dog_app

December 22, 2018

1 Convolutional Neural Networks

1.1 Project: Write an Algorithm for a Dog Identification App

In this notebook, some template code has already been provided for you, and you will need to implement additional functionality to successfully complete this project. You will not need to modify the included code beyond what is requested. Sections that begin with **'(IMPLEMENTATION)'** in the header indicate that the following block of code will require additional functionality which you must provide. Instructions will be provided for each section, and the specifics of the implementation are marked in the code block with a 'TODO' statement. Please be sure to read the instructions carefully!

> **Note**: Once you have completed all of the code implementations, you need to finalize your work by exporting the Jupyter Notebook as an HTML document. Before exporting the notebook to html, all of the code cells need to have been run so that reviewers can see the final implementation and output. You can then export the notebook by using the menu above and navigating to **File -> Download as -> HTML (.html)**. Include the finished document along with this notebook as your submission.

In addition to implementing code, there will be questions that you must answer which relate to the project and your implementation. Each section where you will answer a question is preceded by a **'Question X'** header. Carefully read each question and provide thorough answers in the following text boxes that begin with **'Answer:'**. Your project submission will be evaluated based on your answers to each of the questions and the implementation you provide.

> **Note:** Code and Markdown cells can be executed using the **Shift + Enter** keyboard shortcut. Markdown cells can be edited by double-clicking the cell to enter edit mode.

The rubric contains *optional* "Stand Out Suggestions" for enhancing the project beyond the minimum requirements. If you decide to pursue the "Stand Out Suggestions", you should include the code in this Jupyter notebook.

Step 0: Import Datasets

Make sure that you've downloaded the required human and dog datasets: * Download the dog dataset. Unzip the folder and place it in this project's home directory, at the location `/dog_images`.

- Download the human dataset. Unzip the folder and place it in the home directory, at location `/lfw`.

Note: If you are using a Windows machine, you are encouraged to use 7zip to extract the folder.

In the code cell below, we save the file paths for both the human (LFW) dataset and dog dataset in the numpy arrays `human_files` and `dog_files`.

```
In [110]: import numpy as np
          from glob import glob

          # load filenames for human and dog images
          human_files = np.array(glob("lfw/*/*"))
          dog_files = np.array(glob("dogImages/*/*/*"))

          # print number of images in each dataset
          print('There are %d total human images.' % len(human_files))
          print('There are %d total dog images.' % len(dog_files))
```

```
There are 13233 total human images.
There are 8351 total dog images.
```

Step 1: Detect Humans

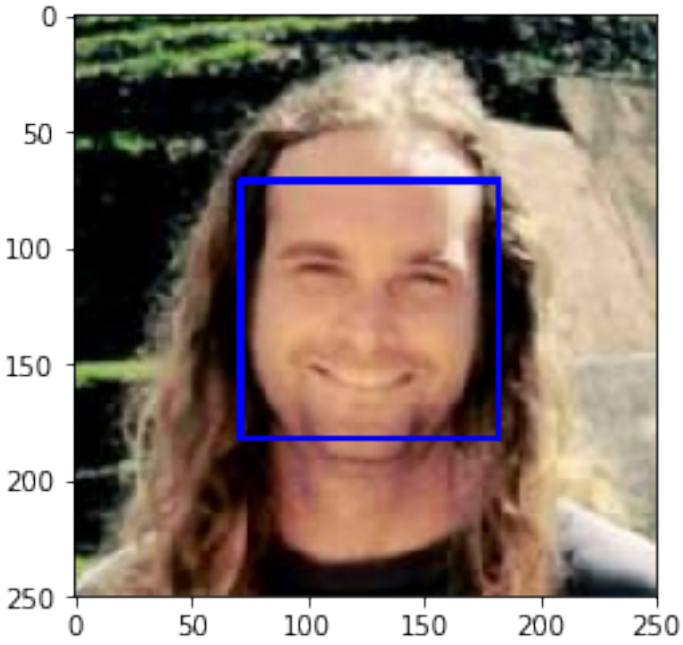
In this section, we use OpenCV's implementation of Haar feature-based cascade classifiers to detect human faces in images.

OpenCV provides many pre-trained face detectors, stored as XML files on github. We have downloaded one of these detectors and stored it in the `haarcascades` directory. In the next code cell, we demonstrate how to use this detector to find human faces in a sample image.

```
In [111]: import cv2
          import matplotlib.pyplot as plt
          %matplotlib inline

          # extract pre-trained face detector
          face_cascade = cv2.CascadeClassifier('haarcascades/haarcascade_frontalface_alt.xml')

          # load color (BGR) image
          img = cv2.imread(human_files[0])
          # convert BGR image to grayscale
          gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)

          # find faces in image
          faces = face_cascade.detectMultiScale(gray)

          # print number of faces detected in the image
          print('Number of faces detected:', len(faces))

          # get bounding box for each detected face
          for (x,y,w,h) in faces:
              # add bounding box to color image
              cv2.rectangle(img,(x,y),(x+w,y+h),(255,0,0),2)
```

```
        # convert BGR image to RGB for plotting
        cv_rgb = cv2.cvtColor(img, cv2.COLOR_BGR2RGB)

        # display the image, along with bounding box
        plt.imshow(cv_rgb)
        plt.show()
```

```
Number of faces detected: 1
```



Before using any of the face detectors, it is standard procedure to convert the images to grayscale. The `detectMultiScale` function executes the classifier stored in `face_cascade` and takes the grayscale image as a parameter.

In the above code, `faces` is a numpy array of detected faces, where each row corresponds to a detected face. Each detected face is a 1D array with four entries that specifies the bounding box of the detected face. The first two entries in the array (extracted in the above code as `x` and `y`) specify the horizontal and vertical positions of the top left corner of the bounding box. The last two entries in the array (extracted here as `w` and `h`) specify the width and height of the box.

1.1.1 Write a Human Face Detector

We can use this procedure to write a function that returns `True` if a human face is detected in an image and `False` otherwise. This function, aptly named `face_detector`, takes a string-valued file path to an image as input and appears in the code block below.

```
In [81]: import cv2
```

```
# returns "True" if face is detected in image stored at img_path
def face_detector(img_path):
    img = cv2.imread(img_path)
    gray = cv2.cvtColor(img, cv2.COLOR_BGR2GRAY)
    faces = face_cascade.detectMultiScale(gray)
    return len(faces) > 0
```

1.1.2 (IMPLEMENTATION) Assess the Human Face Detector

Question 1: Use the code cell below to test the performance of the `face_detector` function.
- What percentage of the first 100 images in `human_files` have a detected human face?
- What percentage of the first 100 images in `dog_files` have a detected human face?

Ideally, we would like 100% of human images with a detected face and 0% of dog images with a detected face. You will see that our algorithm falls short of this goal, but still gives acceptable performance. We extract the file paths for the first 100 images from each of the datasets and store them in the numpy arrays `human_files_short` and `dog_files_short`.

Answer: (You can print out your results and/or write your percentages in this cell) * Human Files: 98% detected human faces * Dog Files: 17% detected human faces

In [4]:
```
from tqdm import tqdm

human_files_short = human_files[:100]
dog_files_short = dog_files[:100]

#-#-# Do NOT modify the code above this line. #-#-#

## TODO: Test the performance of the face_detector algorithm
## on the images in human_files_short and dog_files_short.
human_count = 0
for file in human_files_short:
    human_count += face_detector(file)
print('Human Files: %d%% detected human faces' % (human_count))

human_count = 0
for file in dog_files_short:
    human_count += face_detector(file)
print('Dog Files: %d%% detected human faces' % (human_count))
```

```
Human Files: 98% detected human faces
Dog Files: 17% detected human faces
```

We suggest the face detector from OpenCV as a potential way to detect human images in your algorithm, but you are free to explore other approaches, especially approaches that make use of deep learning :). Please use the code cell below to design and test your own face detection algorithm. If you decide to pursue this *optional* task, report performance on `human_files_short` and `dog_files_short`.

```
In [5]: ### (Optional)
        ### TODO: Test performance of anotherface detection algorithm.
        ### Feel free to use as many code cells as needed.
```

Step 2: Detect Dogs

In this section, we use a pre-trained model to detect dogs in images.

1.1.3 Obtain Pre-trained VGG-16 Model

The code cell below downloads the VGG-16 model, along with weights that have been trained on ImageNet, a very large, very popular dataset used for image classification and other vision tasks. ImageNet contains over 10 million URLs, each linking to an image containing an object from one of 1000 categories.

```
In [75]: import torch
         import torchvision.models as models

         # define VGG16 model
         VGG16 = models.vgg16(pretrained=True)

         # check if CUDA is available
         use_cuda = torch.cuda.is_available()

         # move model to GPU if CUDA is available
         if use_cuda:
             VGG16 = VGG16.cuda()
```

Given an image, this pre-trained VGG-16 model returns a prediction (derived from the 1000 possible categories in ImageNet) for the object that is contained in the image.

1.1.4 (IMPLEMENTATION) Making Predictions with a Pre-trained Model

In the next code cell, you will write a function that accepts a path to an image (such as `'dogImages/train/001.Affenpinscher/Affenpinscher_00001.jpg'`) as input and returns the index corresponding to the ImageNet class that is predicted by the pre-trained VGG-16 model. The output should always be an integer between 0 and 999, inclusive.

Before writing the function, make sure that you take the time to learn how to appropriately pre-process tensors for pre-trained models in the PyTorch documentation.

```
In [14]: #VGG16
```

```
In [70]: from PIL import Image
         import torchvision.transforms as transforms
         import torch.nn.functional as F

         def load_image(img_path):
             '''
```

```
    Load an image from the given path and preprocess it for VGG16.

    Args:
        img_path: path to an image

    Returns:
        Image tensor with shape batch_size, channels, height, width (1, 3, 224, 224)
    '''
    image = Image.open(img_path).convert('RGB')
    in_transform = transforms.Compose([
                        transforms.Resize((224, 224)),
                        transforms.ToTensor(),
                        transforms.Normalize((0.485, 0.456, 0.406),
                                             (0.229, 0.224, 0.225))])
    image = in_transform(image)[:3,:,:].unsqueeze(0)
    return image

def VGG16_predict(img_path):
    '''
    Use pre-trained VGG-16 model to obtain index corresponding to
    predicted ImageNet class for image at specified path

    Args:
        img_path: path to an image

    Returns:
        Index corresponding to VGG-16 model's prediction
    '''

    ## TODO: Complete the function.
    ## Load and pre-process an image from the given img_path
    ## Return the *index* of the predicted class for that image

    # Load an image.
    image = load_image(img_path)

    if use_cuda:
        VGG16.cuda()
        image = image.cuda()

    VGG16.eval()
    output = VGG16(image)
    softmax = F.softmax(output, dim=1)
    predicted_class_index_tensor = torch.max(softmax, dim=1)[1]
    if use_cuda:
        predicted_class_index_tensor = predicted_class_index_tensor.cpu()
    predicted_class_index = predicted_class_index_tensor.numpy()[0]
    return predicted_class_index
```

```
# VGG16_predict("images/American_water_spaniel_00648.jpg")
```

1.1.5 (IMPLEMENTATION) Write a Dog Detector

While looking at the dictionary, you will notice that the categories corresponding to dogs appear in an uninterrupted sequence and correspond to dictionary keys 151-268, inclusive, to include all categories from `'Chihuahua'` to `'Mexican hairless'`. Thus, in order to check to see if an image is predicted to contain a dog by the pre-trained VGG-16 model, we need only check if the pre-trained model predicts an index between 151 and 268 (inclusive).

Use these ideas to complete the `dog_detector` function below, which returns `True` if a dog is detected in an image (and `False` if not).

```
In [66]: ### returns "True" if a dog is detected in the image stored at img_path
         def dog_detector(img_path):
             ## TODO: Complete the function.
             predicted_class_index = VGG16_predict(img_path)
             return (predicted_class_index > 150 and predicted_class_index < 269) # true/false
         # dog_detector("images/American_water_spaniel_00648.jpg")
```

1.1.6 (IMPLEMENTATION) Assess the Dog Detector

Question 2: Use the code cell below to test the performance of your `dog_detector` function.
- What percentage of the images in `human_files_short` have a detected dog?
- What percentage of the images in `dog_files_short` have a detected dog?

Answer:

- Human Files: 0% detected dog faces
- Dog Files: 100% detected dog faces

```
In [13]: ### TODO: Test the performance of the dog_detector function
         ### on the images in human_files_short and dog_files_short.
         dog_count = 0
         for file in human_files_short:
             dog_count += dog_detector(file)
         print('Human Files: %d%% detected dog faces' % (dog_count))

         dog_count = 0
         for file in dog_files_short:
             dog_count += dog_detector(file)
         print('Dog Files: %d%% detected dog faces' % (dog_count))
```

```
Human Files: 0% detected dog faces
Dog Files: 100% detected dog faces
```

We suggest VGG-16 as a potential network to detect dog images in your algorithm, but you are free to explore other pre-trained networks (such as Inception-v3, ResNet-50, etc). Please use the code cell below to test other pre-trained PyTorch models. If you decide to pursue this *optional* task, report performance on `human_files_short` and `dog_files_short`.

```
In [ ]: ### (Optional)
        ### TODO: Report the performance of another pre-trained network.
        ### Feel free to use as many code cells as needed.
```

Step 3: Create a CNN to Classify Dog Breeds (from Scratch)

Now that we have functions for detecting humans and dogs in images, we need a way to predict breed from images. In this step, you will create a CNN that classifies dog breeds. You must create your CNN *from scratch* (so, you can't use transfer learning *yet*!), and you must attain a test accuracy of at least 10%. In Step 4 of this notebook, you will have the opportunity to use transfer learning to create a CNN that attains greatly improved accuracy.

We mention that the task of assigning breed to dogs from images is considered exceptionally challenging. To see why, consider that *even a human* would have trouble distinguishing between a Brittany and a Welsh Springer Spaniel.

Brittany	Welsh Springer Spaniel

It is not difficult to find other dog breed pairs with minimal inter-class variation (for instance, Curly-Coated Retrievers and American Water Spaniels).

Curly-Coated Retriever	American Water Spaniel

Likewise, recall that labradors come in yellow, chocolate, and black. Your vision-based algorithm will have to conquer this high intra-class variation to determine how to classify all of these different shades as the same breed.

Yellow Labrador	Chocolate Labrador

We also mention that random chance presents an exceptionally low bar: setting aside the fact that the classes are slightly imabalanced, a random guess will provide a correct answer roughly 1 in 133 times, which corresponds to an accuracy of less than 1%.

Remember that the practice is far ahead of the theory in deep learning. Experiment with many different architectures, and trust your intuition. And, of course, have fun!

1.1.7 (IMPLEMENTATION) Specify Data Loaders for the Dog Dataset

Use the code cell below to write three separate data loaders for the training, validation, and test datasets of dog images (located at `dog_images/train`, `dog_images/valid`, and `dog_images/test`, respectively). You may find this documentation on custom datasets to be a useful resource. If you are interested in augmenting your training and/or validation data, check out the wide variety of transforms!

In [3]:
```
import os
from torchvision import datasets
from PIL import Image
import torchvision.transforms as transforms
import torch.nn.functional as F
import torch

from PIL import ImageFile
ImageFile.LOAD_TRUNCATED_IMAGES = True

# check if CUDA is available
use_cuda = torch.cuda.is_available()

device = torch.device("cuda:0" if use_cuda else "cpu")
print("Device:", device)

### TODO: Write data loaders for training, validation, and test sets
## Specify appropriate transforms, and batch_sizes
## batch_size is the same as that used in the orignal VGG16 paper.
def create_data_loader(dir_path, is_train=False, batch_size=256):
    if is_train:
        in_transform = transforms.Compose([
                            transforms.RandomRotation(degrees=90),
                            transforms.Resize((224, 224)),
                            transforms.RandomHorizontalFlip(p=0.5),
                            transforms.ToTensor(),
                            transforms.Normalize((0.485, 0.456, 0.406),
                                                 (0.229, 0.224, 0.225))])
    else:
        in_transform = transforms.Compose([
                            transforms.Resize((224, 224)),
                            transforms.ToTensor(),
                            transforms.Normalize((0.485, 0.456, 0.406),
                                                 (0.229, 0.224, 0.225))])
    data = datasets.ImageFolder(dir_path, transform=in_transform, target_transform=None)
    image_loader = torch.utils.data.DataLoader(data, batch_size=batch_size, shuffle=is_t
    return image_loader


def create_loaders(img_path, has_subfolders=True, batch_size=256):
    loaders_scratch = {}
    if has_subfolders:
        train_path = img_path + '/train'
        valid_path = img_path + '/valid'
        test_path = img_path + '/test'
    else:
        train_path = img_path
        valid_path = img_path
```

```
        test_path = img_path
    loaders_scratch['train'] = create_data_loader(train_path, True, batch_size=batch_siz
    loaders_scratch['valid'] = create_data_loader(valid_path, False, batch_size=batch_si
    loaders_scratch['test'] = create_data_loader(test_path, False, batch_size=batch_size
    return loaders_scratch
```

```
Device: cuda:0
```

Question 3: Describe your chosen procedure for preprocessing the data. - How does your code resize the images (by cropping, stretching, etc)? What size did you pick for the input tensor, and why? - Did you decide to augment the dataset? If so, how (through translations, flips, rotations, etc)? If not, why not?

Answer: * Stretching. 224, 224 the same size as VGG16 to compare the performances with the same resolution of photos. * Yes. Flip and rotations. I did not crop in order to make sure I didn't crop out a dog.

1.1.8 (IMPLEMENTATION) Model Architecture

Create a CNN to classify dog breed. Use the template in the code cell below.

1.1.9 More BatchNorms & Learning Rate Decay

```
In [2]: import torch.nn as nn
        import torch.nn.functional as F
        import torch.optim as optim

        # Architecture #2 - WINNER!!!! ###################################
        class Net2(nn.Module):
            ### TODO: choose an architecture, and complete the class
            def __init__(self, num_classes=133):
                super(Net2, self).__init__()
                self.features = nn.Sequential(
                    nn.Conv2d(3, 64, kernel_size=11, stride=4, padding=2),
                    nn.ReLU(inplace=True),
                    nn.MaxPool2d(kernel_size=3, stride=2),
                    nn.BatchNorm2d(64),
                    nn.Conv2d(64, 192, kernel_size=5, padding=2),
                    nn.BatchNorm2d(192),
                    nn.ReLU(inplace=True),
                    nn.MaxPool2d(kernel_size=3, stride=2),
                    nn.BatchNorm2d(192),
                    nn.Conv2d(192, 384, kernel_size=3, padding=1),
                    nn.ReLU(inplace=True),
                    nn.BatchNorm2d(384),
                    nn.Conv2d(384, 256, kernel_size=3, padding=1),
                    nn.ReLU(inplace=True),
                    nn.BatchNorm2d(256),
                    nn.Conv2d(256, 256, kernel_size=3, padding=1),
```

```
            nn.ReLU(inplace=True),
            nn.MaxPool2d(kernel_size=3, stride=2),
        )
        self.FLATTEN_SIZE = 256 * 6 * 6
        self.classifier = nn.Sequential(
            nn.Dropout(),
            nn.Linear(self.FLATTEN_SIZE, 4096),
            nn.ReLU(inplace=True),
            nn.Dropout(),
            nn.Linear(4096, 4096),
            nn.ReLU(inplace=True),
            nn.Linear(4096, num_classes),
        )

    def forward(self, x):
        ## Define forward behavior
        x = self.features(x)
        x = x.view(x.size(0), self.FLATTEN_SIZE)
        x = self.classifier(x)
        return x

# EXPERIMENT 2 ========================================================
#-#-# You do NOT have to modify the code below this line. #-#-#

# instantiate the CNN
model_scratch = Net2(133)

# move tensors to GPU if CUDA is available
if use_cuda:
    model_scratch.cuda()
```

Question 4: Outline the steps you took to get to your final CNN architecture and your reasoning at each step.

Answer:

Final architecture: Net2 above.

I first tried VGG16 architecture and learned that the model was barely learning anything. I think it's because I have too few data for such many parameters to train. Hance, I reduced the size greatly by adapting AlexNet with two extra bath normalization layers added (Experiment #1, Architecture #1 Net). I also configured my ReLU layers to perform in place because I kept running out of CUDA memory without it for the VGG16 training. With AlexNet, I saw a much faster pace of learning with the training and validation loss decreasing greatly at each early epochs. However, the model quickly overfit. So, I ended up applying batch normalization to all convolutional layer inputs except the first one in addition to decreasing the learning rate by a factor of 10 whenever validation loss did not decrease(Experiment #2, Architecture #2 - Net2). The minimum validation loss decreased. However, the model seems to quickly stop learning much. I read somewhere that momemtum doesn't really help with learning. So, I removed the momentum by setting it to the default value 0. And, the final model started to learn for a long time and ended up performing the best! (Experiment #3, Architecture #2 - Net2)

1.1.10 (IMPLEMENTATION) Specify Loss Function and Optimizer

Use the next code cell to specify a loss function and optimizer. Save the chosen loss function as `criterion_scratch`, and the optimizer as `optimizer_scratch` below.

In [6]:
```
### TODO: select loss function
criterion_scratch = nn.CrossEntropyLoss()

### TODO: select optimizer
optimizer_scratch = optim.SGD(model_scratch.parameters(), lr=0.01, weight_decay=0.0005)

if use_cuda:
    criterion_scratch = criterion_scratch.cuda()

loaders_scratch = create_loaders('dogImages', True, batch_size=32)
```

1.1.11 (IMPLEMENTATION) Train and Validate the Model

Train and validate your model in the code cell below. Save the final model parameters at filepath `'model_scratch.pt'`.

In [7]:
```
import numpy as np


def train3(n_epochs, loaders, model, optimizer, criterion, use_cuda, save_path, start_ep
    """returns trained model"""
    # initialize tracker for minimum validation loss
    valid_loss_min = np.Inf

    for epoch in range(start_epoch, n_epochs+1):
        # initialize variables to monitor training and validation loss
        train_loss = 0.0
        valid_loss = 0.0

        ##################
        # train the model #
        ##################
        model.train()
        for batch_idx, (data, target) in enumerate(loaders['train']):
            # move to GPU
            if use_cuda:
                data, target = data.cuda(), target.cuda()
            ## find the loss and update the model parameters accordingly
            ## record the average training loss, using something like
            ## train_loss = train_loss + ((1 / (batch_idx + 1)) * (loss.data - train_los
            optimizer.zero_grad()
            output = model(data)
            loss = criterion(output, target)
            loss.backward()
```

```
            optimizer.step()
            train_loss += ((1 / (batch_idx + 1)) * (loss.data - train_loss))
            if batch_idx%20 == 0:
                print('Epoch: {} \tTrain Batch: {} \tTraining Loss: {:.6f}'.format(
                    epoch,
                    batch_idx,
                    train_loss
                    ))

        ######################
        # validate the model #
        ######################
        model.eval()
        for batch_idx, (data, target) in enumerate(loaders['valid']):
            # move to GPU
            if use_cuda:
                data, target = data.cuda(), target.cuda()
            ## update the average validation loss
            valid_loss += ((1 / (batch_idx + 1)) * (loss.data - valid_loss))
            if batch_idx%20 == 0:
                print('Epoch: {} \tEval Batch: {} \tValidation Loss: {:.6f}'.format(
                    epoch,
                    batch_idx,
                    valid_loss
                    ))


        # print training/validation statistics
        print('Epoch: {} \tTraining Loss: {:.6f} \tValidation Loss: {:.6f}'.format(
            epoch,
            train_loss,
            valid_loss
            ))

        ## TODO: save the model if validation loss has decreased
        if valid_loss < valid_loss_min:
            print('Validation loss decreased ({:.6f} --> {:.6f}).  Saving model ...'.for
                valid_loss_min,
                valid_loss))
            torch.save(model_scratch.state_dict(), save_path)
            valid_loss_min = valid_loss

    # return trained model
    return model

# train the model, epoch=100
model_scratch = train3(100, loaders_scratch, model_scratch, optimizer_scratch,
                       criterion_scratch, use_cuda, 'model_scratch_v3.pt', 1)
```

```
        # load the model that got the best validation accuracy
        model_scratch.load_state_dict(torch.load('model_scratch_v3.pt'))
```

```
Epoch: 1         Train Batch: 0         Training Loss: 4.899663
Epoch: 1         Train Batch: 20         Training Loss: 4.891515
Epoch: 1         Train Batch: 40         Training Loss: 4.886625
Epoch: 1         Train Batch: 60         Training Loss: 4.885201
Epoch: 1         Train Batch: 80         Training Loss: 4.885793
Epoch: 1         Train Batch: 100         Training Loss: 4.883962
Epoch: 1         Train Batch: 120         Training Loss: 4.883111
Epoch: 1         Train Batch: 140         Training Loss: 4.881094
Epoch: 1         Train Batch: 160         Training Loss: 4.878327
Epoch: 1         Train Batch: 180         Training Loss: 4.878655
Epoch: 1         Train Batch: 200         Training Loss: 4.877970
Epoch: 1         Eval Batch: 0         Validation Loss: 4.842398
Epoch: 1         Eval Batch: 20         Validation Loss: 4.842398
Epoch: 1         Training Loss: 4.877651         Validation Loss: 4.842398
Validation loss decreased (inf --> 4.842398).  Saving model ...
Epoch: 2         Train Batch: 0         Training Loss: 4.856989
Epoch: 2         Train Batch: 20         Training Loss: 4.840892
Epoch: 2         Train Batch: 40         Training Loss: 4.836544
Epoch: 2         Train Batch: 60         Training Loss: 4.838330
Epoch: 2         Train Batch: 80         Training Loss: 4.844933
Epoch: 2         Train Batch: 100         Training Loss: 4.843767
Epoch: 2         Train Batch: 120         Training Loss: 4.847188
Epoch: 2         Train Batch: 140         Training Loss: 4.843066
Epoch: 2         Train Batch: 160         Training Loss: 4.840702
Epoch: 2         Train Batch: 180         Training Loss: 4.838713
Epoch: 2         Train Batch: 200         Training Loss: 4.834971
Epoch: 2         Eval Batch: 0         Validation Loss: 4.771731
Epoch: 2         Eval Batch: 20         Validation Loss: 4.771731
Epoch: 2         Training Loss: 4.833445         Validation Loss: 4.771731
Validation loss decreased (4.842398 --> 4.771731).  Saving model ...
Epoch: 3         Train Batch: 0         Training Loss: 4.696023
Epoch: 3         Train Batch: 20         Training Loss: 4.763996
Epoch: 3         Train Batch: 40         Training Loss: 4.774573
Epoch: 3         Train Batch: 60         Training Loss: 4.762744
Epoch: 3         Train Batch: 80         Training Loss: 4.754821
Epoch: 3         Train Batch: 100         Training Loss: 4.745842
Epoch: 3         Train Batch: 120         Training Loss: 4.738996
Epoch: 3         Train Batch: 140         Training Loss: 4.734242
Epoch: 3         Train Batch: 160         Training Loss: 4.725801
Epoch: 3         Train Batch: 180         Training Loss: 4.716602
Epoch: 3         Train Batch: 200         Training Loss: 4.711597
Epoch: 3         Eval Batch: 0         Validation Loss: 4.716656
Epoch: 3         Eval Batch: 20         Validation Loss: 4.716656
Epoch: 3         Training Loss: 4.708155         Validation Loss: 4.716656
```

```
Validation loss decreased (4.771731 --> 4.716656).  Saving model ...
Epoch: 4         Train Batch: 0         Training Loss: 4.592414
Epoch: 4         Train Batch: 20         Training Loss: 4.528234
Epoch: 4         Train Batch: 40         Training Loss: 4.542530
Epoch: 4         Train Batch: 60         Training Loss: 4.541634
Epoch: 4         Train Batch: 80         Training Loss: 4.555929
Epoch: 4         Train Batch: 100         Training Loss: 4.550210
Epoch: 4         Train Batch: 120         Training Loss: 4.530928
Epoch: 4         Train Batch: 140         Training Loss: 4.524861
Epoch: 4         Train Batch: 160         Training Loss: 4.514541
Epoch: 4         Train Batch: 180         Training Loss: 4.510349
Epoch: 4         Train Batch: 200         Training Loss: 4.500598
Epoch: 4         Eval Batch: 0         Validation Loss: 4.444483
Epoch: 4         Eval Batch: 20         Validation Loss: 4.444483
Epoch: 4         Training Loss: 4.494177         Validation Loss: 4.444483
Validation loss decreased (4.716656 --> 4.444483).  Saving model ...
Epoch: 5         Train Batch: 0         Training Loss: 4.333560
Epoch: 5         Train Batch: 20         Training Loss: 4.358593
Epoch: 5         Train Batch: 40         Training Loss: 4.373201
Epoch: 5         Train Batch: 60         Training Loss: 4.366599
Epoch: 5         Train Batch: 80         Training Loss: 4.355165
Epoch: 5         Train Batch: 100         Training Loss: 4.345672
Epoch: 5         Train Batch: 120         Training Loss: 4.336293
Epoch: 5         Train Batch: 140         Training Loss: 4.338703
Epoch: 5         Train Batch: 160         Training Loss: 4.337164
Epoch: 5         Train Batch: 180         Training Loss: 4.324296
Epoch: 5         Train Batch: 200         Training Loss: 4.318849
Epoch: 5         Eval Batch: 0         Validation Loss: 4.172998
Epoch: 5         Eval Batch: 20         Validation Loss: 4.172998
Epoch: 5         Training Loss: 4.319458         Validation Loss: 4.172998
Validation loss decreased (4.444483 --> 4.172998).  Saving model ...
Epoch: 6         Train Batch: 0         Training Loss: 4.067301
Epoch: 6         Train Batch: 20         Training Loss: 4.214815
Epoch: 6         Train Batch: 40         Training Loss: 4.204768
Epoch: 6         Train Batch: 60         Training Loss: 4.206857
Epoch: 6         Train Batch: 80         Training Loss: 4.213040
Epoch: 6         Train Batch: 100         Training Loss: 4.217400
Epoch: 6         Train Batch: 120         Training Loss: 4.224226
Epoch: 6         Train Batch: 140         Training Loss: 4.218793
Epoch: 6         Train Batch: 160         Training Loss: 4.216591
Epoch: 6         Train Batch: 180         Training Loss: 4.210334
Epoch: 6         Train Batch: 200         Training Loss: 4.206092
Epoch: 6         Eval Batch: 0         Validation Loss: 4.392428
Epoch: 6         Eval Batch: 20         Validation Loss: 4.392428
Epoch: 6         Training Loss: 4.205894         Validation Loss: 4.392428
Epoch: 7         Train Batch: 0         Training Loss: 4.073627
Epoch: 7         Train Batch: 20         Training Loss: 4.140383
Epoch: 7         Train Batch: 40         Training Loss: 4.171054
```

```
Epoch: 7         Train Batch: 60         Training Loss: 4.158741
Epoch: 7         Train Batch: 80         Training Loss: 4.145954
Epoch: 7         Train Batch: 100         Training Loss: 4.139270
Epoch: 7         Train Batch: 120         Training Loss: 4.126911
Epoch: 7         Train Batch: 140         Training Loss: 4.132399
Epoch: 7         Train Batch: 160         Training Loss: 4.136737
Epoch: 7         Train Batch: 180         Training Loss: 4.139409
Epoch: 7         Train Batch: 200         Training Loss: 4.135660
Epoch: 7         Eval Batch: 0         Validation Loss: 4.324346
Epoch: 7         Eval Batch: 20         Validation Loss: 4.324346
Epoch: 7         Training Loss: 4.136436         Validation Loss: 4.324346
Epoch: 8         Train Batch: 0         Training Loss: 3.888736
Epoch: 8         Train Batch: 20         Training Loss: 4.014826
Epoch: 8         Train Batch: 40         Training Loss: 4.035375
Epoch: 8         Train Batch: 60         Training Loss: 4.080679
Epoch: 8         Train Batch: 80         Training Loss: 4.075054
Epoch: 8         Train Batch: 100         Training Loss: 4.083127
Epoch: 8         Train Batch: 120         Training Loss: 4.081433
Epoch: 8         Train Batch: 140         Training Loss: 4.081059
Epoch: 8         Train Batch: 160         Training Loss: 4.068830
Epoch: 8         Train Batch: 180         Training Loss: 4.062592
Epoch: 8         Train Batch: 200         Training Loss: 4.061621
Epoch: 8         Eval Batch: 0         Validation Loss: 4.133974
Epoch: 8         Eval Batch: 20         Validation Loss: 4.133974
Epoch: 8         Training Loss: 4.065060         Validation Loss: 4.133974
Validation loss decreased (4.172998 --> 4.133974).  Saving model ...
Epoch: 9         Train Batch: 0         Training Loss: 3.960649
Epoch: 9         Train Batch: 20         Training Loss: 3.984038
Epoch: 9         Train Batch: 40         Training Loss: 3.986036
Epoch: 9         Train Batch: 60         Training Loss: 3.996808
Epoch: 9         Train Batch: 80         Training Loss: 3.995521
Epoch: 9         Train Batch: 100         Training Loss: 3.990098
Epoch: 9         Train Batch: 120         Training Loss: 3.991154
Epoch: 9         Train Batch: 140         Training Loss: 3.990408
Epoch: 9         Train Batch: 160         Training Loss: 3.991417
Epoch: 9         Train Batch: 180         Training Loss: 3.988423
Epoch: 9         Train Batch: 200         Training Loss: 3.988159
Epoch: 9         Eval Batch: 0         Validation Loss: 4.109009
Epoch: 9         Eval Batch: 20         Validation Loss: 4.109009
Epoch: 9         Training Loss: 3.983894         Validation Loss: 4.109009
Validation loss decreased (4.133974 --> 4.109009).  Saving model ...
Epoch: 10         Train Batch: 0         Training Loss: 3.905204
Epoch: 10         Train Batch: 20         Training Loss: 3.953300
Epoch: 10         Train Batch: 40         Training Loss: 3.930944
Epoch: 10         Train Batch: 60         Training Loss: 3.908064
Epoch: 10         Train Batch: 80         Training Loss: 3.906992
Epoch: 10         Train Batch: 100         Training Loss: 3.917348
Epoch: 10         Train Batch: 120         Training Loss: 3.918455
```

```
Epoch: 10         Train Batch: 140         Training Loss: 3.913794
Epoch: 10         Train Batch: 160         Training Loss: 3.914304
Epoch: 10         Train Batch: 180         Training Loss: 3.918154
Epoch: 10         Train Batch: 200         Training Loss: 3.916109
Epoch: 10         Eval Batch: 0         Validation Loss: 3.986746
Epoch: 10         Eval Batch: 20         Validation Loss: 3.986746
Epoch: 10         Training Loss: 3.918339         Validation Loss: 3.986746
Validation loss decreased (4.109009 --> 3.986746).  Saving model ...
Epoch: 11         Train Batch: 0         Training Loss: 3.878609
Epoch: 11         Train Batch: 20         Training Loss: 3.862339
Epoch: 11         Train Batch: 40         Training Loss: 3.830492
Epoch: 11         Train Batch: 60         Training Loss: 3.861417
Epoch: 11         Train Batch: 80         Training Loss: 3.858766
Epoch: 11         Train Batch: 100         Training Loss: 3.850161
Epoch: 11         Train Batch: 120         Training Loss: 3.866227
Epoch: 11         Train Batch: 140         Training Loss: 3.862621
Epoch: 11         Train Batch: 160         Training Loss: 3.862329
Epoch: 11         Train Batch: 180         Training Loss: 3.860484
Epoch: 11         Train Batch: 200         Training Loss: 3.850270
Epoch: 11         Eval Batch: 0         Validation Loss: 3.860701
Epoch: 11         Eval Batch: 20         Validation Loss: 3.860701
Epoch: 11         Training Loss: 3.848233         Validation Loss: 3.860701
Validation loss decreased (3.986746 --> 3.860701).  Saving model ...
Epoch: 12         Train Batch: 0         Training Loss: 4.094330
Epoch: 12         Train Batch: 20         Training Loss: 3.820455
Epoch: 12         Train Batch: 40         Training Loss: 3.780475
Epoch: 12         Train Batch: 60         Training Loss: 3.789708
Epoch: 12         Train Batch: 80         Training Loss: 3.788655
Epoch: 12         Train Batch: 100         Training Loss: 3.790237
Epoch: 12         Train Batch: 120         Training Loss: 3.802256
Epoch: 12         Train Batch: 140         Training Loss: 3.799881
Epoch: 12         Train Batch: 160         Training Loss: 3.786175
Epoch: 12         Train Batch: 180         Training Loss: 3.791195
Epoch: 12         Train Batch: 200         Training Loss: 3.784578
Epoch: 12         Eval Batch: 0         Validation Loss: 3.770950
Epoch: 12         Eval Batch: 20         Validation Loss: 3.770950
Epoch: 12         Training Loss: 3.783393         Validation Loss: 3.770950
Validation loss decreased (3.860701 --> 3.770950).  Saving model ...
Epoch: 13         Train Batch: 0         Training Loss: 3.787499
Epoch: 13         Train Batch: 20         Training Loss: 3.687747
Epoch: 13         Train Batch: 40         Training Loss: 3.717529
Epoch: 13         Train Batch: 60         Training Loss: 3.708595
Epoch: 13         Train Batch: 80         Training Loss: 3.720352
Epoch: 13         Train Batch: 100         Training Loss: 3.739547
Epoch: 13         Train Batch: 120         Training Loss: 3.737522
Epoch: 13         Train Batch: 140         Training Loss: 3.729813
Epoch: 13         Train Batch: 160         Training Loss: 3.720307
Epoch: 13         Train Batch: 180         Training Loss: 3.730067
```

```
Epoch: 13         Train Batch: 200         Training Loss: 3.733501
Epoch: 13         Eval Batch: 0         Validation Loss: 3.548570
Epoch: 13         Eval Batch: 20         Validation Loss: 3.548570
Epoch: 13         Training Loss: 3.730864         Validation Loss: 3.548570
Validation loss decreased (3.770950 --> 3.548570).  Saving model ...
Epoch: 14         Train Batch: 0         Training Loss: 3.792857
Epoch: 14         Train Batch: 20         Training Loss: 3.631748
Epoch: 14         Train Batch: 40         Training Loss: 3.648624
Epoch: 14         Train Batch: 60         Training Loss: 3.672834
Epoch: 14         Train Batch: 80         Training Loss: 3.669865
Epoch: 14         Train Batch: 100         Training Loss: 3.670501
Epoch: 14         Train Batch: 120         Training Loss: 3.673579
Epoch: 14         Train Batch: 140         Training Loss: 3.673194
Epoch: 14         Train Batch: 160         Training Loss: 3.669475
Epoch: 14         Train Batch: 180         Training Loss: 3.669269
Epoch: 14         Train Batch: 200         Training Loss: 3.664168
Epoch: 14         Eval Batch: 0         Validation Loss: 3.219124
Epoch: 14         Eval Batch: 20         Validation Loss: 3.219124
Epoch: 14         Training Loss: 3.664642         Validation Loss: 3.219124
Validation loss decreased (3.548570 --> 3.219124).  Saving model ...
Epoch: 15         Train Batch: 0         Training Loss: 3.618855
Epoch: 15         Train Batch: 20         Training Loss: 3.634901
Epoch: 15         Train Batch: 40         Training Loss: 3.600029
Epoch: 15         Train Batch: 60         Training Loss: 3.613205
Epoch: 15         Train Batch: 80         Training Loss: 3.594705
Epoch: 15         Train Batch: 100         Training Loss: 3.595377
Epoch: 15         Train Batch: 120         Training Loss: 3.597161
Epoch: 15         Train Batch: 140         Training Loss: 3.600609
Epoch: 15         Train Batch: 160         Training Loss: 3.597186
Epoch: 15         Train Batch: 180         Training Loss: 3.592138
Epoch: 15         Train Batch: 200         Training Loss: 3.598974
Epoch: 15         Eval Batch: 0         Validation Loss: 3.594195
Epoch: 15         Eval Batch: 20         Validation Loss: 3.594195
Epoch: 15         Training Loss: 3.596214         Validation Loss: 3.594195
Epoch: 16         Train Batch: 0         Training Loss: 3.558610
Epoch: 16         Train Batch: 20         Training Loss: 3.487426
Epoch: 16         Train Batch: 40         Training Loss: 3.552055
Epoch: 16         Train Batch: 60         Training Loss: 3.558195
Epoch: 16         Train Batch: 80         Training Loss: 3.572336
Epoch: 16         Train Batch: 100         Training Loss: 3.576737
Epoch: 16         Train Batch: 120         Training Loss: 3.562464
Epoch: 16         Train Batch: 140         Training Loss: 3.555259
Epoch: 16         Train Batch: 160         Training Loss: 3.541678
Epoch: 16         Train Batch: 180         Training Loss: 3.548453
Epoch: 16         Train Batch: 200         Training Loss: 3.549680
Epoch: 16         Eval Batch: 0         Validation Loss: 3.427646
Epoch: 16         Eval Batch: 20         Validation Loss: 3.427646
Epoch: 16         Training Loss: 3.550183         Validation Loss: 3.427646
```

```
Epoch: 17         Train Batch: 0         Training Loss: 3.280080
Epoch: 17         Train Batch: 20         Training Loss: 3.472452
Epoch: 17         Train Batch: 40         Training Loss: 3.475152
Epoch: 17         Train Batch: 60         Training Loss: 3.505761
Epoch: 17         Train Batch: 80         Training Loss: 3.514350
Epoch: 17         Train Batch: 100         Training Loss: 3.491560
Epoch: 17         Train Batch: 120         Training Loss: 3.496387
Epoch: 17         Train Batch: 140         Training Loss: 3.485182
Epoch: 17         Train Batch: 160         Training Loss: 3.477576
Epoch: 17         Train Batch: 180         Training Loss: 3.486536
Epoch: 17         Train Batch: 200         Training Loss: 3.493201
Epoch: 17         Eval Batch: 0         Validation Loss: 3.585623
Epoch: 17         Eval Batch: 20         Validation Loss: 3.585623
Epoch: 17         Training Loss: 3.489117         Validation Loss: 3.585623
Epoch: 18         Train Batch: 0         Training Loss: 3.323648
Epoch: 18         Train Batch: 20         Training Loss: 3.478477
Epoch: 18         Train Batch: 40         Training Loss: 3.459556
Epoch: 18         Train Batch: 60         Training Loss: 3.428077
Epoch: 18         Train Batch: 80         Training Loss: 3.429596
Epoch: 18         Train Batch: 100         Training Loss: 3.420693
Epoch: 18         Train Batch: 120         Training Loss: 3.405659
Epoch: 18         Train Batch: 140         Training Loss: 3.417658
Epoch: 18         Train Batch: 160         Training Loss: 3.420790
Epoch: 18         Train Batch: 180         Training Loss: 3.422300
Epoch: 18         Train Batch: 200         Training Loss: 3.428287
Epoch: 18         Eval Batch: 0         Validation Loss: 3.173912
Epoch: 18         Eval Batch: 20         Validation Loss: 3.173912
Epoch: 18         Training Loss: 3.427738         Validation Loss: 3.173912
Validation loss decreased (3.219124 --> 3.173912).  Saving model ...
Epoch: 19         Train Batch: 0         Training Loss: 3.438133
Epoch: 19         Train Batch: 20         Training Loss: 3.367249
Epoch: 19         Train Batch: 40         Training Loss: 3.367041
Epoch: 19         Train Batch: 60         Training Loss: 3.381255
Epoch: 19         Train Batch: 80         Training Loss: 3.390662
Epoch: 19         Train Batch: 100         Training Loss: 3.394500
Epoch: 19         Train Batch: 120         Training Loss: 3.390705
Epoch: 19         Train Batch: 140         Training Loss: 3.377002
Epoch: 19         Train Batch: 160         Training Loss: 3.374961
Epoch: 19         Train Batch: 180         Training Loss: 3.374076
Epoch: 19         Train Batch: 200         Training Loss: 3.377844
Epoch: 19         Eval Batch: 0         Validation Loss: 3.456673
Epoch: 19         Eval Batch: 20         Validation Loss: 3.456673
Epoch: 19         Training Loss: 3.372846         Validation Loss: 3.456673
Epoch: 20         Train Batch: 0         Training Loss: 3.388928
Epoch: 20         Train Batch: 20         Training Loss: 3.313589
Epoch: 20         Train Batch: 40         Training Loss: 3.284246
Epoch: 20         Train Batch: 60         Training Loss: 3.294486
Epoch: 20         Train Batch: 80         Training Loss: 3.304503
```

```
Epoch: 20         Train Batch: 100         Training Loss: 3.305977
Epoch: 20         Train Batch: 120         Training Loss: 3.297347
Epoch: 20         Train Batch: 140         Training Loss: 3.308880
Epoch: 20         Train Batch: 160         Training Loss: 3.312138
Epoch: 20         Train Batch: 180         Training Loss: 3.317299
Epoch: 20         Train Batch: 200         Training Loss: 3.319973
Epoch: 20         Eval Batch: 0         Validation Loss: 2.958544
Epoch: 20         Eval Batch: 20         Validation Loss: 2.958544
Epoch: 20         Training Loss: 3.318333         Validation Loss: 2.958544
Validation loss decreased (3.173912 --> 2.958544).  Saving model ...
Epoch: 21         Train Batch: 0         Training Loss: 3.628781
Epoch: 21         Train Batch: 20         Training Loss: 3.291600
Epoch: 21         Train Batch: 40         Training Loss: 3.328201
Epoch: 21         Train Batch: 60         Training Loss: 3.281039
Epoch: 21         Train Batch: 80         Training Loss: 3.255103
Epoch: 21         Train Batch: 100         Training Loss: 3.251796
Epoch: 21         Train Batch: 120         Training Loss: 3.254513
Epoch: 21         Train Batch: 140         Training Loss: 3.241460
Epoch: 21         Train Batch: 160         Training Loss: 3.253559
Epoch: 21         Train Batch: 180         Training Loss: 3.256999
Epoch: 21         Train Batch: 200         Training Loss: 3.263471
Epoch: 21         Eval Batch: 0         Validation Loss: 3.261299
Epoch: 21         Eval Batch: 20         Validation Loss: 3.261299
Epoch: 21         Training Loss: 3.261100         Validation Loss: 3.261299
Epoch: 22         Train Batch: 0         Training Loss: 3.286849
Epoch: 22         Train Batch: 20         Training Loss: 3.201786
Epoch: 22         Train Batch: 40         Training Loss: 3.223641
Epoch: 22         Train Batch: 60         Training Loss: 3.224791
Epoch: 22         Train Batch: 80         Training Loss: 3.193921
Epoch: 22         Train Batch: 100         Training Loss: 3.197701
Epoch: 22         Train Batch: 120         Training Loss: 3.211090
Epoch: 22         Train Batch: 140         Training Loss: 3.208796
Epoch: 22         Train Batch: 160         Training Loss: 3.214972
Epoch: 22         Train Batch: 180         Training Loss: 3.206065
Epoch: 22         Train Batch: 200         Training Loss: 3.202722
Epoch: 22         Eval Batch: 0         Validation Loss: 2.913784
Epoch: 22         Eval Batch: 20         Validation Loss: 2.913784
Epoch: 22         Training Loss: 3.206239         Validation Loss: 2.913784
Validation loss decreased (2.958544 --> 2.913784).  Saving model ...
Epoch: 23         Train Batch: 0         Training Loss: 3.225441
Epoch: 23         Train Batch: 20         Training Loss: 3.225718
Epoch: 23         Train Batch: 40         Training Loss: 3.205039
Epoch: 23         Train Batch: 60         Training Loss: 3.232528
Epoch: 23         Train Batch: 80         Training Loss: 3.238429
Epoch: 23         Train Batch: 100         Training Loss: 3.219312
Epoch: 23         Train Batch: 120         Training Loss: 3.201975
Epoch: 23         Train Batch: 140         Training Loss: 3.180400
Epoch: 23         Train Batch: 160         Training Loss: 3.167419
```

```
Epoch: 23         Train Batch: 180         Training Loss: 3.172428
Epoch: 23         Train Batch: 200         Training Loss: 3.173315
Epoch: 23         Eval Batch: 0         Validation Loss: 3.190566
Epoch: 23         Eval Batch: 20         Validation Loss: 3.190566
Epoch: 23         Training Loss: 3.171112         Validation Loss: 3.190566
Epoch: 24         Train Batch: 0         Training Loss: 3.141993
Epoch: 24         Train Batch: 20         Training Loss: 3.003468
Epoch: 24         Train Batch: 40         Training Loss: 3.091479
Epoch: 24         Train Batch: 60         Training Loss: 3.102453
Epoch: 24         Train Batch: 80         Training Loss: 3.080372
Epoch: 24         Train Batch: 100         Training Loss: 3.070477
Epoch: 24         Train Batch: 120         Training Loss: 3.072298
Epoch: 24         Train Batch: 140         Training Loss: 3.073602
Epoch: 24         Train Batch: 160         Training Loss: 3.076715
Epoch: 24         Train Batch: 180         Training Loss: 3.081263
Epoch: 24         Train Batch: 200         Training Loss: 3.092008
Epoch: 24         Eval Batch: 0         Validation Loss: 3.385842
Epoch: 24         Eval Batch: 20         Validation Loss: 3.385842
Epoch: 24         Training Loss: 3.096575         Validation Loss: 3.385842
Epoch: 25         Train Batch: 0         Training Loss: 3.217824
Epoch: 25         Train Batch: 20         Training Loss: 3.085257
Epoch: 25         Train Batch: 40         Training Loss: 3.049616
Epoch: 25         Train Batch: 60         Training Loss: 3.052480
Epoch: 25         Train Batch: 80         Training Loss: 3.063970
Epoch: 25         Train Batch: 100         Training Loss: 3.060684
Epoch: 25         Train Batch: 120         Training Loss: 3.052544
Epoch: 25         Train Batch: 140         Training Loss: 3.057661
Epoch: 25         Train Batch: 160         Training Loss: 3.048506
Epoch: 25         Train Batch: 180         Training Loss: 3.053949
Epoch: 25         Train Batch: 200         Training Loss: 3.050633
Epoch: 25         Eval Batch: 0         Validation Loss: 3.031665
Epoch: 25         Eval Batch: 20         Validation Loss: 3.031665
Epoch: 25         Training Loss: 3.047563         Validation Loss: 3.031665
Epoch: 26         Train Batch: 0         Training Loss: 2.846474
Epoch: 26         Train Batch: 20         Training Loss: 2.994807
Epoch: 26         Train Batch: 40         Training Loss: 2.970671
Epoch: 26         Train Batch: 60         Training Loss: 2.989716
Epoch: 26         Train Batch: 80         Training Loss: 2.999427
Epoch: 26         Train Batch: 100         Training Loss: 2.998247
Epoch: 26         Train Batch: 120         Training Loss: 2.998875
Epoch: 26         Train Batch: 140         Training Loss: 2.984790
Epoch: 26         Train Batch: 160         Training Loss: 2.991447
Epoch: 26         Train Batch: 180         Training Loss: 2.989621
Epoch: 26         Train Batch: 200         Training Loss: 3.004874
Epoch: 26         Eval Batch: 0         Validation Loss: 2.922972
Epoch: 26         Eval Batch: 20         Validation Loss: 2.922972
Epoch: 26         Training Loss: 3.004138         Validation Loss: 2.922972
Epoch: 27         Train Batch: 0         Training Loss: 2.534635
```

```
Epoch: 27         Train Batch: 20         Training Loss: 2.983285
Epoch: 27         Train Batch: 40         Training Loss: 2.964975
Epoch: 27         Train Batch: 60         Training Loss: 2.974913
Epoch: 27         Train Batch: 80         Training Loss: 2.967849
Epoch: 27         Train Batch: 100         Training Loss: 2.956983
Epoch: 27         Train Batch: 120         Training Loss: 2.965870
Epoch: 27         Train Batch: 140         Training Loss: 2.960682
Epoch: 27         Train Batch: 160         Training Loss: 2.958625
Epoch: 27         Train Batch: 180         Training Loss: 2.962593
Epoch: 27         Train Batch: 200         Training Loss: 2.954163
Epoch: 27         Eval Batch: 0         Validation Loss: 3.136171
Epoch: 27         Eval Batch: 20         Validation Loss: 3.136171
Epoch: 27         Training Loss: 2.956776         Validation Loss: 3.136171
Epoch: 28         Train Batch: 0         Training Loss: 3.556045
Epoch: 28         Train Batch: 20         Training Loss: 2.959606
Epoch: 28         Train Batch: 40         Training Loss: 2.921069
Epoch: 28         Train Batch: 60         Training Loss: 2.927602
Epoch: 28         Train Batch: 80         Training Loss: 2.931727
Epoch: 28         Train Batch: 100         Training Loss: 2.912486
Epoch: 28         Train Batch: 120         Training Loss: 2.917942
Epoch: 28         Train Batch: 140         Training Loss: 2.917970
Epoch: 28         Train Batch: 160         Training Loss: 2.909794
Epoch: 28         Train Batch: 180         Training Loss: 2.904317
Epoch: 28         Train Batch: 200         Training Loss: 2.901920
Epoch: 28         Eval Batch: 0         Validation Loss: 3.193549
Epoch: 28         Eval Batch: 20         Validation Loss: 3.193549
Epoch: 28         Training Loss: 2.902157         Validation Loss: 3.193549
Epoch: 29         Train Batch: 0         Training Loss: 2.721670
Epoch: 29         Train Batch: 20         Training Loss: 2.871379
Epoch: 29         Train Batch: 40         Training Loss: 2.871311
Epoch: 29         Train Batch: 60         Training Loss: 2.864810
Epoch: 29         Train Batch: 80         Training Loss: 2.823049
Epoch: 29         Train Batch: 100         Training Loss: 2.833133
Epoch: 29         Train Batch: 120         Training Loss: 2.838789
Epoch: 29         Train Batch: 140         Training Loss: 2.841449
Epoch: 29         Train Batch: 160         Training Loss: 2.834870
Epoch: 29         Train Batch: 180         Training Loss: 2.850596
Epoch: 29         Train Batch: 200         Training Loss: 2.839808
Epoch: 29         Eval Batch: 0         Validation Loss: 2.174068
Epoch: 29         Eval Batch: 20         Validation Loss: 2.174068
Epoch: 29         Training Loss: 2.837600         Validation Loss: 2.174068
Validation loss decreased (2.913784 --> 2.174068).  Saving model ...
Epoch: 30         Train Batch: 0         Training Loss: 3.055353
Epoch: 30         Train Batch: 20         Training Loss: 2.715563
Epoch: 30         Train Batch: 40         Training Loss: 2.753263
Epoch: 30         Train Batch: 60         Training Loss: 2.751895
Epoch: 30         Train Batch: 80         Training Loss: 2.769481
Epoch: 30         Train Batch: 100         Training Loss: 2.791065
```

```
Epoch: 30         Train Batch: 120         Training Loss: 2.800690
Epoch: 30         Train Batch: 140         Training Loss: 2.797863
Epoch: 30         Train Batch: 160         Training Loss: 2.824770
Epoch: 30         Train Batch: 180         Training Loss: 2.819146
Epoch: 30         Train Batch: 200         Training Loss: 2.818247
Epoch: 30         Eval Batch: 0         Validation Loss: 2.649293
Epoch: 30         Eval Batch: 20         Validation Loss: 2.649293
Epoch: 30         Training Loss: 2.819884         Validation Loss: 2.649293
Epoch: 31         Train Batch: 0         Training Loss: 3.107172
Epoch: 31         Train Batch: 20         Training Loss: 2.820077
Epoch: 31         Train Batch: 40         Training Loss: 2.766672
Epoch: 31         Train Batch: 60         Training Loss: 2.789214
Epoch: 31         Train Batch: 80         Training Loss: 2.764199
Epoch: 31         Train Batch: 100         Training Loss: 2.769083
Epoch: 31         Train Batch: 120         Training Loss: 2.762064
Epoch: 31         Train Batch: 140         Training Loss: 2.763113
Epoch: 31         Train Batch: 160         Training Loss: 2.760252
Epoch: 31         Train Batch: 180         Training Loss: 2.761630
Epoch: 31         Train Batch: 200         Training Loss: 2.764630
Epoch: 31         Eval Batch: 0         Validation Loss: 2.846816
Epoch: 31         Eval Batch: 20         Validation Loss: 2.846816
Epoch: 31         Training Loss: 2.764298         Validation Loss: 2.846816
Epoch: 32         Train Batch: 0         Training Loss: 2.457906
Epoch: 32         Train Batch: 20         Training Loss: 2.640520
Epoch: 32         Train Batch: 40         Training Loss: 2.662536
Epoch: 32         Train Batch: 60         Training Loss: 2.655368
Epoch: 32         Train Batch: 80         Training Loss: 2.663697
Epoch: 32         Train Batch: 100         Training Loss: 2.683481
Epoch: 32         Train Batch: 120         Training Loss: 2.685737
Epoch: 32         Train Batch: 140         Training Loss: 2.686847
Epoch: 32         Train Batch: 160         Training Loss: 2.692573
Epoch: 32         Train Batch: 180         Training Loss: 2.705573
Epoch: 32         Train Batch: 200         Training Loss: 2.717387
Epoch: 32         Eval Batch: 0         Validation Loss: 2.672999
Epoch: 32         Eval Batch: 20         Validation Loss: 2.672999
Epoch: 32         Training Loss: 2.714322         Validation Loss: 2.672999
Epoch: 33         Train Batch: 0         Training Loss: 2.516093
Epoch: 33         Train Batch: 20         Training Loss: 2.528562
Epoch: 33         Train Batch: 40         Training Loss: 2.565861
Epoch: 33         Train Batch: 60         Training Loss: 2.608754
Epoch: 33         Train Batch: 80         Training Loss: 2.627000
Epoch: 33         Train Batch: 100         Training Loss: 2.646087
Epoch: 33         Train Batch: 120         Training Loss: 2.637311
Epoch: 33         Train Batch: 140         Training Loss: 2.642081
Epoch: 33         Train Batch: 160         Training Loss: 2.655481
Epoch: 33         Train Batch: 180         Training Loss: 2.668723
Epoch: 33         Train Batch: 200         Training Loss: 2.671740
Epoch: 33         Eval Batch: 0         Validation Loss: 2.606550
```

```
Epoch: 33         Eval Batch: 20         Validation Loss: 2.606550
Epoch: 33         Training Loss: 2.672394         Validation Loss: 2.606550
Epoch: 34         Train Batch: 0         Training Loss: 2.621407
Epoch: 34         Train Batch: 20         Training Loss: 2.590626
Epoch: 34         Train Batch: 40         Training Loss: 2.629493
Epoch: 34         Train Batch: 60         Training Loss: 2.615867
Epoch: 34         Train Batch: 80         Training Loss: 2.607555
Epoch: 34         Train Batch: 100         Training Loss: 2.599988
Epoch: 34         Train Batch: 120         Training Loss: 2.607532
Epoch: 34         Train Batch: 140         Training Loss: 2.602628
Epoch: 34         Train Batch: 160         Training Loss: 2.616272
Epoch: 34         Train Batch: 180         Training Loss: 2.607959
Epoch: 34         Train Batch: 200         Training Loss: 2.607646
Epoch: 34         Eval Batch: 0         Validation Loss: 3.409369
Epoch: 34         Eval Batch: 20         Validation Loss: 3.409369
Epoch: 34         Training Loss: 2.619268         Validation Loss: 3.409369
Epoch: 35         Train Batch: 0         Training Loss: 2.730523
Epoch: 35         Train Batch: 20         Training Loss: 2.575775
Epoch: 35         Train Batch: 40         Training Loss: 2.540020
Epoch: 35         Train Batch: 60         Training Loss: 2.541263
Epoch: 35         Train Batch: 80         Training Loss: 2.559929
Epoch: 35         Train Batch: 100         Training Loss: 2.557518
Epoch: 35         Train Batch: 120         Training Loss: 2.574762
Epoch: 35         Train Batch: 140         Training Loss: 2.574361
Epoch: 35         Train Batch: 160         Training Loss: 2.581704
Epoch: 35         Train Batch: 180         Training Loss: 2.582774
Epoch: 35         Train Batch: 200         Training Loss: 2.587787
Epoch: 35         Eval Batch: 0         Validation Loss: 3.163658
Epoch: 35         Eval Batch: 20         Validation Loss: 3.163658
Epoch: 35         Training Loss: 2.593473         Validation Loss: 3.163658
Epoch: 36         Train Batch: 0         Training Loss: 2.837674
Epoch: 36         Train Batch: 20         Training Loss: 2.586354
Epoch: 36         Train Batch: 40         Training Loss: 2.537522
Epoch: 36         Train Batch: 60         Training Loss: 2.570948
Epoch: 36         Train Batch: 80         Training Loss: 2.567358
Epoch: 36         Train Batch: 100         Training Loss: 2.569057
Epoch: 36         Train Batch: 120         Training Loss: 2.559640
Epoch: 36         Train Batch: 140         Training Loss: 2.554379
Epoch: 36         Train Batch: 160         Training Loss: 2.553132
Epoch: 36         Train Batch: 180         Training Loss: 2.550000
Epoch: 36         Train Batch: 200         Training Loss: 2.545922
Epoch: 36         Eval Batch: 0         Validation Loss: 2.870271
Epoch: 36         Eval Batch: 20         Validation Loss: 2.870271
Epoch: 36         Training Loss: 2.540539         Validation Loss: 2.870271
Epoch: 37         Train Batch: 0         Training Loss: 2.281346
Epoch: 37         Train Batch: 20         Training Loss: 2.520209
Epoch: 37         Train Batch: 40         Training Loss: 2.485043
Epoch: 37         Train Batch: 60         Training Loss: 2.485855
```

```
Epoch: 37         Train Batch: 80         Training Loss: 2.474826
Epoch: 37         Train Batch: 100         Training Loss: 2.481773
Epoch: 37         Train Batch: 120         Training Loss: 2.484239
Epoch: 37         Train Batch: 140         Training Loss: 2.482281
Epoch: 37         Train Batch: 160         Training Loss: 2.483524
Epoch: 37         Train Batch: 180         Training Loss: 2.498770
Epoch: 37         Train Batch: 200         Training Loss: 2.495593
Epoch: 37         Eval Batch: 0         Validation Loss: 2.671771
Epoch: 37         Eval Batch: 20         Validation Loss: 2.671771
Epoch: 37         Training Loss: 2.495331         Validation Loss: 2.671771
Epoch: 38         Train Batch: 0         Training Loss: 2.516426
Epoch: 38         Train Batch: 20         Training Loss: 2.408535
Epoch: 38         Train Batch: 40         Training Loss: 2.407216
Epoch: 38         Train Batch: 60         Training Loss: 2.440551
Epoch: 38         Train Batch: 80         Training Loss: 2.440892
Epoch: 38         Train Batch: 100         Training Loss: 2.439970
Epoch: 38         Train Batch: 120         Training Loss: 2.453582
Epoch: 38         Train Batch: 140         Training Loss: 2.464206
Epoch: 38         Train Batch: 160         Training Loss: 2.463252
Epoch: 38         Train Batch: 180         Training Loss: 2.467232
Epoch: 38         Train Batch: 200         Training Loss: 2.466445
Epoch: 38         Eval Batch: 0         Validation Loss: 2.610442
Epoch: 38         Eval Batch: 20         Validation Loss: 2.610442
Epoch: 38         Training Loss: 2.463968         Validation Loss: 2.610442
Epoch: 39         Train Batch: 0         Training Loss: 2.367647
Epoch: 39         Train Batch: 20         Training Loss: 2.393310
Epoch: 39         Train Batch: 40         Training Loss: 2.322592
Epoch: 39         Train Batch: 60         Training Loss: 2.336313
Epoch: 39         Train Batch: 80         Training Loss: 2.338372
Epoch: 39         Train Batch: 100         Training Loss: 2.349333
Epoch: 39         Train Batch: 120         Training Loss: 2.370491
Epoch: 39         Train Batch: 140         Training Loss: 2.378783
Epoch: 39         Train Batch: 160         Training Loss: 2.380615
Epoch: 39         Train Batch: 180         Training Loss: 2.387313
Epoch: 39         Train Batch: 200         Training Loss: 2.396908
Epoch: 39         Eval Batch: 0         Validation Loss: 2.669960
Epoch: 39         Eval Batch: 20         Validation Loss: 2.669960
Epoch: 39         Training Loss: 2.399141         Validation Loss: 2.669960
Epoch: 40         Train Batch: 0         Training Loss: 2.819732
Epoch: 40         Train Batch: 20         Training Loss: 2.366152
Epoch: 40         Train Batch: 40         Training Loss: 2.346035
Epoch: 40         Train Batch: 60         Training Loss: 2.373483
Epoch: 40         Train Batch: 80         Training Loss: 2.401136
Epoch: 40         Train Batch: 100         Training Loss: 2.379135
Epoch: 40         Train Batch: 120         Training Loss: 2.390229
Epoch: 40         Train Batch: 140         Training Loss: 2.373298
Epoch: 40         Train Batch: 160         Training Loss: 2.362965
Epoch: 40         Train Batch: 180         Training Loss: 2.357536
```

```
Epoch: 40         Train Batch: 200         Training Loss: 2.365656
Epoch: 40         Eval Batch: 0         Validation Loss: 2.454102
Epoch: 40         Eval Batch: 20         Validation Loss: 2.454102
Epoch: 40         Training Loss: 2.367312         Validation Loss: 2.454102
Epoch: 41         Train Batch: 0         Training Loss: 2.360018
Epoch: 41         Train Batch: 20         Training Loss: 2.205224
Epoch: 41         Train Batch: 40         Training Loss: 2.247833
Epoch: 41         Train Batch: 60         Training Loss: 2.252614
Epoch: 41         Train Batch: 80         Training Loss: 2.290799
Epoch: 41         Train Batch: 100         Training Loss: 2.323741
Epoch: 41         Train Batch: 120         Training Loss: 2.331513
Epoch: 41         Train Batch: 140         Training Loss: 2.339164
Epoch: 41         Train Batch: 160         Training Loss: 2.333010
Epoch: 41         Train Batch: 180         Training Loss: 2.335787
Epoch: 41         Train Batch: 200         Training Loss: 2.336782
Epoch: 41         Eval Batch: 0         Validation Loss: 2.205732
Epoch: 41         Eval Batch: 20         Validation Loss: 2.205732
Epoch: 41         Training Loss: 2.338932         Validation Loss: 2.205732
Epoch: 42         Train Batch: 0         Training Loss: 2.004151
Epoch: 42         Train Batch: 20         Training Loss: 2.247334
Epoch: 42         Train Batch: 40         Training Loss: 2.235527
Epoch: 42         Train Batch: 60         Training Loss: 2.302308
Epoch: 42         Train Batch: 80         Training Loss: 2.294987
Epoch: 42         Train Batch: 100         Training Loss: 2.300908
Epoch: 42         Train Batch: 120         Training Loss: 2.301610
Epoch: 42         Train Batch: 140         Training Loss: 2.298462
Epoch: 42         Train Batch: 160         Training Loss: 2.304933
Epoch: 42         Train Batch: 180         Training Loss: 2.287097
Epoch: 42         Train Batch: 200         Training Loss: 2.286062
Epoch: 42         Eval Batch: 0         Validation Loss: 2.065732
Epoch: 42         Eval Batch: 20         Validation Loss: 2.065732
Epoch: 42         Training Loss: 2.285068         Validation Loss: 2.065732
Validation loss decreased (2.174068 --> 2.065732).  Saving model ...
Epoch: 43         Train Batch: 0         Training Loss: 2.530864
Epoch: 43         Train Batch: 20         Training Loss: 2.181962
Epoch: 43         Train Batch: 40         Training Loss: 2.253372
Epoch: 43         Train Batch: 60         Training Loss: 2.249789
Epoch: 43         Train Batch: 80         Training Loss: 2.228464
Epoch: 43         Train Batch: 100         Training Loss: 2.240060
Epoch: 43         Train Batch: 120         Training Loss: 2.243221
Epoch: 43         Train Batch: 140         Training Loss: 2.233587
Epoch: 43         Train Batch: 160         Training Loss: 2.235390
Epoch: 43         Train Batch: 180         Training Loss: 2.231835
Epoch: 43         Train Batch: 200         Training Loss: 2.230310
Epoch: 43         Eval Batch: 0         Validation Loss: 2.679307
Epoch: 43         Eval Batch: 20         Validation Loss: 2.679307
Epoch: 43         Training Loss: 2.237393         Validation Loss: 2.679307
Epoch: 44         Train Batch: 0         Training Loss: 2.362163
```

```
Epoch: 44         Train Batch: 20         Training Loss: 2.106246
Epoch: 44         Train Batch: 40         Training Loss: 2.143342
Epoch: 44         Train Batch: 60         Training Loss: 2.202408
Epoch: 44         Train Batch: 80         Training Loss: 2.191857
Epoch: 44         Train Batch: 100         Training Loss: 2.186137
Epoch: 44         Train Batch: 120         Training Loss: 2.191085
Epoch: 44         Train Batch: 140         Training Loss: 2.192054
Epoch: 44         Train Batch: 160         Training Loss: 2.188920
Epoch: 44         Train Batch: 180         Training Loss: 2.200578
Epoch: 44         Train Batch: 200         Training Loss: 2.203408
Epoch: 44         Eval Batch: 0         Validation Loss: 2.003623
Epoch: 44         Eval Batch: 20         Validation Loss: 2.003623
Epoch: 44         Training Loss: 2.203298         Validation Loss: 2.003623
Validation loss decreased (2.065732 --> 2.003623).  Saving model ...
Epoch: 45         Train Batch: 0         Training Loss: 2.183874
Epoch: 45         Train Batch: 20         Training Loss: 2.195235
Epoch: 45         Train Batch: 40         Training Loss: 2.152864
Epoch: 45         Train Batch: 60         Training Loss: 2.176551
Epoch: 45         Train Batch: 80         Training Loss: 2.164947
Epoch: 45         Train Batch: 100         Training Loss: 2.177582
Epoch: 45         Train Batch: 120         Training Loss: 2.200602
Epoch: 45         Train Batch: 140         Training Loss: 2.196641
Epoch: 45         Train Batch: 160         Training Loss: 2.179959
Epoch: 45         Train Batch: 180         Training Loss: 2.181906
Epoch: 45         Train Batch: 200         Training Loss: 2.184926
Epoch: 45         Eval Batch: 0         Validation Loss: 2.479863
Epoch: 45         Eval Batch: 20         Validation Loss: 2.479863
Epoch: 45         Training Loss: 2.180797         Validation Loss: 2.479863
Epoch: 46         Train Batch: 0         Training Loss: 1.992740
Epoch: 46         Train Batch: 20         Training Loss: 2.090252
Epoch: 46         Train Batch: 40         Training Loss: 2.056078
Epoch: 46         Train Batch: 60         Training Loss: 2.113124
Epoch: 46         Train Batch: 80         Training Loss: 2.119809
Epoch: 46         Train Batch: 100         Training Loss: 2.144791
Epoch: 46         Train Batch: 120         Training Loss: 2.145725
Epoch: 46         Train Batch: 140         Training Loss: 2.145235
Epoch: 46         Train Batch: 160         Training Loss: 2.141841
Epoch: 46         Train Batch: 180         Training Loss: 2.146624
Epoch: 46         Train Batch: 200         Training Loss: 2.142116
Epoch: 46         Eval Batch: 0         Validation Loss: 2.095696
Epoch: 46         Eval Batch: 20         Validation Loss: 2.095696
Epoch: 46         Training Loss: 2.141480         Validation Loss: 2.095696
Epoch: 47         Train Batch: 0         Training Loss: 2.404289
Epoch: 47         Train Batch: 20         Training Loss: 1.988176
Epoch: 47         Train Batch: 40         Training Loss: 1.999087
Epoch: 47         Train Batch: 60         Training Loss: 2.028320
Epoch: 47         Train Batch: 80         Training Loss: 2.039133
Epoch: 47         Train Batch: 100         Training Loss: 2.080117
```

```
Epoch: 47         Train Batch: 120         Training Loss: 2.064376
Epoch: 47         Train Batch: 140         Training Loss: 2.079247
Epoch: 47         Train Batch: 160         Training Loss: 2.093969
Epoch: 47         Train Batch: 180         Training Loss: 2.080551
Epoch: 47         Train Batch: 200         Training Loss: 2.094330
Epoch: 47         Eval Batch: 0         Validation Loss: 2.041600
Epoch: 47         Eval Batch: 20         Validation Loss: 2.041600
Epoch: 47         Training Loss: 2.087747         Validation Loss: 2.041600
Epoch: 48         Train Batch: 0         Training Loss: 1.850800
Epoch: 48         Train Batch: 20         Training Loss: 1.908342
Epoch: 48         Train Batch: 40         Training Loss: 1.932627
Epoch: 48         Train Batch: 60         Training Loss: 1.969068
Epoch: 48         Train Batch: 80         Training Loss: 1.976946
Epoch: 48         Train Batch: 100         Training Loss: 2.007730
Epoch: 48         Train Batch: 120         Training Loss: 2.019758
Epoch: 48         Train Batch: 140         Training Loss: 2.023530
Epoch: 48         Train Batch: 160         Training Loss: 2.019892
Epoch: 48         Train Batch: 180         Training Loss: 2.028595
Epoch: 48         Train Batch: 200         Training Loss: 2.037251
Epoch: 48         Eval Batch: 0         Validation Loss: 1.611831
Epoch: 48         Eval Batch: 20         Validation Loss: 1.611831
Epoch: 48         Training Loss: 2.040464         Validation Loss: 1.611831
Validation loss decreased (2.003623 --> 1.611831).  Saving model ...
Epoch: 49         Train Batch: 0         Training Loss: 1.812545
Epoch: 49         Train Batch: 20         Training Loss: 2.012767
Epoch: 49         Train Batch: 40         Training Loss: 2.027576
Epoch: 49         Train Batch: 60         Training Loss: 2.015821
Epoch: 49         Train Batch: 80         Training Loss: 1.992160
Epoch: 49         Train Batch: 100         Training Loss: 2.000428
Epoch: 49         Train Batch: 120         Training Loss: 2.017921
Epoch: 49         Train Batch: 140         Training Loss: 2.020825
Epoch: 49         Train Batch: 160         Training Loss: 2.031028
Epoch: 49         Train Batch: 180         Training Loss: 2.035501
Epoch: 49         Train Batch: 200         Training Loss: 2.035624
Epoch: 49         Eval Batch: 0         Validation Loss: 2.169091
Epoch: 49         Eval Batch: 20         Validation Loss: 2.169091
Epoch: 49         Training Loss: 2.027004         Validation Loss: 2.169091
Epoch: 50         Train Batch: 0         Training Loss: 1.800246
Epoch: 50         Train Batch: 20         Training Loss: 1.902127
Epoch: 50         Train Batch: 40         Training Loss: 1.927509
Epoch: 50         Train Batch: 60         Training Loss: 1.981176
Epoch: 50         Train Batch: 80         Training Loss: 1.957412
Epoch: 50         Train Batch: 100         Training Loss: 1.974005
Epoch: 50         Train Batch: 120         Training Loss: 1.982147
Epoch: 50         Train Batch: 140         Training Loss: 2.003973
Epoch: 50         Train Batch: 160         Training Loss: 2.007968
Epoch: 50         Train Batch: 180         Training Loss: 2.020883
Epoch: 50         Train Batch: 200         Training Loss: 2.015539
```

```
Epoch: 50         Eval Batch: 0         Validation Loss: 1.796491
Epoch: 50         Eval Batch: 20         Validation Loss: 1.796491
Epoch: 50         Training Loss: 2.016052         Validation Loss: 1.796491
Epoch: 51         Train Batch: 0         Training Loss: 2.139799
Epoch: 51         Train Batch: 20         Training Loss: 1.906517
Epoch: 51         Train Batch: 40         Training Loss: 1.928270
Epoch: 51         Train Batch: 60         Training Loss: 1.922044
Epoch: 51         Train Batch: 80         Training Loss: 1.914392
Epoch: 51         Train Batch: 100         Training Loss: 1.930490
Epoch: 51         Train Batch: 120         Training Loss: 1.945178
Epoch: 51         Train Batch: 140         Training Loss: 1.930475
Epoch: 51         Train Batch: 160         Training Loss: 1.936983
Epoch: 51         Train Batch: 180         Training Loss: 1.947606
Epoch: 51         Train Batch: 200         Training Loss: 1.963125
Epoch: 51         Eval Batch: 0         Validation Loss: 2.601744
Epoch: 51         Eval Batch: 20         Validation Loss: 2.601744
Epoch: 51         Training Loss: 1.965051         Validation Loss: 2.601744
Epoch: 52         Train Batch: 0         Training Loss: 1.956174
Epoch: 52         Train Batch: 20         Training Loss: 1.870968
Epoch: 52         Train Batch: 40         Training Loss: 1.868709
Epoch: 52         Train Batch: 60         Training Loss: 1.884663
Epoch: 52         Train Batch: 80         Training Loss: 1.892080
Epoch: 52         Train Batch: 100         Training Loss: 1.895245
Epoch: 52         Train Batch: 120         Training Loss: 1.906241
Epoch: 52         Train Batch: 140         Training Loss: 1.922192
Epoch: 52         Train Batch: 160         Training Loss: 1.932579
Epoch: 52         Train Batch: 180         Training Loss: 1.928221
Epoch: 52         Train Batch: 200         Training Loss: 1.929104
Epoch: 52         Eval Batch: 0         Validation Loss: 2.155913
Epoch: 52         Eval Batch: 20         Validation Loss: 2.155913
Epoch: 52         Training Loss: 1.932618         Validation Loss: 2.155913
Epoch: 53         Train Batch: 0         Training Loss: 1.549233
Epoch: 53         Train Batch: 20         Training Loss: 1.872068
Epoch: 53         Train Batch: 40         Training Loss: 1.942196
Epoch: 53         Train Batch: 60         Training Loss: 1.929074
Epoch: 53         Train Batch: 80         Training Loss: 1.918425
Epoch: 53         Train Batch: 100         Training Loss: 1.874588
Epoch: 53         Train Batch: 120         Training Loss: 1.873413
Epoch: 53         Train Batch: 140         Training Loss: 1.875000
Epoch: 53         Train Batch: 160         Training Loss: 1.874123
Epoch: 53         Train Batch: 180         Training Loss: 1.892356
Epoch: 53         Train Batch: 200         Training Loss: 1.898316
Epoch: 53         Eval Batch: 0         Validation Loss: 2.248645
Epoch: 53         Eval Batch: 20         Validation Loss: 2.248645
Epoch: 53         Training Loss: 1.900057         Validation Loss: 2.248645
Epoch: 54         Train Batch: 0         Training Loss: 1.854739
Epoch: 54         Train Batch: 20         Training Loss: 1.731115
Epoch: 54         Train Batch: 40         Training Loss: 1.785097
```

```
Epoch: 54         Train Batch: 60         Training Loss: 1.852790
Epoch: 54         Train Batch: 80         Training Loss: 1.841501
Epoch: 54         Train Batch: 100         Training Loss: 1.851150
Epoch: 54         Train Batch: 120         Training Loss: 1.871811
Epoch: 54         Train Batch: 140         Training Loss: 1.870744
Epoch: 54         Train Batch: 160         Training Loss: 1.856528
Epoch: 54         Train Batch: 180         Training Loss: 1.855402
Epoch: 54         Train Batch: 200         Training Loss: 1.873502
Epoch: 54         Eval Batch: 0         Validation Loss: 2.090423
Epoch: 54         Eval Batch: 20         Validation Loss: 2.090423
Epoch: 54         Training Loss: 1.870657         Validation Loss: 2.090423
Epoch: 55         Train Batch: 0         Training Loss: 1.716195
Epoch: 55         Train Batch: 20         Training Loss: 1.742178
Epoch: 55         Train Batch: 40         Training Loss: 1.780910
Epoch: 55         Train Batch: 60         Training Loss: 1.815116
Epoch: 55         Train Batch: 80         Training Loss: 1.811785
Epoch: 55         Train Batch: 100         Training Loss: 1.835704
Epoch: 55         Train Batch: 120         Training Loss: 1.836667
Epoch: 55         Train Batch: 140         Training Loss: 1.828812
Epoch: 55         Train Batch: 160         Training Loss: 1.830898
Epoch: 55         Train Batch: 180         Training Loss: 1.840297
Epoch: 55         Train Batch: 200         Training Loss: 1.839571
Epoch: 55         Eval Batch: 0         Validation Loss: 1.870281
Epoch: 55         Eval Batch: 20         Validation Loss: 1.870281
Epoch: 55         Training Loss: 1.842552         Validation Loss: 1.870281
Epoch: 56         Train Batch: 0         Training Loss: 2.198041
Epoch: 56         Train Batch: 20         Training Loss: 1.851072
Epoch: 56         Train Batch: 40         Training Loss: 1.832671
Epoch: 56         Train Batch: 60         Training Loss: 1.791802
Epoch: 56         Train Batch: 80         Training Loss: 1.780764
Epoch: 56         Train Batch: 100         Training Loss: 1.796647
Epoch: 56         Train Batch: 120         Training Loss: 1.795089
Epoch: 56         Train Batch: 140         Training Loss: 1.786258
Epoch: 56         Train Batch: 160         Training Loss: 1.781677
Epoch: 56         Train Batch: 180         Training Loss: 1.791966
Epoch: 56         Train Batch: 200         Training Loss: 1.794953
Epoch: 56         Eval Batch: 0         Validation Loss: 2.012174
Epoch: 56         Eval Batch: 20         Validation Loss: 2.012174
Epoch: 56         Training Loss: 1.800235         Validation Loss: 2.012174
Epoch: 57         Train Batch: 0         Training Loss: 1.502751
Epoch: 57         Train Batch: 20         Training Loss: 1.728096
Epoch: 57         Train Batch: 40         Training Loss: 1.712678
Epoch: 57         Train Batch: 60         Training Loss: 1.710827
Epoch: 57         Train Batch: 80         Training Loss: 1.736980
Epoch: 57         Train Batch: 100         Training Loss: 1.724644
Epoch: 57         Train Batch: 120         Training Loss: 1.722965
Epoch: 57         Train Batch: 140         Training Loss: 1.734800
Epoch: 57         Train Batch: 160         Training Loss: 1.748172
```

```
Epoch: 57         Train Batch: 180         Training Loss: 1.749491
Epoch: 57         Train Batch: 200         Training Loss: 1.756783
Epoch: 57         Eval Batch: 0         Validation Loss: 2.127018
Epoch: 57         Eval Batch: 20         Validation Loss: 2.127018
Epoch: 57         Training Loss: 1.769032         Validation Loss: 2.127018
Epoch: 58         Train Batch: 0         Training Loss: 1.341008
Epoch: 58         Train Batch: 20         Training Loss: 1.567601
Epoch: 58         Train Batch: 40         Training Loss: 1.646345
Epoch: 58         Train Batch: 60         Training Loss: 1.662745
Epoch: 58         Train Batch: 80         Training Loss: 1.685778
Epoch: 58         Train Batch: 100         Training Loss: 1.696788
Epoch: 58         Train Batch: 120         Training Loss: 1.703884
Epoch: 58         Train Batch: 140         Training Loss: 1.712211
Epoch: 58         Train Batch: 160         Training Loss: 1.722405
Epoch: 58         Train Batch: 180         Training Loss: 1.726230
Epoch: 58         Train Batch: 200         Training Loss: 1.725687
Epoch: 58         Eval Batch: 0         Validation Loss: 1.670480
Epoch: 58         Eval Batch: 20         Validation Loss: 1.670480
Epoch: 58         Training Loss: 1.726048         Validation Loss: 1.670480
Epoch: 59         Train Batch: 0         Training Loss: 1.812288
Epoch: 59         Train Batch: 20         Training Loss: 1.686588
Epoch: 59         Train Batch: 40         Training Loss: 1.684427
Epoch: 59         Train Batch: 60         Training Loss: 1.693331
Epoch: 59         Train Batch: 80         Training Loss: 1.693218
Epoch: 59         Train Batch: 100         Training Loss: 1.702882
Epoch: 59         Train Batch: 120         Training Loss: 1.699341
Epoch: 59         Train Batch: 140         Training Loss: 1.692966
Epoch: 59         Train Batch: 160         Training Loss: 1.693896
Epoch: 59         Train Batch: 180         Training Loss: 1.687985
Epoch: 59         Train Batch: 200         Training Loss: 1.699335
Epoch: 59         Eval Batch: 0         Validation Loss: 1.837500
Epoch: 59         Eval Batch: 20         Validation Loss: 1.837500
Epoch: 59         Training Loss: 1.703432         Validation Loss: 1.837500
Epoch: 60         Train Batch: 0         Training Loss: 1.443625
Epoch: 60         Train Batch: 20         Training Loss: 1.552359
Epoch: 60         Train Batch: 40         Training Loss: 1.611624
Epoch: 60         Train Batch: 60         Training Loss: 1.619728
Epoch: 60         Train Batch: 80         Training Loss: 1.626418
Epoch: 60         Train Batch: 100         Training Loss: 1.616338
Epoch: 60         Train Batch: 120         Training Loss: 1.618392
Epoch: 60         Train Batch: 140         Training Loss: 1.629758
Epoch: 60         Train Batch: 160         Training Loss: 1.635545
Epoch: 60         Train Batch: 180         Training Loss: 1.648362
Epoch: 60         Train Batch: 200         Training Loss: 1.655759
Epoch: 60         Eval Batch: 0         Validation Loss: 2.261423
Epoch: 60         Eval Batch: 20         Validation Loss: 2.261423
Epoch: 60         Training Loss: 1.658360         Validation Loss: 2.261423
Epoch: 61         Train Batch: 0         Training Loss: 1.698982
```

```
Epoch: 61         Train Batch: 20         Training Loss: 1.510000
Epoch: 61         Train Batch: 40         Training Loss: 1.576809
Epoch: 61         Train Batch: 60         Training Loss: 1.602011
Epoch: 61         Train Batch: 80         Training Loss: 1.608919
Epoch: 61         Train Batch: 100         Training Loss: 1.610901
Epoch: 61         Train Batch: 120         Training Loss: 1.630519
Epoch: 61         Train Batch: 140         Training Loss: 1.622154
Epoch: 61         Train Batch: 160         Training Loss: 1.615447
Epoch: 61         Train Batch: 180         Training Loss: 1.633802
Epoch: 61         Train Batch: 200         Training Loss: 1.625617
Epoch: 61         Eval Batch: 0         Validation Loss: 1.466495
Epoch: 61         Eval Batch: 20         Validation Loss: 1.466495
Epoch: 61         Training Loss: 1.634671         Validation Loss: 1.466495
Validation loss decreased (1.611831 --> 1.466495).  Saving model ...
Epoch: 62         Train Batch: 0         Training Loss: 1.395077
Epoch: 62         Train Batch: 20         Training Loss: 1.578336
Epoch: 62         Train Batch: 40         Training Loss: 1.536694
Epoch: 62         Train Batch: 60         Training Loss: 1.526656
Epoch: 62         Train Batch: 80         Training Loss: 1.538136
Epoch: 62         Train Batch: 100         Training Loss: 1.545070
Epoch: 62         Train Batch: 120         Training Loss: 1.553603
Epoch: 62         Train Batch: 140         Training Loss: 1.567927
Epoch: 62         Train Batch: 160         Training Loss: 1.565653
Epoch: 62         Train Batch: 180         Training Loss: 1.571030
Epoch: 62         Train Batch: 200         Training Loss: 1.584145
Epoch: 62         Eval Batch: 0         Validation Loss: 2.244132
Epoch: 62         Eval Batch: 20         Validation Loss: 2.244132
Epoch: 62         Training Loss: 1.599662         Validation Loss: 2.244132
Epoch: 63         Train Batch: 0         Training Loss: 1.505582
Epoch: 63         Train Batch: 20         Training Loss: 1.520642
Epoch: 63         Train Batch: 40         Training Loss: 1.523924
Epoch: 63         Train Batch: 60         Training Loss: 1.515644
Epoch: 63         Train Batch: 80         Training Loss: 1.528628
Epoch: 63         Train Batch: 100         Training Loss: 1.536262
Epoch: 63         Train Batch: 120         Training Loss: 1.543700
Epoch: 63         Train Batch: 140         Training Loss: 1.556057
Epoch: 63         Train Batch: 160         Training Loss: 1.563265
Epoch: 63         Train Batch: 180         Training Loss: 1.574322
Epoch: 63         Train Batch: 200         Training Loss: 1.583099
Epoch: 63         Eval Batch: 0         Validation Loss: 1.524938
Epoch: 63         Eval Batch: 20         Validation Loss: 1.524938
Epoch: 63         Training Loss: 1.582481         Validation Loss: 1.524938
Epoch: 64         Train Batch: 0         Training Loss: 1.177104
Epoch: 64         Train Batch: 20         Training Loss: 1.468470
Epoch: 64         Train Batch: 40         Training Loss: 1.459595
Epoch: 64         Train Batch: 60         Training Loss: 1.467175
Epoch: 64         Train Batch: 80         Training Loss: 1.512415
Epoch: 64         Train Batch: 100         Training Loss: 1.515847
```

```
Epoch: 64         Train Batch: 120         Training Loss: 1.517301
Epoch: 64         Train Batch: 140         Training Loss: 1.527812
Epoch: 64         Train Batch: 160         Training Loss: 1.537652
Epoch: 64         Train Batch: 180         Training Loss: 1.528585
Epoch: 64         Train Batch: 200         Training Loss: 1.542646
Epoch: 64         Eval Batch: 0         Validation Loss: 1.298819
Epoch: 64         Eval Batch: 20         Validation Loss: 1.298819
Epoch: 64         Training Loss: 1.538820         Validation Loss: 1.298819
Validation loss decreased (1.466495 --> 1.298819).  Saving model ...
Epoch: 65         Train Batch: 0         Training Loss: 0.996263
Epoch: 65         Train Batch: 20         Training Loss: 1.482794
Epoch: 65         Train Batch: 40         Training Loss: 1.514462
Epoch: 65         Train Batch: 60         Training Loss: 1.507746
Epoch: 65         Train Batch: 80         Training Loss: 1.526173
Epoch: 65         Train Batch: 100         Training Loss: 1.508488
Epoch: 65         Train Batch: 120         Training Loss: 1.516411
Epoch: 65         Train Batch: 140         Training Loss: 1.520734
Epoch: 65         Train Batch: 160         Training Loss: 1.534874
Epoch: 65         Train Batch: 180         Training Loss: 1.526341
Epoch: 65         Train Batch: 200         Training Loss: 1.526986
Epoch: 65         Eval Batch: 0         Validation Loss: 1.028653
Epoch: 65         Eval Batch: 20         Validation Loss: 1.028653
Epoch: 65         Training Loss: 1.531628         Validation Loss: 1.028653
Validation loss decreased (1.298819 --> 1.028653).  Saving model ...
Epoch: 66         Train Batch: 0         Training Loss: 1.468082
Epoch: 66         Train Batch: 20         Training Loss: 1.424338
Epoch: 66         Train Batch: 40         Training Loss: 1.459778
Epoch: 66         Train Batch: 60         Training Loss: 1.479238
Epoch: 66         Train Batch: 80         Training Loss: 1.478020
Epoch: 66         Train Batch: 100         Training Loss: 1.460560
Epoch: 66         Train Batch: 120         Training Loss: 1.460130
Epoch: 66         Train Batch: 140         Training Loss: 1.463044
Epoch: 66         Train Batch: 160         Training Loss: 1.467835
Epoch: 66         Train Batch: 180         Training Loss: 1.469845
Epoch: 66         Train Batch: 200         Training Loss: 1.473848
Epoch: 66         Eval Batch: 0         Validation Loss: 1.790944
Epoch: 66         Eval Batch: 20         Validation Loss: 1.790944
Epoch: 66         Training Loss: 1.479407         Validation Loss: 1.790944
Epoch: 67         Train Batch: 0         Training Loss: 1.593729
Epoch: 67         Train Batch: 20         Training Loss: 1.370957
Epoch: 67         Train Batch: 40         Training Loss: 1.381799
Epoch: 67         Train Batch: 60         Training Loss: 1.399878
Epoch: 67         Train Batch: 80         Training Loss: 1.422242
Epoch: 67         Train Batch: 100         Training Loss: 1.445060
Epoch: 67         Train Batch: 120         Training Loss: 1.445807
Epoch: 67         Train Batch: 140         Training Loss: 1.449578
Epoch: 67         Train Batch: 160         Training Loss: 1.448676
Epoch: 67         Train Batch: 180         Training Loss: 1.444233
```

```
Epoch: 67         Train Batch: 200         Training Loss: 1.462866
Epoch: 67         Eval Batch: 0         Validation Loss: 1.317692
Epoch: 67         Eval Batch: 20         Validation Loss: 1.317692
Epoch: 67         Training Loss: 1.463137         Validation Loss: 1.317692
Epoch: 68         Train Batch: 0         Training Loss: 1.794554
Epoch: 68         Train Batch: 20         Training Loss: 1.416301
Epoch: 68         Train Batch: 40         Training Loss: 1.419605
Epoch: 68         Train Batch: 60         Training Loss: 1.381158
Epoch: 68         Train Batch: 80         Training Loss: 1.418254
Epoch: 68         Train Batch: 100         Training Loss: 1.406985
Epoch: 68         Train Batch: 120         Training Loss: 1.401226
Epoch: 68         Train Batch: 140         Training Loss: 1.413777
Epoch: 68         Train Batch: 160         Training Loss: 1.428839
Epoch: 68         Train Batch: 180         Training Loss: 1.434263
Epoch: 68         Train Batch: 200         Training Loss: 1.438320
Epoch: 68         Eval Batch: 0         Validation Loss: 1.516404
Epoch: 68         Eval Batch: 20         Validation Loss: 1.516404
Epoch: 68         Training Loss: 1.440259         Validation Loss: 1.516404
Epoch: 69         Train Batch: 0         Training Loss: 1.089810
Epoch: 69         Train Batch: 20         Training Loss: 1.250195
Epoch: 69         Train Batch: 40         Training Loss: 1.275052
Epoch: 69         Train Batch: 60         Training Loss: 1.321244
Epoch: 69         Train Batch: 80         Training Loss: 1.359569
Epoch: 69         Train Batch: 100         Training Loss: 1.394304
Epoch: 69         Train Batch: 120         Training Loss: 1.385320
Epoch: 69         Train Batch: 140         Training Loss: 1.389069
Epoch: 69         Train Batch: 160         Training Loss: 1.401403
Epoch: 69         Train Batch: 180         Training Loss: 1.409580
Epoch: 69         Train Batch: 200         Training Loss: 1.415609
Epoch: 69         Eval Batch: 0         Validation Loss: 1.225433
Epoch: 69         Eval Batch: 20         Validation Loss: 1.225433
Epoch: 69         Training Loss: 1.415418         Validation Loss: 1.225433
Epoch: 70         Train Batch: 0         Training Loss: 1.056640
Epoch: 70         Train Batch: 20         Training Loss: 1.311414
Epoch: 70         Train Batch: 40         Training Loss: 1.331911
Epoch: 70         Train Batch: 60         Training Loss: 1.324620
Epoch: 70         Train Batch: 80         Training Loss: 1.323880
Epoch: 70         Train Batch: 100         Training Loss: 1.319668
Epoch: 70         Train Batch: 120         Training Loss: 1.333950
Epoch: 70         Train Batch: 140         Training Loss: 1.343278
Epoch: 70         Train Batch: 160         Training Loss: 1.355338
Epoch: 70         Train Batch: 180         Training Loss: 1.362658
Epoch: 70         Train Batch: 200         Training Loss: 1.368910
Epoch: 70         Eval Batch: 0         Validation Loss: 1.131142
Epoch: 70         Eval Batch: 20         Validation Loss: 1.131142
Epoch: 70         Training Loss: 1.367332         Validation Loss: 1.131142
Epoch: 71         Train Batch: 0         Training Loss: 1.206611
Epoch: 71         Train Batch: 20         Training Loss: 1.250677
```

```
Epoch: 71         Train Batch: 40         Training Loss: 1.275017
Epoch: 71         Train Batch: 60         Training Loss: 1.252952
Epoch: 71         Train Batch: 80         Training Loss: 1.279285
Epoch: 71         Train Batch: 100         Training Loss: 1.294976
Epoch: 71         Train Batch: 120         Training Loss: 1.297742
Epoch: 71         Train Batch: 140         Training Loss: 1.315779
Epoch: 71         Train Batch: 160         Training Loss: 1.325375
Epoch: 71         Train Batch: 180         Training Loss: 1.330886
Epoch: 71         Train Batch: 200         Training Loss: 1.336816
Epoch: 71         Eval Batch: 0         Validation Loss: 1.131425
Epoch: 71         Eval Batch: 20         Validation Loss: 1.131425
Epoch: 71         Training Loss: 1.336251         Validation Loss: 1.131425
Epoch: 72         Train Batch: 0         Training Loss: 1.759456
Epoch: 72         Train Batch: 20         Training Loss: 1.260164
Epoch: 72         Train Batch: 40         Training Loss: 1.274355
Epoch: 72         Train Batch: 60         Training Loss: 1.308672
Epoch: 72         Train Batch: 80         Training Loss: 1.310652
Epoch: 72         Train Batch: 100         Training Loss: 1.301522
Epoch: 72         Train Batch: 120         Training Loss: 1.297872
Epoch: 72         Train Batch: 140         Training Loss: 1.304563
Epoch: 72         Train Batch: 160         Training Loss: 1.307148
Epoch: 72         Train Batch: 180         Training Loss: 1.310980
Epoch: 72         Train Batch: 200         Training Loss: 1.324457
Epoch: 72         Eval Batch: 0         Validation Loss: 1.356273
Epoch: 72         Eval Batch: 20         Validation Loss: 1.356273
Epoch: 72         Training Loss: 1.326070         Validation Loss: 1.356273
Epoch: 73         Train Batch: 0         Training Loss: 1.216242
Epoch: 73         Train Batch: 20         Training Loss: 1.193380
Epoch: 73         Train Batch: 40         Training Loss: 1.234403
Epoch: 73         Train Batch: 60         Training Loss: 1.261285
Epoch: 73         Train Batch: 80         Training Loss: 1.269194
Epoch: 73         Train Batch: 100         Training Loss: 1.270634
Epoch: 73         Train Batch: 120         Training Loss: 1.269333
Epoch: 73         Train Batch: 140         Training Loss: 1.276978
Epoch: 73         Train Batch: 160         Training Loss: 1.293428
Epoch: 73         Train Batch: 180         Training Loss: 1.304060
Epoch: 73         Train Batch: 200         Training Loss: 1.313551
Epoch: 73         Eval Batch: 0         Validation Loss: 1.628425
Epoch: 73         Eval Batch: 20         Validation Loss: 1.628425
Epoch: 73         Training Loss: 1.315955         Validation Loss: 1.628425
Epoch: 74         Train Batch: 0         Training Loss: 1.277889
Epoch: 74         Train Batch: 20         Training Loss: 1.205430
Epoch: 74         Train Batch: 40         Training Loss: 1.199925
Epoch: 74         Train Batch: 60         Training Loss: 1.218580
Epoch: 74         Train Batch: 80         Training Loss: 1.234362
Epoch: 74         Train Batch: 100         Training Loss: 1.217037
Epoch: 74         Train Batch: 120         Training Loss: 1.236329
Epoch: 74         Train Batch: 140         Training Loss: 1.230967
```

```
Epoch: 74         Train Batch: 160         Training Loss: 1.256080
Epoch: 74         Train Batch: 180         Training Loss: 1.264711
Epoch: 74         Train Batch: 200         Training Loss: 1.265616
Epoch: 74         Eval Batch: 0         Validation Loss: 1.394924
Epoch: 74         Eval Batch: 20         Validation Loss: 1.394924
Epoch: 74         Training Loss: 1.266748         Validation Loss: 1.394924
Epoch: 75         Train Batch: 0         Training Loss: 1.099676
Epoch: 75         Train Batch: 20         Training Loss: 1.156557
Epoch: 75         Train Batch: 40         Training Loss: 1.201012
Epoch: 75         Train Batch: 60         Training Loss: 1.219203
Epoch: 75         Train Batch: 80         Training Loss: 1.205133
Epoch: 75         Train Batch: 100         Training Loss: 1.233987
Epoch: 75         Train Batch: 120         Training Loss: 1.238547
Epoch: 75         Train Batch: 140         Training Loss: 1.245044
Epoch: 75         Train Batch: 160         Training Loss: 1.253066
Epoch: 75         Train Batch: 180         Training Loss: 1.255844
Epoch: 75         Train Batch: 200         Training Loss: 1.259486
Epoch: 75         Eval Batch: 0         Validation Loss: 1.515553
Epoch: 75         Eval Batch: 20         Validation Loss: 1.515553
Epoch: 75         Training Loss: 1.256979         Validation Loss: 1.515553
Epoch: 76         Train Batch: 0         Training Loss: 1.107162
Epoch: 76         Train Batch: 20         Training Loss: 1.101576
Epoch: 76         Train Batch: 40         Training Loss: 1.114843
Epoch: 76         Train Batch: 60         Training Loss: 1.137220
Epoch: 76         Train Batch: 80         Training Loss: 1.175406
Epoch: 76         Train Batch: 100         Training Loss: 1.178771
Epoch: 76         Train Batch: 120         Training Loss: 1.192180
Epoch: 76         Train Batch: 140         Training Loss: 1.195661
Epoch: 76         Train Batch: 160         Training Loss: 1.207698
Epoch: 76         Train Batch: 180         Training Loss: 1.204672
Epoch: 76         Train Batch: 200         Training Loss: 1.211045
Epoch: 76         Eval Batch: 0         Validation Loss: 1.011214
Epoch: 76         Eval Batch: 20         Validation Loss: 1.011214
Epoch: 76         Training Loss: 1.215486         Validation Loss: 1.011214
Validation loss decreased (1.028653 --> 1.011214).  Saving model ...
Epoch: 77         Train Batch: 0         Training Loss: 1.051757
Epoch: 77         Train Batch: 20         Training Loss: 1.095429
Epoch: 77         Train Batch: 40         Training Loss: 1.095623
Epoch: 77         Train Batch: 60         Training Loss: 1.132398
Epoch: 77         Train Batch: 80         Training Loss: 1.157507
Epoch: 77         Train Batch: 100         Training Loss: 1.162941
Epoch: 77         Train Batch: 120         Training Loss: 1.162553
Epoch: 77         Train Batch: 140         Training Loss: 1.175487
Epoch: 77         Train Batch: 160         Training Loss: 1.199384
Epoch: 77         Train Batch: 180         Training Loss: 1.198073
Epoch: 77         Train Batch: 200         Training Loss: 1.196757
Epoch: 77         Eval Batch: 0         Validation Loss: 0.980879
Epoch: 77         Eval Batch: 20         Validation Loss: 0.980879
```

```
Epoch: 77         Training Loss: 1.195809         Validation Loss: 0.980879
Validation loss decreased (1.011214 --> 0.980879).  Saving model ...
Epoch: 78         Train Batch: 0         Training Loss: 1.093405
Epoch: 78         Train Batch: 20         Training Loss: 1.160264
Epoch: 78         Train Batch: 40         Training Loss: 1.170931
Epoch: 78         Train Batch: 60         Training Loss: 1.186835
Epoch: 78         Train Batch: 80         Training Loss: 1.189120
Epoch: 78         Train Batch: 100         Training Loss: 1.179509
Epoch: 78         Train Batch: 120         Training Loss: 1.191944
Epoch: 78         Train Batch: 140         Training Loss: 1.203300
Epoch: 78         Train Batch: 160         Training Loss: 1.200173
Epoch: 78         Train Batch: 180         Training Loss: 1.193698
Epoch: 78         Train Batch: 200         Training Loss: 1.197210
Epoch: 78         Eval Batch: 0         Validation Loss: 1.599877
Epoch: 78         Eval Batch: 20         Validation Loss: 1.599877
Epoch: 78         Training Loss: 1.203852         Validation Loss: 1.599877
Epoch: 79         Train Batch: 0         Training Loss: 0.768714
Epoch: 79         Train Batch: 20         Training Loss: 1.129059
Epoch: 79         Train Batch: 40         Training Loss: 1.086902
Epoch: 79         Train Batch: 60         Training Loss: 1.082366
Epoch: 79         Train Batch: 80         Training Loss: 1.103478
Epoch: 79         Train Batch: 100         Training Loss: 1.117104
Epoch: 79         Train Batch: 120         Training Loss: 1.122432
Epoch: 79         Train Batch: 140         Training Loss: 1.116528
Epoch: 79         Train Batch: 160         Training Loss: 1.123864
Epoch: 79         Train Batch: 180         Training Loss: 1.138665
Epoch: 79         Train Batch: 200         Training Loss: 1.152986
Epoch: 79         Eval Batch: 0         Validation Loss: 1.055511
Epoch: 79         Eval Batch: 20         Validation Loss: 1.055511
Epoch: 79         Training Loss: 1.158363         Validation Loss: 1.055511
Epoch: 80         Train Batch: 0         Training Loss: 1.297437
Epoch: 80         Train Batch: 20         Training Loss: 1.118037
Epoch: 80         Train Batch: 40         Training Loss: 1.086487
Epoch: 80         Train Batch: 60         Training Loss: 1.108786
Epoch: 80         Train Batch: 80         Training Loss: 1.123754
Epoch: 80         Train Batch: 100         Training Loss: 1.128540
Epoch: 80         Train Batch: 120         Training Loss: 1.119316
Epoch: 80         Train Batch: 140         Training Loss: 1.118811
Epoch: 80         Train Batch: 160         Training Loss: 1.131380
Epoch: 80         Train Batch: 180         Training Loss: 1.128004
Epoch: 80         Train Batch: 200         Training Loss: 1.131861
Epoch: 80         Eval Batch: 0         Validation Loss: 1.683973
Epoch: 80         Eval Batch: 20         Validation Loss: 1.683973
Epoch: 80         Training Loss: 1.137733         Validation Loss: 1.683973
Epoch: 81         Train Batch: 0         Training Loss: 0.936965
Epoch: 81         Train Batch: 20         Training Loss: 0.995815
Epoch: 81         Train Batch: 40         Training Loss: 1.045850
Epoch: 81         Train Batch: 60         Training Loss: 1.057127
```

```
Epoch: 81         Train Batch: 80         Training Loss: 1.087627
Epoch: 81         Train Batch: 100         Training Loss: 1.116706
Epoch: 81         Train Batch: 120         Training Loss: 1.128176
Epoch: 81         Train Batch: 140         Training Loss: 1.133934
Epoch: 81         Train Batch: 160         Training Loss: 1.140269
Epoch: 81         Train Batch: 180         Training Loss: 1.134969
Epoch: 81         Train Batch: 200         Training Loss: 1.143960
Epoch: 81         Eval Batch: 0         Validation Loss: 0.757605
Epoch: 81         Eval Batch: 20         Validation Loss: 0.757605
Epoch: 81         Training Loss: 1.146610         Validation Loss: 0.757605
Validation loss decreased (0.980879 --> 0.757605).  Saving model ...
Epoch: 82         Train Batch: 0         Training Loss: 0.670183
Epoch: 82         Train Batch: 20         Training Loss: 0.996837
Epoch: 82         Train Batch: 40         Training Loss: 1.020763
Epoch: 82         Train Batch: 60         Training Loss: 1.005750
Epoch: 82         Train Batch: 80         Training Loss: 1.031415
Epoch: 82         Train Batch: 100         Training Loss: 1.028411
Epoch: 82         Train Batch: 120         Training Loss: 1.034459
Epoch: 82         Train Batch: 140         Training Loss: 1.048766
Epoch: 82         Train Batch: 160         Training Loss: 1.053163
Epoch: 82         Train Batch: 180         Training Loss: 1.066956
Epoch: 82         Train Batch: 200         Training Loss: 1.069821
Epoch: 82         Eval Batch: 0         Validation Loss: 1.257121
Epoch: 82         Eval Batch: 20         Validation Loss: 1.257121
Epoch: 82         Training Loss: 1.076634         Validation Loss: 1.257121
Epoch: 83         Train Batch: 0         Training Loss: 0.804059
Epoch: 83         Train Batch: 20         Training Loss: 0.992669
Epoch: 83         Train Batch: 40         Training Loss: 1.050436
Epoch: 83         Train Batch: 60         Training Loss: 1.076040
Epoch: 83         Train Batch: 80         Training Loss: 1.073143
Epoch: 83         Train Batch: 100         Training Loss: 1.063781
Epoch: 83         Train Batch: 120         Training Loss: 1.051602
Epoch: 83         Train Batch: 140         Training Loss: 1.058374
Epoch: 83         Train Batch: 160         Training Loss: 1.066656
Epoch: 83         Train Batch: 180         Training Loss: 1.063236
Epoch: 83         Train Batch: 200         Training Loss: 1.068017
Epoch: 83         Eval Batch: 0         Validation Loss: 1.118510
Epoch: 83         Eval Batch: 20         Validation Loss: 1.118510
Epoch: 83         Training Loss: 1.068310         Validation Loss: 1.118510
Epoch: 84         Train Batch: 0         Training Loss: 1.715784
Epoch: 84         Train Batch: 20         Training Loss: 0.984171
Epoch: 84         Train Batch: 40         Training Loss: 1.020424
Epoch: 84         Train Batch: 60         Training Loss: 1.021457
Epoch: 84         Train Batch: 80         Training Loss: 1.024219
Epoch: 84         Train Batch: 100         Training Loss: 1.003821
Epoch: 84         Train Batch: 120         Training Loss: 1.007633
Epoch: 84         Train Batch: 140         Training Loss: 1.009837
Epoch: 84         Train Batch: 160         Training Loss: 1.009243
```

```
Epoch: 84         Train Batch: 180         Training Loss: 1.012564
Epoch: 84         Train Batch: 200         Training Loss: 1.021967
Epoch: 84         Eval Batch: 0         Validation Loss: 1.301311
Epoch: 84         Eval Batch: 20         Validation Loss: 1.301311
Epoch: 84         Training Loss: 1.033760         Validation Loss: 1.301311
Epoch: 85         Train Batch: 0         Training Loss: 0.852409
Epoch: 85         Train Batch: 20         Training Loss: 1.040360
Epoch: 85         Train Batch: 40         Training Loss: 1.028705
Epoch: 85         Train Batch: 60         Training Loss: 1.018592
Epoch: 85         Train Batch: 80         Training Loss: 1.011538
Epoch: 85         Train Batch: 100         Training Loss: 1.007894
Epoch: 85         Train Batch: 120         Training Loss: 1.010710
Epoch: 85         Train Batch: 140         Training Loss: 1.028283
Epoch: 85         Train Batch: 160         Training Loss: 1.026313
Epoch: 85         Train Batch: 180         Training Loss: 1.025632
Epoch: 85         Train Batch: 200         Training Loss: 1.027771
Epoch: 85         Eval Batch: 0         Validation Loss: 1.242113
Epoch: 85         Eval Batch: 20         Validation Loss: 1.242113
Epoch: 85         Training Loss: 1.028375         Validation Loss: 1.242113
Epoch: 86         Train Batch: 0         Training Loss: 0.777115
Epoch: 86         Train Batch: 20         Training Loss: 0.949890
Epoch: 86         Train Batch: 40         Training Loss: 1.003087
Epoch: 86         Train Batch: 60         Training Loss: 0.983297
Epoch: 86         Train Batch: 80         Training Loss: 0.970536
Epoch: 86         Train Batch: 100         Training Loss: 0.979799
Epoch: 86         Train Batch: 120         Training Loss: 0.981192
Epoch: 86         Train Batch: 140         Training Loss: 0.984060
Epoch: 86         Train Batch: 160         Training Loss: 1.000595
Epoch: 86         Train Batch: 180         Training Loss: 1.003079
Epoch: 86         Train Batch: 200         Training Loss: 1.013184
Epoch: 86         Eval Batch: 0         Validation Loss: 1.089055
Epoch: 86         Eval Batch: 20         Validation Loss: 1.089055
Epoch: 86         Training Loss: 1.015242         Validation Loss: 1.089055
Epoch: 87         Train Batch: 0         Training Loss: 1.003006
Epoch: 87         Train Batch: 20         Training Loss: 0.974303
Epoch: 87         Train Batch: 40         Training Loss: 0.967779
Epoch: 87         Train Batch: 60         Training Loss: 0.973036
Epoch: 87         Train Batch: 80         Training Loss: 0.961878
Epoch: 87         Train Batch: 100         Training Loss: 0.952836
Epoch: 87         Train Batch: 120         Training Loss: 0.960975
Epoch: 87         Train Batch: 140         Training Loss: 0.964565
Epoch: 87         Train Batch: 160         Training Loss: 0.972583
Epoch: 87         Train Batch: 180         Training Loss: 0.972668
Epoch: 87         Train Batch: 200         Training Loss: 0.975317
Epoch: 87         Eval Batch: 0         Validation Loss: 1.350147
Epoch: 87         Eval Batch: 20         Validation Loss: 1.350147
Epoch: 87         Training Loss: 0.983072         Validation Loss: 1.350147
Epoch: 88         Train Batch: 0         Training Loss: 0.895982
```

```
Epoch: 88         Train Batch: 20         Training Loss: 0.928121
Epoch: 88         Train Batch: 40         Training Loss: 0.894193
Epoch: 88         Train Batch: 60         Training Loss: 0.898905
Epoch: 88         Train Batch: 80         Training Loss: 0.904537
Epoch: 88         Train Batch: 100         Training Loss: 0.911803
Epoch: 88         Train Batch: 120         Training Loss: 0.928352
Epoch: 88         Train Batch: 140         Training Loss: 0.931764
Epoch: 88         Train Batch: 160         Training Loss: 0.949883
Epoch: 88         Train Batch: 180         Training Loss: 0.950261
Epoch: 88         Train Batch: 200         Training Loss: 0.951560
Epoch: 88         Eval Batch: 0         Validation Loss: 0.827385
Epoch: 88         Eval Batch: 20         Validation Loss: 0.827385
Epoch: 88         Training Loss: 0.948063         Validation Loss: 0.827385
Epoch: 89         Train Batch: 0         Training Loss: 0.982243
Epoch: 89         Train Batch: 20         Training Loss: 0.867232
Epoch: 89         Train Batch: 40         Training Loss: 0.915172
Epoch: 89         Train Batch: 60         Training Loss: 0.923735
Epoch: 89         Train Batch: 80         Training Loss: 0.934523
Epoch: 89         Train Batch: 100         Training Loss: 0.948164
Epoch: 89         Train Batch: 120         Training Loss: 0.936857
Epoch: 89         Train Batch: 140         Training Loss: 0.948635
Epoch: 89         Train Batch: 160         Training Loss: 0.954074
Epoch: 89         Train Batch: 180         Training Loss: 0.954153
Epoch: 89         Train Batch: 200         Training Loss: 0.951829
Epoch: 89         Eval Batch: 0         Validation Loss: 1.199374
Epoch: 89         Eval Batch: 20         Validation Loss: 1.199374
Epoch: 89         Training Loss: 0.955899         Validation Loss: 1.199374
Epoch: 90         Train Batch: 0         Training Loss: 0.631138
Epoch: 90         Train Batch: 20         Training Loss: 0.865147
Epoch: 90         Train Batch: 40         Training Loss: 0.903019
Epoch: 90         Train Batch: 60         Training Loss: 0.896758
Epoch: 90         Train Batch: 80         Training Loss: 0.890664
Epoch: 90         Train Batch: 100         Training Loss: 0.895632
Epoch: 90         Train Batch: 120         Training Loss: 0.902800
Epoch: 90         Train Batch: 140         Training Loss: 0.919784
Epoch: 90         Train Batch: 160         Training Loss: 0.934967
Epoch: 90         Train Batch: 180         Training Loss: 0.927329
Epoch: 90         Train Batch: 200         Training Loss: 0.926518
Epoch: 90         Eval Batch: 0         Validation Loss: 1.477640
Epoch: 90         Eval Batch: 20         Validation Loss: 1.477640
Epoch: 90         Training Loss: 0.930847         Validation Loss: 1.477640
Epoch: 91         Train Batch: 0         Training Loss: 0.991780
Epoch: 91         Train Batch: 20         Training Loss: 0.831621
Epoch: 91         Train Batch: 40         Training Loss: 0.840773
Epoch: 91         Train Batch: 60         Training Loss: 0.841724
Epoch: 91         Train Batch: 80         Training Loss: 0.858141
Epoch: 91         Train Batch: 100         Training Loss: 0.858536
Epoch: 91         Train Batch: 120         Training Loss: 0.862298
```

```
Epoch: 91         Train Batch: 140         Training Loss: 0.873955
Epoch: 91         Train Batch: 160         Training Loss: 0.877132
Epoch: 91         Train Batch: 180         Training Loss: 0.885222
Epoch: 91         Train Batch: 200         Training Loss: 0.887218
Epoch: 91         Eval Batch: 0         Validation Loss: 1.142412
Epoch: 91         Eval Batch: 20         Validation Loss: 1.142412
Epoch: 91         Training Loss: 0.894531         Validation Loss: 1.142412
Epoch: 92         Train Batch: 0         Training Loss: 0.764341
Epoch: 92         Train Batch: 20         Training Loss: 0.767240
Epoch: 92         Train Batch: 40         Training Loss: 0.817101
Epoch: 92         Train Batch: 60         Training Loss: 0.819677
Epoch: 92         Train Batch: 80         Training Loss: 0.854102
Epoch: 92         Train Batch: 100         Training Loss: 0.863790
Epoch: 92         Train Batch: 120         Training Loss: 0.855632
Epoch: 92         Train Batch: 140         Training Loss: 0.879641
Epoch: 92         Train Batch: 160         Training Loss: 0.876841
Epoch: 92         Train Batch: 180         Training Loss: 0.879796
Epoch: 92         Train Batch: 200         Training Loss: 0.880773
Epoch: 92         Eval Batch: 0         Validation Loss: 0.697310
Epoch: 92         Eval Batch: 20         Validation Loss: 0.697310
Epoch: 92         Training Loss: 0.880594         Validation Loss: 0.697310
Validation loss decreased (0.757605 --> 0.697310).  Saving model ...
Epoch: 93         Train Batch: 0         Training Loss: 0.851947
Epoch: 93         Train Batch: 20         Training Loss: 0.782959
Epoch: 93         Train Batch: 40         Training Loss: 0.776294
Epoch: 93         Train Batch: 60         Training Loss: 0.789810
Epoch: 93         Train Batch: 80         Training Loss: 0.799228
Epoch: 93         Train Batch: 100         Training Loss: 0.822593
Epoch: 93         Train Batch: 120         Training Loss: 0.837237
Epoch: 93         Train Batch: 140         Training Loss: 0.854809
Epoch: 93         Train Batch: 160         Training Loss: 0.859583
Epoch: 93         Train Batch: 180         Training Loss: 0.864745
Epoch: 93         Train Batch: 200         Training Loss: 0.867966
Epoch: 93         Eval Batch: 0         Validation Loss: 0.791782
Epoch: 93         Eval Batch: 20         Validation Loss: 0.791782
Epoch: 93         Training Loss: 0.870855         Validation Loss: 0.791782
Epoch: 94         Train Batch: 0         Training Loss: 0.540076
Epoch: 94         Train Batch: 20         Training Loss: 0.768084
Epoch: 94         Train Batch: 40         Training Loss: 0.793780
Epoch: 94         Train Batch: 60         Training Loss: 0.778986
Epoch: 94         Train Batch: 80         Training Loss: 0.785987
Epoch: 94         Train Batch: 100         Training Loss: 0.812552
Epoch: 94         Train Batch: 120         Training Loss: 0.823088
Epoch: 94         Train Batch: 140         Training Loss: 0.827455
Epoch: 94         Train Batch: 160         Training Loss: 0.831956
Epoch: 94         Train Batch: 180         Training Loss: 0.837838
Epoch: 94         Train Batch: 200         Training Loss: 0.844767
Epoch: 94         Eval Batch: 0         Validation Loss: 0.709061
```

```
Epoch: 94         Eval Batch: 20         Validation Loss: 0.709061
Epoch: 94         Training Loss: 0.844603         Validation Loss: 0.709061
Epoch: 95         Train Batch: 0         Training Loss: 0.769631
Epoch: 95         Train Batch: 20         Training Loss: 0.728988
Epoch: 95         Train Batch: 40         Training Loss: 0.746061
Epoch: 95         Train Batch: 60         Training Loss: 0.745218
Epoch: 95         Train Batch: 80         Training Loss: 0.772559
Epoch: 95         Train Batch: 100         Training Loss: 0.767496
Epoch: 95         Train Batch: 120         Training Loss: 0.782785
Epoch: 95         Train Batch: 140         Training Loss: 0.787938
Epoch: 95         Train Batch: 160         Training Loss: 0.794905
Epoch: 95         Train Batch: 180         Training Loss: 0.804509
Epoch: 95         Train Batch: 200         Training Loss: 0.809797
Epoch: 95         Eval Batch: 0         Validation Loss: 1.408771
Epoch: 95         Eval Batch: 20         Validation Loss: 1.408771
Epoch: 95         Training Loss: 0.812156         Validation Loss: 1.408771
Epoch: 96         Train Batch: 0         Training Loss: 0.558774
Epoch: 96         Train Batch: 20         Training Loss: 0.687874
Epoch: 96         Train Batch: 40         Training Loss: 0.754654
Epoch: 96         Train Batch: 60         Training Loss: 0.736778
Epoch: 96         Train Batch: 80         Training Loss: 0.746020
Epoch: 96         Train Batch: 100         Training Loss: 0.750285
Epoch: 96         Train Batch: 120         Training Loss: 0.759684
Epoch: 96         Train Batch: 140         Training Loss: 0.772804
Epoch: 96         Train Batch: 160         Training Loss: 0.777356
Epoch: 96         Train Batch: 180         Training Loss: 0.787330
Epoch: 96         Train Batch: 200         Training Loss: 0.800590
Epoch: 96         Eval Batch: 0         Validation Loss: 1.118452
Epoch: 96         Eval Batch: 20         Validation Loss: 1.118452
Epoch: 96         Training Loss: 0.803565         Validation Loss: 1.118452
Epoch: 97         Train Batch: 0         Training Loss: 0.407461
Epoch: 97         Train Batch: 20         Training Loss: 0.666904
Epoch: 97         Train Batch: 40         Training Loss: 0.721530
Epoch: 97         Train Batch: 60         Training Loss: 0.730457
Epoch: 97         Train Batch: 80         Training Loss: 0.739744
Epoch: 97         Train Batch: 100         Training Loss: 0.736238
Epoch: 97         Train Batch: 120         Training Loss: 0.752126
Epoch: 97         Train Batch: 140         Training Loss: 0.765584
Epoch: 97         Train Batch: 160         Training Loss: 0.766938
Epoch: 97         Train Batch: 180         Training Loss: 0.775157
Epoch: 97         Train Batch: 200         Training Loss: 0.775844
Epoch: 97         Eval Batch: 0         Validation Loss: 1.144173
Epoch: 97         Eval Batch: 20         Validation Loss: 1.144173
Epoch: 97         Training Loss: 0.776574         Validation Loss: 1.144173
Epoch: 98         Train Batch: 0         Training Loss: 0.730975
Epoch: 98         Train Batch: 20         Training Loss: 0.764455
Epoch: 98         Train Batch: 40         Training Loss: 0.804346
Epoch: 98         Train Batch: 60         Training Loss: 0.790031
```

```
Epoch: 98         Train Batch: 80         Training Loss: 0.761903
Epoch: 98         Train Batch: 100         Training Loss: 0.760365
Epoch: 98         Train Batch: 120         Training Loss: 0.769929
Epoch: 98         Train Batch: 140         Training Loss: 0.774873
Epoch: 98         Train Batch: 160         Training Loss: 0.781982
Epoch: 98         Train Batch: 180         Training Loss: 0.784850
Epoch: 98         Train Batch: 200         Training Loss: 0.794210
Epoch: 98         Eval Batch: 0         Validation Loss: 1.132937
Epoch: 98         Eval Batch: 20         Validation Loss: 1.132937
Epoch: 98         Training Loss: 0.795859         Validation Loss: 1.132937
Epoch: 99         Train Batch: 0         Training Loss: 0.870264
Epoch: 99         Train Batch: 20         Training Loss: 0.678295
Epoch: 99         Train Batch: 40         Training Loss: 0.731584
Epoch: 99         Train Batch: 60         Training Loss: 0.757201
Epoch: 99         Train Batch: 80         Training Loss: 0.770930
Epoch: 99         Train Batch: 100         Training Loss: 0.781055
Epoch: 99         Train Batch: 120         Training Loss: 0.785802
Epoch: 99         Train Batch: 140         Training Loss: 0.775257
Epoch: 99         Train Batch: 160         Training Loss: 0.785038
Epoch: 99         Train Batch: 180         Training Loss: 0.789020
Epoch: 99         Train Batch: 200         Training Loss: 0.782628
Epoch: 99         Eval Batch: 0         Validation Loss: 1.380784
Epoch: 99         Eval Batch: 20         Validation Loss: 1.380784
Epoch: 99         Training Loss: 0.785690         Validation Loss: 1.380784
Epoch: 100         Train Batch: 0         Training Loss: 0.941406
Epoch: 100         Train Batch: 20         Training Loss: 0.613100
Epoch: 100         Train Batch: 40         Training Loss: 0.730489
Epoch: 100         Train Batch: 60         Training Loss: 0.740240
Epoch: 100         Train Batch: 80         Training Loss: 0.749098
Epoch: 100         Train Batch: 100         Training Loss: 0.721532
Epoch: 100         Train Batch: 120         Training Loss: 0.716504
Epoch: 100         Train Batch: 140         Training Loss: 0.716287
Epoch: 100         Train Batch: 160         Training Loss: 0.724266
Epoch: 100         Train Batch: 180         Training Loss: 0.724863
Epoch: 100         Train Batch: 200         Training Loss: 0.730666
Epoch: 100         Eval Batch: 0         Validation Loss: 0.332191
Epoch: 100         Eval Batch: 20         Validation Loss: 0.332191
Epoch: 100         Training Loss: 0.731079         Validation Loss: 0.332191
Validation loss decreased (0.697310 --> 0.332191).  Saving model ...
```

Model learning until the last 100th epoch * Test Loss: 2.788943 * Test Accuracy: 41% (343/836)

1.1.12 (IMPLEMENTATION) Test the Model

Try out your model on the test dataset of dog images. Use the code cell below to calculate and print the test loss and accuracy. Ensure that your test accuracy is greater than 10%.

```
In [85]: import numpy as np

         def test(loaders, model, criterion, use_cuda):

             # monitor test loss and accuracy
             test_loss = 0.
             correct = 0.
             total = 0.

             model.eval()
             for batch_idx, (data, target) in enumerate(loaders['test']):
                 # move to GPU
                 if use_cuda:
                     data, target = data.cuda(), target.cuda()
                 # forward pass: compute predicted outputs by passing inputs to the model
                 output = model(data)
                 # calculate the loss
                 loss = criterion(output, target)
                 # update average test loss
                 test_loss = test_loss + ((1 / (batch_idx + 1)) * (loss.data - test_loss))
                 # convert output probabilities to predicted class
                 pred = output.data.max(1, keepdim=True)[1]
                 # compare predictions to true label
                 correct += np.sum(np.squeeze(pred.eq(target.data.view_as(pred))).cpu().numpy())
                 total += data.size(0)

             print('Test Loss: {:.6f}\n'.format(test_loss))

             print('\nTest Accuracy: %2d%% (%2d/%2d)' % (
                 100. * correct / total, correct, total))

In [9]: # call test function
        test(loaders_scratch, model_scratch, criterion_scratch, use_cuda)

Test Loss: 2.788943


Test Accuracy: 41% (343/836)
```

Step 4: Create a CNN to Classify Dog Breeds (using Transfer Learning)

You will now use transfer learning to create a CNN that can identify dog breed from images. Your CNN must attain at least 60% accuracy on the test set.

1.1.13 (IMPLEMENTATION) Specify Data Loaders for the Dog Dataset

Use the code cell below to write three separate data loaders for the training, validation, and test datasets of dog images (located at `dogImages/train`, `dogImages/valid`, and `dogImages/test`, re-

spectively).

If you like, **you are welcome to use the same data loaders from the previous step**, when you created a CNN from scratch.

```
In [4]: import os
        from torchvision import datasets
        from PIL import Image
        import torchvision.transforms as transforms
        import torch.nn.functional as F
        import torch

        from PIL import ImageFile
        ImageFile.LOAD_TRUNCATED_IMAGES = True

        # check if CUDA is available
        use_cuda = torch.cuda.is_available()

        device = torch.device("cuda:0" if use_cuda else "cpu")
        print("Device:", device)

        data_transfer = {}

        ## batch_size is the same as that used in the orignal VGG16 paper.
        def create_data_loader(dir_path, is_train=False, batch_size=256):
            if is_train:
                in_transform = transforms.Compose([
                                    transforms.RandomRotation(degrees=90),
                                    transforms.Resize((224, 224)),
                                    transforms.RandomHorizontalFlip(p=0.5),
                                    transforms.ToTensor(),
                                    transforms.Normalize((0.485, 0.456, 0.406),
                                                         (0.229, 0.224, 0.225))])
            else:
                in_transform = transforms.Compose([
                                    transforms.Resize((224, 224)),
                                    transforms.ToTensor(),
                                    transforms.Normalize((0.485, 0.456, 0.406),
                                                         (0.229, 0.224, 0.225))])
            data = datasets.ImageFolder(dir_path, transform=in_transform, target_transform=None)
            if is_train:
                data_transfer['train'] = data
            image_loader = torch.utils.data.DataLoader(data, batch_size=batch_size, shuffle=is_t
            return image_loader


        def create_loaders(img_path, has_subfolders=True, batch_size=256):
            loaders_scratch = {}
            if has_subfolders:
```

```
            train_path = img_path + '/train'
            valid_path = img_path + '/valid'
            test_path = img_path + '/test'
        else:
            train_path = img_path
            valid_path = img_path
            test_path = img_path
        loaders_scratch['train'] = create_data_loader(train_path, True, batch_size=batch_siz
        loaders_scratch['valid'] = create_data_loader(valid_path, False, batch_size=batch_si
        loaders_scratch['test'] = create_data_loader(test_path, False, batch_size=batch_size
        return loaders_scratch
loaders_transfer = create_loaders('dogImages', True, batch_size=32)
```

```
Device: cuda:0
```

1.1.14 (IMPLEMENTATION) Model Architecture

Use transfer learning to create a CNN to classify dog breed. Use the code cell below, and save your initialized model as the variable `model_transfer`.

```
In [10]: import torchvision.models as models
         import torch.nn as nn

         model_transfer = models.resnet18(pretrained=True)

         # Freeze the early feature layers
         for param in model_transfer.parameters():
             param.requires_grad = False

         # Change the output class count
         model_transfer.fc = nn.Linear(in_features=512, out_features=133, bias=True)

         if use_cuda:
             model_transfer = model_transfer.cuda()
```

Question 5: Outline the steps you took to get to your final CNN architecture and your reasoning at each step. Describe why you think the architecture is suitable for the current problem.

Answer:

I tried a few different experiments with two different architectures: AlexNet and ResNet18. I tried AlexNet to see the pure effect of pre-training a model of the same architecture to the model_scratch. Adopting what we learned from Style Transfer Learning lesson, I wanted to see the effect of freezing only a few selected layers. And then, I wanted to see the effect of a model deeper. Hence, I finally try ResNet16 model and compare it against AlexNet. (1) Pretrained AlexNet with all layers fine-tuned ended in 55% test accuracy. (2) Pretrained AlexNet with the first three convolutional layers frozen (style-reltaed and probably not content related) ended in 59% test accuracy. (3) Pretrained AlexNet with all layers frozen except the last layer ended in 65% test accuracy. (4) Pretrained ResNet16 with all layers frozen except the last layer ended in 81% test accuracy. Hence, ResNet18 is the chosen winner pre-trained model.

1.1.15 (IMPLEMENTATION) Specify Loss Function and Optimizer

Use the next code cell to specify a loss function and optimizer. Save the chosen loss function as `criterion_transfer`, and the optimizer as `optimizer_transfer` below.

```
In [87]: import torch.optim as optim

         criterion_transfer = nn.CrossEntropyLoss()
         optimizer_transfer = optim.SGD(model_transfer.parameters(), lr=0.01, weight_decay=0.000

         if use_cuda:
             criterion_transfer = criterion_transfer.cuda()
```

1.1.16 (IMPLEMENTATION) Train and Validate the Model

Train and validate your model in the code cell below. Save the final model parameters at filepath `'model_transfer.pt'`.

```
In [8]: # train the model
        model_transfer = train3(100, loaders_transfer, model_transfer, optimizer_transfer, crite
                                use_cuda, 'model_transfer_resnet.pt', 1)

        # load the model that got the best validation accuracy (uncomment the line below)
        model_transfer.load_state_dict(torch.load('model_transfer_resnet.pt'))
```

```
Epoch: 1         Train Batch: 0         Training Loss: 5.023550
Epoch: 1         Train Batch: 20         Training Loss: 5.015308
Epoch: 1         Train Batch: 40         Training Loss: 4.970042
Epoch: 1         Train Batch: 60         Training Loss: 4.912065
Epoch: 1         Train Batch: 80         Training Loss: 4.860831
Epoch: 1         Train Batch: 100         Training Loss: 4.812547
Epoch: 1         Train Batch: 120         Training Loss: 4.769690
Epoch: 1         Train Batch: 140         Training Loss: 4.726911
Epoch: 1         Train Batch: 160         Training Loss: 4.685254
Epoch: 1         Train Batch: 180         Training Loss: 4.641387
Epoch: 1         Train Batch: 200         Training Loss: 4.600260
Epoch: 1         Eval Batch: 0         Validation Loss: 4.037672
Epoch: 1         Eval Batch: 20         Validation Loss: 4.037672
Epoch: 1         Training Loss: 4.583211         Validation Loss: 4.037672
Validation loss decreased (inf --> 4.037672).  Saving model ...
Epoch: 2         Train Batch: 0         Training Loss: 4.089297
Epoch: 2         Train Batch: 20         Training Loss: 4.082132
Epoch: 2         Train Batch: 40         Training Loss: 4.054924
Epoch: 2         Train Batch: 60         Training Loss: 4.028862
Epoch: 2         Train Batch: 80         Training Loss: 3.994467
Epoch: 2         Train Batch: 100         Training Loss: 3.961220
Epoch: 2         Train Batch: 120         Training Loss: 3.938986
Epoch: 2         Train Batch: 140         Training Loss: 3.909961
Epoch: 2         Train Batch: 160         Training Loss: 3.883653
```

```
Epoch: 2         Train Batch: 180         Training Loss: 3.850620
Epoch: 2         Train Batch: 200         Training Loss: 3.827182
Epoch: 2         Eval Batch: 0         Validation Loss: 3.332691
Epoch: 2         Eval Batch: 20         Validation Loss: 3.332691
Epoch: 2         Training Loss: 3.814938         Validation Loss: 3.332691
Validation loss decreased (4.037672 --> 3.332691).  Saving model ...
Epoch: 3         Train Batch: 0         Training Loss: 3.411844
Epoch: 3         Train Batch: 20         Training Loss: 3.469425
Epoch: 3         Train Batch: 40         Training Loss: 3.453404
Epoch: 3         Train Batch: 60         Training Loss: 3.425128
Epoch: 3         Train Batch: 80         Training Loss: 3.384558
Epoch: 3         Train Batch: 100         Training Loss: 3.365254
Epoch: 3         Train Batch: 120         Training Loss: 3.341051
Epoch: 3         Train Batch: 140         Training Loss: 3.317137
Epoch: 3         Train Batch: 160         Training Loss: 3.303152
Epoch: 3         Train Batch: 180         Training Loss: 3.280317
Epoch: 3         Train Batch: 200         Training Loss: 3.263848
Epoch: 3         Eval Batch: 0         Validation Loss: 2.825439
Epoch: 3         Eval Batch: 20         Validation Loss: 2.825439
Epoch: 3         Training Loss: 3.257968         Validation Loss: 2.825439
Validation loss decreased (3.332691 --> 2.825439).  Saving model ...
Epoch: 4         Train Batch: 0         Training Loss: 3.354102
Epoch: 4         Train Batch: 20         Training Loss: 3.031072
Epoch: 4         Train Batch: 40         Training Loss: 3.026166
Epoch: 4         Train Batch: 60         Training Loss: 2.997531
Epoch: 4         Train Batch: 80         Training Loss: 2.960027
Epoch: 4         Train Batch: 100         Training Loss: 2.945592
Epoch: 4         Train Batch: 120         Training Loss: 2.916553
Epoch: 4         Train Batch: 140         Training Loss: 2.901873
Epoch: 4         Train Batch: 160         Training Loss: 2.884095
Epoch: 4         Train Batch: 180         Training Loss: 2.866090
Epoch: 4         Train Batch: 200         Training Loss: 2.847637
Epoch: 4         Eval Batch: 0         Validation Loss: 2.925062
Epoch: 4         Eval Batch: 20         Validation Loss: 2.925062
Epoch: 4         Training Loss: 2.838642         Validation Loss: 2.925062
Epoch: 5         Train Batch: 0         Training Loss: 2.529508
Epoch: 5         Train Batch: 20         Training Loss: 2.666370
Epoch: 5         Train Batch: 40         Training Loss: 2.643345
Epoch: 5         Train Batch: 60         Training Loss: 2.633788
Epoch: 5         Train Batch: 80         Training Loss: 2.610856
Epoch: 5         Train Batch: 100         Training Loss: 2.616416
Epoch: 5         Train Batch: 120         Training Loss: 2.593777
Epoch: 5         Train Batch: 140         Training Loss: 2.574734
Epoch: 5         Train Batch: 160         Training Loss: 2.564675
Epoch: 5         Train Batch: 180         Training Loss: 2.555185
Epoch: 5         Train Batch: 200         Training Loss: 2.552091
Epoch: 5         Eval Batch: 0         Validation Loss: 2.771036
Epoch: 5         Eval Batch: 20         Validation Loss: 2.771036
```

```
Epoch: 5         Training Loss: 2.550780         Validation Loss: 2.771036
Validation loss decreased (2.825439 --> 2.771036).  Saving model ...
Epoch: 6         Train Batch: 0         Training Loss: 2.733330
Epoch: 6         Train Batch: 20         Training Loss: 2.398710
Epoch: 6         Train Batch: 40         Training Loss: 2.367101
Epoch: 6         Train Batch: 60         Training Loss: 2.364183
Epoch: 6         Train Batch: 80         Training Loss: 2.363220
Epoch: 6         Train Batch: 100         Training Loss: 2.350600
Epoch: 6         Train Batch: 120         Training Loss: 2.341160
Epoch: 6         Train Batch: 140         Training Loss: 2.345537
Epoch: 6         Train Batch: 160         Training Loss: 2.344763
Epoch: 6         Train Batch: 180         Training Loss: 2.328730
Epoch: 6         Train Batch: 200         Training Loss: 2.320619
Epoch: 6         Eval Batch: 0         Validation Loss: 2.129217
Epoch: 6         Eval Batch: 20         Validation Loss: 2.129217
Epoch: 6         Training Loss: 2.314165         Validation Loss: 2.129217
Validation loss decreased (2.771036 --> 2.129217).  Saving model ...
Epoch: 7         Train Batch: 0         Training Loss: 1.899545
Epoch: 7         Train Batch: 20         Training Loss: 2.188148
Epoch: 7         Train Batch: 40         Training Loss: 2.171078
Epoch: 7         Train Batch: 60         Training Loss: 2.180557
Epoch: 7         Train Batch: 80         Training Loss: 2.148956
Epoch: 7         Train Batch: 100         Training Loss: 2.151783
Epoch: 7         Train Batch: 120         Training Loss: 2.137347
Epoch: 7         Train Batch: 140         Training Loss: 2.130648
Epoch: 7         Train Batch: 160         Training Loss: 2.129948
Epoch: 7         Train Batch: 180         Training Loss: 2.126424
Epoch: 7         Train Batch: 200         Training Loss: 2.120191
Epoch: 7         Eval Batch: 0         Validation Loss: 1.823386
Epoch: 7         Eval Batch: 20         Validation Loss: 1.823386
Epoch: 7         Training Loss: 2.123107         Validation Loss: 1.823386
Validation loss decreased (2.129217 --> 1.823386).  Saving model ...
Epoch: 8         Train Batch: 0         Training Loss: 2.010787
Epoch: 8         Train Batch: 20         Training Loss: 2.006282
Epoch: 8         Train Batch: 40         Training Loss: 2.023873
Epoch: 8         Train Batch: 60         Training Loss: 2.018048
Epoch: 8         Train Batch: 80         Training Loss: 2.006690
Epoch: 8         Train Batch: 100         Training Loss: 2.013226
Epoch: 8         Train Batch: 120         Training Loss: 2.003063
Epoch: 8         Train Batch: 140         Training Loss: 1.997092
Epoch: 8         Train Batch: 160         Training Loss: 1.981831
Epoch: 8         Train Batch: 180         Training Loss: 1.982394
Epoch: 8         Train Batch: 200         Training Loss: 1.974058
Epoch: 8         Eval Batch: 0         Validation Loss: 1.734937
Epoch: 8         Eval Batch: 20         Validation Loss: 1.734937
Epoch: 8         Training Loss: 1.977113         Validation Loss: 1.734937
Validation loss decreased (1.823386 --> 1.734937).  Saving model ...
Epoch: 9         Train Batch: 0         Training Loss: 1.637565
```

```
Epoch: 9         Train Batch: 20         Training Loss: 1.912665
Epoch: 9         Train Batch: 40         Training Loss: 1.941957
Epoch: 9         Train Batch: 60         Training Loss: 1.920938
Epoch: 9         Train Batch: 80         Training Loss: 1.913909
Epoch: 9         Train Batch: 100         Training Loss: 1.904074
Epoch: 9         Train Batch: 120         Training Loss: 1.895311
Epoch: 9         Train Batch: 140         Training Loss: 1.894387
Epoch: 9         Train Batch: 160         Training Loss: 1.892556
Epoch: 9         Train Batch: 180         Training Loss: 1.890489
Epoch: 9         Train Batch: 200         Training Loss: 1.884422
Epoch: 9         Eval Batch: 0         Validation Loss: 1.654945
Epoch: 9         Eval Batch: 20         Validation Loss: 1.654945
Epoch: 9         Training Loss: 1.883939         Validation Loss: 1.654945
Validation loss decreased (1.734937 --> 1.654945).  Saving model ...
Epoch: 10         Train Batch: 0         Training Loss: 1.833697
Epoch: 10         Train Batch: 20         Training Loss: 1.837613
Epoch: 10         Train Batch: 40         Training Loss: 1.826600
Epoch: 10         Train Batch: 60         Training Loss: 1.837455
Epoch: 10         Train Batch: 80         Training Loss: 1.832375
Epoch: 10         Train Batch: 100         Training Loss: 1.822374
Epoch: 10         Train Batch: 120         Training Loss: 1.826767
Epoch: 10         Train Batch: 140         Training Loss: 1.818525
Epoch: 10         Train Batch: 160         Training Loss: 1.814890
Epoch: 10         Train Batch: 180         Training Loss: 1.802471
Epoch: 10         Train Batch: 200         Training Loss: 1.795188
Epoch: 10         Eval Batch: 0         Validation Loss: 1.807544
Epoch: 10         Eval Batch: 20         Validation Loss: 1.807544
Epoch: 10         Training Loss: 1.797063         Validation Loss: 1.807544
Epoch: 11         Train Batch: 0         Training Loss: 1.729299
Epoch: 11         Train Batch: 20         Training Loss: 1.691916
Epoch: 11         Train Batch: 40         Training Loss: 1.682916
Epoch: 11         Train Batch: 60         Training Loss: 1.701109
Epoch: 11         Train Batch: 80         Training Loss: 1.703354
Epoch: 11         Train Batch: 100         Training Loss: 1.706717
Epoch: 11         Train Batch: 120         Training Loss: 1.707967
Epoch: 11         Train Batch: 140         Training Loss: 1.703079
Epoch: 11         Train Batch: 160         Training Loss: 1.709862
Epoch: 11         Train Batch: 180         Training Loss: 1.710043
Epoch: 11         Train Batch: 200         Training Loss: 1.705536
Epoch: 11         Eval Batch: 0         Validation Loss: 1.460766
Epoch: 11         Eval Batch: 20         Validation Loss: 1.460766
Epoch: 11         Training Loss: 1.700496         Validation Loss: 1.460766
Validation loss decreased (1.654945 --> 1.460766).  Saving model ...
Epoch: 12         Train Batch: 0         Training Loss: 1.832846
Epoch: 12         Train Batch: 20         Training Loss: 1.680337
Epoch: 12         Train Batch: 40         Training Loss: 1.620019
Epoch: 12         Train Batch: 60         Training Loss: 1.606516
Epoch: 12         Train Batch: 80         Training Loss: 1.631805
```

```
Epoch: 12         Train Batch: 100         Training Loss: 1.627364
Epoch: 12         Train Batch: 120         Training Loss: 1.621414
Epoch: 12         Train Batch: 140         Training Loss: 1.613979
Epoch: 12         Train Batch: 160         Training Loss: 1.623652
Epoch: 12         Train Batch: 180         Training Loss: 1.624955
Epoch: 12         Train Batch: 200         Training Loss: 1.629150
Epoch: 12         Eval Batch: 0         Validation Loss: 1.564606
Epoch: 12         Eval Batch: 20         Validation Loss: 1.564606
Epoch: 12         Training Loss: 1.629465         Validation Loss: 1.564606
Epoch: 13         Train Batch: 0         Training Loss: 1.584583
Epoch: 13         Train Batch: 20         Training Loss: 1.518474
Epoch: 13         Train Batch: 40         Training Loss: 1.565107
Epoch: 13         Train Batch: 60         Training Loss: 1.583714
Epoch: 13         Train Batch: 80         Training Loss: 1.586672
Epoch: 13         Train Batch: 100         Training Loss: 1.588184
Epoch: 13         Train Batch: 120         Training Loss: 1.591573
Epoch: 13         Train Batch: 140         Training Loss: 1.595076
Epoch: 13         Train Batch: 160         Training Loss: 1.593486
Epoch: 13         Train Batch: 180         Training Loss: 1.584196
Epoch: 13         Train Batch: 200         Training Loss: 1.581362
Epoch: 13         Eval Batch: 0         Validation Loss: 1.618110
Epoch: 13         Eval Batch: 20         Validation Loss: 1.618110
Epoch: 13         Training Loss: 1.583650         Validation Loss: 1.618110
Epoch: 14         Train Batch: 0         Training Loss: 1.579484
Epoch: 14         Train Batch: 20         Training Loss: 1.516309
Epoch: 14         Train Batch: 40         Training Loss: 1.517623
Epoch: 14         Train Batch: 60         Training Loss: 1.533000
Epoch: 14         Train Batch: 80         Training Loss: 1.533493
Epoch: 14         Train Batch: 100         Training Loss: 1.528628
Epoch: 14         Train Batch: 120         Training Loss: 1.526148
Epoch: 14         Train Batch: 140         Training Loss: 1.522951
Epoch: 14         Train Batch: 160         Training Loss: 1.531928
Epoch: 14         Train Batch: 180         Training Loss: 1.522353
Epoch: 14         Train Batch: 200         Training Loss: 1.519744
Epoch: 14         Eval Batch: 0         Validation Loss: 1.765058
Epoch: 14         Eval Batch: 20         Validation Loss: 1.765058
Epoch: 14         Training Loss: 1.519681         Validation Loss: 1.765058
Epoch: 15         Train Batch: 0         Training Loss: 1.565091
Epoch: 15         Train Batch: 20         Training Loss: 1.456922
Epoch: 15         Train Batch: 40         Training Loss: 1.471311
Epoch: 15         Train Batch: 60         Training Loss: 1.493993
Epoch: 15         Train Batch: 80         Training Loss: 1.473806
Epoch: 15         Train Batch: 100         Training Loss: 1.473596
Epoch: 15         Train Batch: 120         Training Loss: 1.468747
Epoch: 15         Train Batch: 140         Training Loss: 1.469110
Epoch: 15         Train Batch: 160         Training Loss: 1.464006
Epoch: 15         Train Batch: 180         Training Loss: 1.473654
Epoch: 15         Train Batch: 200         Training Loss: 1.476703
```

```
Epoch: 15         Eval Batch: 0         Validation Loss: 1.743337
Epoch: 15         Eval Batch: 20         Validation Loss: 1.743337
Epoch: 15         Training Loss: 1.477976         Validation Loss: 1.743337
Epoch: 16         Train Batch: 0         Training Loss: 1.661556
Epoch: 16         Train Batch: 20         Training Loss: 1.494669
Epoch: 16         Train Batch: 40         Training Loss: 1.477255
Epoch: 16         Train Batch: 60         Training Loss: 1.446481
Epoch: 16         Train Batch: 80         Training Loss: 1.425376
Epoch: 16         Train Batch: 100         Training Loss: 1.429397
Epoch: 16         Train Batch: 120         Training Loss: 1.422506
Epoch: 16         Train Batch: 140         Training Loss: 1.422263
Epoch: 16         Train Batch: 160         Training Loss: 1.420596
Epoch: 16         Train Batch: 180         Training Loss: 1.421646
Epoch: 16         Train Batch: 200         Training Loss: 1.425473
Epoch: 16         Eval Batch: 0         Validation Loss: 1.377062
Epoch: 16         Eval Batch: 20         Validation Loss: 1.377062
Epoch: 16         Training Loss: 1.426028         Validation Loss: 1.377062
Validation loss decreased (1.460766 --> 1.377062).  Saving model ...
Epoch: 17         Train Batch: 0         Training Loss: 1.535766
Epoch: 17         Train Batch: 20         Training Loss: 1.362773
Epoch: 17         Train Batch: 40         Training Loss: 1.377200
Epoch: 17         Train Batch: 60         Training Loss: 1.373913
Epoch: 17         Train Batch: 80         Training Loss: 1.387041
Epoch: 17         Train Batch: 100         Training Loss: 1.393799
Epoch: 17         Train Batch: 120         Training Loss: 1.407352
Epoch: 17         Train Batch: 140         Training Loss: 1.400266
Epoch: 17         Train Batch: 160         Training Loss: 1.401531
Epoch: 17         Train Batch: 180         Training Loss: 1.401196
Epoch: 17         Train Batch: 200         Training Loss: 1.394806
Epoch: 17         Eval Batch: 0         Validation Loss: 1.431812
Epoch: 17         Eval Batch: 20         Validation Loss: 1.431812
Epoch: 17         Training Loss: 1.396993         Validation Loss: 1.431812
Epoch: 18         Train Batch: 0         Training Loss: 1.358704
Epoch: 18         Train Batch: 20         Training Loss: 1.373543
Epoch: 18         Train Batch: 40         Training Loss: 1.366698
Epoch: 18         Train Batch: 60         Training Loss: 1.351061
Epoch: 18         Train Batch: 80         Training Loss: 1.341292
Epoch: 18         Train Batch: 100         Training Loss: 1.330634
Epoch: 18         Train Batch: 120         Training Loss: 1.322456
Epoch: 18         Train Batch: 140         Training Loss: 1.332275
Epoch: 18         Train Batch: 160         Training Loss: 1.337273
Epoch: 18         Train Batch: 180         Training Loss: 1.344500
Epoch: 18         Train Batch: 200         Training Loss: 1.345203
Epoch: 18         Eval Batch: 0         Validation Loss: 1.362558
Epoch: 18         Eval Batch: 20         Validation Loss: 1.362558
Epoch: 18         Training Loss: 1.344566         Validation Loss: 1.362558
Validation loss decreased (1.377062 --> 1.362558).  Saving model ...
Epoch: 19         Train Batch: 0         Training Loss: 1.242618
```

```
Epoch: 19         Train Batch: 20         Training Loss: 1.352548
Epoch: 19         Train Batch: 40         Training Loss: 1.341364
Epoch: 19         Train Batch: 60         Training Loss: 1.328025
Epoch: 19         Train Batch: 80         Training Loss: 1.322073
Epoch: 19         Train Batch: 100         Training Loss: 1.322011
Epoch: 19         Train Batch: 120         Training Loss: 1.311921
Epoch: 19         Train Batch: 140         Training Loss: 1.309446
Epoch: 19         Train Batch: 160         Training Loss: 1.315187
Epoch: 19         Train Batch: 180         Training Loss: 1.317870
Epoch: 19         Train Batch: 200         Training Loss: 1.315118
Epoch: 19         Eval Batch: 0         Validation Loss: 1.222812
Epoch: 19         Eval Batch: 20         Validation Loss: 1.222812
Epoch: 19         Training Loss: 1.314353         Validation Loss: 1.222812
Validation loss decreased (1.362558 --> 1.222812).  Saving model ...
Epoch: 20         Train Batch: 0         Training Loss: 1.170017
Epoch: 20         Train Batch: 20         Training Loss: 1.317476
Epoch: 20         Train Batch: 40         Training Loss: 1.288468
Epoch: 20         Train Batch: 60         Training Loss: 1.290575
Epoch: 20         Train Batch: 80         Training Loss: 1.293495
Epoch: 20         Train Batch: 100         Training Loss: 1.301791
Epoch: 20         Train Batch: 120         Training Loss: 1.308033
Epoch: 20         Train Batch: 140         Training Loss: 1.308211
Epoch: 20         Train Batch: 160         Training Loss: 1.303866
Epoch: 20         Train Batch: 180         Training Loss: 1.306056
Epoch: 20         Train Batch: 200         Training Loss: 1.300688
Epoch: 20         Eval Batch: 0         Validation Loss: 1.131234
Epoch: 20         Eval Batch: 20         Validation Loss: 1.131234
Epoch: 20         Training Loss: 1.300789         Validation Loss: 1.131234
Validation loss decreased (1.222812 --> 1.131234).  Saving model ...
Epoch: 21         Train Batch: 0         Training Loss: 1.073376
Epoch: 21         Train Batch: 20         Training Loss: 1.209723
Epoch: 21         Train Batch: 40         Training Loss: 1.238583
Epoch: 21         Train Batch: 60         Training Loss: 1.230311
Epoch: 21         Train Batch: 80         Training Loss: 1.245280
Epoch: 21         Train Batch: 100         Training Loss: 1.249125
Epoch: 21         Train Batch: 120         Training Loss: 1.254953
Epoch: 21         Train Batch: 140         Training Loss: 1.264491
Epoch: 21         Train Batch: 160         Training Loss: 1.274796
Epoch: 21         Train Batch: 180         Training Loss: 1.269000
Epoch: 21         Train Batch: 200         Training Loss: 1.262735
Epoch: 21         Eval Batch: 0         Validation Loss: 1.371810
Epoch: 21         Eval Batch: 20         Validation Loss: 1.371810
Epoch: 21         Training Loss: 1.263179         Validation Loss: 1.371810
Epoch: 22         Train Batch: 0         Training Loss: 1.498310
Epoch: 22         Train Batch: 20         Training Loss: 1.201248
Epoch: 22         Train Batch: 40         Training Loss: 1.213390
Epoch: 22         Train Batch: 60         Training Loss: 1.208780
Epoch: 22         Train Batch: 80         Training Loss: 1.226754
```

```
Epoch: 22         Train Batch: 100         Training Loss: 1.232288
Epoch: 22         Train Batch: 120         Training Loss: 1.237785
Epoch: 22         Train Batch: 140         Training Loss: 1.235911
Epoch: 22         Train Batch: 160         Training Loss: 1.238721
Epoch: 22         Train Batch: 180         Training Loss: 1.243751
Epoch: 22         Train Batch: 200         Training Loss: 1.251029
Epoch: 22         Eval Batch: 0         Validation Loss: 1.306962
Epoch: 22         Eval Batch: 20         Validation Loss: 1.306962
Epoch: 22         Training Loss: 1.253637         Validation Loss: 1.306962
Epoch: 23         Train Batch: 0         Training Loss: 1.758551
Epoch: 23         Train Batch: 20         Training Loss: 1.292083
Epoch: 23         Train Batch: 40         Training Loss: 1.259817
Epoch: 23         Train Batch: 60         Training Loss: 1.246451
Epoch: 23         Train Batch: 80         Training Loss: 1.262389
Epoch: 23         Train Batch: 100         Training Loss: 1.260586
Epoch: 23         Train Batch: 120         Training Loss: 1.252313
Epoch: 23         Train Batch: 140         Training Loss: 1.236065
Epoch: 23         Train Batch: 160         Training Loss: 1.229918
Epoch: 23         Train Batch: 180         Training Loss: 1.228448
Epoch: 23         Train Batch: 200         Training Loss: 1.225676
Epoch: 23         Eval Batch: 0         Validation Loss: 1.130667
Epoch: 23         Eval Batch: 20         Validation Loss: 1.130667
Epoch: 23         Training Loss: 1.223094         Validation Loss: 1.130667
Validation loss decreased (1.131234 --> 1.130667).  Saving model ...
Epoch: 24         Train Batch: 0         Training Loss: 1.208149
Epoch: 24         Train Batch: 20         Training Loss: 1.192076
Epoch: 24         Train Batch: 40         Training Loss: 1.219674
Epoch: 24         Train Batch: 60         Training Loss: 1.223128
Epoch: 24         Train Batch: 80         Training Loss: 1.216529
Epoch: 24         Train Batch: 100         Training Loss: 1.215506
Epoch: 24         Train Batch: 120         Training Loss: 1.217747
Epoch: 24         Train Batch: 140         Training Loss: 1.223658
Epoch: 24         Train Batch: 160         Training Loss: 1.224556
Epoch: 24         Train Batch: 180         Training Loss: 1.227722
Epoch: 24         Train Batch: 200         Training Loss: 1.224114
Epoch: 24         Eval Batch: 0         Validation Loss: 1.341985
Epoch: 24         Eval Batch: 20         Validation Loss: 1.341985
Epoch: 24         Training Loss: 1.218903         Validation Loss: 1.341985
Epoch: 25         Train Batch: 0         Training Loss: 1.521037
Epoch: 25         Train Batch: 20         Training Loss: 1.250937
Epoch: 25         Train Batch: 40         Training Loss: 1.196918
Epoch: 25         Train Batch: 60         Training Loss: 1.203829
Epoch: 25         Train Batch: 80         Training Loss: 1.210984
Epoch: 25         Train Batch: 100         Training Loss: 1.198617
Epoch: 25         Train Batch: 120         Training Loss: 1.207587
Epoch: 25         Train Batch: 140         Training Loss: 1.206746
Epoch: 25         Train Batch: 160         Training Loss: 1.200979
Epoch: 25         Train Batch: 180         Training Loss: 1.200389
```

```
Epoch: 25         Train Batch: 200         Training Loss: 1.194844
Epoch: 25         Eval Batch: 0         Validation Loss: 1.213590
Epoch: 25         Eval Batch: 20         Validation Loss: 1.213590
Epoch: 25         Training Loss: 1.191890         Validation Loss: 1.213590
Epoch: 26         Train Batch: 0         Training Loss: 1.117214
Epoch: 26         Train Batch: 20         Training Loss: 1.168474
Epoch: 26         Train Batch: 40         Training Loss: 1.191775
Epoch: 26         Train Batch: 60         Training Loss: 1.171265
Epoch: 26         Train Batch: 80         Training Loss: 1.178066
Epoch: 26         Train Batch: 100         Training Loss: 1.184622
Epoch: 26         Train Batch: 120         Training Loss: 1.192596
Epoch: 26         Train Batch: 140         Training Loss: 1.184074
Epoch: 26         Train Batch: 160         Training Loss: 1.182751
Epoch: 26         Train Batch: 180         Training Loss: 1.184935
Epoch: 26         Train Batch: 200         Training Loss: 1.188426
Epoch: 26         Eval Batch: 0         Validation Loss: 1.026947
Epoch: 26         Eval Batch: 20         Validation Loss: 1.026947
Epoch: 26         Training Loss: 1.189190         Validation Loss: 1.026947
Validation loss decreased (1.130667 --> 1.026947).  Saving model ...
Epoch: 27         Train Batch: 0         Training Loss: 1.103639
Epoch: 27         Train Batch: 20         Training Loss: 1.158776
Epoch: 27         Train Batch: 40         Training Loss: 1.180510
Epoch: 27         Train Batch: 60         Training Loss: 1.174130
Epoch: 27         Train Batch: 80         Training Loss: 1.179108
Epoch: 27         Train Batch: 100         Training Loss: 1.164981
Epoch: 27         Train Batch: 120         Training Loss: 1.162043
Epoch: 27         Train Batch: 140         Training Loss: 1.158477
Epoch: 27         Train Batch: 160         Training Loss: 1.146280
Epoch: 27         Train Batch: 180         Training Loss: 1.145134
Epoch: 27         Train Batch: 200         Training Loss: 1.147546
Epoch: 27         Eval Batch: 0         Validation Loss: 1.140981
Epoch: 27         Eval Batch: 20         Validation Loss: 1.140981
Epoch: 27         Training Loss: 1.146299         Validation Loss: 1.140981
Epoch: 28         Train Batch: 0         Training Loss: 1.158128
Epoch: 28         Train Batch: 20         Training Loss: 1.201417
Epoch: 28         Train Batch: 40         Training Loss: 1.212796
Epoch: 28         Train Batch: 60         Training Loss: 1.172081
Epoch: 28         Train Batch: 80         Training Loss: 1.176070
Epoch: 28         Train Batch: 100         Training Loss: 1.167610
Epoch: 28         Train Batch: 120         Training Loss: 1.162581
Epoch: 28         Train Batch: 140         Training Loss: 1.177699
Epoch: 28         Train Batch: 160         Training Loss: 1.170029
Epoch: 28         Train Batch: 180         Training Loss: 1.160386
Epoch: 28         Train Batch: 200         Training Loss: 1.159575
Epoch: 28         Eval Batch: 0         Validation Loss: 1.536517
Epoch: 28         Eval Batch: 20         Validation Loss: 1.536517
Epoch: 28         Training Loss: 1.160496         Validation Loss: 1.536517
Epoch: 29         Train Batch: 0         Training Loss: 1.075693
```

```
Epoch: 29         Train Batch: 20         Training Loss: 1.157548
Epoch: 29         Train Batch: 40         Training Loss: 1.150758
Epoch: 29         Train Batch: 60         Training Loss: 1.131286
Epoch: 29         Train Batch: 80         Training Loss: 1.139794
Epoch: 29         Train Batch: 100         Training Loss: 1.133484
Epoch: 29         Train Batch: 120         Training Loss: 1.130909
Epoch: 29         Train Batch: 140         Training Loss: 1.126619
Epoch: 29         Train Batch: 160         Training Loss: 1.128058
Epoch: 29         Train Batch: 180         Training Loss: 1.128520
Epoch: 29         Train Batch: 200         Training Loss: 1.124273
Epoch: 29         Eval Batch: 0         Validation Loss: 0.825764
Epoch: 29         Eval Batch: 20         Validation Loss: 0.825764
Epoch: 29         Training Loss: 1.120395         Validation Loss: 0.825764
Validation loss decreased (1.026947 --> 0.825764).  Saving model ...
Epoch: 30         Train Batch: 0         Training Loss: 1.146512
Epoch: 30         Train Batch: 20         Training Loss: 1.148855
Epoch: 30         Train Batch: 40         Training Loss: 1.135906
Epoch: 30         Train Batch: 60         Training Loss: 1.141088
Epoch: 30         Train Batch: 80         Training Loss: 1.125813
Epoch: 30         Train Batch: 100         Training Loss: 1.118172
Epoch: 30         Train Batch: 120         Training Loss: 1.123782
Epoch: 30         Train Batch: 140         Training Loss: 1.111999
Epoch: 30         Train Batch: 160         Training Loss: 1.114228
Epoch: 30         Train Batch: 180         Training Loss: 1.112725
Epoch: 30         Train Batch: 200         Training Loss: 1.109885
Epoch: 30         Eval Batch: 0         Validation Loss: 1.080617
Epoch: 30         Eval Batch: 20         Validation Loss: 1.080617
Epoch: 30         Training Loss: 1.108274         Validation Loss: 1.080617
Epoch: 31         Train Batch: 0         Training Loss: 1.003986
Epoch: 31         Train Batch: 20         Training Loss: 1.078443
Epoch: 31         Train Batch: 40         Training Loss: 1.080112
Epoch: 31         Train Batch: 60         Training Loss: 1.065828
Epoch: 31         Train Batch: 80         Training Loss: 1.071033
Epoch: 31         Train Batch: 100         Training Loss: 1.092935
Epoch: 31         Train Batch: 120         Training Loss: 1.095241
Epoch: 31         Train Batch: 140         Training Loss: 1.092084
Epoch: 31         Train Batch: 160         Training Loss: 1.090672
Epoch: 31         Train Batch: 180         Training Loss: 1.087700
Epoch: 31         Train Batch: 200         Training Loss: 1.096300
Epoch: 31         Eval Batch: 0         Validation Loss: 1.008963
Epoch: 31         Eval Batch: 20         Validation Loss: 1.008963
Epoch: 31         Training Loss: 1.098379         Validation Loss: 1.008963
Epoch: 32         Train Batch: 0         Training Loss: 0.894422
Epoch: 32         Train Batch: 20         Training Loss: 1.104686
Epoch: 32         Train Batch: 40         Training Loss: 1.073379
Epoch: 32         Train Batch: 60         Training Loss: 1.086608
Epoch: 32         Train Batch: 80         Training Loss: 1.098700
Epoch: 32         Train Batch: 100         Training Loss: 1.088116
```

```
Epoch: 32         Train Batch: 120         Training Loss: 1.081070
Epoch: 32         Train Batch: 140         Training Loss: 1.077483
Epoch: 32         Train Batch: 160         Training Loss: 1.083246
Epoch: 32         Train Batch: 180         Training Loss: 1.074317
Epoch: 32         Train Batch: 200         Training Loss: 1.078244
Epoch: 32         Eval Batch: 0         Validation Loss: 0.826962
Epoch: 32         Eval Batch: 20         Validation Loss: 0.826962
Epoch: 32         Training Loss: 1.079242         Validation Loss: 0.826962
Epoch: 33         Train Batch: 0         Training Loss: 1.048591
Epoch: 33         Train Batch: 20         Training Loss: 0.984443
Epoch: 33         Train Batch: 40         Training Loss: 1.043509
Epoch: 33         Train Batch: 60         Training Loss: 1.063080
Epoch: 33         Train Batch: 80         Training Loss: 1.052597
Epoch: 33         Train Batch: 100         Training Loss: 1.058416
Epoch: 33         Train Batch: 120         Training Loss: 1.062484
Epoch: 33         Train Batch: 140         Training Loss: 1.070489
Epoch: 33         Train Batch: 160         Training Loss: 1.074296
Epoch: 33         Train Batch: 180         Training Loss: 1.070058
Epoch: 33         Train Batch: 200         Training Loss: 1.071657
Epoch: 33         Eval Batch: 0         Validation Loss: 0.925807
Epoch: 33         Eval Batch: 20         Validation Loss: 0.925807
Epoch: 33         Training Loss: 1.076237         Validation Loss: 0.925807
Epoch: 34         Train Batch: 0         Training Loss: 0.953997
Epoch: 34         Train Batch: 20         Training Loss: 1.030255
Epoch: 34         Train Batch: 40         Training Loss: 1.060549
Epoch: 34         Train Batch: 60         Training Loss: 1.073457
Epoch: 34         Train Batch: 80         Training Loss: 1.081846
Epoch: 34         Train Batch: 100         Training Loss: 1.075677
Epoch: 34         Train Batch: 120         Training Loss: 1.051874
Epoch: 34         Train Batch: 140         Training Loss: 1.055198
Epoch: 34         Train Batch: 160         Training Loss: 1.051732
Epoch: 34         Train Batch: 180         Training Loss: 1.052600
Epoch: 34         Train Batch: 200         Training Loss: 1.048665
Epoch: 34         Eval Batch: 0         Validation Loss: 1.131051
Epoch: 34         Eval Batch: 20         Validation Loss: 1.131051
Epoch: 34         Training Loss: 1.053279         Validation Loss: 1.131051
Epoch: 35         Train Batch: 0         Training Loss: 0.912687
Epoch: 35         Train Batch: 20         Training Loss: 1.016965
Epoch: 35         Train Batch: 40         Training Loss: 1.057606
Epoch: 35         Train Batch: 60         Training Loss: 1.065962
Epoch: 35         Train Batch: 80         Training Loss: 1.056427
Epoch: 35         Train Batch: 100         Training Loss: 1.053201
Epoch: 35         Train Batch: 120         Training Loss: 1.043539
Epoch: 35         Train Batch: 140         Training Loss: 1.044410
Epoch: 35         Train Batch: 160         Training Loss: 1.052533
Epoch: 35         Train Batch: 180         Training Loss: 1.056711
Epoch: 35         Train Batch: 200         Training Loss: 1.060407
Epoch: 35         Eval Batch: 0         Validation Loss: 0.640335
```

```
Epoch: 35         Eval Batch: 20         Validation Loss: 0.640335
Epoch: 35         Training Loss: 1.056596         Validation Loss: 0.640335
Validation loss decreased (0.825764 --> 0.640335).  Saving model ...
Epoch: 36         Train Batch: 0         Training Loss: 0.996310
Epoch: 36         Train Batch: 20         Training Loss: 0.993075
Epoch: 36         Train Batch: 40         Training Loss: 1.020591
Epoch: 36         Train Batch: 60         Training Loss: 1.048905
Epoch: 36         Train Batch: 80         Training Loss: 1.059419
Epoch: 36         Train Batch: 100         Training Loss: 1.061351
Epoch: 36         Train Batch: 120         Training Loss: 1.044287
Epoch: 36         Train Batch: 140         Training Loss: 1.040406
Epoch: 36         Train Batch: 160         Training Loss: 1.046846
Epoch: 36         Train Batch: 180         Training Loss: 1.047567
Epoch: 36         Train Batch: 200         Training Loss: 1.042378
Epoch: 36         Eval Batch: 0         Validation Loss: 0.834421
Epoch: 36         Eval Batch: 20         Validation Loss: 0.834421
Epoch: 36         Training Loss: 1.039414         Validation Loss: 0.834421
Epoch: 37         Train Batch: 0         Training Loss: 1.091532
Epoch: 37         Train Batch: 20         Training Loss: 1.041458
Epoch: 37         Train Batch: 40         Training Loss: 1.051541
Epoch: 37         Train Batch: 60         Training Loss: 1.047673
Epoch: 37         Train Batch: 80         Training Loss: 1.049582
Epoch: 37         Train Batch: 100         Training Loss: 1.046612
Epoch: 37         Train Batch: 120         Training Loss: 1.026288
Epoch: 37         Train Batch: 140         Training Loss: 1.021718
Epoch: 37         Train Batch: 160         Training Loss: 1.022164
Epoch: 37         Train Batch: 180         Training Loss: 1.020234
Epoch: 37         Train Batch: 200         Training Loss: 1.025960
Epoch: 37         Eval Batch: 0         Validation Loss: 1.317415
Epoch: 37         Eval Batch: 20         Validation Loss: 1.317415
Epoch: 37         Training Loss: 1.029695         Validation Loss: 1.317415
Epoch: 38         Train Batch: 0         Training Loss: 1.201083
Epoch: 38         Train Batch: 20         Training Loss: 0.961597
Epoch: 38         Train Batch: 40         Training Loss: 0.992624
Epoch: 38         Train Batch: 60         Training Loss: 0.976278
Epoch: 38         Train Batch: 80         Training Loss: 0.992328
Epoch: 38         Train Batch: 100         Training Loss: 1.002333
Epoch: 38         Train Batch: 120         Training Loss: 1.004220
Epoch: 38         Train Batch: 140         Training Loss: 1.000574
Epoch: 38         Train Batch: 160         Training Loss: 1.003863
Epoch: 38         Train Batch: 180         Training Loss: 1.006840
Epoch: 38         Train Batch: 200         Training Loss: 1.010652
Epoch: 38         Eval Batch: 0         Validation Loss: 1.331526
Epoch: 38         Eval Batch: 20         Validation Loss: 1.331526
Epoch: 38         Training Loss: 1.013314         Validation Loss: 1.331526
Epoch: 39         Train Batch: 0         Training Loss: 0.960847
Epoch: 39         Train Batch: 20         Training Loss: 0.966685
Epoch: 39         Train Batch: 40         Training Loss: 0.963503
```

```
Epoch: 39         Train Batch: 60         Training Loss: 0.976726
Epoch: 39         Train Batch: 80         Training Loss: 1.003539
Epoch: 39         Train Batch: 100         Training Loss: 1.003537
Epoch: 39         Train Batch: 120         Training Loss: 1.007124
Epoch: 39         Train Batch: 140         Training Loss: 0.999760
Epoch: 39         Train Batch: 160         Training Loss: 0.999635
Epoch: 39         Train Batch: 180         Training Loss: 1.001093
Epoch: 39         Train Batch: 200         Training Loss: 1.002743
Epoch: 39         Eval Batch: 0         Validation Loss: 1.308778
Epoch: 39         Eval Batch: 20         Validation Loss: 1.308778
Epoch: 39         Training Loss: 1.006105         Validation Loss: 1.308778
Epoch: 40         Train Batch: 0         Training Loss: 0.996604
Epoch: 40         Train Batch: 20         Training Loss: 0.999277
Epoch: 40         Train Batch: 40         Training Loss: 0.985909
Epoch: 40         Train Batch: 60         Training Loss: 0.984996
Epoch: 40         Train Batch: 80         Training Loss: 0.974791
Epoch: 40         Train Batch: 100         Training Loss: 0.991717
Epoch: 40         Train Batch: 120         Training Loss: 0.991673
Epoch: 40         Train Batch: 140         Training Loss: 0.993640
Epoch: 40         Train Batch: 160         Training Loss: 0.991098
Epoch: 40         Train Batch: 180         Training Loss: 0.988351
Epoch: 40         Train Batch: 200         Training Loss: 0.993136
Epoch: 40         Eval Batch: 0         Validation Loss: 1.121848
Epoch: 40         Eval Batch: 20         Validation Loss: 1.121848
Epoch: 40         Training Loss: 0.996653         Validation Loss: 1.121848
Epoch: 41         Train Batch: 0         Training Loss: 1.200687
Epoch: 41         Train Batch: 20         Training Loss: 1.006801
Epoch: 41         Train Batch: 40         Training Loss: 1.010739
Epoch: 41         Train Batch: 60         Training Loss: 1.013370
Epoch: 41         Train Batch: 80         Training Loss: 0.998938
Epoch: 41         Train Batch: 100         Training Loss: 0.990946
Epoch: 41         Train Batch: 120         Training Loss: 0.990938
Epoch: 41         Train Batch: 140         Training Loss: 0.991714
Epoch: 41         Train Batch: 160         Training Loss: 0.985993
Epoch: 41         Train Batch: 180         Training Loss: 0.986828
Epoch: 41         Train Batch: 200         Training Loss: 0.990718
Epoch: 41         Eval Batch: 0         Validation Loss: 0.848409
Epoch: 41         Eval Batch: 20         Validation Loss: 0.848409
Epoch: 41         Training Loss: 0.987159         Validation Loss: 0.848409
Epoch: 42         Train Batch: 0         Training Loss: 1.029411
Epoch: 42         Train Batch: 20         Training Loss: 0.918854
Epoch: 42         Train Batch: 40         Training Loss: 0.974183
Epoch: 42         Train Batch: 60         Training Loss: 0.987550
Epoch: 42         Train Batch: 80         Training Loss: 0.973727
Epoch: 42         Train Batch: 100         Training Loss: 0.964455
Epoch: 42         Train Batch: 120         Training Loss: 0.972405
Epoch: 42         Train Batch: 140         Training Loss: 0.966246
Epoch: 42         Train Batch: 160         Training Loss: 0.969729
```

```
Epoch: 42         Train Batch: 180         Training Loss: 0.968618
Epoch: 42         Train Batch: 200         Training Loss: 0.973460
Epoch: 42         Eval Batch: 0         Validation Loss: 0.902172
Epoch: 42         Eval Batch: 20         Validation Loss: 0.902172
Epoch: 42         Training Loss: 0.971901         Validation Loss: 0.902172
Epoch: 43         Train Batch: 0         Training Loss: 1.214155
Epoch: 43         Train Batch: 20         Training Loss: 1.005993
Epoch: 43         Train Batch: 40         Training Loss: 1.029403
Epoch: 43         Train Batch: 60         Training Loss: 1.011093
Epoch: 43         Train Batch: 80         Training Loss: 1.009450
Epoch: 43         Train Batch: 100         Training Loss: 1.001116
Epoch: 43         Train Batch: 120         Training Loss: 0.999513
Epoch: 43         Train Batch: 140         Training Loss: 1.000473
Epoch: 43         Train Batch: 160         Training Loss: 0.994437
Epoch: 43         Train Batch: 180         Training Loss: 0.991850
Epoch: 43         Train Batch: 200         Training Loss: 0.988889
Epoch: 43         Eval Batch: 0         Validation Loss: 1.385436
Epoch: 43         Eval Batch: 20         Validation Loss: 1.385436
Epoch: 43         Training Loss: 0.990950         Validation Loss: 1.385436
Epoch: 44         Train Batch: 0         Training Loss: 1.122586
Epoch: 44         Train Batch: 20         Training Loss: 0.944617
Epoch: 44         Train Batch: 40         Training Loss: 0.970384
Epoch: 44         Train Batch: 60         Training Loss: 0.999726
Epoch: 44         Train Batch: 80         Training Loss: 0.993731
Epoch: 44         Train Batch: 100         Training Loss: 0.985761
Epoch: 44         Train Batch: 120         Training Loss: 0.993508
Epoch: 44         Train Batch: 140         Training Loss: 0.988026
Epoch: 44         Train Batch: 160         Training Loss: 0.986813
Epoch: 44         Train Batch: 180         Training Loss: 0.990678
Epoch: 44         Train Batch: 200         Training Loss: 0.981572
Epoch: 44         Eval Batch: 0         Validation Loss: 0.819055
Epoch: 44         Eval Batch: 20         Validation Loss: 0.819055
Epoch: 44         Training Loss: 0.981914         Validation Loss: 0.819055
Epoch: 45         Train Batch: 0         Training Loss: 0.862791
Epoch: 45         Train Batch: 20         Training Loss: 0.944305
Epoch: 45         Train Batch: 40         Training Loss: 0.997502
Epoch: 45         Train Batch: 60         Training Loss: 0.975167
Epoch: 45         Train Batch: 80         Training Loss: 0.973066
Epoch: 45         Train Batch: 100         Training Loss: 0.965349
Epoch: 45         Train Batch: 120         Training Loss: 0.963873
Epoch: 45         Train Batch: 140         Training Loss: 0.969206
Epoch: 45         Train Batch: 160         Training Loss: 0.968774
Epoch: 45         Train Batch: 180         Training Loss: 0.968360
Epoch: 45         Train Batch: 200         Training Loss: 0.968820
Epoch: 45         Eval Batch: 0         Validation Loss: 0.904916
Epoch: 45         Eval Batch: 20         Validation Loss: 0.904916
Epoch: 45         Training Loss: 0.968125         Validation Loss: 0.904916
Epoch: 46         Train Batch: 0         Training Loss: 0.828862
```

```
Epoch: 46         Train Batch: 20         Training Loss: 0.913906
Epoch: 46         Train Batch: 40         Training Loss: 0.909774
Epoch: 46         Train Batch: 60         Training Loss: 0.920414
Epoch: 46         Train Batch: 80         Training Loss: 0.916676
Epoch: 46         Train Batch: 100         Training Loss: 0.923414
Epoch: 46         Train Batch: 120         Training Loss: 0.924719
Epoch: 46         Train Batch: 140         Training Loss: 0.926000
Epoch: 46         Train Batch: 160         Training Loss: 0.923758
Epoch: 46         Train Batch: 180         Training Loss: 0.928495
Epoch: 46         Train Batch: 200         Training Loss: 0.931451
Epoch: 46         Eval Batch: 0         Validation Loss: 1.242669
Epoch: 46         Eval Batch: 20         Validation Loss: 1.242669
Epoch: 46         Training Loss: 0.939760         Validation Loss: 1.242669
Epoch: 47         Train Batch: 0         Training Loss: 0.810491
Epoch: 47         Train Batch: 20         Training Loss: 0.932468
Epoch: 47         Train Batch: 40         Training Loss: 0.935700
Epoch: 47         Train Batch: 60         Training Loss: 0.941983
Epoch: 47         Train Batch: 80         Training Loss: 0.958732
Epoch: 47         Train Batch: 100         Training Loss: 0.952341
Epoch: 47         Train Batch: 120         Training Loss: 0.933976
Epoch: 47         Train Batch: 140         Training Loss: 0.939844
Epoch: 47         Train Batch: 160         Training Loss: 0.942510
Epoch: 47         Train Batch: 180         Training Loss: 0.944307
Epoch: 47         Train Batch: 200         Training Loss: 0.941285
Epoch: 47         Eval Batch: 0         Validation Loss: 0.614379
Epoch: 47         Eval Batch: 20         Validation Loss: 0.614379
Epoch: 47         Training Loss: 0.941837         Validation Loss: 0.614379
Validation loss decreased (0.640335 --> 0.614379).  Saving model ...
Epoch: 48         Train Batch: 0         Training Loss: 0.891555
Epoch: 48         Train Batch: 20         Training Loss: 0.884106
Epoch: 48         Train Batch: 40         Training Loss: 0.902522
Epoch: 48         Train Batch: 60         Training Loss: 0.933365
Epoch: 48         Train Batch: 80         Training Loss: 0.931278
Epoch: 48         Train Batch: 100         Training Loss: 0.932485
Epoch: 48         Train Batch: 120         Training Loss: 0.929951
Epoch: 48         Train Batch: 140         Training Loss: 0.930514
Epoch: 48         Train Batch: 160         Training Loss: 0.929057
Epoch: 48         Train Batch: 180         Training Loss: 0.924723
Epoch: 48         Train Batch: 200         Training Loss: 0.934219
Epoch: 48         Eval Batch: 0         Validation Loss: 1.636785
Epoch: 48         Eval Batch: 20         Validation Loss: 1.636785
Epoch: 48         Training Loss: 0.934654         Validation Loss: 1.636785
Epoch: 49         Train Batch: 0         Training Loss: 1.006143
Epoch: 49         Train Batch: 20         Training Loss: 0.912064
Epoch: 49         Train Batch: 40         Training Loss: 0.940434
Epoch: 49         Train Batch: 60         Training Loss: 0.941717
Epoch: 49         Train Batch: 80         Training Loss: 0.957894
Epoch: 49         Train Batch: 100         Training Loss: 0.950272
```

```
Epoch: 49         Train Batch: 120         Training Loss: 0.932278
Epoch: 49         Train Batch: 140         Training Loss: 0.923628
Epoch: 49         Train Batch: 160         Training Loss: 0.919151
Epoch: 49         Train Batch: 180         Training Loss: 0.920074
Epoch: 49         Train Batch: 200         Training Loss: 0.925415
Epoch: 49         Eval Batch: 0         Validation Loss: 1.069295
Epoch: 49         Eval Batch: 20         Validation Loss: 1.069295
Epoch: 49         Training Loss: 0.930060         Validation Loss: 1.069295
Epoch: 50         Train Batch: 0         Training Loss: 1.125833
Epoch: 50         Train Batch: 20         Training Loss: 0.933789
Epoch: 50         Train Batch: 40         Training Loss: 0.922044
Epoch: 50         Train Batch: 60         Training Loss: 0.915905
Epoch: 50         Train Batch: 80         Training Loss: 0.907504
Epoch: 50         Train Batch: 100         Training Loss: 0.894228
Epoch: 50         Train Batch: 120         Training Loss: 0.897373
Epoch: 50         Train Batch: 140         Training Loss: 0.902547
Epoch: 50         Train Batch: 160         Training Loss: 0.909409
Epoch: 50         Train Batch: 180         Training Loss: 0.911535
Epoch: 50         Train Batch: 200         Training Loss: 0.906602
Epoch: 50         Eval Batch: 0         Validation Loss: 1.014253
Epoch: 50         Eval Batch: 20         Validation Loss: 1.014253
Epoch: 50         Training Loss: 0.907663         Validation Loss: 1.014253
Epoch: 51         Train Batch: 0         Training Loss: 1.172751
Epoch: 51         Train Batch: 20         Training Loss: 0.959491
Epoch: 51         Train Batch: 40         Training Loss: 0.948565
Epoch: 51         Train Batch: 60         Training Loss: 0.943412
Epoch: 51         Train Batch: 80         Training Loss: 0.930385
Epoch: 51         Train Batch: 100         Training Loss: 0.928831
Epoch: 51         Train Batch: 120         Training Loss: 0.931588
Epoch: 51         Train Batch: 140         Training Loss: 0.925346
Epoch: 51         Train Batch: 160         Training Loss: 0.928665
Epoch: 51         Train Batch: 180         Training Loss: 0.924479
Epoch: 51         Train Batch: 200         Training Loss: 0.931512
Epoch: 51         Eval Batch: 0         Validation Loss: 0.769488
Epoch: 51         Eval Batch: 20         Validation Loss: 0.769488
Epoch: 51         Training Loss: 0.930588         Validation Loss: 0.769488
Epoch: 52         Train Batch: 0         Training Loss: 0.719136
Epoch: 52         Train Batch: 20         Training Loss: 0.996469
Epoch: 52         Train Batch: 40         Training Loss: 0.955067
Epoch: 52         Train Batch: 60         Training Loss: 0.930621
Epoch: 52         Train Batch: 80         Training Loss: 0.943792
Epoch: 52         Train Batch: 100         Training Loss: 0.942336
Epoch: 52         Train Batch: 120         Training Loss: 0.941758
Epoch: 52         Train Batch: 140         Training Loss: 0.945772
Epoch: 52         Train Batch: 160         Training Loss: 0.941330
Epoch: 52         Train Batch: 180         Training Loss: 0.931051
Epoch: 52         Train Batch: 200         Training Loss: 0.926770
Epoch: 52         Eval Batch: 0         Validation Loss: 1.087619
```

```
Epoch: 52         Eval Batch: 20         Validation Loss: 1.087619
Epoch: 52         Training Loss: 0.923301         Validation Loss: 1.087619
Epoch: 53         Train Batch: 0         Training Loss: 0.554940
Epoch: 53         Train Batch: 20         Training Loss: 0.859779
Epoch: 53         Train Batch: 40         Training Loss: 0.914012
Epoch: 53         Train Batch: 60         Training Loss: 0.902240
Epoch: 53         Train Batch: 80         Training Loss: 0.880384
Epoch: 53         Train Batch: 100         Training Loss: 0.893429
Epoch: 53         Train Batch: 120         Training Loss: 0.913889
Epoch: 53         Train Batch: 140         Training Loss: 0.916090
Epoch: 53         Train Batch: 160         Training Loss: 0.914043
Epoch: 53         Train Batch: 180         Training Loss: 0.916644
Epoch: 53         Train Batch: 200         Training Loss: 0.918862
Epoch: 53         Eval Batch: 0         Validation Loss: 1.042228
Epoch: 53         Eval Batch: 20         Validation Loss: 1.042228
Epoch: 53         Training Loss: 0.919753         Validation Loss: 1.042228
Epoch: 54         Train Batch: 0         Training Loss: 0.840169
Epoch: 54         Train Batch: 20         Training Loss: 0.899601
Epoch: 54         Train Batch: 40         Training Loss: 0.895870
Epoch: 54         Train Batch: 60         Training Loss: 0.905875
Epoch: 54         Train Batch: 80         Training Loss: 0.908877
Epoch: 54         Train Batch: 100         Training Loss: 0.912810
Epoch: 54         Train Batch: 120         Training Loss: 0.906086
Epoch: 54         Train Batch: 140         Training Loss: 0.906835
Epoch: 54         Train Batch: 160         Training Loss: 0.906609
Epoch: 54         Train Batch: 180         Training Loss: 0.912662
Epoch: 54         Train Batch: 200         Training Loss: 0.910858
Epoch: 54         Eval Batch: 0         Validation Loss: 1.123983
Epoch: 54         Eval Batch: 20         Validation Loss: 1.123983
Epoch: 54         Training Loss: 0.912222         Validation Loss: 1.123983
Epoch: 55         Train Batch: 0         Training Loss: 0.863166
Epoch: 55         Train Batch: 20         Training Loss: 0.919134
Epoch: 55         Train Batch: 40         Training Loss: 0.916080
Epoch: 55         Train Batch: 60         Training Loss: 0.891484
Epoch: 55         Train Batch: 80         Training Loss: 0.913048
Epoch: 55         Train Batch: 100         Training Loss: 0.912141
Epoch: 55         Train Batch: 120         Training Loss: 0.921934
Epoch: 55         Train Batch: 140         Training Loss: 0.918873
Epoch: 55         Train Batch: 160         Training Loss: 0.917205
Epoch: 55         Train Batch: 180         Training Loss: 0.909787
Epoch: 55         Train Batch: 200         Training Loss: 0.911400
Epoch: 55         Eval Batch: 0         Validation Loss: 0.735391
Epoch: 55         Eval Batch: 20         Validation Loss: 0.735391
Epoch: 55         Training Loss: 0.907422         Validation Loss: 0.735391
Epoch: 56         Train Batch: 0         Training Loss: 1.460517
Epoch: 56         Train Batch: 20         Training Loss: 0.869860
Epoch: 56         Train Batch: 40         Training Loss: 0.890787
Epoch: 56         Train Batch: 60         Training Loss: 0.913296
```

```
Epoch: 56         Train Batch: 80         Training Loss: 0.906127
Epoch: 56         Train Batch: 100         Training Loss: 0.914286
Epoch: 56         Train Batch: 120         Training Loss: 0.910778
Epoch: 56         Train Batch: 140         Training Loss: 0.906572
Epoch: 56         Train Batch: 160         Training Loss: 0.904929
Epoch: 56         Train Batch: 180         Training Loss: 0.904952
Epoch: 56         Train Batch: 200         Training Loss: 0.901395
Epoch: 56         Eval Batch: 0         Validation Loss: 0.870257
Epoch: 56         Eval Batch: 20         Validation Loss: 0.870257
Epoch: 56         Training Loss: 0.900722         Validation Loss: 0.870257
Epoch: 57         Train Batch: 0         Training Loss: 0.906975
Epoch: 57         Train Batch: 20         Training Loss: 0.906057
Epoch: 57         Train Batch: 40         Training Loss: 0.894354
Epoch: 57         Train Batch: 60         Training Loss: 0.879443
Epoch: 57         Train Batch: 80         Training Loss: 0.883998
Epoch: 57         Train Batch: 100         Training Loss: 0.884918
Epoch: 57         Train Batch: 120         Training Loss: 0.880807
Epoch: 57         Train Batch: 140         Training Loss: 0.882229
Epoch: 57         Train Batch: 160         Training Loss: 0.876755
Epoch: 57         Train Batch: 180         Training Loss: 0.877424
Epoch: 57         Train Batch: 200         Training Loss: 0.878492
Epoch: 57         Eval Batch: 0         Validation Loss: 1.205946
Epoch: 57         Eval Batch: 20         Validation Loss: 1.205946
Epoch: 57         Training Loss: 0.877406         Validation Loss: 1.205946
Epoch: 58         Train Batch: 0         Training Loss: 0.923905
Epoch: 58         Train Batch: 20         Training Loss: 0.841583
Epoch: 58         Train Batch: 40         Training Loss: 0.849209
Epoch: 58         Train Batch: 60         Training Loss: 0.864150
Epoch: 58         Train Batch: 80         Training Loss: 0.872604
Epoch: 58         Train Batch: 100         Training Loss: 0.882491
Epoch: 58         Train Batch: 120         Training Loss: 0.892765
Epoch: 58         Train Batch: 140         Training Loss: 0.895074
Epoch: 58         Train Batch: 160         Training Loss: 0.889725
Epoch: 58         Train Batch: 180         Training Loss: 0.892820
Epoch: 58         Train Batch: 200         Training Loss: 0.890779
Epoch: 58         Eval Batch: 0         Validation Loss: 1.174918
Epoch: 58         Eval Batch: 20         Validation Loss: 1.174918
Epoch: 58         Training Loss: 0.895283         Validation Loss: 1.174918
Epoch: 59         Train Batch: 0         Training Loss: 0.729715
Epoch: 59         Train Batch: 20         Training Loss: 0.838431
Epoch: 59         Train Batch: 40         Training Loss: 0.856268
Epoch: 59         Train Batch: 60         Training Loss: 0.854824
Epoch: 59         Train Batch: 80         Training Loss: 0.869470
Epoch: 59         Train Batch: 100         Training Loss: 0.859817
Epoch: 59         Train Batch: 120         Training Loss: 0.852138
Epoch: 59         Train Batch: 140         Training Loss: 0.865885
Epoch: 59         Train Batch: 160         Training Loss: 0.875570
Epoch: 59         Train Batch: 180         Training Loss: 0.875899
```

```
Epoch: 59         Train Batch: 200         Training Loss: 0.874311
Epoch: 59         Eval Batch: 0         Validation Loss: 0.707016
Epoch: 59         Eval Batch: 20         Validation Loss: 0.707016
Epoch: 59         Training Loss: 0.876472         Validation Loss: 0.707016
Epoch: 60         Train Batch: 0         Training Loss: 0.797787
Epoch: 60         Train Batch: 20         Training Loss: 0.895618
Epoch: 60         Train Batch: 40         Training Loss: 0.892839
Epoch: 60         Train Batch: 60         Training Loss: 0.885546
Epoch: 60         Train Batch: 80         Training Loss: 0.875521
Epoch: 60         Train Batch: 100         Training Loss: 0.892989
Epoch: 60         Train Batch: 120         Training Loss: 0.882396
Epoch: 60         Train Batch: 140         Training Loss: 0.884292
Epoch: 60         Train Batch: 160         Training Loss: 0.885457
Epoch: 60         Train Batch: 180         Training Loss: 0.885331
Epoch: 60         Train Batch: 200         Training Loss: 0.885963
Epoch: 60         Eval Batch: 0         Validation Loss: 1.003081
Epoch: 60         Eval Batch: 20         Validation Loss: 1.003081
Epoch: 60         Training Loss: 0.889016         Validation Loss: 1.003081
Epoch: 61         Train Batch: 0         Training Loss: 0.999407
Epoch: 61         Train Batch: 20         Training Loss: 0.838490
Epoch: 61         Train Batch: 40         Training Loss: 0.844193
Epoch: 61         Train Batch: 60         Training Loss: 0.850098
Epoch: 61         Train Batch: 80         Training Loss: 0.868293
Epoch: 61         Train Batch: 100         Training Loss: 0.879913
Epoch: 61         Train Batch: 120         Training Loss: 0.879446
Epoch: 61         Train Batch: 140         Training Loss: 0.872887
Epoch: 61         Train Batch: 160         Training Loss: 0.871547
Epoch: 61         Train Batch: 180         Training Loss: 0.874885
Epoch: 61         Train Batch: 200         Training Loss: 0.879119
Epoch: 61         Eval Batch: 0         Validation Loss: 1.233456
Epoch: 61         Eval Batch: 20         Validation Loss: 1.233456
Epoch: 61         Training Loss: 0.880475         Validation Loss: 1.233456
Epoch: 62         Train Batch: 0         Training Loss: 1.001419
Epoch: 62         Train Batch: 20         Training Loss: 0.813440
Epoch: 62         Train Batch: 40         Training Loss: 0.825244
Epoch: 62         Train Batch: 60         Training Loss: 0.859469
Epoch: 62         Train Batch: 80         Training Loss: 0.858575
Epoch: 62         Train Batch: 100         Training Loss: 0.853532
Epoch: 62         Train Batch: 120         Training Loss: 0.859089
Epoch: 62         Train Batch: 140         Training Loss: 0.859638
Epoch: 62         Train Batch: 160         Training Loss: 0.853580
Epoch: 62         Train Batch: 180         Training Loss: 0.848411
Epoch: 62         Train Batch: 200         Training Loss: 0.851417
Epoch: 62         Eval Batch: 0         Validation Loss: 1.122277
Epoch: 62         Eval Batch: 20         Validation Loss: 1.122277
Epoch: 62         Training Loss: 0.856537         Validation Loss: 1.122277
Epoch: 63         Train Batch: 0         Training Loss: 0.594283
Epoch: 63         Train Batch: 20         Training Loss: 0.832708
```

```
Epoch: 63         Train Batch: 40         Training Loss: 0.846124
Epoch: 63         Train Batch: 60         Training Loss: 0.856441
Epoch: 63         Train Batch: 80         Training Loss: 0.857669
Epoch: 63         Train Batch: 100         Training Loss: 0.852833
Epoch: 63         Train Batch: 120         Training Loss: 0.852998
Epoch: 63         Train Batch: 140         Training Loss: 0.852052
Epoch: 63         Train Batch: 160         Training Loss: 0.854454
Epoch: 63         Train Batch: 180         Training Loss: 0.853394
Epoch: 63         Train Batch: 200         Training Loss: 0.857867
Epoch: 63         Eval Batch: 0         Validation Loss: 1.155640
Epoch: 63         Eval Batch: 20         Validation Loss: 1.155640
Epoch: 63         Training Loss: 0.859016         Validation Loss: 1.155640
Epoch: 64         Train Batch: 0         Training Loss: 0.531287
Epoch: 64         Train Batch: 20         Training Loss: 0.817960
Epoch: 64         Train Batch: 40         Training Loss: 0.820157
Epoch: 64         Train Batch: 60         Training Loss: 0.836954
Epoch: 64         Train Batch: 80         Training Loss: 0.843824
Epoch: 64         Train Batch: 100         Training Loss: 0.835540
Epoch: 64         Train Batch: 120         Training Loss: 0.839868
Epoch: 64         Train Batch: 140         Training Loss: 0.838576
Epoch: 64         Train Batch: 160         Training Loss: 0.839254
Epoch: 64         Train Batch: 180         Training Loss: 0.842681
Epoch: 64         Train Batch: 200         Training Loss: 0.852214
Epoch: 64         Eval Batch: 0         Validation Loss: 0.874846
Epoch: 64         Eval Batch: 20         Validation Loss: 0.874846
Epoch: 64         Training Loss: 0.852157         Validation Loss: 0.874846
Epoch: 65         Train Batch: 0         Training Loss: 0.786365
Epoch: 65         Train Batch: 20         Training Loss: 0.802161
Epoch: 65         Train Batch: 40         Training Loss: 0.831747
Epoch: 65         Train Batch: 60         Training Loss: 0.833991
Epoch: 65         Train Batch: 80         Training Loss: 0.845876
Epoch: 65         Train Batch: 100         Training Loss: 0.852181
Epoch: 65         Train Batch: 120         Training Loss: 0.859224
Epoch: 65         Train Batch: 140         Training Loss: 0.863289
Epoch: 65         Train Batch: 160         Training Loss: 0.863365
Epoch: 65         Train Batch: 180         Training Loss: 0.858254
Epoch: 65         Train Batch: 200         Training Loss: 0.862840
Epoch: 65         Eval Batch: 0         Validation Loss: 0.616739
Epoch: 65         Eval Batch: 20         Validation Loss: 0.616739
Epoch: 65         Training Loss: 0.857821         Validation Loss: 0.616739
Epoch: 66         Train Batch: 0         Training Loss: 0.915289
Epoch: 66         Train Batch: 20         Training Loss: 0.808216
Epoch: 66         Train Batch: 40         Training Loss: 0.826667
Epoch: 66         Train Batch: 60         Training Loss: 0.838068
Epoch: 66         Train Batch: 80         Training Loss: 0.844278
Epoch: 66         Train Batch: 100         Training Loss: 0.843677
Epoch: 66         Train Batch: 120         Training Loss: 0.843184
Epoch: 66         Train Batch: 140         Training Loss: 0.847789
```

```
Epoch: 66         Train Batch: 160         Training Loss: 0.843401
Epoch: 66         Train Batch: 180         Training Loss: 0.849969
Epoch: 66         Train Batch: 200         Training Loss: 0.852583
Epoch: 66         Eval Batch: 0         Validation Loss: 1.214841
Epoch: 66         Eval Batch: 20         Validation Loss: 1.214841
Epoch: 66         Training Loss: 0.850959         Validation Loss: 1.214841
Epoch: 67         Train Batch: 0         Training Loss: 0.922314
Epoch: 67         Train Batch: 20         Training Loss: 0.791963
Epoch: 67         Train Batch: 40         Training Loss: 0.789850
Epoch: 67         Train Batch: 60         Training Loss: 0.804077
Epoch: 67         Train Batch: 80         Training Loss: 0.815120
Epoch: 67         Train Batch: 100         Training Loss: 0.828684
Epoch: 67         Train Batch: 120         Training Loss: 0.826624
Epoch: 67         Train Batch: 140         Training Loss: 0.832315
Epoch: 67         Train Batch: 160         Training Loss: 0.839585
Epoch: 67         Train Batch: 180         Training Loss: 0.845939
Epoch: 67         Train Batch: 200         Training Loss: 0.845499
Epoch: 67         Eval Batch: 0         Validation Loss: 0.795884
Epoch: 67         Eval Batch: 20         Validation Loss: 0.795884
Epoch: 67         Training Loss: 0.847458         Validation Loss: 0.795884
Epoch: 68         Train Batch: 0         Training Loss: 0.767566
Epoch: 68         Train Batch: 20         Training Loss: 0.791660
Epoch: 68         Train Batch: 40         Training Loss: 0.803301
Epoch: 68         Train Batch: 60         Training Loss: 0.825299
Epoch: 68         Train Batch: 80         Training Loss: 0.827228
Epoch: 68         Train Batch: 100         Training Loss: 0.825342
Epoch: 68         Train Batch: 120         Training Loss: 0.829813
Epoch: 68         Train Batch: 140         Training Loss: 0.828428
Epoch: 68         Train Batch: 160         Training Loss: 0.833384
Epoch: 68         Train Batch: 180         Training Loss: 0.826494
Epoch: 68         Train Batch: 200         Training Loss: 0.829358
Epoch: 68         Eval Batch: 0         Validation Loss: 0.903700
Epoch: 68         Eval Batch: 20         Validation Loss: 0.903700
Epoch: 68         Training Loss: 0.828984         Validation Loss: 0.903700
Epoch: 69         Train Batch: 0         Training Loss: 0.830450
Epoch: 69         Train Batch: 20         Training Loss: 0.819454
Epoch: 69         Train Batch: 40         Training Loss: 0.829051
Epoch: 69         Train Batch: 60         Training Loss: 0.852670
Epoch: 69         Train Batch: 80         Training Loss: 0.840254
Epoch: 69         Train Batch: 100         Training Loss: 0.827845
Epoch: 69         Train Batch: 120         Training Loss: 0.821369
Epoch: 69         Train Batch: 140         Training Loss: 0.824852
Epoch: 69         Train Batch: 160         Training Loss: 0.828939
Epoch: 69         Train Batch: 180         Training Loss: 0.825829
Epoch: 69         Train Batch: 200         Training Loss: 0.823990
Epoch: 69         Eval Batch: 0         Validation Loss: 1.166280
Epoch: 69         Eval Batch: 20         Validation Loss: 1.166280
Epoch: 69         Training Loss: 0.828641         Validation Loss: 1.166280
```

```
Epoch: 70         Train Batch: 0         Training Loss: 0.855678
Epoch: 70         Train Batch: 20         Training Loss: 0.854668
Epoch: 70         Train Batch: 40         Training Loss: 0.857748
Epoch: 70         Train Batch: 60         Training Loss: 0.863984
Epoch: 70         Train Batch: 80         Training Loss: 0.862168
Epoch: 70         Train Batch: 100         Training Loss: 0.857786
Epoch: 70         Train Batch: 120         Training Loss: 0.854467
Epoch: 70         Train Batch: 140         Training Loss: 0.838907
Epoch: 70         Train Batch: 160         Training Loss: 0.834736
Epoch: 70         Train Batch: 180         Training Loss: 0.835525
Epoch: 70         Train Batch: 200         Training Loss: 0.832617
Epoch: 70         Eval Batch: 0         Validation Loss: 0.617536
Epoch: 70         Eval Batch: 20         Validation Loss: 0.617536
Epoch: 70         Training Loss: 0.831050         Validation Loss: 0.617536
Epoch: 71         Train Batch: 0         Training Loss: 0.802919
Epoch: 71         Train Batch: 20         Training Loss: 0.794026
Epoch: 71         Train Batch: 40         Training Loss: 0.822183
Epoch: 71         Train Batch: 60         Training Loss: 0.817304
Epoch: 71         Train Batch: 80         Training Loss: 0.832362
Epoch: 71         Train Batch: 100         Training Loss: 0.835209
Epoch: 71         Train Batch: 120         Training Loss: 0.838125
Epoch: 71         Train Batch: 140         Training Loss: 0.830716
Epoch: 71         Train Batch: 160         Training Loss: 0.829416
Epoch: 71         Train Batch: 180         Training Loss: 0.830801
Epoch: 71         Train Batch: 200         Training Loss: 0.835404
Epoch: 71         Eval Batch: 0         Validation Loss: 0.755997
Epoch: 71         Eval Batch: 20         Validation Loss: 0.755997
Epoch: 71         Training Loss: 0.830345         Validation Loss: 0.755997
Epoch: 72         Train Batch: 0         Training Loss: 1.074307
Epoch: 72         Train Batch: 20         Training Loss: 0.839334
Epoch: 72         Train Batch: 40         Training Loss: 0.842056
Epoch: 72         Train Batch: 60         Training Loss: 0.816062
Epoch: 72         Train Batch: 80         Training Loss: 0.813488
Epoch: 72         Train Batch: 100         Training Loss: 0.815870
Epoch: 72         Train Batch: 120         Training Loss: 0.809316
Epoch: 72         Train Batch: 140         Training Loss: 0.809889
Epoch: 72         Train Batch: 160         Training Loss: 0.811991
Epoch: 72         Train Batch: 180         Training Loss: 0.815895
Epoch: 72         Train Batch: 200         Training Loss: 0.811850
Epoch: 72         Eval Batch: 0         Validation Loss: 1.181954
Epoch: 72         Eval Batch: 20         Validation Loss: 1.181954
Epoch: 72         Training Loss: 0.813666         Validation Loss: 1.181954
Epoch: 73         Train Batch: 0         Training Loss: 0.948049
Epoch: 73         Train Batch: 20         Training Loss: 0.771998
Epoch: 73         Train Batch: 40         Training Loss: 0.798592
Epoch: 73         Train Batch: 60         Training Loss: 0.815577
Epoch: 73         Train Batch: 80         Training Loss: 0.826257
Epoch: 73         Train Batch: 100         Training Loss: 0.829296
```

```
Epoch: 73         Train Batch: 120         Training Loss: 0.838276
Epoch: 73         Train Batch: 140         Training Loss: 0.845429
Epoch: 73         Train Batch: 160         Training Loss: 0.838592
Epoch: 73         Train Batch: 180         Training Loss: 0.835348
Epoch: 73         Train Batch: 200         Training Loss: 0.836592
Epoch: 73         Eval Batch: 0         Validation Loss: 0.890671
Epoch: 73         Eval Batch: 20         Validation Loss: 0.890671
Epoch: 73         Training Loss: 0.839066         Validation Loss: 0.890671
Epoch: 74         Train Batch: 0         Training Loss: 0.981816
Epoch: 74         Train Batch: 20         Training Loss: 0.841214
Epoch: 74         Train Batch: 40         Training Loss: 0.840903
Epoch: 74         Train Batch: 60         Training Loss: 0.844350
Epoch: 74         Train Batch: 80         Training Loss: 0.837687
Epoch: 74         Train Batch: 100         Training Loss: 0.820631
Epoch: 74         Train Batch: 120         Training Loss: 0.819098
Epoch: 74         Train Batch: 140         Training Loss: 0.825695
Epoch: 74         Train Batch: 160         Training Loss: 0.827652
Epoch: 74         Train Batch: 180         Training Loss: 0.821728
Epoch: 74         Train Batch: 200         Training Loss: 0.813396
Epoch: 74         Eval Batch: 0         Validation Loss: 0.683960
Epoch: 74         Eval Batch: 20         Validation Loss: 0.683960
Epoch: 74         Training Loss: 0.812860         Validation Loss: 0.683960
Epoch: 75         Train Batch: 0         Training Loss: 0.704537
Epoch: 75         Train Batch: 20         Training Loss: 0.743473
Epoch: 75         Train Batch: 40         Training Loss: 0.788121
Epoch: 75         Train Batch: 60         Training Loss: 0.817536
Epoch: 75         Train Batch: 80         Training Loss: 0.822627
Epoch: 75         Train Batch: 100         Training Loss: 0.804269
Epoch: 75         Train Batch: 120         Training Loss: 0.814881
Epoch: 75         Train Batch: 140         Training Loss: 0.820422
Epoch: 75         Train Batch: 160         Training Loss: 0.821755
Epoch: 75         Train Batch: 180         Training Loss: 0.824895
Epoch: 75         Train Batch: 200         Training Loss: 0.819211
Epoch: 75         Eval Batch: 0         Validation Loss: 0.670229
Epoch: 75         Eval Batch: 20         Validation Loss: 0.670229
Epoch: 75         Training Loss: 0.818082         Validation Loss: 0.670229
Epoch: 76         Train Batch: 0         Training Loss: 0.823649
Epoch: 76         Train Batch: 20         Training Loss: 0.823652
Epoch: 76         Train Batch: 40         Training Loss: 0.794596
Epoch: 76         Train Batch: 60         Training Loss: 0.787539
Epoch: 76         Train Batch: 80         Training Loss: 0.802520
Epoch: 76         Train Batch: 100         Training Loss: 0.813141
Epoch: 76         Train Batch: 120         Training Loss: 0.806280
Epoch: 76         Train Batch: 140         Training Loss: 0.814188
Epoch: 76         Train Batch: 160         Training Loss: 0.810999
Epoch: 76         Train Batch: 180         Training Loss: 0.813473
Epoch: 76         Train Batch: 200         Training Loss: 0.813669
Epoch: 76         Eval Batch: 0         Validation Loss: 1.049115
```

```
Epoch: 76         Eval Batch: 20         Validation Loss: 1.049115
Epoch: 76         Training Loss: 0.812245         Validation Loss: 1.049115
Epoch: 77         Train Batch: 0         Training Loss: 0.879458
Epoch: 77         Train Batch: 20         Training Loss: 0.783063
Epoch: 77         Train Batch: 40         Training Loss: 0.795714
Epoch: 77         Train Batch: 60         Training Loss: 0.796921
Epoch: 77         Train Batch: 80         Training Loss: 0.800561
Epoch: 77         Train Batch: 100         Training Loss: 0.800571
Epoch: 77         Train Batch: 120         Training Loss: 0.797190
Epoch: 77         Train Batch: 140         Training Loss: 0.808256
Epoch: 77         Train Batch: 160         Training Loss: 0.813230
Epoch: 77         Train Batch: 180         Training Loss: 0.810621
Epoch: 77         Train Batch: 200         Training Loss: 0.815128
Epoch: 77         Eval Batch: 0         Validation Loss: 0.804971
Epoch: 77         Eval Batch: 20         Validation Loss: 0.804971
Epoch: 77         Training Loss: 0.813956         Validation Loss: 0.804971
Epoch: 78         Train Batch: 0         Training Loss: 0.546466
Epoch: 78         Train Batch: 20         Training Loss: 0.767908
Epoch: 78         Train Batch: 40         Training Loss: 0.782457
Epoch: 78         Train Batch: 60         Training Loss: 0.814968
Epoch: 78         Train Batch: 80         Training Loss: 0.846647
Epoch: 78         Train Batch: 100         Training Loss: 0.835785
Epoch: 78         Train Batch: 120         Training Loss: 0.824582
Epoch: 78         Train Batch: 140         Training Loss: 0.826466
Epoch: 78         Train Batch: 160         Training Loss: 0.820260
Epoch: 78         Train Batch: 180         Training Loss: 0.818877
Epoch: 78         Train Batch: 200         Training Loss: 0.816336
Epoch: 78         Eval Batch: 0         Validation Loss: 0.622970
Epoch: 78         Eval Batch: 20         Validation Loss: 0.622970
Epoch: 78         Training Loss: 0.817209         Validation Loss: 0.622970
Epoch: 79         Train Batch: 0         Training Loss: 0.743323
Epoch: 79         Train Batch: 20         Training Loss: 0.878979
Epoch: 79         Train Batch: 40         Training Loss: 0.816130
Epoch: 79         Train Batch: 60         Training Loss: 0.797989
Epoch: 79         Train Batch: 80         Training Loss: 0.800629
Epoch: 79         Train Batch: 100         Training Loss: 0.792922
Epoch: 79         Train Batch: 120         Training Loss: 0.800591
Epoch: 79         Train Batch: 140         Training Loss: 0.806267
Epoch: 79         Train Batch: 160         Training Loss: 0.800920
Epoch: 79         Train Batch: 180         Training Loss: 0.796861
Epoch: 79         Train Batch: 200         Training Loss: 0.797248
Epoch: 79         Eval Batch: 0         Validation Loss: 0.457625
Epoch: 79         Eval Batch: 20         Validation Loss: 0.457625
Epoch: 79         Training Loss: 0.797630         Validation Loss: 0.457625
Validation loss decreased (0.614379 --> 0.457625).  Saving model ...
Epoch: 80         Train Batch: 0         Training Loss: 0.563847
Epoch: 80         Train Batch: 20         Training Loss: 0.815662
Epoch: 80         Train Batch: 40         Training Loss: 0.805178
```

```
Epoch: 80         Train Batch: 60         Training Loss: 0.784338
Epoch: 80         Train Batch: 80         Training Loss: 0.778938
Epoch: 80         Train Batch: 100         Training Loss: 0.790604
Epoch: 80         Train Batch: 120         Training Loss: 0.782789
Epoch: 80         Train Batch: 140         Training Loss: 0.779193
Epoch: 80         Train Batch: 160         Training Loss: 0.783839
Epoch: 80         Train Batch: 180         Training Loss: 0.787395
Epoch: 80         Train Batch: 200         Training Loss: 0.791360
Epoch: 80         Eval Batch: 0         Validation Loss: 0.498971
Epoch: 80         Eval Batch: 20         Validation Loss: 0.498971
Epoch: 80         Training Loss: 0.787042         Validation Loss: 0.498971
Epoch: 81         Train Batch: 0         Training Loss: 0.549030
Epoch: 81         Train Batch: 20         Training Loss: 0.793577
Epoch: 81         Train Batch: 40         Training Loss: 0.769605
Epoch: 81         Train Batch: 60         Training Loss: 0.758211
Epoch: 81         Train Batch: 80         Training Loss: 0.762089
Epoch: 81         Train Batch: 100         Training Loss: 0.779767
Epoch: 81         Train Batch: 120         Training Loss: 0.783851
Epoch: 81         Train Batch: 140         Training Loss: 0.784465
Epoch: 81         Train Batch: 160         Training Loss: 0.792582
Epoch: 81         Train Batch: 180         Training Loss: 0.788397
Epoch: 81         Train Batch: 200         Training Loss: 0.788351
Epoch: 81         Eval Batch: 0         Validation Loss: 0.919490
Epoch: 81         Eval Batch: 20         Validation Loss: 0.919490
Epoch: 81         Training Loss: 0.791367         Validation Loss: 0.919490
Epoch: 82         Train Batch: 0         Training Loss: 0.879223
Epoch: 82         Train Batch: 20         Training Loss: 0.841939
Epoch: 82         Train Batch: 40         Training Loss: 0.810697
Epoch: 82         Train Batch: 60         Training Loss: 0.801680
Epoch: 82         Train Batch: 80         Training Loss: 0.812473
Epoch: 82         Train Batch: 100         Training Loss: 0.815948
Epoch: 82         Train Batch: 120         Training Loss: 0.818802
Epoch: 82         Train Batch: 140         Training Loss: 0.811652
Epoch: 82         Train Batch: 160         Training Loss: 0.809697
Epoch: 82         Train Batch: 180         Training Loss: 0.807062
Epoch: 82         Train Batch: 200         Training Loss: 0.806475
Epoch: 82         Eval Batch: 0         Validation Loss: 0.623309
Epoch: 82         Eval Batch: 20         Validation Loss: 0.623309
Epoch: 82         Training Loss: 0.805420         Validation Loss: 0.623309
Epoch: 83         Train Batch: 0         Training Loss: 0.575752
Epoch: 83         Train Batch: 20         Training Loss: 0.759599
Epoch: 83         Train Batch: 40         Training Loss: 0.781046
Epoch: 83         Train Batch: 60         Training Loss: 0.776534
Epoch: 83         Train Batch: 80         Training Loss: 0.768943
Epoch: 83         Train Batch: 100         Training Loss: 0.759452
Epoch: 83         Train Batch: 120         Training Loss: 0.769818
Epoch: 83         Train Batch: 140         Training Loss: 0.776309
Epoch: 83         Train Batch: 160         Training Loss: 0.774936
```

```
Epoch: 83         Train Batch: 180         Training Loss: 0.781670
Epoch: 83         Train Batch: 200         Training Loss: 0.779579
Epoch: 83         Eval Batch: 0         Validation Loss: 1.177475
Epoch: 83         Eval Batch: 20         Validation Loss: 1.177475
Epoch: 83         Training Loss: 0.782826         Validation Loss: 1.177475
Epoch: 84         Train Batch: 0         Training Loss: 0.811987
Epoch: 84         Train Batch: 20         Training Loss: 0.818826
Epoch: 84         Train Batch: 40         Training Loss: 0.814440
Epoch: 84         Train Batch: 60         Training Loss: 0.833788
Epoch: 84         Train Batch: 80         Training Loss: 0.807992
Epoch: 84         Train Batch: 100         Training Loss: 0.801538
Epoch: 84         Train Batch: 120         Training Loss: 0.781414
Epoch: 84         Train Batch: 140         Training Loss: 0.780663
Epoch: 84         Train Batch: 160         Training Loss: 0.779958
Epoch: 84         Train Batch: 180         Training Loss: 0.783949
Epoch: 84         Train Batch: 200         Training Loss: 0.790309
Epoch: 84         Eval Batch: 0         Validation Loss: 0.816240
Epoch: 84         Eval Batch: 20         Validation Loss: 0.816240
Epoch: 84         Training Loss: 0.789239         Validation Loss: 0.816240
Epoch: 85         Train Batch: 0         Training Loss: 0.996330
Epoch: 85         Train Batch: 20         Training Loss: 0.746930
Epoch: 85         Train Batch: 40         Training Loss: 0.761409
Epoch: 85         Train Batch: 60         Training Loss: 0.785064
Epoch: 85         Train Batch: 80         Training Loss: 0.759089
Epoch: 85         Train Batch: 100         Training Loss: 0.759902
Epoch: 85         Train Batch: 120         Training Loss: 0.764212
Epoch: 85         Train Batch: 140         Training Loss: 0.782979
Epoch: 85         Train Batch: 160         Training Loss: 0.785403
Epoch: 85         Train Batch: 180         Training Loss: 0.778705
Epoch: 85         Train Batch: 200         Training Loss: 0.780725
Epoch: 85         Eval Batch: 0         Validation Loss: 0.545513
Epoch: 85         Eval Batch: 20         Validation Loss: 0.545513
Epoch: 85         Training Loss: 0.781246         Validation Loss: 0.545513
Epoch: 86         Train Batch: 0         Training Loss: 1.054475
Epoch: 86         Train Batch: 20         Training Loss: 0.807821
Epoch: 86         Train Batch: 40         Training Loss: 0.797839
Epoch: 86         Train Batch: 60         Training Loss: 0.780087
Epoch: 86         Train Batch: 80         Training Loss: 0.777526
Epoch: 86         Train Batch: 100         Training Loss: 0.788844
Epoch: 86         Train Batch: 120         Training Loss: 0.792117
Epoch: 86         Train Batch: 140         Training Loss: 0.787584
Epoch: 86         Train Batch: 160         Training Loss: 0.783026
Epoch: 86         Train Batch: 180         Training Loss: 0.783047
Epoch: 86         Train Batch: 200         Training Loss: 0.780884
Epoch: 86         Eval Batch: 0         Validation Loss: 0.677664
Epoch: 86         Eval Batch: 20         Validation Loss: 0.677664
Epoch: 86         Training Loss: 0.776674         Validation Loss: 0.677664
Epoch: 87         Train Batch: 0         Training Loss: 0.655287
```

```
Epoch: 87         Train Batch: 20         Training Loss: 0.783081
Epoch: 87         Train Batch: 40         Training Loss: 0.755284
Epoch: 87         Train Batch: 60         Training Loss: 0.759894
Epoch: 87         Train Batch: 80         Training Loss: 0.770419
Epoch: 87         Train Batch: 100         Training Loss: 0.775670
Epoch: 87         Train Batch: 120         Training Loss: 0.780273
Epoch: 87         Train Batch: 140         Training Loss: 0.780594
Epoch: 87         Train Batch: 160         Training Loss: 0.777262
Epoch: 87         Train Batch: 180         Training Loss: 0.778924
Epoch: 87         Train Batch: 200         Training Loss: 0.776216
Epoch: 87         Eval Batch: 0         Validation Loss: 0.930805
Epoch: 87         Eval Batch: 20         Validation Loss: 0.930805
Epoch: 87         Training Loss: 0.772826         Validation Loss: 0.930805
Epoch: 88         Train Batch: 0         Training Loss: 0.797716
Epoch: 88         Train Batch: 20         Training Loss: 0.718616
Epoch: 88         Train Batch: 40         Training Loss: 0.724054
Epoch: 88         Train Batch: 60         Training Loss: 0.746460
Epoch: 88         Train Batch: 80         Training Loss: 0.755786
Epoch: 88         Train Batch: 100         Training Loss: 0.759448
Epoch: 88         Train Batch: 120         Training Loss: 0.753807
Epoch: 88         Train Batch: 140         Training Loss: 0.756106
Epoch: 88         Train Batch: 160         Training Loss: 0.758995
Epoch: 88         Train Batch: 180         Training Loss: 0.767101
Epoch: 88         Train Batch: 200         Training Loss: 0.771366
Epoch: 88         Eval Batch: 0         Validation Loss: 0.476298
Epoch: 88         Eval Batch: 20         Validation Loss: 0.476298
Epoch: 88         Training Loss: 0.769391         Validation Loss: 0.476298
Epoch: 89         Train Batch: 0         Training Loss: 0.967131
Epoch: 89         Train Batch: 20         Training Loss: 0.821539
Epoch: 89         Train Batch: 40         Training Loss: 0.778535
Epoch: 89         Train Batch: 60         Training Loss: 0.766705
Epoch: 89         Train Batch: 80         Training Loss: 0.755344
Epoch: 89         Train Batch: 100         Training Loss: 0.743707
Epoch: 89         Train Batch: 120         Training Loss: 0.758700
Epoch: 89         Train Batch: 140         Training Loss: 0.764674
Epoch: 89         Train Batch: 160         Training Loss: 0.765268
Epoch: 89         Train Batch: 180         Training Loss: 0.772040
Epoch: 89         Train Batch: 200         Training Loss: 0.773964
Epoch: 89         Eval Batch: 0         Validation Loss: 1.528908
Epoch: 89         Eval Batch: 20         Validation Loss: 1.528908
Epoch: 89         Training Loss: 0.779457         Validation Loss: 1.528908
Epoch: 90         Train Batch: 0         Training Loss: 0.564490
Epoch: 90         Train Batch: 20         Training Loss: 0.776968
Epoch: 90         Train Batch: 40         Training Loss: 0.787268
Epoch: 90         Train Batch: 60         Training Loss: 0.777762
Epoch: 90         Train Batch: 80         Training Loss: 0.776994
Epoch: 90         Train Batch: 100         Training Loss: 0.776981
Epoch: 90         Train Batch: 120         Training Loss: 0.776193
```

```
Epoch: 90         Train Batch: 140         Training Loss: 0.769284
Epoch: 90         Train Batch: 160         Training Loss: 0.770167
Epoch: 90         Train Batch: 180         Training Loss: 0.773473
Epoch: 90         Train Batch: 200         Training Loss: 0.773211
Epoch: 90         Eval Batch: 0         Validation Loss: 0.748833
Epoch: 90         Eval Batch: 20         Validation Loss: 0.748833
Epoch: 90         Training Loss: 0.772812         Validation Loss: 0.748833
Epoch: 91         Train Batch: 0         Training Loss: 0.682435
Epoch: 91         Train Batch: 20         Training Loss: 0.760992
Epoch: 91         Train Batch: 40         Training Loss: 0.755366
Epoch: 91         Train Batch: 60         Training Loss: 0.772891
Epoch: 91         Train Batch: 80         Training Loss: 0.761813
Epoch: 91         Train Batch: 100         Training Loss: 0.759673
Epoch: 91         Train Batch: 120         Training Loss: 0.769271
Epoch: 91         Train Batch: 140         Training Loss: 0.764814
Epoch: 91         Train Batch: 160         Training Loss: 0.762667
Epoch: 91         Train Batch: 180         Training Loss: 0.768372
Epoch: 91         Train Batch: 200         Training Loss: 0.770049
Epoch: 91         Eval Batch: 0         Validation Loss: 0.751436
Epoch: 91         Eval Batch: 20         Validation Loss: 0.751436
Epoch: 91         Training Loss: 0.769764         Validation Loss: 0.751436
Epoch: 92         Train Batch: 0         Training Loss: 0.658910
Epoch: 92         Train Batch: 20         Training Loss: 0.719964
Epoch: 92         Train Batch: 40         Training Loss: 0.762487
Epoch: 92         Train Batch: 60         Training Loss: 0.766500
Epoch: 92         Train Batch: 80         Training Loss: 0.759559
Epoch: 92         Train Batch: 100         Training Loss: 0.763709
Epoch: 92         Train Batch: 120         Training Loss: 0.762856
Epoch: 92         Train Batch: 140         Training Loss: 0.762551
Epoch: 92         Train Batch: 160         Training Loss: 0.759737
Epoch: 92         Train Batch: 180         Training Loss: 0.762134
Epoch: 92         Train Batch: 200         Training Loss: 0.763438
Epoch: 92         Eval Batch: 0         Validation Loss: 0.792141
Epoch: 92         Eval Batch: 20         Validation Loss: 0.792141
Epoch: 92         Training Loss: 0.763655         Validation Loss: 0.792141
Epoch: 93         Train Batch: 0         Training Loss: 0.684077
Epoch: 93         Train Batch: 20         Training Loss: 0.700918
Epoch: 93         Train Batch: 40         Training Loss: 0.708751
Epoch: 93         Train Batch: 60         Training Loss: 0.722930
Epoch: 93         Train Batch: 80         Training Loss: 0.729514
Epoch: 93         Train Batch: 100         Training Loss: 0.744846
Epoch: 93         Train Batch: 120         Training Loss: 0.746998
Epoch: 93         Train Batch: 140         Training Loss: 0.759974
Epoch: 93         Train Batch: 160         Training Loss: 0.756561
Epoch: 93         Train Batch: 180         Training Loss: 0.760982
Epoch: 93         Train Batch: 200         Training Loss: 0.759468
Epoch: 93         Eval Batch: 0         Validation Loss: 0.806283
Epoch: 93         Eval Batch: 20         Validation Loss: 0.806283
```

```
Epoch: 93         Training Loss: 0.755984         Validation Loss: 0.806283
Epoch: 94         Train Batch: 0         Training Loss: 0.569089
Epoch: 94         Train Batch: 20         Training Loss: 0.735254
Epoch: 94         Train Batch: 40         Training Loss: 0.748764
Epoch: 94         Train Batch: 60         Training Loss: 0.753872
Epoch: 94         Train Batch: 80         Training Loss: 0.740422
Epoch: 94         Train Batch: 100         Training Loss: 0.756790
Epoch: 94         Train Batch: 120         Training Loss: 0.749319
Epoch: 94         Train Batch: 140         Training Loss: 0.750659
Epoch: 94         Train Batch: 160         Training Loss: 0.758141
Epoch: 94         Train Batch: 180         Training Loss: 0.759673
Epoch: 94         Train Batch: 200         Training Loss: 0.759678
Epoch: 94         Eval Batch: 0         Validation Loss: 0.818878
Epoch: 94         Eval Batch: 20         Validation Loss: 0.818878
Epoch: 94         Training Loss: 0.758795         Validation Loss: 0.818878
Epoch: 95         Train Batch: 0         Training Loss: 0.973263
Epoch: 95         Train Batch: 20         Training Loss: 0.805650
Epoch: 95         Train Batch: 40         Training Loss: 0.791997
Epoch: 95         Train Batch: 60         Training Loss: 0.773109
Epoch: 95         Train Batch: 80         Training Loss: 0.747983
Epoch: 95         Train Batch: 100         Training Loss: 0.753375
Epoch: 95         Train Batch: 120         Training Loss: 0.754449
Epoch: 95         Train Batch: 140         Training Loss: 0.752806
Epoch: 95         Train Batch: 160         Training Loss: 0.752139
Epoch: 95         Train Batch: 180         Training Loss: 0.745238
Epoch: 95         Train Batch: 200         Training Loss: 0.750540
Epoch: 95         Eval Batch: 0         Validation Loss: 0.842208
Epoch: 95         Eval Batch: 20         Validation Loss: 0.842208
Epoch: 95         Training Loss: 0.748068         Validation Loss: 0.842208
Epoch: 96         Train Batch: 0         Training Loss: 0.541988
Epoch: 96         Train Batch: 20         Training Loss: 0.742404
Epoch: 96         Train Batch: 40         Training Loss: 0.750505
Epoch: 96         Train Batch: 60         Training Loss: 0.758138
Epoch: 96         Train Batch: 80         Training Loss: 0.757853
Epoch: 96         Train Batch: 100         Training Loss: 0.748087
Epoch: 96         Train Batch: 120         Training Loss: 0.750000
Epoch: 96         Train Batch: 140         Training Loss: 0.754432
Epoch: 96         Train Batch: 160         Training Loss: 0.754170
Epoch: 96         Train Batch: 180         Training Loss: 0.753223
Epoch: 96         Train Batch: 200         Training Loss: 0.754427
Epoch: 96         Eval Batch: 0         Validation Loss: 0.877078
Epoch: 96         Eval Batch: 20         Validation Loss: 0.877078
Epoch: 96         Training Loss: 0.755196         Validation Loss: 0.877078
Epoch: 97         Train Batch: 0         Training Loss: 0.564201
Epoch: 97         Train Batch: 20         Training Loss: 0.803290
Epoch: 97         Train Batch: 40         Training Loss: 0.816453
Epoch: 97         Train Batch: 60         Training Loss: 0.783780
Epoch: 97         Train Batch: 80         Training Loss: 0.764306
```

```
Epoch: 97         Train Batch: 100         Training Loss: 0.764648
Epoch: 97         Train Batch: 120         Training Loss: 0.759954
Epoch: 97         Train Batch: 140         Training Loss: 0.758734
Epoch: 97         Train Batch: 160         Training Loss: 0.762739
Epoch: 97         Train Batch: 180         Training Loss: 0.768262
Epoch: 97         Train Batch: 200         Training Loss: 0.773849
Epoch: 97         Eval Batch: 0         Validation Loss: 0.809195
Epoch: 97         Eval Batch: 20         Validation Loss: 0.809195
Epoch: 97         Training Loss: 0.772742         Validation Loss: 0.809195
Epoch: 98         Train Batch: 0         Training Loss: 0.668808
Epoch: 98         Train Batch: 20         Training Loss: 0.729092
Epoch: 98         Train Batch: 40         Training Loss: 0.755027
Epoch: 98         Train Batch: 60         Training Loss: 0.754168
Epoch: 98         Train Batch: 80         Training Loss: 0.755038
Epoch: 98         Train Batch: 100         Training Loss: 0.757211
Epoch: 98         Train Batch: 120         Training Loss: 0.755910
Epoch: 98         Train Batch: 140         Training Loss: 0.760097
Epoch: 98         Train Batch: 160         Training Loss: 0.761912
Epoch: 98         Train Batch: 180         Training Loss: 0.758432
Epoch: 98         Train Batch: 200         Training Loss: 0.751926
Epoch: 98         Eval Batch: 0         Validation Loss: 0.715062
Epoch: 98         Eval Batch: 20         Validation Loss: 0.715062
Epoch: 98         Training Loss: 0.749197         Validation Loss: 0.715062
Epoch: 99         Train Batch: 0         Training Loss: 0.592850
Epoch: 99         Train Batch: 20         Training Loss: 0.755092
Epoch: 99         Train Batch: 40         Training Loss: 0.787695
Epoch: 99         Train Batch: 60         Training Loss: 0.764359
Epoch: 99         Train Batch: 80         Training Loss: 0.750583
Epoch: 99         Train Batch: 100         Training Loss: 0.741942
Epoch: 99         Train Batch: 120         Training Loss: 0.754153
Epoch: 99         Train Batch: 140         Training Loss: 0.748103
Epoch: 99         Train Batch: 160         Training Loss: 0.745415
Epoch: 99         Train Batch: 180         Training Loss: 0.751075
Epoch: 99         Train Batch: 200         Training Loss: 0.754295
Epoch: 99         Eval Batch: 0         Validation Loss: 0.767772
Epoch: 99         Eval Batch: 20         Validation Loss: 0.767772
Epoch: 99         Training Loss: 0.754378         Validation Loss: 0.767772
Epoch: 100         Train Batch: 0         Training Loss: 0.594043
Epoch: 100         Train Batch: 20         Training Loss: 0.691135
Epoch: 100         Train Batch: 40         Training Loss: 0.706350
Epoch: 100         Train Batch: 60         Training Loss: 0.714674
Epoch: 100         Train Batch: 80         Training Loss: 0.733336
Epoch: 100         Train Batch: 100         Training Loss: 0.744217
Epoch: 100         Train Batch: 120         Training Loss: 0.749552
Epoch: 100         Train Batch: 140         Training Loss: 0.750403
Epoch: 100         Train Batch: 160         Training Loss: 0.757323
Epoch: 100         Train Batch: 180         Training Loss: 0.755218
Epoch: 100         Train Batch: 200         Training Loss: 0.758565
```

```
Epoch: 100         Eval Batch: 0         Validation Loss: 0.788743
Epoch: 100         Eval Batch: 20         Validation Loss: 0.788743
Epoch: 100         Training Loss: 0.758788         Validation Loss: 0.788743
```

ResNet16 had the min valid loss @ 79th Epoch * Test Loss: 0.682454 * Test Accuracy: 81% (678/836)

1.1.17 (IMPLEMENTATION) Test the Model

Try out your model on the test dataset of dog images. Use the code cell below to calculate and print the test loss and accuracy. Ensure that your test accuracy is greater than 60%.

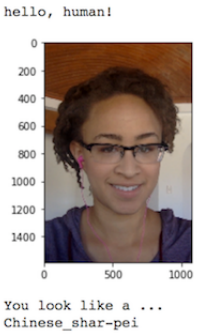
```
In [11]: test(loaders_transfer, model_transfer, criterion_transfer, use_cuda)
```

```
Test Loss: 0.682454


Test Accuracy: 81% (678/836)
```

1.1.18 (IMPLEMENTATION) Predict Dog Breed with the Model

Write a function that takes an image path as input and returns the dog breed (`Affenpinscher`, `Afghan hound`, etc) that is predicted by your model.

```
In [91]: ### TODO: Write a function that takes a path to an image as input
         ### and returns the dog breed that is predicted by the model.

         # list of class names by index, i.e. a name can be accessed like class_names[0]
         class_names = [item[4:].replace("_", " ") for item in data_transfer['train'].classes]

         def load_image(img_path):
             # load the image
             in_transform = transforms.Compose([
                                transforms.Resize((224, 224)),
                                transforms.ToTensor(),
                                transforms.Normalize((0.485, 0.456, 0.406),
                                                     (0.229, 0.224, 0.225))])
             image = Image.open(img_path)
             image = in_transform(image).float()
             image = image.unsqueeze(0)  #this is for VGG, may not be needed for ResNet
             if use_cuda:
                 image = image.cuda()
             return image

         def predict_breed_transfer(img_path, model):
             # load the image and return the predicted breed
             image = load_image(img_path)
```



Sample Human Output

```
    output = model(image)
    _, indices = F.softmax(output, dim=1).max(dim=1)

    if use_cuda:
        indices = indices.cpu()
    predicted_breed_name = class_names[indices.numpy()[0]]
    return predicted_breed_name
```

```
In [103]: model_transfer.load_state_dict(torch.load('model_transfer_resnet.pt'))

          predict_breed_transfer("images/Labrador_retriever_06449.jpg", model_transfer)
```

```
Out[103]: 'Labrador retriever'
```

Step 5: Write your Algorithm

Write an algorithm that accepts a file path to an image and first determines whether the image contains a human, dog, or neither. Then, - if a **dog** is detected in the image, return the predicted breed. - if a **human** is detected in the image, return the resembling dog breed. - if **neither** is detected in the image, provide output that indicates an error.

You are welcome to write your own functions for detecting humans and dogs in images, but feel free to use the `face_detector` and `human_detector` functions developed above. You are **required** to use your CNN from Step 4 to predict dog breed.

Some sample output for our algorithm is provided below, but feel free to design your own user experience!

1.1.19 (IMPLEMENTATION) Write your Algorithm

```
In [93]: ### TODO: Write your algorithm.
         ### Feel free to use as many code cells as needed.

         def run_app(img_path, model):
             ## handle cases for a human face, dog, and neither
             if dog_detector(img_path) or face_detector(img_path):
                 dog_breed = predict_breed_transfer(img_path, model)
```

```
            print("You look like a ...")
            print(dog_breed)
        else:
            print("ERROR: The input image is neither human face nor a dog.")
```

```
In [102]: run_app("images/Curly-coated_retriever_03896.jpg", model_transfer)

You look like a ...
Curly-coated retriever
```

Step 6: Test Your Algorithm

In this section, you will take your new algorithm for a spin! What kind of dog does the algorithm think that *you* look like? If you have a dog, does it predict your dog's breed accurately? If you have a cat, does it mistakenly think that your cat is a dog?

1.1.20 (IMPLEMENTATION) Test Your Algorithm on Sample Images!

Test your algorithm at least six images on your computer. Feel free to use any images you like. Use at least two human and two dog images.

Question 6: Is the output better than you expected :) ? Or worse :(? Provide at least three possible points of improvement for your algorithm.

The output meets my expectation. I think it is having difficulty classifying between American water spaniel and Boykin spaniel. I think if I train a separate binary classifier distinguishing just between the two breeds, I might be able to tune all the hyperparameters to classify the dog between the two breeds much better. Or, I could also gather more training data set for those breeds to retrain my model on the updated dataset.

```
In [112]: run_app("lfw/Aaron_Peirsol/Aaron_Peirsol_0001.jpg", model_transfer)

You look like a ...
Dachshund
```

```
In [115]: run_app("lfw/Abdel_Nasser_Assidi/Abdel_Nasser_Assidi_0002.jpg", model_transfer)

You look like a ...
American water spaniel
```

```
In [101]: run_app("images/American_water_spaniel_00648.jpg", model_transfer)

You look like a ...
Boykin spaniel
```

```
In [98]: run_app("images/Brittany_02625.jpg", model_transfer)
```

```
You look like a ...
Brittany
```

```
In [99]: run_app("images/cat1.jpg", model_transfer)
```

```
ERROR: The input image is neither human face nor a dog.
```

```
In [100]: run_app("images/cat2.jpg", model_transfer)
```

```
ERROR: The input image is neither human face nor a dog.
```

Answer: (Three possible points for improvement)

```
In [114]: ## TODO: Execute your algorithm from Step 6 on
          ## at least 6 images on your computer.
          ## Feel free to use as many code cells as needed.

          ## suggested code, below
          for file in np.hstack((human_files[:3], dog_files[:3])):
              run_app(file, model_transfer)
```

```
You look like a ...
Dogue de bordeaux
You look like a ...
Chinese crested
You look like a ...
Dogue de bordeaux
You look like a ...
Leonberger
You look like a ...
Leonberger
You look like a ...
Leonberger
```

```
In [ ]:
```