  6.6177e-01,  3.4091e-02,
         -1.4906e-01, -1.5009e-01,  5.8519e-01, -2.5255e-01,  1.4975e-01,
          7.0765e-01,  5.0860e-01,  3.5802e-01, -5.6471e-02, -6.4993e-02,
         -9.8368e-03, -1.1208e-01, -1.5518e-01,  6.0399e-01,  5.0071e-01,
         -2.7540e-01, -2.5394e-01,  4.3116e-01,  6.7467e-02, -1.2505e-01,
         -2.1849e-01, -1.6951e-01, -1.0391e-01, -2.4049e-01,  1.6538e-01,
         -4.3109e-02, -2.2503e-01, -1.1801e-01,  2.6394e-01, -2.8796e-01,
         -2.6867e-01,  1.5537e-01,  4.9630e-01, -6.9490e-02, -1.4031e-01,
         -2.9698e-01, -3.2645e-01, -1.7759e-02, -1.2855e-01,  3.7534e-01,
         -2.2055e-02, -1.6288e-01, -2.0429e-02, -1.6032e-02, -1.6787e-01,
         -5.3949e-02, -2.4362e-01, -5.9726e-02, -1.8437e-01, -6.0740e-02,
          4.0212e-01, -3.1784e-01,  3.7678e-01, -7.6779e-02,  3.9906e-01,
         -4.6646e-04,  1.2145e-01, -1.3220e-01,  2.4004e-02],
        [ 4.3218e-01, -2.2053e-01, -1.4727e-01, -2.7430e-01, -3.0975e-01,
         -1.7404e-01, -1.8606e-01, -1.4458e-01, -7.4440e-02, -2.5513e-01,
          3.5979e-01, -1.2656e-01, -7.6645e-02,  3.4274e-01,  3.6340e-01,
          9.9818e-02, -1.0891e-01, -3.0954e-01, -2.1559e-01, -2.6891e-01,
         -6.9285e-02, -2.3029e-01, -2.1416e-01,  6.3743e-01, -8.3363e-02,
         -2.8156e-01, -2.0645e-01,  6.4889e-01,  3.2541e-01, -2.6820e-03,
         -2.6029e-01,  1.6260e-01,  3.7760e-01, -3.0107e-01, -2.5582e-01,
          1.6983e-01,  1.6201e-01, -1.7698e-01,  1.4060e-01, -1.4303e-01,
         -1.4093e-01, -2.5366e-01,  5.8451e-02, -9.9402e-02,  5.6660e-01,
         -1.1234e-01,  3.3843e-01, -2.3948e-01,  9.5276e-02, -1.4596e-01,
         -1.1002e-01,  5.6948e-01, -7.2906e-02, -1.0169e-01, -2.5810e-01,
          3.9855e-02, -3.0532e-01,  6.4734e-01,  4.5967e-02, -8.6700e-02,
         -3.2079e-01,  1.2647e-01, -1.6112e-01,  5.8503e-01]])), ('out.bias', tensor([
0.0902,  0.0709,  0.0331, -0.0473, -0.1082, -0.0926, -0.0551, -0.0961,
        -0.1117,  0.1010])), ('layer_norm1.weight', tensor([0.9790, 1.0460, 0.8685, 1.04
83, 1.0223, 1.0333, 1.0588, 0.9583, 0.9854,
        1.1330, 1.0208, 1.0124, 1.0041, 0.8098, 1.0678, 0.6619, 1.0815, 0.9554,
        0.9711, 0.9457, 1.0378, 0.9609, 0.9586, 0.9747, 0.9950, 0.8134, 0.9142,
        0.8460, 0.9583, 1.0811, 1.0459, 0.8242, 1.0428, 0.9377, 0.9645, 0.9640,
        0.9543, 0.9857, 1.0033, 0.9217, 0.9580, 0.9905, 1.0870, 0.9657, 0.7445,
        0.8782, 0.9660, 0.9856, 1.0049, 1.0452, 0.8781, 0.9911, 0.7509, 0.9408,
        0.9892, 1.0019, 1.0210, 0.9682, 1.0064, 0.9173, 1.0754, 1.0414, 0.9680,
        0.9316, 1.0209, 1.0508, 1.0123, 0.8057, 0.9585, 1.0518, 0.9498, 0.9558,
```

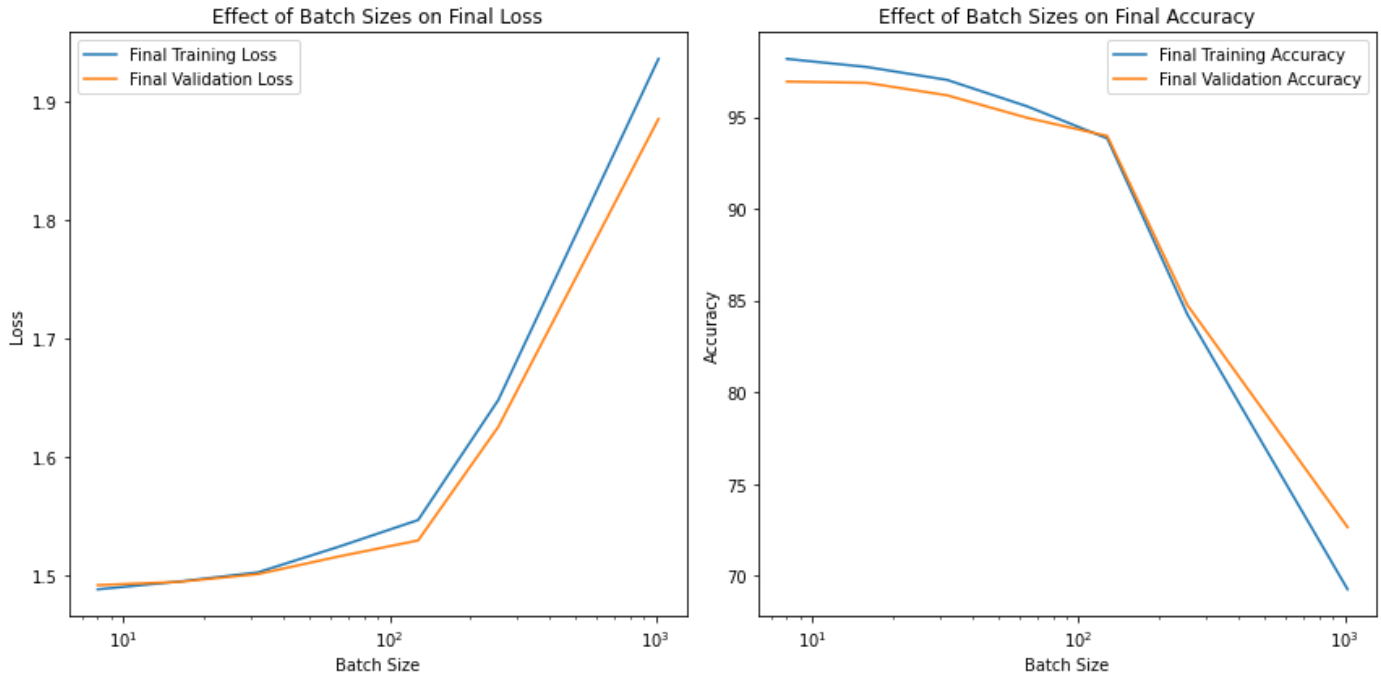
```
        0.9455, 1.0588, 1.1041, 0.8943, 0.9264, 1.0292, 0.9519, 1.0250, 0.9125,
        0.9425, 1.0040, 0.9304, 1.0254, 0.9876, 0.9864, 0.9486, 0.9680, 1.0060,
        0.9971, 1.0737, 0.9431, 0.9835, 1.0355, 1.0395, 0.9264, 1.0013, 1.0775,
        1.0349, 1.0270, 0.9944, 0.9696, 1.0256, 1.0282, 0.9696, 0.9692, 1.1176,
        0.7973, 0.9610, 1.0660, 1.0555, 1.0878, 1.0377, 0.8417, 0.9809, 0.9892,
        1.0390, 0.9513, 0.9703, 1.0911, 0.9073, 0.9346, 0.9757, 1.0364, 0.9452,
        1.0861, 1.0548])), ('layer_norm1.bias', tensor([-0.1162, -0.1484, -0.1676, -0.15
09, -0.1054, -0.1349, -0.0926, -0.1222,
        -0.1051, -0.0842, -0.1001, -0.1329, -0.0979, -0.1644, -0.0897, -0.2170,
        -0.1167, -0.1616, -0.1898, -0.1174, -0.1178, -0.1425, -0.1736, -0.1106,
        -0.1240, -0.2105, -0.1833, -0.1848, -0.1809, -0.1207, -0.1269, -0.1711,
        -0.0921, -0.1342, -0.1513, -0.1486, -0.1271, -0.1424, -0.0674, -0.1433,
        -0.1750, -0.1298, -0.1009, -0.1542, -0.1815, -0.2013, -0.1212, -0.1265,
        -0.1281, -0.1013, -0.1554, -0.1015, -0.1733, -0.1265, -0.1538, -0.1224,
        -0.0867, -0.1677, -0.0804, -0.1461, -0.1100, -0.1208, -0.1192, -0.1995,
        -0.1143, -0.1165, -0.0979, -0.1784, -0.2252, -0.0667, -0.1909, -0.1376,
        -0.1142, -0.1051, -0.0635, -0.1753, -0.2025, -0.0977, -0.2042, -0.1273,
        -0.1606, -0.1204, -0.1342, -0.1339, -0.1335, -0.1367, -0.1468, -0.1491,
        -0.0993, -0.1410, -0.1195, -0.0895, -0.1589, -0.1266, -0.1317, -0.1076,
        -0.1686, -0.1106, -0.0917, -0.1217, -0.1029, -0.1484, -0.1058, -0.0880,
        -0.1014, -0.1441, -0.1340, -0.0801, -0.1763, -0.1329, -0.0937, -0.1141,
        -0.0895, -0.0846, -0.1833, -0.1350, -0.1313, -0.1061, -0.1816, -0.1274,
        -0.1156, -0.1336, -0.1928, -0.1101, -0.0923, -0.1898, -0.0859, -0.1131])), ('lay
er_norm2.weight', tensor([1.4073, 1.1976, 1.2518, 1.3674, 1.2509, 1.2554, 1.1809, 1.162
6, 1.1940,
        1.4280, 1.2926, 1.3357, 1.1993, 1.0710, 1.3131, 1.2600, 1.2618, 1.2544,
        1.3800, 1.3392, 1.3607, 1.4594, 1.1532, 1.2410, 1.3482, 1.4532, 1.3473,
        1.4575, 1.3703, 1.0946, 1.2477, 1.2733, 1.1827, 1.2927, 1.3590, 1.3448,
        1.1393, 1.3575, 1.1565, 1.3183, 1.4723, 1.2616, 1.2309, 1.2250, 1.2999,
        1.2594, 1.2045, 1.2758, 1.0937, 1.0355, 1.2954, 1.4922, 1.1646, 1.3079,
        1.2445, 1.0908, 1.3293, 1.3883, 1.0484, 1.2742, 1.3185, 1.0660, 1.4102,
        1.3686])), ('layer_norm2.bias', tensor([ 0.0625,  0.0206,  0.0567,  0.0796,  0.0
338,  0.0169,  0.0282,  0.0093,
         0.1022,  0.1054,  0.0649,  0.0699,  0.1056,  0.0410,  0.1305,  0.0804,
         0.0682,  0.0896,  0.0741,  0.1509,  0.0554,  0.0645,  0.0762,  0.0104,
         0.0469,  0.1093,  0.0633,  0.0235,  0.0923,  0.0420,  0.1224,  0.0935,
         0.0842,  0.1169,  0.0950,  0.1127,  0.0791,  0.1366,  0.0631,  0.0598,
         0.0931,  0.1642,  0.0126,  0.0931,  0.1063,  0.0130,  0.0969,  0.0885,
         0.0431,  0.0602,  0.0706,  0.1106,  0.0347,  0.0508,  0.0929,  0.0488,
         0.1042,  0.1073,  0.0469,  0.0973, -0.0033,  0.0702,  0.1542,  0.0575]))])
```

Plot Effect of changing batch size on the accuracy and loss

In [35]:

```python
# Plot the effect of changing batch sizes on final accuracy and loss
plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)
plt.plot(batch_sizes, final_train_losses, label='Final Training Loss')
plt.plot(batch_sizes, final_val_losses, label='Final Validation Loss')
plt.title('Effect of Batch Sizes on Final Loss')
plt.xlabel('Batch Size')
plt.ylabel('Loss')
plt.xscale('log')  # Use log scale for better visualization
plt.legend()

plt.subplot(1, 2, 2)
plt.plot(batch_sizes, final_train_accuracies, label='Final Training Accuracy')
plt.plot(batch_sizes, final_val_accuracies, label='Final Validation Accuracy')
plt.title('Effect of Batch Sizes on Final Accuracy')
plt.xlabel('Batch Size')
plt.ylabel('Accuracy')
plt.xscale('log')  # Use log scale for better visualization
```

```
plt.tight_layout()
plt.show()
```



get best model (choose best model according to learning rate and batch size together)

In [37]:
```
from itertools import product

# Define the hyperparameter grid
param_grid = {
    'lr': [0.001, 0.01, 0.03, 0.05, 0.1, 0.5],
    'batch_size': [8, 16, 32, 64, 128],
}

# Initialize variables to keep track of the best model
best_model = None
best_val_accuracy = 0.0
best_hyperparameters = None

# Iterate over all combinations of hyperparameters
for params in product(*param_grid.values()):
    hyperparameters = dict(zip(param_grid.keys(), params))

    # Create data loaders with the specified batch size
    train_loader = DataLoader(train_set, batch_size=hyperparameters['batch_size'], shuff
    val_loader = DataLoader(val_set, batch_size=hyperparameters['batch_size'], shuffle=F

    # Instantiate the model
    model = model = Model(in_features, h1, h2, out_features)

    # Set a new learning rate
    model.optimizer = optim.SGD(model.parameters(), lr=hyperparameters['lr'], weight_dec

    # Lists to store metrics for plotting
    train_loss_history = []
    val_loss_history = []
    train_acc_history = []
    val_acc_history = []
```

```python
    # Training loop
    for epoch in range(num_epochs):
        # Training
        average_loss, train_accuracy = train(model, train_loader)
        train_loss_history.append(average_loss)
        train_acc_history.append(train_accuracy)

        # Validation
        average_val_loss, val_accuracy = validate(model, val_loader)
        val_loss_history.append(average_val_loss)
        val_acc_history.append(val_accuracy)

        # Print and visualize metrics
        print(f'Epoch [{epoch + 1}/{num_epochs}], '
              f'Training Loss: {average_loss:.4f}, Training Accuracy: {train_accuracy:.4
              f'Validation Loss: {average_val_loss:.4f}, Validation Accuracy: {val_accur

    # Check if current model has the best validation accuracy
    if val_accuracy > best_val_accuracy:
        best_val_accuracy = val_accuracy
        best_model = model.state_dict().copy()
        best_hyperparameters = hyperparameters
```

```
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 460.69it/s]
Epoch [1/20], Training Loss: 2.1794, Training Accuracy: 45.6726%, Validation Loss: 1.992
7, Validation Accuracy: 68.7143%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 446.62it/s]
Epoch [2/20], Training Loss: 1.9209, Training Accuracy: 69.9464%, Validation Loss: 1.797
6, Validation Accuracy: 74.4286%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 489.46it/s]
Epoch [3/20], Training Loss: 1.7962, Training Accuracy: 74.4613%, Validation Loss: 1.718
8, Validation Accuracy: 79.3452%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 485.66it/s]
Epoch [4/20], Training Loss: 1.7172, Training Accuracy: 83.2321%, Validation Loss: 1.638
5, Validation Accuracy: 88.0476%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.21it/s]
Epoch [5/20], Training Loss: 1.6545, Training Accuracy: 88.3423%, Validation Loss: 1.595
0, Validation Accuracy: 90.5357%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 485.13it/s]
Epoch [6/20], Training Loss: 1.6191, Training Accuracy: 90.1042%, Validation Loss: 1.575
1, Validation Accuracy: 91.2381%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 490.77it/s]
Epoch [7/20], Training Loss: 1.6005, Training Accuracy: 91.2024%, Validation Loss: 1.563
0, Validation Accuracy: 91.7500%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 489.85it/s]
Epoch [8/20], Training Loss: 1.5845, Training Accuracy: 91.8810%, Validation Loss: 1.554
1, Validation Accuracy: 92.2500%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 485.92it/s]
Epoch [9/20], Training Loss: 1.5752, Training Accuracy: 92.2976%, Validation Loss: 1.547
3, Validation Accuracy: 92.8333%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 481.24it/s]
```

```
Epoch [10/20], Training Loss: 1.5654, Training Accuracy: 92.8423%, Validation Loss: 1.54
20, Validation Accuracy: 93.1310%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 473.76it/s]
Epoch [11/20], Training Loss: 1.5596, Training Accuracy: 93.1518%, Validation Loss: 1.53
84, Validation Accuracy: 93.3333%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 490.11it/s]
Epoch [12/20], Training Loss: 1.5546, Training Accuracy: 93.6071%, Validation Loss: 1.53
43, Validation Accuracy: 93.6310%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 493.22it/s]
Epoch [13/20], Training Loss: 1.5496, Training Accuracy: 93.8274%, Validation Loss: 1.53
19, Validation Accuracy: 93.7619%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.33it/s]
Epoch [14/20], Training Loss: 1.5455, Training Accuracy: 94.1726%, Validation Loss: 1.52
92, Validation Accuracy: 94.0119%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 491.66it/s]
Epoch [15/20], Training Loss: 1.5411, Training Accuracy: 94.3839%, Validation Loss: 1.52
70, Validation Accuracy: 94.0952%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 494.04it/s]
Epoch [16/20], Training Loss: 1.5387, Training Accuracy: 94.6815%, Validation Loss: 1.52
51, Validation Accuracy: 94.2262%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 490.50it/s]
Epoch [17/20], Training Loss: 1.5351, Training Accuracy: 94.7024%, Validation Loss: 1.52
22, Validation Accuracy: 94.4881%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 490.48it/s]
Epoch [18/20], Training Loss: 1.5329, Training Accuracy: 94.9256%, Validation Loss: 1.52
12, Validation Accuracy: 94.5714%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 491.05it/s]
Epoch [19/20], Training Loss: 1.5290, Training Accuracy: 95.1042%, Validation Loss: 1.51
98, Validation Accuracy: 94.5952%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.92it/s]
Epoch [20/20], Training Loss: 1.5275, Training Accuracy: 95.1994%, Validation Loss: 1.51
79, Validation Accuracy: 94.7143%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 469.34it/s]
Epoch [1/20], Training Loss: 2.2556, Training Accuracy: 29.5804%, Validation Loss: 2.172
6, Validation Accuracy: 56.3929%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 455.20it/s]
Epoch [2/20], Training Loss: 2.1172, Training Accuracy: 59.0804%, Validation Loss: 1.997
7, Validation Accuracy: 67.2738%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 406.78it/s]
Epoch [3/20], Training Loss: 1.9709, Training Accuracy: 68.2530%, Validation Loss: 1.866
0, Validation Accuracy: 73.5476%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 468.08it/s]
Epoch [4/20], Training Loss: 1.8675, Training Accuracy: 74.5714%, Validation Loss: 1.779
8, Validation Accuracy: 79.7381%
Training: 100%|██████████████████████████████████████████████████████████| 210
<00:00, 463.83it/s]
```

```
Epoch [5/20], Training Loss: 1.7956, Training Accuracy: 79.0863%, Validation Loss: 1.720
4, Validation Accuracy: 82.2143%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 465.70it/s]
Epoch [6/20], Training Loss: 1.7451, Training Accuracy: 81.5119%, Validation Loss: 1.683
2, Validation Accuracy: 83.6429%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 464.93it/s]
Epoch [7/20], Training Loss: 1.7135, Training Accuracy: 83.1548%, Validation Loss: 1.656
8, Validation Accuracy: 85.3929%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 470.57it/s]
Epoch [8/20], Training Loss: 1.6847, Training Accuracy: 85.3244%, Validation Loss: 1.630
3, Validation Accuracy: 88.2619%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 465.98it/s]
Epoch [9/20], Training Loss: 1.6581, Training Accuracy: 87.8482%, Validation Loss: 1.607
5, Validation Accuracy: 90.0357%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 466.23it/s]
Epoch [10/20], Training Loss: 1.6381, Training Accuracy: 89.2411%, Validation Loss: 1.59
18, Validation Accuracy: 90.9048%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 467.92it/s]
Epoch [11/20], Training Loss: 1.6220, Training Accuracy: 89.9494%, Validation Loss: 1.58
13, Validation Accuracy: 91.2500%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 467.43it/s]
Epoch [12/20], Training Loss: 1.6089, Training Accuracy: 90.6071%, Validation Loss: 1.57
29, Validation Accuracy: 91.6071%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 471.54it/s]
Epoch [13/20], Training Loss: 1.6014, Training Accuracy: 90.8304%, Validation Loss: 1.56
66, Validation Accuracy: 91.8214%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 469.11it/s]
Epoch [14/20], Training Loss: 1.5929, Training Accuracy: 91.3244%, Validation Loss: 1.56
09, Validation Accuracy: 91.9524%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 467.76it/s]
Epoch [15/20], Training Loss: 1.5859, Training Accuracy: 91.6548%, Validation Loss: 1.55
65, Validation Accuracy: 92.2024%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 469.83it/s]
Epoch [16/20], Training Loss: 1.5797, Training Accuracy: 92.0208%, Validation Loss: 1.55
27, Validation Accuracy: 92.4643%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 469.56it/s]
Epoch [17/20], Training Loss: 1.5747, Training Accuracy: 92.2173%, Validation Loss: 1.54
95, Validation Accuracy: 92.5714%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 466.46it/s]
Epoch [18/20], Training Loss: 1.5702, Training Accuracy: 92.5298%, Validation Loss: 1.54
67, Validation Accuracy: 92.7738%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 471.07it/s]
Epoch [19/20], Training Loss: 1.5669, Training Accuracy: 92.6399%, Validation Loss: 1.54
40, Validation Accuracy: 93.0000%
Training: 100%|██████████████████████████████████████████████████████████| 210
<00:00, 467.78it/s]
```

```
Epoch [20/20], Training Loss: 1.5621, Training Accuracy: 92.9048%, Validation Loss: 1.54
15, Validation Accuracy: 93.1429%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 451.59it/s]
Epoch [1/20], Training Loss: 2.2811, Training Accuracy: 19.7113%, Validation Loss: 2.244
0, Validation Accuracy: 32.6905%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.16it/s]
Epoch [2/20], Training Loss: 2.2174, Training Accuracy: 36.7113%, Validation Loss: 2.161
3, Validation Accuracy: 49.0238%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.15it/s]
Epoch [3/20], Training Loss: 2.1471, Training Accuracy: 50.8244%, Validation Loss: 2.083
6, Validation Accuracy: 61.6071%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.60it/s]
Epoch [4/20], Training Loss: 2.0805, Training Accuracy: 58.9435%, Validation Loss: 2.013
2, Validation Accuracy: 63.8929%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 417.97it/s]
Epoch [5/20], Training Loss: 2.0203, Training Accuracy: 62.2143%, Validation Loss: 1.954
0, Validation Accuracy: 65.5952%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 408.19it/s]
Epoch [6/20], Training Loss: 1.9672, Training Accuracy: 65.4702%, Validation Loss: 1.903
6, Validation Accuracy: 68.5595%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 404.29it/s]
Epoch [7/20], Training Loss: 1.9222, Training Accuracy: 69.0536%, Validation Loss: 1.858
7, Validation Accuracy: 73.1071%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.06it/s]
Epoch [8/20], Training Loss: 1.8807, Training Accuracy: 73.4196%, Validation Loss: 1.814
6, Validation Accuracy: 78.8929%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.06it/s]
Epoch [9/20], Training Loss: 1.8388, Training Accuracy: 78.6935%, Validation Loss: 1.769
4, Validation Accuracy: 83.7738%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 438.56it/s]
Epoch [10/20], Training Loss: 1.7963, Training Accuracy: 82.9792%, Validation Loss: 1.72
70, Validation Accuracy: 86.2619%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.06it/s]
Epoch [11/20], Training Loss: 1.7575, Training Accuracy: 84.9286%, Validation Loss: 1.69
23, Validation Accuracy: 87.5119%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.45it/s]
Epoch [12/20], Training Loss: 1.7280, Training Accuracy: 86.3065%, Validation Loss: 1.66
64, Validation Accuracy: 88.5476%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 451.47it/s]
Epoch [13/20], Training Loss: 1.7041, Training Accuracy: 87.2738%, Validation Loss: 1.64
61, Validation Accuracy: 89.0595%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 440.99it/s]
Epoch [14/20], Training Loss: 1.6841, Training Accuracy: 87.7262%, Validation Loss: 1.63
02, Validation Accuracy: 89.7262%
Training: 100%|██████████████████████████████████████████████████████████| 105
<00:00, 446.31it/s]
```

```
Epoch [15/20], Training Loss: 1.6680, Training Accuracy: 88.3423%, Validation Loss: 1.61
76, Validation Accuracy: 90.2024%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.68it/s]
Epoch [16/20], Training Loss: 1.6541, Training Accuracy: 88.9702%, Validation Loss: 1.60
76, Validation Accuracy: 90.3333%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.50it/s]
Epoch [17/20], Training Loss: 1.6444, Training Accuracy: 89.4107%, Validation Loss: 1.59
88, Validation Accuracy: 90.6548%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.06it/s]
Epoch [18/20], Training Loss: 1.6352, Training Accuracy: 89.6339%, Validation Loss: 1.59
20, Validation Accuracy: 90.8690%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.33it/s]
Epoch [19/20], Training Loss: 1.6255, Training Accuracy: 89.7887%, Validation Loss: 1.58
61, Validation Accuracy: 91.1071%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 451.46it/s]
Epoch [20/20], Training Loss: 1.6191, Training Accuracy: 90.0952%, Validation Loss: 1.58
10, Validation Accuracy: 91.2738%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 370.38it/s]
Epoch [1/20], Training Loss: 2.2991, Training Accuracy: 12.5060%, Validation Loss: 2.285
9, Validation Accuracy: 23.9405%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 371.69it/s]
Epoch [2/20], Training Loss: 2.2751, Training Accuracy: 24.3810%, Validation Loss: 2.252
9, Validation Accuracy: 29.5238%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 360.05it/s]
Epoch [3/20], Training Loss: 2.2453, Training Accuracy: 30.1071%, Validation Loss: 2.216
9, Validation Accuracy: 34.3810%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 368.84it/s]
Epoch [4/20], Training Loss: 2.2123, Training Accuracy: 35.2679%, Validation Loss: 2.179
8, Validation Accuracy: 40.5833%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 362.67it/s]
Epoch [5/20], Training Loss: 2.1806, Training Accuracy: 40.4196%, Validation Loss: 2.143
9, Validation Accuracy: 46.2619%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 371.19it/s]
Epoch [6/20], Training Loss: 2.1491, Training Accuracy: 45.0268%, Validation Loss: 2.109
9, Validation Accuracy: 50.5952%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 366.21it/s]
Epoch [7/20], Training Loss: 2.1192, Training Accuracy: 48.6845%, Validation Loss: 2.077
8, Validation Accuracy: 54.6429%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 370.51it/s]
Epoch [8/20], Training Loss: 2.0906, Training Accuracy: 52.4792%, Validation Loss: 2.047
4, Validation Accuracy: 57.6548%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 360.24it/s]
Epoch [9/20], Training Loss: 2.0627, Training Accuracy: 55.4702%, Validation Loss: 2.018
0, Validation Accuracy: 60.5595%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
00:00, 367.28it/s]
```

Loading [MathJax]/extensions/Safe.js

```
Epoch [10/20], Training Loss: 2.0360, Training Accuracy: 58.4375%, Validation Loss: 1.98
91, Validation Accuracy: 63.6190%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 369.34it/s]
Epoch [11/20], Training Loss: 2.0091, Training Accuracy: 61.3542%, Validation Loss: 1.96
05, Validation Accuracy: 66.0000%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 364.11it/s]
Epoch [12/20], Training Loss: 1.9822, Training Accuracy: 63.6310%, Validation Loss: 1.93
25, Validation Accuracy: 68.3452%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 371.15it/s]
Epoch [13/20], Training Loss: 1.9564, Training Accuracy: 65.9405%, Validation Loss: 1.90
56, Validation Accuracy: 70.5476%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 371.60it/s]
Epoch [14/20], Training Loss: 1.9312, Training Accuracy: 68.7232%, Validation Loss: 1.87
98, Validation Accuracy: 73.1429%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 363.98it/s]
Epoch [15/20], Training Loss: 1.9088, Training Accuracy: 71.0923%, Validation Loss: 1.85
49, Validation Accuracy: 75.3690%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 277.38it/s]
Epoch [16/20], Training Loss: 1.8860, Training Accuracy: 73.3810%, Validation Loss: 1.83
09, Validation Accuracy: 77.5238%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 318.78it/s]
Epoch [17/20], Training Loss: 1.8633, Training Accuracy: 75.4107%, Validation Loss: 1.80
80, Validation Accuracy: 79.2143%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 340.32it/s]
Epoch [18/20], Training Loss: 1.8432, Training Accuracy: 77.0000%, Validation Loss: 1.78
68, Validation Accuracy: 80.4286%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 336.37it/s]
Epoch [19/20], Training Loss: 1.8244, Training Accuracy: 78.2173%, Validation Loss: 1.76
78, Validation Accuracy: 81.0119%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 362.10it/s]
Epoch [20/20], Training Loss: 1.8063, Training Accuracy: 78.8661%, Validation Loss: 1.75
08, Validation Accuracy: 81.4405%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 276.66it/s]
Epoch [1/20], Training Loss: 2.2981, Training Accuracy: 13.1935%, Validation Loss: 2.291
6, Validation Accuracy: 19.4881%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 256.05it/s]
Epoch [2/20], Training Loss: 2.2875, Training Accuracy: 18.9524%, Validation Loss: 2.278
0, Validation Accuracy: 25.7857%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 251.21it/s]
Epoch [3/20], Training Loss: 2.2750, Training Accuracy: 25.4256%, Validation Loss: 2.262
0, Validation Accuracy: 32.5476%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 271.25it/s]
Epoch [4/20], Training Loss: 2.2605, Training Accuracy: 30.3423%, Validation Loss: 2.243
8, Validation Accuracy: 37.8095%
Training: 100%|██████████████████████████████████████████████████████████| 26
00:00, 282.46it/s]
```

```
Epoch [5/20], Training Loss: 2.2440, Training Accuracy: 34.1726%, Validation Loss: 2.224
0, Validation Accuracy: 41.0476%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 253.13it/s]
Epoch [6/20], Training Loss: 2.2270, Training Accuracy: 37.1607%, Validation Loss: 2.203
6, Validation Accuracy: 43.0714%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 278.56it/s]
Epoch [7/20], Training Loss: 2.2093, Training Accuracy: 40.1250%, Validation Loss: 2.183
3, Validation Accuracy: 44.8452%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 234.11it/s]
Epoch [8/20], Training Loss: 2.1918, Training Accuracy: 42.4583%, Validation Loss: 2.163
2, Validation Accuracy: 46.6786%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 230.77it/s]
Epoch [9/20], Training Loss: 2.1741, Training Accuracy: 44.6012%, Validation Loss: 2.143
5, Validation Accuracy: 48.8095%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 272.09it/s]
Epoch [10/20], Training Loss: 2.1558, Training Accuracy: 46.8482%, Validation Loss: 2.12
42, Validation Accuracy: 51.2262%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.78it/s]
Epoch [11/20], Training Loss: 2.1394, Training Accuracy: 48.9345%, Validation Loss: 2.10
57, Validation Accuracy: 53.8095%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 278.26it/s]
Epoch [12/20], Training Loss: 2.1230, Training Accuracy: 51.4077%, Validation Loss: 2.08
74, Validation Accuracy: 56.5238%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 279.80it/s]
Epoch [13/20], Training Loss: 2.1077, Training Accuracy: 53.6815%, Validation Loss: 2.06
95, Validation Accuracy: 58.9524%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.03it/s]
Epoch [14/20], Training Loss: 2.0912, Training Accuracy: 56.1071%, Validation Loss: 2.05
16, Validation Accuracy: 61.5833%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 281.37it/s]
Epoch [15/20], Training Loss: 2.0742, Training Accuracy: 58.5625%, Validation Loss: 2.03
37, Validation Accuracy: 63.8690%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 270.84it/s]
Epoch [16/20], Training Loss: 2.0578, Training Accuracy: 61.1310%, Validation Loss: 2.01
55, Validation Accuracy: 66.6905%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 281.26it/s]
Epoch [17/20], Training Loss: 2.0411, Training Accuracy: 63.5893%, Validation Loss: 1.99
70, Validation Accuracy: 68.9405%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.06it/s]
Epoch [18/20], Training Loss: 2.0236, Training Accuracy: 65.9464%, Validation Loss: 1.97
82, Validation Accuracy: 71.2024%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 281.06it/s]
Epoch [19/20], Training Loss: 2.0074, Training Accuracy: 68.1101%, Validation Loss: 1.95
92, Validation Accuracy: 73.1786%
Training: 100%|██████████████████████████████████████████████████████████| 26
00:00, 280.62it/s]
```

```
Epoch [20/20], Training Loss: 1.9904, Training Accuracy: 69.9881%, Validation Loss: 1.94
03, Validation Accuracy: 74.5714%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.45it/s]
Epoch [1/20], Training Loss: 1.7552, Training Accuracy: 80.9137%, Validation Loss: 1.565
2, Validation Accuracy: 91.2738%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 474.52it/s]
Epoch [2/20], Training Loss: 1.5684, Training Accuracy: 93.9286%, Validation Loss: 1.535
8, Validation Accuracy: 93.0000%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.42it/s]
Epoch [3/20], Training Loss: 1.5423, Training Accuracy: 95.1161%, Validation Loss: 1.522
1, Validation Accuracy: 94.4167%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 485.64it/s]
Epoch [4/20], Training Loss: 1.5304, Training Accuracy: 95.7232%, Validation Loss: 1.516
5, Validation Accuracy: 94.7976%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.56it/s]
Epoch [5/20], Training Loss: 1.5207, Training Accuracy: 96.1935%, Validation Loss: 1.512
8, Validation Accuracy: 95.1667%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 485.60it/s]
Epoch [6/20], Training Loss: 1.5163, Training Accuracy: 96.4524%, Validation Loss: 1.509
7, Validation Accuracy: 95.4048%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.35it/s]
Epoch [7/20], Training Loss: 1.5125, Training Accuracy: 96.6250%, Validation Loss: 1.510
3, Validation Accuracy: 95.3571%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 478.53it/s]
Epoch [8/20], Training Loss: 1.5086, Training Accuracy: 96.7917%, Validation Loss: 1.502
5, Validation Accuracy: 95.9881%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 482.67it/s]
Epoch [9/20], Training Loss: 1.5060, Training Accuracy: 97.0417%, Validation Loss: 1.502
2, Validation Accuracy: 96.0476%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 490.42it/s]
Epoch [10/20], Training Loss: 1.5025, Training Accuracy: 97.1548%, Validation Loss: 1.50
15, Validation Accuracy: 96.0833%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 485.12it/s]
Epoch [11/20], Training Loss: 1.5009, Training Accuracy: 97.2768%, Validation Loss: 1.49
94, Validation Accuracy: 96.2857%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 471.21it/s]
Epoch [12/20], Training Loss: 1.5005, Training Accuracy: 97.3631%, Validation Loss: 1.49
74, Validation Accuracy: 96.5000%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 468.21it/s]
Epoch [13/20], Training Loss: 1.4973, Training Accuracy: 97.4613%, Validation Loss: 1.49
57, Validation Accuracy: 96.6667%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 463.88it/s]
Epoch [14/20], Training Loss: 1.4972, Training Accuracy: 97.6637%, Validation Loss: 1.49
43, Validation Accuracy: 96.8810%
Training: 100%|██████████████████████████████████████████████████████████| 420
<00:00, 478.77it/s]
```

```
Epoch [15/20], Training Loss: 1.4936, Training Accuracy: 97.7202%, Validation Loss: 1.49
31, Validation Accuracy: 97.0238%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 464.21it/s]
Epoch [16/20], Training Loss: 1.4934, Training Accuracy: 97.8452%, Validation Loss: 1.49
52, Validation Accuracy: 96.8214%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 441.51it/s]
Epoch [17/20], Training Loss: 1.4927, Training Accuracy: 97.9524%, Validation Loss: 1.49
27, Validation Accuracy: 96.9524%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 445.00it/s]
Epoch [18/20], Training Loss: 1.4911, Training Accuracy: 97.9107%, Validation Loss: 1.49
31, Validation Accuracy: 96.9167%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 449.39it/s]
Epoch [19/20], Training Loss: 1.4896, Training Accuracy: 97.9315%, Validation Loss: 1.49
30, Validation Accuracy: 96.9167%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 469.37it/s]
Epoch [20/20], Training Loss: 1.4902, Training Accuracy: 98.0565%, Validation Loss: 1.49
18, Validation Accuracy: 97.0595%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 447.73it/s]
Epoch [1/20], Training Loss: 1.8945, Training Accuracy: 67.7708%, Validation Loss: 1.698
2, Validation Accuracy: 78.9881%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 452.01it/s]
Epoch [2/20], Training Loss: 1.6736, Training Accuracy: 83.5565%, Validation Loss: 1.627
6, Validation Accuracy: 84.2262%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 468.72it/s]
Epoch [3/20], Training Loss: 1.6368, Training Accuracy: 85.4435%, Validation Loss: 1.614
1, Validation Accuracy: 85.1905%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 446.99it/s]
Epoch [4/20], Training Loss: 1.6229, Training Accuracy: 86.0506%, Validation Loss: 1.607
1, Validation Accuracy: 85.6310%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 465.90it/s]
Epoch [5/20], Training Loss: 1.6159, Training Accuracy: 86.4821%, Validation Loss: 1.605
4, Validation Accuracy: 85.7857%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 443.13it/s]
Epoch [6/20], Training Loss: 1.6090, Training Accuracy: 86.9464%, Validation Loss: 1.598
1, Validation Accuracy: 86.3690%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 371.86it/s]
Epoch [7/20], Training Loss: 1.6056, Training Accuracy: 87.1101%, Validation Loss: 1.596
1, Validation Accuracy: 86.5238%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 398.60it/s]
Epoch [8/20], Training Loss: 1.6008, Training Accuracy: 87.4018%, Validation Loss: 1.594
9, Validation Accuracy: 86.6548%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 409.96it/s]
Epoch [9/20], Training Loss: 1.5979, Training Accuracy: 87.4881%, Validation Loss: 1.590
8, Validation Accuracy: 86.7619%
Training: 100%|██████████████████████████████████████████████████████████| 210
<00:00, 414.09it/s]
```

```
Epoch [10/20], Training Loss: 1.5473, Training Accuracy: 94.1310%, Validation Loss: 1.51
54, Validation Accuracy: 95.0952%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 412.46it/s]
Epoch [11/20], Training Loss: 1.5217, Training Accuracy: 96.2946%, Validation Loss: 1.50
86, Validation Accuracy: 95.7143%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 406.63it/s]
Epoch [12/20], Training Loss: 1.5147, Training Accuracy: 96.5357%, Validation Loss: 1.50
65, Validation Accuracy: 95.8452%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 405.47it/s]
Epoch [13/20], Training Loss: 1.5118, Training Accuracy: 96.7024%, Validation Loss: 1.50
49, Validation Accuracy: 95.7976%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 405.54it/s]
Epoch [14/20], Training Loss: 1.5084, Training Accuracy: 96.9107%, Validation Loss: 1.50
17, Validation Accuracy: 96.2976%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 414.14it/s]
Epoch [15/20], Training Loss: 1.5042, Training Accuracy: 96.9583%, Validation Loss: 1.50
30, Validation Accuracy: 95.9762%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 431.02it/s]
Epoch [16/20], Training Loss: 1.5031, Training Accuracy: 97.1399%, Validation Loss: 1.50
05, Validation Accuracy: 96.2143%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 434.30it/s]
Epoch [17/20], Training Loss: 1.5014, Training Accuracy: 97.4077%, Validation Loss: 1.49
89, Validation Accuracy: 96.4524%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 448.01it/s]
Epoch [18/20], Training Loss: 1.4999, Training Accuracy: 97.3571%, Validation Loss: 1.49
70, Validation Accuracy: 96.6190%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 451.23it/s]
Epoch [19/20], Training Loss: 1.4994, Training Accuracy: 97.4940%, Validation Loss: 1.49
69, Validation Accuracy: 96.5357%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 440.77it/s]
Epoch [20/20], Training Loss: 1.4961, Training Accuracy: 97.6726%, Validation Loss: 1.49
60, Validation Accuracy: 96.6429%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 400.73it/s]
Epoch [1/20], Training Loss: 1.9883, Training Accuracy: 63.8185%, Validation Loss: 1.726
8, Validation Accuracy: 81.0357%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 416.63it/s]
Epoch [2/20], Training Loss: 1.7039, Training Accuracy: 82.9643%, Validation Loss: 1.639
3, Validation Accuracy: 84.1548%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 433.71it/s]
Epoch [3/20], Training Loss: 1.6524, Training Accuracy: 84.8571%, Validation Loss: 1.620
9, Validation Accuracy: 84.9405%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.19it/s]
Epoch [4/20], Training Loss: 1.6288, Training Accuracy: 86.4226%, Validation Loss: 1.584
2, Validation Accuracy: 89.8214%
Training: 100%|██████████████████████████████████████████████████████████| 105
<00:00, 446.42it/s]
```

```
Epoch [5/20], Training Loss: 1.5812, Training Accuracy: 92.2827%, Validation Loss: 1.547
2, Validation Accuracy: 92.7500%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 388.34it/s]
Epoch [6/20], Training Loss: 1.5584, Training Accuracy: 93.6815%, Validation Loss: 1.536
1, Validation Accuracy: 93.2857%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 400.84it/s]
Epoch [7/20], Training Loss: 1.5475, Training Accuracy: 94.2589%, Validation Loss: 1.528
8, Validation Accuracy: 93.9405%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 407.23it/s]
Epoch [8/20], Training Loss: 1.5392, Training Accuracy: 94.7440%, Validation Loss: 1.524
4, Validation Accuracy: 94.3452%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 384.37it/s]
Epoch [9/20], Training Loss: 1.5337, Training Accuracy: 95.1875%, Validation Loss: 1.520
3, Validation Accuracy: 94.7976%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 399.73it/s]
Epoch [10/20], Training Loss: 1.5285, Training Accuracy: 95.4464%, Validation Loss: 1.51
86, Validation Accuracy: 94.6786%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 403.29it/s]
Epoch [11/20], Training Loss: 1.5243, Training Accuracy: 95.7887%, Validation Loss: 1.51
40, Validation Accuracy: 95.1071%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 390.34it/s]
Epoch [12/20], Training Loss: 1.5205, Training Accuracy: 95.9821%, Validation Loss: 1.51
22, Validation Accuracy: 95.3333%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 437.26it/s]
Epoch [13/20], Training Loss: 1.5174, Training Accuracy: 96.1250%, Validation Loss: 1.51
06, Validation Accuracy: 95.3214%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.33it/s]
Epoch [14/20], Training Loss: 1.5151, Training Accuracy: 96.4940%, Validation Loss: 1.50
75, Validation Accuracy: 95.7262%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.39it/s]
Epoch [15/20], Training Loss: 1.5124, Training Accuracy: 96.4762%, Validation Loss: 1.50
68, Validation Accuracy: 95.8214%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 432.31it/s]
Epoch [16/20], Training Loss: 1.5107, Training Accuracy: 96.6607%, Validation Loss: 1.50
65, Validation Accuracy: 95.8333%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 437.52it/s]
Epoch [17/20], Training Loss: 1.5090, Training Accuracy: 96.6875%, Validation Loss: 1.50
39, Validation Accuracy: 96.1667%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 432.11it/s]
Epoch [18/20], Training Loss: 1.5073, Training Accuracy: 96.8899%, Validation Loss: 1.50
45, Validation Accuracy: 95.9762%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 421.98it/s]
Epoch [19/20], Training Loss: 1.5061, Training Accuracy: 96.9226%, Validation Loss: 1.50
16, Validation Accuracy: 96.2262%
Training: 100%|██████████████████████████████████████████████████████████| 105
<00:00, 428.59it/s]
```

```
Epoch [20/20], Training Loss: 1.5047, Training Accuracy: 97.0565%, Validation Loss: 1.50
18, Validation Accuracy: 96.1905%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 339.55it/s]
Epoch [1/20], Training Loss: 2.1533, Training Accuracy: 49.2440%, Validation Loss: 1.935
0, Validation Accuracy: 70.2857%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 365.04it/s]
Epoch [2/20], Training Loss: 1.8686, Training Accuracy: 72.9970%, Validation Loss: 1.748
6, Validation Accuracy: 79.5000%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 367.06it/s]
Epoch [3/20], Training Loss: 1.7412, Training Accuracy: 80.8601%, Validation Loss: 1.667
3, Validation Accuracy: 83.2619%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 350.82it/s]
Epoch [4/20], Training Loss: 1.6844, Training Accuracy: 83.1964%, Validation Loss: 1.641
5, Validation Accuracy: 84.0952%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 356.27it/s]
Epoch [5/20], Training Loss: 1.6595, Training Accuracy: 84.1964%, Validation Loss: 1.626
9, Validation Accuracy: 84.7738%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 363.35it/s]
Epoch [6/20], Training Loss: 1.6430, Training Accuracy: 85.1786%, Validation Loss: 1.609
8, Validation Accuracy: 87.1071%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 364.15it/s]
Epoch [7/20], Training Loss: 1.6175, Training Accuracy: 88.6607%, Validation Loss: 1.576
7, Validation Accuracy: 90.5119%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 334.66it/s]
Epoch [8/20], Training Loss: 1.5892, Training Accuracy: 91.5119%, Validation Loss: 1.555
2, Validation Accuracy: 92.0952%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 339.76it/s]
Epoch [9/20], Training Loss: 1.5709, Training Accuracy: 92.5804%, Validation Loss: 1.545
3, Validation Accuracy: 92.8214%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 327.51it/s]
Epoch [10/20], Training Loss: 1.5611, Training Accuracy: 93.3690%, Validation Loss: 1.53
82, Validation Accuracy: 93.3810%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 349.12it/s]
Epoch [11/20], Training Loss: 1.5535, Training Accuracy: 93.6429%, Validation Loss: 1.53
34, Validation Accuracy: 93.7143%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 351.58it/s]
Epoch [12/20], Training Loss: 1.5467, Training Accuracy: 94.0238%, Validation Loss: 1.52
89, Validation Accuracy: 94.0000%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 361.35it/s]
Epoch [13/20], Training Loss: 1.5414, Training Accuracy: 94.5357%, Validation Loss: 1.52
65, Validation Accuracy: 94.2857%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 354.76it/s]
Epoch [14/20], Training Loss: 1.5368, Training Accuracy: 94.8185%, Validation Loss: 1.52
36, Validation Accuracy: 94.4405%
Training: 100%|██████████████████████████████████████████████████████████| 52
00:00, 363.60it/s]
```

```
Epoch [15/20], Training Loss: 1.5338, Training Accuracy: 94.9762%, Validation Loss: 1.52
21, Validation Accuracy: 94.4643%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 333.13it/s]
Epoch [16/20], Training Loss: 1.5315, Training Accuracy: 95.2143%, Validation Loss: 1.51
94, Validation Accuracy: 94.7619%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 361.22it/s]
Epoch [17/20], Training Loss: 1.5281, Training Accuracy: 95.2589%, Validation Loss: 1.51
81, Validation Accuracy: 94.8333%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 366.52it/s]
Epoch [18/20], Training Loss: 1.5253, Training Accuracy: 95.4256%, Validation Loss: 1.51
52, Validation Accuracy: 95.0833%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 370.62it/s]
Epoch [19/20], Training Loss: 1.5238, Training Accuracy: 95.5387%, Validation Loss: 1.51
39, Validation Accuracy: 95.1429%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 367.54it/s]
Epoch [20/20], Training Loss: 1.5217, Training Accuracy: 95.7738%, Validation Loss: 1.51
26, Validation Accuracy: 95.2500%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 283.34it/s]
Epoch [1/20], Training Loss: 2.2284, Training Accuracy: 36.1518%, Validation Loss: 2.107
0, Validation Accuracy: 54.7738%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 283.04it/s]
Epoch [2/20], Training Loss: 2.0485, Training Accuracy: 58.6875%, Validation Loss: 1.933
5, Validation Accuracy: 65.0357%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 270.80it/s]
Epoch [3/20], Training Loss: 1.9205, Training Accuracy: 65.5863%, Validation Loss: 1.838
0, Validation Accuracy: 68.8333%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 277.82it/s]
Epoch [4/20], Training Loss: 1.8454, Training Accuracy: 70.4524%, Validation Loss: 1.778
0, Validation Accuracy: 74.4048%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 277.10it/s]
Epoch [5/20], Training Loss: 1.7924, Training Accuracy: 74.7411%, Validation Loss: 1.733
6, Validation Accuracy: 77.2857%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 279.59it/s]
Epoch [6/20], Training Loss: 1.7519, Training Accuracy: 78.0446%, Validation Loss: 1.700
5, Validation Accuracy: 80.5595%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 283.54it/s]
Epoch [7/20], Training Loss: 1.7207, Training Accuracy: 80.6458%, Validation Loss: 1.674
8, Validation Accuracy: 81.9881%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 282.13it/s]
Epoch [8/20], Training Loss: 1.6984, Training Accuracy: 81.7440%, Validation Loss: 1.659
3, Validation Accuracy: 82.6905%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 277.63it/s]
Epoch [9/20], Training Loss: 1.6822, Training Accuracy: 82.3780%, Validation Loss: 1.648
0, Validation Accuracy: 83.2262%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 226.78it/s]
```

```
Epoch [10/20], Training Loss: 1.6708, Training Accuracy: 82.9673%, Validation Loss: 1.63
68, Validation Accuracy: 83.6905%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 252.41it/s]
Epoch [11/20], Training Loss: 1.6526, Training Accuracy: 85.3214%, Validation Loss: 1.60
36, Validation Accuracy: 89.6429%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 281.22it/s]
Epoch [12/20], Training Loss: 1.6201, Training Accuracy: 90.0387%, Validation Loss: 1.57
70, Validation Accuracy: 91.6310%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 245.60it/s]
Epoch [13/20], Training Loss: 1.5997, Training Accuracy: 91.2500%, Validation Loss: 1.56
51, Validation Accuracy: 92.0119%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 241.09it/s]
Epoch [14/20], Training Loss: 1.5879, Training Accuracy: 91.7143%, Validation Loss: 1.55
75, Validation Accuracy: 92.3810%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.33it/s]
Epoch [15/20], Training Loss: 1.5791, Training Accuracy: 92.0149%, Validation Loss: 1.55
20, Validation Accuracy: 92.5476%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 236.97it/s]
Epoch [16/20], Training Loss: 1.5715, Training Accuracy: 92.3125%, Validation Loss: 1.54
76, Validation Accuracy: 92.7976%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 258.14it/s]
Epoch [17/20], Training Loss: 1.5669, Training Accuracy: 92.6458%, Validation Loss: 1.54
43, Validation Accuracy: 92.9762%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 251.30it/s]
Epoch [18/20], Training Loss: 1.5619, Training Accuracy: 92.8720%, Validation Loss: 1.54
12, Validation Accuracy: 93.2143%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 245.48it/s]
Epoch [19/20], Training Loss: 1.5588, Training Accuracy: 93.2440%, Validation Loss: 1.53
89, Validation Accuracy: 93.3810%
Training: 100%|██████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 242.99it/s]
Epoch [20/20], Training Loss: 1.5548, Training Accuracy: 93.2262%, Validation Loss: 1.53
61, Validation Accuracy: 93.5119%
Training: 100%|██████████████████████████████████████████████████| 420
0/4200 [00:10<00:00, 416.69it/s]
Epoch [1/20], Training Loss: 1.6546, Training Accuracy: 88.5298%, Validation Loss: 1.544
1, Validation Accuracy: 92.2976%
Training: 100%|██████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 433.60it/s]
Epoch [2/20], Training Loss: 1.5464, Training Accuracy: 94.7262%, Validation Loss: 1.520
4, Validation Accuracy: 94.3214%
Training: 100%|██████████████████████████████████████████████████| 420
0/4200 [00:10<00:00, 417.50it/s]
Epoch [3/20], Training Loss: 1.5294, Training Accuracy: 95.6399%, Validation Loss: 1.515
3, Validation Accuracy: 94.7381%
Training: 100%|██████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 433.48it/s]
Epoch [4/20], Training Loss: 1.5201, Training Accuracy: 96.1726%, Validation Loss: 1.529
3, Validation Accuracy: 93.4167%
Training: 100%|██████████████████████████████████████████████████| 420
<00:00, 448.01it/s]
```

```
Epoch [5/20], Training Loss: 1.5137, Training Accuracy: 96.3780%, Validation Loss: 1.507
6, Validation Accuracy: 95.6190%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.59it/s]
Epoch [6/20], Training Loss: 1.5092, Training Accuracy: 96.6518%, Validation Loss: 1.507
4, Validation Accuracy: 95.5476%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 489.87it/s]
Epoch [7/20], Training Loss: 1.5054, Training Accuracy: 96.8988%, Validation Loss: 1.501
2, Validation Accuracy: 96.1190%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 484.29it/s]
Epoch [8/20], Training Loss: 1.5030, Training Accuracy: 97.0446%, Validation Loss: 1.496
4, Validation Accuracy: 96.4881%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 482.66it/s]
Epoch [9/20], Training Loss: 1.4997, Training Accuracy: 97.2500%, Validation Loss: 1.496
0, Validation Accuracy: 96.7500%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 489.50it/s]
Epoch [10/20], Training Loss: 1.4980, Training Accuracy: 97.4583%, Validation Loss: 1.49
54, Validation Accuracy: 96.6905%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 484.39it/s]
Epoch [11/20], Training Loss: 1.4958, Training Accuracy: 97.5417%, Validation Loss: 1.49
44, Validation Accuracy: 96.7143%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.29it/s]
Epoch [12/20], Training Loss: 1.4933, Training Accuracy: 97.6607%, Validation Loss: 1.49
51, Validation Accuracy: 96.6429%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 436.77it/s]
Epoch [13/20], Training Loss: 1.4924, Training Accuracy: 97.7857%, Validation Loss: 1.49
42, Validation Accuracy: 96.7500%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 476.39it/s]
Epoch [14/20], Training Loss: 1.4909, Training Accuracy: 97.8304%, Validation Loss: 1.49
51, Validation Accuracy: 96.6429%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 487.05it/s]
Epoch [15/20], Training Loss: 1.4898, Training Accuracy: 97.8780%, Validation Loss: 1.49
29, Validation Accuracy: 96.9286%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 487.74it/s]
Epoch [16/20], Training Loss: 1.4889, Training Accuracy: 98.0357%, Validation Loss: 1.49
29, Validation Accuracy: 96.8214%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 486.21it/s]
Epoch [17/20], Training Loss: 1.4883, Training Accuracy: 98.0238%, Validation Loss: 1.48
88, Validation Accuracy: 97.3452%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 477.18it/s]
Epoch [18/20], Training Loss: 1.4876, Training Accuracy: 98.0685%, Validation Loss: 1.49
03, Validation Accuracy: 97.1429%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 450.09it/s]
Epoch [19/20], Training Loss: 1.4868, Training Accuracy: 98.1250%, Validation Loss: 1.49
23, Validation Accuracy: 96.9643%
Training: 100%|██████████████████████████████████████████████████████████| 420
<00:00, 461.03it/s]
```

```
Epoch [20/20], Training Loss: 1.4850, Training Accuracy: 98.2381%, Validation Loss: 1.49
11, Validation Accuracy: 97.0833%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 437.32it/s]
Epoch [1/20], Training Loss: 1.7342, Training Accuracy: 80.8423%, Validation Loss: 1.614
1, Validation Accuracy: 85.2976%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 464.30it/s]
Epoch [2/20], Training Loss: 1.6219, Training Accuracy: 87.0030%, Validation Loss: 1.601
6, Validation Accuracy: 86.0595%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 433.98it/s]
Epoch [3/20], Training Loss: 1.5830, Training Accuracy: 90.5982%, Validation Loss: 1.530
9, Validation Accuracy: 93.6667%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 463.36it/s]
Epoch [4/20], Training Loss: 1.5338, Training Accuracy: 95.4673%, Validation Loss: 1.519
7, Validation Accuracy: 94.4048%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 454.24it/s]
Epoch [5/20], Training Loss: 1.5230, Training Accuracy: 96.2113%, Validation Loss: 1.507
4, Validation Accuracy: 95.6548%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 478.08it/s]
Epoch [6/20], Training Loss: 1.5142, Training Accuracy: 96.5149%, Validation Loss: 1.505
5, Validation Accuracy: 95.8095%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 474.02it/s]
Epoch [7/20], Training Loss: 1.5102, Training Accuracy: 96.8036%, Validation Loss: 1.505
6, Validation Accuracy: 95.7143%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 456.51it/s]
Epoch [8/20], Training Loss: 1.5067, Training Accuracy: 97.0000%, Validation Loss: 1.505
4, Validation Accuracy: 95.7143%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 435.79it/s]
Epoch [9/20], Training Loss: 1.5033, Training Accuracy: 97.1607%, Validation Loss: 1.500
9, Validation Accuracy: 96.1667%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 416.22it/s]
Epoch [10/20], Training Loss: 1.5009, Training Accuracy: 97.4315%, Validation Loss: 1.49
98, Validation Accuracy: 96.2619%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 420.03it/s]
Epoch [11/20], Training Loss: 1.4995, Training Accuracy: 97.3750%, Validation Loss: 1.49
46, Validation Accuracy: 96.8333%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 432.22it/s]
Epoch [12/20], Training Loss: 1.4964, Training Accuracy: 97.5387%, Validation Loss: 1.49
48, Validation Accuracy: 96.6786%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 436.70it/s]
Epoch [13/20], Training Loss: 1.4962, Training Accuracy: 97.6875%, Validation Loss: 1.49
48, Validation Accuracy: 96.7381%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 440.12it/s]
Epoch [14/20], Training Loss: 1.4938, Training Accuracy: 97.6845%, Validation Loss: 1.49
15, Validation Accuracy: 97.0833%
Training: 100%|██████████████████████████████████████████████████████████| 210
<00:00, 459.61it/s]
```

```
Epoch [15/20], Training Loss: 1.4930, Training Accuracy: 97.7827%, Validation Loss: 1.49
29, Validation Accuracy: 96.8452%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 480.07it/s]
Epoch [16/20], Training Loss: 1.4918, Training Accuracy: 97.8452%, Validation Loss: 1.49
34, Validation Accuracy: 96.9167%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 482.90it/s]
Epoch [17/20], Training Loss: 1.4902, Training Accuracy: 98.0030%, Validation Loss: 1.49
35, Validation Accuracy: 96.9048%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 474.81it/s]
Epoch [18/20], Training Loss: 1.4887, Training Accuracy: 98.1310%, Validation Loss: 1.49
30, Validation Accuracy: 96.9167%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 478.91it/s]
Epoch [19/20], Training Loss: 1.4878, Training Accuracy: 98.1310%, Validation Loss: 1.49
04, Validation Accuracy: 97.1905%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 477.50it/s]
Epoch [20/20], Training Loss: 1.4871, Training Accuracy: 98.1845%, Validation Loss: 1.48
99, Validation Accuracy: 97.2381%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 450.30it/s]
Epoch [1/20], Training Loss: 1.8217, Training Accuracy: 73.4762%, Validation Loss: 1.637
4, Validation Accuracy: 84.4524%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.65it/s]
Epoch [2/20], Training Loss: 1.5970, Training Accuracy: 91.9911%, Validation Loss: 1.544
8, Validation Accuracy: 92.6667%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.08it/s]
Epoch [3/20], Training Loss: 1.5536, Training Accuracy: 94.3512%, Validation Loss: 1.530
1, Validation Accuracy: 93.7619%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.90it/s]
Epoch [4/20], Training Loss: 1.5365, Training Accuracy: 95.2113%, Validation Loss: 1.519
1, Validation Accuracy: 94.6667%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 432.72it/s]
Epoch [5/20], Training Loss: 1.5264, Training Accuracy: 95.9554%, Validation Loss: 1.514
7, Validation Accuracy: 95.0119%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 395.23it/s]
Epoch [6/20], Training Loss: 1.5213, Training Accuracy: 96.1399%, Validation Loss: 1.509
9, Validation Accuracy: 95.4643%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 410.83it/s]
Epoch [7/20], Training Loss: 1.5152, Training Accuracy: 96.4970%, Validation Loss: 1.510
9, Validation Accuracy: 95.2381%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.94it/s]
Epoch [8/20], Training Loss: 1.5116, Training Accuracy: 96.7173%, Validation Loss: 1.504
2, Validation Accuracy: 96.0000%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 406.31it/s]
Epoch [9/20], Training Loss: 1.5082, Training Accuracy: 96.9315%, Validation Loss: 1.505
6, Validation Accuracy: 95.8095%
Training: 100%|██████████████████████████████████████████████████████████| 105
<00:00, 373.71it/s]
```

```
Epoch [10/20], Training Loss: 1.5056, Training Accuracy: 97.0685%, Validation Loss: 1.50
05, Validation Accuracy: 96.2500%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 428.22it/s]
Epoch [11/20], Training Loss: 1.5035, Training Accuracy: 97.1190%, Validation Loss: 1.50
09, Validation Accuracy: 96.1667%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 437.87it/s]
Epoch [12/20], Training Loss: 1.5016, Training Accuracy: 97.3899%, Validation Loss: 1.49
88, Validation Accuracy: 96.3810%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.39it/s]
Epoch [13/20], Training Loss: 1.4992, Training Accuracy: 97.3869%, Validation Loss: 1.49
92, Validation Accuracy: 96.2619%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 370.59it/s]
Epoch [14/20], Training Loss: 1.4970, Training Accuracy: 97.5744%, Validation Loss: 1.49
68, Validation Accuracy: 96.6190%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 436.49it/s]
Epoch [15/20], Training Loss: 1.4967, Training Accuracy: 97.5923%, Validation Loss: 1.49
65, Validation Accuracy: 96.6310%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 443.64it/s]
Epoch [16/20], Training Loss: 1.4953, Training Accuracy: 97.6399%, Validation Loss: 1.49
61, Validation Accuracy: 96.6786%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.51it/s]
Epoch [17/20], Training Loss: 1.4930, Training Accuracy: 97.7768%, Validation Loss: 1.49
69, Validation Accuracy: 96.5595%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 438.17it/s]
Epoch [18/20], Training Loss: 1.4932, Training Accuracy: 97.9048%, Validation Loss: 1.49
52, Validation Accuracy: 96.6071%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.91it/s]
Epoch [19/20], Training Loss: 1.4919, Training Accuracy: 97.8780%, Validation Loss: 1.49
47, Validation Accuracy: 96.7024%
Training: 100%|████████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 454.21it/s]
Epoch [20/20], Training Loss: 1.4908, Training Accuracy: 97.9673%, Validation Loss: 1.49
62, Validation Accuracy: 96.5476%
Training: 100%|██████████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 359.62it/s]
Epoch [1/20], Training Loss: 1.9245, Training Accuracy: 66.8155%, Validation Loss: 1.670
6, Validation Accuracy: 83.0476%
Training: 100%|██████████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 358.92it/s]
Epoch [2/20], Training Loss: 1.6680, Training Accuracy: 84.5714%, Validation Loss: 1.625
3, Validation Accuracy: 84.7381%
Training: 100%|██████████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 368.47it/s]
Epoch [3/20], Training Loss: 1.6283, Training Accuracy: 86.8214%, Validation Loss: 1.571
4, Validation Accuracy: 91.2143%
Training: 100%|██████████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 371.06it/s]
Epoch [4/20], Training Loss: 1.5709, Training Accuracy: 93.1458%, Validation Loss: 1.543
8, Validation Accuracy: 92.6310%
Training: 100%|██████████████████████████████████████████████████████████████████████████| 52
Loading [MathJax]/extensions/Safe.js 00:00, 357.00it/s]
```

```
Epoch [5/20], Training Loss: 1.5498, Training Accuracy: 94.5000%, Validation Loss: 1.530
6, Validation Accuracy: 93.8810%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 332.87it/s]
Epoch [6/20], Training Loss: 1.5394, Training Accuracy: 94.8958%, Validation Loss: 1.523
3, Validation Accuracy: 94.4286%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 364.08it/s]
Epoch [7/20], Training Loss: 1.5319, Training Accuracy: 95.3542%, Validation Loss: 1.520
0, Validation Accuracy: 94.6548%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 362.68it/s]
Epoch [8/20], Training Loss: 1.5263, Training Accuracy: 95.6875%, Validation Loss: 1.518
2, Validation Accuracy: 94.7738%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 349.44it/s]
Epoch [9/20], Training Loss: 1.5209, Training Accuracy: 96.0655%, Validation Loss: 1.512
6, Validation Accuracy: 95.3571%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 349.59it/s]
Epoch [10/20], Training Loss: 1.5177, Training Accuracy: 96.1488%, Validation Loss: 1.51
15, Validation Accuracy: 95.3333%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 343.62it/s]
Epoch [11/20], Training Loss: 1.5137, Training Accuracy: 96.4821%, Validation Loss: 1.50
74, Validation Accuracy: 95.6667%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 337.62it/s]
Epoch [12/20], Training Loss: 1.5105, Training Accuracy: 96.5655%, Validation Loss: 1.50
62, Validation Accuracy: 95.7976%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 351.32it/s]
Epoch [13/20], Training Loss: 1.5083, Training Accuracy: 96.8601%, Validation Loss: 1.50
37, Validation Accuracy: 96.0000%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 346.11it/s]
Epoch [14/20], Training Loss: 1.5060, Training Accuracy: 96.9345%, Validation Loss: 1.50
38, Validation Accuracy: 96.0714%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 356.30it/s]
Epoch [15/20], Training Loss: 1.5051, Training Accuracy: 97.0595%, Validation Loss: 1.50
39, Validation Accuracy: 95.9643%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 359.33it/s]
Epoch [16/20], Training Loss: 1.5030, Training Accuracy: 97.1875%, Validation Loss: 1.50
11, Validation Accuracy: 96.1905%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 364.99it/s]
Epoch [17/20], Training Loss: 1.5018, Training Accuracy: 97.3006%, Validation Loss: 1.50
11, Validation Accuracy: 96.3095%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 336.22it/s]
Epoch [18/20], Training Loss: 1.4998, Training Accuracy: 97.4256%, Validation Loss: 1.49
79, Validation Accuracy: 96.6190%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 357.97it/s]
Epoch [19/20], Training Loss: 1.4978, Training Accuracy: 97.5446%, Validation Loss: 1.49
86, Validation Accuracy: 96.5238%
Training: 100%|██████████████████████████████████████████████████████████| 52
00:00, 350.18it/s]
```

```
Epoch [20/20], Training Loss: 1.4975, Training Accuracy: 97.5536%, Validation Loss: 1.49
62, Validation Accuracy: 96.6786%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 262.74it/s]
Epoch [1/20], Training Loss: 2.0646, Training Accuracy: 55.2917%, Validation Loss: 1.826
9, Validation Accuracy: 71.5000%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 236.95it/s]
Epoch [2/20], Training Loss: 1.7960, Training Accuracy: 74.1815%, Validation Loss: 1.709
0, Validation Accuracy: 79.0000%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 254.83it/s]
Epoch [3/20], Training Loss: 1.6990, Training Accuracy: 82.2738%, Validation Loss: 1.635
7, Validation Accuracy: 84.8571%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 233.60it/s]
Epoch [4/20], Training Loss: 1.6327, Training Accuracy: 88.7292%, Validation Loss: 1.575
4, Validation Accuracy: 91.0833%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 250.23it/s]
Epoch [5/20], Training Loss: 1.5902, Training Accuracy: 92.0089%, Validation Loss: 1.555
8, Validation Accuracy: 92.0476%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 227.01it/s]
Epoch [6/20], Training Loss: 1.5707, Training Accuracy: 93.0655%, Validation Loss: 1.544
9, Validation Accuracy: 92.8929%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 257.90it/s]
Epoch [7/20], Training Loss: 1.5591, Training Accuracy: 93.5625%, Validation Loss: 1.541
4, Validation Accuracy: 92.8571%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 274.70it/s]
Epoch [8/20], Training Loss: 1.5489, Training Accuracy: 94.1637%, Validation Loss: 1.530
8, Validation Accuracy: 93.9405%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 270.65it/s]
Epoch [9/20], Training Loss: 1.5415, Training Accuracy: 94.4375%, Validation Loss: 1.526
3, Validation Accuracy: 94.2500%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 272.60it/s]
Epoch [10/20], Training Loss: 1.5367, Training Accuracy: 94.8065%, Validation Loss: 1.52
35, Validation Accuracy: 94.5357%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 278.25it/s]
Epoch [11/20], Training Loss: 1.5337, Training Accuracy: 95.0446%, Validation Loss: 1.51
95, Validation Accuracy: 94.7738%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 277.08it/s]
Epoch [12/20], Training Loss: 1.5288, Training Accuracy: 95.4256%, Validation Loss: 1.51
71, Validation Accuracy: 94.9286%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 270.30it/s]
Epoch [13/20], Training Loss: 1.5255, Training Accuracy: 95.6042%, Validation Loss: 1.51
45, Validation Accuracy: 95.0357%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 274.62it/s]
Epoch [14/20], Training Loss: 1.5216, Training Accuracy: 95.7738%, Validation Loss: 1.51
18, Validation Accuracy: 95.4405%
Training: 100%|██████████████████████████████████████████████████████████| 26
00:00, 273.37it/s]
```

```
Epoch [15/20], Training Loss: 1.5208, Training Accuracy: 95.8631%, Validation Loss: 1.51
13, Validation Accuracy: 95.2738%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 270.90it/s]
Epoch [16/20], Training Loss: 1.5174, Training Accuracy: 96.0327%, Validation Loss: 1.50
86, Validation Accuracy: 95.5119%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 276.02it/s]
Epoch [17/20], Training Loss: 1.5153, Training Accuracy: 96.3036%, Validation Loss: 1.50
92, Validation Accuracy: 95.5119%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 256.05it/s]
Epoch [18/20], Training Loss: 1.5129, Training Accuracy: 96.3631%, Validation Loss: 1.50
63, Validation Accuracy: 95.7857%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 271.90it/s]
Epoch [19/20], Training Loss: 1.5111, Training Accuracy: 96.4970%, Validation Loss: 1.50
66, Validation Accuracy: 95.8214%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 272.20it/s]
Epoch [20/20], Training Loss: 1.5096, Training Accuracy: 96.6786%, Validation Loss: 1.50
50, Validation Accuracy: 95.7857%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 500.08it/s]
Epoch [1/20], Training Loss: 1.6324, Training Accuracy: 89.7887%, Validation Loss: 1.535
6, Validation Accuracy: 92.8452%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 503.58it/s]
Epoch [2/20], Training Loss: 1.5420, Training Accuracy: 94.8958%, Validation Loss: 1.529
7, Validation Accuracy: 93.3810%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 505.40it/s]
Epoch [3/20], Training Loss: 1.5244, Training Accuracy: 95.7113%, Validation Loss: 1.513
7, Validation Accuracy: 94.9286%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 502.55it/s]
Epoch [4/20], Training Loss: 1.5175, Training Accuracy: 96.1935%, Validation Loss: 1.508
2, Validation Accuracy: 95.4167%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 489.22it/s]
Epoch [5/20], Training Loss: 1.5106, Training Accuracy: 96.5714%, Validation Loss: 1.502
8, Validation Accuracy: 95.9762%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 491.16it/s]
Epoch [6/20], Training Loss: 1.5063, Training Accuracy: 96.8929%, Validation Loss: 1.501
4, Validation Accuracy: 96.0952%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 478.33it/s]
Epoch [7/20], Training Loss: 1.5035, Training Accuracy: 96.9673%, Validation Loss: 1.502
4, Validation Accuracy: 95.9524%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 475.71it/s]
Epoch [8/20], Training Loss: 1.5009, Training Accuracy: 97.2589%, Validation Loss: 1.501
8, Validation Accuracy: 95.9524%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 478.92it/s]
Epoch [9/20], Training Loss: 1.4996, Training Accuracy: 97.3601%, Validation Loss: 1.503
5, Validation Accuracy: 95.7857%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
            <00:00, 468.58it/s]
```

```
Epoch [10/20], Training Loss: 1.4972, Training Accuracy: 97.4464%, Validation Loss: 1.49
54, Validation Accuracy: 96.6071%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 509.92it/s]
Epoch [11/20], Training Loss: 1.4940, Training Accuracy: 97.6250%, Validation Loss: 1.49
39, Validation Accuracy: 96.7976%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 509.65it/s]
Epoch [12/20], Training Loss: 1.4929, Training Accuracy: 97.5893%, Validation Loss: 1.49
26, Validation Accuracy: 96.9286%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 467.72it/s]
Epoch [13/20], Training Loss: 1.4916, Training Accuracy: 97.7262%, Validation Loss: 1.49
52, Validation Accuracy: 96.6071%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 478.19it/s]
Epoch [14/20], Training Loss: 1.4909, Training Accuracy: 97.8571%, Validation Loss: 1.49
14, Validation Accuracy: 97.1071%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 492.61it/s]
Epoch [15/20], Training Loss: 1.4893, Training Accuracy: 97.8839%, Validation Loss: 1.49
19, Validation Accuracy: 96.9643%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 476.74it/s]
Epoch [16/20], Training Loss: 1.4872, Training Accuracy: 98.0298%, Validation Loss: 1.49
17, Validation Accuracy: 96.9762%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 508.67it/s]
Epoch [17/20], Training Loss: 1.4878, Training Accuracy: 98.0744%, Validation Loss: 1.49
26, Validation Accuracy: 96.8929%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 511.63it/s]
Epoch [18/20], Training Loss: 1.4868, Training Accuracy: 98.1280%, Validation Loss: 1.49
19, Validation Accuracy: 96.8810%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 489.77it/s]
Epoch [19/20], Training Loss: 1.4859, Training Accuracy: 98.1577%, Validation Loss: 1.48
81, Validation Accuracy: 97.3452%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 465.74it/s]
Epoch [20/20], Training Loss: 1.4853, Training Accuracy: 98.1607%, Validation Loss: 1.49
07, Validation Accuracy: 97.1190%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 446.69it/s]
Epoch [1/20], Training Loss: 1.6490, Training Accuracy: 89.2679%, Validation Loss: 1.539
2, Validation Accuracy: 92.6429%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 440.80it/s]
Epoch [2/20], Training Loss: 1.5459, Training Accuracy: 94.7202%, Validation Loss: 1.524
7, Validation Accuracy: 93.9524%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 441.33it/s]
Epoch [3/20], Training Loss: 1.5305, Training Accuracy: 95.4345%, Validation Loss: 1.518
2, Validation Accuracy: 94.5714%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 439.79it/s]
Epoch [4/20], Training Loss: 1.5200, Training Accuracy: 95.9643%, Validation Loss: 1.506
6, Validation Accuracy: 95.6905%
Training: 100%|██████████████████████████████████████████████████████████| 210
<00:00, 428.27it/s]
```

```
Epoch [5/20], Training Loss: 1.5151, Training Accuracy: 96.4167%, Validation Loss: 1.507
3, Validation Accuracy: 95.3810%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 476.41it/s]
Epoch [6/20], Training Loss: 1.5100, Training Accuracy: 96.6518%, Validation Loss: 1.501
9, Validation Accuracy: 96.0119%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 479.58it/s]
Epoch [7/20], Training Loss: 1.5064, Training Accuracy: 96.8125%, Validation Loss: 1.501
8, Validation Accuracy: 96.0238%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 482.33it/s]
Epoch [8/20], Training Loss: 1.5045, Training Accuracy: 96.9911%, Validation Loss: 1.502
7, Validation Accuracy: 95.7976%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 473.35it/s]
Epoch [9/20], Training Loss: 1.5025, Training Accuracy: 97.1994%, Validation Loss: 1.498
6, Validation Accuracy: 96.3452%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 483.23it/s]
Epoch [10/20], Training Loss: 1.4987, Training Accuracy: 97.4137%, Validation Loss: 1.49
65, Validation Accuracy: 96.6190%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 468.35it/s]
Epoch [11/20], Training Loss: 1.4963, Training Accuracy: 97.5179%, Validation Loss: 1.49
59, Validation Accuracy: 96.5714%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 483.50it/s]
Epoch [12/20], Training Loss: 1.4962, Training Accuracy: 97.4881%, Validation Loss: 1.49
56, Validation Accuracy: 96.6310%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 476.40it/s]
Epoch [13/20], Training Loss: 1.4946, Training Accuracy: 97.7143%, Validation Loss: 1.49
45, Validation Accuracy: 96.7262%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 483.91it/s]
Epoch [14/20], Training Loss: 1.4927, Training Accuracy: 97.7351%, Validation Loss: 1.49
45, Validation Accuracy: 96.7262%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 475.16it/s]
Epoch [15/20], Training Loss: 1.4903, Training Accuracy: 97.8958%, Validation Loss: 1.49
17, Validation Accuracy: 96.9643%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 481.46it/s]
Epoch [16/20], Training Loss: 1.4903, Training Accuracy: 97.8690%, Validation Loss: 1.49
28, Validation Accuracy: 96.8333%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 474.22it/s]
Epoch [17/20], Training Loss: 1.4888, Training Accuracy: 97.9911%, Validation Loss: 1.49
29, Validation Accuracy: 96.8333%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 458.23it/s]
Epoch [18/20], Training Loss: 1.4881, Training Accuracy: 98.1250%, Validation Loss: 1.49
05, Validation Accuracy: 97.1667%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 468.65it/s]
Epoch [19/20], Training Loss: 1.4871, Training Accuracy: 98.1190%, Validation Loss: 1.49
15, Validation Accuracy: 97.0238%
Training: 100%|██████████████████████████████████████████████████████████| 210
<00:00, 479.74it/s]
```

```
Epoch [20/20], Training Loss: 1.4861, Training Accuracy: 98.1696%, Validation Loss: 1.49
08, Validation Accuracy: 97.0357%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.95it/s]
Epoch [1/20], Training Loss: 1.7395, Training Accuracy: 81.1696%, Validation Loss: 1.614
9, Validation Accuracy: 85.1429%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 436.43it/s]
Epoch [2/20], Training Loss: 1.5772, Training Accuracy: 92.7440%, Validation Loss: 1.533
4, Validation Accuracy: 93.4167%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.50it/s]
Epoch [3/20], Training Loss: 1.5412, Training Accuracy: 95.2381%, Validation Loss: 1.520
2, Validation Accuracy: 94.5595%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.36it/s]
Epoch [4/20], Training Loss: 1.5288, Training Accuracy: 95.7798%, Validation Loss: 1.514
4, Validation Accuracy: 95.0238%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 387.12it/s]
Epoch [5/20], Training Loss: 1.5197, Training Accuracy: 96.3006%, Validation Loss: 1.507
4, Validation Accuracy: 95.6310%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 385.14it/s]
Epoch [6/20], Training Loss: 1.5140, Training Accuracy: 96.5357%, Validation Loss: 1.504
7, Validation Accuracy: 95.9167%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 389.24it/s]
Epoch [7/20], Training Loss: 1.5108, Training Accuracy: 96.7470%, Validation Loss: 1.502
5, Validation Accuracy: 96.0952%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 408.70it/s]
Epoch [8/20], Training Loss: 1.5070, Training Accuracy: 96.9613%, Validation Loss: 1.499
8, Validation Accuracy: 96.2143%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 364.42it/s]
Epoch [9/20], Training Loss: 1.5034, Training Accuracy: 97.1488%, Validation Loss: 1.498
6, Validation Accuracy: 96.3571%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 424.24it/s]
Epoch [10/20], Training Loss: 1.5013, Training Accuracy: 97.2292%, Validation Loss: 1.49
85, Validation Accuracy: 96.3929%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.82it/s]
Epoch [11/20], Training Loss: 1.5001, Training Accuracy: 97.3631%, Validation Loss: 1.49
87, Validation Accuracy: 96.2976%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.69it/s]
Epoch [12/20], Training Loss: 1.4978, Training Accuracy: 97.4345%, Validation Loss: 1.49
51, Validation Accuracy: 96.7738%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 372.98it/s]
Epoch [13/20], Training Loss: 1.4971, Training Accuracy: 97.6042%, Validation Loss: 1.49
60, Validation Accuracy: 96.6310%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 419.87it/s]
Epoch [14/20], Training Loss: 1.4957, Training Accuracy: 97.7202%, Validation Loss: 1.49
55, Validation Accuracy: 96.7024%
Training: 100%|██████████████████████████████████████████████████████████| 105
<00:00, 420.23it/s]
```

```
Epoch [15/20], Training Loss: 1.4938, Training Accuracy: 97.7917%, Validation Loss: 1.49
42, Validation Accuracy: 96.7381%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 438.11it/s]
Epoch [16/20], Training Loss: 1.4920, Training Accuracy: 97.8065%, Validation Loss: 1.49
61, Validation Accuracy: 96.5952%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 436.12it/s]
Epoch [17/20], Training Loss: 1.4909, Training Accuracy: 97.9315%, Validation Loss: 1.49
46, Validation Accuracy: 96.7381%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 407.59it/s]
Epoch [18/20], Training Loss: 1.4901, Training Accuracy: 97.9375%, Validation Loss: 1.49
39, Validation Accuracy: 96.8214%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.26it/s]
Epoch [19/20], Training Loss: 1.4899, Training Accuracy: 98.0149%, Validation Loss: 1.49
04, Validation Accuracy: 97.0952%
Training: 100%|███████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.32it/s]
Epoch [20/20], Training Loss: 1.4880, Training Accuracy: 98.1577%, Validation Loss: 1.49
12, Validation Accuracy: 97.0000%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 300.27it/s]
Epoch [1/20], Training Loss: 1.8335, Training Accuracy: 73.4613%, Validation Loss: 1.595
3, Validation Accuracy: 89.5000%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 316.26it/s]
Epoch [2/20], Training Loss: 1.5892, Training Accuracy: 92.5952%, Validation Loss: 1.550
1, Validation Accuracy: 92.1071%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 369.65it/s]
Epoch [3/20], Training Loss: 1.5570, Training Accuracy: 94.0804%, Validation Loss: 1.531
4, Validation Accuracy: 93.4524%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 368.82it/s]
Epoch [4/20], Training Loss: 1.5413, Training Accuracy: 94.9286%, Validation Loss: 1.522
7, Validation Accuracy: 94.1905%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 362.87it/s]
Epoch [5/20], Training Loss: 1.5305, Training Accuracy: 95.5625%, Validation Loss: 1.517
1, Validation Accuracy: 94.6905%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 366.01it/s]
Epoch [6/20], Training Loss: 1.5230, Training Accuracy: 95.9405%, Validation Loss: 1.514
8, Validation Accuracy: 95.0357%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 361.09it/s]
Epoch [7/20], Training Loss: 1.5178, Training Accuracy: 96.2411%, Validation Loss: 1.510
6, Validation Accuracy: 95.3333%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 362.01it/s]
Epoch [8/20], Training Loss: 1.5148, Training Accuracy: 96.3780%, Validation Loss: 1.506
3, Validation Accuracy: 95.7381%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 368.55it/s]
Epoch [9/20], Training Loss: 1.5108, Training Accuracy: 96.6726%, Validation Loss: 1.506
8, Validation Accuracy: 95.7143%
Training: 100%|█████████████████████████████████████████████████████████████████████| 52
<00:00, 372.47it/s]
```

```
Epoch [10/20], Training Loss: 1.5073, Training Accuracy: 96.8988%, Validation Loss: 1.50
57, Validation Accuracy: 95.7976%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 363.74it/s]
Epoch [11/20], Training Loss: 1.5042, Training Accuracy: 97.0744%, Validation Loss: 1.50
20, Validation Accuracy: 96.0952%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 325.43it/s]
Epoch [12/20], Training Loss: 1.5017, Training Accuracy: 97.1607%, Validation Loss: 1.50
04, Validation Accuracy: 96.2976%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 364.52it/s]
Epoch [13/20], Training Loss: 1.5003, Training Accuracy: 97.4137%, Validation Loss: 1.49
88, Validation Accuracy: 96.4762%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 369.08it/s]
Epoch [14/20], Training Loss: 1.4987, Training Accuracy: 97.4315%, Validation Loss: 1.49
91, Validation Accuracy: 96.3571%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 335.51it/s]
Epoch [15/20], Training Loss: 1.4979, Training Accuracy: 97.5238%, Validation Loss: 1.49
97, Validation Accuracy: 96.3214%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 352.74it/s]
Epoch [16/20], Training Loss: 1.4962, Training Accuracy: 97.6012%, Validation Loss: 1.49
75, Validation Accuracy: 96.5238%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 348.76it/s]
Epoch [17/20], Training Loss: 1.4963, Training Accuracy: 97.7173%, Validation Loss: 1.49
46, Validation Accuracy: 96.8452%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 305.97it/s]
Epoch [18/20], Training Loss: 1.4940, Training Accuracy: 97.7738%, Validation Loss: 1.49
46, Validation Accuracy: 96.8095%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 345.99it/s]
Epoch [19/20], Training Loss: 1.4935, Training Accuracy: 97.7887%, Validation Loss: 1.49
55, Validation Accuracy: 96.7976%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 362.71it/s]
Epoch [20/20], Training Loss: 1.4921, Training Accuracy: 97.8363%, Validation Loss: 1.49
50, Validation Accuracy: 96.7500%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 276.64it/s]
Epoch [1/20], Training Loss: 1.9692, Training Accuracy: 65.9018%, Validation Loss: 1.686
7, Validation Accuracy: 86.0952%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 263.00it/s]
Epoch [2/20], Training Loss: 1.6460, Training Accuracy: 89.8333%, Validation Loss: 1.572
9, Validation Accuracy: 91.1429%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 276.36it/s]
Epoch [3/20], Training Loss: 1.5851, Training Accuracy: 92.5833%, Validation Loss: 1.547
2, Validation Accuracy: 92.7500%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 265.98it/s]
Epoch [4/20], Training Loss: 1.5619, Training Accuracy: 93.4732%, Validation Loss: 1.537
5, Validation Accuracy: 93.1667%
Training: 100%|██████████████████████████████████████████████████████████| 26
00:00, 238.36it/s]
```

```
Epoch [5/20], Training Loss: 1.5501, Training Accuracy: 94.2381%, Validation Loss: 1.533
5, Validation Accuracy: 93.5119%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 268.79it/s]
Epoch [6/20], Training Loss: 1.5399, Training Accuracy: 94.7679%, Validation Loss: 1.532
0, Validation Accuracy: 93.5714%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 259.98it/s]
Epoch [7/20], Training Loss: 1.5350, Training Accuracy: 95.1548%, Validation Loss: 1.523
8, Validation Accuracy: 94.2024%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 250.04it/s]
Epoch [8/20], Training Loss: 1.5288, Training Accuracy: 95.3899%, Validation Loss: 1.522
1, Validation Accuracy: 94.3929%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 277.91it/s]
Epoch [9/20], Training Loss: 1.5253, Training Accuracy: 95.5714%, Validation Loss: 1.514
6, Validation Accuracy: 95.0833%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 257.93it/s]
Epoch [10/20], Training Loss: 1.5204, Training Accuracy: 96.0506%, Validation Loss: 1.51
13, Validation Accuracy: 95.3333%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 266.73it/s]
Epoch [11/20], Training Loss: 1.5180, Training Accuracy: 96.2679%, Validation Loss: 1.51
27, Validation Accuracy: 95.2738%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 261.57it/s]
Epoch [12/20], Training Loss: 1.5134, Training Accuracy: 96.3839%, Validation Loss: 1.50
86, Validation Accuracy: 95.5476%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 230.63it/s]
Epoch [13/20], Training Loss: 1.5111, Training Accuracy: 96.5149%, Validation Loss: 1.50
61, Validation Accuracy: 95.8571%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 224.90it/s]
Epoch [14/20], Training Loss: 1.5085, Training Accuracy: 96.7321%, Validation Loss: 1.50
68, Validation Accuracy: 95.6667%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 236.64it/s]
Epoch [15/20], Training Loss: 1.5066, Training Accuracy: 96.8810%, Validation Loss: 1.50
64, Validation Accuracy: 95.7857%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 219.71it/s]
Epoch [16/20], Training Loss: 1.5044, Training Accuracy: 97.0446%, Validation Loss: 1.50
23, Validation Accuracy: 96.1310%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 239.34it/s]
Epoch [17/20], Training Loss: 1.5038, Training Accuracy: 97.1369%, Validation Loss: 1.50
32, Validation Accuracy: 96.0476%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 275.39it/s]
Epoch [18/20], Training Loss: 1.5017, Training Accuracy: 97.2500%, Validation Loss: 1.49
88, Validation Accuracy: 96.4286%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 246.16it/s]
Epoch [19/20], Training Loss: 1.5007, Training Accuracy: 97.3958%, Validation Loss: 1.50
51, Validation Accuracy: 95.9048%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 284.27it/s]
```

```
Epoch [20/20], Training Loss: 1.4988, Training Accuracy: 97.4524%, Validation Loss: 1.50
04, Validation Accuracy: 96.2500%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 508.95it/s]
Epoch [1/20], Training Loss: 1.6701, Training Accuracy: 84.0327%, Validation Loss: 1.615
0, Validation Accuracy: 84.6310%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 508.77it/s]
Epoch [2/20], Training Loss: 1.6126, Training Accuracy: 86.8155%, Validation Loss: 1.597
9, Validation Accuracy: 86.4048%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 505.57it/s]
Epoch [3/20], Training Loss: 1.6009, Training Accuracy: 87.4524%, Validation Loss: 1.590
3, Validation Accuracy: 87.1190%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 508.20it/s]
Epoch [4/20], Training Loss: 1.5944, Training Accuracy: 87.6964%, Validation Loss: 1.590
0, Validation Accuracy: 87.1548%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 437.96it/s]
Epoch [5/20], Training Loss: 1.5904, Training Accuracy: 88.0893%, Validation Loss: 1.587
9, Validation Accuracy: 87.3095%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 434.09it/s]
Epoch [6/20], Training Loss: 1.5878, Training Accuracy: 88.2946%, Validation Loss: 1.584
7, Validation Accuracy: 87.5952%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 420.25it/s]
Epoch [7/20], Training Loss: 1.5487, Training Accuracy: 92.8988%, Validation Loss: 1.509
4, Validation Accuracy: 95.2500%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 456.75it/s]
Epoch [8/20], Training Loss: 1.5127, Training Accuracy: 96.2619%, Validation Loss: 1.503
5, Validation Accuracy: 95.8214%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 447.40it/s]
Epoch [9/20], Training Loss: 1.5066, Training Accuracy: 96.6488%, Validation Loss: 1.500
4, Validation Accuracy: 96.1071%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 444.71it/s]
Epoch [10/20], Training Loss: 1.5043, Training Accuracy: 96.8631%, Validation Loss: 1.50
08, Validation Accuracy: 96.0833%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 449.35it/s]
Epoch [11/20], Training Loss: 1.5026, Training Accuracy: 96.9792%, Validation Loss: 1.49
70, Validation Accuracy: 96.4762%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 461.93it/s]
Epoch [12/20], Training Loss: 1.4989, Training Accuracy: 97.2054%, Validation Loss: 1.49
71, Validation Accuracy: 96.3810%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 488.09it/s]
Epoch [13/20], Training Loss: 1.4954, Training Accuracy: 97.4345%, Validation Loss: 1.49
97, Validation Accuracy: 96.1905%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 473.71it/s]
Epoch [14/20], Training Loss: 1.4932, Training Accuracy: 97.5893%, Validation Loss: 1.49
53, Validation Accuracy: 96.6071%
Training: 100%|██████████████████████████████████████████████████████████| 420
<00:00, 509.82it/s]
```

```
Epoch [15/20], Training Loss: 1.4933, Training Accuracy: 97.5982%, Validation Loss: 1.49
34, Validation Accuracy: 96.8095%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 509.31it/s]
Epoch [16/20], Training Loss: 1.4921, Training Accuracy: 97.6399%, Validation Loss: 1.49
27, Validation Accuracy: 96.9167%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 509.85it/s]
Epoch [17/20], Training Loss: 1.4895, Training Accuracy: 97.8601%, Validation Loss: 1.49
28, Validation Accuracy: 96.9524%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 506.42it/s]
Epoch [18/20], Training Loss: 1.4898, Training Accuracy: 97.9315%, Validation Loss: 1.49
38, Validation Accuracy: 96.7976%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 510.69it/s]
Epoch [19/20], Training Loss: 1.4893, Training Accuracy: 97.8512%, Validation Loss: 1.49
46, Validation Accuracy: 96.6905%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 510.69it/s]
Epoch [20/20], Training Loss: 1.4877, Training Accuracy: 97.9018%, Validation Loss: 1.49
05, Validation Accuracy: 97.1190%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 475.31it/s]
Epoch [1/20], Training Loss: 1.7014, Training Accuracy: 81.0595%, Validation Loss: 1.613
6, Validation Accuracy: 84.6548%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 480.11it/s]
Epoch [2/20], Training Loss: 1.5566, Training Accuracy: 94.0208%, Validation Loss: 1.530
1, Validation Accuracy: 93.4524%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 476.90it/s]
Epoch [3/20], Training Loss: 1.5290, Training Accuracy: 95.4315%, Validation Loss: 1.518
4, Validation Accuracy: 94.4762%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 481.64it/s]
Epoch [4/20], Training Loss: 1.5204, Training Accuracy: 95.7857%, Validation Loss: 1.514
7, Validation Accuracy: 94.8571%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 473.84it/s]
Epoch [5/20], Training Loss: 1.5136, Training Accuracy: 96.3125%, Validation Loss: 1.503
4, Validation Accuracy: 95.8810%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 480.83it/s]
Epoch [6/20], Training Loss: 1.5088, Training Accuracy: 96.6488%, Validation Loss: 1.505
0, Validation Accuracy: 95.6905%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 477.04it/s]
Epoch [7/20], Training Loss: 1.5058, Training Accuracy: 96.6518%, Validation Loss: 1.501
2, Validation Accuracy: 96.1548%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 479.16it/s]
Epoch [8/20], Training Loss: 1.5006, Training Accuracy: 97.1220%, Validation Loss: 1.499
1, Validation Accuracy: 96.2024%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 484.50it/s]
Epoch [9/20], Training Loss: 1.4993, Training Accuracy: 97.2411%, Validation Loss: 1.496
6, Validation Accuracy: 96.5476%
Training: 100%|██████████████████████████████████████████████████████████| 210
<00:00, 475.07it/s]
```

```
Epoch [10/20], Training Loss: 1.4973, Training Accuracy: 97.4613%, Validation Loss: 1.49
70, Validation Accuracy: 96.4405%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 482.47it/s]
Epoch [11/20], Training Loss: 1.4952, Training Accuracy: 97.6280%, Validation Loss: 1.49
58, Validation Accuracy: 96.5119%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 480.16it/s]
Epoch [12/20], Training Loss: 1.4937, Training Accuracy: 97.6875%, Validation Loss: 1.49
33, Validation Accuracy: 96.8333%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 478.61it/s]
Epoch [13/20], Training Loss: 1.4924, Training Accuracy: 97.7917%, Validation Loss: 1.49
35, Validation Accuracy: 96.7857%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 452.14it/s]
Epoch [14/20], Training Loss: 1.4902, Training Accuracy: 97.7887%, Validation Loss: 1.49
30, Validation Accuracy: 96.8810%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 431.33it/s]
Epoch [15/20], Training Loss: 1.4890, Training Accuracy: 98.0000%, Validation Loss: 1.49
51, Validation Accuracy: 96.6429%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 444.44it/s]
Epoch [16/20], Training Loss: 1.4883, Training Accuracy: 98.0417%, Validation Loss: 1.49
22, Validation Accuracy: 96.9405%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 466.55it/s]
Epoch [17/20], Training Loss: 1.4881, Training Accuracy: 97.9345%, Validation Loss: 1.49
09, Validation Accuracy: 97.1190%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:05<00:00, 399.45it/s]
Epoch [18/20], Training Loss: 1.4880, Training Accuracy: 98.1369%, Validation Loss: 1.49
15, Validation Accuracy: 97.0238%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 458.95it/s]
Epoch [19/20], Training Loss: 1.4852, Training Accuracy: 98.1637%, Validation Loss: 1.49
07, Validation Accuracy: 97.0238%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 480.65it/s]
Epoch [20/20], Training Loss: 1.4846, Training Accuracy: 98.2827%, Validation Loss: 1.48
87, Validation Accuracy: 97.3810%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 439.60it/s]
Epoch [1/20], Training Loss: 1.7038, Training Accuracy: 82.7024%, Validation Loss: 1.613
9, Validation Accuracy: 84.9643%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 442.66it/s]
Epoch [2/20], Training Loss: 1.6166, Training Accuracy: 86.9524%, Validation Loss: 1.603
3, Validation Accuracy: 85.9405%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.12it/s]
Epoch [3/20], Training Loss: 1.6035, Training Accuracy: 87.5446%, Validation Loss: 1.595
1, Validation Accuracy: 86.5119%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.32it/s]
Epoch [4/20], Training Loss: 1.5425, Training Accuracy: 94.5685%, Validation Loss: 1.516
8, Validation Accuracy: 94.6548%
Training: 100%|██████████████████████████████████████████████████████████| 105
<00:00, 447.04it/s]
```

```
Epoch [5/20], Training Loss: 1.5219, Training Accuracy: 95.9315%, Validation Loss: 1.511
4, Validation Accuracy: 95.2024%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 441.75it/s]
Epoch [6/20], Training Loss: 1.5136, Training Accuracy: 96.3958%, Validation Loss: 1.506
2, Validation Accuracy: 95.5476%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 439.62it/s]
Epoch [7/20], Training Loss: 1.5107, Training Accuracy: 96.7321%, Validation Loss: 1.522
0, Validation Accuracy: 94.1071%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.71it/s]
Epoch [8/20], Training Loss: 1.5058, Training Accuracy: 96.9018%, Validation Loss: 1.499
2, Validation Accuracy: 96.3690%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.56it/s]
Epoch [9/20], Training Loss: 1.5023, Training Accuracy: 97.1726%, Validation Loss: 1.498
6, Validation Accuracy: 96.3452%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 439.68it/s]
Epoch [10/20], Training Loss: 1.5004, Training Accuracy: 97.2589%, Validation Loss: 1.49
93, Validation Accuracy: 96.2381%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.63it/s]
Epoch [11/20], Training Loss: 1.4986, Training Accuracy: 97.4524%, Validation Loss: 1.49
70, Validation Accuracy: 96.3929%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 433.05it/s]
Epoch [12/20], Training Loss: 1.4953, Training Accuracy: 97.6577%, Validation Loss: 1.49
53, Validation Accuracy: 96.6548%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 398.32it/s]
Epoch [13/20], Training Loss: 1.4948, Training Accuracy: 97.6637%, Validation Loss: 1.49
27, Validation Accuracy: 96.9762%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 384.05it/s]
Epoch [14/20], Training Loss: 1.4919, Training Accuracy: 97.7679%, Validation Loss: 1.49
70, Validation Accuracy: 96.4286%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 404.66it/s]
Epoch [15/20], Training Loss: 1.4909, Training Accuracy: 97.8423%, Validation Loss: 1.49
63, Validation Accuracy: 96.5476%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 408.51it/s]
Epoch [16/20], Training Loss: 1.4893, Training Accuracy: 97.8958%, Validation Loss: 1.49
25, Validation Accuracy: 96.9286%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 409.97it/s]
Epoch [17/20], Training Loss: 1.4902, Training Accuracy: 97.9881%, Validation Loss: 1.49
41, Validation Accuracy: 96.7262%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 392.10it/s]
Epoch [18/20], Training Loss: 1.4887, Training Accuracy: 98.0833%, Validation Loss: 1.49
12, Validation Accuracy: 97.1429%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 396.09it/s]
Epoch [19/20], Training Loss: 1.4871, Training Accuracy: 98.1696%, Validation Loss: 1.49
16, Validation Accuracy: 96.9881%
Training: 100%|██████████████████████████████████████████████████████████| 105
<00:00, 425.07it/s]
```

```
Epoch [20/20], Training Loss: 1.4862, Training Accuracy: 98.3304%, Validation Loss: 1.49
07, Validation Accuracy: 97.1667%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 341.65it/s]
Epoch [1/20], Training Loss: 1.7438, Training Accuracy: 80.7321%, Validation Loss: 1.618
7, Validation Accuracy: 84.8452%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 353.97it/s]
Epoch [2/20], Training Loss: 1.5981, Training Accuracy: 89.9732%, Validation Loss: 1.539
0, Validation Accuracy: 93.0238%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 356.37it/s]
Epoch [3/20], Training Loss: 1.5442, Training Accuracy: 94.9821%, Validation Loss: 1.525
1, Validation Accuracy: 94.0000%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 359.07it/s]
Epoch [4/20], Training Loss: 1.5292, Training Accuracy: 95.7113%, Validation Loss: 1.520
9, Validation Accuracy: 94.4405%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 353.79it/s]
Epoch [5/20], Training Loss: 1.5211, Training Accuracy: 96.1012%, Validation Loss: 1.510
6, Validation Accuracy: 95.3095%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 345.81it/s]
Epoch [6/20], Training Loss: 1.5164, Training Accuracy: 96.3988%, Validation Loss: 1.504
8, Validation Accuracy: 95.7976%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 318.15it/s]
Epoch [7/20], Training Loss: 1.5115, Training Accuracy: 96.6399%, Validation Loss: 1.505
3, Validation Accuracy: 95.7500%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 344.06it/s]
Epoch [8/20], Training Loss: 1.5070, Training Accuracy: 96.9405%, Validation Loss: 1.504
3, Validation Accuracy: 95.8810%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 340.74it/s]
Epoch [9/20], Training Loss: 1.5048, Training Accuracy: 97.0625%, Validation Loss: 1.500
3, Validation Accuracy: 96.2500%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 350.66it/s]
Epoch [10/20], Training Loss: 1.5022, Training Accuracy: 97.2381%, Validation Loss: 1.49
99, Validation Accuracy: 96.2619%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 342.68it/s]
Epoch [11/20], Training Loss: 1.4996, Training Accuracy: 97.3631%, Validation Loss: 1.49
79, Validation Accuracy: 96.5595%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 351.30it/s]
Epoch [12/20], Training Loss: 1.4980, Training Accuracy: 97.4196%, Validation Loss: 1.49
83, Validation Accuracy: 96.3333%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 353.18it/s]
Epoch [13/20], Training Loss: 1.4973, Training Accuracy: 97.4375%, Validation Loss: 1.49
63, Validation Accuracy: 96.5714%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 348.72it/s]
Epoch [14/20], Training Loss: 1.4944, Training Accuracy: 97.5923%, Validation Loss: 1.49
50, Validation Accuracy: 96.5833%
Training: 100%|██████████████████████████████████████████████████████████| 52
00:00, 346.96it/s]
```

```
Epoch [15/20], Training Loss: 1.4937, Training Accuracy: 97.6458%, Validation Loss: 1.49
44, Validation Accuracy: 96.7857%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 346.25it/s]
Epoch [16/20], Training Loss: 1.4924, Training Accuracy: 97.7560%, Validation Loss: 1.49
43, Validation Accuracy: 96.8333%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 325.36it/s]
Epoch [17/20], Training Loss: 1.4913, Training Accuracy: 97.8780%, Validation Loss: 1.49
24, Validation Accuracy: 97.0714%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 342.70it/s]
Epoch [18/20], Training Loss: 1.4902, Training Accuracy: 97.9464%, Validation Loss: 1.49
32, Validation Accuracy: 97.0595%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 323.29it/s]
Epoch [19/20], Training Loss: 1.4895, Training Accuracy: 98.0149%, Validation Loss: 1.49
20, Validation Accuracy: 97.0714%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 323.72it/s]
Epoch [20/20], Training Loss: 1.4880, Training Accuracy: 98.0446%, Validation Loss: 1.49
37, Validation Accuracy: 96.7857%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 253.90it/s]
Epoch [1/20], Training Loss: 1.8271, Training Accuracy: 76.5446%, Validation Loss: 1.574
6, Validation Accuracy: 91.2738%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 252.31it/s]
Epoch [2/20], Training Loss: 1.5832, Training Accuracy: 92.7798%, Validation Loss: 1.549
1, Validation Accuracy: 92.2024%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 253.67it/s]
Epoch [3/20], Training Loss: 1.5543, Training Accuracy: 94.1280%, Validation Loss: 1.529
9, Validation Accuracy: 93.7738%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 262.11it/s]
Epoch [4/20], Training Loss: 1.5405, Training Accuracy: 94.8958%, Validation Loss: 1.523
4, Validation Accuracy: 94.1667%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 257.69it/s]
Epoch [5/20], Training Loss: 1.5308, Training Accuracy: 95.3095%, Validation Loss: 1.520
8, Validation Accuracy: 94.4762%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 256.85it/s]
Epoch [6/20], Training Loss: 1.5240, Training Accuracy: 95.6994%, Validation Loss: 1.513
1, Validation Accuracy: 95.0357%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 271.33it/s]
Epoch [7/20], Training Loss: 1.5188, Training Accuracy: 96.1845%, Validation Loss: 1.508
9, Validation Accuracy: 95.4762%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 275.90it/s]
Epoch [8/20], Training Loss: 1.5143, Training Accuracy: 96.3214%, Validation Loss: 1.511
4, Validation Accuracy: 95.3214%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 272.79it/s]
Epoch [9/20], Training Loss: 1.5124, Training Accuracy: 96.5982%, Validation Loss: 1.524
0, Validation Accuracy: 93.9762%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 271.86it/s]
```

```
Epoch [10/20], Training Loss: 1.5090, Training Accuracy: 96.8542%, Validation Loss: 1.50
42, Validation Accuracy: 95.9286%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 273.24it/s]
Epoch [11/20], Training Loss: 1.5056, Training Accuracy: 96.9464%, Validation Loss: 1.50
01, Validation Accuracy: 96.2619%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 248.61it/s]
Epoch [12/20], Training Loss: 1.5036, Training Accuracy: 97.2202%, Validation Loss: 1.50
32, Validation Accuracy: 95.9524%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 257.71it/s]
Epoch [13/20], Training Loss: 1.5022, Training Accuracy: 97.2232%, Validation Loss: 1.50
12, Validation Accuracy: 96.2500%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 240.36it/s]
Epoch [14/20], Training Loss: 1.5005, Training Accuracy: 97.3601%, Validation Loss: 1.50
09, Validation Accuracy: 96.2500%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 233.97it/s]
Epoch [15/20], Training Loss: 1.4981, Training Accuracy: 97.4494%, Validation Loss: 1.49
83, Validation Accuracy: 96.5238%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 231.30it/s]
Epoch [16/20], Training Loss: 1.4978, Training Accuracy: 97.5804%, Validation Loss: 1.49
93, Validation Accuracy: 96.3333%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 232.83it/s]
Epoch [17/20], Training Loss: 1.4953, Training Accuracy: 97.6042%, Validation Loss: 1.50
16, Validation Accuracy: 96.0952%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 238.99it/s]
Epoch [18/20], Training Loss: 1.4939, Training Accuracy: 97.6756%, Validation Loss: 1.49
54, Validation Accuracy: 96.8214%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 233.21it/s]
Epoch [19/20], Training Loss: 1.4930, Training Accuracy: 97.8244%, Validation Loss: 1.49
43, Validation Accuracy: 96.8333%
Training: 100%|█████████████████████████████████████████████████████████████████████| 26
3/263 [00:01<00:00, 242.66it/s]
Epoch [20/20], Training Loss: 1.4921, Training Accuracy: 97.9167%, Validation Loss: 1.49
56, Validation Accuracy: 96.7500%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 448.08it/s]
Epoch [1/20], Training Loss: 1.6851, Training Accuracy: 80.8036%, Validation Loss: 1.617
7, Validation Accuracy: 84.2262%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:09<00:00, 464.46it/s]
Epoch [2/20], Training Loss: 1.6304, Training Accuracy: 84.8423%, Validation Loss: 1.610
0, Validation Accuracy: 85.0238%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 500.27it/s]
Epoch [3/20], Training Loss: 1.6173, Training Accuracy: 85.6458%, Validation Loss: 1.605
7, Validation Accuracy: 85.5000%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 506.58it/s]
Epoch [4/20], Training Loss: 1.6090, Training Accuracy: 86.2232%, Validation Loss: 1.611
1, Validation Accuracy: 85.0119%
Training: 100%|███████████████████████████████████████████████████████████████████| 420
<00:00, 511.46it/s]
```

```
Epoch [5/20], Training Loss: 1.6049, Training Accuracy: 86.4673%, Validation Loss: 1.600
7, Validation Accuracy: 86.0833%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 513.23it/s]
Epoch [6/20], Training Loss: 1.6011, Training Accuracy: 86.7887%, Validation Loss: 1.599
1, Validation Accuracy: 86.1548%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 512.57it/s]
Epoch [7/20], Training Loss: 1.5982, Training Accuracy: 87.1250%, Validation Loss: 1.596
2, Validation Accuracy: 86.4643%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 511.25it/s]
Epoch [8/20], Training Loss: 1.5969, Training Accuracy: 87.2827%, Validation Loss: 1.588
9, Validation Accuracy: 87.2619%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 467.31it/s]
Epoch [9/20], Training Loss: 1.5913, Training Accuracy: 87.5655%, Validation Loss: 1.592
0, Validation Accuracy: 86.8690%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 495.79it/s]
Epoch [10/20], Training Loss: 1.5907, Training Accuracy: 87.7262%, Validation Loss: 1.58
92, Validation Accuracy: 87.1548%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 502.58it/s]
Epoch [11/20], Training Loss: 1.5908, Training Accuracy: 87.6458%, Validation Loss: 1.58
86, Validation Accuracy: 87.2024%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 509.39it/s]
Epoch [12/20], Training Loss: 1.5886, Training Accuracy: 87.9226%, Validation Loss: 1.58
92, Validation Accuracy: 87.2381%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 492.84it/s]
Epoch [13/20], Training Loss: 1.5548, Training Accuracy: 91.8006%, Validation Loss: 1.52
11, Validation Accuracy: 94.0833%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 510.24it/s]
Epoch [14/20], Training Loss: 1.5177, Training Accuracy: 95.4226%, Validation Loss: 1.51
17, Validation Accuracy: 94.8810%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 509.44it/s]
Epoch [15/20], Training Loss: 1.5133, Training Accuracy: 95.7440%, Validation Loss: 1.50
80, Validation Accuracy: 95.3333%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 507.35it/s]
Epoch [16/20], Training Loss: 1.5079, Training Accuracy: 96.0863%, Validation Loss: 1.50
63, Validation Accuracy: 95.4881%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 487.48it/s]
Epoch [17/20], Training Loss: 1.5093, Training Accuracy: 96.0893%, Validation Loss: 1.50
11, Validation Accuracy: 96.0476%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 510.15it/s]
Epoch [18/20], Training Loss: 1.5071, Training Accuracy: 96.0595%, Validation Loss: 1.50
37, Validation Accuracy: 95.7500%
Training: 100%|██████████████████████████████████████████████████████████| 420
0/4200 [00:08<00:00, 490.01it/s]
Epoch [19/20], Training Loss: 1.5060, Training Accuracy: 96.0685%, Validation Loss: 1.50
99, Validation Accuracy: 95.1786%
Training: 100%|██████████████████████████████████████████████████████████| 420
<00:00, 510.02it/s]
```

```
Epoch [20/20], Training Loss: 1.5068, Training Accuracy: 96.1756%, Validation Loss: 1.50
08, Validation Accuracy: 96.0238%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 478.68it/s]
Epoch [1/20], Training Loss: 1.6237, Training Accuracy: 88.4554%, Validation Loss: 1.543
9, Validation Accuracy: 91.8333%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 483.02it/s]
Epoch [2/20], Training Loss: 1.5481, Training Accuracy: 93.5327%, Validation Loss: 1.539
5, Validation Accuracy: 92.1310%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 479.01it/s]
Epoch [3/20], Training Loss: 1.5335, Training Accuracy: 94.4494%, Validation Loss: 1.514
6, Validation Accuracy: 94.6548%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 474.49it/s]
Epoch [4/20], Training Loss: 1.5248, Training Accuracy: 95.0714%, Validation Loss: 1.509
0, Validation Accuracy: 95.2381%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 482.01it/s]
Epoch [5/20], Training Loss: 1.5181, Training Accuracy: 95.4375%, Validation Loss: 1.521
3, Validation Accuracy: 93.9643%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 482.39it/s]
Epoch [6/20], Training Loss: 1.5138, Training Accuracy: 95.7411%, Validation Loss: 1.504
3, Validation Accuracy: 95.7738%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 483.03it/s]
Epoch [7/20], Training Loss: 1.5101, Training Accuracy: 96.0536%, Validation Loss: 1.502
0, Validation Accuracy: 96.0476%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 481.36it/s]
Epoch [8/20], Training Loss: 1.5078, Training Accuracy: 96.2470%, Validation Loss: 1.504
7, Validation Accuracy: 95.7024%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 479.48it/s]
Epoch [9/20], Training Loss: 1.5066, Training Accuracy: 96.4315%, Validation Loss: 1.504
7, Validation Accuracy: 95.6310%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 481.86it/s]
Epoch [10/20], Training Loss: 1.5047, Training Accuracy: 96.3899%, Validation Loss: 1.50
01, Validation Accuracy: 96.2024%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 480.89it/s]
Epoch [11/20], Training Loss: 1.5024, Training Accuracy: 96.7857%, Validation Loss: 1.49
99, Validation Accuracy: 96.1429%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 480.65it/s]
Epoch [12/20], Training Loss: 1.4998, Training Accuracy: 96.7887%, Validation Loss: 1.51
21, Validation Accuracy: 94.8571%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 482.16it/s]
Epoch [13/20], Training Loss: 1.4999, Training Accuracy: 96.8750%, Validation Loss: 1.50
25, Validation Accuracy: 95.8214%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 481.28it/s]
Epoch [14/20], Training Loss: 1.4969, Training Accuracy: 97.1667%, Validation Loss: 1.49
92, Validation Accuracy: 96.2381%
Training: 100%|██████████████████████████████████████████████████████████| 210
<00:00, 485.12it/s]
```

```
Epoch [15/20], Training Loss: 1.4958, Training Accuracy: 97.1815%, Validation Loss: 1.50
14, Validation Accuracy: 95.9762%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 438.14it/s]
Epoch [16/20], Training Loss: 1.4959, Training Accuracy: 97.2173%, Validation Loss: 1.49
82, Validation Accuracy: 96.3214%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 463.93it/s]
Epoch [17/20], Training Loss: 1.4951, Training Accuracy: 97.2708%, Validation Loss: 1.49
79, Validation Accuracy: 96.3452%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 481.37it/s]
Epoch [18/20], Training Loss: 1.4925, Training Accuracy: 97.3780%, Validation Loss: 1.49
67, Validation Accuracy: 96.4286%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 484.53it/s]
Epoch [19/20], Training Loss: 1.4914, Training Accuracy: 97.5446%, Validation Loss: 1.49
24, Validation Accuracy: 96.9048%
Training: 100%|██████████████████████████████████████████████████████████| 210
0/2100 [00:04<00:00, 479.01it/s]
Epoch [20/20], Training Loss: 1.4913, Training Accuracy: 97.6190%, Validation Loss: 1.49
22, Validation Accuracy: 96.9167%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.86it/s]
Epoch [1/20], Training Loss: 1.6423, Training Accuracy: 86.7917%, Validation Loss: 1.555
7, Validation Accuracy: 90.8452%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.85it/s]
Epoch [2/20], Training Loss: 1.5442, Training Accuracy: 94.1012%, Validation Loss: 1.522
7, Validation Accuracy: 94.0000%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 444.35it/s]
Epoch [3/20], Training Loss: 1.5290, Training Accuracy: 94.8750%, Validation Loss: 1.523
5, Validation Accuracy: 93.9048%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.89it/s]
Epoch [4/20], Training Loss: 1.5199, Training Accuracy: 95.7262%, Validation Loss: 1.512
3, Validation Accuracy: 94.9286%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 451.01it/s]
Epoch [5/20], Training Loss: 1.5137, Training Accuracy: 96.0655%, Validation Loss: 1.506
1, Validation Accuracy: 95.5714%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 447.48it/s]
Epoch [6/20], Training Loss: 1.5098, Training Accuracy: 96.3095%, Validation Loss: 1.504
7, Validation Accuracy: 95.7024%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.30it/s]
Epoch [7/20], Training Loss: 1.5065, Training Accuracy: 96.5179%, Validation Loss: 1.506
9, Validation Accuracy: 95.5357%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 456.17it/s]
Epoch [8/20], Training Loss: 1.5035, Training Accuracy: 96.7649%, Validation Loss: 1.498
8, Validation Accuracy: 96.2262%
Training: 100%|██████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 451.55it/s]
Epoch [9/20], Training Loss: 1.5006, Training Accuracy: 96.9196%, Validation Loss: 1.500
0, Validation Accuracy: 96.1548%
Training: 100%|██████████████████████████████████████████████████████████| 105
<00:00, 449.64it/s]
```

```
Epoch [10/20], Training Loss: 1.4986, Training Accuracy: 97.0506%, Validation Loss: 1.49
73, Validation Accuracy: 96.4405%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 453.48it/s]
Epoch [11/20], Training Loss: 1.4964, Training Accuracy: 97.3482%, Validation Loss: 1.49
82, Validation Accuracy: 96.3452%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 454.42it/s]
Epoch [12/20], Training Loss: 1.4950, Training Accuracy: 97.3929%, Validation Loss: 1.49
80, Validation Accuracy: 96.2738%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 455.48it/s]
Epoch [13/20], Training Loss: 1.4939, Training Accuracy: 97.5030%, Validation Loss: 1.49
27, Validation Accuracy: 96.8810%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 446.87it/s]
Epoch [14/20], Training Loss: 1.4932, Training Accuracy: 97.5833%, Validation Loss: 1.49
21, Validation Accuracy: 96.9643%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 448.33it/s]
Epoch [15/20], Training Loss: 1.4915, Training Accuracy: 97.6786%, Validation Loss: 1.49
20, Validation Accuracy: 96.9286%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 452.84it/s]
Epoch [16/20], Training Loss: 1.4929, Training Accuracy: 97.6310%, Validation Loss: 1.50
06, Validation Accuracy: 96.1786%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 445.20it/s]
Epoch [17/20], Training Loss: 1.4901, Training Accuracy: 97.6399%, Validation Loss: 1.49
19, Validation Accuracy: 96.9881%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 458.11it/s]
Epoch [18/20], Training Loss: 1.4893, Training Accuracy: 97.7262%, Validation Loss: 1.49
09, Validation Accuracy: 97.1190%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 449.60it/s]
Epoch [19/20], Training Loss: 1.4877, Training Accuracy: 97.8929%, Validation Loss: 1.49
11, Validation Accuracy: 97.0833%
Training: 100%|████████████████████████████████████████████████████████████████████| 105
0/1050 [00:02<00:00, 454.35it/s]
Epoch [20/20], Training Loss: 1.4868, Training Accuracy: 98.0268%, Validation Loss: 1.49
16, Validation Accuracy: 96.9762%
Training: 100%|██████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 371.60it/s]
Epoch [1/20], Training Loss: 1.6857, Training Accuracy: 81.9940%, Validation Loss: 1.615
9, Validation Accuracy: 84.7143%
Training: 100%|██████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 369.58it/s]
Epoch [2/20], Training Loss: 1.5817, Training Accuracy: 90.7976%, Validation Loss: 1.531
8, Validation Accuracy: 93.1667%
Training: 100%|██████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 369.84it/s]
Epoch [3/20], Training Loss: 1.5360, Training Accuracy: 94.7321%, Validation Loss: 1.522
1, Validation Accuracy: 94.1429%
Training: 100%|██████████████████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 372.38it/s]
Epoch [4/20], Training Loss: 1.5237, Training Accuracy: 95.5476%, Validation Loss: 1.517
5, Validation Accuracy: 94.4405%
Training: 100%|██████████████████████████████████████████████████████████████████████| 52
00:00, 361.63it/s]
```

```
Epoch [5/20], Training Loss: 1.5154, Training Accuracy: 96.1726%, Validation Loss: 1.511
7, Validation Accuracy: 95.0000%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 364.23it/s]
Epoch [6/20], Training Loss: 1.5085, Training Accuracy: 96.6012%, Validation Loss: 1.511
0, Validation Accuracy: 95.0595%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 369.53it/s]
Epoch [7/20], Training Loss: 1.5068, Training Accuracy: 96.7649%, Validation Loss: 1.505
8, Validation Accuracy: 95.5595%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 366.86it/s]
Epoch [8/20], Training Loss: 1.5021, Training Accuracy: 97.0536%, Validation Loss: 1.502
1, Validation Accuracy: 96.1071%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 372.16it/s]
Epoch [9/20], Training Loss: 1.4991, Training Accuracy: 97.0893%, Validation Loss: 1.496
1, Validation Accuracy: 96.6310%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 374.55it/s]
Epoch [10/20], Training Loss: 1.4992, Training Accuracy: 97.1756%, Validation Loss: 1.50
22, Validation Accuracy: 96.0476%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 375.20it/s]
Epoch [11/20], Training Loss: 1.4960, Training Accuracy: 97.4405%, Validation Loss: 1.49
66, Validation Accuracy: 96.5238%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 374.36it/s]
Epoch [12/20], Training Loss: 1.4934, Training Accuracy: 97.5060%, Validation Loss: 1.49
49, Validation Accuracy: 96.6905%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 374.73it/s]
Epoch [13/20], Training Loss: 1.4926, Training Accuracy: 97.5804%, Validation Loss: 1.49
55, Validation Accuracy: 96.5952%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 372.21it/s]
Epoch [14/20], Training Loss: 1.4921, Training Accuracy: 97.7530%, Validation Loss: 1.49
97, Validation Accuracy: 96.1667%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 371.80it/s]
Epoch [15/20], Training Loss: 1.4912, Training Accuracy: 97.8512%, Validation Loss: 1.49
66, Validation Accuracy: 96.5357%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 372.47it/s]
Epoch [16/20], Training Loss: 1.4897, Training Accuracy: 97.8423%, Validation Loss: 1.49
19, Validation Accuracy: 96.9643%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 359.34it/s]
Epoch [17/20], Training Loss: 1.4886, Training Accuracy: 97.9702%, Validation Loss: 1.49
26, Validation Accuracy: 96.9048%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 373.60it/s]
Epoch [18/20], Training Loss: 1.4874, Training Accuracy: 98.0060%, Validation Loss: 1.49
30, Validation Accuracy: 96.9405%
Training: 100%|██████████████████████████████████████████████████████████| 52
5/525 [00:01<00:00, 376.63it/s]
Epoch [19/20], Training Loss: 1.4879, Training Accuracy: 97.9940%, Validation Loss: 1.49
04, Validation Accuracy: 97.1429%
Training: 100%|██████████████████████████████████████████████████████████| 52
00:00, 375.31it/s]
```

```
Epoch [20/20], Training Loss: 1.4857, Training Accuracy: 98.1042%, Validation Loss: 1.49
16, Validation Accuracy: 97.0119%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 288.51it/s]
Epoch [1/20], Training Loss: 1.6967, Training Accuracy: 82.1339%, Validation Loss: 1.562
4, Validation Accuracy: 90.5357%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 280.17it/s]
Epoch [2/20], Training Loss: 1.5550, Training Accuracy: 93.4583%, Validation Loss: 1.535
7, Validation Accuracy: 92.8571%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 285.81it/s]
Epoch [3/20], Training Loss: 1.5321, Training Accuracy: 94.9375%, Validation Loss: 1.541
1, Validation Accuracy: 92.2262%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 284.67it/s]
Epoch [4/20], Training Loss: 1.5229, Training Accuracy: 95.6994%, Validation Loss: 1.509
7, Validation Accuracy: 95.3690%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 288.85it/s]
Epoch [5/20], Training Loss: 1.5162, Training Accuracy: 96.1101%, Validation Loss: 1.509
3, Validation Accuracy: 95.3095%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 281.40it/s]
Epoch [6/20], Training Loss: 1.5127, Training Accuracy: 96.5060%, Validation Loss: 1.513
3, Validation Accuracy: 94.9881%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 286.76it/s]
Epoch [7/20], Training Loss: 1.5071, Training Accuracy: 96.7024%, Validation Loss: 1.513
5, Validation Accuracy: 94.8929%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 287.81it/s]
Epoch [8/20], Training Loss: 1.5025, Training Accuracy: 96.8542%, Validation Loss: 1.533
2, Validation Accuracy: 92.8810%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 286.61it/s]
Epoch [9/20], Training Loss: 1.5024, Training Accuracy: 97.0625%, Validation Loss: 1.499
2, Validation Accuracy: 96.2262%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 285.67it/s]
Epoch [10/20], Training Loss: 1.4991, Training Accuracy: 97.2381%, Validation Loss: 1.49
77, Validation Accuracy: 96.3690%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 287.38it/s]
Epoch [11/20], Training Loss: 1.4958, Training Accuracy: 97.4375%, Validation Loss: 1.49
46, Validation Accuracy: 96.7857%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 288.53it/s]
Epoch [12/20], Training Loss: 1.4966, Training Accuracy: 97.5030%, Validation Loss: 1.49
72, Validation Accuracy: 96.4762%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 286.32it/s]
Epoch [13/20], Training Loss: 1.4947, Training Accuracy: 97.6071%, Validation Loss: 1.49
57, Validation Accuracy: 96.6667%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 286.63it/s]
Epoch [14/20], Training Loss: 1.4917, Training Accuracy: 97.8125%, Validation Loss: 1.50
05, Validation Accuracy: 96.1786%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 286.41it/s]
```

```
Epoch [15/20], Training Loss: 1.4918, Training Accuracy: 97.7381%, Validation Loss: 1.49
37, Validation Accuracy: 96.8095%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 286.94it/s]
Epoch [16/20], Training Loss: 1.4908, Training Accuracy: 97.8214%, Validation Loss: 1.49
72, Validation Accuracy: 96.4762%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 285.21it/s]
Epoch [17/20], Training Loss: 1.4886, Training Accuracy: 97.9792%, Validation Loss: 1.49
24, Validation Accuracy: 96.8929%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 290.30it/s]
Epoch [18/20], Training Loss: 1.4869, Training Accuracy: 98.0595%, Validation Loss: 1.49
24, Validation Accuracy: 96.9167%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 282.09it/s]
Epoch [19/20], Training Loss: 1.4867, Training Accuracy: 98.0744%, Validation Loss: 1.49
73, Validation Accuracy: 96.4167%
Training: 100%|██████████████████████████████████████████████████████████| 26
3/263 [00:00<00:00, 291.58it/s]
Epoch [20/20], Training Loss: 1.4880, Training Accuracy: 98.0923%, Validation Loss: 1.48
82, Validation Accuracy: 97.3333%
```

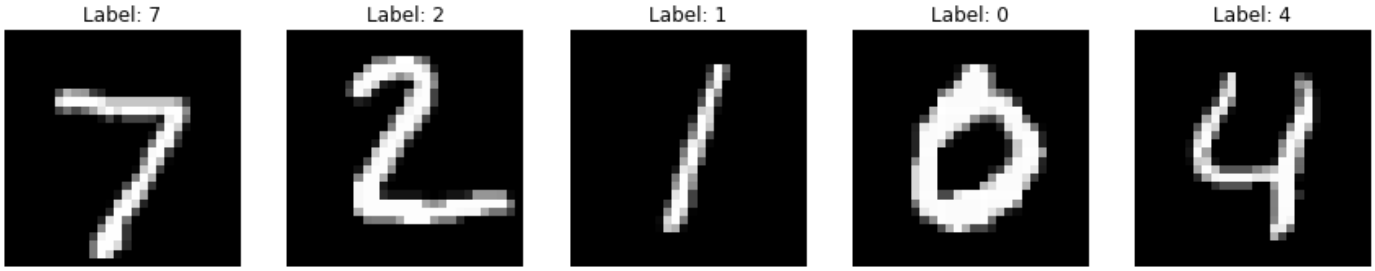
In [40]:

```python
# Load the best model parameters into the model
model.load_state_dict(best_model)

# Print the best hyperparameters
print(f"Best Hyperparameters: {best_hyperparameters}")
print(f"Validation Accuracy: {val_accuracy:.4f}%")

# Plot training and validation loss for the best model
plt.figure(figsize=(12, 6))

plt.subplot(1, 2, 1)
plt.plot(train_loss_history, label='Training Loss')
plt.plot(val_loss_history, label='Validation Loss')
plt.title('Training and Validation Loss')
plt.xlabel('Epochs')
plt.ylabel('Loss')
plt.legend()

# Plot training and validation accuracy for the best model
plt.subplot(1, 2, 2)
plt.plot(train_acc_history, label='Training Accuracy')
plt.plot(val_acc_history, label='Validation Accuracy')
plt.title('Training and Validation Accuracy')
plt.xlabel('Epochs')
plt.ylabel('Accuracy')
plt.legend()

plt.tight_layout()  # Adjust layout for better spacing
plt.show()
```

```
Best Hyperparameters: {'lr': 0.1, 'batch_size': 16}
Validation Accuracy: 97.3333%
```

Load Images to test our best model and prepare it

In [61]:
```
# Assuming you have a DataLoader for the test set named test_loader
# and a model that you want to evaluate named model

# Define the data transformation
transform = transforms.Compose([transforms.ToTensor(), transforms.Normalize((0.5,), (0.5

# Load the MNIST test set
test_set = datasets.MNIST(root='./data', train=False, download=True, transform=transform
test_loader = DataLoader(test_set, batch_size=64, shuffle=False)

# Evaluate the model on the test set
test_loss, test_accuracy = evaluate(model, test_loader)

# Visualize 5 samples from the test set
for images, labels in test_loader:
    fig, axes = plt.subplots(1, 5, figsize=(15, 3))

    for i in range(5):
        img = images[i].numpy().squeeze()
        label = labels[i].item()

        axes[i].imshow(img, cmap='gray')
        axes[i].set_title(f"Label: {label}")
        axes[i].axis('off')

    plt.show()
    break  # Break to only visualize one batch
```



In [ ]:

Function For Testing

In [62]:
```
# Evaluate the model on the test set
def evaluate(model, data_loader):
    model.eval()  # Set the model to evaluation mode
    total_loss = 0.0
    correct_predictions = 0

    with torch.no_grad():
        for images, labels in data_loader:
            # Forward pass
            predictions = model.forward(images)
            # Compute loss
            loss = model.loss(predictions, labels)
            total_loss += loss.item()

            # Count correct predictions
            correct_predictions += (model.predict(images) == labels).sum().item()

    # Calculate average loss and accuracy
    average_loss = total_loss / len(data_loader)
    accuracy = 100 * correct_predictions / len(data_loader.dataset)

    return average_loss, accuracy
```

Make Predictions with our model (test 20 From Internet (unseeeen data))

In [64]:
```
def visualize_samples_with_predictions(data_loader, num_samples=20, model=None):
    model.eval()  # Set the model to evaluation mode

    samples_visualized = 0

    for images, labels in data_loader:
        batch_size = images.size(0)
        num_samples_to_visualize = min(num_samples - samples_visualized, batch_size)
        fig, axes = plt.subplots(1, num_samples_to_visualize, figsize=(15, 3))

        for i in range(num_samples_to_visualize):
            img = images[i].numpy().squeeze()
            label = labels[i].item()

            # Make predictions using the model
            if model is not None:
                prediction = model.predict(images[i].unsqueeze(0)).item()
                title = f"True: {label}\nPred: {prediction}"
            else:
                title = f"Label: {label}"

            axes[i].imshow(img, cmap='gray')
            axes[i].set_title(title)
            axes[i].axis('off')

        plt.tight_layout()  # Adjust layout for better spacing
        plt.show()

        samples_visualized += num_samples_to_visualize
```

```
        if samples_visualized >= num_samples:
            break  # Break if the desired number of samples has been visualized


# Visualize 10 samples with model predictions and correct labels
visualize_samples_with_predictions(test_loader, num_samples=20, model=model)
```

In [66]:

```
# Evaluate the model on the test set
test_loss, test_accuracy = evaluate(model, test_loader)

# Print the test results
print(f"Test Loss: {test_loss:.4f}, Test Accuracy: {test_accuracy:.4f}%")
```

```
Test Loss: 1.7249, Test Accuracy: 73.7800%
```