coursework_02

February 13, 2025

1 Coursework 2: Image segmentation

In this coursework you will develop and train a convolutional neural network for brain tumour image segmentation. Please read both the text and the code in this notebook to get an idea what you are expected to implement. Pay attention to the missing code blocks that look like this:

```
### Insert your code ###
...
### End of your code ###
```

1.1 What to do?

- Complete and run the code using `jupyter-lab` or `jupyter-notebook` to get the results.

- Export (File | Save and Export Notebook As…) the notebook as a PDF file, which contains your code, results and answers, and upload the PDF file onto Scientia.

- Instead of clicking the Export button, you can also run the following command instead: `jupyter nbconvert coursework.ipynb --to pdf`

- If Jupyter complains about some problems in exporting, it is likely that pandoc (https://pandoc.org/installing.html) or latex is not installed, or their paths have not been included. You can install the relevant libraries and retry.

- If Jupyter-lab does not work for you at the end, you can use Google Colab to write the code and export the PDF file.

1.2 Dependencies

You need to install Jupyter-Lab (https://jupyterlab.readthedocs.io/en/stable/getting_started/installation.html) and other libraries used in this coursework, such as by running the command: `pip3 install [package_name]`

1.3 GPU resource

The coursework is developed to be able to run on CPU, as all images have been pre-processed to be 2D and of a smaller size, compared to original 3D volumes.

However, to save training time, you may want to use GPU. In that case, you can run this notebook on Google Colab. On Google Colab, go to the menu, Runtime - Change runtime type, and select **GPU** as the hardware acceleartor. At the end, please still export everything and submit as a PDF file on Scientia.

```
[2]:  # Import libraries
      # These libraries should be sufficient for this tutorial.
      # However, if any other library is needed, please install by yourself.
      import tarfile
      import imageio
      import torch
      import torch.nn as nn
      import torch.nn.functional as F
      import torch.optim as optim
      from torch.utils.data import Dataset
      import numpy as np
      import time
      import os
      import random
      import matplotlib.pyplot as plt
      from matplotlib import colors
```

1.4 1. Download and visualise the imaging dataset.

The dataset is curated from the brain imaging dataset in Medical Decathlon Challenge. To save
the storage and reduce the computational cost for this tutorial, we extract 2D image slices from
T1-Gd contrast enhanced 3D brain volumes and downsample the images.

The dataset consists of a training set and a test set. Each image is of dimension 120 x 120, with
a corresponding label map of the same dimension. There are four number of classes in the label
map:

- 0: background
- 1: edema
- 2: non-enhancing tumour
- 3: enhancing tumour

```
[3]:  # Download the dataset
      !wget https://www.dropbox.com/s/zmytk2yu284af6t/Task01_BrainTumour_2D.tar.gz

      # Unzip the '.tar.gz' file to the current directory
      datafile = tarfile.open('Task01_BrainTumour_2D.tar.gz')
      datafile.extractall()
      datafile.close()
```

--2025-02-13 17:18:33--
https://www.dropbox.com/s/zmytk2yu284af6t/Task01_BrainTumour_2D.tar.gz
Resolving www.dropbox.com (www.dropbox.com)… 162.125.64.18,
2620:100:6020:18::a27d:4012
Connecting to www.dropbox.com (www.dropbox.com)|162.125.64.18|:443… connected.
HTTP request sent, awaiting response… 302 Found
Location: https://www.dropbox.com/scl/fi/4bf8fqcfgf3lebiv2in99/Task01_BrainTumou
r_2D.tar.gz?rlkey=ceq898g2tr3aaxjxn4xjxbob1 [following]
--2025-02-13 17:18:33-- https://www.dropbox.com/scl/fi/4bf8fqcfgf3lebiv2in99/Ta

sk01_BrainTumour_2D.tar.gz?rlkey=ceq898g2tr3aaxjxn4xjxbob1
Reusing existing connection to www.dropbox.com:443.
HTTP request sent, awaiting response… 302 Found
Location: https://ucb01e9f6e3d86c04d5be0907ad0.dl.dropboxusercontent.com/cd/0/in
line/CkD6mCx_qSLOCWECuD0ycmrmlsRbIT4ACuCcURxdh85YXMPjsYOFqehRipWTBidNHSGaQ_12b2H
LLtRaDZicokjfxmq0xdwfT-pic4xmiv1ZX8KUL2L_MTH2BPD6ZmzEq40/file# [following]
--2025-02-13 17:18:34-- https://ucb01e9f6e3d86c04d5be0907ad0.dl.dropboxusercont
ent.com/cd/0/inline/CkD6mCx_qSLOCWECuD0ycmrmlsRbIT4ACuCcURxdh85YXMPjsYOFqehRipWT
BidNHSGaQ_12b2HLLtRaDZicokjfxmq0xdwfT-pic4xmiv1ZX8KUL2L_MTH2BPD6ZmzEq40/file
Resolving ucb01e9f6e3d86c04d5be0907ad0.dl.dropboxusercontent.com
(ucb01e9f6e3d86c04d5be0907ad0.dl.dropboxusercontent.com)… 162.125.64.15,
2620:100:6020:15::a27d:400f
Connecting to ucb01e9f6e3d86c04d5be0907ad0.dl.dropboxusercontent.com
(ucb01e9f6e3d86c04d5be0907ad0.dl.dropboxusercontent.com)|162.125.64.15|:443…
connected.
HTTP request sent, awaiting response… 302 Found
Location: /cd/0/inline2/CkD5HO7P7pziMosRVmccncymrwzTktapF9zuPAyyubAcDDDItFwl8dlf
duvsztxgxqz7lIVmhmsDoAceZb61fc5H724knyA4pOoWibUU3JhNXJRMDhiUbCX1YComPCxjeSECkwmT
Yi3y-pokYiLJh5F7x_I-zEkhWCwl44LUGhzdKEE2vii1L2Lv4OWYVwAcDkzrbHOBNT7LjX5ut_XwwIOI
GJgpcwT3_h-B3to0CNszYCCpNn90CMTU1HBmvaaiC5iD2yamfWtxjtQ7W-
G6zlwUnpb6y160DJqWLrKp3zG28LBQwTIbqdZBI_nJ9q-
FpGQws3rkcQKAFCNAUHMWMiCwnO0d6uy2vri0yUW3Xkwe4Q/file [following]
--2025-02-13 17:18:34-- https://ucb01e9f6e3d86c04d5be0907ad0.dl.dropboxusercont
ent.com/cd/0/inline2/CkD5HO7P7pziMosRVmccncymrwzTktapF9zuPAyyubAcDDDItFwl8dlfduv
sztxgxqz7lIVmhmsDoAceZb61fc5H724knyA4pOoWibUU3JhNXJRMDhiUbCX1YComPCxjeSECkwmTYi3
y-pokYiLJh5F7x_I-zEkhWCwl44LUGhzdKEE2vii1L2Lv4OWYVwAcDkzrbHOBNT7LjX5ut_XwwIOIGJg
pcwT3_h-B3to0CNszYCCpNn90CMTU1HBmvaaiC5iD2yamfWtxjtQ7W-
G6zlwUnpb6y160DJqWLrKp3zG28LBQwTIbqdZBI_nJ9q-
FpGQws3rkcQKAFCNAUHMWMiCwnO0d6uy2vri0yUW3Xkwe4Q/file
Reusing existing connection to
ucb01e9f6e3d86c04d5be0907ad0.dl.dropboxusercontent.com:443.
HTTP request sent, awaiting response… 200 OK
Length: 9251149 (8.8M) [application/octet-stream]
Saving to: 'Task01_BrainTumour_2D.tar.gz'

Task01_BrainTumour_ 100%[===================>]   8.82M 11.7MB/s   in 0.8s

2025-02-13 17:18:36 (11.7 MB/s) - 'Task01_BrainTumour_2D.tar.gz' saved
[9251149/9251149]

1.5 Visualise a random set of 4 training images along with their label maps.

Suggested colour map for brain MR image:

cmap = 'gray'

Suggested colour map for segmentation map:

```
    cmap = colors.ListedColormap(['black', 'green', 'blue', 'red'])
```

```
[4]:  ### Insert your code ###

      # Visualise a random set of 4 training images along with their label maps
      fig, axes = plt.subplots(4, 2, figsize=(10, 20))

      # Suggested colour maps
      image_cmap = 'gray'
      label_cmap = colors.ListedColormap(['black', 'green', 'blue', 'red'])

      # Load the images and labels
      image_path = 'Task01_BrainTumour_2D/training_images'
      label_path = 'Task01_BrainTumour_2D/training_labels'
      image_names = sorted(os.listdir(image_path))
      label_names = sorted(os.listdir(label_path))

      # Randomly select 4 images
      random_indices = random.sample(range(len(image_names)), 4)

      for i, idx in enumerate(random_indices):
          image = imageio.imread(os.path.join(image_path, image_names[idx]))
          label = imageio.imread(os.path.join(label_path, label_names[idx]))

          axes[i, 0].imshow(image, cmap=image_cmap)
          axes[i, 0].set_title(f'Image {idx}')
          axes[i, 0].axis('off')

          axes[i, 1].imshow(label, cmap=label_cmap)
          axes[i, 1].set_title(f'Label {idx}')
          axes[i, 1].axis('off')

      plt.tight_layout()
      plt.show()


      ### End of your code ###
```

/tmp/ipykernel_8866/3903710960.py:20: DeprecationWarning: Starting with ImageIO
v3 the behavior of this function will switch to that of iio.v3.imread. To keep
the current behavior (and make this warning disappear) use `import imageio.v2 as
imageio` or call `imageio.v2.imread` directly.
  image = imageio.imread(os.path.join(image_path, image_names[idx]))
/tmp/ipykernel_8866/3903710960.py:21: DeprecationWarning: Starting with ImageIO
v3 the behavior of this function will switch to that of iio.v3.imread. To keep
the current behavior (and make this warning disappear) use `import imageio.v2 as
imageio` or call `imageio.v2.imread` directly.
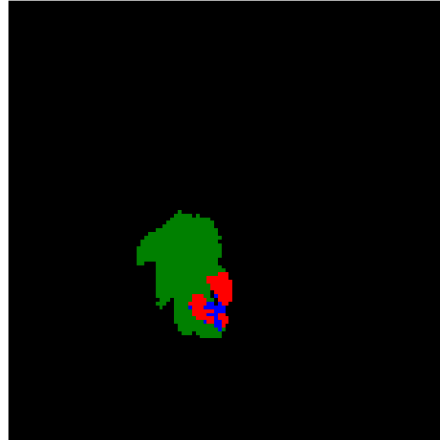 label = imageio.imread(os.path.join(label_path, label_names[idx]))
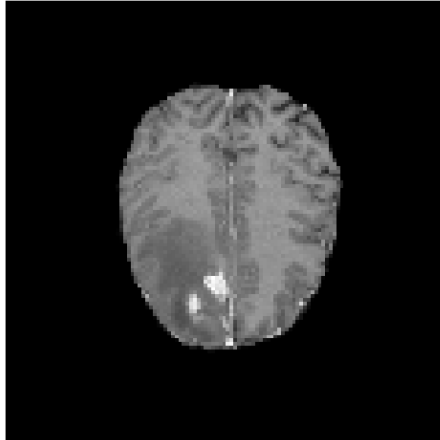
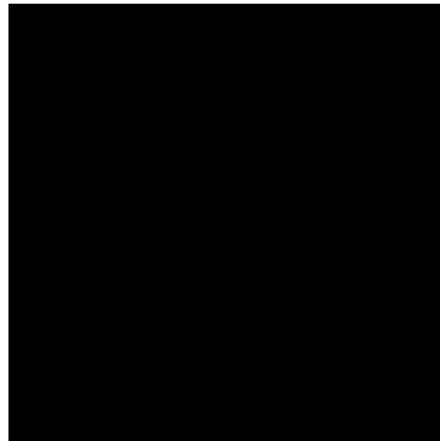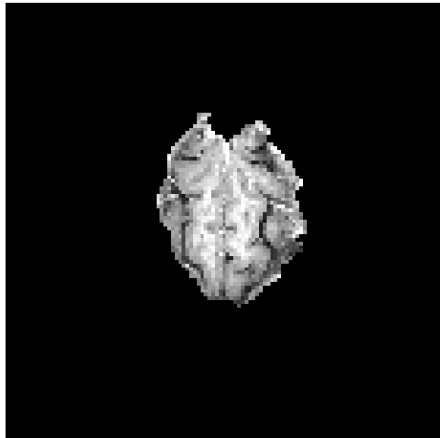Image 66 · Label 66 · Image 1430 · Label 1430 · Image 1643 · Label 1643 · Image 1837 · Label 1837

1.6   2. Implement a dataset class.

It can read the imaging dataset and get items, pairs of images and label maps, as training batches.

```
[5]:  def normalise_intensity(image, thres_roi=1.0):
          """ Normalise the image intensity by the mean and standard deviation """
          # ROI defines the image foreground
          val_l = np.percentile(image, thres_roi)
          roi = (image >= val_l)
          mu, sigma = np.mean(image[roi]), np.std(image[roi])
          eps = 1e-6
          image2 = (image - mu) / (sigma + eps)
          return image2


      class BrainImageSet(Dataset):
          """ Brain image set """
          def __init__(self, image_path, label_path='', deploy=False):
              self.image_path = image_path
              self.deploy = deploy
              self.images = []
              self.labels = []

              image_names = sorted(os.listdir(image_path))
              for image_name in image_names:
                  # Read the image
                  image = imageio.imread(os.path.join(image_path, image_name))
                  self.images += [image]

                  # Read the label map
                  if not self.deploy:
                      label_name = os.path.join(label_path, image_name)
                      label = imageio.imread(label_name)
                      self.labels += [label]

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

          def __getitem__(self, idx):
              # Get an image and perform intensity normalisation
              # Dimension: XY
              image = normalise_intensity(self.images[idx])

              # Get its label map
              # Dimension: XY
```

```
            label = self.labels[idx]
            return image, label

    def get_random_batch(self, batch_size):
        # Get a batch of paired images and label maps
        # Dimension of images: NCXY
        # Dimension of labels: NXY
        images, labels = [], []

        ### Insert your code ###
        indices = random.sample(range(len(self.images)), batch_size)
        for idx in indices:
            image, label = self.__getitem__(idx)
            images.append(image)
            labels.append(label)

        images = np.expand_dims(np.array(images), axis=1)  # Add channel␣
↪dimension
        labels = np.array(labels)
        ### End of your code ###
        return images, labels
```

1.7 3. Build a U-net architecture.

You will implement a U-net architecture. If you are not familiar with U-net, please read this paper:

[1] Olaf Ronneberger et al. U-Net: Convolutional networks for biomedical image segmentation. MICCAI, 2015.

For the first convolutional layer, you can start with 16 filters. We have implemented the encoder path. Please complete the decoder path.

```
[6]: class UNet(nn.Module):
    def __init__(self, input_channel=1, output_channel=1, num_filter=16):
        super(UNet, self).__init__()

        # BatchNorm: by default during training this layer keeps running␣
↪estimates
        # of its computed mean and variance, which are then used for␣
↪normalization
        # during evaluation.

        # Encoder path
        n = num_filter  # 16
        self.conv1 = nn.Sequential(
            nn.Conv2d(input_channel, n, kernel_size=3, padding=1),
            nn.BatchNorm2d(n),
            nn.ReLU(),
```

```python
            nn.Conv2d(n, n, kernel_size=3, padding=1),
            nn.BatchNorm2d(n),
            nn.ReLU()
        )

        self.conv2 = nn.Sequential(
            nn.Conv2d(n, 2*n, kernel_size=3, stride=2, padding=1),
            nn.BatchNorm2d(2*n),
            nn.ReLU(),
            nn.Conv2d(2*n, 2*n, kernel_size=3, padding=1),
            nn.BatchNorm2d(2*n),
            nn.ReLU()
        )

        self.conv3 = nn.Sequential(
            nn.Conv2d(2*n, 4*n, kernel_size=3, stride=2, padding=1),
            nn.BatchNorm2d(4*n),
            nn.ReLU(),
            nn.Conv2d(4*n, 4*n, kernel_size=3, padding=1),
            nn.BatchNorm2d(4*n),
            nn.ReLU()
        )

        self.conv4 = nn.Sequential(
            nn.Conv2d(4*n, 8*n, kernel_size=3, stride=2, padding=1),
            nn.BatchNorm2d(8*n),
            nn.ReLU(),
            nn.Conv2d(8*n, 8*n, kernel_size=3, padding=1),
            nn.BatchNorm2d(8*n),
            nn.ReLU()
        )

        # Decoder path
        self.upconv4 = nn.ConvTranspose2d(8*n, 4*n, kernel_size=2, stride=2)
        self.conv_up4 = nn.Sequential(
            nn.Conv2d(8*n, 4*n, kernel_size=3, padding=1),  # 8*n because of
↪concatenation
            nn.BatchNorm2d(4*n),
            nn.ReLU(),
            nn.Conv2d(4*n, 4*n, kernel_size=3, padding=1),
            nn.BatchNorm2d(4*n),
            nn.ReLU()
        )

        self.upconv3 = nn.ConvTranspose2d(4*n, 2*n, kernel_size=2, stride=2)
        self.conv_up3 = nn.Sequential(
```

```python
            nn.Conv2d(4*n, 2*n, kernel_size=3, padding=1),  # 4*n because of
↪concatenation
            nn.BatchNorm2d(2*n),
            nn.ReLU(),
            nn.Conv2d(2*n, 2*n, kernel_size=3, padding=1),
            nn.BatchNorm2d(2*n),
            nn.ReLU()
        )

        self.upconv2 = nn.ConvTranspose2d(2*n, n, kernel_size=2, stride=2)
        self.conv_up2 = nn.Sequential(
            nn.Conv2d(2*n, n, kernel_size=3, padding=1),  # 2*n because of
↪concatenation
            nn.BatchNorm2d(n),
            nn.ReLU(),
            nn.Conv2d(n, n, kernel_size=3, padding=1),
            nn.BatchNorm2d(n),
            nn.ReLU()
        )

        self.final_conv = nn.Conv2d(n, output_channel, kernel_size=1)

    def forward(self, x):
        # Encoder
        x1 = self.conv1(x)
        x2 = self.conv2(x1)
        x3 = self.conv3(x2)
        x4 = self.conv4(x3)

        # Decoder
        x = self.upconv4(x4)
        x = torch.cat((x, x3), dim=1)
        x = self.conv_up4(x)

        x = self.upconv3(x)
        x = torch.cat((x, x2), dim=1)
        x = self.conv_up3(x)

        x = self.upconv2(x)
        x = torch.cat((x, x1), dim=1)
        x = self.conv_up2(x)

        x = self.final_conv(x)

        return x
```

1.8 4. Train the segmentation model.

```
[7]: # CUDA device
     device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')
     print('Device: {0}'.format(device))

     # Build the model
     num_class = 4
     model = UNet(input_channel=1, output_channel=num_class, num_filter=16)
     model = model.to(device)
     params = list(model.parameters())

     model_dir = 'saved_models'
     if not os.path.exists(model_dir):
         os.makedirs(model_dir)

     # Optimizer
     optimizer = optim.Adam(params, lr=1e-3)

     # Segmentation loss
     criterion = nn.CrossEntropyLoss()

     # Datasets
     train_set = BrainImageSet('Task01_BrainTumour_2D/training_images',␣
      ↪'Task01_BrainTumour_2D/training_labels')
     test_set = BrainImageSet('Task01_BrainTumour_2D/test_images',␣
      ↪'Task01_BrainTumour_2D/test_labels')

     # Train the model
     # Note: when you debug the model, you may reduce the number of iterations or␣
      ↪batch size to save time.
     num_iter = 10000
     train_batch_size = 16
     eval_batch_size = 16
     start = time.time()
     for it in range(1, 1 + num_iter):
         # Set the modules in training mode, which will have effects on certain␣
      ↪modules, e.g. dropout or batchnorm.
         start_iter = time.time()
         model.train()

         # Get a batch of images and labels
         images, labels = train_set.get_random_batch(train_batch_size)
         images, labels = torch.from_numpy(images), torch.from_numpy(labels)
         images, labels = images.to(device, dtype=torch.float32), labels.to(device,␣
      ↪dtype=torch.long)
         logits = model(images)
```

```python
        # Perform optimisation and print out the training loss
        ### Insert your code ###
        optimizer.zero_grad()
        loss = criterion(logits, labels)
        loss.backward()
        optimizer.step()
        print(f'Iteration {it}, Loss: {loss.item():.4f}, Time: {time.time() -␣
    ↪start_iter:.2f}s')
        ### End of your code ###

        # Evaluate
        if it % 100 == 0:
            print(f'Iteration {it}, Loss: {loss.item():.4f}, Time: {time.time() -␣
    ↪start_iter:.2f}s')
            # Disabling gradient calculation during reference to reduce memory␣
    ↪consumption
            with torch.no_grad():
                # Evaluate on a batch of test images and print out the test loss
                ### Insert your code ###
                test_images, test_labels = test_set.
    ↪get_random_batch(eval_batch_size)
                test_images, test_labels = torch.from_numpy(test_images), torch.
    ↪from_numpy(test_labels)
                test_images, test_labels = test_images.to(device, dtype=torch.
    ↪float32), test_labels.to(device, dtype=torch.long)
                test_logits = model(test_images)
                test_loss = criterion(test_logits, test_labels)
                print(f'Test Loss: {test_loss.item():.4f}')
                ### End of your code ###

        # Save the model
        if it % 5000 == 0:
            torch.save(model.state_dict(), os.path.join(model_dir, 'model_{0}.pt'.
    ↪format(it)))
print('Training took {:.3f}s in total.'.format(time.time() - start))
```

Device: cuda

/tmp/ipykernel_8866/3272815305.py:23: DeprecationWarning: Starting with ImageIO
v3 the behavior of this function will switch to that of iio.v3.imread. To keep
the current behavior (and make this warning disappear) use `import imageio.v2 as
imageio` or call `imageio.v2.imread` directly.
  image = imageio.imread(os.path.join(image_path, image_name))
/tmp/ipykernel_8866/3272815305.py:29: DeprecationWarning: Starting with ImageIO
v3 the behavior of this function will switch to that of iio.v3.imread. To keep
the current behavior (and make this warning disappear) use `import imageio.v2 as
imageio` or call `imageio.v2.imread` directly.

```
    label = imageio.imread(label_name)
```

Iteration 1, Loss: 1.4807, Time: 1.56s
Iteration 2, Loss: 1.4378, Time: 0.06s
Iteration 3, Loss: 1.3880, Time: 0.04s
Iteration 4, Loss: 1.3599, Time: 0.05s
Iteration 5, Loss: 1.3373, Time: 0.04s
Iteration 6, Loss: 1.3158, Time: 0.05s
Iteration 7, Loss: 1.2933, Time: 0.05s
Iteration 8, Loss: 1.2833, Time: 0.05s
Iteration 9, Loss: 1.2549, Time: 0.04s
Iteration 10, Loss: 1.2401, Time: 0.05s
Iteration 11, Loss: 1.2156, Time: 0.05s
Iteration 12, Loss: 1.1972, Time: 0.04s
Iteration 13, Loss: 1.1840, Time: 0.05s
Iteration 14, Loss: 1.1786, Time: 0.04s
Iteration 15, Loss: 1.1507, Time: 0.05s
Iteration 16, Loss: 1.1313, Time: 0.05s
Iteration 17, Loss: 1.1204, Time: 0.04s
Iteration 18, Loss: 1.0957, Time: 0.05s
Iteration 19, Loss: 1.0923, Time: 0.05s
Iteration 20, Loss: 1.0720, Time: 0.05s
Iteration 21, Loss: 1.0675, Time: 0.06s
Iteration 22, Loss: 1.0549, Time: 0.05s
Iteration 23, Loss: 1.0472, Time: 0.06s
Iteration 24, Loss: 1.0275, Time: 0.04s
Iteration 25, Loss: 1.0199, Time: 0.04s
Iteration 26, Loss: 1.0081, Time: 0.09s
Iteration 27, Loss: 1.0011, Time: 0.07s
Iteration 28, Loss: 0.9832, Time: 0.06s
Iteration 29, Loss: 0.9790, Time: 0.05s
Iteration 30, Loss: 0.9711, Time: 0.05s
Iteration 31, Loss: 0.9508, Time: 0.05s
Iteration 32, Loss: 0.9508, Time: 0.05s
Iteration 33, Loss: 0.9383, Time: 0.05s
Iteration 34, Loss: 0.9339, Time: 0.05s
Iteration 35, Loss: 0.9394, Time: 0.04s
Iteration 36, Loss: 0.9141, Time: 0.05s
Iteration 37, Loss: 0.9106, Time: 0.06s
Iteration 38, Loss: 0.9008, Time: 0.05s
Iteration 39, Loss: 0.8930, Time: 0.05s
Iteration 40, Loss: 0.8894, Time: 0.05s
Iteration 41, Loss: 0.8762, Time: 0.05s
Iteration 42, Loss: 0.8660, Time: 0.05s
Iteration 43, Loss: 0.8679, Time: -0.09s
Iteration 44, Loss: 0.8599, Time: 0.05s
Iteration 45, Loss: 0.8646, Time: 0.04s
Iteration 46, Loss: 0.8396, Time: 0.05s

```
Iteration 47, Loss: 0.8409, Time: 0.04s
Iteration 48, Loss: 0.8300, Time: 0.04s
Iteration 49, Loss: 0.8137, Time: 0.04s
Iteration 50, Loss: 0.8096, Time: 0.05s
Iteration 51, Loss: 0.8088, Time: 0.04s
Iteration 52, Loss: 0.8042, Time: 0.04s
Iteration 53, Loss: 0.7925, Time: 0.05s
Iteration 54, Loss: 0.7778, Time: 0.05s
Iteration 55, Loss: 0.7814, Time: 0.05s
Iteration 56, Loss: 0.7750, Time: 0.04s
Iteration 57, Loss: 0.7578, Time: 0.04s
Iteration 58, Loss: 0.7516, Time: 0.05s
Iteration 59, Loss: 0.7667, Time: 0.04s
Iteration 60, Loss: 0.7468, Time: 0.04s
Iteration 61, Loss: 0.7424, Time: 0.04s
Iteration 62, Loss: 0.7399, Time: 0.04s
Iteration 63, Loss: 0.7197, Time: 0.04s
Iteration 64, Loss: 0.7131, Time: 0.04s
Iteration 65, Loss: 0.7007, Time: 0.04s
Iteration 66, Loss: 0.7004, Time: 0.04s
Iteration 67, Loss: 0.6970, Time: 0.04s
Iteration 68, Loss: 0.6854, Time: 0.04s
Iteration 69, Loss: 0.6773, Time: 0.05s
Iteration 70, Loss: 0.6723, Time: 0.05s
Iteration 71, Loss: 0.6708, Time: 0.04s
Iteration 72, Loss: 0.6628, Time: 0.04s
Iteration 73, Loss: 0.6579, Time: 0.04s
Iteration 74, Loss: 0.6439, Time: 0.04s
Iteration 75, Loss: 0.6283, Time: 0.05s
Iteration 76, Loss: 0.6318, Time: 0.04s
Iteration 77, Loss: 0.6174, Time: 0.04s
Iteration 78, Loss: 0.6246, Time: 0.04s
Iteration 79, Loss: 0.6132, Time: 0.04s
Iteration 80, Loss: 0.5975, Time: 0.06s
Iteration 81, Loss: 0.6057, Time: 0.04s
Iteration 82, Loss: 0.5887, Time: 0.04s
Iteration 83, Loss: 0.5777, Time: 0.04s
Iteration 84, Loss: 0.5718, Time: 0.04s
Iteration 85, Loss: 0.5693, Time: 0.04s
Iteration 86, Loss: 0.5657, Time: 0.04s
Iteration 87, Loss: 0.5638, Time: 0.05s
Iteration 88, Loss: 0.5554, Time: 0.04s
Iteration 89, Loss: 0.5494, Time: 0.04s
Iteration 90, Loss: 0.5425, Time: 0.04s
Iteration 91, Loss: 0.5310, Time: 0.05s
Iteration 92, Loss: 0.5292, Time: 0.04s
Iteration 93, Loss: 0.5200, Time: 0.04s
Iteration 94, Loss: 0.5170, Time: 0.04s
```

```
Iteration 95, Loss: 0.5202, Time: 0.04s
Iteration 96, Loss: 0.5075, Time: 0.06s
Iteration 97, Loss: 0.4918, Time: 0.06s
Iteration 98, Loss: 0.4868, Time: 0.06s
Iteration 99, Loss: 0.4815, Time: 0.04s
Iteration 100, Loss: 0.4912, Time: 0.04s
Iteration 100, Loss: 0.4912, Time: 0.04s
Test Loss: 0.4791
Iteration 101, Loss: 0.4742, Time: 0.04s
Iteration 102, Loss: 0.4637, Time: 0.04s
Iteration 103, Loss: 0.4588, Time: 0.04s
Iteration 104, Loss: 0.4568, Time: 0.05s
Iteration 105, Loss: 0.4464, Time: 0.04s
Iteration 106, Loss: 0.4613, Time: 0.05s
Iteration 107, Loss: 0.4422, Time: 0.05s
Iteration 108, Loss: 0.4388, Time: 0.05s
Iteration 109, Loss: 0.4185, Time: 0.05s
Iteration 110, Loss: 0.4327, Time: 0.04s
Iteration 111, Loss: 0.4300, Time: 0.05s
Iteration 112, Loss: 0.4139, Time: 0.04s
Iteration 113, Loss: 0.4110, Time: 0.04s
Iteration 114, Loss: 0.4082, Time: 0.05s
Iteration 115, Loss: 0.3991, Time: 0.04s
Iteration 116, Loss: 0.4039, Time: 0.05s
Iteration 117, Loss: 0.3880, Time: 0.05s
Iteration 118, Loss: 0.3959, Time: 0.05s
Iteration 119, Loss: 0.3770, Time: 0.04s
Iteration 120, Loss: 0.3737, Time: 0.05s
Iteration 121, Loss: 0.3749, Time: 0.04s
Iteration 122, Loss: 0.3795, Time: 0.04s
Iteration 123, Loss: 0.3658, Time: 0.04s
Iteration 124, Loss: 0.3644, Time: 0.05s
Iteration 125, Loss: 0.3599, Time: 0.04s
Iteration 126, Loss: 0.3527, Time: 0.04s
Iteration 127, Loss: 0.3535, Time: 0.05s
Iteration 128, Loss: 0.3346, Time: 0.04s
Iteration 129, Loss: 0.3411, Time: 0.04s
Iteration 130, Loss: 0.3404, Time: 0.05s
Iteration 131, Loss: 0.3188, Time: 0.04s
Iteration 132, Loss: 0.3475, Time: 0.04s
Iteration 133, Loss: 0.3207, Time: 0.05s
Iteration 134, Loss: 0.3193, Time: 0.04s
Iteration 135, Loss: 0.3438, Time: 0.04s
Iteration 136, Loss: 0.3318, Time: 0.04s
Iteration 137, Loss: 0.2984, Time: 0.05s
Iteration 138, Loss: 0.3182, Time: 0.05s
Iteration 139, Loss: 0.3111, Time: 0.04s
Iteration 140, Loss: 0.2931, Time: 0.05s
```

```
Iteration 141, Loss: 0.3118, Time: 0.04s
Iteration 142, Loss: 0.3168, Time: 0.05s
Iteration 143, Loss: 0.2912, Time: 0.04s
Iteration 144, Loss: 0.2984, Time: 0.05s
Iteration 145, Loss: 0.3030, Time: 0.04s
Iteration 146, Loss: 0.2871, Time: 0.05s
Iteration 147, Loss: 0.2630, Time: 0.05s
Iteration 148, Loss: 0.2966, Time: 0.05s
Iteration 149, Loss: 0.2794, Time: 0.05s
Iteration 150, Loss: 0.2684, Time: 0.05s
Iteration 151, Loss: 0.2825, Time: 0.04s
Iteration 152, Loss: 0.2995, Time: 0.05s
Iteration 153, Loss: 0.2625, Time: 0.05s
Iteration 154, Loss: 0.2582, Time: 0.04s
Iteration 155, Loss: 0.2653, Time: 0.04s
Iteration 156, Loss: 0.2466, Time: 0.04s
Iteration 157, Loss: 0.2484, Time: 0.05s
Iteration 158, Loss: 0.2542, Time: 0.04s
Iteration 159, Loss: 0.2520, Time: 0.05s
Iteration 160, Loss: 0.2320, Time: 0.05s
Iteration 161, Loss: 0.2694, Time: 0.04s
Iteration 162, Loss: 0.2263, Time: 0.04s
Iteration 163, Loss: 0.2528, Time: 0.04s
Iteration 164, Loss: 0.2369, Time: 0.05s
Iteration 165, Loss: 0.2618, Time: 0.04s
Iteration 166, Loss: 0.2362, Time: 0.05s
Iteration 167, Loss: 0.2281, Time: 0.04s
Iteration 168, Loss: 0.2355, Time: 0.04s
Iteration 169, Loss: 0.2152, Time: 0.04s
Iteration 170, Loss: 0.2218, Time: 0.04s
Iteration 171, Loss: 0.2205, Time: 0.04s
Iteration 172, Loss: 0.2416, Time: 0.04s
Iteration 173, Loss: 0.2283, Time: 0.04s
Iteration 174, Loss: 0.2373, Time: 0.04s
Iteration 175, Loss: 0.2205, Time: 0.04s
Iteration 176, Loss: 0.2119, Time: 0.04s
Iteration 177, Loss: 0.2334, Time: 0.04s
Iteration 178, Loss: 0.2118, Time: 0.04s
Iteration 179, Loss: 0.2124, Time: 0.05s
Iteration 180, Loss: 0.2061, Time: 0.05s
Iteration 181, Loss: 0.2116, Time: 0.04s
Iteration 182, Loss: 0.2117, Time: 0.04s
Iteration 183, Loss: 0.1974, Time: 0.05s
Iteration 184, Loss: 0.1773, Time: 0.04s
Iteration 185, Loss: 0.2001, Time: 0.04s
Iteration 186, Loss: 0.2069, Time: 0.05s
Iteration 187, Loss: 0.2045, Time: 0.04s
Iteration 188, Loss: 0.2148, Time: 0.05s
```

```
Iteration 189, Loss: 0.2191, Time: 0.04s
Iteration 190, Loss: 0.1996, Time: 0.05s
Iteration 191, Loss: 0.1834, Time: 0.05s
Iteration 192, Loss: 0.1748, Time: 0.04s
Iteration 193, Loss: 0.1854, Time: 0.04s
Iteration 194, Loss: 0.1786, Time: 0.05s
Iteration 195, Loss: 0.1891, Time: 0.05s
Iteration 196, Loss: 0.1787, Time: 0.04s
Iteration 197, Loss: 0.2028, Time: 0.05s
Iteration 198, Loss: 0.2126, Time: 0.04s
Iteration 199, Loss: 0.1839, Time: 0.04s
Iteration 200, Loss: 0.1813, Time: 0.04s
Iteration 200, Loss: 0.1813, Time: 0.04s
Test Loss: 0.1795
Iteration 201, Loss: 0.1832, Time: 0.04s
Iteration 202, Loss: 0.1943, Time: 0.05s
Iteration 203, Loss: 0.1839, Time: 0.05s
Iteration 204, Loss: 0.1916, Time: 0.04s
Iteration 205, Loss: 0.1841, Time: 0.04s
Iteration 206, Loss: 0.1815, Time: 0.04s
Iteration 207, Loss: 0.1680, Time: 0.04s
Iteration 208, Loss: 0.1617, Time: 0.04s
Iteration 209, Loss: 0.1714, Time: 0.04s
Iteration 210, Loss: 0.1621, Time: 0.04s
Iteration 211, Loss: 0.1623, Time: 0.04s
Iteration 212, Loss: 0.1703, Time: 0.04s
Iteration 213, Loss: 0.1578, Time: 0.05s
Iteration 214, Loss: 0.1677, Time: 0.04s
Iteration 215, Loss: 0.1703, Time: 0.04s
Iteration 216, Loss: 0.1597, Time: 0.04s
Iteration 217, Loss: 0.1494, Time: 0.05s
Iteration 218, Loss: 0.1605, Time: 0.05s
Iteration 219, Loss: 0.1614, Time: 0.04s
Iteration 220, Loss: 0.1473, Time: 0.05s
Iteration 221, Loss: 0.1886, Time: 0.04s
Iteration 222, Loss: 0.1604, Time: 0.04s
Iteration 223, Loss: 0.1550, Time: 0.04s
Iteration 224, Loss: 0.1753, Time: 0.05s
Iteration 225, Loss: 0.1454, Time: 0.04s
Iteration 226, Loss: 0.1431, Time: 0.04s
Iteration 227, Loss: 0.1421, Time: 0.05s
Iteration 228, Loss: 0.1507, Time: 0.04s
Iteration 229, Loss: 0.1494, Time: 0.05s
Iteration 230, Loss: 0.1454, Time: 0.04s
Iteration 231, Loss: 0.1406, Time: 0.05s
Iteration 232, Loss: 0.1390, Time: 0.05s
Iteration 233, Loss: 0.1565, Time: 0.04s
Iteration 234, Loss: 0.1524, Time: 0.05s
```

```
Iteration 235, Loss: 0.1561, Time: 0.04s
Iteration 236, Loss: 0.1432, Time: 0.04s
Iteration 237, Loss: 0.1608, Time: 0.05s
Iteration 238, Loss: 0.1380, Time: 0.04s
Iteration 239, Loss: 0.1583, Time: 0.04s
Iteration 240, Loss: 0.1435, Time: 0.04s
Iteration 241, Loss: 0.1470, Time: 0.05s
Iteration 242, Loss: 0.1542, Time: 0.04s
Iteration 243, Loss: 0.1546, Time: 0.05s
Iteration 244, Loss: 0.1382, Time: 0.06s
Iteration 245, Loss: 0.1354, Time: 0.04s
Iteration 246, Loss: 0.1427, Time: 0.05s
Iteration 247, Loss: 0.1387, Time: 0.05s
Iteration 248, Loss: 0.1358, Time: 0.06s
Iteration 249, Loss: 0.1307, Time: 0.05s
Iteration 250, Loss: 0.1394, Time: 0.05s
Iteration 251, Loss: 0.1185, Time: 0.04s
Iteration 252, Loss: 0.1476, Time: 0.04s
Iteration 253, Loss: 0.1360, Time: 0.04s
Iteration 254, Loss: 0.1379, Time: 0.04s
Iteration 255, Loss: 0.1344, Time: 0.04s
Iteration 256, Loss: 0.1493, Time: 0.04s
Iteration 257, Loss: 0.1251, Time: 0.05s
Iteration 258, Loss: 0.1277, Time: 0.05s
Iteration 259, Loss: 0.1157, Time: 0.05s
Iteration 260, Loss: 0.1100, Time: 0.04s
Iteration 261, Loss: 0.1293, Time: 0.04s
Iteration 262, Loss: 0.1314, Time: 0.04s
Iteration 263, Loss: 0.1241, Time: 0.04s
Iteration 264, Loss: 0.1547, Time: 0.05s
Iteration 265, Loss: 0.1386, Time: 0.04s
Iteration 266, Loss: 0.1397, Time: 0.05s
Iteration 267, Loss: 0.1338, Time: 0.05s
Iteration 268, Loss: 0.1276, Time: 0.04s
Iteration 269, Loss: 0.1262, Time: 0.04s
Iteration 270, Loss: 0.1204, Time: 0.05s
Iteration 271, Loss: 0.1113, Time: 0.04s
Iteration 272, Loss: 0.1198, Time: 0.04s
Iteration 273, Loss: 0.1265, Time: 0.04s
Iteration 274, Loss: 0.1078, Time: 0.04s
Iteration 275, Loss: 0.1158, Time: 0.04s
Iteration 276, Loss: 0.1302, Time: 0.04s
Iteration 277, Loss: 0.1152, Time: 0.04s
Iteration 278, Loss: 0.1149, Time: 0.04s
Iteration 279, Loss: 0.1300, Time: 0.05s
Iteration 280, Loss: 0.1193, Time: 0.04s
Iteration 281, Loss: 0.1129, Time: 0.05s
Iteration 282, Loss: 0.1170, Time: 0.04s
```

```
Iteration 283, Loss: 0.1213, Time: 0.05s
Iteration 284, Loss: 0.1094, Time: 0.04s
Iteration 285, Loss: 0.1192, Time: 0.04s
Iteration 286, Loss: 0.1089, Time: 0.04s
Iteration 287, Loss: 0.0994, Time: 0.05s
Iteration 288, Loss: 0.1101, Time: 0.05s
Iteration 289, Loss: 0.1284, Time: 0.04s
Iteration 290, Loss: 0.1235, Time: 0.04s
Iteration 291, Loss: 0.1187, Time: 0.04s
Iteration 292, Loss: 0.1036, Time: 0.04s
Iteration 293, Loss: 0.0966, Time: 0.05s
Iteration 294, Loss: 0.1284, Time: 0.04s
Iteration 295, Loss: 0.0978, Time: 0.04s
Iteration 296, Loss: 0.1040, Time: 0.04s
Iteration 297, Loss: 0.0975, Time: 0.04s
Iteration 298, Loss: 0.1201, Time: 0.04s
Iteration 299, Loss: 0.1184, Time: 0.04s
Iteration 300, Loss: 0.1167, Time: 0.05s
Iteration 300, Loss: 0.1167, Time: 0.05s
Test Loss: 0.1062
Iteration 301, Loss: 0.1197, Time: 0.04s
Iteration 302, Loss: 0.1172, Time: 0.05s
Iteration 303, Loss: 0.1267, Time: 0.04s
Iteration 304, Loss: 0.0854, Time: 0.06s
Iteration 305, Loss: 0.1060, Time: 0.04s
Iteration 306, Loss: 0.1303, Time: 0.04s
Iteration 307, Loss: 0.1198, Time: 0.04s
Iteration 308, Loss: 0.1143, Time: 0.05s
Iteration 309, Loss: 0.1084, Time: 0.05s
Iteration 310, Loss: 0.0989, Time: 0.04s
Iteration 311, Loss: 0.1177, Time: 0.04s
Iteration 312, Loss: 0.0985, Time: 0.04s
Iteration 313, Loss: 0.1315, Time: 0.04s
Iteration 314, Loss: 0.1167, Time: 0.04s
Iteration 315, Loss: 0.1141, Time: 0.05s
Iteration 316, Loss: 0.1186, Time: 0.04s
Iteration 317, Loss: 0.1079, Time: 0.04s
Iteration 318, Loss: 0.1144, Time: 0.04s
Iteration 319, Loss: 0.0826, Time: 0.04s
Iteration 320, Loss: 0.0896, Time: 0.05s
Iteration 321, Loss: 0.0983, Time: 0.05s
Iteration 322, Loss: 0.0695, Time: 0.05s
Iteration 323, Loss: 0.0856, Time: 0.05s
Iteration 324, Loss: 0.0900, Time: 0.04s
Iteration 325, Loss: 0.1110, Time: 0.04s
Iteration 326, Loss: 0.1089, Time: 0.04s
Iteration 327, Loss: 0.1053, Time: 0.04s
Iteration 328, Loss: 0.1209, Time: 0.04s
```

```
Iteration 329, Loss: 0.0799, Time: 0.04s
Iteration 330, Loss: 0.0946, Time: 0.04s
Iteration 331, Loss: 0.1177, Time: 0.04s
Iteration 332, Loss: 0.1004, Time: 0.04s
Iteration 333, Loss: 0.1134, Time: 0.04s
Iteration 334, Loss: 0.0902, Time: 0.04s
Iteration 335, Loss: 0.0991, Time: 0.04s
Iteration 336, Loss: 0.1105, Time: 0.04s
Iteration 337, Loss: 0.0844, Time: 0.05s
Iteration 338, Loss: 0.1084, Time: 0.05s
Iteration 339, Loss: 0.0997, Time: 0.04s
Iteration 340, Loss: 0.0862, Time: 0.04s
Iteration 341, Loss: 0.1052, Time: 0.04s
Iteration 342, Loss: 0.1007, Time: 0.05s
Iteration 343, Loss: 0.0965, Time: 0.04s
Iteration 344, Loss: 0.0979, Time: 0.05s
Iteration 345, Loss: 0.1125, Time: 0.04s
Iteration 346, Loss: 0.0875, Time: 0.04s
Iteration 347, Loss: 0.0885, Time: 0.04s
Iteration 348, Loss: 0.0946, Time: 0.05s
Iteration 349, Loss: 0.1380, Time: 0.04s
Iteration 350, Loss: 0.0968, Time: 0.04s
Iteration 351, Loss: 0.1095, Time: 0.04s
Iteration 352, Loss: 0.1251, Time: 0.05s
Iteration 353, Loss: 0.1012, Time: 0.05s
Iteration 354, Loss: 0.0953, Time: 0.04s
Iteration 355, Loss: 0.1213, Time: 0.05s
Iteration 356, Loss: 0.0950, Time: 0.04s
Iteration 357, Loss: 0.0776, Time: 0.04s
Iteration 358, Loss: 0.0851, Time: 0.05s
Iteration 359, Loss: 0.1016, Time: 0.05s
Iteration 360, Loss: 0.0842, Time: 0.04s
Iteration 361, Loss: 0.0799, Time: 0.04s
Iteration 362, Loss: 0.0707, Time: 0.05s
Iteration 363, Loss: 0.0885, Time: 0.04s
Iteration 364, Loss: 0.0797, Time: 0.04s
Iteration 365, Loss: 0.0939, Time: 0.04s
Iteration 366, Loss: 0.0941, Time: 0.05s
Iteration 367, Loss: 0.0980, Time: 0.05s
Iteration 368, Loss: 0.0678, Time: 0.05s
Iteration 369, Loss: 0.0662, Time: 0.04s
Iteration 370, Loss: 0.0899, Time: 0.04s
Iteration 371, Loss: 0.0880, Time: 0.05s
Iteration 372, Loss: 0.0853, Time: 0.04s
Iteration 373, Loss: 0.0793, Time: 0.04s
Iteration 374, Loss: 0.0668, Time: 0.04s
Iteration 375, Loss: 0.0651, Time: 0.04s
Iteration 376, Loss: 0.1262, Time: 0.05s
```

```
Iteration 377, Loss: 0.0891, Time: 0.05s
Iteration 378, Loss: 0.1018, Time: 0.05s
Iteration 379, Loss: 0.0748, Time: 0.04s
Iteration 380, Loss: 0.0805, Time: 0.04s
Iteration 381, Loss: 0.0833, Time: 0.05s
Iteration 382, Loss: 0.0801, Time: 0.04s
Iteration 383, Loss: 0.0808, Time: 0.05s
Iteration 384, Loss: 0.0878, Time: 0.05s
Iteration 385, Loss: 0.0919, Time: 0.06s
Iteration 386, Loss: 0.1106, Time: 0.07s
Iteration 387, Loss: 0.1001, Time: 0.07s
Iteration 388, Loss: 0.0977, Time: 0.08s
Iteration 389, Loss: 0.0901, Time: 0.06s
Iteration 390, Loss: 0.0954, Time: 0.07s
Iteration 391, Loss: 0.0798, Time: 0.06s
Iteration 392, Loss: 0.1016, Time: 0.06s
Iteration 393, Loss: 0.0852, Time: 0.06s
Iteration 394, Loss: 0.0976, Time: 0.07s
Iteration 395, Loss: 0.0929, Time: 0.06s
Iteration 396, Loss: 0.0783, Time: 0.05s
Iteration 397, Loss: 0.0828, Time: 0.06s
Iteration 398, Loss: 0.0890, Time: 0.05s
Iteration 399, Loss: 0.1047, Time: 0.06s
Iteration 400, Loss: 0.1080, Time: 0.05s
Iteration 400, Loss: 0.1080, Time: 0.05s
Test Loss: 0.0964
Iteration 401, Loss: 0.0819, Time: 0.06s
Iteration 402, Loss: 0.0773, Time: 0.05s
Iteration 403, Loss: 0.0852, Time: 0.05s
Iteration 404, Loss: 0.0882, Time: 0.05s
Iteration 405, Loss: 0.0788, Time: 0.05s
Iteration 406, Loss: 0.0789, Time: 0.04s
Iteration 407, Loss: 0.0931, Time: 0.04s
Iteration 408, Loss: 0.0729, Time: 0.04s
Iteration 409, Loss: 0.0703, Time: 0.05s
Iteration 410, Loss: 0.0783, Time: 0.04s
Iteration 411, Loss: 0.0596, Time: 0.04s
Iteration 412, Loss: 0.1063, Time: 0.05s
Iteration 413, Loss: 0.0857, Time: 0.04s
Iteration 414, Loss: 0.0662, Time: 0.05s
Iteration 415, Loss: 0.0770, Time: 0.04s
Iteration 416, Loss: 0.1028, Time: 0.04s
Iteration 417, Loss: 0.0942, Time: 0.06s
Iteration 418, Loss: 0.0774, Time: 0.04s
Iteration 419, Loss: 0.0754, Time: 0.04s
Iteration 420, Loss: 0.0619, Time: 0.04s
Iteration 421, Loss: 0.0681, Time: 0.04s
Iteration 422, Loss: 0.0882, Time: 0.05s
```

```
Iteration 423, Loss: 0.0914, Time: 0.04s
Iteration 424, Loss: 0.0809, Time: 0.05s
Iteration 425, Loss: 0.0651, Time: 0.04s
Iteration 426, Loss: 0.0649, Time: 0.05s
Iteration 427, Loss: 0.0797, Time: 0.05s
Iteration 428, Loss: 0.0853, Time: 0.04s
Iteration 429, Loss: 0.0872, Time: 0.04s
Iteration 430, Loss: 0.0933, Time: 0.05s
Iteration 431, Loss: 0.0860, Time: 0.04s
Iteration 432, Loss: 0.0814, Time: 0.05s
Iteration 433, Loss: 0.0844, Time: 0.05s
Iteration 434, Loss: 0.0952, Time: 0.04s
Iteration 435, Loss: 0.0989, Time: 0.04s
Iteration 436, Loss: 0.0882, Time: 0.05s
Iteration 437, Loss: 0.0688, Time: 0.05s
Iteration 438, Loss: 0.0877, Time: 0.05s
Iteration 439, Loss: 0.0703, Time: 0.04s
Iteration 440, Loss: 0.0653, Time: 0.04s
Iteration 441, Loss: 0.0931, Time: 0.04s
Iteration 442, Loss: 0.0769, Time: 0.05s
Iteration 443, Loss: 0.0795, Time: 0.04s
Iteration 444, Loss: 0.0899, Time: 0.05s
Iteration 445, Loss: 0.0778, Time: 0.05s
Iteration 446, Loss: 0.0692, Time: 0.05s
Iteration 447, Loss: 0.0691, Time: 0.05s
Iteration 448, Loss: 0.0694, Time: 0.04s
Iteration 449, Loss: 0.0760, Time: 0.05s
Iteration 450, Loss: 0.0836, Time: 0.04s
Iteration 451, Loss: 0.0922, Time: 0.04s
Iteration 452, Loss: 0.0882, Time: 0.06s
Iteration 453, Loss: 0.0629, Time: 0.05s
Iteration 454, Loss: 0.0740, Time: 0.05s
Iteration 455, Loss: 0.0696, Time: 0.04s
Iteration 456, Loss: 0.0825, Time: 0.05s
Iteration 457, Loss: 0.0821, Time: 0.04s
Iteration 458, Loss: 0.0928, Time: 0.06s
Iteration 459, Loss: 0.0484, Time: 0.07s
Iteration 460, Loss: 0.0738, Time: 0.07s
Iteration 461, Loss: 0.0651, Time: 0.06s
Iteration 462, Loss: 0.0757, Time: 0.07s
Iteration 463, Loss: 0.0773, Time: 0.06s
Iteration 464, Loss: 0.0679, Time: 0.07s
Iteration 465, Loss: 0.0453, Time: 0.06s
Iteration 466, Loss: 0.0656, Time: 0.06s
Iteration 467, Loss: 0.0721, Time: 0.06s
Iteration 468, Loss: 0.0676, Time: 0.05s
Iteration 469, Loss: 0.0689, Time: 0.05s
Iteration 470, Loss: 0.0807, Time: 0.05s
```

```
Iteration 471, Loss: 0.0606, Time: 0.04s
Iteration 472, Loss: 0.0839, Time: 0.06s
Iteration 473, Loss: 0.0763, Time: 0.07s
Iteration 474, Loss: 0.0699, Time: 0.05s
Iteration 475, Loss: 0.0657, Time: 0.05s
Iteration 476, Loss: 0.0868, Time: 0.04s
Iteration 477, Loss: 0.0604, Time: 0.05s
Iteration 478, Loss: 0.0879, Time: 0.05s
Iteration 479, Loss: 0.0683, Time: 0.05s
Iteration 480, Loss: 0.0950, Time: 0.05s
Iteration 481, Loss: 0.0929, Time: 0.04s
Iteration 482, Loss: 0.0762, Time: 0.04s
Iteration 483, Loss: 0.0929, Time: 0.05s
Iteration 484, Loss: 0.1032, Time: 0.04s
Iteration 485, Loss: 0.0642, Time: 0.04s
Iteration 486, Loss: 0.0572, Time: 0.04s
Iteration 487, Loss: 0.0564, Time: 0.04s
Iteration 488, Loss: 0.0831, Time: 0.05s
Iteration 489, Loss: 0.0742, Time: 0.05s
Iteration 490, Loss: 0.0685, Time: 0.05s
Iteration 491, Loss: 0.0684, Time: 0.06s
Iteration 492, Loss: 0.0549, Time: 0.04s
Iteration 493, Loss: 0.0830, Time: 0.06s
Iteration 494, Loss: 0.0601, Time: 0.06s
Iteration 495, Loss: 0.0769, Time: 0.05s
Iteration 496, Loss: 0.0649, Time: 0.05s
Iteration 497, Loss: 0.0984, Time: 0.04s
Iteration 498, Loss: 0.0638, Time: 0.05s
Iteration 499, Loss: 0.0764, Time: 0.04s
Iteration 500, Loss: 0.0925, Time: 0.04s
Iteration 500, Loss: 0.0925, Time: 0.04s
Test Loss: 0.0709
Iteration 501, Loss: 0.0574, Time: 0.04s
Iteration 502, Loss: 0.0806, Time: 0.06s
Iteration 503, Loss: 0.0780, Time: 0.05s
Iteration 504, Loss: 0.0634, Time: 0.04s
Iteration 505, Loss: 0.0884, Time: 0.04s
Iteration 506, Loss: 0.0893, Time: 0.04s
Iteration 507, Loss: 0.0757, Time: 0.04s
Iteration 508, Loss: 0.0716, Time: 0.05s
Iteration 509, Loss: 0.0772, Time: 0.04s
Iteration 510, Loss: 0.0551, Time: 0.05s
Iteration 511, Loss: 0.0603, Time: 0.05s
Iteration 512, Loss: 0.0624, Time: 0.04s
Iteration 513, Loss: 0.1022, Time: 0.04s
Iteration 514, Loss: 0.0911, Time: 0.06s
Iteration 515, Loss: 0.0588, Time: 0.04s
Iteration 516, Loss: 0.0817, Time: 0.05s
```

```
Iteration 517, Loss: 0.0619, Time: 0.04s
Iteration 518, Loss: 0.0803, Time: 0.05s
Iteration 519, Loss: 0.0798, Time: 0.04s
Iteration 520, Loss: 0.0537, Time: 0.05s
Iteration 521, Loss: 0.0663, Time: 0.04s
Iteration 522, Loss: 0.0647, Time: 0.05s
Iteration 523, Loss: 0.0717, Time: 0.05s
Iteration 524, Loss: 0.0752, Time: 0.05s
Iteration 525, Loss: 0.0726, Time: 0.04s
Iteration 526, Loss: 0.0541, Time: 0.04s
Iteration 527, Loss: 0.0519, Time: 0.04s
Iteration 528, Loss: 0.0512, Time: 0.04s
Iteration 529, Loss: 0.0700, Time: 0.04s
Iteration 530, Loss: 0.0946, Time: 0.05s
Iteration 531, Loss: 0.0506, Time: 0.05s
Iteration 532, Loss: 0.0447, Time: 0.04s
Iteration 533, Loss: 0.0514, Time: 0.05s
Iteration 534, Loss: 0.0895, Time: 0.05s
Iteration 535, Loss: 0.0588, Time: 0.04s
Iteration 536, Loss: 0.0607, Time: 0.04s
Iteration 537, Loss: 0.0700, Time: 0.05s
Iteration 538, Loss: 0.0509, Time: 0.04s
Iteration 539, Loss: 0.0543, Time: 0.04s
Iteration 540, Loss: 0.0682, Time: 0.05s
Iteration 541, Loss: 0.0592, Time: 0.05s
Iteration 542, Loss: 0.0811, Time: 0.04s
Iteration 543, Loss: 0.0498, Time: 0.04s
Iteration 544, Loss: 0.0531, Time: 0.04s
Iteration 545, Loss: 0.0618, Time: 0.04s
Iteration 546, Loss: 0.0647, Time: 0.04s
Iteration 547, Loss: 0.0856, Time: 0.05s
Iteration 548, Loss: 0.0673, Time: 0.05s
Iteration 549, Loss: 0.0735, Time: 0.05s
Iteration 550, Loss: 0.0628, Time: 0.04s
Iteration 551, Loss: 0.0802, Time: 0.04s
Iteration 552, Loss: 0.0756, Time: 0.04s
Iteration 553, Loss: 0.0546, Time: 0.04s
Iteration 554, Loss: 0.0844, Time: 0.05s
Iteration 555, Loss: 0.0731, Time: 0.04s
Iteration 556, Loss: 0.0861, Time: 0.05s
Iteration 557, Loss: 0.0708, Time: 0.05s
Iteration 558, Loss: 0.0659, Time: 0.05s
Iteration 559, Loss: 0.0731, Time: 0.04s
Iteration 560, Loss: 0.0729, Time: 0.04s
Iteration 561, Loss: 0.0764, Time: 0.04s
Iteration 562, Loss: 0.0678, Time: 0.05s
Iteration 563, Loss: 0.0704, Time: 0.07s
Iteration 564, Loss: 0.0885, Time: 0.05s
```

```
Iteration 565, Loss: 0.0802, Time: 0.06s
Iteration 566, Loss: 0.0515, Time: 0.05s
Iteration 567, Loss: 0.0668, Time: 0.04s
Iteration 568, Loss: 0.0594, Time: 0.04s
Iteration 569, Loss: 0.0604, Time: 0.05s
Iteration 570, Loss: 0.0645, Time: 0.04s
Iteration 571, Loss: 0.0429, Time: 0.04s
Iteration 572, Loss: 0.0703, Time: 0.04s
Iteration 573, Loss: 0.0545, Time: 0.04s
Iteration 574, Loss: 0.0877, Time: 0.04s
Iteration 575, Loss: 0.0722, Time: 0.05s
Iteration 576, Loss: 0.0767, Time: 0.04s
Iteration 577, Loss: 0.0519, Time: 0.05s
Iteration 578, Loss: 0.0661, Time: 0.05s
Iteration 579, Loss: 0.0623, Time: 0.04s
Iteration 580, Loss: 0.0453, Time: 0.05s
Iteration 581, Loss: 0.0770, Time: 0.04s
Iteration 582, Loss: 0.0461, Time: 0.05s
Iteration 583, Loss: 0.0782, Time: 0.05s
Iteration 584, Loss: 0.0609, Time: 0.04s
Iteration 585, Loss: 0.0666, Time: 0.06s
Iteration 586, Loss: 0.0800, Time: 0.05s
Iteration 587, Loss: 0.0561, Time: 0.05s
Iteration 588, Loss: 0.0667, Time: 0.05s
Iteration 589, Loss: 0.0661, Time: 0.05s
Iteration 590, Loss: 0.0449, Time: 0.06s
Iteration 591, Loss: 0.0552, Time: 0.06s
Iteration 592, Loss: 0.0586, Time: 0.04s
Iteration 593, Loss: 0.0564, Time: 0.04s
Iteration 594, Loss: 0.0879, Time: 0.04s
Iteration 595, Loss: 0.0611, Time: 0.05s
Iteration 596, Loss: 0.0560, Time: 0.05s
Iteration 597, Loss: 0.0552, Time: 0.05s
Iteration 598, Loss: 0.0582, Time: 0.04s
Iteration 599, Loss: 0.0696, Time: 0.04s
Iteration 600, Loss: 0.0531, Time: 0.05s
Iteration 600, Loss: 0.0531, Time: 0.05s
Test Loss: 0.0625
Iteration 601, Loss: 0.0568, Time: 0.05s
Iteration 602, Loss: 0.0823, Time: 0.05s
Iteration 603, Loss: 0.0608, Time: 0.04s
Iteration 604, Loss: 0.0824, Time: 0.05s
Iteration 605, Loss: 0.0522, Time: 0.04s
Iteration 606, Loss: 0.0517, Time: 0.05s
Iteration 607, Loss: 0.0713, Time: 0.04s
Iteration 608, Loss: 0.0630, Time: 0.05s
Iteration 609, Loss: 0.0591, Time: 0.05s
Iteration 610, Loss: 0.0629, Time: 0.04s
```

```
Iteration 611, Loss: 0.0716, Time: 0.05s
Iteration 612, Loss: 0.0629, Time: 0.04s
Iteration 613, Loss: 0.0589, Time: 0.04s
Iteration 614, Loss: 0.0620, Time: 0.05s
Iteration 615, Loss: 0.0596, Time: 0.04s
Iteration 616, Loss: 0.0633, Time: 0.05s
Iteration 617, Loss: 0.0628, Time: 0.04s
Iteration 618, Loss: 0.0507, Time: 0.05s
Iteration 619, Loss: 0.0465, Time: 0.04s
Iteration 620, Loss: 0.0685, Time: 0.04s
Iteration 621, Loss: 0.0412, Time: 0.04s
Iteration 622, Loss: 0.0619, Time: 0.04s
Iteration 623, Loss: 0.0584, Time: 0.05s
Iteration 624, Loss: 0.0721, Time: 0.05s
Iteration 625, Loss: 0.0443, Time: 0.05s
Iteration 626, Loss: 0.0614, Time: 0.04s
Iteration 627, Loss: 0.0707, Time: 0.05s
Iteration 628, Loss: 0.0515, Time: 0.04s
Iteration 629, Loss: 0.0647, Time: 0.04s
Iteration 630, Loss: 0.0623, Time: 0.05s
Iteration 631, Loss: 0.0536, Time: 0.04s
Iteration 632, Loss: 0.0657, Time: 0.06s
Iteration 633, Loss: 0.0774, Time: 0.06s
Iteration 634, Loss: 0.0641, Time: 0.04s
Iteration 635, Loss: 0.0451, Time: 0.04s
Iteration 636, Loss: 0.0611, Time: 0.06s
Iteration 637, Loss: 0.0428, Time: 0.05s
Iteration 638, Loss: 0.0684, Time: 0.05s
Iteration 639, Loss: 0.0617, Time: 0.05s
Iteration 640, Loss: 0.0660, Time: 0.04s
Iteration 641, Loss: 0.0528, Time: 0.05s
Iteration 642, Loss: 0.0532, Time: 0.04s
Iteration 643, Loss: 0.0695, Time: 0.05s
Iteration 644, Loss: 0.0483, Time: 0.04s
Iteration 645, Loss: 0.0651, Time: 0.06s
Iteration 646, Loss: 0.0774, Time: 0.04s
Iteration 647, Loss: 0.0671, Time: 0.05s
Iteration 648, Loss: 0.0360, Time: 0.04s
Iteration 649, Loss: 0.0629, Time: 0.04s
Iteration 650, Loss: 0.0552, Time: 0.05s
Iteration 651, Loss: 0.0653, Time: 0.04s
Iteration 652, Loss: 0.0447, Time: 0.04s
Iteration 653, Loss: 0.0524, Time: 0.05s
Iteration 654, Loss: 0.0548, Time: 0.04s
Iteration 655, Loss: 0.0558, Time: 0.04s
Iteration 656, Loss: 0.0642, Time: 0.05s
Iteration 657, Loss: 0.0651, Time: 0.04s
Iteration 658, Loss: 0.0464, Time: 0.05s
```

```
Iteration 659, Loss: 0.0615, Time: 0.04s
Iteration 660, Loss: 0.0548, Time: 0.04s
Iteration 661, Loss: 0.0927, Time: 0.04s
Iteration 662, Loss: 0.0613, Time: 0.05s
Iteration 663, Loss: 0.0678, Time: 0.05s
Iteration 664, Loss: 0.0430, Time: 0.04s
Iteration 665, Loss: 0.0450, Time: 0.04s
Iteration 666, Loss: 0.0762, Time: 0.04s
Iteration 667, Loss: 0.0525, Time: 0.05s
Iteration 668, Loss: 0.0454, Time: 0.04s
Iteration 669, Loss: 0.0439, Time: 0.05s
Iteration 670, Loss: 0.0408, Time: 0.05s
Iteration 671, Loss: 0.0896, Time: 0.06s
Iteration 672, Loss: 0.0819, Time: 0.06s
Iteration 673, Loss: 0.0537, Time: 0.06s
Iteration 674, Loss: 0.0703, Time: 0.06s
Iteration 675, Loss: 0.0412, Time: 0.06s
Iteration 676, Loss: 0.0467, Time: 0.04s
Iteration 677, Loss: 0.0618, Time: 0.05s
Iteration 678, Loss: 0.0618, Time: 0.04s
Iteration 679, Loss: 0.0513, Time: 0.04s
Iteration 680, Loss: 0.0644, Time: 0.05s
Iteration 681, Loss: 0.0456, Time: 0.06s
Iteration 682, Loss: 0.0663, Time: 0.04s
Iteration 683, Loss: 0.0632, Time: 0.05s
Iteration 684, Loss: 0.0573, Time: 0.05s
Iteration 685, Loss: 0.0585, Time: 0.05s
Iteration 686, Loss: 0.0710, Time: -0.10s
Iteration 687, Loss: 0.0464, Time: 0.05s
Iteration 688, Loss: 0.0744, Time: 0.04s
Iteration 689, Loss: 0.0462, Time: 0.05s
Iteration 690, Loss: 0.0772, Time: 0.04s
Iteration 691, Loss: 0.0671, Time: 0.05s
Iteration 692, Loss: 0.0619, Time: 0.05s
Iteration 693, Loss: 0.0624, Time: 0.04s
Iteration 694, Loss: 0.0581, Time: 0.04s
Iteration 695, Loss: 0.0634, Time: 0.05s
Iteration 696, Loss: 0.0632, Time: 0.05s
Iteration 697, Loss: 0.0554, Time: 0.05s
Iteration 698, Loss: 0.0584, Time: 0.05s
Iteration 699, Loss: 0.0726, Time: 0.04s
Iteration 700, Loss: 0.0641, Time: 0.04s
Iteration 700, Loss: 0.0641, Time: 0.04s
Test Loss: 0.0747
Iteration 701, Loss: 0.0550, Time: 0.04s
Iteration 702, Loss: 0.0516, Time: 0.04s
Iteration 703, Loss: 0.0691, Time: 0.04s
Iteration 704, Loss: 0.0493, Time: 0.04s
```

```
Iteration 705, Loss: 0.0487, Time: 0.04s
Iteration 706, Loss: 0.0624, Time: 0.05s
Iteration 707, Loss: 0.0517, Time: 0.05s
Iteration 708, Loss: 0.0573, Time: 0.05s
Iteration 709, Loss: 0.0757, Time: 0.04s
Iteration 710, Loss: 0.0493, Time: 0.05s
Iteration 711, Loss: 0.0412, Time: 0.04s
Iteration 712, Loss: 0.0516, Time: 0.05s
Iteration 713, Loss: 0.0740, Time: 0.04s
Iteration 714, Loss: 0.0597, Time: 0.05s
Iteration 715, Loss: 0.0487, Time: 0.04s
Iteration 716, Loss: 0.0513, Time: 0.05s
Iteration 717, Loss: 0.0473, Time: 0.05s
Iteration 718, Loss: 0.0394, Time: 0.06s
Iteration 719, Loss: 0.0557, Time: 0.04s
Iteration 720, Loss: 0.0427, Time: 0.05s
Iteration 721, Loss: 0.0637, Time: 0.05s
Iteration 722, Loss: 0.0485, Time: 0.06s
Iteration 723, Loss: 0.0701, Time: 0.05s
Iteration 724, Loss: 0.0584, Time: 0.06s
Iteration 725, Loss: 0.0398, Time: 0.05s
Iteration 726, Loss: 0.0546, Time: 0.04s
Iteration 727, Loss: 0.0574, Time: 0.05s
Iteration 728, Loss: 0.0433, Time: 0.04s
Iteration 729, Loss: 0.0657, Time: 0.05s
Iteration 730, Loss: 0.0571, Time: 0.05s
Iteration 731, Loss: 0.0640, Time: 0.05s
Iteration 732, Loss: 0.0429, Time: 0.05s
Iteration 733, Loss: 0.0673, Time: 0.05s
Iteration 734, Loss: 0.0526, Time: 0.05s
Iteration 735, Loss: 0.0470, Time: 0.05s
Iteration 736, Loss: 0.0449, Time: 0.04s
Iteration 737, Loss: 0.0546, Time: 0.04s
Iteration 738, Loss: 0.0660, Time: 0.05s
Iteration 739, Loss: 0.0421, Time: 0.04s
Iteration 740, Loss: 0.0738, Time: 0.04s
Iteration 741, Loss: 0.0874, Time: 0.04s
Iteration 742, Loss: 0.0459, Time: 0.04s
Iteration 743, Loss: 0.0529, Time: 0.04s
Iteration 744, Loss: 0.0547, Time: 0.04s
Iteration 745, Loss: 0.0580, Time: 0.05s
Iteration 746, Loss: 0.0469, Time: 0.05s
Iteration 747, Loss: 0.0350, Time: 0.04s
Iteration 748, Loss: 0.0547, Time: 0.04s
Iteration 749, Loss: 0.0455, Time: 0.05s
Iteration 750, Loss: 0.0788, Time: 0.04s
Iteration 751, Loss: 0.0557, Time: 0.05s
Iteration 752, Loss: 0.0592, Time: 0.04s
```

```
Iteration 753, Loss: 0.0576, Time: 0.05s
Iteration 754, Loss: 0.0477, Time: 0.04s
Iteration 755, Loss: 0.0696, Time: 0.05s
Iteration 756, Loss: 0.0810, Time: 0.04s
Iteration 757, Loss: 0.0643, Time: 0.04s
Iteration 758, Loss: 0.0667, Time: 0.04s
Iteration 759, Loss: 0.0742, Time: 0.04s
Iteration 760, Loss: 0.0739, Time: 0.04s
Iteration 761, Loss: 0.0609, Time: 0.05s
Iteration 762, Loss: 0.0678, Time: 0.04s
Iteration 763, Loss: 0.0617, Time: 0.04s
Iteration 764, Loss: 0.0618, Time: 0.04s
Iteration 765, Loss: 0.0506, Time: 0.05s
Iteration 766, Loss: 0.0455, Time: 0.05s
Iteration 767, Loss: 0.0648, Time: 0.04s
Iteration 768, Loss: 0.0653, Time: 0.04s
Iteration 769, Loss: 0.0637, Time: 0.05s
Iteration 770, Loss: 0.0538, Time: 0.05s
Iteration 771, Loss: 0.0977, Time: 0.04s
Iteration 772, Loss: 0.0542, Time: 0.05s
Iteration 773, Loss: 0.0535, Time: 0.04s
Iteration 774, Loss: 0.0601, Time: 0.05s
Iteration 775, Loss: 0.0582, Time: 0.05s
Iteration 776, Loss: 0.0516, Time: 0.05s
Iteration 777, Loss: 0.0540, Time: 0.05s
Iteration 778, Loss: 0.0868, Time: 0.04s
Iteration 779, Loss: 0.0461, Time: 0.05s
Iteration 780, Loss: 0.0620, Time: 0.04s
Iteration 781, Loss: 0.0622, Time: 0.04s
Iteration 782, Loss: 0.0666, Time: 0.06s
Iteration 783, Loss: 0.0362, Time: 0.05s
Iteration 784, Loss: 0.0439, Time: 0.05s
Iteration 785, Loss: 0.0580, Time: 0.05s
Iteration 786, Loss: 0.0567, Time: 0.04s
Iteration 787, Loss: 0.0591, Time: 0.05s
Iteration 788, Loss: 0.0446, Time: 0.05s
Iteration 789, Loss: 0.0666, Time: 0.04s
Iteration 790, Loss: 0.0512, Time: 0.05s
Iteration 791, Loss: 0.0527, Time: 0.05s
Iteration 792, Loss: 0.0547, Time: 0.05s
Iteration 793, Loss: 0.0539, Time: 0.04s
Iteration 794, Loss: 0.0460, Time: 0.04s
Iteration 795, Loss: 0.0438, Time: 0.05s
Iteration 796, Loss: 0.0614, Time: 0.04s
Iteration 797, Loss: 0.0643, Time: 0.04s
Iteration 798, Loss: 0.0632, Time: 0.04s
Iteration 799, Loss: 0.0411, Time: 0.04s
Iteration 800, Loss: 0.0690, Time: 0.05s
```

```
Iteration 800, Loss: 0.0690, Time: 0.05s
Test Loss: 0.0703
Iteration 801, Loss: 0.0523, Time: 0.04s
Iteration 802, Loss: 0.0474, Time: 0.05s
Iteration 803, Loss: 0.0451, Time: 0.04s
Iteration 804, Loss: 0.0719, Time: 0.05s
Iteration 805, Loss: 0.0463, Time: 0.04s
Iteration 806, Loss: 0.0865, Time: 0.04s
Iteration 807, Loss: 0.0650, Time: 0.04s
Iteration 808, Loss: 0.0532, Time: 0.04s
Iteration 809, Loss: 0.0686, Time: 0.05s
Iteration 810, Loss: 0.0616, Time: 0.05s
Iteration 811, Loss: 0.0533, Time: 0.05s
Iteration 812, Loss: 0.0547, Time: 0.05s
Iteration 813, Loss: 0.0763, Time: 0.05s
Iteration 814, Loss: 0.0428, Time: 0.05s
Iteration 815, Loss: 0.0631, Time: 0.06s
Iteration 816, Loss: 0.0566, Time: 0.05s
Iteration 817, Loss: 0.0500, Time: 0.05s
Iteration 818, Loss: 0.0469, Time: 0.05s
Iteration 819, Loss: 0.0573, Time: 0.05s
Iteration 820, Loss: 0.0484, Time: 0.05s
Iteration 821, Loss: 0.0558, Time: 0.04s
Iteration 822, Loss: 0.0357, Time: 0.05s
Iteration 823, Loss: 0.0420, Time: 0.04s
Iteration 824, Loss: 0.0554, Time: 0.05s
Iteration 825, Loss: 0.0524, Time: 0.05s
Iteration 826, Loss: 0.0718, Time: 0.04s
Iteration 827, Loss: 0.0587, Time: 0.05s
Iteration 828, Loss: 0.0547, Time: 0.05s
Iteration 829, Loss: 0.0524, Time: 0.05s
Iteration 830, Loss: 0.0574, Time: 0.05s
Iteration 831, Loss: 0.0581, Time: 0.05s
Iteration 832, Loss: 0.0559, Time: 0.04s
Iteration 833, Loss: 0.0650, Time: 0.04s
Iteration 834, Loss: 0.0658, Time: 0.05s
Iteration 835, Loss: 0.0421, Time: 0.05s
Iteration 836, Loss: 0.0624, Time: 0.04s
Iteration 837, Loss: 0.0472, Time: 0.05s
Iteration 838, Loss: 0.0703, Time: 0.05s
Iteration 839, Loss: 0.0614, Time: 0.04s
Iteration 840, Loss: 0.0444, Time: 0.05s
Iteration 841, Loss: 0.0717, Time: 0.05s
Iteration 842, Loss: 0.0527, Time: 0.05s
Iteration 843, Loss: 0.0599, Time: 0.04s
Iteration 844, Loss: 0.0499, Time: 0.04s
Iteration 845, Loss: 0.0592, Time: 0.05s
Iteration 846, Loss: 0.0400, Time: 0.05s
```

```
Iteration 847, Loss: 0.0568, Time: 0.05s
Iteration 848, Loss: 0.0501, Time: 0.05s
Iteration 849, Loss: 0.0589, Time: 0.05s
Iteration 850, Loss: 0.0606, Time: 0.05s
Iteration 851, Loss: 0.0931, Time: 0.05s
Iteration 852, Loss: 0.0541, Time: 0.04s
Iteration 853, Loss: 0.0574, Time: 0.06s
Iteration 854, Loss: 0.0350, Time: 0.06s
Iteration 855, Loss: 0.0614, Time: 0.05s
Iteration 856, Loss: 0.0490, Time: 0.05s
Iteration 857, Loss: 0.0528, Time: 0.05s
Iteration 858, Loss: 0.0657, Time: 0.04s
Iteration 859, Loss: 0.0612, Time: 0.06s
Iteration 860, Loss: 0.0740, Time: 0.06s
Iteration 861, Loss: 0.0704, Time: 0.05s
Iteration 862, Loss: 0.0364, Time: 0.04s
Iteration 863, Loss: 0.0658, Time: 0.04s
Iteration 864, Loss: 0.0314, Time: 0.05s
Iteration 865, Loss: 0.0418, Time: 0.04s
Iteration 866, Loss: 0.0729, Time: 0.06s
Iteration 867, Loss: 0.0461, Time: 0.05s
Iteration 868, Loss: 0.0767, Time: 0.05s
Iteration 869, Loss: 0.0409, Time: 0.05s
Iteration 870, Loss: 0.0419, Time: 0.05s
Iteration 871, Loss: 0.0417, Time: 0.04s
Iteration 872, Loss: 0.0311, Time: 0.05s
Iteration 873, Loss: 0.0413, Time: 0.04s
Iteration 874, Loss: 0.0500, Time: 0.05s
Iteration 875, Loss: 0.0445, Time: 0.05s
Iteration 876, Loss: 0.0575, Time: 0.05s
Iteration 877, Loss: 0.0602, Time: 0.05s
Iteration 878, Loss: 0.0476, Time: 0.04s
Iteration 879, Loss: 0.0588, Time: 0.05s
Iteration 880, Loss: 0.0630, Time: 0.05s
Iteration 881, Loss: 0.0461, Time: 0.07s
Iteration 882, Loss: 0.0461, Time: 0.07s
Iteration 883, Loss: 0.0622, Time: 0.07s
Iteration 884, Loss: 0.0481, Time: 0.05s
Iteration 885, Loss: 0.0489, Time: 0.05s
Iteration 886, Loss: 0.0496, Time: 0.06s
Iteration 887, Loss: 0.0601, Time: 0.05s
Iteration 888, Loss: 0.0710, Time: 0.04s
Iteration 889, Loss: 0.0366, Time: 0.04s
Iteration 890, Loss: 0.0529, Time: 0.05s
Iteration 891, Loss: 0.0715, Time: 0.06s
Iteration 892, Loss: 0.0370, Time: 0.05s
Iteration 893, Loss: 0.0473, Time: 0.05s
Iteration 894, Loss: 0.0598, Time: 0.06s
```

```
Iteration 895, Loss: 0.0340, Time: 0.05s
Iteration 896, Loss: 0.0514, Time: 0.04s
Iteration 897, Loss: 0.0607, Time: 0.05s
Iteration 898, Loss: 0.0716, Time: 0.05s
Iteration 899, Loss: 0.0525, Time: 0.04s
Iteration 900, Loss: 0.0516, Time: 0.05s
Iteration 900, Loss: 0.0516, Time: 0.05s
Test Loss: 0.0494
Iteration 901, Loss: 0.0797, Time: 0.04s
Iteration 902, Loss: 0.0420, Time: 0.05s
Iteration 903, Loss: 0.0508, Time: 0.04s
Iteration 904, Loss: 0.0524, Time: 0.05s
Iteration 905, Loss: 0.0378, Time: 0.06s
Iteration 906, Loss: 0.0515, Time: 0.05s
Iteration 907, Loss: 0.0564, Time: 0.06s
Iteration 908, Loss: 0.0633, Time: 0.06s
Iteration 909, Loss: 0.0446, Time: 0.05s
Iteration 910, Loss: 0.0740, Time: 0.07s
Iteration 911, Loss: 0.0510, Time: 0.05s
Iteration 912, Loss: 0.0441, Time: 0.06s
Iteration 913, Loss: 0.0313, Time: 0.05s
Iteration 914, Loss: 0.0579, Time: 0.05s
Iteration 915, Loss: 0.0456, Time: 0.07s
Iteration 916, Loss: 0.0554, Time: 0.06s
Iteration 917, Loss: 0.0459, Time: 0.06s
Iteration 918, Loss: 0.0570, Time: 0.06s
Iteration 919, Loss: 0.0488, Time: 0.06s
Iteration 920, Loss: 0.0438, Time: 0.05s
Iteration 921, Loss: 0.0664, Time: 0.05s
Iteration 922, Loss: 0.0439, Time: 0.07s
Iteration 923, Loss: 0.0620, Time: 0.05s
Iteration 924, Loss: 0.0478, Time: 0.05s
Iteration 925, Loss: 0.0463, Time: 0.06s
Iteration 926, Loss: 0.0480, Time: 0.05s
Iteration 927, Loss: 0.0556, Time: 0.05s
Iteration 928, Loss: 0.0382, Time: 0.05s
Iteration 929, Loss: 0.0409, Time: 0.05s
Iteration 930, Loss: 0.0717, Time: 0.06s
Iteration 931, Loss: 0.0414, Time: 0.05s
Iteration 932, Loss: 0.0523, Time: 0.05s
Iteration 933, Loss: 0.0493, Time: 0.06s
Iteration 934, Loss: 0.0476, Time: 0.05s
Iteration 935, Loss: 0.0487, Time: 0.06s
Iteration 936, Loss: 0.0441, Time: 0.05s
Iteration 937, Loss: 0.0578, Time: 0.04s
Iteration 938, Loss: 0.0594, Time: 0.04s
Iteration 939, Loss: 0.0562, Time: 0.06s
Iteration 940, Loss: 0.0661, Time: 0.05s
```

```
Iteration 941, Loss: 0.0397, Time: 0.04s
Iteration 942, Loss: 0.0613, Time: 0.04s
Iteration 943, Loss: 0.0535, Time: 0.04s
Iteration 944, Loss: 0.0412, Time: 0.04s
Iteration 945, Loss: 0.0454, Time: 0.05s
Iteration 946, Loss: 0.0570, Time: 0.05s
Iteration 947, Loss: 0.0633, Time: 0.04s
Iteration 948, Loss: 0.0448, Time: 0.04s
Iteration 949, Loss: 0.0491, Time: 0.05s
Iteration 950, Loss: 0.0422, Time: 0.04s
Iteration 951, Loss: 0.0491, Time: 0.04s
Iteration 952, Loss: 0.0468, Time: 0.04s
Iteration 953, Loss: 0.0523, Time: 0.04s
Iteration 954, Loss: 0.0436, Time: 0.05s
Iteration 955, Loss: 0.0584, Time: 0.05s
Iteration 956, Loss: 0.0468, Time: 0.04s
Iteration 957, Loss: 0.0580, Time: 0.04s
Iteration 958, Loss: 0.0663, Time: 0.05s
Iteration 959, Loss: 0.0725, Time: 0.05s
Iteration 960, Loss: 0.0424, Time: 0.05s
Iteration 961, Loss: 0.0637, Time: 0.04s
Iteration 962, Loss: 0.0561, Time: 0.04s
Iteration 963, Loss: 0.0375, Time: 0.05s
Iteration 964, Loss: 0.0461, Time: 0.04s
Iteration 965, Loss: 0.0526, Time: 0.04s
Iteration 966, Loss: 0.0580, Time: 0.04s
Iteration 967, Loss: 0.0582, Time: 0.05s
Iteration 968, Loss: 0.0430, Time: 0.05s
Iteration 969, Loss: 0.0447, Time: 0.04s
Iteration 970, Loss: 0.0568, Time: 0.05s
Iteration 971, Loss: 0.0447, Time: 0.04s
Iteration 972, Loss: 0.0433, Time: 0.05s
Iteration 973, Loss: 0.0465, Time: 0.05s
Iteration 974, Loss: 0.0555, Time: 0.05s
Iteration 975, Loss: 0.0562, Time: 0.05s
Iteration 976, Loss: 0.0437, Time: 0.04s
Iteration 977, Loss: 0.0532, Time: 0.05s
Iteration 978, Loss: 0.0524, Time: 0.05s
Iteration 979, Loss: 0.0560, Time: 0.04s
Iteration 980, Loss: 0.0347, Time: 0.05s
Iteration 981, Loss: 0.0542, Time: 0.05s
Iteration 982, Loss: 0.0594, Time: 0.04s
Iteration 983, Loss: 0.0333, Time: 0.04s
Iteration 984, Loss: 0.0510, Time: 0.05s
Iteration 985, Loss: 0.0561, Time: 0.04s
Iteration 986, Loss: 0.0409, Time: 0.04s
Iteration 987, Loss: 0.0378, Time: 0.04s
Iteration 988, Loss: 0.0273, Time: 0.04s
```

```
Iteration 989, Loss: 0.0618, Time: 0.05s
Iteration 990, Loss: 0.0390, Time: 0.04s
Iteration 991, Loss: 0.0322, Time: 0.05s
Iteration 992, Loss: 0.0600, Time: 0.04s
Iteration 993, Loss: 0.0421, Time: 0.05s
Iteration 994, Loss: 0.0532, Time: 0.05s
Iteration 995, Loss: 0.0510, Time: 0.04s
Iteration 996, Loss: 0.0674, Time: 0.04s
Iteration 997, Loss: 0.0444, Time: 0.04s
Iteration 998, Loss: 0.0527, Time: 0.05s
Iteration 999, Loss: 0.0501, Time: 0.04s
Iteration 1000, Loss: 0.0648, Time: 0.05s
Iteration 1000, Loss: 0.0648, Time: 0.05s
Test Loss: 0.0496
Iteration 1001, Loss: 0.0611, Time: 0.05s
Iteration 1002, Loss: 0.0544, Time: 0.07s
Iteration 1003, Loss: 0.0595, Time: 0.04s
Iteration 1004, Loss: 0.0459, Time: 0.04s
Iteration 1005, Loss: 0.0466, Time: 0.05s
Iteration 1006, Loss: 0.0523, Time: 0.04s
Iteration 1007, Loss: 0.0526, Time: 0.05s
Iteration 1008, Loss: 0.0438, Time: 0.05s
Iteration 1009, Loss: 0.0388, Time: 0.04s
Iteration 1010, Loss: 0.0593, Time: 0.05s
Iteration 1011, Loss: 0.0624, Time: 0.05s
Iteration 1012, Loss: 0.0686, Time: 0.05s
Iteration 1013, Loss: 0.0403, Time: 0.05s
Iteration 1014, Loss: 0.0450, Time: 0.05s
Iteration 1015, Loss: 0.0512, Time: 0.06s
Iteration 1016, Loss: 0.0369, Time: 0.04s
Iteration 1017, Loss: 0.0465, Time: 0.06s
Iteration 1018, Loss: 0.0317, Time: 0.04s
Iteration 1019, Loss: 0.0448, Time: 0.04s
Iteration 1020, Loss: 0.0519, Time: 0.05s
Iteration 1021, Loss: 0.0571, Time: 0.04s
Iteration 1022, Loss: 0.0501, Time: 0.05s
Iteration 1023, Loss: 0.0608, Time: 0.06s
Iteration 1024, Loss: 0.0457, Time: 0.05s
Iteration 1025, Loss: 0.0495, Time: 0.05s
Iteration 1026, Loss: 0.0563, Time: 0.05s
Iteration 1027, Loss: 0.0882, Time: 0.06s
Iteration 1028, Loss: 0.0412, Time: 0.04s
Iteration 1029, Loss: 0.0435, Time: 0.04s
Iteration 1030, Loss: 0.0377, Time: 0.05s
Iteration 1031, Loss: 0.0475, Time: 0.06s
Iteration 1032, Loss: 0.0493, Time: 0.06s
Iteration 1033, Loss: 0.0579, Time: 0.04s
Iteration 1034, Loss: 0.0591, Time: 0.05s
```

```
Iteration 1035, Loss: 0.0594, Time: 0.05s
Iteration 1036, Loss: 0.0529, Time: 0.05s
Iteration 1037, Loss: 0.0624, Time: 0.05s
Iteration 1038, Loss: 0.0490, Time: 0.04s
Iteration 1039, Loss: 0.0587, Time: 0.05s
Iteration 1040, Loss: 0.0589, Time: 0.04s
Iteration 1041, Loss: 0.0647, Time: 0.05s
Iteration 1042, Loss: 0.0469, Time: 0.04s
Iteration 1043, Loss: 0.0486, Time: 0.04s
Iteration 1044, Loss: 0.0477, Time: 0.05s
Iteration 1045, Loss: 0.0469, Time: 0.04s
Iteration 1046, Loss: 0.0342, Time: 0.05s
Iteration 1047, Loss: 0.0742, Time: 0.04s
Iteration 1048, Loss: 0.0396, Time: 0.06s
Iteration 1049, Loss: 0.0538, Time: 0.05s
Iteration 1050, Loss: 0.0359, Time: 0.06s
Iteration 1051, Loss: 0.0699, Time: 0.05s
Iteration 1052, Loss: 0.0447, Time: 0.05s
Iteration 1053, Loss: 0.0521, Time: 0.04s
Iteration 1054, Loss: 0.0393, Time: 0.05s
Iteration 1055, Loss: 0.0548, Time: 0.04s
Iteration 1056, Loss: 0.0703, Time: 0.05s
Iteration 1057, Loss: 0.0401, Time: 0.05s
Iteration 1058, Loss: 0.0575, Time: 0.05s
Iteration 1059, Loss: 0.0569, Time: 0.05s
Iteration 1060, Loss: 0.0617, Time: 0.05s
Iteration 1061, Loss: 0.0338, Time: 0.05s
Iteration 1062, Loss: 0.0409, Time: 0.05s
Iteration 1063, Loss: 0.0370, Time: 0.05s
Iteration 1064, Loss: 0.0568, Time: 0.04s
Iteration 1065, Loss: 0.0689, Time: 0.05s
Iteration 1066, Loss: 0.0355, Time: 0.05s
Iteration 1067, Loss: 0.0592, Time: 0.04s
Iteration 1068, Loss: 0.0467, Time: 0.04s
Iteration 1069, Loss: 0.0487, Time: 0.05s
Iteration 1070, Loss: 0.0513, Time: 0.04s
Iteration 1071, Loss: 0.0461, Time: 0.04s
Iteration 1072, Loss: 0.0574, Time: 0.05s
Iteration 1073, Loss: 0.0585, Time: 0.04s
Iteration 1074, Loss: 0.0371, Time: 0.05s
Iteration 1075, Loss: 0.0543, Time: 0.04s
Iteration 1076, Loss: 0.0325, Time: 0.06s
Iteration 1077, Loss: 0.0344, Time: 0.04s
Iteration 1078, Loss: 0.0509, Time: 0.04s
Iteration 1079, Loss: 0.0495, Time: 0.05s
Iteration 1080, Loss: 0.0347, Time: 0.04s
Iteration 1081, Loss: 0.0615, Time: 0.04s
Iteration 1082, Loss: 0.0651, Time: 0.04s
```

```
Iteration 1083, Loss: 0.0536, Time: 0.05s
Iteration 1084, Loss: 0.0583, Time: 0.05s
Iteration 1085, Loss: 0.0488, Time: 0.05s
Iteration 1086, Loss: 0.0510, Time: 0.07s
Iteration 1087, Loss: 0.0512, Time: 0.06s
Iteration 1088, Loss: 0.0407, Time: 0.05s
Iteration 1089, Loss: 0.0464, Time: 0.06s
Iteration 1090, Loss: 0.0566, Time: 0.07s
Iteration 1091, Loss: 0.0504, Time: 0.05s
Iteration 1092, Loss: 0.0428, Time: 0.06s
Iteration 1093, Loss: 0.0425, Time: 0.05s
Iteration 1094, Loss: 0.0557, Time: 0.05s
Iteration 1095, Loss: 0.0379, Time: 0.04s
Iteration 1096, Loss: 0.0604, Time: 0.04s
Iteration 1097, Loss: 0.0515, Time: 0.05s
Iteration 1098, Loss: 0.0494, Time: 0.05s
Iteration 1099, Loss: 0.0359, Time: 0.04s
Iteration 1100, Loss: 0.0352, Time: 0.05s
Iteration 1100, Loss: 0.0352, Time: 0.05s
Test Loss: 0.0767
Iteration 1101, Loss: 0.0533, Time: 0.05s
Iteration 1102, Loss: 0.0361, Time: 0.04s
Iteration 1103, Loss: 0.0548, Time: 0.04s
Iteration 1104, Loss: 0.0351, Time: 0.04s
Iteration 1105, Loss: 0.0423, Time: 0.04s
Iteration 1106, Loss: 0.0512, Time: 0.04s
Iteration 1107, Loss: 0.0610, Time: 0.04s
Iteration 1108, Loss: 0.0590, Time: 0.05s
Iteration 1109, Loss: 0.0593, Time: 0.04s
Iteration 1110, Loss: 0.0481, Time: 0.04s
Iteration 1111, Loss: 0.0404, Time: 0.05s
Iteration 1112, Loss: 0.0585, Time: 0.04s
Iteration 1113, Loss: 0.0401, Time: 0.05s
Iteration 1114, Loss: 0.0515, Time: 0.05s
Iteration 1115, Loss: 0.0432, Time: 0.05s
Iteration 1116, Loss: 0.0535, Time: 0.05s
Iteration 1117, Loss: 0.0431, Time: 0.05s
Iteration 1118, Loss: 0.0712, Time: 0.05s
Iteration 1119, Loss: 0.0489, Time: 0.04s
Iteration 1120, Loss: 0.0482, Time: 0.04s
Iteration 1121, Loss: 0.0480, Time: 0.04s
Iteration 1122, Loss: 0.0732, Time: 0.04s
Iteration 1123, Loss: 0.0577, Time: 0.05s
Iteration 1124, Loss: 0.0533, Time: 0.04s
Iteration 1125, Loss: 0.0380, Time: 0.04s
Iteration 1126, Loss: 0.0304, Time: 0.05s
Iteration 1127, Loss: 0.0390, Time: 0.04s
Iteration 1128, Loss: 0.0381, Time: 0.05s
```

```
Iteration 1129, Loss: 0.0632, Time: 0.04s
Iteration 1130, Loss: 0.0525, Time: 0.04s
Iteration 1131, Loss: 0.0308, Time: 0.04s
Iteration 1132, Loss: 0.0540, Time: 0.05s
Iteration 1133, Loss: 0.0427, Time: 0.05s
Iteration 1134, Loss: 0.0414, Time: 0.05s
Iteration 1135, Loss: 0.0430, Time: 0.05s
Iteration 1136, Loss: 0.0384, Time: 0.04s
Iteration 1137, Loss: 0.0384, Time: 0.04s
Iteration 1138, Loss: 0.0442, Time: 0.04s
Iteration 1139, Loss: 0.0466, Time: 0.04s
Iteration 1140, Loss: 0.0594, Time: 0.04s
Iteration 1141, Loss: 0.0514, Time: 0.04s
Iteration 1142, Loss: 0.0333, Time: 0.04s
Iteration 1143, Loss: 0.0440, Time: 0.05s
Iteration 1144, Loss: 0.0402, Time: 0.05s
Iteration 1145, Loss: 0.0448, Time: 0.04s
Iteration 1146, Loss: 0.0339, Time: 0.05s
Iteration 1147, Loss: 0.0566, Time: 0.04s
Iteration 1148, Loss: 0.0336, Time: 0.04s
Iteration 1149, Loss: 0.0380, Time: 0.05s
Iteration 1150, Loss: 0.0541, Time: 0.04s
Iteration 1151, Loss: 0.0401, Time: 0.05s
Iteration 1152, Loss: 0.0453, Time: 0.04s
Iteration 1153, Loss: 0.0487, Time: 0.05s
Iteration 1154, Loss: 0.0496, Time: 0.04s
Iteration 1155, Loss: 0.0333, Time: 0.04s
Iteration 1156, Loss: 0.0716, Time: 0.04s
Iteration 1157, Loss: 0.0407, Time: 0.04s
Iteration 1158, Loss: 0.0353, Time: 0.05s
Iteration 1159, Loss: 0.0483, Time: 0.04s
Iteration 1160, Loss: 0.0296, Time: 0.04s
Iteration 1161, Loss: 0.0419, Time: 0.05s
Iteration 1162, Loss: 0.0322, Time: 0.04s
Iteration 1163, Loss: 0.0437, Time: 0.05s
Iteration 1164, Loss: 0.0538, Time: 0.04s
Iteration 1165, Loss: 0.0376, Time: 0.04s
Iteration 1166, Loss: 0.0397, Time: 0.05s
Iteration 1167, Loss: 0.0558, Time: 0.04s
Iteration 1168, Loss: 0.0635, Time: 0.05s
Iteration 1169, Loss: 0.0784, Time: 0.04s
Iteration 1170, Loss: 0.0483, Time: 0.04s
Iteration 1171, Loss: 0.0676, Time: 0.04s
Iteration 1172, Loss: 0.0381, Time: 0.05s
Iteration 1173, Loss: 0.0442, Time: 0.05s
Iteration 1174, Loss: 0.0509, Time: 0.05s
Iteration 1175, Loss: 0.0454, Time: 0.04s
Iteration 1176, Loss: 0.0425, Time: 0.04s
```

```
Iteration 1177, Loss: 0.0466, Time: 0.05s
Iteration 1178, Loss: 0.0388, Time: 0.04s
Iteration 1179, Loss: 0.0594, Time: 0.05s
Iteration 1180, Loss: 0.0425, Time: 0.05s
Iteration 1181, Loss: 0.0696, Time: 0.04s
Iteration 1182, Loss: 0.0286, Time: 0.05s
Iteration 1183, Loss: 0.0418, Time: 0.04s
Iteration 1184, Loss: 0.0650, Time: 0.04s
Iteration 1185, Loss: 0.0550, Time: 0.04s
Iteration 1186, Loss: 0.0340, Time: 0.05s
Iteration 1187, Loss: 0.0455, Time: 0.04s
Iteration 1188, Loss: 0.0447, Time: 0.05s
Iteration 1189, Loss: 0.0476, Time: 0.04s
Iteration 1190, Loss: 0.0421, Time: 0.04s
Iteration 1191, Loss: 0.0445, Time: 0.05s
Iteration 1192, Loss: 0.0443, Time: 0.04s
Iteration 1193, Loss: 0.0570, Time: 0.05s
Iteration 1194, Loss: 0.0430, Time: 0.04s
Iteration 1195, Loss: 0.0439, Time: 0.04s
Iteration 1196, Loss: 0.0602, Time: 0.04s
Iteration 1197, Loss: 0.0364, Time: 0.05s
Iteration 1198, Loss: 0.0413, Time: 0.06s
Iteration 1199, Loss: 0.0400, Time: 0.04s
Iteration 1200, Loss: 0.0403, Time: 0.05s
Iteration 1200, Loss: 0.0403, Time: 0.05s
Test Loss: 0.0286
Iteration 1201, Loss: 0.0463, Time: 0.05s
Iteration 1202, Loss: 0.0418, Time: 0.04s
Iteration 1203, Loss: 0.0566, Time: 0.05s
Iteration 1204, Loss: 0.0339, Time: 0.05s
Iteration 1205, Loss: 0.0382, Time: 0.04s
Iteration 1206, Loss: 0.0760, Time: 0.05s
Iteration 1207, Loss: 0.0488, Time: 0.04s
Iteration 1208, Loss: 0.0412, Time: 0.04s
Iteration 1209, Loss: 0.0441, Time: 0.04s
Iteration 1210, Loss: 0.0451, Time: 0.04s
Iteration 1211, Loss: 0.0479, Time: 0.05s
Iteration 1212, Loss: 0.0487, Time: 0.04s
Iteration 1213, Loss: 0.0413, Time: 0.04s
Iteration 1214, Loss: 0.0468, Time: 0.05s
Iteration 1215, Loss: 0.0611, Time: 0.04s
Iteration 1216, Loss: 0.0325, Time: 0.05s
Iteration 1217, Loss: 0.0594, Time: 0.05s
Iteration 1218, Loss: 0.0317, Time: 0.04s
Iteration 1219, Loss: 0.0457, Time: 0.05s
Iteration 1220, Loss: 0.0631, Time: 0.04s
Iteration 1221, Loss: 0.0384, Time: 0.05s
Iteration 1222, Loss: 0.0536, Time: 0.04s
```

```
Iteration 1223, Loss: 0.0602, Time: 0.05s
Iteration 1224, Loss: 0.0402, Time: 0.04s
Iteration 1225, Loss: 0.0559, Time: 0.06s
Iteration 1226, Loss: 0.0391, Time: 0.06s
Iteration 1227, Loss: 0.0418, Time: 0.05s
Iteration 1228, Loss: 0.0446, Time: 0.05s
Iteration 1229, Loss: 0.0450, Time: 0.04s
Iteration 1230, Loss: 0.0514, Time: 0.05s
Iteration 1231, Loss: 0.0314, Time: 0.04s
Iteration 1232, Loss: 0.0387, Time: 0.05s
Iteration 1233, Loss: 0.0396, Time: 0.04s
Iteration 1234, Loss: 0.0364, Time: 0.04s
Iteration 1235, Loss: 0.0542, Time: 0.04s
Iteration 1236, Loss: 0.0451, Time: 0.04s
Iteration 1237, Loss: 0.0566, Time: 0.04s
Iteration 1238, Loss: 0.0696, Time: 0.05s
Iteration 1239, Loss: 0.0678, Time: 0.05s
Iteration 1240, Loss: 0.0466, Time: 0.04s
Iteration 1241, Loss: 0.0505, Time: 0.05s
Iteration 1242, Loss: 0.0453, Time: 0.05s
Iteration 1243, Loss: 0.0524, Time: 0.05s
Iteration 1244, Loss: 0.0367, Time: 0.05s
Iteration 1245, Loss: 0.0654, Time: 0.05s
Iteration 1246, Loss: 0.0421, Time: 0.05s
Iteration 1247, Loss: 0.0379, Time: 0.05s
Iteration 1248, Loss: 0.0519, Time: 0.06s
Iteration 1249, Loss: 0.0386, Time: 0.05s
Iteration 1250, Loss: 0.0392, Time: 0.05s
Iteration 1251, Loss: 0.0346, Time: 0.06s
Iteration 1252, Loss: 0.0330, Time: 0.05s
Iteration 1253, Loss: 0.0438, Time: 0.05s
Iteration 1254, Loss: 0.0384, Time: 0.05s
Iteration 1255, Loss: 0.0226, Time: 0.05s
Iteration 1256, Loss: 0.0407, Time: 0.05s
Iteration 1257, Loss: 0.0662, Time: 0.04s
Iteration 1258, Loss: 0.0471, Time: 0.05s
Iteration 1259, Loss: 0.0315, Time: 0.05s
Iteration 1260, Loss: 0.0524, Time: 0.05s
Iteration 1261, Loss: 0.0404, Time: 0.05s
Iteration 1262, Loss: 0.0328, Time: 0.05s
Iteration 1263, Loss: 0.0653, Time: 0.05s
Iteration 1264, Loss: 0.0590, Time: 0.05s
Iteration 1265, Loss: 0.0532, Time: 0.05s
Iteration 1266, Loss: 0.0257, Time: 0.05s
Iteration 1267, Loss: 0.0358, Time: 0.05s
Iteration 1268, Loss: 0.0502, Time: 0.04s
Iteration 1269, Loss: 0.0343, Time: 0.05s
Iteration 1270, Loss: 0.0386, Time: 0.05s
```

```
Iteration 1271, Loss: 0.0644, Time: 0.05s
Iteration 1272, Loss: 0.0480, Time: 0.05s
Iteration 1273, Loss: 0.0364, Time: 0.04s
Iteration 1274, Loss: 0.0369, Time: 0.04s
Iteration 1275, Loss: 0.0548, Time: 0.05s
Iteration 1276, Loss: 0.0351, Time: 0.05s
Iteration 1277, Loss: 0.0421, Time: 0.04s
Iteration 1278, Loss: 0.0482, Time: 0.05s
Iteration 1279, Loss: 0.0418, Time: 0.04s
Iteration 1280, Loss: 0.0356, Time: 0.04s
Iteration 1281, Loss: 0.0400, Time: 0.05s
Iteration 1282, Loss: 0.0321, Time: 0.05s
Iteration 1283, Loss: 0.0457, Time: 0.05s
Iteration 1284, Loss: 0.0418, Time: 0.05s
Iteration 1285, Loss: 0.0532, Time: 0.05s
Iteration 1286, Loss: 0.0466, Time: 0.05s
Iteration 1287, Loss: 0.0403, Time: 0.05s
Iteration 1288, Loss: 0.0393, Time: 0.05s
Iteration 1289, Loss: 0.0408, Time: 0.04s
Iteration 1290, Loss: 0.0343, Time: 0.04s
Iteration 1291, Loss: 0.0435, Time: 0.05s
Iteration 1292, Loss: 0.0461, Time: 0.05s
Iteration 1293, Loss: 0.0452, Time: 0.04s
Iteration 1294, Loss: 0.0455, Time: 0.04s
Iteration 1295, Loss: 0.0485, Time: 0.05s
Iteration 1296, Loss: 0.0350, Time: 0.04s
Iteration 1297, Loss: 0.0349, Time: 0.04s
Iteration 1298, Loss: 0.0454, Time: 0.04s
Iteration 1299, Loss: 0.0634, Time: 0.05s
Iteration 1300, Loss: 0.0464, Time: 0.04s
Iteration 1300, Loss: 0.0464, Time: 0.04s
Test Loss: 0.0462
Iteration 1301, Loss: 0.0745, Time: 0.06s
Iteration 1302, Loss: 0.0351, Time: 0.06s
Iteration 1303, Loss: 0.0653, Time: 0.05s
Iteration 1304, Loss: 0.0564, Time: 0.06s
Iteration 1305, Loss: 0.0476, Time: 0.07s
Iteration 1306, Loss: 0.0554, Time: 0.05s
Iteration 1307, Loss: 0.0529, Time: 0.05s
Iteration 1308, Loss: 0.0566, Time: 0.04s
Iteration 1309, Loss: 0.0453, Time: 0.05s
Iteration 1310, Loss: 0.0410, Time: 0.04s
Iteration 1311, Loss: 0.0414, Time: 0.04s
Iteration 1312, Loss: 0.0580, Time: 0.04s
Iteration 1313, Loss: 0.0363, Time: -0.10s
Iteration 1314, Loss: 0.0468, Time: 0.04s
Iteration 1315, Loss: 0.0240, Time: 0.04s
Iteration 1316, Loss: 0.0375, Time: 0.04s
```

```
Iteration 1317, Loss: 0.0371, Time: 0.04s
Iteration 1318, Loss: 0.0648, Time: 0.04s
Iteration 1319, Loss: 0.0540, Time: 0.04s
Iteration 1320, Loss: 0.0322, Time: 0.04s
Iteration 1321, Loss: 0.0495, Time: 0.04s
Iteration 1322, Loss: 0.0533, Time: 0.04s
Iteration 1323, Loss: 0.0394, Time: 0.04s
Iteration 1324, Loss: 0.0439, Time: 0.04s
Iteration 1325, Loss: 0.0430, Time: 0.04s
Iteration 1326, Loss: 0.0302, Time: 0.05s
Iteration 1327, Loss: 0.0528, Time: 0.05s
Iteration 1328, Loss: 0.0542, Time: 0.04s
Iteration 1329, Loss: 0.0430, Time: 0.04s
Iteration 1330, Loss: 0.0418, Time: 0.04s
Iteration 1331, Loss: 0.0341, Time: 0.04s
Iteration 1332, Loss: 0.0400, Time: 0.04s
Iteration 1333, Loss: 0.0390, Time: 0.04s
Iteration 1334, Loss: 0.0585, Time: 0.04s
Iteration 1335, Loss: 0.0476, Time: 0.04s
Iteration 1336, Loss: 0.0255, Time: 0.04s
Iteration 1337, Loss: 0.0411, Time: 0.05s
Iteration 1338, Loss: 0.0461, Time: 0.04s
Iteration 1339, Loss: 0.0402, Time: 0.04s
Iteration 1340, Loss: 0.0449, Time: 0.04s
Iteration 1341, Loss: 0.0277, Time: 0.04s
Iteration 1342, Loss: 0.0520, Time: 0.04s
Iteration 1343, Loss: 0.0530, Time: 0.04s
Iteration 1344, Loss: 0.0454, Time: 0.05s
Iteration 1345, Loss: 0.0494, Time: 0.04s
Iteration 1346, Loss: 0.0358, Time: 0.04s
Iteration 1347, Loss: 0.0493, Time: 0.04s
Iteration 1348, Loss: 0.0349, Time: 0.04s
Iteration 1349, Loss: 0.0352, Time: 0.04s
Iteration 1350, Loss: 0.0413, Time: 0.05s
Iteration 1351, Loss: 0.0370, Time: 0.04s
Iteration 1352, Loss: 0.0378, Time: 0.04s
Iteration 1353, Loss: 0.0458, Time: 0.04s
Iteration 1354, Loss: 0.0449, Time: 0.04s
Iteration 1355, Loss: 0.0500, Time: 0.04s
Iteration 1356, Loss: 0.0468, Time: 0.04s
Iteration 1357, Loss: 0.0400, Time: 0.05s
Iteration 1358, Loss: 0.0355, Time: 0.04s
Iteration 1359, Loss: 0.0510, Time: 0.04s
Iteration 1360, Loss: 0.0424, Time: 0.04s
Iteration 1361, Loss: 0.0489, Time: 0.04s
Iteration 1362, Loss: 0.0434, Time: 0.05s
Iteration 1363, Loss: 0.0404, Time: 0.05s
Iteration 1364, Loss: 0.0526, Time: 0.04s
```

```
Iteration 1365, Loss: 0.0394, Time: 0.04s
Iteration 1366, Loss: 0.0468, Time: 0.04s
Iteration 1367, Loss: 0.0378, Time: 0.04s
Iteration 1368, Loss: 0.0416, Time: 0.04s
Iteration 1369, Loss: 0.0295, Time: 0.04s
Iteration 1370, Loss: 0.0585, Time: 0.04s
Iteration 1371, Loss: 0.0357, Time: 0.04s
Iteration 1372, Loss: 0.0265, Time: 0.04s
Iteration 1373, Loss: 0.0366, Time: 0.05s
Iteration 1374, Loss: 0.0422, Time: 0.04s
Iteration 1375, Loss: 0.0644, Time: 0.05s
Iteration 1376, Loss: 0.0478, Time: 0.04s
Iteration 1377, Loss: 0.0641, Time: 0.04s
Iteration 1378, Loss: 0.0388, Time: 0.04s
Iteration 1379, Loss: 0.0436, Time: 0.05s
Iteration 1380, Loss: 0.0403, Time: 0.04s
Iteration 1381, Loss: 0.0420, Time: 0.04s
Iteration 1382, Loss: 0.0426, Time: 0.04s
Iteration 1383, Loss: 0.0478, Time: 0.04s
Iteration 1384, Loss: 0.0422, Time: 0.05s
Iteration 1385, Loss: 0.0376, Time: 0.04s
Iteration 1386, Loss: 0.0379, Time: 0.04s
Iteration 1387, Loss: 0.0351, Time: 0.05s
Iteration 1388, Loss: 0.0377, Time: 0.04s
Iteration 1389, Loss: 0.0381, Time: 0.05s
Iteration 1390, Loss: 0.0380, Time: 0.04s
Iteration 1391, Loss: 0.0504, Time: 0.05s
Iteration 1392, Loss: 0.0350, Time: 0.05s
Iteration 1393, Loss: 0.0446, Time: 0.04s
Iteration 1394, Loss: 0.0342, Time: 0.04s
Iteration 1395, Loss: 0.0586, Time: 0.04s
Iteration 1396, Loss: 0.0345, Time: 0.05s
Iteration 1397, Loss: 0.0358, Time: 0.04s
Iteration 1398, Loss: 0.0355, Time: 0.04s
Iteration 1399, Loss: 0.0542, Time: 0.04s
Iteration 1400, Loss: 0.0354, Time: 0.04s
Iteration 1400, Loss: 0.0354, Time: 0.04s
Test Loss: 0.0432
Iteration 1401, Loss: 0.0311, Time: 0.04s
Iteration 1402, Loss: 0.0418, Time: 0.04s
Iteration 1403, Loss: 0.0428, Time: 0.05s
Iteration 1404, Loss: 0.0600, Time: 0.05s
Iteration 1405, Loss: 0.0439, Time: 0.05s
Iteration 1406, Loss: 0.0306, Time: 0.05s
Iteration 1407, Loss: 0.0325, Time: 0.05s
Iteration 1408, Loss: 0.0364, Time: 0.04s
Iteration 1409, Loss: 0.0338, Time: 0.04s
Iteration 1410, Loss: 0.0346, Time: 0.04s
```

```
Iteration 1411, Loss: 0.0377, Time: 0.04s
Iteration 1412, Loss: 0.0423, Time: 0.05s
Iteration 1413, Loss: 0.0657, Time: 0.04s
Iteration 1414, Loss: 0.0448, Time: 0.05s
Iteration 1415, Loss: 0.0317, Time: 0.04s
Iteration 1416, Loss: 0.0328, Time: 0.05s
Iteration 1417, Loss: 0.0305, Time: 0.05s
Iteration 1418, Loss: 0.0339, Time: 0.04s
Iteration 1419, Loss: 0.0488, Time: 0.05s
Iteration 1420, Loss: 0.0413, Time: 0.05s
Iteration 1421, Loss: 0.0522, Time: 0.05s
Iteration 1422, Loss: 0.0464, Time: 0.04s
Iteration 1423, Loss: 0.0389, Time: 0.04s
Iteration 1424, Loss: 0.0413, Time: 0.05s
Iteration 1425, Loss: 0.0424, Time: 0.05s
Iteration 1426, Loss: 0.0481, Time: 0.04s
Iteration 1427, Loss: 0.0301, Time: 0.04s
Iteration 1428, Loss: 0.0509, Time: 0.04s
Iteration 1429, Loss: 0.0508, Time: 0.05s
Iteration 1430, Loss: 0.0396, Time: 0.04s
Iteration 1431, Loss: 0.0498, Time: 0.04s
Iteration 1432, Loss: 0.0424, Time: 0.04s
Iteration 1433, Loss: 0.0421, Time: 0.04s
Iteration 1434, Loss: 0.0515, Time: 0.04s
Iteration 1435, Loss: 0.0262, Time: 0.05s
Iteration 1436, Loss: 0.0365, Time: 0.05s
Iteration 1437, Loss: 0.0303, Time: 0.05s
Iteration 1438, Loss: 0.0267, Time: 0.05s
Iteration 1439, Loss: 0.0489, Time: 0.05s
Iteration 1440, Loss: 0.0567, Time: 0.05s
Iteration 1441, Loss: 0.0368, Time: 0.05s
Iteration 1442, Loss: 0.0485, Time: 0.04s
Iteration 1443, Loss: 0.0412, Time: 0.04s
Iteration 1444, Loss: 0.0451, Time: 0.04s
Iteration 1445, Loss: 0.0519, Time: 0.04s
Iteration 1446, Loss: 0.0432, Time: 0.04s
Iteration 1447, Loss: 0.0329, Time: 0.05s
Iteration 1448, Loss: 0.0419, Time: 0.04s
Iteration 1449, Loss: 0.0573, Time: 0.04s
Iteration 1450, Loss: 0.0364, Time: 0.04s
Iteration 1451, Loss: 0.0427, Time: 0.04s
Iteration 1452, Loss: 0.0479, Time: 0.05s
Iteration 1453, Loss: 0.0526, Time: 0.05s
Iteration 1454, Loss: 0.0353, Time: 0.05s
Iteration 1455, Loss: 0.0377, Time: 0.05s
Iteration 1456, Loss: 0.0398, Time: 0.04s
Iteration 1457, Loss: 0.0440, Time: 0.04s
Iteration 1458, Loss: 0.0481, Time: 0.04s
```

```
Iteration 1459, Loss: 0.0407, Time: 0.05s
Iteration 1460, Loss: 0.0444, Time: 0.05s
Iteration 1461, Loss: 0.0705, Time: 0.04s
Iteration 1462, Loss: 0.0364, Time: 0.04s
Iteration 1463, Loss: 0.0441, Time: 0.04s
Iteration 1464, Loss: 0.0445, Time: 0.04s
Iteration 1465, Loss: 0.0364, Time: 0.04s
Iteration 1466, Loss: 0.0454, Time: 0.04s
Iteration 1467, Loss: 0.0369, Time: 0.05s
Iteration 1468, Loss: 0.0373, Time: 0.05s
Iteration 1469, Loss: 0.0256, Time: 0.05s
Iteration 1470, Loss: 0.0297, Time: 0.05s
Iteration 1471, Loss: 0.0314, Time: 0.04s
Iteration 1472, Loss: 0.0384, Time: 0.05s
Iteration 1473, Loss: 0.0434, Time: 0.04s
Iteration 1474, Loss: 0.0498, Time: 0.04s
Iteration 1475, Loss: 0.0359, Time: 0.04s
Iteration 1476, Loss: 0.0411, Time: 0.04s
Iteration 1477, Loss: 0.0439, Time: 0.05s
Iteration 1478, Loss: 0.0433, Time: 0.05s
Iteration 1479, Loss: 0.0378, Time: 0.05s
Iteration 1480, Loss: 0.0379, Time: 0.05s
Iteration 1481, Loss: 0.0486, Time: 0.05s
Iteration 1482, Loss: 0.0393, Time: 0.04s
Iteration 1483, Loss: 0.0418, Time: 0.05s
Iteration 1484, Loss: 0.0409, Time: 0.05s
Iteration 1485, Loss: 0.0327, Time: 0.04s
Iteration 1486, Loss: 0.0503, Time: 0.04s
Iteration 1487, Loss: 0.0393, Time: 0.04s
Iteration 1488, Loss: 0.0461, Time: 0.05s
Iteration 1489, Loss: 0.0391, Time: 0.04s
Iteration 1490, Loss: 0.0362, Time: 0.04s
Iteration 1491, Loss: 0.0382, Time: 0.04s
Iteration 1492, Loss: 0.0417, Time: 0.06s
Iteration 1493, Loss: 0.0274, Time: 0.04s
Iteration 1494, Loss: 0.0385, Time: 0.04s
Iteration 1495, Loss: 0.0317, Time: 0.05s
Iteration 1496, Loss: 0.0601, Time: 0.04s
Iteration 1497, Loss: 0.0331, Time: 0.04s
Iteration 1498, Loss: 0.0385, Time: 0.05s
Iteration 1499, Loss: 0.0425, Time: 0.04s
Iteration 1500, Loss: 0.0494, Time: 0.04s
Iteration 1500, Loss: 0.0494, Time: 0.04s
Test Loss: 0.0550
Iteration 1501, Loss: 0.0630, Time: 0.04s
Iteration 1502, Loss: 0.0347, Time: 0.04s
Iteration 1503, Loss: 0.0391, Time: 0.05s
Iteration 1504, Loss: 0.0416, Time: 0.04s
```

```
Iteration 1505, Loss: 0.0418, Time: 0.05s
Iteration 1506, Loss: 0.0518, Time: 0.05s
Iteration 1507, Loss: 0.0495, Time: 0.05s
Iteration 1508, Loss: 0.0410, Time: 0.04s
Iteration 1509, Loss: 0.0327, Time: 0.04s
Iteration 1510, Loss: 0.0423, Time: 0.04s
Iteration 1511, Loss: 0.0615, Time: 0.05s
Iteration 1512, Loss: 0.0402, Time: 0.04s
Iteration 1513, Loss: 0.0538, Time: 0.05s
Iteration 1514, Loss: 0.0411, Time: 0.04s
Iteration 1515, Loss: 0.0531, Time: 0.04s
Iteration 1516, Loss: 0.0345, Time: 0.04s
Iteration 1517, Loss: 0.0513, Time: 0.05s
Iteration 1518, Loss: 0.0566, Time: 0.04s
Iteration 1519, Loss: 0.0549, Time: 0.04s
Iteration 1520, Loss: 0.0397, Time: 0.05s
Iteration 1521, Loss: 0.0416, Time: 0.04s
Iteration 1522, Loss: 0.0426, Time: 0.06s
Iteration 1523, Loss: 0.0394, Time: 0.05s
Iteration 1524, Loss: 0.0386, Time: 0.06s
Iteration 1525, Loss: 0.0367, Time: 0.06s
Iteration 1526, Loss: 0.0371, Time: 0.05s
Iteration 1527, Loss: 0.0393, Time: 0.05s
Iteration 1528, Loss: 0.0461, Time: 0.06s
Iteration 1529, Loss: 0.0430, Time: 0.05s
Iteration 1530, Loss: 0.0568, Time: 0.05s
Iteration 1531, Loss: 0.0217, Time: 0.04s
Iteration 1532, Loss: 0.0448, Time: 0.04s
Iteration 1533, Loss: 0.0708, Time: 0.05s
Iteration 1534, Loss: 0.0439, Time: 0.04s
Iteration 1535, Loss: 0.0510, Time: 0.05s
Iteration 1536, Loss: 0.0509, Time: 0.04s
Iteration 1537, Loss: 0.0455, Time: 0.05s
Iteration 1538, Loss: 0.0560, Time: 0.05s
Iteration 1539, Loss: 0.0471, Time: 0.06s
Iteration 1540, Loss: 0.0427, Time: 0.05s
Iteration 1541, Loss: 0.0581, Time: 0.04s
Iteration 1542, Loss: 0.0388, Time: 0.05s
Iteration 1543, Loss: 0.0510, Time: 0.04s
Iteration 1544, Loss: 0.0415, Time: 0.04s
Iteration 1545, Loss: 0.0337, Time: 0.05s
Iteration 1546, Loss: 0.0363, Time: 0.05s
Iteration 1547, Loss: 0.0509, Time: 0.05s
Iteration 1548, Loss: 0.0384, Time: 0.05s
Iteration 1549, Loss: 0.0307, Time: 0.06s
Iteration 1550, Loss: 0.0397, Time: 0.05s
Iteration 1551, Loss: 0.0527, Time: 0.05s
Iteration 1552, Loss: 0.0467, Time: 0.06s
```

```
Iteration 1553, Loss: 0.0246, Time: 0.04s
Iteration 1554, Loss: 0.0398, Time: 0.05s
Iteration 1555, Loss: 0.0385, Time: 0.04s
Iteration 1556, Loss: 0.0491, Time: 0.05s
Iteration 1557, Loss: 0.0244, Time: 0.05s
Iteration 1558, Loss: 0.0300, Time: 0.04s
Iteration 1559, Loss: 0.0586, Time: 0.04s
Iteration 1560, Loss: 0.0476, Time: 0.05s
Iteration 1561, Loss: 0.0198, Time: 0.04s
Iteration 1562, Loss: 0.0482, Time: 0.04s
Iteration 1563, Loss: 0.0385, Time: 0.05s
Iteration 1564, Loss: 0.0471, Time: 0.04s
Iteration 1565, Loss: 0.0533, Time: 0.04s
Iteration 1566, Loss: 0.0256, Time: 0.05s
Iteration 1567, Loss: 0.0422, Time: 0.05s
Iteration 1568, Loss: 0.0396, Time: 0.05s
Iteration 1569, Loss: 0.0522, Time: 0.05s
Iteration 1570, Loss: 0.0330, Time: 0.04s
Iteration 1571, Loss: 0.0416, Time: 0.05s
Iteration 1572, Loss: 0.0346, Time: 0.04s
Iteration 1573, Loss: 0.0438, Time: 0.05s
Iteration 1574, Loss: 0.0324, Time: 0.05s
Iteration 1575, Loss: 0.0365, Time: 0.04s
Iteration 1576, Loss: 0.0388, Time: 0.05s
Iteration 1577, Loss: 0.0474, Time: 0.04s
Iteration 1578, Loss: 0.0561, Time: 0.05s
Iteration 1579, Loss: 0.0484, Time: 0.05s
Iteration 1580, Loss: 0.0366, Time: 0.05s
Iteration 1581, Loss: 0.0330, Time: 0.04s
Iteration 1582, Loss: 0.0379, Time: 0.04s
Iteration 1583, Loss: 0.0358, Time: 0.04s
Iteration 1584, Loss: 0.0632, Time: 0.05s
Iteration 1585, Loss: 0.0216, Time: 0.05s
Iteration 1586, Loss: 0.0338, Time: 0.05s
Iteration 1587, Loss: 0.0459, Time: 0.04s
Iteration 1588, Loss: 0.0326, Time: 0.05s
Iteration 1589, Loss: 0.0374, Time: 0.05s
Iteration 1590, Loss: 0.0412, Time: 0.05s
Iteration 1591, Loss: 0.0432, Time: 0.05s
Iteration 1592, Loss: 0.0453, Time: 0.04s
Iteration 1593, Loss: 0.0336, Time: 0.05s
Iteration 1594, Loss: 0.0270, Time: 0.05s
Iteration 1595, Loss: 0.0491, Time: 0.04s
Iteration 1596, Loss: 0.0484, Time: 0.05s
Iteration 1597, Loss: 0.0424, Time: 0.04s
Iteration 1598, Loss: 0.0339, Time: 0.05s
Iteration 1599, Loss: 0.0466, Time: 0.04s
Iteration 1600, Loss: 0.0474, Time: 0.04s
```

```
Iteration 1600, Loss: 0.0474, Time: 0.04s
Test Loss: 0.0393
Iteration 1601, Loss: 0.0549, Time: 0.04s
Iteration 1602, Loss: 0.0297, Time: 0.05s
Iteration 1603, Loss: 0.0434, Time: 0.05s
Iteration 1604, Loss: 0.0537, Time: 0.05s
Iteration 1605, Loss: 0.0555, Time: 0.04s
Iteration 1606, Loss: 0.0449, Time: 0.04s
Iteration 1607, Loss: 0.0313, Time: 0.04s
Iteration 1608, Loss: 0.0351, Time: 0.04s
Iteration 1609, Loss: 0.0487, Time: 0.04s
Iteration 1610, Loss: 0.0462, Time: 0.04s
Iteration 1611, Loss: 0.0566, Time: 0.04s
Iteration 1612, Loss: 0.0627, Time: 0.04s
Iteration 1613, Loss: 0.0435, Time: 0.04s
Iteration 1614, Loss: 0.0336, Time: 0.05s
Iteration 1615, Loss: 0.0369, Time: 0.04s
Iteration 1616, Loss: 0.0297, Time: 0.05s
Iteration 1617, Loss: 0.0422, Time: 0.04s
Iteration 1618, Loss: 0.0524, Time: 0.04s
Iteration 1619, Loss: 0.0301, Time: 0.04s
Iteration 1620, Loss: 0.0436, Time: 0.04s
Iteration 1621, Loss: 0.0484, Time: 0.05s
Iteration 1622, Loss: 0.0481, Time: 0.05s
Iteration 1623, Loss: 0.0380, Time: 0.04s
Iteration 1624, Loss: 0.0357, Time: 0.04s
Iteration 1625, Loss: 0.0349, Time: 0.04s
Iteration 1626, Loss: 0.0473, Time: 0.05s
Iteration 1627, Loss: 0.0397, Time: 0.05s
Iteration 1628, Loss: 0.0302, Time: 0.04s
Iteration 1629, Loss: 0.0410, Time: 0.04s
Iteration 1630, Loss: 0.0605, Time: 0.04s
Iteration 1631, Loss: 0.0503, Time: 0.04s
Iteration 1632, Loss: 0.0356, Time: 0.04s
Iteration 1633, Loss: 0.0621, Time: 0.04s
Iteration 1634, Loss: 0.0334, Time: 0.04s
Iteration 1635, Loss: 0.0386, Time: 0.05s
Iteration 1636, Loss: 0.0421, Time: 0.04s
Iteration 1637, Loss: 0.0413, Time: 0.04s
Iteration 1638, Loss: 0.0344, Time: 0.04s
Iteration 1639, Loss: 0.0373, Time: 0.04s
Iteration 1640, Loss: 0.0285, Time: 0.05s
Iteration 1641, Loss: 0.0240, Time: 0.04s
Iteration 1642, Loss: 0.0376, Time: 0.05s
Iteration 1643, Loss: 0.0454, Time: 0.05s
Iteration 1644, Loss: 0.0498, Time: 0.04s
Iteration 1645, Loss: 0.0315, Time: 0.04s
Iteration 1646, Loss: 0.0399, Time: 0.04s
```

```
Iteration 1647, Loss: 0.0470, Time: 0.05s
Iteration 1648, Loss: 0.0350, Time: 0.04s
Iteration 1649, Loss: 0.0329, Time: 0.04s
Iteration 1650, Loss: 0.0471, Time: 0.04s
Iteration 1651, Loss: 0.0336, Time: 0.04s
Iteration 1652, Loss: 0.0223, Time: 0.05s
Iteration 1653, Loss: 0.0310, Time: 0.04s
Iteration 1654, Loss: 0.0444, Time: 0.05s
Iteration 1655, Loss: 0.0471, Time: 0.04s
Iteration 1656, Loss: 0.0469, Time: 0.05s
Iteration 1657, Loss: 0.0254, Time: 0.04s
Iteration 1658, Loss: 0.0565, Time: 0.05s
Iteration 1659, Loss: 0.0418, Time: 0.04s
Iteration 1660, Loss: 0.0445, Time: 0.05s
Iteration 1661, Loss: 0.0621, Time: 0.04s
Iteration 1662, Loss: 0.0409, Time: 0.04s
Iteration 1663, Loss: 0.0487, Time: 0.04s
Iteration 1664, Loss: 0.0508, Time: 0.05s
Iteration 1665, Loss: 0.0436, Time: 0.04s
Iteration 1666, Loss: 0.0486, Time: 0.04s
Iteration 1667, Loss: 0.0303, Time: 0.04s
Iteration 1668, Loss: 0.0377, Time: 0.04s
Iteration 1669, Loss: 0.0565, Time: 0.04s
Iteration 1670, Loss: 0.0411, Time: 0.04s
Iteration 1671, Loss: 0.0462, Time: 0.05s
Iteration 1672, Loss: 0.0396, Time: 0.04s
Iteration 1673, Loss: 0.0479, Time: 0.04s
Iteration 1674, Loss: 0.0478, Time: 0.04s
Iteration 1675, Loss: 0.0339, Time: 0.05s
Iteration 1676, Loss: 0.0344, Time: 0.04s
Iteration 1677, Loss: 0.0433, Time: 0.04s
Iteration 1678, Loss: 0.0383, Time: 0.05s
Iteration 1679, Loss: 0.0403, Time: 0.05s
Iteration 1680, Loss: 0.0386, Time: 0.05s
Iteration 1681, Loss: 0.0355, Time: 0.04s
Iteration 1682, Loss: 0.0421, Time: 0.04s
Iteration 1683, Loss: 0.0325, Time: 0.04s
Iteration 1684, Loss: 0.0360, Time: 0.04s
Iteration 1685, Loss: 0.0511, Time: 0.04s
Iteration 1686, Loss: 0.0284, Time: 0.05s
Iteration 1687, Loss: 0.0448, Time: 0.04s
Iteration 1688, Loss: 0.0354, Time: 0.04s
Iteration 1689, Loss: 0.0483, Time: 0.05s
Iteration 1690, Loss: 0.0354, Time: 0.04s
Iteration 1691, Loss: 0.0417, Time: 0.04s
Iteration 1692, Loss: 0.0388, Time: 0.04s
Iteration 1693, Loss: 0.0365, Time: 0.04s
Iteration 1694, Loss: 0.0518, Time: 0.05s
```

```
Iteration 1695, Loss: 0.0516, Time: 0.05s
Iteration 1696, Loss: 0.0292, Time: 0.04s
Iteration 1697, Loss: 0.0376, Time: 0.05s
Iteration 1698, Loss: 0.0300, Time: 0.05s
Iteration 1699, Loss: 0.0586, Time: 0.04s
Iteration 1700, Loss: 0.0391, Time: 0.04s
Iteration 1700, Loss: 0.0391, Time: 0.04s
Test Loss: 0.0426
Iteration 1701, Loss: 0.0398, Time: 0.04s
Iteration 1702, Loss: 0.0434, Time: 0.05s
Iteration 1703, Loss: 0.0380, Time: 0.05s
Iteration 1704, Loss: 0.0325, Time: 0.04s
Iteration 1705, Loss: 0.0372, Time: 0.05s
Iteration 1706, Loss: 0.0382, Time: 0.05s
Iteration 1707, Loss: 0.0468, Time: 0.04s
Iteration 1708, Loss: 0.0307, Time: 0.05s
Iteration 1709, Loss: 0.0228, Time: 0.05s
Iteration 1710, Loss: 0.0440, Time: 0.05s
Iteration 1711, Loss: 0.0195, Time: 0.04s
Iteration 1712, Loss: 0.0547, Time: 0.05s
Iteration 1713, Loss: 0.0491, Time: 0.04s
Iteration 1714, Loss: 0.0410, Time: 0.04s
Iteration 1715, Loss: 0.0545, Time: 0.04s
Iteration 1716, Loss: 0.0271, Time: 0.04s
Iteration 1717, Loss: 0.0419, Time: 0.04s
Iteration 1718, Loss: 0.0431, Time: 0.04s
Iteration 1719, Loss: 0.0451, Time: 0.05s
Iteration 1720, Loss: 0.0201, Time: 0.04s
Iteration 1721, Loss: 0.0333, Time: 0.04s
Iteration 1722, Loss: 0.0342, Time: 0.04s
Iteration 1723, Loss: 0.0325, Time: 0.04s
Iteration 1724, Loss: 0.0641, Time: 0.05s
Iteration 1725, Loss: 0.0438, Time: 0.05s
Iteration 1726, Loss: 0.0479, Time: 0.04s
Iteration 1727, Loss: 0.0401, Time: 0.04s
Iteration 1728, Loss: 0.0310, Time: 0.04s
Iteration 1729, Loss: 0.0272, Time: 0.04s
Iteration 1730, Loss: 0.0351, Time: 0.05s
Iteration 1731, Loss: 0.0380, Time: 0.04s
Iteration 1732, Loss: 0.0509, Time: 0.04s
Iteration 1733, Loss: 0.0333, Time: 0.05s
Iteration 1734, Loss: 0.0514, Time: 0.04s
Iteration 1735, Loss: 0.0306, Time: 0.05s
Iteration 1736, Loss: 0.0422, Time: 0.04s
Iteration 1737, Loss: 0.0632, Time: 0.04s
Iteration 1738, Loss: 0.0432, Time: 0.05s
Iteration 1739, Loss: 0.0327, Time: 0.04s
Iteration 1740, Loss: 0.0435, Time: 0.05s
```

```
Iteration 1741, Loss: 0.0246, Time: 0.05s
Iteration 1742, Loss: 0.0378, Time: 0.07s
Iteration 1743, Loss: 0.0395, Time: 0.06s
Iteration 1744, Loss: 0.0321, Time: 0.07s
Iteration 1745, Loss: 0.0346, Time: 0.05s
Iteration 1746, Loss: 0.0310, Time: 0.07s
Iteration 1747, Loss: 0.0232, Time: 0.06s
Iteration 1748, Loss: 0.0395, Time: 0.04s
Iteration 1749, Loss: 0.0343, Time: 0.05s
Iteration 1750, Loss: 0.0269, Time: 0.04s
Iteration 1751, Loss: 0.0367, Time: 0.05s
Iteration 1752, Loss: 0.0304, Time: 0.04s
Iteration 1753, Loss: 0.0566, Time: 0.04s
Iteration 1754, Loss: 0.0402, Time: 0.05s
Iteration 1755, Loss: 0.0278, Time: 0.04s
Iteration 1756, Loss: 0.0355, Time: 0.05s
Iteration 1757, Loss: 0.0471, Time: 0.05s
Iteration 1758, Loss: 0.0456, Time: 0.04s
Iteration 1759, Loss: 0.0347, Time: 0.05s
Iteration 1760, Loss: 0.0531, Time: 0.05s
Iteration 1761, Loss: 0.0372, Time: 0.05s
Iteration 1762, Loss: 0.0244, Time: 0.04s
Iteration 1763, Loss: 0.0417, Time: 0.04s
Iteration 1764, Loss: 0.0363, Time: 0.05s
Iteration 1765, Loss: 0.0611, Time: 0.05s
Iteration 1766, Loss: 0.0245, Time: 0.05s
Iteration 1767, Loss: 0.0356, Time: 0.04s
Iteration 1768, Loss: 0.0344, Time: 0.05s
Iteration 1769, Loss: 0.0269, Time: 0.04s
Iteration 1770, Loss: 0.0259, Time: 0.05s
Iteration 1771, Loss: 0.0424, Time: 0.05s
Iteration 1772, Loss: 0.0279, Time: 0.05s
Iteration 1773, Loss: 0.0441, Time: 0.04s
Iteration 1774, Loss: 0.0346, Time: 0.05s
Iteration 1775, Loss: 0.0281, Time: 0.04s
Iteration 1776, Loss: 0.0242, Time: 0.04s
Iteration 1777, Loss: 0.0339, Time: 0.05s
Iteration 1778, Loss: 0.0267, Time: 0.04s
Iteration 1779, Loss: 0.0271, Time: 0.05s
Iteration 1780, Loss: 0.0417, Time: 0.04s
Iteration 1781, Loss: 0.0384, Time: 0.05s
Iteration 1782, Loss: 0.0318, Time: 0.04s
Iteration 1783, Loss: 0.0451, Time: 0.04s
Iteration 1784, Loss: 0.0356, Time: 0.05s
Iteration 1785, Loss: 0.0308, Time: 0.04s
Iteration 1786, Loss: 0.0415, Time: 0.04s
Iteration 1787, Loss: 0.0451, Time: 0.05s
Iteration 1788, Loss: 0.0368, Time: 0.04s
```

```
Iteration 1789, Loss: 0.0424, Time: 0.04s
Iteration 1790, Loss: 0.0258, Time: 0.05s
Iteration 1791, Loss: 0.0398, Time: 0.05s
Iteration 1792, Loss: 0.0407, Time: 0.05s
Iteration 1793, Loss: 0.0300, Time: 0.04s
Iteration 1794, Loss: 0.0379, Time: 0.04s
Iteration 1795, Loss: 0.0396, Time: 0.05s
Iteration 1796, Loss: 0.0436, Time: 0.04s
Iteration 1797, Loss: 0.0535, Time: 0.05s
Iteration 1798, Loss: 0.0321, Time: 0.05s
Iteration 1799, Loss: 0.0586, Time: 0.05s
Iteration 1800, Loss: 0.0373, Time: 0.04s
Iteration 1800, Loss: 0.0373, Time: 0.04s
Test Loss: 0.0299
Iteration 1801, Loss: 0.0343, Time: 0.05s
Iteration 1802, Loss: 0.0395, Time: 0.04s
Iteration 1803, Loss: 0.0379, Time: 0.05s
Iteration 1804, Loss: 0.0384, Time: 0.04s
Iteration 1805, Loss: 0.0453, Time: 0.05s
Iteration 1806, Loss: 0.0285, Time: 0.05s
Iteration 1807, Loss: 0.0255, Time: 0.04s
Iteration 1808, Loss: 0.0272, Time: 0.05s
Iteration 1809, Loss: 0.0237, Time: 0.05s
Iteration 1810, Loss: 0.0439, Time: 0.05s
Iteration 1811, Loss: 0.0374, Time: 0.05s
Iteration 1812, Loss: 0.0343, Time: 0.04s
Iteration 1813, Loss: 0.0462, Time: 0.04s
Iteration 1814, Loss: 0.0427, Time: 0.04s
Iteration 1815, Loss: 0.0342, Time: 0.05s
Iteration 1816, Loss: 0.0374, Time: 0.04s
Iteration 1817, Loss: 0.0609, Time: 0.04s
Iteration 1818, Loss: 0.0289, Time: 0.05s
Iteration 1819, Loss: 0.0338, Time: 0.04s
Iteration 1820, Loss: 0.0327, Time: 0.05s
Iteration 1821, Loss: 0.0458, Time: 0.05s
Iteration 1822, Loss: 0.0326, Time: 0.05s
Iteration 1823, Loss: 0.0227, Time: 0.04s
Iteration 1824, Loss: 0.0438, Time: 0.05s
Iteration 1825, Loss: 0.0639, Time: 0.05s
Iteration 1826, Loss: 0.0300, Time: 0.04s
Iteration 1827, Loss: 0.0315, Time: 0.05s
Iteration 1828, Loss: 0.0364, Time: 0.04s
Iteration 1829, Loss: 0.0408, Time: 0.04s
Iteration 1830, Loss: 0.0353, Time: 0.04s
Iteration 1831, Loss: 0.0379, Time: 0.05s
Iteration 1832, Loss: 0.0440, Time: 0.04s
Iteration 1833, Loss: 0.0445, Time: 0.05s
Iteration 1834, Loss: 0.0255, Time: 0.05s
```

```
Iteration 1835, Loss: 0.0505, Time: 0.05s
Iteration 1836, Loss: 0.0294, Time: 0.04s
Iteration 1837, Loss: 0.0278, Time: 0.04s
Iteration 1838, Loss: 0.0473, Time: 0.04s
Iteration 1839, Loss: 0.0294, Time: 0.05s
Iteration 1840, Loss: 0.0465, Time: 0.04s
Iteration 1841, Loss: 0.0351, Time: 0.05s
Iteration 1842, Loss: 0.0327, Time: 0.05s
Iteration 1843, Loss: 0.0459, Time: 0.04s
Iteration 1844, Loss: 0.0514, Time: 0.04s
Iteration 1845, Loss: 0.0332, Time: 0.04s
Iteration 1846, Loss: 0.0279, Time: 0.05s
Iteration 1847, Loss: 0.0399, Time: 0.04s
Iteration 1848, Loss: 0.0250, Time: 0.05s
Iteration 1849, Loss: 0.0199, Time: 0.04s
Iteration 1850, Loss: 0.0189, Time: 0.05s
Iteration 1851, Loss: 0.0484, Time: 0.04s
Iteration 1852, Loss: 0.0496, Time: 0.04s
Iteration 1853, Loss: 0.0363, Time: 0.05s
Iteration 1854, Loss: 0.0348, Time: 0.04s
Iteration 1855, Loss: 0.0356, Time: 0.04s
Iteration 1856, Loss: 0.0488, Time: 0.05s
Iteration 1857, Loss: 0.0429, Time: 0.04s
Iteration 1858, Loss: 0.0517, Time: 0.05s
Iteration 1859, Loss: 0.0375, Time: 0.04s
Iteration 1860, Loss: 0.0403, Time: 0.04s
Iteration 1861, Loss: 0.0437, Time: 0.04s
Iteration 1862, Loss: 0.0434, Time: 0.04s
Iteration 1863, Loss: 0.0323, Time: 0.04s
Iteration 1864, Loss: 0.0377, Time: 0.04s
Iteration 1865, Loss: 0.0164, Time: 0.05s
Iteration 1866, Loss: 0.0188, Time: 0.04s
Iteration 1867, Loss: 0.0347, Time: 0.04s
Iteration 1868, Loss: 0.0228, Time: 0.04s
Iteration 1869, Loss: 0.0456, Time: 0.05s
Iteration 1870, Loss: 0.0238, Time: 0.04s
Iteration 1871, Loss: 0.0432, Time: 0.05s
Iteration 1872, Loss: 0.0565, Time: 0.05s
Iteration 1873, Loss: 0.0285, Time: 0.05s
Iteration 1874, Loss: 0.0433, Time: 0.05s
Iteration 1875, Loss: 0.0493, Time: 0.04s
Iteration 1876, Loss: 0.0393, Time: 0.04s
Iteration 1877, Loss: 0.0370, Time: 0.04s
Iteration 1878, Loss: 0.0462, Time: 0.04s
Iteration 1879, Loss: 0.0445, Time: 0.04s
Iteration 1880, Loss: 0.0344, Time: 0.05s
Iteration 1881, Loss: 0.0342, Time: 0.05s
Iteration 1882, Loss: 0.0388, Time: 0.04s
```

```
Iteration 1883, Loss: 0.0190, Time: 0.04s
Iteration 1884, Loss: 0.0501, Time: 0.04s
Iteration 1885, Loss: 0.0281, Time: 0.04s
Iteration 1886, Loss: 0.0238, Time: 0.05s
Iteration 1887, Loss: 0.0445, Time: 0.04s
Iteration 1888, Loss: 0.0279, Time: 0.04s
Iteration 1889, Loss: 0.0181, Time: 0.05s
Iteration 1890, Loss: 0.0304, Time: 0.05s
Iteration 1891, Loss: 0.0279, Time: 0.05s
Iteration 1892, Loss: 0.0463, Time: 0.04s
Iteration 1893, Loss: 0.0323, Time: 0.05s
Iteration 1894, Loss: 0.0241, Time: 0.05s
Iteration 1895, Loss: 0.0403, Time: 0.05s
Iteration 1896, Loss: 0.0329, Time: 0.04s
Iteration 1897, Loss: 0.0428, Time: 0.04s
Iteration 1898, Loss: 0.0291, Time: 0.04s
Iteration 1899, Loss: 0.0244, Time: 0.04s
Iteration 1900, Loss: 0.0385, Time: 0.05s
Iteration 1900, Loss: 0.0385, Time: 0.05s
Test Loss: 0.0487
Iteration 1901, Loss: 0.0498, Time: 0.04s
Iteration 1902, Loss: 0.0245, Time: 0.04s
Iteration 1903, Loss: 0.0273, Time: 0.04s
Iteration 1904, Loss: 0.0378, Time: 0.04s
Iteration 1905, Loss: 0.0429, Time: 0.04s
Iteration 1906, Loss: 0.0314, Time: 0.04s
Iteration 1907, Loss: 0.0307, Time: 0.04s
Iteration 1908, Loss: 0.0345, Time: 0.04s
Iteration 1909, Loss: 0.0450, Time: 0.04s
Iteration 1910, Loss: 0.0336, Time: 0.05s
Iteration 1911, Loss: 0.0408, Time: 0.04s
Iteration 1912, Loss: 0.0364, Time: 0.04s
Iteration 1913, Loss: 0.0391, Time: 0.04s
Iteration 1914, Loss: 0.0289, Time: 0.05s
Iteration 1915, Loss: 0.0309, Time: 0.04s
Iteration 1916, Loss: 0.0358, Time: 0.04s
Iteration 1917, Loss: 0.0352, Time: 0.04s
Iteration 1918, Loss: 0.0387, Time: 0.04s
Iteration 1919, Loss: 0.0394, Time: 0.04s
Iteration 1920, Loss: 0.0324, Time: 0.05s
Iteration 1921, Loss: 0.0274, Time: 0.05s
Iteration 1922, Loss: 0.0311, Time: 0.04s
Iteration 1923, Loss: 0.0483, Time: 0.04s
Iteration 1924, Loss: 0.0376, Time: 0.04s
Iteration 1925, Loss: 0.0241, Time: 0.04s
Iteration 1926, Loss: 0.0429, Time: 0.04s
Iteration 1927, Loss: 0.0280, Time: 0.04s
Iteration 1928, Loss: 0.0293, Time: 0.04s
```

```
Iteration 1929, Loss: 0.0320, Time: 0.05s
Iteration 1930, Loss: 0.0355, Time: 0.05s
Iteration 1931, Loss: 0.0386, Time: 0.05s
Iteration 1932, Loss: 0.0476, Time: 0.04s
Iteration 1933, Loss: 0.0327, Time: 0.05s
Iteration 1934, Loss: 0.0348, Time: 0.04s
Iteration 1935, Loss: 0.0400, Time: 0.04s
Iteration 1936, Loss: 0.0347, Time: 0.05s
Iteration 1937, Loss: 0.0231, Time: 0.05s
Iteration 1938, Loss: 0.0308, Time: 0.04s
Iteration 1939, Loss: 0.0587, Time: 0.04s
Iteration 1940, Loss: 0.0314, Time: 0.05s
Iteration 1941, Loss: 0.0526, Time: 0.05s
Iteration 1942, Loss: 0.0302, Time: 0.05s
Iteration 1943, Loss: 0.0378, Time: 0.05s
Iteration 1944, Loss: 0.0451, Time: 0.04s
Iteration 1945, Loss: 0.0410, Time: 0.05s
Iteration 1946, Loss: 0.0290, Time: 0.04s
Iteration 1947, Loss: 0.0272, Time: 0.04s
Iteration 1948, Loss: 0.0467, Time: 0.04s
Iteration 1949, Loss: 0.0279, Time: 0.04s
Iteration 1950, Loss: 0.0332, Time: 0.05s
Iteration 1951, Loss: 0.0594, Time: 0.05s
Iteration 1952, Loss: 0.0363, Time: 0.05s
Iteration 1953, Loss: 0.0474, Time: 0.05s
Iteration 1954, Loss: 0.0445, Time: 0.04s
Iteration 1955, Loss: 0.0318, Time: 0.04s
Iteration 1956, Loss: 0.0525, Time: 0.04s
Iteration 1957, Loss: 0.0234, Time: 0.05s
Iteration 1958, Loss: 0.0401, Time: 0.05s
Iteration 1959, Loss: 0.0266, Time: 0.05s
Iteration 1960, Loss: 0.0342, Time: 0.06s
Iteration 1961, Loss: 0.0490, Time: 0.05s
Iteration 1962, Loss: 0.0262, Time: 0.06s
Iteration 1963, Loss: 0.0257, Time: 0.05s
Iteration 1964, Loss: 0.0263, Time: 0.08s
Iteration 1965, Loss: 0.0438, Time: 0.05s
Iteration 1966, Loss: 0.0451, Time: 0.05s
Iteration 1967, Loss: 0.0366, Time: 0.04s
Iteration 1968, Loss: 0.0482, Time: -0.10s
Iteration 1969, Loss: 0.0393, Time: 0.05s
Iteration 1970, Loss: 0.0403, Time: 0.05s
Iteration 1971, Loss: 0.0250, Time: 0.05s
Iteration 1972, Loss: 0.0212, Time: 0.05s
Iteration 1973, Loss: 0.0320, Time: 0.05s
Iteration 1974, Loss: 0.0331, Time: 0.05s
Iteration 1975, Loss: 0.0429, Time: 0.04s
Iteration 1976, Loss: 0.0562, Time: 0.04s
```

```
Iteration 1977, Loss: 0.0354, Time: 0.05s
Iteration 1978, Loss: 0.0350, Time: 0.04s
Iteration 1979, Loss: 0.0258, Time: 0.05s
Iteration 1980, Loss: 0.0439, Time: 0.04s
Iteration 1981, Loss: 0.0418, Time: 0.04s
Iteration 1982, Loss: 0.0362, Time: 0.05s
Iteration 1983, Loss: 0.0514, Time: 0.04s
Iteration 1984, Loss: 0.0357, Time: 0.05s
Iteration 1985, Loss: 0.0269, Time: 0.04s
Iteration 1986, Loss: 0.0291, Time: 0.04s
Iteration 1987, Loss: 0.0294, Time: 0.04s
Iteration 1988, Loss: 0.0305, Time: 0.04s
Iteration 1989, Loss: 0.0554, Time: 0.04s
Iteration 1990, Loss: 0.0498, Time: 0.04s
Iteration 1991, Loss: 0.0330, Time: 0.04s
Iteration 1992, Loss: 0.0328, Time: 0.04s
Iteration 1993, Loss: 0.0341, Time: 0.04s
Iteration 1994, Loss: 0.0269, Time: 0.04s
Iteration 1995, Loss: 0.0339, Time: 0.05s
Iteration 1996, Loss: 0.0384, Time: 0.04s
Iteration 1997, Loss: 0.0293, Time: 0.04s
Iteration 1998, Loss: 0.0269, Time: 0.04s
Iteration 1999, Loss: 0.0233, Time: 0.04s
Iteration 2000, Loss: 0.0256, Time: 0.05s
Iteration 2000, Loss: 0.0256, Time: 0.05s
Test Loss: 0.0457
Iteration 2001, Loss: 0.0394, Time: 0.04s
Iteration 2002, Loss: 0.0373, Time: 0.05s
Iteration 2003, Loss: 0.0460, Time: 0.04s
Iteration 2004, Loss: 0.0390, Time: 0.05s
Iteration 2005, Loss: 0.0217, Time: 0.04s
Iteration 2006, Loss: 0.0475, Time: 0.05s
Iteration 2007, Loss: 0.0370, Time: 0.04s
Iteration 2008, Loss: 0.0420, Time: 0.04s
Iteration 2009, Loss: 0.0272, Time: 0.05s
Iteration 2010, Loss: 0.0450, Time: 0.04s
Iteration 2011, Loss: 0.0467, Time: 0.05s
Iteration 2012, Loss: 0.0530, Time: 0.04s
Iteration 2013, Loss: 0.0393, Time: 0.05s
Iteration 2014, Loss: 0.0505, Time: 0.04s
Iteration 2015, Loss: 0.0301, Time: 0.05s
Iteration 2016, Loss: 0.0497, Time: 0.05s
Iteration 2017, Loss: 0.0462, Time: 0.05s
Iteration 2018, Loss: 0.0366, Time: 0.05s
Iteration 2019, Loss: 0.0423, Time: 0.05s
Iteration 2020, Loss: 0.0455, Time: 0.04s
Iteration 2021, Loss: 0.0321, Time: 0.05s
Iteration 2022, Loss: 0.0395, Time: 0.04s
```

```
Iteration 2023, Loss: 0.0336, Time: 0.05s
Iteration 2024, Loss: 0.0364, Time: 0.04s
Iteration 2025, Loss: 0.0235, Time: 0.04s
Iteration 2026, Loss: 0.0273, Time: 0.04s
Iteration 2027, Loss: 0.0245, Time: 0.05s
Iteration 2028, Loss: 0.0380, Time: 0.05s
Iteration 2029, Loss: 0.0251, Time: 0.06s
Iteration 2030, Loss: 0.0492, Time: 0.07s
Iteration 2031, Loss: 0.0481, Time: 0.06s
Iteration 2032, Loss: 0.0411, Time: 0.05s
Iteration 2033, Loss: 0.0360, Time: 0.05s
Iteration 2034, Loss: 0.0394, Time: 0.04s
Iteration 2035, Loss: 0.0444, Time: 0.04s
Iteration 2036, Loss: 0.0325, Time: 0.06s
Iteration 2037, Loss: 0.0330, Time: 0.05s
Iteration 2038, Loss: 0.0453, Time: 0.06s
Iteration 2039, Loss: 0.0467, Time: 0.04s
Iteration 2040, Loss: 0.0276, Time: 0.05s
Iteration 2041, Loss: 0.0247, Time: 0.05s
Iteration 2042, Loss: 0.0353, Time: 0.06s
Iteration 2043, Loss: 0.0303, Time: 0.05s
Iteration 2044, Loss: 0.0524, Time: 0.05s
Iteration 2045, Loss: 0.0207, Time: 0.04s
Iteration 2046, Loss: 0.0359, Time: 0.05s
Iteration 2047, Loss: 0.0366, Time: 0.04s
Iteration 2048, Loss: 0.0308, Time: 0.06s
Iteration 2049, Loss: 0.0246, Time: 0.05s
Iteration 2050, Loss: 0.0270, Time: 0.05s
Iteration 2051, Loss: 0.0396, Time: 0.06s
Iteration 2052, Loss: 0.0276, Time: 0.06s
Iteration 2053, Loss: 0.0469, Time: 0.06s
Iteration 2054, Loss: 0.0255, Time: 0.06s
Iteration 2055, Loss: 0.0264, Time: 0.06s
Iteration 2056, Loss: 0.0247, Time: 0.05s
Iteration 2057, Loss: 0.0566, Time: 0.04s
Iteration 2058, Loss: 0.0452, Time: 0.06s
Iteration 2059, Loss: 0.0211, Time: 0.05s
Iteration 2060, Loss: 0.0338, Time: 0.05s
Iteration 2061, Loss: 0.0277, Time: 0.04s
Iteration 2062, Loss: 0.0367, Time: 0.05s
Iteration 2063, Loss: 0.0612, Time: 0.05s
Iteration 2064, Loss: 0.0322, Time: 0.04s
Iteration 2065, Loss: 0.0330, Time: 0.04s
Iteration 2066, Loss: 0.0239, Time: 0.05s
Iteration 2067, Loss: 0.0341, Time: 0.05s
Iteration 2068, Loss: 0.0362, Time: 0.04s
Iteration 2069, Loss: 0.0383, Time: 0.04s
Iteration 2070, Loss: 0.0220, Time: 0.04s
```

```
Iteration 2071, Loss: 0.0333, Time: 0.05s
Iteration 2072, Loss: 0.0431, Time: 0.05s
Iteration 2073, Loss: 0.0350, Time: 0.05s
Iteration 2074, Loss: 0.0436, Time: 0.04s
Iteration 2075, Loss: 0.0340, Time: 0.05s
Iteration 2076, Loss: 0.0313, Time: 0.04s
Iteration 2077, Loss: 0.0414, Time: 0.04s
Iteration 2078, Loss: 0.0264, Time: 0.04s
Iteration 2079, Loss: 0.0343, Time: 0.05s
Iteration 2080, Loss: 0.0375, Time: 0.04s
Iteration 2081, Loss: 0.0297, Time: 0.05s
Iteration 2082, Loss: 0.0311, Time: 0.04s
Iteration 2083, Loss: 0.0372, Time: 0.05s
Iteration 2084, Loss: 0.0304, Time: 0.05s
Iteration 2085, Loss: 0.0295, Time: 0.04s
Iteration 2086, Loss: 0.0377, Time: 0.04s
Iteration 2087, Loss: 0.0477, Time: 0.05s
Iteration 2088, Loss: 0.0281, Time: 0.05s
Iteration 2089, Loss: 0.0266, Time: 0.06s
Iteration 2090, Loss: 0.0403, Time: 0.04s
Iteration 2091, Loss: 0.0444, Time: 0.04s
Iteration 2092, Loss: 0.0407, Time: 0.04s
Iteration 2093, Loss: 0.0207, Time: 0.05s
Iteration 2094, Loss: 0.0301, Time: 0.04s
Iteration 2095, Loss: 0.0343, Time: 0.04s
Iteration 2096, Loss: 0.0466, Time: 0.05s
Iteration 2097, Loss: 0.0251, Time: 0.04s
Iteration 2098, Loss: 0.0425, Time: 0.05s
Iteration 2099, Loss: 0.0187, Time: 0.04s
Iteration 2100, Loss: 0.0456, Time: 0.05s
Iteration 2100, Loss: 0.0456, Time: 0.05s
Test Loss: 0.0562
Iteration 2101, Loss: 0.0433, Time: 0.04s
Iteration 2102, Loss: 0.0171, Time: 0.06s
Iteration 2103, Loss: 0.0359, Time: 0.05s
Iteration 2104, Loss: 0.0227, Time: 0.04s
Iteration 2105, Loss: 0.0303, Time: 0.04s
Iteration 2106, Loss: 0.0371, Time: 0.05s
Iteration 2107, Loss: 0.0340, Time: 0.05s
Iteration 2108, Loss: 0.0234, Time: 0.05s
Iteration 2109, Loss: 0.0380, Time: 0.05s
Iteration 2110, Loss: 0.0349, Time: 0.05s
Iteration 2111, Loss: 0.0352, Time: 0.05s
Iteration 2112, Loss: 0.0422, Time: 0.05s
Iteration 2113, Loss: 0.0434, Time: 0.05s
Iteration 2114, Loss: 0.0304, Time: 0.05s
Iteration 2115, Loss: 0.0279, Time: 0.04s
Iteration 2116, Loss: 0.0369, Time: 0.05s
```

```
Iteration 2117, Loss: 0.0315, Time: 0.04s
Iteration 2118, Loss: 0.0252, Time: 0.05s
Iteration 2119, Loss: 0.0308, Time: 0.05s
Iteration 2120, Loss: 0.0400, Time: 0.04s
Iteration 2121, Loss: 0.0329, Time: 0.04s
Iteration 2122, Loss: 0.0324, Time: 0.04s
Iteration 2123, Loss: 0.0405, Time: 0.04s
Iteration 2124, Loss: 0.0370, Time: 0.04s
Iteration 2125, Loss: 0.0240, Time: 0.04s
Iteration 2126, Loss: 0.0430, Time: 0.05s
Iteration 2127, Loss: 0.0384, Time: 0.04s
Iteration 2128, Loss: 0.0403, Time: 0.04s
Iteration 2129, Loss: 0.0294, Time: 0.04s
Iteration 2130, Loss: 0.0390, Time: 0.05s
Iteration 2131, Loss: 0.0325, Time: 0.04s
Iteration 2132, Loss: 0.0387, Time: 0.05s
Iteration 2133, Loss: 0.0350, Time: 0.05s
Iteration 2134, Loss: 0.0274, Time: 0.04s
Iteration 2135, Loss: 0.0373, Time: 0.04s
Iteration 2136, Loss: 0.0477, Time: 0.04s
Iteration 2137, Loss: 0.0321, Time: 0.04s
Iteration 2138, Loss: 0.0459, Time: 0.04s
Iteration 2139, Loss: 0.0282, Time: 0.04s
Iteration 2140, Loss: 0.0360, Time: 0.04s
Iteration 2141, Loss: 0.0307, Time: 0.04s
Iteration 2142, Loss: 0.0192, Time: 0.04s
Iteration 2143, Loss: 0.0408, Time: 0.04s
Iteration 2144, Loss: 0.0222, Time: 0.04s
Iteration 2145, Loss: 0.0262, Time: 0.04s
Iteration 2146, Loss: 0.0479, Time: 0.04s
Iteration 2147, Loss: 0.0326, Time: 0.04s
Iteration 2148, Loss: 0.0410, Time: 0.04s
Iteration 2149, Loss: 0.0450, Time: 0.04s
Iteration 2150, Loss: 0.0368, Time: 0.04s
Iteration 2151, Loss: 0.0369, Time: 0.04s
Iteration 2152, Loss: 0.0339, Time: 0.04s
Iteration 2153, Loss: 0.0439, Time: 0.05s
Iteration 2154, Loss: 0.0378, Time: 0.04s
Iteration 2155, Loss: 0.0489, Time: 0.04s
Iteration 2156, Loss: 0.0467, Time: 0.04s
Iteration 2157, Loss: 0.0333, Time: 0.05s
Iteration 2158, Loss: 0.0500, Time: 0.05s
Iteration 2159, Loss: 0.0439, Time: 0.04s
Iteration 2160, Loss: 0.0295, Time: 0.04s
Iteration 2161, Loss: 0.0271, Time: 0.04s
Iteration 2162, Loss: 0.0386, Time: 0.04s
Iteration 2163, Loss: 0.0321, Time: 0.05s
Iteration 2164, Loss: 0.0270, Time: 0.04s
```

```
Iteration 2165, Loss: 0.0260, Time: 0.04s
Iteration 2166, Loss: 0.0311, Time: 0.05s
Iteration 2167, Loss: 0.0470, Time: 0.05s
Iteration 2168, Loss: 0.0284, Time: 0.04s
Iteration 2169, Loss: 0.0379, Time: 0.05s
Iteration 2170, Loss: 0.0381, Time: 0.05s
Iteration 2171, Loss: 0.0551, Time: 0.04s
Iteration 2172, Loss: 0.0461, Time: 0.05s
Iteration 2173, Loss: 0.0386, Time: 0.06s
Iteration 2174, Loss: 0.0372, Time: 0.05s
Iteration 2175, Loss: 0.0338, Time: 0.05s
Iteration 2176, Loss: 0.0424, Time: 0.04s
Iteration 2177, Loss: 0.0279, Time: 0.05s
Iteration 2178, Loss: 0.0288, Time: 0.05s
Iteration 2179, Loss: 0.0412, Time: 0.06s
Iteration 2180, Loss: 0.0428, Time: 0.04s
Iteration 2181, Loss: 0.0181, Time: 0.05s
Iteration 2182, Loss: 0.0279, Time: 0.04s
Iteration 2183, Loss: 0.0233, Time: 0.04s
Iteration 2184, Loss: 0.0427, Time: 0.04s
Iteration 2185, Loss: 0.0304, Time: 0.04s
Iteration 2186, Loss: 0.0211, Time: 0.04s
Iteration 2187, Loss: 0.0194, Time: 0.04s
Iteration 2188, Loss: 0.0291, Time: 0.05s
Iteration 2189, Loss: 0.0343, Time: 0.04s
Iteration 2190, Loss: 0.0399, Time: 0.06s
Iteration 2191, Loss: 0.0306, Time: 0.05s
Iteration 2192, Loss: 0.0314, Time: 0.05s
Iteration 2193, Loss: 0.0241, Time: 0.04s
Iteration 2194, Loss: 0.0432, Time: 0.04s
Iteration 2195, Loss: 0.0235, Time: 0.05s
Iteration 2196, Loss: 0.0327, Time: 0.04s
Iteration 2197, Loss: 0.0356, Time: 0.05s
Iteration 2198, Loss: 0.0340, Time: 0.04s
Iteration 2199, Loss: 0.0375, Time: 0.04s
Iteration 2200, Loss: 0.0444, Time: 0.04s
Iteration 2200, Loss: 0.0444, Time: 0.04s
Test Loss: 0.0410
Iteration 2201, Loss: 0.0261, Time: 0.04s
Iteration 2202, Loss: 0.0265, Time: 0.05s
Iteration 2203, Loss: 0.0262, Time: 0.04s
Iteration 2204, Loss: 0.0393, Time: 0.04s
Iteration 2205, Loss: 0.0361, Time: 0.05s
Iteration 2206, Loss: 0.0391, Time: 0.05s
Iteration 2207, Loss: 0.0282, Time: 0.05s
Iteration 2208, Loss: 0.0262, Time: 0.04s
Iteration 2209, Loss: 0.0286, Time: 0.04s
Iteration 2210, Loss: 0.0277, Time: 0.04s
```

```
Iteration 2211, Loss: 0.0289, Time: 0.04s
Iteration 2212, Loss: 0.0508, Time: 0.05s
Iteration 2213, Loss: 0.0278, Time: 0.05s
Iteration 2214, Loss: 0.0307, Time: 0.05s
Iteration 2215, Loss: 0.0320, Time: 0.05s
Iteration 2216, Loss: 0.0270, Time: 0.04s
Iteration 2217, Loss: 0.0283, Time: 0.05s
Iteration 2218, Loss: 0.0452, Time: 0.05s
Iteration 2219, Loss: 0.0350, Time: 0.06s
Iteration 2220, Loss: 0.0494, Time: 0.05s
Iteration 2221, Loss: 0.0354, Time: 0.05s
Iteration 2222, Loss: 0.0239, Time: 0.05s
Iteration 2223, Loss: 0.0440, Time: 0.04s
Iteration 2224, Loss: 0.0381, Time: 0.04s
Iteration 2225, Loss: 0.0371, Time: 0.05s
Iteration 2226, Loss: 0.0355, Time: 0.04s
Iteration 2227, Loss: 0.0254, Time: 0.04s
Iteration 2228, Loss: 0.0386, Time: 0.05s
Iteration 2229, Loss: 0.0325, Time: 0.04s
Iteration 2230, Loss: 0.0269, Time: 0.04s
Iteration 2231, Loss: 0.0251, Time: 0.05s
Iteration 2232, Loss: 0.0256, Time: 0.04s
Iteration 2233, Loss: 0.0232, Time: 0.05s
Iteration 2234, Loss: 0.0512, Time: 0.05s
Iteration 2235, Loss: 0.0232, Time: 0.04s
Iteration 2236, Loss: 0.0440, Time: 0.04s
Iteration 2237, Loss: 0.0477, Time: 0.04s
Iteration 2238, Loss: 0.0126, Time: 0.05s
Iteration 2239, Loss: 0.0176, Time: 0.05s
Iteration 2240, Loss: 0.0386, Time: 0.05s
Iteration 2241, Loss: 0.0266, Time: 0.04s
Iteration 2242, Loss: 0.0427, Time: 0.04s
Iteration 2243, Loss: 0.0356, Time: 0.05s
Iteration 2244, Loss: 0.0239, Time: 0.05s
Iteration 2245, Loss: 0.0288, Time: 0.04s
Iteration 2246, Loss: 0.0267, Time: 0.04s
Iteration 2247, Loss: 0.0366, Time: 0.05s
Iteration 2248, Loss: 0.0366, Time: 0.05s
Iteration 2249, Loss: 0.0414, Time: 0.05s
Iteration 2250, Loss: 0.0377, Time: 0.04s
Iteration 2251, Loss: 0.0277, Time: 0.05s
Iteration 2252, Loss: 0.0297, Time: 0.04s
Iteration 2253, Loss: 0.0364, Time: 0.05s
Iteration 2254, Loss: 0.0307, Time: 0.05s
Iteration 2255, Loss: 0.0358, Time: 0.05s
Iteration 2256, Loss: 0.0260, Time: 0.04s
Iteration 2257, Loss: 0.0403, Time: 0.05s
Iteration 2258, Loss: 0.0331, Time: 0.05s
```

```
Iteration 2259, Loss: 0.0203, Time: 0.04s
Iteration 2260, Loss: 0.0234, Time: 0.04s
Iteration 2261, Loss: 0.0378, Time: 0.04s
Iteration 2262, Loss: 0.0291, Time: 0.05s
Iteration 2263, Loss: 0.0313, Time: 0.05s
Iteration 2264, Loss: 0.0402, Time: 0.04s
Iteration 2265, Loss: 0.0294, Time: 0.05s
Iteration 2266, Loss: 0.0358, Time: 0.04s
Iteration 2267, Loss: 0.0394, Time: 0.05s
Iteration 2268, Loss: 0.0327, Time: 0.04s
Iteration 2269, Loss: 0.0292, Time: 0.04s
Iteration 2270, Loss: 0.0398, Time: 0.04s
Iteration 2271, Loss: 0.0249, Time: 0.04s
Iteration 2272, Loss: 0.0282, Time: 0.04s
Iteration 2273, Loss: 0.0202, Time: 0.05s
Iteration 2274, Loss: 0.0369, Time: 0.05s
Iteration 2275, Loss: 0.0218, Time: 0.05s
Iteration 2276, Loss: 0.0351, Time: 0.04s
Iteration 2277, Loss: 0.0511, Time: 0.05s
Iteration 2278, Loss: 0.0267, Time: 0.05s
Iteration 2279, Loss: 0.0271, Time: 0.05s
Iteration 2280, Loss: 0.0263, Time: 0.04s
Iteration 2281, Loss: 0.0300, Time: 0.04s
Iteration 2282, Loss: 0.0322, Time: 0.04s
Iteration 2283, Loss: 0.0267, Time: 0.04s
Iteration 2284, Loss: 0.0377, Time: 0.04s
Iteration 2285, Loss: 0.0266, Time: 0.04s
Iteration 2286, Loss: 0.0428, Time: 0.04s
Iteration 2287, Loss: 0.0301, Time: 0.04s
Iteration 2288, Loss: 0.0255, Time: 0.05s
Iteration 2289, Loss: 0.0265, Time: 0.05s
Iteration 2290, Loss: 0.0299, Time: 0.05s
Iteration 2291, Loss: 0.0312, Time: 0.04s
Iteration 2292, Loss: 0.0227, Time: 0.05s
Iteration 2293, Loss: 0.0463, Time: 0.04s
Iteration 2294, Loss: 0.0278, Time: 0.05s
Iteration 2295, Loss: 0.0237, Time: 0.05s
Iteration 2296, Loss: 0.0526, Time: 0.04s
Iteration 2297, Loss: 0.0407, Time: 0.05s
Iteration 2298, Loss: 0.0221, Time: 0.04s
Iteration 2299, Loss: 0.0364, Time: 0.05s
Iteration 2300, Loss: 0.0231, Time: 0.04s
Iteration 2300, Loss: 0.0231, Time: 0.04s
Test Loss: 0.0397
Iteration 2301, Loss: 0.0351, Time: 0.05s
Iteration 2302, Loss: 0.0406, Time: 0.04s
Iteration 2303, Loss: 0.0236, Time: 0.04s
Iteration 2304, Loss: 0.0214, Time: 0.04s
```

```
Iteration 2305, Loss: 0.0392, Time: 0.04s
Iteration 2306, Loss: 0.0376, Time: 0.04s
Iteration 2307, Loss: 0.0317, Time: 0.05s
Iteration 2308, Loss: 0.0238, Time: 0.04s
Iteration 2309, Loss: 0.0333, Time: 0.05s
Iteration 2310, Loss: 0.0566, Time: 0.04s
Iteration 2311, Loss: 0.0483, Time: 0.04s
Iteration 2312, Loss: 0.0363, Time: 0.04s
Iteration 2313, Loss: 0.0288, Time: 0.05s
Iteration 2314, Loss: 0.0394, Time: 0.05s
Iteration 2315, Loss: 0.0408, Time: 0.05s
Iteration 2316, Loss: 0.0324, Time: 0.05s
Iteration 2317, Loss: 0.0244, Time: 0.05s
Iteration 2318, Loss: 0.0335, Time: 0.04s
Iteration 2319, Loss: 0.0271, Time: 0.04s
Iteration 2320, Loss: 0.0286, Time: 0.04s
Iteration 2321, Loss: 0.0216, Time: 0.04s
Iteration 2322, Loss: 0.0377, Time: 0.04s
Iteration 2323, Loss: 0.0285, Time: 0.05s
Iteration 2324, Loss: 0.0271, Time: 0.04s
Iteration 2325, Loss: 0.0307, Time: 0.05s
Iteration 2326, Loss: 0.0308, Time: 0.04s
Iteration 2327, Loss: 0.0253, Time: 0.05s
Iteration 2328, Loss: 0.0351, Time: 0.05s
Iteration 2329, Loss: 0.0206, Time: 0.04s
Iteration 2330, Loss: 0.0273, Time: 0.04s
Iteration 2331, Loss: 0.0360, Time: 0.05s
Iteration 2332, Loss: 0.0258, Time: 0.05s
Iteration 2333, Loss: 0.0407, Time: 0.04s
Iteration 2334, Loss: 0.0277, Time: 0.05s
Iteration 2335, Loss: 0.0207, Time: 0.05s
Iteration 2336, Loss: 0.0325, Time: 0.04s
Iteration 2337, Loss: 0.0348, Time: 0.05s
Iteration 2338, Loss: 0.0287, Time: 0.04s
Iteration 2339, Loss: 0.0312, Time: 0.04s
Iteration 2340, Loss: 0.0190, Time: 0.04s
Iteration 2341, Loss: 0.0397, Time: 0.05s
Iteration 2342, Loss: 0.0291, Time: 0.04s
Iteration 2343, Loss: 0.0264, Time: 0.04s
Iteration 2344, Loss: 0.0376, Time: 0.04s
Iteration 2345, Loss: 0.0251, Time: 0.05s
Iteration 2346, Loss: 0.0325, Time: 0.04s
Iteration 2347, Loss: 0.0280, Time: 0.05s
Iteration 2348, Loss: 0.0254, Time: 0.05s
Iteration 2349, Loss: 0.0230, Time: 0.04s
Iteration 2350, Loss: 0.0301, Time: 0.05s
Iteration 2351, Loss: 0.0218, Time: 0.04s
Iteration 2352, Loss: 0.0375, Time: 0.04s
```

```
Iteration 2353, Loss: 0.0340, Time: 0.04s
Iteration 2354, Loss: 0.0415, Time: 0.04s
Iteration 2355, Loss: 0.0338, Time: 0.05s
Iteration 2356, Loss: 0.0327, Time: 0.04s
Iteration 2357, Loss: 0.0278, Time: 0.05s
Iteration 2358, Loss: 0.0305, Time: 0.04s
Iteration 2359, Loss: 0.0314, Time: 0.05s
Iteration 2360, Loss: 0.0302, Time: 0.04s
Iteration 2361, Loss: 0.0422, Time: 0.04s
Iteration 2362, Loss: 0.0202, Time: 0.04s
Iteration 2363, Loss: 0.0336, Time: 0.04s
Iteration 2364, Loss: 0.0355, Time: 0.04s
Iteration 2365, Loss: 0.0391, Time: 0.04s
Iteration 2366, Loss: 0.0384, Time: 0.04s
Iteration 2367, Loss: 0.0337, Time: 0.05s
Iteration 2368, Loss: 0.0216, Time: 0.04s
Iteration 2369, Loss: 0.0371, Time: 0.05s
Iteration 2370, Loss: 0.0196, Time: 0.05s
Iteration 2371, Loss: 0.0361, Time: 0.05s
Iteration 2372, Loss: 0.0248, Time: 0.04s
Iteration 2373, Loss: 0.0339, Time: 0.05s
Iteration 2374, Loss: 0.0388, Time: 0.05s
Iteration 2375, Loss: 0.0303, Time: 0.04s
Iteration 2376, Loss: 0.0334, Time: 0.04s
Iteration 2377, Loss: 0.0356, Time: 0.05s
Iteration 2378, Loss: 0.0337, Time: 0.04s
Iteration 2379, Loss: 0.0289, Time: 0.04s
Iteration 2380, Loss: 0.0404, Time: 0.04s
Iteration 2381, Loss: 0.0338, Time: 0.05s
Iteration 2382, Loss: 0.0296, Time: 0.06s
Iteration 2383, Loss: 0.0338, Time: 0.05s
Iteration 2384, Loss: 0.0298, Time: 0.04s
Iteration 2385, Loss: 0.0291, Time: 0.04s
Iteration 2386, Loss: 0.0311, Time: 0.04s
Iteration 2387, Loss: 0.0392, Time: 0.04s
Iteration 2388, Loss: 0.0539, Time: 0.05s
Iteration 2389, Loss: 0.0256, Time: 0.04s
Iteration 2390, Loss: 0.0360, Time: 0.04s
Iteration 2391, Loss: 0.0300, Time: 0.04s
Iteration 2392, Loss: 0.0276, Time: 0.05s
Iteration 2393, Loss: 0.0292, Time: 0.05s
Iteration 2394, Loss: 0.0353, Time: 0.05s
Iteration 2395, Loss: 0.0268, Time: 0.06s
Iteration 2396, Loss: 0.0608, Time: 0.06s
Iteration 2397, Loss: 0.0276, Time: 0.05s
Iteration 2398, Loss: 0.0251, Time: 0.05s
Iteration 2399, Loss: 0.0234, Time: 0.05s
Iteration 2400, Loss: 0.0294, Time: 0.05s
```

```
Iteration 2400, Loss: 0.0294, Time: 0.05s
Test Loss: 0.0383
Iteration 2401, Loss: 0.0303, Time: 0.05s
Iteration 2402, Loss: 0.0264, Time: 0.04s
Iteration 2403, Loss: 0.0532, Time: 0.05s
Iteration 2404, Loss: 0.0254, Time: 0.05s
Iteration 2405, Loss: 0.0313, Time: 0.04s
Iteration 2406, Loss: 0.0341, Time: 0.05s
Iteration 2407, Loss: 0.0191, Time: 0.04s
Iteration 2408, Loss: 0.0377, Time: 0.04s
Iteration 2409, Loss: 0.0357, Time: 0.04s
Iteration 2410, Loss: 0.0245, Time: 0.04s
Iteration 2411, Loss: 0.0331, Time: 0.04s
Iteration 2412, Loss: 0.0289, Time: 0.04s
Iteration 2413, Loss: 0.0250, Time: 0.06s
Iteration 2414, Loss: 0.0448, Time: 0.04s
Iteration 2415, Loss: 0.0423, Time: 0.05s
Iteration 2416, Loss: 0.0429, Time: 0.05s
Iteration 2417, Loss: 0.0301, Time: 0.05s
Iteration 2418, Loss: 0.0365, Time: 0.05s
Iteration 2419, Loss: 0.0198, Time: 0.05s
Iteration 2420, Loss: 0.0265, Time: 0.04s
Iteration 2421, Loss: 0.0313, Time: 0.05s
Iteration 2422, Loss: 0.0323, Time: 0.05s
Iteration 2423, Loss: 0.0505, Time: 0.04s
Iteration 2424, Loss: 0.0346, Time: 0.06s
Iteration 2425, Loss: 0.0317, Time: 0.05s
Iteration 2426, Loss: 0.0316, Time: 0.04s
Iteration 2427, Loss: 0.0373, Time: 0.04s
Iteration 2428, Loss: 0.0289, Time: 0.04s
Iteration 2429, Loss: 0.0304, Time: 0.05s
Iteration 2430, Loss: 0.0421, Time: 0.04s
Iteration 2431, Loss: 0.0325, Time: 0.06s
Iteration 2432, Loss: 0.0224, Time: 0.05s
Iteration 2433, Loss: 0.0280, Time: 0.05s
Iteration 2434, Loss: 0.0379, Time: 0.04s
Iteration 2435, Loss: 0.0286, Time: 0.05s
Iteration 2436, Loss: 0.0351, Time: 0.05s
Iteration 2437, Loss: 0.0357, Time: 0.05s
Iteration 2438, Loss: 0.0318, Time: 0.05s
Iteration 2439, Loss: 0.0263, Time: 0.05s
Iteration 2440, Loss: 0.0359, Time: 0.05s
Iteration 2441, Loss: 0.0572, Time: 0.04s
Iteration 2442, Loss: 0.0459, Time: 0.06s
Iteration 2443, Loss: 0.0325, Time: 0.04s
Iteration 2444, Loss: 0.0347, Time: 0.04s
Iteration 2445, Loss: 0.0400, Time: 0.04s
Iteration 2446, Loss: 0.0242, Time: 0.05s
```

```
Iteration 2447, Loss: 0.0172, Time: 0.04s
Iteration 2448, Loss: 0.0299, Time: 0.04s
Iteration 2449, Loss: 0.0262, Time: 0.04s
Iteration 2450, Loss: 0.0193, Time: 0.04s
Iteration 2451, Loss: 0.0260, Time: 0.05s
Iteration 2452, Loss: 0.0287, Time: 0.04s
Iteration 2453, Loss: 0.0450, Time: 0.05s
Iteration 2454, Loss: 0.0274, Time: 0.04s
Iteration 2455, Loss: 0.0323, Time: 0.05s
Iteration 2456, Loss: 0.0256, Time: 0.05s
Iteration 2457, Loss: 0.0267, Time: 0.04s
Iteration 2458, Loss: 0.0244, Time: 0.05s
Iteration 2459, Loss: 0.0375, Time: 0.05s
Iteration 2460, Loss: 0.0397, Time: 0.05s
Iteration 2461, Loss: 0.0265, Time: 0.04s
Iteration 2462, Loss: 0.0332, Time: 0.05s
Iteration 2463, Loss: 0.0425, Time: 0.05s
Iteration 2464, Loss: 0.0222, Time: 0.04s
Iteration 2465, Loss: 0.0367, Time: 0.05s
Iteration 2466, Loss: 0.0339, Time: 0.04s
Iteration 2467, Loss: 0.0373, Time: 0.04s
Iteration 2468, Loss: 0.0248, Time: 0.04s
Iteration 2469, Loss: 0.0350, Time: 0.05s
Iteration 2470, Loss: 0.0277, Time: 0.04s
Iteration 2471, Loss: 0.0378, Time: 0.06s
Iteration 2472, Loss: 0.0335, Time: 0.04s
Iteration 2473, Loss: 0.0337, Time: 0.06s
Iteration 2474, Loss: 0.0255, Time: 0.04s
Iteration 2475, Loss: 0.0284, Time: 0.05s
Iteration 2476, Loss: 0.0357, Time: 0.04s
Iteration 2477, Loss: 0.0314, Time: 0.05s
Iteration 2478, Loss: 0.0187, Time: 0.05s
Iteration 2479, Loss: 0.0312, Time: 0.05s
Iteration 2480, Loss: 0.0260, Time: 0.05s
Iteration 2481, Loss: 0.0377, Time: 0.04s
Iteration 2482, Loss: 0.0313, Time: 0.04s
Iteration 2483, Loss: 0.0300, Time: 0.05s
Iteration 2484, Loss: 0.0331, Time: 0.04s
Iteration 2485, Loss: 0.0272, Time: 0.04s
Iteration 2486, Loss: 0.0324, Time: 0.05s
Iteration 2487, Loss: 0.0387, Time: 0.05s
Iteration 2488, Loss: 0.0437, Time: 0.04s
Iteration 2489, Loss: 0.0233, Time: 0.04s
Iteration 2490, Loss: 0.0256, Time: 0.04s
Iteration 2491, Loss: 0.0283, Time: 0.04s
Iteration 2492, Loss: 0.0299, Time: 0.04s
Iteration 2493, Loss: 0.0229, Time: 0.04s
Iteration 2494, Loss: 0.0251, Time: 0.05s
```

```
Iteration 2495, Loss: 0.0281, Time: 0.05s
Iteration 2496, Loss: 0.0280, Time: 0.05s
Iteration 2497, Loss: 0.0204, Time: 0.04s
Iteration 2498, Loss: 0.0314, Time: 0.04s
Iteration 2499, Loss: 0.0346, Time: 0.04s
Iteration 2500, Loss: 0.0291, Time: 0.04s
Iteration 2500, Loss: 0.0291, Time: 0.04s
Test Loss: 0.0431
Iteration 2501, Loss: 0.0250, Time: 0.04s
Iteration 2502, Loss: 0.0279, Time: 0.04s
Iteration 2503, Loss: 0.0323, Time: 0.04s
Iteration 2504, Loss: 0.0217, Time: 0.04s
Iteration 2505, Loss: 0.0258, Time: 0.04s
Iteration 2506, Loss: 0.0382, Time: 0.04s
Iteration 2507, Loss: 0.0364, Time: 0.04s
Iteration 2508, Loss: 0.0165, Time: 0.04s
Iteration 2509, Loss: 0.0390, Time: 0.05s
Iteration 2510, Loss: 0.0307, Time: 0.04s
Iteration 2511, Loss: 0.0247, Time: 0.05s
Iteration 2512, Loss: 0.0207, Time: 0.04s
Iteration 2513, Loss: 0.0189, Time: 0.04s
Iteration 2514, Loss: 0.0469, Time: 0.05s
Iteration 2515, Loss: 0.0460, Time: 0.05s
Iteration 2516, Loss: 0.0246, Time: 0.04s
Iteration 2517, Loss: 0.0295, Time: 0.04s
Iteration 2518, Loss: 0.0474, Time: 0.05s
Iteration 2519, Loss: 0.0290, Time: 0.05s
Iteration 2520, Loss: 0.0352, Time: 0.04s
Iteration 2521, Loss: 0.0325, Time: 0.05s
Iteration 2522, Loss: 0.0319, Time: 0.05s
Iteration 2523, Loss: 0.0343, Time: 0.05s
Iteration 2524, Loss: 0.0343, Time: 0.04s
Iteration 2525, Loss: 0.0205, Time: 0.05s
Iteration 2526, Loss: 0.0247, Time: 0.05s
Iteration 2527, Loss: 0.0355, Time: 0.04s
Iteration 2528, Loss: 0.0342, Time: 0.05s
Iteration 2529, Loss: 0.0260, Time: 0.05s
Iteration 2530, Loss: 0.0382, Time: 0.04s
Iteration 2531, Loss: 0.0147, Time: 0.05s
Iteration 2532, Loss: 0.0246, Time: 0.04s
Iteration 2533, Loss: 0.0389, Time: 0.05s
Iteration 2534, Loss: 0.0284, Time: 0.04s
Iteration 2535, Loss: 0.0251, Time: 0.05s
Iteration 2536, Loss: 0.0371, Time: 0.04s
Iteration 2537, Loss: 0.0343, Time: 0.04s
Iteration 2538, Loss: 0.0347, Time: 0.05s
Iteration 2539, Loss: 0.0290, Time: 0.04s
Iteration 2540, Loss: 0.0344, Time: 0.05s
```

```
Iteration 2541, Loss: 0.0367, Time: 0.05s
Iteration 2542, Loss: 0.0317, Time: 0.04s
Iteration 2543, Loss: 0.0242, Time: 0.04s
Iteration 2544, Loss: 0.0227, Time: 0.04s
Iteration 2545, Loss: 0.0275, Time: 0.04s
Iteration 2546, Loss: 0.0222, Time: 0.04s
Iteration 2547, Loss: 0.0275, Time: 0.04s
Iteration 2548, Loss: 0.0294, Time: 0.04s
Iteration 2549, Loss: 0.0251, Time: 0.05s
Iteration 2550, Loss: 0.0273, Time: 0.04s
Iteration 2551, Loss: 0.0450, Time: 0.05s
Iteration 2552, Loss: 0.0297, Time: 0.05s
Iteration 2553, Loss: 0.0282, Time: 0.04s
Iteration 2554, Loss: 0.0307, Time: 0.05s
Iteration 2555, Loss: 0.0390, Time: 0.05s
Iteration 2556, Loss: 0.0269, Time: 0.05s
Iteration 2557, Loss: 0.0295, Time: 0.04s
Iteration 2558, Loss: 0.0365, Time: 0.04s
Iteration 2559, Loss: 0.0304, Time: 0.04s
Iteration 2560, Loss: 0.0229, Time: 0.04s
Iteration 2561, Loss: 0.0203, Time: 0.05s
Iteration 2562, Loss: 0.0258, Time: 0.05s
Iteration 2563, Loss: 0.0359, Time: 0.04s
Iteration 2564, Loss: 0.0229, Time: 0.04s
Iteration 2565, Loss: 0.0190, Time: 0.04s
Iteration 2566, Loss: 0.0286, Time: 0.04s
Iteration 2567, Loss: 0.0242, Time: 0.05s
Iteration 2568, Loss: 0.0366, Time: 0.04s
Iteration 2569, Loss: 0.0186, Time: 0.05s
Iteration 2570, Loss: 0.0471, Time: 0.04s
Iteration 2571, Loss: 0.0242, Time: 0.04s
Iteration 2572, Loss: 0.0367, Time: 0.04s
Iteration 2573, Loss: 0.0308, Time: 0.04s
Iteration 2574, Loss: 0.0322, Time: 0.04s
Iteration 2575, Loss: 0.0288, Time: 0.04s
Iteration 2576, Loss: 0.0391, Time: 0.04s
Iteration 2577, Loss: 0.0368, Time: 0.04s
Iteration 2578, Loss: 0.0414, Time: 0.05s
Iteration 2579, Loss: 0.0319, Time: 0.04s
Iteration 2580, Loss: 0.0453, Time: 0.05s
Iteration 2581, Loss: 0.0367, Time: 0.05s
Iteration 2582, Loss: 0.0354, Time: 0.04s
Iteration 2583, Loss: 0.0230, Time: 0.04s
Iteration 2584, Loss: 0.0262, Time: 0.05s
Iteration 2585, Loss: 0.0314, Time: 0.05s
Iteration 2586, Loss: 0.0263, Time: 0.04s
Iteration 2587, Loss: 0.0221, Time: 0.05s
Iteration 2588, Loss: 0.0355, Time: 0.04s
```

```
Iteration 2589, Loss: 0.0338, Time: 0.04s
Iteration 2590, Loss: 0.0330, Time: 0.04s
Iteration 2591, Loss: 0.0384, Time: 0.05s
Iteration 2592, Loss: 0.0313, Time: 0.04s
Iteration 2593, Loss: 0.0503, Time: 0.04s
Iteration 2594, Loss: 0.0320, Time: 0.05s
Iteration 2595, Loss: 0.0394, Time: 0.05s
Iteration 2596, Loss: 0.0303, Time: 0.05s
Iteration 2597, Loss: 0.0202, Time: 0.04s
Iteration 2598, Loss: 0.0274, Time: 0.05s
Iteration 2599, Loss: 0.0499, Time: 0.04s
Iteration 2600, Loss: 0.0289, Time: 0.05s
Iteration 2600, Loss: 0.0289, Time: 0.05s
Test Loss: 0.0528
Iteration 2601, Loss: 0.0251, Time: 0.04s
Iteration 2602, Loss: 0.0265, Time: 0.04s
Iteration 2603, Loss: 0.0297, Time: 0.04s
Iteration 2604, Loss: 0.0403, Time: 0.04s
Iteration 2605, Loss: 0.0353, Time: 0.04s
Iteration 2606, Loss: 0.0232, Time: 0.04s
Iteration 2607, Loss: 0.0222, Time: 0.04s
Iteration 2608, Loss: 0.0321, Time: 0.04s
Iteration 2609, Loss: 0.0311, Time: 0.05s
Iteration 2610, Loss: 0.0347, Time: 0.05s
Iteration 2611, Loss: 0.0207, Time: 0.05s
Iteration 2612, Loss: 0.0155, Time: 0.06s
Iteration 2613, Loss: 0.0307, Time: 0.05s
Iteration 2614, Loss: 0.0395, Time: 0.06s
Iteration 2615, Loss: 0.0469, Time: 0.07s
Iteration 2616, Loss: 0.0290, Time: 0.09s
Iteration 2617, Loss: 0.0236, Time: -0.09s
Iteration 2618, Loss: 0.0253, Time: 0.07s
Iteration 2619, Loss: 0.0279, Time: 0.05s
Iteration 2620, Loss: 0.0180, Time: 0.05s
Iteration 2621, Loss: 0.0255, Time: 0.04s
Iteration 2622, Loss: 0.0206, Time: 0.05s
Iteration 2623, Loss: 0.0293, Time: 0.04s
Iteration 2624, Loss: 0.0228, Time: 0.05s
Iteration 2625, Loss: 0.0256, Time: 0.06s
Iteration 2626, Loss: 0.0460, Time: 0.05s
Iteration 2627, Loss: 0.0308, Time: 0.05s
Iteration 2628, Loss: 0.0393, Time: 0.04s
Iteration 2629, Loss: 0.0181, Time: 0.04s
Iteration 2630, Loss: 0.0292, Time: 0.04s
Iteration 2631, Loss: 0.0340, Time: 0.04s
Iteration 2632, Loss: 0.0386, Time: 0.05s
Iteration 2633, Loss: 0.0309, Time: 0.04s
Iteration 2634, Loss: 0.0378, Time: 0.04s
```

```
Iteration 2635, Loss: 0.0200, Time: 0.04s
Iteration 2636, Loss: 0.0309, Time: 0.04s
Iteration 2637, Loss: 0.0286, Time: 0.04s
Iteration 2638, Loss: 0.0223, Time: 0.04s
Iteration 2639, Loss: 0.0437, Time: 0.04s
Iteration 2640, Loss: 0.0296, Time: 0.04s
Iteration 2641, Loss: 0.0282, Time: 0.04s
Iteration 2642, Loss: 0.0264, Time: 0.05s
Iteration 2643, Loss: 0.0254, Time: 0.04s
Iteration 2644, Loss: 0.0303, Time: 0.05s
Iteration 2645, Loss: 0.0270, Time: 0.04s
Iteration 2646, Loss: 0.0253, Time: 0.05s
Iteration 2647, Loss: 0.0270, Time: 0.04s
Iteration 2648, Loss: 0.0309, Time: 0.05s
Iteration 2649, Loss: 0.0396, Time: 0.04s
Iteration 2650, Loss: 0.0383, Time: 0.04s
Iteration 2651, Loss: 0.0344, Time: 0.04s
Iteration 2652, Loss: 0.0285, Time: 0.05s
Iteration 2653, Loss: 0.0210, Time: 0.04s
Iteration 2654, Loss: 0.0233, Time: 0.04s
Iteration 2655, Loss: 0.0233, Time: 0.04s
Iteration 2656, Loss: 0.0380, Time: 0.05s
Iteration 2657, Loss: 0.0231, Time: 0.04s
Iteration 2658, Loss: 0.0246, Time: 0.04s
Iteration 2659, Loss: 0.0368, Time: 0.06s
Iteration 2660, Loss: 0.0281, Time: 0.04s
Iteration 2661, Loss: 0.0296, Time: 0.05s
Iteration 2662, Loss: 0.0187, Time: 0.04s
Iteration 2663, Loss: 0.0334, Time: 0.04s
Iteration 2664, Loss: 0.0315, Time: 0.05s
Iteration 2665, Loss: 0.0390, Time: 0.05s
Iteration 2666, Loss: 0.0215, Time: 0.04s
Iteration 2667, Loss: 0.0369, Time: 0.04s
Iteration 2668, Loss: 0.0214, Time: 0.05s
Iteration 2669, Loss: 0.0221, Time: 0.04s
Iteration 2670, Loss: 0.0286, Time: 0.05s
Iteration 2671, Loss: 0.0478, Time: 0.04s
Iteration 2672, Loss: 0.0220, Time: 0.05s
Iteration 2673, Loss: 0.0319, Time: 0.04s
Iteration 2674, Loss: 0.0409, Time: 0.04s
Iteration 2675, Loss: 0.0222, Time: 0.05s
Iteration 2676, Loss: 0.0276, Time: 0.04s
Iteration 2677, Loss: 0.0383, Time: 0.04s
Iteration 2678, Loss: 0.0420, Time: 0.05s
Iteration 2679, Loss: 0.0236, Time: 0.04s
Iteration 2680, Loss: 0.0224, Time: 0.04s
Iteration 2681, Loss: 0.0293, Time: 0.04s
Iteration 2682, Loss: 0.0225, Time: 0.05s
```

```
Iteration 2683, Loss: 0.0233, Time: 0.04s
Iteration 2684, Loss: 0.0188, Time: 0.05s
Iteration 2685, Loss: 0.0296, Time: 0.04s
Iteration 2686, Loss: 0.0252, Time: 0.04s
Iteration 2687, Loss: 0.0347, Time: 0.04s
Iteration 2688, Loss: 0.0171, Time: 0.04s
Iteration 2689, Loss: 0.0260, Time: 0.04s
Iteration 2690, Loss: 0.0252, Time: 0.05s
Iteration 2691, Loss: 0.0212, Time: 0.05s
Iteration 2692, Loss: 0.0289, Time: 0.05s
Iteration 2693, Loss: 0.0213, Time: 0.04s
Iteration 2694, Loss: 0.0223, Time: 0.04s
Iteration 2695, Loss: 0.0308, Time: 0.05s
Iteration 2696, Loss: 0.0377, Time: 0.05s
Iteration 2697, Loss: 0.0295, Time: 0.05s
Iteration 2698, Loss: 0.0257, Time: 0.05s
Iteration 2699, Loss: 0.0328, Time: 0.05s
Iteration 2700, Loss: 0.0276, Time: 0.05s
Iteration 2700, Loss: 0.0276, Time: 0.05s
Test Loss: 0.0359
Iteration 2701, Loss: 0.0344, Time: 0.05s
Iteration 2702, Loss: 0.0333, Time: 0.04s
Iteration 2703, Loss: 0.0288, Time: 0.05s
Iteration 2704, Loss: 0.0317, Time: 0.05s
Iteration 2705, Loss: 0.0407, Time: 0.04s
Iteration 2706, Loss: 0.0285, Time: 0.05s
Iteration 2707, Loss: 0.0233, Time: 0.04s
Iteration 2708, Loss: 0.0364, Time: 0.05s
Iteration 2709, Loss: 0.0304, Time: 0.05s
Iteration 2710, Loss: 0.0277, Time: 0.05s
Iteration 2711, Loss: 0.0394, Time: 0.04s
Iteration 2712, Loss: 0.0280, Time: 0.05s
Iteration 2713, Loss: 0.0381, Time: 0.04s
Iteration 2714, Loss: 0.0261, Time: 0.04s
Iteration 2715, Loss: 0.0272, Time: 0.05s
Iteration 2716, Loss: 0.0312, Time: 0.05s
Iteration 2717, Loss: 0.0298, Time: 0.05s
Iteration 2718, Loss: 0.0323, Time: 0.05s
Iteration 2719, Loss: 0.0278, Time: 0.04s
Iteration 2720, Loss: 0.0241, Time: 0.05s
Iteration 2721, Loss: 0.0197, Time: 0.04s
Iteration 2722, Loss: 0.0248, Time: 0.04s
Iteration 2723, Loss: 0.0336, Time: 0.04s
Iteration 2724, Loss: 0.0358, Time: 0.04s
Iteration 2725, Loss: 0.0258, Time: 0.05s
Iteration 2726, Loss: 0.0246, Time: 0.05s
Iteration 2727, Loss: 0.0360, Time: 0.06s
Iteration 2728, Loss: 0.0274, Time: 0.04s
```

```
Iteration 2729, Loss: 0.0332, Time: 0.05s
Iteration 2730, Loss: 0.0279, Time: 0.04s
Iteration 2731, Loss: 0.0369, Time: 0.06s
Iteration 2732, Loss: 0.0290, Time: 0.04s
Iteration 2733, Loss: 0.0231, Time: 0.05s
Iteration 2734, Loss: 0.0336, Time: 0.04s
Iteration 2735, Loss: 0.0232, Time: 0.05s
Iteration 2736, Loss: 0.0336, Time: 0.04s
Iteration 2737, Loss: 0.0241, Time: 0.04s
Iteration 2738, Loss: 0.0331, Time: 0.05s
Iteration 2739, Loss: 0.0264, Time: 0.04s
Iteration 2740, Loss: 0.0214, Time: 0.04s
Iteration 2741, Loss: 0.0344, Time: 0.05s
Iteration 2742, Loss: 0.0285, Time: 0.04s
Iteration 2743, Loss: 0.0314, Time: 0.04s
Iteration 2744, Loss: 0.0432, Time: 0.05s
Iteration 2745, Loss: 0.0246, Time: 0.05s
Iteration 2746, Loss: 0.0357, Time: 0.04s
Iteration 2747, Loss: 0.0193, Time: 0.04s
Iteration 2748, Loss: 0.0298, Time: 0.04s
Iteration 2749, Loss: 0.0333, Time: 0.05s
Iteration 2750, Loss: 0.0270, Time: 0.04s
Iteration 2751, Loss: 0.0270, Time: 0.05s
Iteration 2752, Loss: 0.0337, Time: 0.05s
Iteration 2753, Loss: 0.0254, Time: 0.05s
Iteration 2754, Loss: 0.0260, Time: 0.04s
Iteration 2755, Loss: 0.0346, Time: 0.04s
Iteration 2756, Loss: 0.0170, Time: 0.05s
Iteration 2757, Loss: 0.0207, Time: 0.04s
Iteration 2758, Loss: 0.0165, Time: 0.05s
Iteration 2759, Loss: 0.0129, Time: 0.04s
Iteration 2760, Loss: 0.0455, Time: 0.04s
Iteration 2761, Loss: 0.0183, Time: 0.04s
Iteration 2762, Loss: 0.0272, Time: 0.05s
Iteration 2763, Loss: 0.0237, Time: 0.04s
Iteration 2764, Loss: 0.0324, Time: 0.05s
Iteration 2765, Loss: 0.0331, Time: 0.05s
Iteration 2766, Loss: 0.0338, Time: 0.04s
Iteration 2767, Loss: 0.0256, Time: 0.04s
Iteration 2768, Loss: 0.0315, Time: 0.04s
Iteration 2769, Loss: 0.0400, Time: 0.04s
Iteration 2770, Loss: 0.0175, Time: 0.04s
Iteration 2771, Loss: 0.0261, Time: 0.05s
Iteration 2772, Loss: 0.0259, Time: 0.05s
Iteration 2773, Loss: 0.0181, Time: 0.04s
Iteration 2774, Loss: 0.0337, Time: 0.04s
Iteration 2775, Loss: 0.0387, Time: 0.05s
Iteration 2776, Loss: 0.0327, Time: 0.05s
```

```
Iteration 2777, Loss: 0.0322, Time: 0.04s
Iteration 2778, Loss: 0.0274, Time: 0.04s
Iteration 2779, Loss: 0.0197, Time: 0.04s
Iteration 2780, Loss: 0.0235, Time: 0.04s
Iteration 2781, Loss: 0.0307, Time: 0.05s
Iteration 2782, Loss: 0.0228, Time: 0.04s
Iteration 2783, Loss: 0.0317, Time: 0.04s
Iteration 2784, Loss: 0.0116, Time: 0.04s
Iteration 2785, Loss: 0.0289, Time: 0.04s
Iteration 2786, Loss: 0.0280, Time: 0.04s
Iteration 2787, Loss: 0.0397, Time: 0.05s
Iteration 2788, Loss: 0.0197, Time: 0.04s
Iteration 2789, Loss: 0.0278, Time: 0.04s
Iteration 2790, Loss: 0.0269, Time: 0.04s
Iteration 2791, Loss: 0.0305, Time: 0.04s
Iteration 2792, Loss: 0.0301, Time: 0.04s
Iteration 2793, Loss: 0.0356, Time: 0.04s
Iteration 2794, Loss: 0.0310, Time: 0.05s
Iteration 2795, Loss: 0.0292, Time: 0.05s
Iteration 2796, Loss: 0.0386, Time: 0.04s
Iteration 2797, Loss: 0.0313, Time: 0.04s
Iteration 2798, Loss: 0.0255, Time: 0.04s
Iteration 2799, Loss: 0.0208, Time: 0.04s
Iteration 2800, Loss: 0.0270, Time: 0.04s
Iteration 2800, Loss: 0.0270, Time: 0.05s
Test Loss: 0.0320
Iteration 2801, Loss: 0.0214, Time: 0.04s
Iteration 2802, Loss: 0.0336, Time: 0.05s
Iteration 2803, Loss: 0.0395, Time: 0.04s
Iteration 2804, Loss: 0.0219, Time: 0.05s
Iteration 2805, Loss: 0.0276, Time: 0.05s
Iteration 2806, Loss: 0.0328, Time: 0.05s
Iteration 2807, Loss: 0.0397, Time: 0.04s
Iteration 2808, Loss: 0.0198, Time: 0.05s
Iteration 2809, Loss: 0.0378, Time: 0.04s
Iteration 2810, Loss: 0.0332, Time: 0.05s
Iteration 2811, Loss: 0.0286, Time: 0.05s
Iteration 2812, Loss: 0.0360, Time: 0.05s
Iteration 2813, Loss: 0.0430, Time: 0.04s
Iteration 2814, Loss: 0.0264, Time: 0.05s
Iteration 2815, Loss: 0.0436, Time: 0.04s
Iteration 2816, Loss: 0.0248, Time: 0.04s
Iteration 2817, Loss: 0.0262, Time: 0.04s
Iteration 2818, Loss: 0.0270, Time: 0.04s
Iteration 2819, Loss: 0.0350, Time: 0.04s
Iteration 2820, Loss: 0.0303, Time: 0.05s
Iteration 2821, Loss: 0.0285, Time: 0.04s
Iteration 2822, Loss: 0.0268, Time: 0.05s
```

```
Iteration 2823, Loss: 0.0297, Time: 0.05s
Iteration 2824, Loss: 0.0259, Time: 0.05s
Iteration 2825, Loss: 0.0403, Time: 0.05s
Iteration 2826, Loss: 0.0239, Time: 0.04s
Iteration 2827, Loss: 0.0195, Time: 0.04s
Iteration 2828, Loss: 0.0286, Time: 0.04s
Iteration 2829, Loss: 0.0220, Time: 0.05s
Iteration 2830, Loss: 0.0222, Time: 0.05s
Iteration 2831, Loss: 0.0350, Time: 0.06s
Iteration 2832, Loss: 0.0210, Time: 0.05s
Iteration 2833, Loss: 0.0337, Time: 0.06s
Iteration 2834, Loss: 0.0278, Time: 0.06s
Iteration 2835, Loss: 0.0216, Time: 0.05s
Iteration 2836, Loss: 0.0253, Time: 0.07s
Iteration 2837, Loss: 0.0157, Time: 0.05s
Iteration 2838, Loss: 0.0240, Time: 0.05s
Iteration 2839, Loss: 0.0104, Time: 0.05s
Iteration 2840, Loss: 0.0329, Time: 0.04s
Iteration 2841, Loss: 0.0282, Time: 0.05s
Iteration 2842, Loss: 0.0276, Time: 0.05s
Iteration 2843, Loss: 0.0512, Time: 0.04s
Iteration 2844, Loss: 0.0202, Time: 0.04s
Iteration 2845, Loss: 0.0238, Time: 0.04s
Iteration 2846, Loss: 0.0379, Time: 0.04s
Iteration 2847, Loss: 0.0268, Time: 0.04s
Iteration 2848, Loss: 0.0280, Time: 0.05s
Iteration 2849, Loss: 0.0238, Time: 0.05s
Iteration 2850, Loss: 0.0197, Time: 0.04s
Iteration 2851, Loss: 0.0358, Time: 0.05s
Iteration 2852, Loss: 0.0164, Time: 0.04s
Iteration 2853, Loss: 0.0186, Time: 0.05s
Iteration 2854, Loss: 0.0326, Time: 0.05s
Iteration 2855, Loss: 0.0208, Time: 0.05s
Iteration 2856, Loss: 0.0252, Time: 0.04s
Iteration 2857, Loss: 0.0333, Time: 0.04s
Iteration 2858, Loss: 0.0223, Time: 0.05s
Iteration 2859, Loss: 0.0300, Time: 0.05s
Iteration 2860, Loss: 0.0280, Time: 0.04s
Iteration 2861, Loss: 0.0333, Time: 0.05s
Iteration 2862, Loss: 0.0244, Time: 0.04s
Iteration 2863, Loss: 0.0233, Time: 0.04s
Iteration 2864, Loss: 0.0287, Time: 0.04s
Iteration 2865, Loss: 0.0316, Time: 0.04s
Iteration 2866, Loss: 0.0329, Time: 0.04s
Iteration 2867, Loss: 0.0448, Time: 0.04s
Iteration 2868, Loss: 0.0266, Time: 0.04s
Iteration 2869, Loss: 0.0321, Time: 0.04s
Iteration 2870, Loss: 0.0165, Time: 0.05s
```

```
Iteration 2871, Loss: 0.0188, Time: 0.04s
Iteration 2872, Loss: 0.0239, Time: 0.04s
Iteration 2873, Loss: 0.0213, Time: 0.05s
Iteration 2874, Loss: 0.0262, Time: 0.04s
Iteration 2875, Loss: 0.0271, Time: 0.05s
Iteration 2876, Loss: 0.0239, Time: 0.05s
Iteration 2877, Loss: 0.0318, Time: 0.05s
Iteration 2878, Loss: 0.0243, Time: 0.04s
Iteration 2879, Loss: 0.0414, Time: 0.05s
Iteration 2880, Loss: 0.0223, Time: 0.04s
Iteration 2881, Loss: 0.0260, Time: 0.04s
Iteration 2882, Loss: 0.0421, Time: 0.05s
Iteration 2883, Loss: 0.0264, Time: 0.05s
Iteration 2884, Loss: 0.0277, Time: 0.04s
Iteration 2885, Loss: 0.0151, Time: 0.04s
Iteration 2886, Loss: 0.0307, Time: 0.04s
Iteration 2887, Loss: 0.0350, Time: 0.04s
Iteration 2888, Loss: 0.0335, Time: 0.04s
Iteration 2889, Loss: 0.0114, Time: 0.05s
Iteration 2890, Loss: 0.0228, Time: 0.05s
Iteration 2891, Loss: 0.0236, Time: 0.04s
Iteration 2892, Loss: 0.0299, Time: 0.05s
Iteration 2893, Loss: 0.0266, Time: 0.05s
Iteration 2894, Loss: 0.0335, Time: 0.04s
Iteration 2895, Loss: 0.0237, Time: 0.04s
Iteration 2896, Loss: 0.0295, Time: 0.05s
Iteration 2897, Loss: 0.0336, Time: 0.05s
Iteration 2898, Loss: 0.0263, Time: 0.04s
Iteration 2899, Loss: 0.0284, Time: 0.05s
Iteration 2900, Loss: 0.0379, Time: 0.05s
Iteration 2900, Loss: 0.0379, Time: 0.05s
Test Loss: 0.0329
Iteration 2901, Loss: 0.0185, Time: 0.05s
Iteration 2902, Loss: 0.0121, Time: 0.05s
Iteration 2903, Loss: 0.0284, Time: 0.06s
Iteration 2904, Loss: 0.0175, Time: 0.05s
Iteration 2905, Loss: 0.0222, Time: 0.05s
Iteration 2906, Loss: 0.0303, Time: 0.04s
Iteration 2907, Loss: 0.0233, Time: 0.04s
Iteration 2908, Loss: 0.0250, Time: 0.04s
Iteration 2909, Loss: 0.0199, Time: 0.04s
Iteration 2910, Loss: 0.0424, Time: 0.05s
Iteration 2911, Loss: 0.0371, Time: 0.05s
Iteration 2912, Loss: 0.0307, Time: 0.05s
Iteration 2913, Loss: 0.0256, Time: 0.05s
Iteration 2914, Loss: 0.0232, Time: 0.05s
Iteration 2915, Loss: 0.0260, Time: 0.05s
Iteration 2916, Loss: 0.0219, Time: 0.06s
```

```
Iteration 2917, Loss: 0.0221, Time: 0.06s
Iteration 2918, Loss: 0.0173, Time: 0.04s
Iteration 2919, Loss: 0.0244, Time: 0.04s
Iteration 2920, Loss: 0.0249, Time: 0.05s
Iteration 2921, Loss: 0.0377, Time: 0.05s
Iteration 2922, Loss: 0.0254, Time: 0.04s
Iteration 2923, Loss: 0.0300, Time: 0.05s
Iteration 2924, Loss: 0.0231, Time: 0.05s
Iteration 2925, Loss: 0.0247, Time: 0.06s
Iteration 2926, Loss: 0.0245, Time: 0.04s
Iteration 2927, Loss: 0.0296, Time: 0.04s
Iteration 2928, Loss: 0.0330, Time: 0.04s
Iteration 2929, Loss: 0.0234, Time: 0.04s
Iteration 2930, Loss: 0.0216, Time: 0.05s
Iteration 2931, Loss: 0.0273, Time: 0.05s
Iteration 2932, Loss: 0.0258, Time: 0.04s
Iteration 2933, Loss: 0.0475, Time: 0.04s
Iteration 2934, Loss: 0.0305, Time: 0.04s
Iteration 2935, Loss: 0.0271, Time: 0.05s
Iteration 2936, Loss: 0.0389, Time: 0.05s
Iteration 2937, Loss: 0.0229, Time: 0.05s
Iteration 2938, Loss: 0.0244, Time: 0.05s
Iteration 2939, Loss: 0.0256, Time: 0.04s
Iteration 2940, Loss: 0.0195, Time: 0.05s
Iteration 2941, Loss: 0.0182, Time: 0.04s
Iteration 2942, Loss: 0.0196, Time: 0.05s
Iteration 2943, Loss: 0.0325, Time: 0.05s
Iteration 2944, Loss: 0.0256, Time: 0.05s
Iteration 2945, Loss: 0.0316, Time: 0.05s
Iteration 2946, Loss: 0.0224, Time: 0.06s
Iteration 2947, Loss: 0.0280, Time: 0.05s
Iteration 2948, Loss: 0.0255, Time: 0.05s
Iteration 2949, Loss: 0.0256, Time: 0.05s
Iteration 2950, Loss: 0.0344, Time: 0.05s
Iteration 2951, Loss: 0.0267, Time: 0.04s
Iteration 2952, Loss: 0.0177, Time: 0.04s
Iteration 2953, Loss: 0.0252, Time: 0.04s
Iteration 2954, Loss: 0.0217, Time: 0.04s
Iteration 2955, Loss: 0.0269, Time: 0.05s
Iteration 2956, Loss: 0.0214, Time: 0.04s
Iteration 2957, Loss: 0.0319, Time: 0.06s
Iteration 2958, Loss: 0.0325, Time: 0.05s
Iteration 2959, Loss: 0.0436, Time: 0.05s
Iteration 2960, Loss: 0.0280, Time: 0.05s
Iteration 2961, Loss: 0.0323, Time: 0.04s
Iteration 2962, Loss: 0.0168, Time: 0.05s
Iteration 2963, Loss: 0.0232, Time: 0.05s
Iteration 2964, Loss: 0.0258, Time: 0.05s
```

```
Iteration 2965, Loss: 0.0303, Time: 0.05s
Iteration 2966, Loss: 0.0399, Time: 0.04s
Iteration 2967, Loss: 0.0297, Time: 0.04s
Iteration 2968, Loss: 0.0157, Time: 0.04s
Iteration 2969, Loss: 0.0296, Time: 0.04s
Iteration 2970, Loss: 0.0231, Time: 0.04s
Iteration 2971, Loss: 0.0212, Time: 0.05s
Iteration 2972, Loss: 0.0307, Time: 0.04s
Iteration 2973, Loss: 0.0314, Time: 0.05s
Iteration 2974, Loss: 0.0244, Time: 0.05s
Iteration 2975, Loss: 0.0405, Time: 0.04s
Iteration 2976, Loss: 0.0234, Time: 0.04s
Iteration 2977, Loss: 0.0239, Time: 0.04s
Iteration 2978, Loss: 0.0216, Time: 0.04s
Iteration 2979, Loss: 0.0231, Time: 0.05s
Iteration 2980, Loss: 0.0191, Time: 0.04s
Iteration 2981, Loss: 0.0218, Time: 0.04s
Iteration 2982, Loss: 0.0363, Time: 0.04s
Iteration 2983, Loss: 0.0325, Time: 0.04s
Iteration 2984, Loss: 0.0189, Time: 0.05s
Iteration 2985, Loss: 0.0366, Time: 0.05s
Iteration 2986, Loss: 0.0191, Time: 0.05s
Iteration 2987, Loss: 0.0242, Time: 0.05s
Iteration 2988, Loss: 0.0264, Time: 0.05s
Iteration 2989, Loss: 0.0272, Time: 0.05s
Iteration 2990, Loss: 0.0334, Time: 0.05s
Iteration 2991, Loss: 0.0271, Time: 0.05s
Iteration 2992, Loss: 0.0261, Time: 0.04s
Iteration 2993, Loss: 0.0269, Time: 0.04s
Iteration 2994, Loss: 0.0293, Time: 0.04s
Iteration 2995, Loss: 0.0299, Time: 0.04s
Iteration 2996, Loss: 0.0370, Time: 0.05s
Iteration 2997, Loss: 0.0206, Time: 0.05s
Iteration 2998, Loss: 0.0201, Time: 0.05s
Iteration 2999, Loss: 0.0195, Time: 0.05s
Iteration 3000, Loss: 0.0237, Time: 0.04s
Iteration 3000, Loss: 0.0237, Time: 0.04s
Test Loss: 0.0346
Iteration 3001, Loss: 0.0192, Time: 0.04s
Iteration 3002, Loss: 0.0311, Time: 0.05s
Iteration 3003, Loss: 0.0337, Time: 0.05s
Iteration 3004, Loss: 0.0312, Time: 0.04s
Iteration 3005, Loss: 0.0249, Time: 0.04s
Iteration 3006, Loss: 0.0297, Time: 0.04s
Iteration 3007, Loss: 0.0204, Time: 0.04s
Iteration 3008, Loss: 0.0247, Time: 0.04s
Iteration 3009, Loss: 0.0256, Time: 0.05s
Iteration 3010, Loss: 0.0290, Time: 0.04s
```

```
Iteration 3011, Loss: 0.0341, Time: 0.04s
Iteration 3012, Loss: 0.0276, Time: 0.05s
Iteration 3013, Loss: 0.0345, Time: 0.05s
Iteration 3014, Loss: 0.0317, Time: 0.05s
Iteration 3015, Loss: 0.0370, Time: 0.04s
Iteration 3016, Loss: 0.0274, Time: 0.05s
Iteration 3017, Loss: 0.0214, Time: 0.05s
Iteration 3018, Loss: 0.0188, Time: 0.04s
Iteration 3019, Loss: 0.0327, Time: 0.04s
Iteration 3020, Loss: 0.0404, Time: 0.04s
Iteration 3021, Loss: 0.0414, Time: 0.04s
Iteration 3022, Loss: 0.0237, Time: 0.05s
Iteration 3023, Loss: 0.0208, Time: 0.04s
Iteration 3024, Loss: 0.0319, Time: 0.04s
Iteration 3025, Loss: 0.0367, Time: 0.05s
Iteration 3026, Loss: 0.0291, Time: 0.05s
Iteration 3027, Loss: 0.0346, Time: 0.04s
Iteration 3028, Loss: 0.0373, Time: 0.04s
Iteration 3029, Loss: 0.0298, Time: 0.04s
Iteration 3030, Loss: 0.0355, Time: 0.04s
Iteration 3031, Loss: 0.0171, Time: 0.05s
Iteration 3032, Loss: 0.0263, Time: 0.04s
Iteration 3033, Loss: 0.0195, Time: 0.04s
Iteration 3034, Loss: 0.0192, Time: 0.05s
Iteration 3035, Loss: 0.0247, Time: 0.04s
Iteration 3036, Loss: 0.0243, Time: 0.05s
Iteration 3037, Loss: 0.0173, Time: 0.05s
Iteration 3038, Loss: 0.0270, Time: 0.04s
Iteration 3039, Loss: 0.0248, Time: 0.04s
Iteration 3040, Loss: 0.0306, Time: 0.04s
Iteration 3041, Loss: 0.0365, Time: 0.04s
Iteration 3042, Loss: 0.0154, Time: 0.04s
Iteration 3043, Loss: 0.0205, Time: 0.04s
Iteration 3044, Loss: 0.0253, Time: 0.05s
Iteration 3045, Loss: 0.0249, Time: 0.05s
Iteration 3046, Loss: 0.0145, Time: 0.05s
Iteration 3047, Loss: 0.0277, Time: 0.05s
Iteration 3048, Loss: 0.0314, Time: 0.06s
Iteration 3049, Loss: 0.0296, Time: 0.06s
Iteration 3050, Loss: 0.0172, Time: 0.05s
Iteration 3051, Loss: 0.0284, Time: 0.06s
Iteration 3052, Loss: 0.0294, Time: 0.05s
Iteration 3053, Loss: 0.0449, Time: 0.05s
Iteration 3054, Loss: 0.0283, Time: 0.04s
Iteration 3055, Loss: 0.0209, Time: 0.05s
Iteration 3056, Loss: 0.0272, Time: 0.05s
Iteration 3057, Loss: 0.0426, Time: 0.04s
Iteration 3058, Loss: 0.0200, Time: 0.05s
```

```
Iteration 3059, Loss: 0.0234, Time: 0.05s
Iteration 3060, Loss: 0.0237, Time: 0.04s
Iteration 3061, Loss: 0.0370, Time: 0.05s
Iteration 3062, Loss: 0.0252, Time: 0.05s
Iteration 3063, Loss: 0.0213, Time: 0.05s
Iteration 3064, Loss: 0.0155, Time: 0.04s
Iteration 3065, Loss: 0.0279, Time: 0.04s
Iteration 3066, Loss: 0.0403, Time: 0.05s
Iteration 3067, Loss: 0.0300, Time: 0.05s
Iteration 3068, Loss: 0.0226, Time: 0.05s
Iteration 3069, Loss: 0.0179, Time: 0.05s
Iteration 3070, Loss: 0.0311, Time: 0.05s
Iteration 3071, Loss: 0.0248, Time: 0.04s
Iteration 3072, Loss: 0.0204, Time: 0.04s
Iteration 3073, Loss: 0.0253, Time: 0.04s
Iteration 3074, Loss: 0.0208, Time: 0.05s
Iteration 3075, Loss: 0.0279, Time: 0.05s
Iteration 3076, Loss: 0.0156, Time: 0.05s
Iteration 3077, Loss: 0.0209, Time: 0.05s
Iteration 3078, Loss: 0.0263, Time: 0.05s
Iteration 3079, Loss: 0.0258, Time: 0.06s
Iteration 3080, Loss: 0.0224, Time: 0.05s
Iteration 3081, Loss: 0.0379, Time: 0.05s
Iteration 3082, Loss: 0.0193, Time: 0.04s
Iteration 3083, Loss: 0.0317, Time: 0.05s
Iteration 3084, Loss: 0.0240, Time: 0.05s
Iteration 3085, Loss: 0.0241, Time: 0.05s
Iteration 3086, Loss: 0.0228, Time: 0.05s
Iteration 3087, Loss: 0.0255, Time: 0.04s
Iteration 3088, Loss: 0.0278, Time: 0.05s
Iteration 3089, Loss: 0.0169, Time: 0.05s
Iteration 3090, Loss: 0.0233, Time: 0.05s
Iteration 3091, Loss: 0.0162, Time: 0.06s
Iteration 3092, Loss: 0.0297, Time: 0.04s
Iteration 3093, Loss: 0.0255, Time: 0.05s
Iteration 3094, Loss: 0.0274, Time: 0.04s
Iteration 3095, Loss: 0.0252, Time: 0.04s
Iteration 3096, Loss: 0.0226, Time: 0.04s
Iteration 3097, Loss: 0.0342, Time: 0.05s
Iteration 3098, Loss: 0.0230, Time: 0.05s
Iteration 3099, Loss: 0.0167, Time: 0.04s
Iteration 3100, Loss: 0.0306, Time: 0.05s
Iteration 3100, Loss: 0.0306, Time: 0.05s
Test Loss: 0.0331
Iteration 3101, Loss: 0.0284, Time: 0.04s
Iteration 3102, Loss: 0.0301, Time: 0.04s
Iteration 3103, Loss: 0.0214, Time: 0.04s
Iteration 3104, Loss: 0.0280, Time: 0.04s
```

```
Iteration 3105, Loss: 0.0178, Time: 0.05s
Iteration 3106, Loss: 0.0218, Time: 0.04s
Iteration 3107, Loss: 0.0324, Time: 0.05s
Iteration 3108, Loss: 0.0340, Time: 0.04s
Iteration 3109, Loss: 0.0208, Time: 0.04s
Iteration 3110, Loss: 0.0349, Time: 0.05s
Iteration 3111, Loss: 0.0297, Time: 0.05s
Iteration 3112, Loss: 0.0249, Time: 0.05s
Iteration 3113, Loss: 0.0246, Time: 0.04s
Iteration 3114, Loss: 0.0223, Time: 0.06s
Iteration 3115, Loss: 0.0327, Time: 0.05s
Iteration 3116, Loss: 0.0191, Time: 0.04s
Iteration 3117, Loss: 0.0226, Time: 0.05s
Iteration 3118, Loss: 0.0179, Time: 0.04s
Iteration 3119, Loss: 0.0240, Time: 0.04s
Iteration 3120, Loss: 0.0235, Time: 0.04s
Iteration 3121, Loss: 0.0237, Time: 0.04s
Iteration 3122, Loss: 0.0257, Time: 0.05s
Iteration 3123, Loss: 0.0269, Time: 0.04s
Iteration 3124, Loss: 0.0315, Time: 0.06s
Iteration 3125, Loss: 0.0318, Time: 0.06s
Iteration 3126, Loss: 0.0151, Time: 0.07s
Iteration 3127, Loss: 0.0243, Time: 0.05s
Iteration 3128, Loss: 0.0324, Time: 0.05s
Iteration 3129, Loss: 0.0187, Time: 0.05s
Iteration 3130, Loss: 0.0243, Time: 0.05s
Iteration 3131, Loss: 0.0172, Time: 0.04s
Iteration 3132, Loss: 0.0284, Time: 0.04s
Iteration 3133, Loss: 0.0375, Time: 0.04s
Iteration 3134, Loss: 0.0301, Time: 0.04s
Iteration 3135, Loss: 0.0266, Time: 0.04s
Iteration 3136, Loss: 0.0192, Time: 0.04s
Iteration 3137, Loss: 0.0309, Time: 0.04s
Iteration 3138, Loss: 0.0190, Time: 0.05s
Iteration 3139, Loss: 0.0279, Time: 0.04s
Iteration 3140, Loss: 0.0198, Time: 0.05s
Iteration 3141, Loss: 0.0334, Time: 0.04s
Iteration 3142, Loss: 0.0223, Time: 0.04s
Iteration 3143, Loss: 0.0269, Time: 0.05s
Iteration 3144, Loss: 0.0242, Time: 0.04s
Iteration 3145, Loss: 0.0233, Time: 0.05s
Iteration 3146, Loss: 0.0296, Time: 0.04s
Iteration 3147, Loss: 0.0271, Time: 0.04s
Iteration 3148, Loss: 0.0295, Time: 0.05s
Iteration 3149, Loss: 0.0263, Time: 0.05s
Iteration 3150, Loss: 0.0230, Time: 0.05s
Iteration 3151, Loss: 0.0227, Time: 0.05s
Iteration 3152, Loss: 0.0272, Time: 0.04s
```

```
Iteration 3153, Loss: 0.0326, Time: 0.05s
Iteration 3154, Loss: 0.0260, Time: 0.05s
Iteration 3155, Loss: 0.0293, Time: 0.05s
Iteration 3156, Loss: 0.0264, Time: 0.04s
Iteration 3157, Loss: 0.0220, Time: 0.04s
Iteration 3158, Loss: 0.0335, Time: 0.04s
Iteration 3159, Loss: 0.0186, Time: 0.04s
Iteration 3160, Loss: 0.0187, Time: 0.05s
Iteration 3161, Loss: 0.0294, Time: 0.04s
Iteration 3162, Loss: 0.0196, Time: 0.05s
Iteration 3163, Loss: 0.0274, Time: 0.05s
Iteration 3164, Loss: 0.0143, Time: 0.05s
Iteration 3165, Loss: 0.0247, Time: 0.04s
Iteration 3166, Loss: 0.0201, Time: 0.05s
Iteration 3167, Loss: 0.0265, Time: 0.05s
Iteration 3168, Loss: 0.0236, Time: 0.04s
Iteration 3169, Loss: 0.0202, Time: 0.05s
Iteration 3170, Loss: 0.0194, Time: 0.05s
Iteration 3171, Loss: 0.0356, Time: 0.05s
Iteration 3172, Loss: 0.0113, Time: 0.04s
Iteration 3173, Loss: 0.0244, Time: 0.04s
Iteration 3174, Loss: 0.0207, Time: 0.05s
Iteration 3175, Loss: 0.0207, Time: 0.05s
Iteration 3176, Loss: 0.0302, Time: 0.04s
Iteration 3177, Loss: 0.0237, Time: 0.04s
Iteration 3178, Loss: 0.0320, Time: 0.04s
Iteration 3179, Loss: 0.0185, Time: 0.05s
Iteration 3180, Loss: 0.0269, Time: 0.06s
Iteration 3181, Loss: 0.0235, Time: 0.06s
Iteration 3182, Loss: 0.0224, Time: 0.06s
Iteration 3183, Loss: 0.0240, Time: 0.05s
Iteration 3184, Loss: 0.0373, Time: 0.05s
Iteration 3185, Loss: 0.0255, Time: 0.04s
Iteration 3186, Loss: 0.0250, Time: 0.05s
Iteration 3187, Loss: 0.0155, Time: 0.05s
Iteration 3188, Loss: 0.0269, Time: 0.05s
Iteration 3189, Loss: 0.0203, Time: 0.05s
Iteration 3190, Loss: 0.0291, Time: 0.04s
Iteration 3191, Loss: 0.0168, Time: 0.04s
Iteration 3192, Loss: 0.0205, Time: 0.04s
Iteration 3193, Loss: 0.0242, Time: 0.05s
Iteration 3194, Loss: 0.0185, Time: 0.05s
Iteration 3195, Loss: 0.0275, Time: 0.05s
Iteration 3196, Loss: 0.0204, Time: 0.04s
Iteration 3197, Loss: 0.0221, Time: 0.04s
Iteration 3198, Loss: 0.0183, Time: 0.04s
Iteration 3199, Loss: 0.0303, Time: 0.05s
Iteration 3200, Loss: 0.0255, Time: 0.04s
```

```
Iteration 3200, Loss: 0.0255, Time: 0.04s
Test Loss: 0.0435
Iteration 3201, Loss: 0.0144, Time: 0.04s
Iteration 3202, Loss: 0.0232, Time: 0.05s
Iteration 3203, Loss: 0.0164, Time: 0.05s
Iteration 3204, Loss: 0.0306, Time: 0.05s
Iteration 3205, Loss: 0.0251, Time: 0.04s
Iteration 3206, Loss: 0.0229, Time: 0.04s
Iteration 3207, Loss: 0.0234, Time: 0.05s
Iteration 3208, Loss: 0.0256, Time: 0.05s
Iteration 3209, Loss: 0.0210, Time: 0.05s
Iteration 3210, Loss: 0.0260, Time: 0.05s
Iteration 3211, Loss: 0.0303, Time: 0.05s
Iteration 3212, Loss: 0.0204, Time: 0.05s
Iteration 3213, Loss: 0.0302, Time: 0.04s
Iteration 3214, Loss: 0.0235, Time: 0.07s
Iteration 3215, Loss: 0.0392, Time: 0.07s
Iteration 3216, Loss: 0.0251, Time: 0.07s
Iteration 3217, Loss: 0.0233, Time: 0.07s
Iteration 3218, Loss: 0.0197, Time: 0.06s
Iteration 3219, Loss: 0.0176, Time: 0.05s
Iteration 3220, Loss: 0.0236, Time: 0.04s
Iteration 3221, Loss: 0.0156, Time: 0.05s
Iteration 3222, Loss: 0.0333, Time: 0.05s
Iteration 3223, Loss: 0.0214, Time: 0.05s
Iteration 3224, Loss: 0.0186, Time: 0.05s
Iteration 3225, Loss: 0.0164, Time: 0.04s
Iteration 3226, Loss: 0.0341, Time: 0.04s
Iteration 3227, Loss: 0.0261, Time: 0.04s
Iteration 3228, Loss: 0.0204, Time: 0.05s
Iteration 3229, Loss: 0.0360, Time: 0.04s
Iteration 3230, Loss: 0.0157, Time: 0.04s
Iteration 3231, Loss: 0.0279, Time: 0.05s
Iteration 3232, Loss: 0.0192, Time: 0.04s
Iteration 3233, Loss: 0.0295, Time: 0.04s
Iteration 3234, Loss: 0.0267, Time: 0.04s
Iteration 3235, Loss: 0.0173, Time: 0.04s
Iteration 3236, Loss: 0.0133, Time: 0.05s
Iteration 3237, Loss: 0.0277, Time: 0.04s
Iteration 3238, Loss: 0.0214, Time: 0.04s
Iteration 3239, Loss: 0.0174, Time: 0.04s
Iteration 3240, Loss: 0.0217, Time: 0.05s
Iteration 3241, Loss: 0.0199, Time: 0.04s
Iteration 3242, Loss: 0.0235, Time: 0.05s
Iteration 3243, Loss: 0.0295, Time: 0.04s
Iteration 3244, Loss: 0.0176, Time: 0.05s
Iteration 3245, Loss: 0.0207, Time: 0.04s
Iteration 3246, Loss: 0.0283, Time: 0.05s
```

```
Iteration 3247, Loss: 0.0227, Time: 0.04s
Iteration 3248, Loss: 0.0254, Time: 0.04s
Iteration 3249, Loss: 0.0330, Time: 0.05s
Iteration 3250, Loss: 0.0248, Time: 0.05s
Iteration 3251, Loss: 0.0210, Time: 0.05s
Iteration 3252, Loss: 0.0222, Time: 0.04s
Iteration 3253, Loss: 0.0220, Time: 0.05s
Iteration 3254, Loss: 0.0263, Time: 0.05s
Iteration 3255, Loss: 0.0328, Time: 0.04s
Iteration 3256, Loss: 0.0108, Time: 0.05s
Iteration 3257, Loss: 0.0375, Time: 0.05s
Iteration 3258, Loss: 0.0286, Time: 0.07s
Iteration 3259, Loss: 0.0200, Time: -0.08s
Iteration 3260, Loss: 0.0283, Time: 0.06s
Iteration 3261, Loss: 0.0239, Time: 0.05s
Iteration 3262, Loss: 0.0205, Time: 0.04s
Iteration 3263, Loss: 0.0208, Time: 0.05s
Iteration 3264, Loss: 0.0230, Time: 0.05s
Iteration 3265, Loss: 0.0205, Time: 0.04s
Iteration 3266, Loss: 0.0289, Time: 0.04s
Iteration 3267, Loss: 0.0276, Time: 0.04s
Iteration 3268, Loss: 0.0149, Time: 0.04s
Iteration 3269, Loss: 0.0201, Time: 0.04s
Iteration 3270, Loss: 0.0316, Time: 0.04s
Iteration 3271, Loss: 0.0304, Time: 0.04s
Iteration 3272, Loss: 0.0247, Time: 0.04s
Iteration 3273, Loss: 0.0215, Time: 0.04s
Iteration 3274, Loss: 0.0193, Time: 0.04s
Iteration 3275, Loss: 0.0330, Time: 0.04s
Iteration 3276, Loss: 0.0182, Time: 0.04s
Iteration 3277, Loss: 0.0345, Time: 0.05s
Iteration 3278, Loss: 0.0202, Time: 0.04s
Iteration 3279, Loss: 0.0209, Time: 0.04s
Iteration 3280, Loss: 0.0198, Time: 0.04s
Iteration 3281, Loss: 0.0155, Time: 0.04s
Iteration 3282, Loss: 0.0251, Time: 0.05s
Iteration 3283, Loss: 0.0214, Time: 0.04s
Iteration 3284, Loss: 0.0230, Time: 0.04s
Iteration 3285, Loss: 0.0248, Time: 0.04s
Iteration 3286, Loss: 0.0165, Time: 0.04s
Iteration 3287, Loss: 0.0154, Time: 0.04s
Iteration 3288, Loss: 0.0279, Time: 0.04s
Iteration 3289, Loss: 0.0345, Time: 0.04s
Iteration 3290, Loss: 0.0242, Time: 0.04s
Iteration 3291, Loss: 0.0360, Time: 0.04s
Iteration 3292, Loss: 0.0129, Time: 0.04s
Iteration 3293, Loss: 0.0222, Time: 0.04s
Iteration 3294, Loss: 0.0246, Time: 0.04s
```

```
Iteration 3295, Loss: 0.0270, Time: 0.05s
Iteration 3296, Loss: 0.0223, Time: 0.04s
Iteration 3297, Loss: 0.0207, Time: 0.04s
Iteration 3298, Loss: 0.0234, Time: 0.04s
Iteration 3299, Loss: 0.0199, Time: 0.04s
Iteration 3300, Loss: 0.0203, Time: 0.04s
Iteration 3300, Loss: 0.0203, Time: 0.04s
Test Loss: 0.0414
Iteration 3301, Loss: 0.0207, Time: 0.04s
Iteration 3302, Loss: 0.0232, Time: 0.05s
Iteration 3303, Loss: 0.0198, Time: 0.04s
Iteration 3304, Loss: 0.0250, Time: 0.05s
Iteration 3305, Loss: 0.0349, Time: 0.04s
Iteration 3306, Loss: 0.0306, Time: 0.05s
Iteration 3307, Loss: 0.0294, Time: 0.05s
Iteration 3308, Loss: 0.0224, Time: 0.05s
Iteration 3309, Loss: 0.0250, Time: 0.04s
Iteration 3310, Loss: 0.0313, Time: 0.06s
Iteration 3311, Loss: 0.0236, Time: 0.05s
Iteration 3312, Loss: 0.0156, Time: 0.06s
Iteration 3313, Loss: 0.0201, Time: 0.04s
Iteration 3314, Loss: 0.0183, Time: 0.04s
Iteration 3315, Loss: 0.0204, Time: 0.04s
Iteration 3316, Loss: 0.0384, Time: 0.04s
Iteration 3317, Loss: 0.0274, Time: 0.04s
Iteration 3318, Loss: 0.0238, Time: 0.05s
Iteration 3319, Loss: 0.0242, Time: 0.04s
Iteration 3320, Loss: 0.0210, Time: 0.04s
Iteration 3321, Loss: 0.0302, Time: 0.04s
Iteration 3322, Loss: 0.0166, Time: 0.04s
Iteration 3323, Loss: 0.0198, Time: 0.04s
Iteration 3324, Loss: 0.0226, Time: 0.05s
Iteration 3325, Loss: 0.0242, Time: 0.04s
Iteration 3326, Loss: 0.0195, Time: 0.04s
Iteration 3327, Loss: 0.0290, Time: 0.04s
Iteration 3328, Loss: 0.0128, Time: 0.04s
Iteration 3329, Loss: 0.0149, Time: 0.04s
Iteration 3330, Loss: 0.0156, Time: 0.04s
Iteration 3331, Loss: 0.0124, Time: 0.05s
Iteration 3332, Loss: 0.0255, Time: 0.04s
Iteration 3333, Loss: 0.0258, Time: 0.04s
Iteration 3334, Loss: 0.0174, Time: 0.04s
Iteration 3335, Loss: 0.0349, Time: 0.04s
Iteration 3336, Loss: 0.0416, Time: 0.04s
Iteration 3337, Loss: 0.0264, Time: 0.04s
Iteration 3338, Loss: 0.0246, Time: 0.05s
Iteration 3339, Loss: 0.0337, Time: 0.04s
Iteration 3340, Loss: 0.0282, Time: 0.04s
```

```
Iteration 3341, Loss: 0.0308, Time: 0.04s
Iteration 3342, Loss: 0.0155, Time: 0.04s
Iteration 3343, Loss: 0.0225, Time: 0.05s
Iteration 3344, Loss: 0.0252, Time: 0.04s
Iteration 3345, Loss: 0.0309, Time: 0.04s
Iteration 3346, Loss: 0.0253, Time: 0.04s
Iteration 3347, Loss: 0.0214, Time: 0.04s
Iteration 3348, Loss: 0.0189, Time: 0.04s
Iteration 3349, Loss: 0.0408, Time: 0.05s
Iteration 3350, Loss: 0.0318, Time: 0.05s
Iteration 3351, Loss: 0.0284, Time: 0.04s
Iteration 3352, Loss: 0.0144, Time: 0.05s
Iteration 3353, Loss: 0.0238, Time: 0.05s
Iteration 3354, Loss: 0.0229, Time: 0.04s
Iteration 3355, Loss: 0.0207, Time: 0.04s
Iteration 3356, Loss: 0.0213, Time: 0.06s
Iteration 3357, Loss: 0.0323, Time: 0.05s
Iteration 3358, Loss: 0.0275, Time: 0.04s
Iteration 3359, Loss: 0.0238, Time: 0.05s
Iteration 3360, Loss: 0.0321, Time: 0.04s
Iteration 3361, Loss: 0.0320, Time: 0.04s
Iteration 3362, Loss: 0.0124, Time: 0.04s
Iteration 3363, Loss: 0.0230, Time: 0.05s
Iteration 3364, Loss: 0.0176, Time: 0.04s
Iteration 3365, Loss: 0.0285, Time: 0.04s
Iteration 3366, Loss: 0.0195, Time: 0.05s
Iteration 3367, Loss: 0.0184, Time: 0.05s
Iteration 3368, Loss: 0.0345, Time: 0.04s
Iteration 3369, Loss: 0.0221, Time: 0.04s
Iteration 3370, Loss: 0.0166, Time: 0.04s
Iteration 3371, Loss: 0.0269, Time: 0.04s
Iteration 3372, Loss: 0.0268, Time: 0.04s
Iteration 3373, Loss: 0.0186, Time: 0.04s
Iteration 3374, Loss: 0.0191, Time: 0.04s
Iteration 3375, Loss: 0.0203, Time: 0.05s
Iteration 3376, Loss: 0.0306, Time: 0.04s
Iteration 3377, Loss: 0.0197, Time: 0.05s
Iteration 3378, Loss: 0.0368, Time: 0.04s
Iteration 3379, Loss: 0.0276, Time: 0.04s
Iteration 3380, Loss: 0.0301, Time: 0.04s
Iteration 3381, Loss: 0.0298, Time: 0.05s
Iteration 3382, Loss: 0.0284, Time: 0.04s
Iteration 3383, Loss: 0.0259, Time: 0.04s
Iteration 3384, Loss: 0.0219, Time: 0.04s
Iteration 3385, Loss: 0.0206, Time: 0.04s
Iteration 3386, Loss: 0.0215, Time: 0.04s
Iteration 3387, Loss: 0.0250, Time: 0.05s
Iteration 3388, Loss: 0.0174, Time: 0.04s
```

```
Iteration 3389, Loss: 0.0226, Time: 0.04s
Iteration 3390, Loss: 0.0309, Time: 0.04s
Iteration 3391, Loss: 0.0286, Time: 0.04s
Iteration 3392, Loss: 0.0223, Time: 0.04s
Iteration 3393, Loss: 0.0227, Time: 0.04s
Iteration 3394, Loss: 0.0188, Time: 0.04s
Iteration 3395, Loss: 0.0321, Time: 0.04s
Iteration 3396, Loss: 0.0176, Time: 0.04s
Iteration 3397, Loss: 0.0236, Time: 0.05s
Iteration 3398, Loss: 0.0152, Time: 0.04s
Iteration 3399, Loss: 0.0148, Time: 0.05s
Iteration 3400, Loss: 0.0325, Time: 0.04s
Iteration 3400, Loss: 0.0325, Time: 0.04s
Test Loss: 0.0382
Iteration 3401, Loss: 0.0281, Time: 0.04s
Iteration 3402, Loss: 0.0349, Time: 0.04s
Iteration 3403, Loss: 0.0247, Time: 0.04s
Iteration 3404, Loss: 0.0341, Time: 0.04s
Iteration 3405, Loss: 0.0214, Time: 0.04s
Iteration 3406, Loss: 0.0133, Time: 0.04s
Iteration 3407, Loss: 0.0174, Time: 0.04s
Iteration 3408, Loss: 0.0235, Time: 0.04s
Iteration 3409, Loss: 0.0268, Time: 0.04s
Iteration 3410, Loss: 0.0324, Time: 0.04s
Iteration 3411, Loss: 0.0170, Time: 0.04s
Iteration 3412, Loss: 0.0212, Time: 0.04s
Iteration 3413, Loss: 0.0141, Time: 0.04s
Iteration 3414, Loss: 0.0321, Time: 0.04s
Iteration 3415, Loss: 0.0399, Time: 0.05s
Iteration 3416, Loss: 0.0195, Time: 0.05s
Iteration 3417, Loss: 0.0265, Time: 0.04s
Iteration 3418, Loss: 0.0181, Time: 0.04s
Iteration 3419, Loss: 0.0312, Time: 0.05s
Iteration 3420, Loss: 0.0240, Time: 0.04s
Iteration 3421, Loss: 0.0222, Time: 0.05s
Iteration 3422, Loss: 0.0295, Time: 0.04s
Iteration 3423, Loss: 0.0211, Time: 0.05s
Iteration 3424, Loss: 0.0150, Time: 0.04s
Iteration 3425, Loss: 0.0251, Time: 0.05s
Iteration 3426, Loss: 0.0223, Time: 0.04s
Iteration 3427, Loss: 0.0213, Time: 0.04s
Iteration 3428, Loss: 0.0255, Time: 0.04s
Iteration 3429, Loss: 0.0242, Time: 0.04s
Iteration 3430, Loss: 0.0233, Time: 0.04s
Iteration 3431, Loss: 0.0266, Time: 0.05s
Iteration 3432, Loss: 0.0282, Time: 0.04s
Iteration 3433, Loss: 0.0195, Time: 0.04s
Iteration 3434, Loss: 0.0145, Time: 0.04s
```

```
Iteration 3435, Loss: 0.0249, Time: 0.05s
Iteration 3436, Loss: 0.0368, Time: 0.04s
Iteration 3437, Loss: 0.0225, Time: 0.04s
Iteration 3438, Loss: 0.0297, Time: 0.04s
Iteration 3439, Loss: 0.0340, Time: 0.05s
Iteration 3440, Loss: 0.0282, Time: 0.04s
Iteration 3441, Loss: 0.0228, Time: 0.05s
Iteration 3442, Loss: 0.0209, Time: 0.05s
Iteration 3443, Loss: 0.0115, Time: 0.04s
Iteration 3444, Loss: 0.0283, Time: 0.04s
Iteration 3445, Loss: 0.0210, Time: 0.05s
Iteration 3446, Loss: 0.0219, Time: 0.04s
Iteration 3447, Loss: 0.0348, Time: 0.04s
Iteration 3448, Loss: 0.0145, Time: 0.04s
Iteration 3449, Loss: 0.0227, Time: 0.05s
Iteration 3450, Loss: 0.0194, Time: 0.05s
Iteration 3451, Loss: 0.0345, Time: 0.05s
Iteration 3452, Loss: 0.0163, Time: 0.04s
Iteration 3453, Loss: 0.0198, Time: 0.04s
Iteration 3454, Loss: 0.0208, Time: 0.04s
Iteration 3455, Loss: 0.0229, Time: 0.04s
Iteration 3456, Loss: 0.0232, Time: 0.05s
Iteration 3457, Loss: 0.0367, Time: 0.04s
Iteration 3458, Loss: 0.0264, Time: 0.05s
Iteration 3459, Loss: 0.0182, Time: 0.04s
Iteration 3460, Loss: 0.0205, Time: 0.04s
Iteration 3461, Loss: 0.0172, Time: 0.05s
Iteration 3462, Loss: 0.0207, Time: 0.04s
Iteration 3463, Loss: 0.0344, Time: 0.04s
Iteration 3464, Loss: 0.0214, Time: 0.05s
Iteration 3465, Loss: 0.0162, Time: 0.04s
Iteration 3466, Loss: 0.0318, Time: 0.05s
Iteration 3467, Loss: 0.0250, Time: 0.05s
Iteration 3468, Loss: 0.0197, Time: 0.04s
Iteration 3469, Loss: 0.0129, Time: 0.04s
Iteration 3470, Loss: 0.0199, Time: 0.05s
Iteration 3471, Loss: 0.0118, Time: 0.04s
Iteration 3472, Loss: 0.0226, Time: 0.04s
Iteration 3473, Loss: 0.0298, Time: 0.05s
Iteration 3474, Loss: 0.0218, Time: 0.04s
Iteration 3475, Loss: 0.0344, Time: 0.04s
Iteration 3476, Loss: 0.0227, Time: 0.04s
Iteration 3477, Loss: 0.0230, Time: 0.04s
Iteration 3478, Loss: 0.0252, Time: 0.04s
Iteration 3479, Loss: 0.0268, Time: 0.04s
Iteration 3480, Loss: 0.0213, Time: 0.04s
Iteration 3481, Loss: 0.0231, Time: 0.05s
Iteration 3482, Loss: 0.0313, Time: 0.05s
```

```
Iteration 3483, Loss: 0.0194, Time: 0.05s
Iteration 3484, Loss: 0.0269, Time: 0.05s
Iteration 3485, Loss: 0.0189, Time: 0.05s
Iteration 3486, Loss: 0.0214, Time: 0.04s
Iteration 3487, Loss: 0.0184, Time: 0.04s
Iteration 3488, Loss: 0.0265, Time: 0.05s
Iteration 3489, Loss: 0.0233, Time: 0.05s
Iteration 3490, Loss: 0.0242, Time: 0.05s
Iteration 3491, Loss: 0.0158, Time: 0.04s
Iteration 3492, Loss: 0.0293, Time: 0.05s
Iteration 3493, Loss: 0.0237, Time: 0.04s
Iteration 3494, Loss: 0.0313, Time: 0.04s
Iteration 3495, Loss: 0.0195, Time: 0.05s
Iteration 3496, Loss: 0.0177, Time: 0.04s
Iteration 3497, Loss: 0.0263, Time: 0.04s
Iteration 3498, Loss: 0.0241, Time: 0.04s
Iteration 3499, Loss: 0.0280, Time: 0.04s
Iteration 3500, Loss: 0.0245, Time: 0.04s
Iteration 3500, Loss: 0.0245, Time: 0.04s
Test Loss: 0.0299
Iteration 3501, Loss: 0.0229, Time: 0.05s
Iteration 3502, Loss: 0.0267, Time: 0.04s
Iteration 3503, Loss: 0.0311, Time: 0.04s
Iteration 3504, Loss: 0.0130, Time: 0.04s
Iteration 3505, Loss: 0.0193, Time: 0.05s
Iteration 3506, Loss: 0.0215, Time: 0.04s
Iteration 3507, Loss: 0.0237, Time: 0.04s
Iteration 3508, Loss: 0.0286, Time: 0.05s
Iteration 3509, Loss: 0.0228, Time: 0.04s
Iteration 3510, Loss: 0.0235, Time: 0.04s
Iteration 3511, Loss: 0.0300, Time: 0.04s
Iteration 3512, Loss: 0.0224, Time: 0.05s
Iteration 3513, Loss: 0.0352, Time: 0.05s
Iteration 3514, Loss: 0.0216, Time: 0.04s
Iteration 3515, Loss: 0.0192, Time: 0.05s
Iteration 3516, Loss: 0.0245, Time: 0.04s
Iteration 3517, Loss: 0.0241, Time: 0.05s
Iteration 3518, Loss: 0.0254, Time: 0.04s
Iteration 3519, Loss: 0.0282, Time: 0.05s
Iteration 3520, Loss: 0.0246, Time: 0.04s
Iteration 3521, Loss: 0.0220, Time: 0.04s
Iteration 3522, Loss: 0.0175, Time: 0.04s
Iteration 3523, Loss: 0.0224, Time: 0.05s
Iteration 3524, Loss: 0.0214, Time: 0.04s
Iteration 3525, Loss: 0.0262, Time: 0.04s
Iteration 3526, Loss: 0.0339, Time: 0.04s
Iteration 3527, Loss: 0.0148, Time: 0.04s
Iteration 3528, Loss: 0.0219, Time: 0.05s
```

```
Iteration 3529, Loss: 0.0200, Time: 0.04s
Iteration 3530, Loss: 0.0282, Time: 0.05s
Iteration 3531, Loss: 0.0256, Time: 0.05s
Iteration 3532, Loss: 0.0210, Time: 0.04s
Iteration 3533, Loss: 0.0216, Time: 0.04s
Iteration 3534, Loss: 0.0150, Time: 0.04s
Iteration 3535, Loss: 0.0218, Time: 0.05s
Iteration 3536, Loss: 0.0230, Time: 0.04s
Iteration 3537, Loss: 0.0212, Time: 0.04s
Iteration 3538, Loss: 0.0220, Time: 0.04s
Iteration 3539, Loss: 0.0143, Time: 0.05s
Iteration 3540, Loss: 0.0280, Time: 0.04s
Iteration 3541, Loss: 0.0188, Time: 0.05s
Iteration 3542, Loss: 0.0139, Time: 0.05s
Iteration 3543, Loss: 0.0201, Time: 0.05s
Iteration 3544, Loss: 0.0336, Time: 0.04s
Iteration 3545, Loss: 0.0229, Time: 0.05s
Iteration 3546, Loss: 0.0190, Time: 0.05s
Iteration 3547, Loss: 0.0195, Time: 0.05s
Iteration 3548, Loss: 0.0179, Time: 0.04s
Iteration 3549, Loss: 0.0170, Time: 0.04s
Iteration 3550, Loss: 0.0232, Time: 0.06s
Iteration 3551, Loss: 0.0218, Time: 0.07s
Iteration 3552, Loss: 0.0219, Time: 0.05s
Iteration 3553, Loss: 0.0228, Time: 0.05s
Iteration 3554, Loss: 0.0167, Time: 0.05s
Iteration 3555, Loss: 0.0182, Time: 0.05s
Iteration 3556, Loss: 0.0209, Time: 0.04s
Iteration 3557, Loss: 0.0190, Time: 0.07s
Iteration 3558, Loss: 0.0263, Time: 0.06s
Iteration 3559, Loss: 0.0176, Time: 0.05s
Iteration 3560, Loss: 0.0230, Time: 0.05s
Iteration 3561, Loss: 0.0202, Time: 0.04s
Iteration 3562, Loss: 0.0171, Time: 0.06s
Iteration 3563, Loss: 0.0234, Time: 0.04s
Iteration 3564, Loss: 0.0194, Time: 0.04s
Iteration 3565, Loss: 0.0190, Time: 0.05s
Iteration 3566, Loss: 0.0252, Time: 0.04s
Iteration 3567, Loss: 0.0204, Time: 0.05s
Iteration 3568, Loss: 0.0198, Time: 0.05s
Iteration 3569, Loss: 0.0239, Time: 0.04s
Iteration 3570, Loss: 0.0171, Time: 0.04s
Iteration 3571, Loss: 0.0200, Time: 0.05s
Iteration 3572, Loss: 0.0201, Time: 0.04s
Iteration 3573, Loss: 0.0177, Time: 0.05s
Iteration 3574, Loss: 0.0260, Time: 0.04s
Iteration 3575, Loss: 0.0180, Time: 0.05s
Iteration 3576, Loss: 0.0248, Time: 0.05s
```

```
Iteration 3577, Loss: 0.0240, Time: 0.05s
Iteration 3578, Loss: 0.0247, Time: 0.04s
Iteration 3579, Loss: 0.0233, Time: 0.05s
Iteration 3580, Loss: 0.0156, Time: 0.05s
Iteration 3581, Loss: 0.0215, Time: 0.06s
Iteration 3582, Loss: 0.0234, Time: 0.05s
Iteration 3583, Loss: 0.0214, Time: 0.05s
Iteration 3584, Loss: 0.0243, Time: 0.05s
Iteration 3585, Loss: 0.0162, Time: 0.05s
Iteration 3586, Loss: 0.0252, Time: 0.05s
Iteration 3587, Loss: 0.0232, Time: 0.04s
Iteration 3588, Loss: 0.0192, Time: 0.05s
Iteration 3589, Loss: 0.0195, Time: 0.05s
Iteration 3590, Loss: 0.0180, Time: 0.05s
Iteration 3591, Loss: 0.0158, Time: 0.04s
Iteration 3592, Loss: 0.0284, Time: 0.05s
Iteration 3593, Loss: 0.0223, Time: 0.04s
Iteration 3594, Loss: 0.0166, Time: 0.05s
Iteration 3595, Loss: 0.0221, Time: 0.06s
Iteration 3596, Loss: 0.0190, Time: 0.04s
Iteration 3597, Loss: 0.0202, Time: 0.04s
Iteration 3598, Loss: 0.0239, Time: 0.04s
Iteration 3599, Loss: 0.0242, Time: 0.04s
Iteration 3600, Loss: 0.0310, Time: 0.04s
Iteration 3600, Loss: 0.0310, Time: 0.04s
Test Loss: 0.0336
Iteration 3601, Loss: 0.0185, Time: 0.04s
Iteration 3602, Loss: 0.0144, Time: 0.05s
Iteration 3603, Loss: 0.0269, Time: 0.04s
Iteration 3604, Loss: 0.0202, Time: 0.05s
Iteration 3605, Loss: 0.0209, Time: 0.05s
Iteration 3606, Loss: 0.0143, Time: 0.04s
Iteration 3607, Loss: 0.0183, Time: 0.05s
Iteration 3608, Loss: 0.0204, Time: 0.05s
Iteration 3609, Loss: 0.0226, Time: 0.05s
Iteration 3610, Loss: 0.0104, Time: 0.05s
Iteration 3611, Loss: 0.0208, Time: 0.04s
Iteration 3612, Loss: 0.0131, Time: 0.04s
Iteration 3613, Loss: 0.0176, Time: 0.04s
Iteration 3614, Loss: 0.0146, Time: 0.04s
Iteration 3615, Loss: 0.0218, Time: 0.04s
Iteration 3616, Loss: 0.0199, Time: 0.05s
Iteration 3617, Loss: 0.0200, Time: 0.04s
Iteration 3618, Loss: 0.0151, Time: 0.04s
Iteration 3619, Loss: 0.0130, Time: 0.04s
Iteration 3620, Loss: 0.0153, Time: 0.04s
Iteration 3621, Loss: 0.0192, Time: 0.04s
Iteration 3622, Loss: 0.0290, Time: 0.04s
```

```
Iteration 3623, Loss: 0.0130, Time: 0.05s
Iteration 3624, Loss: 0.0223, Time: 0.05s
Iteration 3625, Loss: 0.0227, Time: 0.04s
Iteration 3626, Loss: 0.0214, Time: 0.04s
Iteration 3627, Loss: 0.0121, Time: 0.05s
Iteration 3628, Loss: 0.0176, Time: 0.04s
Iteration 3629, Loss: 0.0147, Time: 0.04s
Iteration 3630, Loss: 0.0153, Time: 0.04s
Iteration 3631, Loss: 0.0166, Time: 0.04s
Iteration 3632, Loss: 0.0189, Time: 0.05s
Iteration 3633, Loss: 0.0241, Time: 0.07s
Iteration 3634, Loss: 0.0201, Time: 0.06s
Iteration 3635, Loss: 0.0157, Time: 0.06s
Iteration 3636, Loss: 0.0160, Time: 0.06s
Iteration 3637, Loss: 0.0175, Time: 0.05s
Iteration 3638, Loss: 0.0232, Time: 0.05s
Iteration 3639, Loss: 0.0141, Time: 0.05s
Iteration 3640, Loss: 0.0206, Time: 0.04s
Iteration 3641, Loss: 0.0269, Time: 0.04s
Iteration 3642, Loss: 0.0210, Time: 0.04s
Iteration 3643, Loss: 0.0183, Time: 0.05s
Iteration 3644, Loss: 0.0211, Time: 0.05s
Iteration 3645, Loss: 0.0201, Time: 0.05s
Iteration 3646, Loss: 0.0286, Time: 0.05s
Iteration 3647, Loss: 0.0232, Time: 0.05s
Iteration 3648, Loss: 0.0172, Time: 0.04s
Iteration 3649, Loss: 0.0112, Time: 0.05s
Iteration 3650, Loss: 0.0209, Time: 0.04s
Iteration 3651, Loss: 0.0240, Time: 0.05s
Iteration 3652, Loss: 0.0203, Time: 0.04s
Iteration 3653, Loss: 0.0188, Time: 0.05s
Iteration 3654, Loss: 0.0171, Time: 0.04s
Iteration 3655, Loss: 0.0166, Time: 0.05s
Iteration 3656, Loss: 0.0175, Time: 0.04s
Iteration 3657, Loss: 0.0153, Time: 0.04s
Iteration 3658, Loss: 0.0165, Time: 0.05s
Iteration 3659, Loss: 0.0148, Time: 0.05s
Iteration 3660, Loss: 0.0221, Time: 0.05s
Iteration 3661, Loss: 0.0180, Time: 0.05s
Iteration 3662, Loss: 0.0164, Time: 0.05s
Iteration 3663, Loss: 0.0259, Time: 0.06s
Iteration 3664, Loss: 0.0151, Time: 0.04s
Iteration 3665, Loss: 0.0211, Time: 0.05s
Iteration 3666, Loss: 0.0172, Time: 0.05s
Iteration 3667, Loss: 0.0194, Time: 0.05s
Iteration 3668, Loss: 0.0228, Time: 0.04s
Iteration 3669, Loss: 0.0177, Time: 0.04s
Iteration 3670, Loss: 0.0221, Time: 0.04s
```

```
Iteration 3671, Loss: 0.0166, Time: 0.04s
Iteration 3672, Loss: 0.0201, Time: 0.04s
Iteration 3673, Loss: 0.0156, Time: 0.04s
Iteration 3674, Loss: 0.0278, Time: 0.04s
Iteration 3675, Loss: 0.0243, Time: 0.04s
Iteration 3676, Loss: 0.0261, Time: 0.05s
Iteration 3677, Loss: 0.0163, Time: 0.05s
Iteration 3678, Loss: 0.0151, Time: 0.04s
Iteration 3679, Loss: 0.0214, Time: 0.04s
Iteration 3680, Loss: 0.0285, Time: 0.04s
Iteration 3681, Loss: 0.0185, Time: 0.04s
Iteration 3682, Loss: 0.0269, Time: 0.04s
Iteration 3683, Loss: 0.0279, Time: 0.04s
Iteration 3684, Loss: 0.0189, Time: 0.04s
Iteration 3685, Loss: 0.0202, Time: 0.04s
Iteration 3686, Loss: 0.0208, Time: 0.04s
Iteration 3687, Loss: 0.0139, Time: 0.04s
Iteration 3688, Loss: 0.0229, Time: 0.05s
Iteration 3689, Loss: 0.0258, Time: 0.04s
Iteration 3690, Loss: 0.0195, Time: 0.04s
Iteration 3691, Loss: 0.0141, Time: 0.04s
Iteration 3692, Loss: 0.0190, Time: 0.04s
Iteration 3693, Loss: 0.0209, Time: 0.04s
Iteration 3694, Loss: 0.0149, Time: 0.04s
Iteration 3695, Loss: 0.0126, Time: 0.04s
Iteration 3696, Loss: 0.0143, Time: 0.04s
Iteration 3697, Loss: 0.0206, Time: 0.05s
Iteration 3698, Loss: 0.0229, Time: 0.05s
Iteration 3699, Loss: 0.0218, Time: 0.05s
Iteration 3700, Loss: 0.0099, Time: 0.05s
Iteration 3700, Loss: 0.0099, Time: 0.05s
Test Loss: 0.0495
Iteration 3701, Loss: 0.0274, Time: 0.06s
Iteration 3702, Loss: 0.0161, Time: 0.06s
Iteration 3703, Loss: 0.0152, Time: 0.05s
Iteration 3704, Loss: 0.0279, Time: 0.05s
Iteration 3705, Loss: 0.0193, Time: 0.05s
Iteration 3706, Loss: 0.0219, Time: 0.05s
Iteration 3707, Loss: 0.0232, Time: 0.05s
Iteration 3708, Loss: 0.0272, Time: 0.05s
Iteration 3709, Loss: 0.0161, Time: 0.04s
Iteration 3710, Loss: 0.0198, Time: 0.05s
Iteration 3711, Loss: 0.0177, Time: 0.05s
Iteration 3712, Loss: 0.0224, Time: 0.06s
Iteration 3713, Loss: 0.0183, Time: 0.04s
Iteration 3714, Loss: 0.0135, Time: 0.04s
Iteration 3715, Loss: 0.0180, Time: 0.05s
Iteration 3716, Loss: 0.0262, Time: 0.04s
```

```
Iteration 3717, Loss: 0.0216, Time: 0.05s
Iteration 3718, Loss: 0.0144, Time: 0.04s
Iteration 3719, Loss: 0.0164, Time: 0.04s
Iteration 3720, Loss: 0.0171, Time: 0.05s
Iteration 3721, Loss: 0.0261, Time: 0.05s
Iteration 3722, Loss: 0.0148, Time: 0.04s
Iteration 3723, Loss: 0.0153, Time: 0.05s
Iteration 3724, Loss: 0.0289, Time: 0.04s
Iteration 3725, Loss: 0.0165, Time: 0.04s
Iteration 3726, Loss: 0.0273, Time: 0.05s
Iteration 3727, Loss: 0.0275, Time: 0.04s
Iteration 3728, Loss: 0.0228, Time: 0.04s
Iteration 3729, Loss: 0.0220, Time: 0.04s
Iteration 3730, Loss: 0.0212, Time: 0.05s
Iteration 3731, Loss: 0.0206, Time: 0.04s
Iteration 3732, Loss: 0.0221, Time: 0.04s
Iteration 3733, Loss: 0.0192, Time: 0.05s
Iteration 3734, Loss: 0.0140, Time: 0.04s
Iteration 3735, Loss: 0.0166, Time: 0.04s
Iteration 3736, Loss: 0.0232, Time: 0.04s
Iteration 3737, Loss: 0.0237, Time: 0.05s
Iteration 3738, Loss: 0.0173, Time: 0.04s
Iteration 3739, Loss: 0.0211, Time: 0.04s
Iteration 3740, Loss: 0.0165, Time: 0.04s
Iteration 3741, Loss: 0.0264, Time: 0.04s
Iteration 3742, Loss: 0.0205, Time: 0.04s
Iteration 3743, Loss: 0.0224, Time: 0.04s
Iteration 3744, Loss: 0.0139, Time: 0.04s
Iteration 3745, Loss: 0.0274, Time: 0.06s
Iteration 3746, Loss: 0.0136, Time: 0.05s
Iteration 3747, Loss: 0.0231, Time: 0.05s
Iteration 3748, Loss: 0.0198, Time: 0.04s
Iteration 3749, Loss: 0.0142, Time: 0.04s
Iteration 3750, Loss: 0.0121, Time: 0.04s
Iteration 3751, Loss: 0.0235, Time: 0.05s
Iteration 3752, Loss: 0.0188, Time: 0.04s
Iteration 3753, Loss: 0.0217, Time: 0.04s
Iteration 3754, Loss: 0.0320, Time: 0.05s
Iteration 3755, Loss: 0.0207, Time: 0.04s
Iteration 3756, Loss: 0.0161, Time: 0.04s
Iteration 3757, Loss: 0.0204, Time: 0.05s
Iteration 3758, Loss: 0.0170, Time: 0.04s
Iteration 3759, Loss: 0.0141, Time: 0.05s
Iteration 3760, Loss: 0.0167, Time: 0.04s
Iteration 3761, Loss: 0.0196, Time: 0.04s
Iteration 3762, Loss: 0.0240, Time: 0.04s
Iteration 3763, Loss: 0.0224, Time: 0.04s
Iteration 3764, Loss: 0.0172, Time: 0.05s
```

```
Iteration 3765, Loss: 0.0250, Time: 0.04s
Iteration 3766, Loss: 0.0255, Time: 0.05s
Iteration 3767, Loss: 0.0147, Time: 0.05s
Iteration 3768, Loss: 0.0196, Time: 0.04s
Iteration 3769, Loss: 0.0204, Time: 0.04s
Iteration 3770, Loss: 0.0209, Time: 0.05s
Iteration 3771, Loss: 0.0191, Time: 0.04s
Iteration 3772, Loss: 0.0204, Time: 0.05s
Iteration 3773, Loss: 0.0154, Time: 0.04s
Iteration 3774, Loss: 0.0116, Time: 0.04s
Iteration 3775, Loss: 0.0207, Time: 0.04s
Iteration 3776, Loss: 0.0270, Time: 0.04s
Iteration 3777, Loss: 0.0243, Time: 0.04s
Iteration 3778, Loss: 0.0227, Time: 0.05s
Iteration 3779, Loss: 0.0117, Time: 0.05s
Iteration 3780, Loss: 0.0242, Time: 0.05s
Iteration 3781, Loss: 0.0218, Time: 0.04s
Iteration 3782, Loss: 0.0148, Time: 0.04s
Iteration 3783, Loss: 0.0190, Time: 0.05s
Iteration 3784, Loss: 0.0172, Time: 0.04s
Iteration 3785, Loss: 0.0177, Time: 0.05s
Iteration 3786, Loss: 0.0251, Time: 0.04s
Iteration 3787, Loss: 0.0179, Time: 0.04s
Iteration 3788, Loss: 0.0217, Time: 0.04s
Iteration 3789, Loss: 0.0176, Time: 0.05s
Iteration 3790, Loss: 0.0234, Time: 0.04s
Iteration 3791, Loss: 0.0192, Time: 0.05s
Iteration 3792, Loss: 0.0193, Time: 0.04s
Iteration 3793, Loss: 0.0151, Time: 0.05s
Iteration 3794, Loss: 0.0241, Time: 0.05s
Iteration 3795, Loss: 0.0145, Time: 0.04s
Iteration 3796, Loss: 0.0162, Time: 0.05s
Iteration 3797, Loss: 0.0169, Time: 0.04s
Iteration 3798, Loss: 0.0155, Time: 0.04s
Iteration 3799, Loss: 0.0258, Time: 0.04s
Iteration 3800, Loss: 0.0258, Time: 0.05s
Iteration 3800, Loss: 0.0258, Time: 0.05s
Test Loss: 0.0411
Iteration 3801, Loss: 0.0159, Time: 0.04s
Iteration 3802, Loss: 0.0167, Time: 0.04s
Iteration 3803, Loss: 0.0244, Time: 0.04s
Iteration 3804, Loss: 0.0233, Time: 0.04s
Iteration 3805, Loss: 0.0121, Time: 0.04s
Iteration 3806, Loss: 0.0156, Time: 0.04s
Iteration 3807, Loss: 0.0282, Time: 0.04s
Iteration 3808, Loss: 0.0198, Time: 0.04s
Iteration 3809, Loss: 0.0217, Time: 0.04s
Iteration 3810, Loss: 0.0244, Time: 0.04s
```

```
Iteration 3811, Loss: 0.0385, Time: 0.04s
Iteration 3812, Loss: 0.0134, Time: 0.04s
Iteration 3813, Loss: 0.0223, Time: 0.04s
Iteration 3814, Loss: 0.0193, Time: 0.04s
Iteration 3815, Loss: 0.0256, Time: 0.04s
Iteration 3816, Loss: 0.0240, Time: 0.04s
Iteration 3817, Loss: 0.0195, Time: 0.04s
Iteration 3818, Loss: 0.0233, Time: 0.04s
Iteration 3819, Loss: 0.0137, Time: 0.05s
Iteration 3820, Loss: 0.0155, Time: 0.05s
Iteration 3821, Loss: 0.0216, Time: 0.04s
Iteration 3822, Loss: 0.0120, Time: 0.04s
Iteration 3823, Loss: 0.0204, Time: 0.05s
Iteration 3824, Loss: 0.0227, Time: 0.05s
Iteration 3825, Loss: 0.0124, Time: 0.05s
Iteration 3826, Loss: 0.0278, Time: 0.04s
Iteration 3827, Loss: 0.0223, Time: 0.05s
Iteration 3828, Loss: 0.0241, Time: 0.04s
Iteration 3829, Loss: 0.0164, Time: 0.04s
Iteration 3830, Loss: 0.0188, Time: 0.04s
Iteration 3831, Loss: 0.0169, Time: 0.04s
Iteration 3832, Loss: 0.0188, Time: 0.04s
Iteration 3833, Loss: 0.0228, Time: 0.04s
Iteration 3834, Loss: 0.0249, Time: 0.04s
Iteration 3835, Loss: 0.0189, Time: 0.04s
Iteration 3836, Loss: 0.0158, Time: 0.04s
Iteration 3837, Loss: 0.0158, Time: 0.04s
Iteration 3838, Loss: 0.0141, Time: 0.04s
Iteration 3839, Loss: 0.0196, Time: 0.04s
Iteration 3840, Loss: 0.0264, Time: 0.05s
Iteration 3841, Loss: 0.0123, Time: 0.04s
Iteration 3842, Loss: 0.0133, Time: 0.04s
Iteration 3843, Loss: 0.0203, Time: 0.04s
Iteration 3844, Loss: 0.0193, Time: 0.04s
Iteration 3845, Loss: 0.0156, Time: 0.05s
Iteration 3846, Loss: 0.0243, Time: 0.05s
Iteration 3847, Loss: 0.0169, Time: 0.04s
Iteration 3848, Loss: 0.0148, Time: 0.04s
Iteration 3849, Loss: 0.0138, Time: 0.04s
Iteration 3850, Loss: 0.0092, Time: 0.04s
Iteration 3851, Loss: 0.0220, Time: 0.04s
Iteration 3852, Loss: 0.0194, Time: 0.04s
Iteration 3853, Loss: 0.0133, Time: 0.04s
Iteration 3854, Loss: 0.0324, Time: 0.04s
Iteration 3855, Loss: 0.0180, Time: 0.04s
Iteration 3856, Loss: 0.0279, Time: 0.04s
Iteration 3857, Loss: 0.0222, Time: 0.04s
Iteration 3858, Loss: 0.0260, Time: 0.04s
```

```
Iteration 3859, Loss: 0.0246, Time: 0.04s
Iteration 3860, Loss: 0.0217, Time: 0.05s
Iteration 3861, Loss: 0.0198, Time: 0.05s
Iteration 3862, Loss: 0.0156, Time: 0.04s
Iteration 3863, Loss: 0.0260, Time: 0.04s
Iteration 3864, Loss: 0.0234, Time: 0.04s
Iteration 3865, Loss: 0.0310, Time: 0.05s
Iteration 3866, Loss: 0.0224, Time: 0.04s
Iteration 3867, Loss: 0.0183, Time: 0.04s
Iteration 3868, Loss: 0.0169, Time: 0.04s
Iteration 3869, Loss: 0.0186, Time: 0.04s
Iteration 3870, Loss: 0.0135, Time: 0.05s
Iteration 3871, Loss: 0.0187, Time: 0.05s
Iteration 3872, Loss: 0.0206, Time: 0.05s
Iteration 3873, Loss: 0.0161, Time: 0.04s
Iteration 3874, Loss: 0.0172, Time: 0.04s
Iteration 3875, Loss: 0.0151, Time: 0.04s
Iteration 3876, Loss: 0.0188, Time: 0.05s
Iteration 3877, Loss: 0.0194, Time: 0.04s
Iteration 3878, Loss: 0.0130, Time: 0.05s
Iteration 3879, Loss: 0.0240, Time: 0.04s
Iteration 3880, Loss: 0.0140, Time: 0.04s
Iteration 3881, Loss: 0.0181, Time: 0.04s
Iteration 3882, Loss: 0.0161, Time: 0.04s
Iteration 3883, Loss: 0.0079, Time: 0.05s
Iteration 3884, Loss: 0.0238, Time: 0.04s
Iteration 3885, Loss: 0.0144, Time: 0.04s
Iteration 3886, Loss: 0.0155, Time: 0.05s
Iteration 3887, Loss: 0.0139, Time: 0.04s
Iteration 3888, Loss: 0.0163, Time: 0.05s
Iteration 3889, Loss: 0.0213, Time: 0.04s
Iteration 3890, Loss: 0.0143, Time: 0.05s
Iteration 3891, Loss: 0.0168, Time: 0.05s
Iteration 3892, Loss: 0.0130, Time: 0.05s
Iteration 3893, Loss: 0.0170, Time: 0.04s
Iteration 3894, Loss: 0.0126, Time: 0.05s
Iteration 3895, Loss: 0.0157, Time: 0.04s
Iteration 3896, Loss: 0.0145, Time: 0.04s
Iteration 3897, Loss: 0.0119, Time: 0.05s
Iteration 3898, Loss: 0.0101, Time: 0.04s
Iteration 3899, Loss: 0.0270, Time: 0.04s
Iteration 3900, Loss: 0.0185, Time: 0.05s
Iteration 3900, Loss: 0.0185, Time: 0.05s
Test Loss: 0.0299
Iteration 3901, Loss: 0.0152, Time: 0.04s
Iteration 3902, Loss: 0.0154, Time: 0.04s
Iteration 3903, Loss: 0.0098, Time: 0.04s
Iteration 3904, Loss: 0.0222, Time: 0.05s
```

```
Iteration 3905, Loss: 0.0178, Time: 0.05s
Iteration 3906, Loss: 0.0193, Time: 0.04s
Iteration 3907, Loss: 0.0227, Time: 0.04s
Iteration 3908, Loss: 0.0145, Time: 0.04s
Iteration 3909, Loss: 0.0091, Time: 0.04s
Iteration 3910, Loss: 0.0134, Time: 0.06s
Iteration 3911, Loss: 0.0164, Time: 0.04s
Iteration 3912, Loss: 0.0182, Time: 0.04s
Iteration 3913, Loss: 0.0183, Time: 0.04s
Iteration 3914, Loss: 0.0143, Time: 0.04s
Iteration 3915, Loss: 0.0222, Time: 0.04s
Iteration 3916, Loss: 0.0207, Time: 0.04s
Iteration 3917, Loss: 0.0165, Time: 0.04s
Iteration 3918, Loss: 0.0172, Time: 0.04s
Iteration 3919, Loss: 0.0250, Time: 0.05s
Iteration 3920, Loss: 0.0181, Time: 0.05s
Iteration 3921, Loss: 0.0113, Time: -0.10s
Iteration 3922, Loss: 0.0140, Time: 0.05s
Iteration 3923, Loss: 0.0133, Time: 0.05s
Iteration 3924, Loss: 0.0192, Time: 0.05s
Iteration 3925, Loss: 0.0185, Time: 0.05s
Iteration 3926, Loss: 0.0185, Time: 0.04s
Iteration 3927, Loss: 0.0284, Time: 0.04s
Iteration 3928, Loss: 0.0219, Time: 0.05s
Iteration 3929, Loss: 0.0221, Time: 0.04s
Iteration 3930, Loss: 0.0123, Time: 0.06s
Iteration 3931, Loss: 0.0124, Time: 0.05s
Iteration 3932, Loss: 0.0191, Time: 0.04s
Iteration 3933, Loss: 0.0147, Time: 0.04s
Iteration 3934, Loss: 0.0190, Time: 0.05s
Iteration 3935, Loss: 0.0103, Time: 0.04s
Iteration 3936, Loss: 0.0136, Time: 0.04s
Iteration 3937, Loss: 0.0118, Time: 0.04s
Iteration 3938, Loss: 0.0178, Time: 0.04s
Iteration 3939, Loss: 0.0219, Time: 0.04s
Iteration 3940, Loss: 0.0230, Time: 0.04s
Iteration 3941, Loss: 0.0281, Time: 0.05s
Iteration 3942, Loss: 0.0207, Time: 0.04s
Iteration 3943, Loss: 0.0172, Time: 0.04s
Iteration 3944, Loss: 0.0205, Time: 0.05s
Iteration 3945, Loss: 0.0147, Time: 0.04s
Iteration 3946, Loss: 0.0185, Time: 0.05s
Iteration 3947, Loss: 0.0132, Time: 0.04s
Iteration 3948, Loss: 0.0186, Time: 0.04s
Iteration 3949, Loss: 0.0182, Time: 0.04s
Iteration 3950, Loss: 0.0212, Time: 0.04s
Iteration 3951, Loss: 0.0127, Time: 0.04s
Iteration 3952, Loss: 0.0216, Time: 0.04s
```

```
Iteration 3953, Loss: 0.0109, Time: 0.04s
Iteration 3954, Loss: 0.0171, Time: 0.04s
Iteration 3955, Loss: 0.0268, Time: 0.04s
Iteration 3956, Loss: 0.0132, Time: 0.04s
Iteration 3957, Loss: 0.0182, Time: 0.04s
Iteration 3958, Loss: 0.0174, Time: 0.04s
Iteration 3959, Loss: 0.0136, Time: 0.05s
Iteration 3960, Loss: 0.0132, Time: 0.04s
Iteration 3961, Loss: 0.0216, Time: 0.04s
Iteration 3962, Loss: 0.0202, Time: 0.04s
Iteration 3963, Loss: 0.0125, Time: 0.04s
Iteration 3964, Loss: 0.0171, Time: 0.04s
Iteration 3965, Loss: 0.0155, Time: 0.04s
Iteration 3966, Loss: 0.0135, Time: 0.05s
Iteration 3967, Loss: 0.0175, Time: 0.04s
Iteration 3968, Loss: 0.0218, Time: 0.04s
Iteration 3969, Loss: 0.0185, Time: 0.04s
Iteration 3970, Loss: 0.0149, Time: 0.04s
Iteration 3971, Loss: 0.0207, Time: 0.05s
Iteration 3972, Loss: 0.0136, Time: 0.05s
Iteration 3973, Loss: 0.0147, Time: 0.05s
Iteration 3974, Loss: 0.0137, Time: 0.05s
Iteration 3975, Loss: 0.0166, Time: 0.05s
Iteration 3976, Loss: 0.0136, Time: 0.05s
Iteration 3977, Loss: 0.0114, Time: 0.05s
Iteration 3978, Loss: 0.0137, Time: 0.04s
Iteration 3979, Loss: 0.0149, Time: 0.04s
Iteration 3980, Loss: 0.0196, Time: 0.05s
Iteration 3981, Loss: 0.0183, Time: 0.04s
Iteration 3982, Loss: 0.0216, Time: 0.04s
Iteration 3983, Loss: 0.0200, Time: 0.05s
Iteration 3984, Loss: 0.0138, Time: 0.05s
Iteration 3985, Loss: 0.0123, Time: 0.05s
Iteration 3986, Loss: 0.0137, Time: 0.04s
Iteration 3987, Loss: 0.0197, Time: 0.04s
Iteration 3988, Loss: 0.0156, Time: 0.04s
Iteration 3989, Loss: 0.0190, Time: 0.04s
Iteration 3990, Loss: 0.0192, Time: 0.04s
Iteration 3991, Loss: 0.0170, Time: 0.04s
Iteration 3992, Loss: 0.0247, Time: 0.04s
Iteration 3993, Loss: 0.0185, Time: 0.04s
Iteration 3994, Loss: 0.0154, Time: 0.04s
Iteration 3995, Loss: 0.0196, Time: 0.05s
Iteration 3996, Loss: 0.0207, Time: 0.04s
Iteration 3997, Loss: 0.0177, Time: 0.04s
Iteration 3998, Loss: 0.0136, Time: 0.04s
Iteration 3999, Loss: 0.0171, Time: 0.04s
Iteration 4000, Loss: 0.0245, Time: 0.05s
```

```
Iteration 4000, Loss: 0.0245, Time: 0.05s
Test Loss: 0.0171
Iteration 4001, Loss: 0.0119, Time: 0.04s
Iteration 4002, Loss: 0.0229, Time: 0.04s
Iteration 4003, Loss: 0.0212, Time: 0.05s
Iteration 4004, Loss: 0.0157, Time: 0.04s
Iteration 4005, Loss: 0.0211, Time: 0.04s
Iteration 4006, Loss: 0.0196, Time: 0.04s
Iteration 4007, Loss: 0.0195, Time: 0.04s
Iteration 4008, Loss: 0.0149, Time: 0.04s
Iteration 4009, Loss: 0.0210, Time: 0.04s
Iteration 4010, Loss: 0.0180, Time: 0.04s
Iteration 4011, Loss: 0.0121, Time: 0.04s
Iteration 4012, Loss: 0.0120, Time: 0.04s
Iteration 4013, Loss: 0.0135, Time: 0.04s
Iteration 4014, Loss: 0.0212, Time: 0.04s
Iteration 4015, Loss: 0.0155, Time: 0.05s
Iteration 4016, Loss: 0.0188, Time: 0.05s
Iteration 4017, Loss: 0.0154, Time: 0.04s
Iteration 4018, Loss: 0.0192, Time: 0.04s
Iteration 4019, Loss: 0.0143, Time: 0.04s
Iteration 4020, Loss: 0.0162, Time: 0.05s
Iteration 4021, Loss: 0.0171, Time: 0.04s
Iteration 4022, Loss: 0.0143, Time: 0.05s
Iteration 4023, Loss: 0.0142, Time: 0.04s
Iteration 4024, Loss: 0.0186, Time: 0.04s
Iteration 4025, Loss: 0.0176, Time: 0.04s
Iteration 4026, Loss: 0.0167, Time: 0.04s
Iteration 4027, Loss: 0.0165, Time: 0.04s
Iteration 4028, Loss: 0.0226, Time: 0.04s
Iteration 4029, Loss: 0.0091, Time: 0.04s
Iteration 4030, Loss: 0.0188, Time: 0.04s
Iteration 4031, Loss: 0.0217, Time: 0.04s
Iteration 4032, Loss: 0.0173, Time: 0.04s
Iteration 4033, Loss: 0.0197, Time: 0.05s
Iteration 4034, Loss: 0.0223, Time: 0.04s
Iteration 4035, Loss: 0.0263, Time: 0.04s
Iteration 4036, Loss: 0.0196, Time: 0.04s
Iteration 4037, Loss: 0.0182, Time: 0.04s
Iteration 4038, Loss: 0.0211, Time: 0.04s
Iteration 4039, Loss: 0.0237, Time: 0.04s
Iteration 4040, Loss: 0.0182, Time: 0.04s
Iteration 4041, Loss: 0.0214, Time: 0.05s
Iteration 4042, Loss: 0.0110, Time: 0.04s
Iteration 4043, Loss: 0.0124, Time: 0.04s
Iteration 4044, Loss: 0.0223, Time: 0.04s
Iteration 4045, Loss: 0.0165, Time: 0.04s
Iteration 4046, Loss: 0.0175, Time: 0.04s
```

```
Iteration 4047, Loss: 0.0220, Time: 0.04s
Iteration 4048, Loss: 0.0157, Time: 0.04s
Iteration 4049, Loss: 0.0183, Time: 0.04s
Iteration 4050, Loss: 0.0203, Time: 0.05s
Iteration 4051, Loss: 0.0097, Time: 0.05s
Iteration 4052, Loss: 0.0196, Time: 0.05s
Iteration 4053, Loss: 0.0157, Time: 0.04s
Iteration 4054, Loss: 0.0236, Time: 0.05s
Iteration 4055, Loss: 0.0249, Time: 0.04s
Iteration 4056, Loss: 0.0193, Time: 0.04s
Iteration 4057, Loss: 0.0193, Time: 0.04s
Iteration 4058, Loss: 0.0245, Time: 0.05s
Iteration 4059, Loss: 0.0161, Time: 0.04s
Iteration 4060, Loss: 0.0183, Time: 0.04s
Iteration 4061, Loss: 0.0156, Time: 0.04s
Iteration 4062, Loss: 0.0168, Time: 0.04s
Iteration 4063, Loss: 0.0109, Time: 0.05s
Iteration 4064, Loss: 0.0158, Time: 0.04s
Iteration 4065, Loss: 0.0226, Time: 0.05s
Iteration 4066, Loss: 0.0208, Time: 0.04s
Iteration 4067, Loss: 0.0183, Time: 0.04s
Iteration 4068, Loss: 0.0162, Time: 0.04s
Iteration 4069, Loss: 0.0128, Time: 0.05s
Iteration 4070, Loss: 0.0193, Time: 0.05s
Iteration 4071, Loss: 0.0158, Time: 0.04s
Iteration 4072, Loss: 0.0176, Time: 0.04s
Iteration 4073, Loss: 0.0194, Time: 0.04s
Iteration 4074, Loss: 0.0122, Time: 0.04s
Iteration 4075, Loss: 0.0194, Time: 0.04s
Iteration 4076, Loss: 0.0125, Time: 0.04s
Iteration 4077, Loss: 0.0204, Time: 0.05s
Iteration 4078, Loss: 0.0178, Time: 0.04s
Iteration 4079, Loss: 0.0119, Time: 0.04s
Iteration 4080, Loss: 0.0203, Time: 0.04s
Iteration 4081, Loss: 0.0177, Time: 0.04s
Iteration 4082, Loss: 0.0100, Time: 0.04s
Iteration 4083, Loss: 0.0162, Time: 0.04s
Iteration 4084, Loss: 0.0201, Time: 0.04s
Iteration 4085, Loss: 0.0161, Time: 0.04s
Iteration 4086, Loss: 0.0166, Time: 0.04s
Iteration 4087, Loss: 0.0255, Time: 0.04s
Iteration 4088, Loss: 0.0143, Time: 0.04s
Iteration 4089, Loss: 0.0184, Time: 0.04s
Iteration 4090, Loss: 0.0140, Time: 0.04s
Iteration 4091, Loss: 0.0152, Time: 0.05s
Iteration 4092, Loss: 0.0208, Time: 0.05s
Iteration 4093, Loss: 0.0195, Time: 0.04s
Iteration 4094, Loss: 0.0320, Time: 0.04s
```

```
Iteration 4095, Loss: 0.0132, Time: 0.04s
Iteration 4096, Loss: 0.0177, Time: 0.04s
Iteration 4097, Loss: 0.0177, Time: 0.04s
Iteration 4098, Loss: 0.0269, Time: 0.04s
Iteration 4099, Loss: 0.0111, Time: 0.05s
Iteration 4100, Loss: 0.0139, Time: 0.04s
Iteration 4100, Loss: 0.0139, Time: 0.04s
Test Loss: 0.0315
Iteration 4101, Loss: 0.0153, Time: 0.05s
Iteration 4102, Loss: 0.0173, Time: 0.04s
Iteration 4103, Loss: 0.0142, Time: 0.04s
Iteration 4104, Loss: 0.0186, Time: 0.05s
Iteration 4105, Loss: 0.0156, Time: 0.04s
Iteration 4106, Loss: 0.0302, Time: 0.05s
Iteration 4107, Loss: 0.0136, Time: 0.04s
Iteration 4108, Loss: 0.0159, Time: 0.04s
Iteration 4109, Loss: 0.0176, Time: 0.05s
Iteration 4110, Loss: 0.0179, Time: 0.04s
Iteration 4111, Loss: 0.0217, Time: 0.04s
Iteration 4112, Loss: 0.0136, Time: 0.04s
Iteration 4113, Loss: 0.0129, Time: 0.04s
Iteration 4114, Loss: 0.0248, Time: 0.05s
Iteration 4115, Loss: 0.0189, Time: 0.04s
Iteration 4116, Loss: 0.0232, Time: 0.05s
Iteration 4117, Loss: 0.0199, Time: 0.04s
Iteration 4118, Loss: 0.0124, Time: 0.05s
Iteration 4119, Loss: 0.0135, Time: 0.04s
Iteration 4120, Loss: 0.0197, Time: 0.04s
Iteration 4121, Loss: 0.0175, Time: 0.04s
Iteration 4122, Loss: 0.0208, Time: 0.04s
Iteration 4123, Loss: 0.0173, Time: 0.04s
Iteration 4124, Loss: 0.0145, Time: 0.04s
Iteration 4125, Loss: 0.0099, Time: 0.04s
Iteration 4126, Loss: 0.0110, Time: 0.04s
Iteration 4127, Loss: 0.0076, Time: 0.04s
Iteration 4128, Loss: 0.0145, Time: 0.04s
Iteration 4129, Loss: 0.0275, Time: 0.04s
Iteration 4130, Loss: 0.0319, Time: 0.05s
Iteration 4131, Loss: 0.0170, Time: 0.05s
Iteration 4132, Loss: 0.0151, Time: 0.04s
Iteration 4133, Loss: 0.0194, Time: 0.04s
Iteration 4134, Loss: 0.0213, Time: 0.04s
Iteration 4135, Loss: 0.0214, Time: 0.04s
Iteration 4136, Loss: 0.0114, Time: 0.04s
Iteration 4137, Loss: 0.0107, Time: 0.04s
Iteration 4138, Loss: 0.0210, Time: 0.05s
Iteration 4139, Loss: 0.0170, Time: 0.04s
Iteration 4140, Loss: 0.0120, Time: 0.04s
```

```
Iteration 4141, Loss: 0.0174, Time: 0.04s
Iteration 4142, Loss: 0.0116, Time: 0.04s
Iteration 4143, Loss: 0.0215, Time: 0.04s
Iteration 4144, Loss: 0.0183, Time: 0.05s
Iteration 4145, Loss: 0.0145, Time: 0.04s
Iteration 4146, Loss: 0.0127, Time: 0.04s
Iteration 4147, Loss: 0.0248, Time: 0.04s
Iteration 4148, Loss: 0.0232, Time: 0.05s
Iteration 4149, Loss: 0.0192, Time: 0.05s
Iteration 4150, Loss: 0.0242, Time: 0.06s
Iteration 4151, Loss: 0.0179, Time: 0.06s
Iteration 4152, Loss: 0.0185, Time: 0.04s
Iteration 4153, Loss: 0.0200, Time: 0.04s
Iteration 4154, Loss: 0.0128, Time: 0.05s
Iteration 4155, Loss: 0.0077, Time: 0.04s
Iteration 4156, Loss: 0.0287, Time: 0.05s
Iteration 4157, Loss: 0.0117, Time: 0.04s
Iteration 4158, Loss: 0.0159, Time: 0.04s
Iteration 4159, Loss: 0.0139, Time: 0.04s
Iteration 4160, Loss: 0.0120, Time: 0.04s
Iteration 4161, Loss: 0.0240, Time: 0.05s
Iteration 4162, Loss: 0.0157, Time: 0.04s
Iteration 4163, Loss: 0.0137, Time: 0.04s
Iteration 4164, Loss: 0.0129, Time: 0.04s
Iteration 4165, Loss: 0.0159, Time: 0.04s
Iteration 4166, Loss: 0.0180, Time: 0.05s
Iteration 4167, Loss: 0.0189, Time: 0.04s
Iteration 4168, Loss: 0.0137, Time: 0.04s
Iteration 4169, Loss: 0.0103, Time: 0.04s
Iteration 4170, Loss: 0.0120, Time: 0.04s
Iteration 4171, Loss: 0.0196, Time: 0.05s
Iteration 4172, Loss: 0.0151, Time: 0.05s
Iteration 4173, Loss: 0.0200, Time: 0.04s
Iteration 4174, Loss: 0.0131, Time: 0.04s
Iteration 4175, Loss: 0.0154, Time: 0.05s
Iteration 4176, Loss: 0.0235, Time: 0.04s
Iteration 4177, Loss: 0.0175, Time: 0.04s
Iteration 4178, Loss: 0.0198, Time: 0.04s
Iteration 4179, Loss: 0.0156, Time: 0.05s
Iteration 4180, Loss: 0.0183, Time: 0.05s
Iteration 4181, Loss: 0.0166, Time: 0.04s
Iteration 4182, Loss: 0.0192, Time: 0.05s
Iteration 4183, Loss: 0.0166, Time: 0.04s
Iteration 4184, Loss: 0.0170, Time: 0.04s
Iteration 4185, Loss: 0.0145, Time: 0.05s
Iteration 4186, Loss: 0.0182, Time: 0.04s
Iteration 4187, Loss: 0.0134, Time: 0.04s
Iteration 4188, Loss: 0.0190, Time: 0.05s
```

```
Iteration 4189, Loss: 0.0117, Time: 0.04s
Iteration 4190, Loss: 0.0153, Time: 0.04s
Iteration 4191, Loss: 0.0145, Time: 0.04s
Iteration 4192, Loss: 0.0180, Time: 0.04s
Iteration 4193, Loss: 0.0171, Time: 0.05s
Iteration 4194, Loss: 0.0182, Time: 0.04s
Iteration 4195, Loss: 0.0182, Time: 0.04s
Iteration 4196, Loss: 0.0148, Time: 0.04s
Iteration 4197, Loss: 0.0109, Time: 0.04s
Iteration 4198, Loss: 0.0171, Time: 0.06s
Iteration 4199, Loss: 0.0139, Time: 0.04s
Iteration 4200, Loss: 0.0229, Time: 0.04s
Iteration 4200, Loss: 0.0229, Time: 0.04s
Test Loss: 0.0440
Iteration 4201, Loss: 0.0105, Time: 0.04s
Iteration 4202, Loss: 0.0058, Time: 0.04s
Iteration 4203, Loss: 0.0082, Time: 0.04s
Iteration 4204, Loss: 0.0180, Time: 0.05s
Iteration 4205, Loss: 0.0168, Time: 0.04s
Iteration 4206, Loss: 0.0260, Time: 0.04s
Iteration 4207, Loss: 0.0285, Time: 0.05s
Iteration 4208, Loss: 0.0143, Time: 0.04s
Iteration 4209, Loss: 0.0212, Time: 0.04s
Iteration 4210, Loss: 0.0122, Time: 0.05s
Iteration 4211, Loss: 0.0221, Time: 0.04s
Iteration 4212, Loss: 0.0185, Time: 0.04s
Iteration 4213, Loss: 0.0190, Time: 0.04s
Iteration 4214, Loss: 0.0110, Time: 0.04s
Iteration 4215, Loss: 0.0140, Time: 0.04s
Iteration 4216, Loss: 0.0201, Time: 0.04s
Iteration 4217, Loss: 0.0139, Time: 0.04s
Iteration 4218, Loss: 0.0184, Time: 0.04s
Iteration 4219, Loss: 0.0144, Time: 0.04s
Iteration 4220, Loss: 0.0197, Time: 0.04s
Iteration 4221, Loss: 0.0174, Time: 0.05s
Iteration 4222, Loss: 0.0233, Time: 0.04s
Iteration 4223, Loss: 0.0138, Time: 0.04s
Iteration 4224, Loss: 0.0194, Time: 0.04s
Iteration 4225, Loss: 0.0176, Time: 0.05s
Iteration 4226, Loss: 0.0133, Time: 0.04s
Iteration 4227, Loss: 0.0174, Time: 0.05s
Iteration 4228, Loss: 0.0193, Time: 0.05s
Iteration 4229, Loss: 0.0193, Time: 0.05s
Iteration 4230, Loss: 0.0190, Time: 0.05s
Iteration 4231, Loss: 0.0166, Time: 0.05s
Iteration 4232, Loss: 0.0148, Time: 0.05s
Iteration 4233, Loss: 0.0220, Time: 0.05s
Iteration 4234, Loss: 0.0229, Time: 0.05s
```

```
Iteration 4235, Loss: 0.0192, Time: 0.04s
Iteration 4236, Loss: 0.0130, Time: 0.04s
Iteration 4237, Loss: 0.0184, Time: 0.05s
Iteration 4238, Loss: 0.0170, Time: 0.05s
Iteration 4239, Loss: 0.0108, Time: 0.05s
Iteration 4240, Loss: 0.0118, Time: 0.04s
Iteration 4241, Loss: 0.0131, Time: 0.04s
Iteration 4242, Loss: 0.0114, Time: 0.04s
Iteration 4243, Loss: 0.0276, Time: 0.04s
Iteration 4244, Loss: 0.0189, Time: 0.04s
Iteration 4245, Loss: 0.0150, Time: 0.04s
Iteration 4246, Loss: 0.0169, Time: 0.04s
Iteration 4247, Loss: 0.0121, Time: 0.04s
Iteration 4248, Loss: 0.0195, Time: 0.04s
Iteration 4249, Loss: 0.0156, Time: 0.04s
Iteration 4250, Loss: 0.0164, Time: 0.04s
Iteration 4251, Loss: 0.0110, Time: 0.04s
Iteration 4252, Loss: 0.0178, Time: 0.06s
Iteration 4253, Loss: 0.0150, Time: 0.04s
Iteration 4254, Loss: 0.0127, Time: 0.05s
Iteration 4255, Loss: 0.0177, Time: 0.04s
Iteration 4256, Loss: 0.0128, Time: 0.04s
Iteration 4257, Loss: 0.0152, Time: 0.04s
Iteration 4258, Loss: 0.0189, Time: 0.05s
Iteration 4259, Loss: 0.0157, Time: 0.05s
Iteration 4260, Loss: 0.0182, Time: 0.04s
Iteration 4261, Loss: 0.0192, Time: 0.04s
Iteration 4262, Loss: 0.0198, Time: 0.05s
Iteration 4263, Loss: 0.0105, Time: 0.04s
Iteration 4264, Loss: 0.0121, Time: 0.05s
Iteration 4265, Loss: 0.0187, Time: 0.04s
Iteration 4266, Loss: 0.0159, Time: 0.04s
Iteration 4267, Loss: 0.0140, Time: 0.05s
Iteration 4268, Loss: 0.0155, Time: 0.04s
Iteration 4269, Loss: 0.0182, Time: 0.04s
Iteration 4270, Loss: 0.0076, Time: 0.05s
Iteration 4271, Loss: 0.0198, Time: 0.04s
Iteration 4272, Loss: 0.0176, Time: 0.04s
Iteration 4273, Loss: 0.0207, Time: 0.04s
Iteration 4274, Loss: 0.0142, Time: 0.05s
Iteration 4275, Loss: 0.0150, Time: 0.05s
Iteration 4276, Loss: 0.0112, Time: 0.05s
Iteration 4277, Loss: 0.0162, Time: 0.05s
Iteration 4278, Loss: 0.0153, Time: 0.04s
Iteration 4279, Loss: 0.0141, Time: 0.04s
Iteration 4280, Loss: 0.0185, Time: 0.04s
Iteration 4281, Loss: 0.0148, Time: 0.05s
Iteration 4282, Loss: 0.0186, Time: 0.05s
```

```
Iteration 4283, Loss: 0.0158, Time: 0.04s
Iteration 4284, Loss: 0.0170, Time: 0.05s
Iteration 4285, Loss: 0.0142, Time: 0.05s
Iteration 4286, Loss: 0.0186, Time: 0.04s
Iteration 4287, Loss: 0.0218, Time: 0.04s
Iteration 4288, Loss: 0.0187, Time: 0.05s
Iteration 4289, Loss: 0.0351, Time: 0.05s
Iteration 4290, Loss: 0.0108, Time: 0.05s
Iteration 4291, Loss: 0.0162, Time: 0.04s
Iteration 4292, Loss: 0.0119, Time: 0.04s
Iteration 4293, Loss: 0.0199, Time: 0.05s
Iteration 4294, Loss: 0.0194, Time: 0.04s
Iteration 4295, Loss: 0.0160, Time: 0.04s
Iteration 4296, Loss: 0.0144, Time: 0.04s
Iteration 4297, Loss: 0.0115, Time: 0.04s
Iteration 4298, Loss: 0.0110, Time: 0.05s
Iteration 4299, Loss: 0.0133, Time: 0.04s
Iteration 4300, Loss: 0.0164, Time: 0.05s
Iteration 4300, Loss: 0.0164, Time: 0.05s
Test Loss: 0.0403
Iteration 4301, Loss: 0.0132, Time: 0.04s
Iteration 4302, Loss: 0.0182, Time: 0.04s
Iteration 4303, Loss: 0.0155, Time: 0.05s
Iteration 4304, Loss: 0.0208, Time: 0.04s
Iteration 4305, Loss: 0.0171, Time: 0.04s
Iteration 4306, Loss: 0.0165, Time: 0.04s
Iteration 4307, Loss: 0.0168, Time: 0.04s
Iteration 4308, Loss: 0.0108, Time: 0.04s
Iteration 4309, Loss: 0.0205, Time: 0.04s
Iteration 4310, Loss: 0.0188, Time: 0.04s
Iteration 4311, Loss: 0.0138, Time: 0.04s
Iteration 4312, Loss: 0.0171, Time: 0.04s
Iteration 4313, Loss: 0.0184, Time: 0.04s
Iteration 4314, Loss: 0.0204, Time: 0.04s
Iteration 4315, Loss: 0.0155, Time: 0.04s
Iteration 4316, Loss: 0.0120, Time: 0.05s
Iteration 4317, Loss: 0.0162, Time: 0.04s
Iteration 4318, Loss: 0.0153, Time: 0.05s
Iteration 4319, Loss: 0.0170, Time: 0.05s
Iteration 4320, Loss: 0.0282, Time: 0.04s
Iteration 4321, Loss: 0.0155, Time: 0.05s
Iteration 4322, Loss: 0.0128, Time: 0.04s
Iteration 4323, Loss: 0.0212, Time: 0.04s
Iteration 4324, Loss: 0.0125, Time: 0.04s
Iteration 4325, Loss: 0.0203, Time: 0.04s
Iteration 4326, Loss: 0.0144, Time: 0.04s
Iteration 4327, Loss: 0.0165, Time: 0.04s
Iteration 4328, Loss: 0.0165, Time: 0.04s
```

```
Iteration 4329, Loss: 0.0215, Time: 0.05s
Iteration 4330, Loss: 0.0125, Time: 0.04s
Iteration 4331, Loss: 0.0153, Time: 0.05s
Iteration 4332, Loss: 0.0111, Time: 0.05s
Iteration 4333, Loss: 0.0173, Time: 0.04s
Iteration 4334, Loss: 0.0136, Time: 0.05s
Iteration 4335, Loss: 0.0166, Time: 0.05s
Iteration 4336, Loss: 0.0177, Time: 0.04s
Iteration 4337, Loss: 0.0205, Time: 0.04s
Iteration 4338, Loss: 0.0167, Time: 0.05s
Iteration 4339, Loss: 0.0148, Time: 0.05s
Iteration 4340, Loss: 0.0137, Time: 0.05s
Iteration 4341, Loss: 0.0136, Time: 0.04s
Iteration 4342, Loss: 0.0182, Time: 0.05s
Iteration 4343, Loss: 0.0200, Time: 0.04s
Iteration 4344, Loss: 0.0092, Time: 0.04s
Iteration 4345, Loss: 0.0151, Time: 0.05s
Iteration 4346, Loss: 0.0144, Time: 0.05s
Iteration 4347, Loss: 0.0129, Time: 0.04s
Iteration 4348, Loss: 0.0150, Time: 0.05s
Iteration 4349, Loss: 0.0139, Time: 0.05s
Iteration 4350, Loss: 0.0200, Time: 0.04s
Iteration 4351, Loss: 0.0143, Time: 0.04s
Iteration 4352, Loss: 0.0161, Time: 0.05s
Iteration 4353, Loss: 0.0209, Time: 0.04s
Iteration 4354, Loss: 0.0148, Time: 0.04s
Iteration 4355, Loss: 0.0150, Time: 0.04s
Iteration 4356, Loss: 0.0187, Time: 0.05s
Iteration 4357, Loss: 0.0136, Time: 0.04s
Iteration 4358, Loss: 0.0182, Time: 0.05s
Iteration 4359, Loss: 0.0200, Time: 0.05s
Iteration 4360, Loss: 0.0164, Time: 0.04s
Iteration 4361, Loss: 0.0116, Time: 0.05s
Iteration 4362, Loss: 0.0305, Time: 0.05s
Iteration 4363, Loss: 0.0111, Time: 0.04s
Iteration 4364, Loss: 0.0172, Time: 0.05s
Iteration 4365, Loss: 0.0149, Time: 0.04s
Iteration 4366, Loss: 0.0153, Time: 0.04s
Iteration 4367, Loss: 0.0132, Time: 0.04s
Iteration 4368, Loss: 0.0175, Time: 0.04s
Iteration 4369, Loss: 0.0141, Time: 0.04s
Iteration 4370, Loss: 0.0205, Time: 0.05s
Iteration 4371, Loss: 0.0140, Time: 0.04s
Iteration 4372, Loss: 0.0194, Time: 0.04s
Iteration 4373, Loss: 0.0124, Time: 0.05s
Iteration 4374, Loss: 0.0192, Time: 0.05s
Iteration 4375, Loss: 0.0274, Time: 0.05s
Iteration 4376, Loss: 0.0183, Time: 0.05s
```

```
Iteration 4377, Loss: 0.0163, Time: 0.05s
Iteration 4378, Loss: 0.0228, Time: 0.05s
Iteration 4379, Loss: 0.0095, Time: 0.05s
Iteration 4380, Loss: 0.0147, Time: 0.04s
Iteration 4381, Loss: 0.0224, Time: 0.04s
Iteration 4382, Loss: 0.0114, Time: 0.04s
Iteration 4383, Loss: 0.0214, Time: 0.05s
Iteration 4384, Loss: 0.0205, Time: 0.04s
Iteration 4385, Loss: 0.0160, Time: 0.04s
Iteration 4386, Loss: 0.0155, Time: 0.04s
Iteration 4387, Loss: 0.0165, Time: 0.05s
Iteration 4388, Loss: 0.0130, Time: 0.04s
Iteration 4389, Loss: 0.0133, Time: 0.05s
Iteration 4390, Loss: 0.0123, Time: 0.04s
Iteration 4391, Loss: 0.0166, Time: 0.04s
Iteration 4392, Loss: 0.0112, Time: 0.04s
Iteration 4393, Loss: 0.0188, Time: 0.04s
Iteration 4394, Loss: 0.0119, Time: 0.04s
Iteration 4395, Loss: 0.0128, Time: 0.04s
Iteration 4396, Loss: 0.0143, Time: 0.05s
Iteration 4397, Loss: 0.0120, Time: 0.04s
Iteration 4398, Loss: 0.0222, Time: 0.05s
Iteration 4399, Loss: 0.0111, Time: 0.05s
Iteration 4400, Loss: 0.0158, Time: 0.04s
Iteration 4400, Loss: 0.0158, Time: 0.04s
Test Loss: 0.0345
Iteration 4401, Loss: 0.0163, Time: 0.04s
Iteration 4402, Loss: 0.0193, Time: 0.04s
Iteration 4403, Loss: 0.0133, Time: 0.04s
Iteration 4404, Loss: 0.0163, Time: 0.05s
Iteration 4405, Loss: 0.0162, Time: 0.05s
Iteration 4406, Loss: 0.0190, Time: 0.04s
Iteration 4407, Loss: 0.0166, Time: 0.04s
Iteration 4408, Loss: 0.0143, Time: 0.04s
Iteration 4409, Loss: 0.0171, Time: 0.04s
Iteration 4410, Loss: 0.0215, Time: 0.04s
Iteration 4411, Loss: 0.0179, Time: 0.05s
Iteration 4412, Loss: 0.0158, Time: 0.04s
Iteration 4413, Loss: 0.0205, Time: 0.04s
Iteration 4414, Loss: 0.0167, Time: 0.04s
Iteration 4415, Loss: 0.0082, Time: 0.04s
Iteration 4416, Loss: 0.0134, Time: 0.04s
Iteration 4417, Loss: 0.0170, Time: 0.04s
Iteration 4418, Loss: 0.0196, Time: 0.04s
Iteration 4419, Loss: 0.0134, Time: 0.04s
Iteration 4420, Loss: 0.0169, Time: 0.04s
Iteration 4421, Loss: 0.0180, Time: 0.05s
Iteration 4422, Loss: 0.0230, Time: 0.05s
```

```
Iteration 4423, Loss: 0.0145, Time: 0.04s
Iteration 4424, Loss: 0.0110, Time: 0.04s
Iteration 4425, Loss: 0.0177, Time: 0.04s
Iteration 4426, Loss: 0.0179, Time: 0.04s
Iteration 4427, Loss: 0.0169, Time: 0.05s
Iteration 4428, Loss: 0.0164, Time: 0.04s
Iteration 4429, Loss: 0.0175, Time: 0.04s
Iteration 4430, Loss: 0.0199, Time: 0.05s
Iteration 4431, Loss: 0.0123, Time: 0.04s
Iteration 4432, Loss: 0.0204, Time: 0.04s
Iteration 4433, Loss: 0.0153, Time: 0.05s
Iteration 4434, Loss: 0.0162, Time: 0.05s
Iteration 4435, Loss: 0.0131, Time: 0.06s
Iteration 4436, Loss: 0.0087, Time: 0.04s
Iteration 4437, Loss: 0.0183, Time: 0.05s
Iteration 4438, Loss: 0.0172, Time: 0.04s
Iteration 4439, Loss: 0.0206, Time: 0.05s
Iteration 4440, Loss: 0.0269, Time: 0.04s
Iteration 4441, Loss: 0.0211, Time: 0.04s
Iteration 4442, Loss: 0.0141, Time: 0.04s
Iteration 4443, Loss: 0.0158, Time: 0.04s
Iteration 4444, Loss: 0.0150, Time: 0.05s
Iteration 4445, Loss: 0.0090, Time: 0.04s
Iteration 4446, Loss: 0.0159, Time: 0.04s
Iteration 4447, Loss: 0.0079, Time: 0.04s
Iteration 4448, Loss: 0.0169, Time: 0.04s
Iteration 4449, Loss: 0.0127, Time: 0.04s
Iteration 4450, Loss: 0.0134, Time: 0.04s
Iteration 4451, Loss: 0.0162, Time: 0.05s
Iteration 4452, Loss: 0.0186, Time: 0.04s
Iteration 4453, Loss: 0.0126, Time: 0.05s
Iteration 4454, Loss: 0.0141, Time: 0.04s
Iteration 4455, Loss: 0.0144, Time: 0.04s
Iteration 4456, Loss: 0.0155, Time: 0.04s
Iteration 4457, Loss: 0.0234, Time: 0.05s
Iteration 4458, Loss: 0.0212, Time: 0.05s
Iteration 4459, Loss: 0.0158, Time: 0.04s
Iteration 4460, Loss: 0.0167, Time: 0.04s
Iteration 4461, Loss: 0.0092, Time: 0.04s
Iteration 4462, Loss: 0.0183, Time: 0.05s
Iteration 4463, Loss: 0.0201, Time: 0.04s
Iteration 4464, Loss: 0.0170, Time: 0.04s
Iteration 4465, Loss: 0.0221, Time: 0.04s
Iteration 4466, Loss: 0.0121, Time: 0.04s
Iteration 4467, Loss: 0.0177, Time: 0.04s
Iteration 4468, Loss: 0.0168, Time: 0.04s
Iteration 4469, Loss: 0.0203, Time: 0.04s
Iteration 4470, Loss: 0.0124, Time: 0.05s
```

```
Iteration 4471, Loss: 0.0167, Time: 0.05s
Iteration 4472, Loss: 0.0124, Time: 0.04s
Iteration 4473, Loss: 0.0205, Time: 0.04s
Iteration 4474, Loss: 0.0124, Time: 0.04s
Iteration 4475, Loss: 0.0155, Time: 0.04s
Iteration 4476, Loss: 0.0162, Time: 0.04s
Iteration 4477, Loss: 0.0108, Time: 0.05s
Iteration 4478, Loss: 0.0204, Time: 0.04s
Iteration 4479, Loss: 0.0198, Time: 0.04s
Iteration 4480, Loss: 0.0151, Time: 0.04s
Iteration 4481, Loss: 0.0123, Time: 0.04s
Iteration 4482, Loss: 0.0184, Time: 0.04s
Iteration 4483, Loss: 0.0126, Time: 0.04s
Iteration 4484, Loss: 0.0176, Time: 0.05s
Iteration 4485, Loss: 0.0166, Time: 0.04s
Iteration 4486, Loss: 0.0245, Time: 0.05s
Iteration 4487, Loss: 0.0215, Time: 0.04s
Iteration 4488, Loss: 0.0105, Time: 0.05s
Iteration 4489, Loss: 0.0197, Time: 0.04s
Iteration 4490, Loss: 0.0153, Time: 0.04s
Iteration 4491, Loss: 0.0181, Time: 0.05s
Iteration 4492, Loss: 0.0112, Time: 0.04s
Iteration 4493, Loss: 0.0168, Time: 0.04s
Iteration 4494, Loss: 0.0113, Time: 0.04s
Iteration 4495, Loss: 0.0153, Time: 0.04s
Iteration 4496, Loss: 0.0153, Time: 0.05s
Iteration 4497, Loss: 0.0209, Time: 0.05s
Iteration 4498, Loss: 0.0184, Time: 0.05s
Iteration 4499, Loss: 0.0168, Time: 0.04s
Iteration 4500, Loss: 0.0166, Time: 0.04s
Iteration 4500, Loss: 0.0166, Time: 0.04s
Test Loss: 0.0352
Iteration 4501, Loss: 0.0152, Time: 0.05s
Iteration 4502, Loss: 0.0159, Time: 0.04s
Iteration 4503, Loss: 0.0170, Time: 0.04s
Iteration 4504, Loss: 0.0137, Time: 0.04s
Iteration 4505, Loss: 0.0169, Time: 0.04s
Iteration 4506, Loss: 0.0127, Time: 0.05s
Iteration 4507, Loss: 0.0199, Time: 0.04s
Iteration 4508, Loss: 0.0177, Time: 0.04s
Iteration 4509, Loss: 0.0187, Time: 0.05s
Iteration 4510, Loss: 0.0208, Time: 0.04s
Iteration 4511, Loss: 0.0182, Time: 0.04s
Iteration 4512, Loss: 0.0188, Time: 0.04s
Iteration 4513, Loss: 0.0165, Time: 0.04s
Iteration 4514, Loss: 0.0187, Time: 0.04s
Iteration 4515, Loss: 0.0132, Time: 0.04s
Iteration 4516, Loss: 0.0111, Time: 0.05s
```

```
Iteration 4517, Loss: 0.0076, Time: 0.04s
Iteration 4518, Loss: 0.0173, Time: 0.04s
Iteration 4519, Loss: 0.0219, Time: 0.04s
Iteration 4520, Loss: 0.0183, Time: 0.05s
Iteration 4521, Loss: 0.0213, Time: 0.04s
Iteration 4522, Loss: 0.0115, Time: 0.04s
Iteration 4523, Loss: 0.0155, Time: 0.05s
Iteration 4524, Loss: 0.0143, Time: 0.04s
Iteration 4525, Loss: 0.0154, Time: 0.04s
Iteration 4526, Loss: 0.0157, Time: 0.04s
Iteration 4527, Loss: 0.0163, Time: 0.05s
Iteration 4528, Loss: 0.0161, Time: 0.04s
Iteration 4529, Loss: 0.0130, Time: 0.04s
Iteration 4530, Loss: 0.0093, Time: 0.04s
Iteration 4531, Loss: 0.0087, Time: 0.04s
Iteration 4532, Loss: 0.0133, Time: 0.04s
Iteration 4533, Loss: 0.0125, Time: 0.04s
Iteration 4534, Loss: 0.0153, Time: 0.04s
Iteration 4535, Loss: 0.0122, Time: 0.04s
Iteration 4536, Loss: 0.0184, Time: 0.05s
Iteration 4537, Loss: 0.0167, Time: 0.04s
Iteration 4538, Loss: 0.0204, Time: 0.04s
Iteration 4539, Loss: 0.0202, Time: 0.05s
Iteration 4540, Loss: 0.0183, Time: 0.04s
Iteration 4541, Loss: 0.0217, Time: 0.04s
Iteration 4542, Loss: 0.0156, Time: 0.04s
Iteration 4543, Loss: 0.0171, Time: 0.04s
Iteration 4544, Loss: 0.0159, Time: 0.04s
Iteration 4545, Loss: 0.0177, Time: 0.05s
Iteration 4546, Loss: 0.0166, Time: 0.04s
Iteration 4547, Loss: 0.0149, Time: 0.05s
Iteration 4548, Loss: 0.0153, Time: 0.04s
Iteration 4549, Loss: 0.0150, Time: 0.05s
Iteration 4550, Loss: 0.0098, Time: 0.05s
Iteration 4551, Loss: 0.0222, Time: 0.05s
Iteration 4552, Loss: 0.0168, Time: 0.04s
Iteration 4553, Loss: 0.0139, Time: 0.04s
Iteration 4554, Loss: 0.0084, Time: 0.04s
Iteration 4555, Loss: 0.0171, Time: 0.05s
Iteration 4556, Loss: 0.0191, Time: 0.04s
Iteration 4557, Loss: 0.0124, Time: 0.04s
Iteration 4558, Loss: 0.0267, Time: 0.05s
Iteration 4559, Loss: 0.0188, Time: 0.04s
Iteration 4560, Loss: 0.0176, Time: 0.04s
Iteration 4561, Loss: 0.0116, Time: 0.04s
Iteration 4562, Loss: 0.0164, Time: 0.04s
Iteration 4563, Loss: 0.0128, Time: 0.05s
Iteration 4564, Loss: 0.0144, Time: 0.04s
```

```
Iteration 4565, Loss: 0.0150, Time: 0.05s
Iteration 4566, Loss: 0.0170, Time: 0.04s
Iteration 4567, Loss: 0.0118, Time: 0.04s
Iteration 4568, Loss: 0.0159, Time: 0.04s
Iteration 4569, Loss: 0.0151, Time: 0.04s
Iteration 4570, Loss: 0.0294, Time: 0.04s
Iteration 4571, Loss: 0.0174, Time: 0.04s
Iteration 4572, Loss: 0.0099, Time: 0.04s
Iteration 4573, Loss: 0.0146, Time: 0.04s
Iteration 4574, Loss: 0.0166, Time: 0.04s
Iteration 4575, Loss: 0.0132, Time: 0.04s
Iteration 4576, Loss: 0.0186, Time: 0.05s
Iteration 4577, Loss: 0.0216, Time: 0.05s
Iteration 4578, Loss: 0.0153, Time: 0.04s
Iteration 4579, Loss: 0.0200, Time: 0.04s
Iteration 4580, Loss: 0.0163, Time: 0.05s
Iteration 4581, Loss: 0.0148, Time: 0.05s
Iteration 4582, Loss: 0.0215, Time: 0.04s
Iteration 4583, Loss: 0.0159, Time: 0.04s
Iteration 4584, Loss: 0.0165, Time: 0.04s
Iteration 4585, Loss: 0.0174, Time: 0.04s
Iteration 4586, Loss: 0.0134, Time: 0.04s
Iteration 4587, Loss: 0.0124, Time: 0.04s
Iteration 4588, Loss: 0.0167, Time: 0.04s
Iteration 4589, Loss: 0.0120, Time: 0.05s
Iteration 4590, Loss: 0.0153, Time: 0.05s
Iteration 4591, Loss: 0.0154, Time: 0.05s
Iteration 4592, Loss: 0.0104, Time: 0.04s
Iteration 4593, Loss: 0.0129, Time: 0.05s
Iteration 4594, Loss: 0.0126, Time: 0.05s
Iteration 4595, Loss: 0.0202, Time: 0.05s
Iteration 4596, Loss: 0.0135, Time: -0.09s
Iteration 4597, Loss: 0.0117, Time: 0.06s
Iteration 4598, Loss: 0.0139, Time: 0.05s
Iteration 4599, Loss: 0.0184, Time: 0.06s
Iteration 4600, Loss: 0.0108, Time: 0.09s
Iteration 4600, Loss: 0.0108, Time: 0.10s
Test Loss: 0.0411
Iteration 4601, Loss: 0.0175, Time: 0.08s
Iteration 4602, Loss: 0.0115, Time: 0.07s
Iteration 4603, Loss: 0.0171, Time: 0.07s
Iteration 4604, Loss: 0.0218, Time: 0.05s
Iteration 4605, Loss: 0.0213, Time: 0.06s
Iteration 4606, Loss: 0.0160, Time: 0.05s
Iteration 4607, Loss: 0.0127, Time: 0.05s
Iteration 4608, Loss: 0.0207, Time: 0.05s
Iteration 4609, Loss: 0.0250, Time: 0.05s
Iteration 4610, Loss: 0.0204, Time: 0.04s
```

```
Iteration 4611, Loss: 0.0178, Time: 0.04s
Iteration 4612, Loss: 0.0142, Time: 0.04s
Iteration 4613, Loss: 0.0135, Time: 0.04s
Iteration 4614, Loss: 0.0269, Time: 0.04s
Iteration 4615, Loss: 0.0095, Time: 0.05s
Iteration 4616, Loss: 0.0178, Time: 0.05s
Iteration 4617, Loss: 0.0182, Time: 0.05s
Iteration 4618, Loss: 0.0162, Time: 0.04s
Iteration 4619, Loss: 0.0065, Time: 0.05s
Iteration 4620, Loss: 0.0193, Time: 0.05s
Iteration 4621, Loss: 0.0183, Time: 0.09s
Iteration 4622, Loss: 0.0152, Time: 0.06s
Iteration 4623, Loss: 0.0213, Time: 0.05s
Iteration 4624, Loss: 0.0187, Time: 0.05s
Iteration 4625, Loss: 0.0180, Time: 0.04s
Iteration 4626, Loss: 0.0172, Time: 0.05s
Iteration 4627, Loss: 0.0154, Time: 0.04s
Iteration 4628, Loss: 0.0131, Time: 0.04s
Iteration 4629, Loss: 0.0163, Time: 0.05s
Iteration 4630, Loss: 0.0141, Time: 0.06s
Iteration 4631, Loss: 0.0124, Time: 0.07s
Iteration 4632, Loss: 0.0111, Time: 0.05s
Iteration 4633, Loss: 0.0097, Time: 0.05s
Iteration 4634, Loss: 0.0242, Time: 0.05s
Iteration 4635, Loss: 0.0103, Time: 0.05s
Iteration 4636, Loss: 0.0136, Time: 0.05s
Iteration 4637, Loss: 0.0156, Time: 0.04s
Iteration 4638, Loss: 0.0125, Time: 0.05s
Iteration 4639, Loss: 0.0130, Time: 0.05s
Iteration 4640, Loss: 0.0178, Time: 0.06s
Iteration 4641, Loss: 0.0169, Time: 0.16s
Iteration 4642, Loss: 0.0164, Time: 0.05s
Iteration 4643, Loss: 0.0228, Time: 0.04s
Iteration 4644, Loss: 0.0146, Time: 0.04s
Iteration 4645, Loss: 0.0127, Time: 0.04s
Iteration 4646, Loss: 0.0144, Time: 0.08s
Iteration 4647, Loss: 0.0160, Time: 0.08s
Iteration 4648, Loss: 0.0112, Time: 0.04s
Iteration 4649, Loss: 0.0078, Time: 0.05s
Iteration 4650, Loss: 0.0079, Time: 0.05s
Iteration 4651, Loss: 0.0202, Time: 0.05s
Iteration 4652, Loss: 0.0165, Time: 0.04s
Iteration 4653, Loss: 0.0123, Time: 0.05s
Iteration 4654, Loss: 0.0195, Time: 0.05s
Iteration 4655, Loss: 0.0145, Time: 0.04s
Iteration 4656, Loss: 0.0156, Time: 0.05s
Iteration 4657, Loss: 0.0149, Time: 0.04s
Iteration 4658, Loss: 0.0161, Time: 0.05s
```

```
Iteration 4659, Loss: 0.0150, Time: 0.05s
Iteration 4660, Loss: 0.0170, Time: 0.05s
Iteration 4661, Loss: 0.0143, Time: 0.05s
Iteration 4662, Loss: 0.0205, Time: 0.05s
Iteration 4663, Loss: 0.0145, Time: 0.04s
Iteration 4664, Loss: 0.0196, Time: 0.04s
Iteration 4665, Loss: 0.0160, Time: 0.04s
Iteration 4666, Loss: 0.0170, Time: 0.05s
Iteration 4667, Loss: 0.0184, Time: 0.05s
Iteration 4668, Loss: 0.0165, Time: 0.05s
Iteration 4669, Loss: 0.0158, Time: 0.05s
Iteration 4670, Loss: 0.0112, Time: 0.07s
Iteration 4671, Loss: 0.0226, Time: 0.04s
Iteration 4672, Loss: 0.0149, Time: 0.04s
Iteration 4673, Loss: 0.0124, Time: 0.05s
Iteration 4674, Loss: 0.0148, Time: 0.04s
Iteration 4675, Loss: 0.0163, Time: 0.05s
Iteration 4676, Loss: 0.0128, Time: 0.04s
Iteration 4677, Loss: 0.0182, Time: 0.04s
Iteration 4678, Loss: 0.0163, Time: 0.04s
Iteration 4679, Loss: 0.0142, Time: 0.06s
Iteration 4680, Loss: 0.0145, Time: 0.05s
Iteration 4681, Loss: 0.0174, Time: 0.04s
Iteration 4682, Loss: 0.0120, Time: 0.05s
Iteration 4683, Loss: 0.0203, Time: 0.05s
Iteration 4684, Loss: 0.0160, Time: 0.04s
Iteration 4685, Loss: 0.0141, Time: 0.04s
Iteration 4686, Loss: 0.0185, Time: 0.04s
Iteration 4687, Loss: 0.0247, Time: 0.05s
Iteration 4688, Loss: 0.0167, Time: 0.04s
Iteration 4689, Loss: 0.0144, Time: 0.05s
Iteration 4690, Loss: 0.0158, Time: 0.05s
Iteration 4691, Loss: 0.0155, Time: 0.04s
Iteration 4692, Loss: 0.0193, Time: 0.04s
Iteration 4693, Loss: 0.0140, Time: 0.06s
Iteration 4694, Loss: 0.0221, Time: 0.05s
Iteration 4695, Loss: 0.0111, Time: 0.05s
Iteration 4696, Loss: 0.0184, Time: 0.05s
Iteration 4697, Loss: 0.0143, Time: 0.05s
Iteration 4698, Loss: 0.0161, Time: 0.05s
Iteration 4699, Loss: 0.0128, Time: 0.05s
Iteration 4700, Loss: 0.0090, Time: 0.05s
Iteration 4700, Loss: 0.0090, Time: 0.06s
Test Loss: 0.0404
Iteration 4701, Loss: 0.0179, Time: 0.05s
Iteration 4702, Loss: 0.0138, Time: 0.05s
Iteration 4703, Loss: 0.0122, Time: 0.05s
Iteration 4704, Loss: 0.0141, Time: 0.05s
```

```
Iteration 4705, Loss: 0.0152, Time: 0.05s
Iteration 4706, Loss: 0.0170, Time: 0.05s
Iteration 4707, Loss: 0.0131, Time: 0.05s
Iteration 4708, Loss: 0.0125, Time: 0.05s
Iteration 4709, Loss: 0.0168, Time: 0.05s
Iteration 4710, Loss: 0.0135, Time: 0.04s
Iteration 4711, Loss: 0.0102, Time: 0.05s
Iteration 4712, Loss: 0.0130, Time: 0.04s
Iteration 4713, Loss: 0.0146, Time: 0.04s
Iteration 4714, Loss: 0.0127, Time: 0.05s
Iteration 4715, Loss: 0.0199, Time: 0.05s
Iteration 4716, Loss: 0.0210, Time: 0.04s
Iteration 4717, Loss: 0.0141, Time: 0.05s
Iteration 4718, Loss: 0.0140, Time: 0.04s
Iteration 4719, Loss: 0.0117, Time: 0.04s
Iteration 4720, Loss: 0.0120, Time: 0.05s
Iteration 4721, Loss: 0.0266, Time: 0.04s
Iteration 4722, Loss: 0.0155, Time: 0.04s
Iteration 4723, Loss: 0.0150, Time: 0.04s
Iteration 4724, Loss: 0.0119, Time: 0.05s
Iteration 4725, Loss: 0.0177, Time: 0.05s
Iteration 4726, Loss: 0.0182, Time: 0.04s
Iteration 4727, Loss: 0.0173, Time: 0.04s
Iteration 4728, Loss: 0.0141, Time: 0.04s
Iteration 4729, Loss: 0.0138, Time: 0.04s
Iteration 4730, Loss: 0.0193, Time: 0.05s
Iteration 4731, Loss: 0.0172, Time: 0.04s
Iteration 4732, Loss: 0.0159, Time: 0.04s
Iteration 4733, Loss: 0.0183, Time: 0.04s
Iteration 4734, Loss: 0.0157, Time: 0.05s
Iteration 4735, Loss: 0.0232, Time: 0.05s
Iteration 4736, Loss: 0.0211, Time: 0.04s
Iteration 4737, Loss: 0.0105, Time: 0.04s
Iteration 4738, Loss: 0.0153, Time: 0.05s
Iteration 4739, Loss: 0.0157, Time: 0.04s
Iteration 4740, Loss: 0.0208, Time: 0.05s
Iteration 4741, Loss: 0.0195, Time: 0.05s
Iteration 4742, Loss: 0.0152, Time: 0.05s
Iteration 4743, Loss: 0.0175, Time: 0.05s
Iteration 4744, Loss: 0.0152, Time: 0.05s
Iteration 4745, Loss: 0.0145, Time: 0.05s
Iteration 4746, Loss: 0.0173, Time: 0.05s
Iteration 4747, Loss: 0.0196, Time: 0.06s
Iteration 4748, Loss: 0.0116, Time: 0.04s
Iteration 4749, Loss: 0.0150, Time: 0.05s
Iteration 4750, Loss: 0.0225, Time: 0.05s
Iteration 4751, Loss: 0.0095, Time: 0.05s
Iteration 4752, Loss: 0.0186, Time: 0.05s
```

```
Iteration 4753, Loss: 0.0155, Time: 0.04s
Iteration 4754, Loss: 0.0185, Time: 0.04s
Iteration 4755, Loss: 0.0221, Time: 0.05s
Iteration 4756, Loss: 0.0184, Time: 0.04s
Iteration 4757, Loss: 0.0147, Time: 0.05s
Iteration 4758, Loss: 0.0224, Time: 0.04s
Iteration 4759, Loss: 0.0134, Time: 0.05s
Iteration 4760, Loss: 0.0164, Time: 0.05s
Iteration 4761, Loss: 0.0207, Time: 0.05s
Iteration 4762, Loss: 0.0249, Time: 0.04s
Iteration 4763, Loss: 0.0234, Time: 0.04s
Iteration 4764, Loss: 0.0132, Time: 0.05s
Iteration 4765, Loss: 0.0153, Time: 0.05s
Iteration 4766, Loss: 0.0170, Time: 0.05s
Iteration 4767, Loss: 0.0228, Time: 0.05s
Iteration 4768, Loss: 0.0174, Time: 0.05s
Iteration 4769, Loss: 0.0162, Time: 0.04s
Iteration 4770, Loss: 0.0141, Time: 0.05s
Iteration 4771, Loss: 0.0146, Time: 0.04s
Iteration 4772, Loss: 0.0119, Time: 0.05s
Iteration 4773, Loss: 0.0154, Time: 0.05s
Iteration 4774, Loss: 0.0216, Time: 0.05s
Iteration 4775, Loss: 0.0183, Time: 0.04s
Iteration 4776, Loss: 0.0120, Time: 0.04s
Iteration 4777, Loss: 0.0249, Time: 0.05s
Iteration 4778, Loss: 0.0226, Time: 0.05s
Iteration 4779, Loss: 0.0134, Time: 0.04s
Iteration 4780, Loss: 0.0101, Time: 0.05s
Iteration 4781, Loss: 0.0223, Time: 0.05s
Iteration 4782, Loss: 0.0233, Time: 0.05s
Iteration 4783, Loss: 0.0177, Time: 0.04s
Iteration 4784, Loss: 0.0174, Time: 0.05s
Iteration 4785, Loss: 0.0156, Time: 0.05s
Iteration 4786, Loss: 0.0128, Time: 0.05s
Iteration 4787, Loss: 0.0131, Time: 0.05s
Iteration 4788, Loss: 0.0186, Time: 0.04s
Iteration 4789, Loss: 0.0124, Time: 0.04s
Iteration 4790, Loss: 0.0094, Time: 0.05s
Iteration 4791, Loss: 0.0161, Time: 0.06s
Iteration 4792, Loss: 0.0182, Time: 0.05s
Iteration 4793, Loss: 0.0168, Time: 0.05s
Iteration 4794, Loss: 0.0177, Time: 0.06s
Iteration 4795, Loss: 0.0152, Time: 0.07s
Iteration 4796, Loss: 0.0163, Time: 0.07s
Iteration 4797, Loss: 0.0160, Time: 0.06s
Iteration 4798, Loss: 0.0163, Time: 0.05s
Iteration 4799, Loss: 0.0158, Time: 0.21s
Iteration 4800, Loss: 0.0141, Time: 0.09s
```

```
Iteration 4800, Loss: 0.0141, Time: 0.09s
Test Loss: 0.0421
Iteration 4801, Loss: 0.0160, Time: 0.07s
Iteration 4802, Loss: 0.0113, Time: 0.09s
Iteration 4803, Loss: 0.0220, Time: 0.08s
Iteration 4804, Loss: 0.0145, Time: 0.08s
Iteration 4805, Loss: 0.0101, Time: 0.06s
Iteration 4806, Loss: 0.0170, Time: 0.06s
Iteration 4807, Loss: 0.0097, Time: 0.05s
Iteration 4808, Loss: 0.0182, Time: 0.05s
Iteration 4809, Loss: 0.0136, Time: 0.05s
Iteration 4810, Loss: 0.0251, Time: 0.05s
Iteration 4811, Loss: 0.0124, Time: 0.05s
Iteration 4812, Loss: 0.0122, Time: 0.05s
Iteration 4813, Loss: 0.0143, Time: 0.05s
Iteration 4814, Loss: 0.0118, Time: 0.06s
Iteration 4815, Loss: 0.0117, Time: 0.06s
Iteration 4816, Loss: 0.0158, Time: 0.05s
Iteration 4817, Loss: 0.0192, Time: 0.05s
Iteration 4818, Loss: 0.0148, Time: 0.05s
Iteration 4819, Loss: 0.0132, Time: 0.05s
Iteration 4820, Loss: 0.0192, Time: 0.05s
Iteration 4821, Loss: 0.0220, Time: 0.05s
Iteration 4822, Loss: 0.0263, Time: 0.07s
Iteration 4823, Loss: 0.0153, Time: 0.07s
Iteration 4824, Loss: 0.0168, Time: 0.06s
Iteration 4825, Loss: 0.0109, Time: 0.06s
Iteration 4826, Loss: 0.0194, Time: 0.06s
Iteration 4827, Loss: 0.0130, Time: 0.06s
Iteration 4828, Loss: 0.0148, Time: 0.08s
Iteration 4829, Loss: 0.0112, Time: 0.05s
Iteration 4830, Loss: 0.0175, Time: 0.06s
Iteration 4831, Loss: 0.0118, Time: 0.07s
Iteration 4832, Loss: 0.0115, Time: 0.09s
Iteration 4833, Loss: 0.0169, Time: 0.06s
Iteration 4834, Loss: 0.0111, Time: 0.14s
Iteration 4835, Loss: 0.0179, Time: 0.06s
Iteration 4836, Loss: 0.0216, Time: 0.24s
Iteration 4837, Loss: 0.0143, Time: 0.09s
Iteration 4838, Loss: 0.0175, Time: 0.09s
Iteration 4839, Loss: 0.0231, Time: 0.06s
Iteration 4840, Loss: 0.0205, Time: 0.09s
Iteration 4841, Loss: 0.0139, Time: 0.13s
Iteration 4842, Loss: 0.0208, Time: 0.17s
Iteration 4843, Loss: 0.0139, Time: 0.07s
Iteration 4844, Loss: 0.0188, Time: 0.07s
Iteration 4845, Loss: 0.0151, Time: 0.06s
Iteration 4846, Loss: 0.0189, Time: 0.06s
```

```
Iteration 4847, Loss: 0.0134, Time: 0.07s
Iteration 4848, Loss: 0.0182, Time: 0.11s
Iteration 4849, Loss: 0.0169, Time: 0.19s
Iteration 4850, Loss: 0.0133, Time: 0.09s
Iteration 4851, Loss: 0.0155, Time: 0.06s
Iteration 4852, Loss: 0.0167, Time: 0.06s
Iteration 4853, Loss: 0.0121, Time: 0.07s
Iteration 4854, Loss: 0.0152, Time: 0.07s
Iteration 4855, Loss: 0.0146, Time: 0.07s
Iteration 4856, Loss: 0.0150, Time: 0.07s
Iteration 4857, Loss: 0.0127, Time: 0.07s
Iteration 4858, Loss: 0.0122, Time: 0.07s
Iteration 4859, Loss: 0.0153, Time: 0.07s
Iteration 4860, Loss: 0.0183, Time: 0.05s
Iteration 4861, Loss: 0.0192, Time: 0.06s
Iteration 4862, Loss: 0.0142, Time: 0.07s
Iteration 4863, Loss: 0.0137, Time: 0.06s
Iteration 4864, Loss: 0.0202, Time: 0.06s
Iteration 4865, Loss: 0.0157, Time: 0.07s
Iteration 4866, Loss: 0.0222, Time: 0.07s
Iteration 4867, Loss: 0.0167, Time: 0.07s
Iteration 4868, Loss: 0.0172, Time: 0.06s
Iteration 4869, Loss: 0.0186, Time: 0.05s
Iteration 4870, Loss: 0.0152, Time: 0.07s
Iteration 4871, Loss: 0.0159, Time: 0.07s
Iteration 4872, Loss: 0.0138, Time: 0.06s
Iteration 4873, Loss: 0.0151, Time: 0.06s
Iteration 4874, Loss: 0.0153, Time: 0.07s
Iteration 4875, Loss: 0.0200, Time: 0.09s
Iteration 4876, Loss: 0.0087, Time: 0.06s
Iteration 4877, Loss: 0.0127, Time: 0.05s
Iteration 4878, Loss: 0.0108, Time: 0.06s
Iteration 4879, Loss: 0.0197, Time: 0.05s
Iteration 4880, Loss: 0.0145, Time: 0.05s
Iteration 4881, Loss: 0.0161, Time: 0.06s
Iteration 4882, Loss: 0.0156, Time: 0.05s
Iteration 4883, Loss: 0.0184, Time: 0.05s
Iteration 4884, Loss: 0.0288, Time: 0.04s
Iteration 4885, Loss: 0.0135, Time: 0.05s
Iteration 4886, Loss: 0.0167, Time: 0.04s
Iteration 4887, Loss: 0.0202, Time: 0.06s
Iteration 4888, Loss: 0.0135, Time: 0.07s
Iteration 4889, Loss: 0.0186, Time: 0.06s
Iteration 4890, Loss: 0.0191, Time: 0.05s
Iteration 4891, Loss: 0.0151, Time: 0.04s
Iteration 4892, Loss: 0.0139, Time: 0.04s
Iteration 4893, Loss: 0.0141, Time: 0.05s
Iteration 4894, Loss: 0.0134, Time: 0.05s
```

```
Iteration 4895, Loss: 0.0174, Time: 0.04s
Iteration 4896, Loss: 0.0187, Time: 0.04s
Iteration 4897, Loss: 0.0169, Time: 0.05s
Iteration 4898, Loss: 0.0128, Time: 0.05s
Iteration 4899, Loss: 0.0185, Time: 0.04s
Iteration 4900, Loss: 0.0149, Time: 0.05s
Iteration 4900, Loss: 0.0149, Time: 0.05s
Test Loss: 0.0401
Iteration 4901, Loss: 0.0151, Time: 0.06s
Iteration 4902, Loss: 0.0099, Time: 0.05s
Iteration 4903, Loss: 0.0133, Time: 0.06s
Iteration 4904, Loss: 0.0122, Time: 0.04s
Iteration 4905, Loss: 0.0161, Time: 0.05s
Iteration 4906, Loss: 0.0120, Time: 0.05s
Iteration 4907, Loss: 0.0153, Time: 0.05s
Iteration 4908, Loss: 0.0194, Time: 0.04s
Iteration 4909, Loss: 0.0116, Time: 0.06s
Iteration 4910, Loss: 0.0162, Time: 0.05s
Iteration 4911, Loss: 0.0138, Time: 0.06s
Iteration 4912, Loss: 0.0126, Time: 0.05s
Iteration 4913, Loss: 0.0165, Time: 0.06s
Iteration 4914, Loss: 0.0207, Time: 0.05s
Iteration 4915, Loss: 0.0185, Time: 0.05s
Iteration 4916, Loss: 0.0100, Time: 0.04s
Iteration 4917, Loss: 0.0169, Time: 0.05s
Iteration 4918, Loss: 0.0197, Time: 0.05s
Iteration 4919, Loss: 0.0130, Time: 0.05s
Iteration 4920, Loss: 0.0151, Time: 0.06s
Iteration 4921, Loss: 0.0195, Time: 0.05s
Iteration 4922, Loss: 0.0240, Time: 0.04s
Iteration 4923, Loss: 0.0103, Time: 0.06s
Iteration 4924, Loss: 0.0137, Time: 0.05s
Iteration 4925, Loss: 0.0166, Time: 0.05s
Iteration 4926, Loss: 0.0099, Time: 0.05s
Iteration 4927, Loss: 0.0144, Time: 0.05s
Iteration 4928, Loss: 0.0142, Time: 0.06s
Iteration 4929, Loss: 0.0196, Time: 0.05s
Iteration 4930, Loss: 0.0123, Time: 0.04s
Iteration 4931, Loss: 0.0124, Time: 0.05s
Iteration 4932, Loss: 0.0158, Time: 0.04s
Iteration 4933, Loss: 0.0137, Time: 0.05s
Iteration 4934, Loss: 0.0149, Time: 0.05s
Iteration 4935, Loss: 0.0088, Time: 0.05s
Iteration 4936, Loss: 0.0167, Time: 0.05s
Iteration 4937, Loss: 0.0120, Time: 0.05s
Iteration 4938, Loss: 0.0108, Time: 0.05s
Iteration 4939, Loss: 0.0193, Time: 0.04s
Iteration 4940, Loss: 0.0139, Time: 0.05s
```

```
Iteration 4941, Loss: 0.0149, Time: 0.04s
Iteration 4942, Loss: 0.0118, Time: 0.05s
Iteration 4943, Loss: 0.0165, Time: 0.05s
Iteration 4944, Loss: 0.0126, Time: 0.05s
Iteration 4945, Loss: 0.0170, Time: 0.04s
Iteration 4946, Loss: 0.0155, Time: 0.05s
Iteration 4947, Loss: 0.0153, Time: 0.06s
Iteration 4948, Loss: 0.0114, Time: 0.06s
Iteration 4949, Loss: 0.0202, Time: 0.05s
Iteration 4950, Loss: 0.0202, Time: 0.04s
Iteration 4951, Loss: 0.0161, Time: 0.04s
Iteration 4952, Loss: 0.0187, Time: 0.05s
Iteration 4953, Loss: 0.0178, Time: 0.05s
Iteration 4954, Loss: 0.0166, Time: 0.04s
Iteration 4955, Loss: 0.0117, Time: 0.05s
Iteration 4956, Loss: 0.0227, Time: 0.04s
Iteration 4957, Loss: 0.0156, Time: 0.04s
Iteration 4958, Loss: 0.0138, Time: 0.05s
Iteration 4959, Loss: 0.0207, Time: 0.05s
Iteration 4960, Loss: 0.0178, Time: 0.05s
Iteration 4961, Loss: 0.0169, Time: 0.05s
Iteration 4962, Loss: 0.0146, Time: 0.07s
Iteration 4963, Loss: 0.0130, Time: 0.05s
Iteration 4964, Loss: 0.0143, Time: 0.06s
Iteration 4965, Loss: 0.0227, Time: 0.07s
Iteration 4966, Loss: 0.0149, Time: 0.05s
Iteration 4967, Loss: 0.0093, Time: 0.07s
Iteration 4968, Loss: 0.0132, Time: 0.06s
Iteration 4969, Loss: 0.0136, Time: 0.06s
Iteration 4970, Loss: 0.0122, Time: 0.04s
Iteration 4971, Loss: 0.0113, Time: 0.05s
Iteration 4972, Loss: 0.0121, Time: 0.05s
Iteration 4973, Loss: 0.0077, Time: 0.05s
Iteration 4974, Loss: 0.0209, Time: 0.05s
Iteration 4975, Loss: 0.0093, Time: 0.04s
Iteration 4976, Loss: 0.0199, Time: 0.04s
Iteration 4977, Loss: 0.0106, Time: 0.04s
Iteration 4978, Loss: 0.0182, Time: 0.03s
Iteration 4979, Loss: 0.0151, Time: 0.03s
Iteration 4980, Loss: 0.0129, Time: 0.03s
Iteration 4981, Loss: 0.0157, Time: 0.03s
Iteration 4982, Loss: 0.0129, Time: 0.03s
Iteration 4983, Loss: 0.0153, Time: 0.03s
Iteration 4984, Loss: 0.0150, Time: 0.03s
Iteration 4985, Loss: 0.0211, Time: 0.03s
Iteration 4986, Loss: 0.0182, Time: 0.03s
Iteration 4987, Loss: 0.0103, Time: 0.03s
Iteration 4988, Loss: 0.0138, Time: 0.03s
```

```
Iteration 4989, Loss: 0.0184, Time: 0.03s
Iteration 4990, Loss: 0.0123, Time: 0.03s
Iteration 4991, Loss: 0.0168, Time: 0.03s
Iteration 4992, Loss: 0.0158, Time: 0.03s
Iteration 4993, Loss: 0.0140, Time: 0.03s
Iteration 4994, Loss: 0.0193, Time: 0.03s
Iteration 4995, Loss: 0.0164, Time: 0.03s
Iteration 4996, Loss: 0.0152, Time: 0.03s
Iteration 4997, Loss: 0.0178, Time: 0.03s
Iteration 4998, Loss: 0.0232, Time: 0.03s
Iteration 4999, Loss: 0.0086, Time: 0.03s
Iteration 5000, Loss: 0.0185, Time: 0.03s
Iteration 5000, Loss: 0.0185, Time: 0.03s
Test Loss: 0.0358
Iteration 5001, Loss: 0.0182, Time: 0.03s
Iteration 5002, Loss: 0.0213, Time: 0.03s
Iteration 5003, Loss: 0.0127, Time: 0.03s
Iteration 5004, Loss: 0.0145, Time: 0.03s
Iteration 5005, Loss: 0.0138, Time: 0.03s
Iteration 5006, Loss: 0.0159, Time: 0.03s
Iteration 5007, Loss: 0.0163, Time: 0.03s
Iteration 5008, Loss: 0.0156, Time: 0.03s
Iteration 5009, Loss: 0.0141, Time: 0.03s
Iteration 5010, Loss: 0.0167, Time: 0.03s
Iteration 5011, Loss: 0.0209, Time: 0.03s
Iteration 5012, Loss: 0.0173, Time: 0.03s
Iteration 5013, Loss: 0.0123, Time: 0.03s
Iteration 5014, Loss: 0.0176, Time: 0.03s
Iteration 5015, Loss: 0.0124, Time: 0.03s
Iteration 5016, Loss: 0.0146, Time: 0.03s
Iteration 5017, Loss: 0.0102, Time: 0.03s
Iteration 5018, Loss: 0.0231, Time: 0.03s
Iteration 5019, Loss: 0.0178, Time: 0.03s
Iteration 5020, Loss: 0.0200, Time: 0.03s
Iteration 5021, Loss: 0.0113, Time: 0.03s
Iteration 5022, Loss: 0.0152, Time: 0.03s
Iteration 5023, Loss: 0.0152, Time: 0.03s
Iteration 5024, Loss: 0.0080, Time: 0.03s
Iteration 5025, Loss: 0.0157, Time: 0.03s
Iteration 5026, Loss: 0.0208, Time: 0.03s
Iteration 5027, Loss: 0.0168, Time: 0.03s
Iteration 5028, Loss: 0.0129, Time: 0.03s
Iteration 5029, Loss: 0.0148, Time: 0.03s
Iteration 5030, Loss: 0.0181, Time: 0.03s
Iteration 5031, Loss: 0.0121, Time: 0.03s
Iteration 5032, Loss: 0.0129, Time: 0.03s
Iteration 5033, Loss: 0.0136, Time: 0.03s
Iteration 5034, Loss: 0.0138, Time: 0.03s
```

```
Iteration 5035, Loss: 0.0139, Time: 0.03s
Iteration 5036, Loss: 0.0089, Time: 0.03s
Iteration 5037, Loss: 0.0112, Time: 0.03s
Iteration 5038, Loss: 0.0163, Time: 0.03s
Iteration 5039, Loss: 0.0155, Time: 0.03s
Iteration 5040, Loss: 0.0156, Time: 0.03s
Iteration 5041, Loss: 0.0103, Time: 0.03s
Iteration 5042, Loss: 0.0231, Time: 0.03s
Iteration 5043, Loss: 0.0096, Time: 0.03s
Iteration 5044, Loss: 0.0130, Time: 0.03s
Iteration 5045, Loss: 0.0127, Time: 0.03s
Iteration 5046, Loss: 0.0086, Time: 0.03s
Iteration 5047, Loss: 0.0157, Time: 0.03s
Iteration 5048, Loss: 0.0153, Time: 0.03s
Iteration 5049, Loss: 0.0103, Time: 0.03s
Iteration 5050, Loss: 0.0107, Time: 0.03s
Iteration 5051, Loss: 0.0166, Time: 0.03s
Iteration 5052, Loss: 0.0128, Time: 0.03s
Iteration 5053, Loss: 0.0154, Time: 0.03s
Iteration 5054, Loss: 0.0091, Time: 0.03s
Iteration 5055, Loss: 0.0120, Time: 0.03s
Iteration 5056, Loss: 0.0133, Time: 0.03s
Iteration 5057, Loss: 0.0136, Time: 0.03s
Iteration 5058, Loss: 0.0157, Time: 0.03s
Iteration 5059, Loss: 0.0171, Time: 0.03s
Iteration 5060, Loss: 0.0114, Time: 0.03s
Iteration 5061, Loss: 0.0153, Time: 0.03s
Iteration 5062, Loss: 0.0235, Time: 0.03s
Iteration 5063, Loss: 0.0125, Time: 0.03s
Iteration 5064, Loss: 0.0126, Time: 0.03s
Iteration 5065, Loss: 0.0112, Time: 0.03s
Iteration 5066, Loss: 0.0141, Time: 0.03s
Iteration 5067, Loss: 0.0099, Time: 0.03s
Iteration 5068, Loss: 0.0162, Time: 0.03s
Iteration 5069, Loss: 0.0178, Time: 0.03s
Iteration 5070, Loss: 0.0258, Time: 0.03s
Iteration 5071, Loss: 0.0095, Time: 0.03s
Iteration 5072, Loss: 0.0134, Time: 0.03s
Iteration 5073, Loss: 0.0096, Time: 0.03s
Iteration 5074, Loss: 0.0140, Time: 0.03s
Iteration 5075, Loss: 0.0203, Time: 0.03s
Iteration 5076, Loss: 0.0114, Time: 0.03s
Iteration 5077, Loss: 0.0166, Time: 0.03s
Iteration 5078, Loss: 0.0143, Time: 0.03s
Iteration 5079, Loss: 0.0139, Time: 0.03s
Iteration 5080, Loss: 0.0165, Time: 0.03s
Iteration 5081, Loss: 0.0154, Time: 0.04s
Iteration 5082, Loss: 0.0151, Time: 0.04s
```

```
Iteration 5083, Loss: 0.0091, Time: 0.04s
Iteration 5084, Loss: 0.0140, Time: 0.03s
Iteration 5085, Loss: 0.0173, Time: 0.04s
Iteration 5086, Loss: 0.0126, Time: 0.04s
Iteration 5087, Loss: 0.0140, Time: 0.03s
Iteration 5088, Loss: 0.0120, Time: 0.03s
Iteration 5089, Loss: 0.0136, Time: 0.03s
Iteration 5090, Loss: 0.0143, Time: 0.03s
Iteration 5091, Loss: 0.0106, Time: 0.04s
Iteration 5092, Loss: 0.0129, Time: 0.03s
Iteration 5093, Loss: 0.0121, Time: 0.03s
Iteration 5094, Loss: 0.0148, Time: 0.03s
Iteration 5095, Loss: 0.0129, Time: 0.03s
Iteration 5096, Loss: 0.0174, Time: 0.03s
Iteration 5097, Loss: 0.0164, Time: 0.03s
Iteration 5098, Loss: 0.0131, Time: 0.03s
Iteration 5099, Loss: 0.0161, Time: 0.03s
Iteration 5100, Loss: 0.0127, Time: 0.03s
Iteration 5100, Loss: 0.0127, Time: 0.03s
Test Loss: 0.0316
Iteration 5101, Loss: 0.0159, Time: 0.03s
Iteration 5102, Loss: 0.0166, Time: 0.03s
Iteration 5103, Loss: 0.0126, Time: 0.03s
Iteration 5104, Loss: 0.0173, Time: 0.03s
Iteration 5105, Loss: 0.0106, Time: 0.03s
Iteration 5106, Loss: 0.0133, Time: 0.03s
Iteration 5107, Loss: 0.0169, Time: 0.03s
Iteration 5108, Loss: 0.0153, Time: 0.03s
Iteration 5109, Loss: 0.0153, Time: 0.03s
Iteration 5110, Loss: 0.0077, Time: 0.03s
Iteration 5111, Loss: 0.0190, Time: 0.04s
Iteration 5112, Loss: 0.0112, Time: 0.03s
Iteration 5113, Loss: 0.0116, Time: 0.03s
Iteration 5114, Loss: 0.0132, Time: 0.03s
Iteration 5115, Loss: 0.0119, Time: 0.04s
Iteration 5116, Loss: 0.0118, Time: 0.04s
Iteration 5117, Loss: 0.0136, Time: 0.03s
Iteration 5118, Loss: 0.0198, Time: 0.03s
Iteration 5119, Loss: 0.0109, Time: 0.03s
Iteration 5120, Loss: 0.0183, Time: 0.04s
Iteration 5121, Loss: 0.0176, Time: 0.04s
Iteration 5122, Loss: 0.0131, Time: 0.03s
Iteration 5123, Loss: 0.0184, Time: 0.03s
Iteration 5124, Loss: 0.0151, Time: 0.03s
Iteration 5125, Loss: 0.0104, Time: 0.03s
Iteration 5126, Loss: 0.0220, Time: 0.03s
Iteration 5127, Loss: 0.0151, Time: 0.03s
Iteration 5128, Loss: 0.0149, Time: 0.03s
```

```
Iteration 5129, Loss: 0.0118, Time: 0.03s
Iteration 5130, Loss: 0.0150, Time: 0.04s
Iteration 5131, Loss: 0.0152, Time: 0.03s
Iteration 5132, Loss: 0.0153, Time: 0.03s
Iteration 5133, Loss: 0.0144, Time: 0.03s
Iteration 5134, Loss: 0.0149, Time: 0.03s
Iteration 5135, Loss: 0.0089, Time: 0.03s
Iteration 5136, Loss: 0.0140, Time: 0.03s
Iteration 5137, Loss: 0.0147, Time: 0.03s
Iteration 5138, Loss: 0.0172, Time: 0.03s
Iteration 5139, Loss: 0.0125, Time: 0.03s
Iteration 5140, Loss: 0.0134, Time: 0.03s
Iteration 5141, Loss: 0.0134, Time: 0.03s
Iteration 5142, Loss: 0.0180, Time: 0.03s
Iteration 5143, Loss: 0.0148, Time: 0.03s
Iteration 5144, Loss: 0.0125, Time: 0.03s
Iteration 5145, Loss: 0.0124, Time: 0.03s
Iteration 5146, Loss: 0.0151, Time: 0.03s
Iteration 5147, Loss: 0.0077, Time: 0.03s
Iteration 5148, Loss: 0.0165, Time: 0.03s
Iteration 5149, Loss: 0.0214, Time: 0.03s
Iteration 5150, Loss: 0.0213, Time: 0.03s
Iteration 5151, Loss: 0.0111, Time: 0.03s
Iteration 5152, Loss: 0.0070, Time: 0.03s
Iteration 5153, Loss: 0.0225, Time: 0.03s
Iteration 5154, Loss: 0.0170, Time: 0.03s
Iteration 5155, Loss: 0.0181, Time: 0.03s
Iteration 5156, Loss: 0.0134, Time: 0.03s
Iteration 5157, Loss: 0.0086, Time: 0.03s
Iteration 5158, Loss: 0.0127, Time: 0.03s
Iteration 5159, Loss: 0.0172, Time: 0.03s
Iteration 5160, Loss: 0.0136, Time: 0.03s
Iteration 5161, Loss: 0.0133, Time: 0.03s
Iteration 5162, Loss: 0.0217, Time: 0.03s
Iteration 5163, Loss: 0.0125, Time: 0.03s
Iteration 5164, Loss: 0.0158, Time: 0.03s
Iteration 5165, Loss: 0.0160, Time: 0.03s
Iteration 5166, Loss: 0.0188, Time: 0.03s
Iteration 5167, Loss: 0.0178, Time: 0.03s
Iteration 5168, Loss: 0.0178, Time: 0.03s
Iteration 5169, Loss: 0.0197, Time: 0.03s
Iteration 5170, Loss: 0.0120, Time: 0.03s
Iteration 5171, Loss: 0.0128, Time: 0.03s
Iteration 5172, Loss: 0.0111, Time: 0.03s
Iteration 5173, Loss: 0.0134, Time: 0.03s
Iteration 5174, Loss: 0.0162, Time: 0.03s
Iteration 5175, Loss: 0.0101, Time: 0.03s
Iteration 5176, Loss: 0.0151, Time: 0.03s
```

```
Iteration 5177, Loss: 0.0135, Time: 0.03s
Iteration 5178, Loss: 0.0143, Time: 0.03s
Iteration 5179, Loss: 0.0109, Time: 0.03s
Iteration 5180, Loss: 0.0142, Time: 0.03s
Iteration 5181, Loss: 0.0185, Time: 0.03s
Iteration 5182, Loss: 0.0082, Time: 0.03s
Iteration 5183, Loss: 0.0119, Time: 0.03s
Iteration 5184, Loss: 0.0112, Time: 0.03s
Iteration 5185, Loss: 0.0145, Time: 0.03s
Iteration 5186, Loss: 0.0131, Time: 0.03s
Iteration 5187, Loss: 0.0111, Time: 0.03s
Iteration 5188, Loss: 0.0207, Time: 0.03s
Iteration 5189, Loss: 0.0119, Time: 0.03s
Iteration 5190, Loss: 0.0175, Time: 0.03s
Iteration 5191, Loss: 0.0085, Time: 0.03s
Iteration 5192, Loss: 0.0173, Time: 0.03s
Iteration 5193, Loss: 0.0128, Time: 0.03s
Iteration 5194, Loss: 0.0104, Time: 0.03s
Iteration 5195, Loss: 0.0093, Time: 0.03s
Iteration 5196, Loss: 0.0165, Time: 0.03s
Iteration 5197, Loss: 0.0175, Time: 0.03s
Iteration 5198, Loss: 0.0128, Time: 0.03s
Iteration 5199, Loss: 0.0125, Time: 0.03s
Iteration 5200, Loss: 0.0108, Time: 0.03s
Iteration 5200, Loss: 0.0108, Time: 0.03s
Test Loss: 0.0358
Iteration 5201, Loss: 0.0160, Time: 0.03s
Iteration 5202, Loss: 0.0144, Time: 0.03s
Iteration 5203, Loss: 0.0111, Time: 0.03s
Iteration 5204, Loss: 0.0181, Time: 0.03s
Iteration 5205, Loss: 0.0107, Time: 0.03s
Iteration 5206, Loss: 0.0141, Time: 0.03s
Iteration 5207, Loss: 0.0210, Time: 0.04s
Iteration 5208, Loss: 0.0141, Time: 0.03s
Iteration 5209, Loss: 0.0137, Time: 0.04s
Iteration 5210, Loss: 0.0092, Time: 0.03s
Iteration 5211, Loss: 0.0150, Time: 0.03s
Iteration 5212, Loss: 0.0168, Time: 0.03s
Iteration 5213, Loss: 0.0163, Time: 0.03s
Iteration 5214, Loss: 0.0127, Time: 0.03s
Iteration 5215, Loss: 0.0100, Time: 0.03s
Iteration 5216, Loss: 0.0158, Time: 0.03s
Iteration 5217, Loss: 0.0111, Time: 0.03s
Iteration 5218, Loss: 0.0190, Time: 0.03s
Iteration 5219, Loss: 0.0146, Time: 0.03s
Iteration 5220, Loss: 0.0149, Time: 0.03s
Iteration 5221, Loss: 0.0089, Time: 0.03s
Iteration 5222, Loss: 0.0097, Time: 0.03s
```

```
Iteration 5223, Loss: 0.0124, Time: 0.04s
Iteration 5224, Loss: 0.0115, Time: 0.04s
Iteration 5225, Loss: 0.0116, Time: 0.03s
Iteration 5226, Loss: 0.0192, Time: 0.04s
Iteration 5227, Loss: 0.0161, Time: 0.04s
Iteration 5228, Loss: 0.0144, Time: 0.04s
Iteration 5229, Loss: 0.0167, Time: 0.04s
Iteration 5230, Loss: 0.0183, Time: 0.04s
Iteration 5231, Loss: 0.0082, Time: 0.04s
Iteration 5232, Loss: 0.0149, Time: 0.03s
Iteration 5233, Loss: 0.0154, Time: 0.03s
Iteration 5234, Loss: 0.0150, Time: 0.04s
Iteration 5235, Loss: 0.0127, Time: 0.04s
Iteration 5236, Loss: 0.0108, Time: 0.03s
Iteration 5237, Loss: 0.0135, Time: 0.04s
Iteration 5238, Loss: 0.0154, Time: 0.04s
Iteration 5239, Loss: 0.0149, Time: 0.03s
Iteration 5240, Loss: 0.0139, Time: 0.03s
Iteration 5241, Loss: 0.0155, Time: 0.04s
Iteration 5242, Loss: 0.0125, Time: 0.04s
Iteration 5243, Loss: 0.0123, Time: 0.04s
Iteration 5244, Loss: 0.0159, Time: 0.03s
Iteration 5245, Loss: 0.0171, Time: 0.03s
Iteration 5246, Loss: 0.0142, Time: 0.03s
Iteration 5247, Loss: 0.0153, Time: 0.03s
Iteration 5248, Loss: 0.0184, Time: 0.03s
Iteration 5249, Loss: 0.0174, Time: 0.03s
Iteration 5250, Loss: 0.0097, Time: 0.03s
Iteration 5251, Loss: 0.0088, Time: 0.03s
Iteration 5252, Loss: 0.0128, Time: 0.03s
Iteration 5253, Loss: 0.0227, Time: 0.04s
Iteration 5254, Loss: 0.0164, Time: 0.03s
Iteration 5255, Loss: 0.0130, Time: 0.03s
Iteration 5256, Loss: 0.0127, Time: 0.03s
Iteration 5257, Loss: 0.0138, Time: 0.03s
Iteration 5258, Loss: 0.0135, Time: 0.03s
Iteration 5259, Loss: 0.0152, Time: 0.03s
Iteration 5260, Loss: 0.0153, Time: 0.03s
Iteration 5261, Loss: 0.0173, Time: 0.03s
Iteration 5262, Loss: 0.0160, Time: -0.12s
Iteration 5263, Loss: 0.0127, Time: 0.03s
Iteration 5264, Loss: 0.0143, Time: 0.03s
Iteration 5265, Loss: 0.0246, Time: 0.03s
Iteration 5266, Loss: 0.0192, Time: 0.03s
Iteration 5267, Loss: 0.0138, Time: 0.03s
Iteration 5268, Loss: 0.0135, Time: 0.03s
Iteration 5269, Loss: 0.0115, Time: 0.03s
Iteration 5270, Loss: 0.0214, Time: 0.04s
```

```
Iteration 5271, Loss: 0.0166, Time: 0.03s
Iteration 5272, Loss: 0.0191, Time: 0.03s
Iteration 5273, Loss: 0.0114, Time: 0.03s
Iteration 5274, Loss: 0.0204, Time: 0.03s
Iteration 5275, Loss: 0.0100, Time: 0.03s
Iteration 5276, Loss: 0.0160, Time: 0.03s
Iteration 5277, Loss: 0.0136, Time: 0.03s
Iteration 5278, Loss: 0.0114, Time: 0.03s
Iteration 5279, Loss: 0.0188, Time: 0.03s
Iteration 5280, Loss: 0.0162, Time: 0.03s
Iteration 5281, Loss: 0.0164, Time: 0.03s
Iteration 5282, Loss: 0.0098, Time: 0.03s
Iteration 5283, Loss: 0.0118, Time: 0.03s
Iteration 5284, Loss: 0.0134, Time: 0.03s
Iteration 5285, Loss: 0.0110, Time: 0.03s
Iteration 5286, Loss: 0.0127, Time: 0.03s
Iteration 5287, Loss: 0.0146, Time: 0.03s
Iteration 5288, Loss: 0.0112, Time: 0.03s
Iteration 5289, Loss: 0.0286, Time: 0.03s
Iteration 5290, Loss: 0.0064, Time: 0.03s
Iteration 5291, Loss: 0.0164, Time: 0.03s
Iteration 5292, Loss: 0.0119, Time: 0.03s
Iteration 5293, Loss: 0.0122, Time: 0.03s
Iteration 5294, Loss: 0.0148, Time: 0.03s
Iteration 5295, Loss: 0.0186, Time: 0.03s
Iteration 5296, Loss: 0.0125, Time: 0.03s
Iteration 5297, Loss: 0.0155, Time: 0.03s
Iteration 5298, Loss: 0.0142, Time: 0.03s
Iteration 5299, Loss: 0.0231, Time: 0.03s
Iteration 5300, Loss: 0.0145, Time: 0.03s
Iteration 5300, Loss: 0.0145, Time: 0.03s
Test Loss: 0.0662
Iteration 5301, Loss: 0.0200, Time: 0.03s
Iteration 5302, Loss: 0.0161, Time: 0.03s
Iteration 5303, Loss: 0.0115, Time: 0.03s
Iteration 5304, Loss: 0.0129, Time: 0.03s
Iteration 5305, Loss: 0.0142, Time: 0.03s
Iteration 5306, Loss: 0.0189, Time: 0.03s
Iteration 5307, Loss: 0.0123, Time: 0.03s
Iteration 5308, Loss: 0.0144, Time: 0.03s
Iteration 5309, Loss: 0.0134, Time: 0.03s
Iteration 5310, Loss: 0.0148, Time: 0.03s
Iteration 5311, Loss: 0.0159, Time: 0.03s
Iteration 5312, Loss: 0.0126, Time: 0.03s
Iteration 5313, Loss: 0.0130, Time: 0.03s
Iteration 5314, Loss: 0.0167, Time: 0.03s
Iteration 5315, Loss: 0.0133, Time: 0.03s
Iteration 5316, Loss: 0.0118, Time: 0.03s
```

```
Iteration 5317, Loss: 0.0224, Time: 0.03s
Iteration 5318, Loss: 0.0128, Time: 0.03s
Iteration 5319, Loss: 0.0079, Time: 0.03s
Iteration 5320, Loss: 0.0173, Time: 0.03s
Iteration 5321, Loss: 0.0176, Time: 0.03s
Iteration 5322, Loss: 0.0127, Time: 0.03s
Iteration 5323, Loss: 0.0110, Time: 0.03s
Iteration 5324, Loss: 0.0121, Time: 0.03s
Iteration 5325, Loss: 0.0112, Time: 0.03s
Iteration 5326, Loss: 0.0120, Time: 0.03s
Iteration 5327, Loss: 0.0129, Time: 0.03s
Iteration 5328, Loss: 0.0107, Time: 0.03s
Iteration 5329, Loss: 0.0145, Time: 0.03s
Iteration 5330, Loss: 0.0101, Time: 0.03s
Iteration 5331, Loss: 0.0177, Time: 0.03s
Iteration 5332, Loss: 0.0132, Time: 0.03s
Iteration 5333, Loss: 0.0105, Time: 0.03s
Iteration 5334, Loss: 0.0272, Time: 0.03s
Iteration 5335, Loss: 0.0107, Time: 0.03s
Iteration 5336, Loss: 0.0099, Time: 0.03s
Iteration 5337, Loss: 0.0170, Time: 0.03s
Iteration 5338, Loss: 0.0122, Time: 0.03s
Iteration 5339, Loss: 0.0201, Time: 0.03s
Iteration 5340, Loss: 0.0153, Time: 0.03s
Iteration 5341, Loss: 0.0116, Time: 0.03s
Iteration 5342, Loss: 0.0106, Time: 0.03s
Iteration 5343, Loss: 0.0148, Time: 0.03s
Iteration 5344, Loss: 0.0133, Time: 0.03s
Iteration 5345, Loss: 0.0103, Time: 0.03s
Iteration 5346, Loss: 0.0143, Time: 0.03s
Iteration 5347, Loss: 0.0115, Time: 0.03s
Iteration 5348, Loss: 0.0113, Time: 0.04s
Iteration 5349, Loss: 0.0201, Time: 0.03s
Iteration 5350, Loss: 0.0177, Time: 0.03s
Iteration 5351, Loss: 0.0210, Time: 0.03s
Iteration 5352, Loss: 0.0057, Time: 0.03s
Iteration 5353, Loss: 0.0128, Time: 0.03s
Iteration 5354, Loss: 0.0119, Time: 0.03s
Iteration 5355, Loss: 0.0176, Time: 0.03s
Iteration 5356, Loss: 0.0136, Time: 0.03s
Iteration 5357, Loss: 0.0114, Time: 0.03s
Iteration 5358, Loss: 0.0126, Time: 0.03s
Iteration 5359, Loss: 0.0146, Time: 0.03s
Iteration 5360, Loss: 0.0200, Time: 0.03s
Iteration 5361, Loss: 0.0169, Time: 0.03s
Iteration 5362, Loss: 0.0143, Time: 0.03s
Iteration 5363, Loss: 0.0146, Time: 0.04s
Iteration 5364, Loss: 0.0111, Time: 0.03s
```

```
Iteration 5365, Loss: 0.0096, Time: 0.03s
Iteration 5366, Loss: 0.0234, Time: 0.03s
Iteration 5367, Loss: 0.0121, Time: 0.03s
Iteration 5368, Loss: 0.0118, Time: 0.03s
Iteration 5369, Loss: 0.0151, Time: 0.03s
Iteration 5370, Loss: 0.0107, Time: 0.03s
Iteration 5371, Loss: 0.0260, Time: 0.03s
Iteration 5372, Loss: 0.0131, Time: 0.03s
Iteration 5373, Loss: 0.0110, Time: 0.03s
Iteration 5374, Loss: 0.0186, Time: 0.03s
Iteration 5375, Loss: 0.0129, Time: 0.03s
Iteration 5376, Loss: 0.0193, Time: 0.03s
Iteration 5377, Loss: 0.0152, Time: 0.03s
Iteration 5378, Loss: 0.0139, Time: 0.03s
Iteration 5379, Loss: 0.0102, Time: 0.03s
Iteration 5380, Loss: 0.0099, Time: 0.03s
Iteration 5381, Loss: 0.0137, Time: 0.04s
Iteration 5382, Loss: 0.0128, Time: 0.03s
Iteration 5383, Loss: 0.0129, Time: 0.04s
Iteration 5384, Loss: 0.0168, Time: 0.03s
Iteration 5385, Loss: 0.0143, Time: 0.03s
Iteration 5386, Loss: 0.0125, Time: 0.03s
Iteration 5387, Loss: 0.0111, Time: 0.03s
Iteration 5388, Loss: 0.0110, Time: 0.03s
Iteration 5389, Loss: 0.0083, Time: 0.03s
Iteration 5390, Loss: 0.0080, Time: 0.03s
Iteration 5391, Loss: 0.0103, Time: 0.03s
Iteration 5392, Loss: 0.0116, Time: 0.03s
Iteration 5393, Loss: 0.0150, Time: 0.03s
Iteration 5394, Loss: 0.0221, Time: 0.03s
Iteration 5395, Loss: 0.0219, Time: 0.03s
Iteration 5396, Loss: 0.0138, Time: 0.03s
Iteration 5397, Loss: 0.0098, Time: 0.03s
Iteration 5398, Loss: 0.0136, Time: 0.03s
Iteration 5399, Loss: 0.0157, Time: 0.03s
Iteration 5400, Loss: 0.0099, Time: 0.03s
Iteration 5400, Loss: 0.0099, Time: 0.03s
Test Loss: 0.0452
Iteration 5401, Loss: 0.0154, Time: 0.03s
Iteration 5402, Loss: 0.0088, Time: 0.03s
Iteration 5403, Loss: 0.0145, Time: 0.03s
Iteration 5404, Loss: 0.0128, Time: 0.03s
Iteration 5405, Loss: 0.0047, Time: 0.03s
Iteration 5406, Loss: 0.0146, Time: 0.03s
Iteration 5407, Loss: 0.0171, Time: 0.03s
Iteration 5408, Loss: 0.0165, Time: 0.03s
Iteration 5409, Loss: 0.0128, Time: 0.03s
Iteration 5410, Loss: 0.0104, Time: 0.03s
```

```
Iteration 5411, Loss: 0.0164, Time: 0.03s
Iteration 5412, Loss: 0.0158, Time: 0.03s
Iteration 5413, Loss: 0.0112, Time: 0.03s
Iteration 5414, Loss: 0.0124, Time: 0.03s
Iteration 5415, Loss: 0.0122, Time: 0.03s
Iteration 5416, Loss: 0.0127, Time: 0.03s
Iteration 5417, Loss: 0.0113, Time: 0.03s
Iteration 5418, Loss: 0.0159, Time: 0.03s
Iteration 5419, Loss: 0.0093, Time: 0.03s
Iteration 5420, Loss: 0.0099, Time: 0.03s
Iteration 5421, Loss: 0.0170, Time: 0.03s
Iteration 5422, Loss: 0.0146, Time: 0.03s
Iteration 5423, Loss: 0.0153, Time: 0.03s
Iteration 5424, Loss: 0.0146, Time: 0.03s
Iteration 5425, Loss: 0.0126, Time: 0.03s
Iteration 5426, Loss: 0.0149, Time: 0.03s
Iteration 5427, Loss: 0.0112, Time: 0.03s
Iteration 5428, Loss: 0.0070, Time: 0.03s
Iteration 5429, Loss: 0.0184, Time: 0.03s
Iteration 5430, Loss: 0.0097, Time: 0.03s
Iteration 5431, Loss: 0.0163, Time: 0.03s
Iteration 5432, Loss: 0.0112, Time: 0.03s
Iteration 5433, Loss: 0.0154, Time: 0.03s
Iteration 5434, Loss: 0.0092, Time: 0.03s
Iteration 5435, Loss: 0.0134, Time: 0.03s
Iteration 5436, Loss: 0.0177, Time: 0.03s
Iteration 5437, Loss: 0.0086, Time: 0.03s
Iteration 5438, Loss: 0.0155, Time: 0.03s
Iteration 5439, Loss: 0.0159, Time: 0.03s
Iteration 5440, Loss: 0.0133, Time: 0.03s
Iteration 5441, Loss: 0.0133, Time: 0.03s
Iteration 5442, Loss: 0.0131, Time: 0.03s
Iteration 5443, Loss: 0.0174, Time: 0.03s
Iteration 5444, Loss: 0.0150, Time: 0.03s
Iteration 5445, Loss: 0.0121, Time: 0.04s
Iteration 5446, Loss: 0.0116, Time: 0.03s
Iteration 5447, Loss: 0.0083, Time: 0.03s
Iteration 5448, Loss: 0.0143, Time: 0.03s
Iteration 5449, Loss: 0.0186, Time: 0.03s
Iteration 5450, Loss: 0.0118, Time: 0.03s
Iteration 5451, Loss: 0.0157, Time: 0.03s
Iteration 5452, Loss: 0.0107, Time: 0.03s
Iteration 5453, Loss: 0.0108, Time: 0.03s
Iteration 5454, Loss: 0.0096, Time: 0.03s
Iteration 5455, Loss: 0.0165, Time: 0.03s
Iteration 5456, Loss: 0.0130, Time: 0.03s
Iteration 5457, Loss: 0.0112, Time: 0.03s
Iteration 5458, Loss: 0.0129, Time: 0.03s
```

```
Iteration 5459, Loss: 0.0137, Time: 0.03s
Iteration 5460, Loss: 0.0155, Time: 0.03s
Iteration 5461, Loss: 0.0192, Time: 0.03s
Iteration 5462, Loss: 0.0143, Time: 0.03s
Iteration 5463, Loss: 0.0125, Time: 0.03s
Iteration 5464, Loss: 0.0106, Time: 0.03s
Iteration 5465, Loss: 0.0109, Time: 0.03s
Iteration 5466, Loss: 0.0155, Time: 0.03s
Iteration 5467, Loss: 0.0106, Time: 0.03s
Iteration 5468, Loss: 0.0119, Time: 0.03s
Iteration 5469, Loss: 0.0096, Time: 0.03s
Iteration 5470, Loss: 0.0139, Time: 0.03s
Iteration 5471, Loss: 0.0154, Time: 0.03s
Iteration 5472, Loss: 0.0140, Time: 0.03s
Iteration 5473, Loss: 0.0104, Time: 0.03s
Iteration 5474, Loss: 0.0068, Time: 0.03s
Iteration 5475, Loss: 0.0157, Time: 0.03s
Iteration 5476, Loss: 0.0125, Time: 0.04s
Iteration 5477, Loss: 0.0199, Time: 0.03s
Iteration 5478, Loss: 0.0147, Time: 0.03s
Iteration 5479, Loss: 0.0135, Time: 0.03s
Iteration 5480, Loss: 0.0101, Time: 0.03s
Iteration 5481, Loss: 0.0117, Time: 0.03s
Iteration 5482, Loss: 0.0158, Time: 0.03s
Iteration 5483, Loss: 0.0162, Time: 0.03s
Iteration 5484, Loss: 0.0106, Time: 0.03s
Iteration 5485, Loss: 0.0163, Time: 0.03s
Iteration 5486, Loss: 0.0148, Time: 0.03s
Iteration 5487, Loss: 0.0176, Time: 0.03s
Iteration 5488, Loss: 0.0221, Time: 0.03s
Iteration 5489, Loss: 0.0078, Time: 0.03s
Iteration 5490, Loss: 0.0064, Time: 0.03s
Iteration 5491, Loss: 0.0091, Time: 0.03s
Iteration 5492, Loss: 0.0143, Time: 0.03s
Iteration 5493, Loss: 0.0090, Time: 0.03s
Iteration 5494, Loss: 0.0112, Time: 0.03s
Iteration 5495, Loss: 0.0108, Time: 0.03s
Iteration 5496, Loss: 0.0127, Time: 0.03s
Iteration 5497, Loss: 0.0119, Time: 0.03s
Iteration 5498, Loss: 0.0148, Time: 0.03s
Iteration 5499, Loss: 0.0098, Time: 0.03s
Iteration 5500, Loss: 0.0176, Time: 0.03s
Iteration 5500, Loss: 0.0176, Time: 0.03s
Test Loss: 0.0314
Iteration 5501, Loss: 0.0146, Time: 0.03s
Iteration 5502, Loss: 0.0173, Time: 0.03s
Iteration 5503, Loss: 0.0147, Time: 0.03s
Iteration 5504, Loss: 0.0112, Time: 0.03s
```

```
Iteration 5505, Loss: 0.0133, Time: 0.03s
Iteration 5506, Loss: 0.0217, Time: 0.03s
Iteration 5507, Loss: 0.0108, Time: 0.03s
Iteration 5508, Loss: 0.0111, Time: 0.03s
Iteration 5509, Loss: 0.0159, Time: 0.03s
Iteration 5510, Loss: 0.0129, Time: 0.03s
Iteration 5511, Loss: 0.0148, Time: 0.03s
Iteration 5512, Loss: 0.0089, Time: 0.03s
Iteration 5513, Loss: 0.0161, Time: 0.03s
Iteration 5514, Loss: 0.0114, Time: 0.03s
Iteration 5515, Loss: 0.0191, Time: 0.03s
Iteration 5516, Loss: 0.0161, Time: 0.03s
Iteration 5517, Loss: 0.0161, Time: 0.03s
Iteration 5518, Loss: 0.0141, Time: 0.03s
Iteration 5519, Loss: 0.0112, Time: 0.03s
Iteration 5520, Loss: 0.0152, Time: 0.03s
Iteration 5521, Loss: 0.0172, Time: 0.03s
Iteration 5522, Loss: 0.0180, Time: 0.03s
Iteration 5523, Loss: 0.0186, Time: 0.03s
Iteration 5524, Loss: 0.0147, Time: 0.03s
Iteration 5525, Loss: 0.0158, Time: 0.03s
Iteration 5526, Loss: 0.0221, Time: 0.03s
Iteration 5527, Loss: 0.0075, Time: 0.03s
Iteration 5528, Loss: 0.0140, Time: 0.03s
Iteration 5529, Loss: 0.0150, Time: 0.03s
Iteration 5530, Loss: 0.0102, Time: 0.03s
Iteration 5531, Loss: 0.0183, Time: 0.03s
Iteration 5532, Loss: 0.0160, Time: 0.03s
Iteration 5533, Loss: 0.0170, Time: 0.03s
Iteration 5534, Loss: 0.0062, Time: 0.03s
Iteration 5535, Loss: 0.0167, Time: 0.03s
Iteration 5536, Loss: 0.0117, Time: 0.03s
Iteration 5537, Loss: 0.0107, Time: 0.03s
Iteration 5538, Loss: 0.0155, Time: 0.03s
Iteration 5539, Loss: 0.0175, Time: 0.03s
Iteration 5540, Loss: 0.0124, Time: 0.03s
Iteration 5541, Loss: 0.0150, Time: 0.03s
Iteration 5542, Loss: 0.0131, Time: 0.03s
Iteration 5543, Loss: 0.0091, Time: 0.03s
Iteration 5544, Loss: 0.0133, Time: 0.03s
Iteration 5545, Loss: 0.0103, Time: 0.03s
Iteration 5546, Loss: 0.0192, Time: 0.03s
Iteration 5547, Loss: 0.0095, Time: 0.03s
Iteration 5548, Loss: 0.0155, Time: 0.03s
Iteration 5549, Loss: 0.0140, Time: 0.03s
Iteration 5550, Loss: 0.0085, Time: 0.03s
Iteration 5551, Loss: 0.0148, Time: 0.03s
Iteration 5552, Loss: 0.0113, Time: 0.03s
```

```
Iteration 5553, Loss: 0.0124, Time: 0.03s
Iteration 5554, Loss: 0.0089, Time: 0.03s
Iteration 5555, Loss: 0.0160, Time: 0.03s
Iteration 5556, Loss: 0.0172, Time: 0.03s
Iteration 5557, Loss: 0.0211, Time: 0.03s
Iteration 5558, Loss: 0.0156, Time: 0.03s
Iteration 5559, Loss: 0.0107, Time: 0.03s
Iteration 5560, Loss: 0.0114, Time: 0.03s
Iteration 5561, Loss: 0.0198, Time: 0.03s
Iteration 5562, Loss: 0.0157, Time: 0.03s
Iteration 5563, Loss: 0.0107, Time: 0.03s
Iteration 5564, Loss: 0.0128, Time: 0.03s
Iteration 5565, Loss: 0.0123, Time: 0.03s
Iteration 5566, Loss: 0.0123, Time: 0.03s
Iteration 5567, Loss: 0.0140, Time: 0.03s
Iteration 5568, Loss: 0.0171, Time: 0.03s
Iteration 5569, Loss: 0.0137, Time: 0.03s
Iteration 5570, Loss: 0.0158, Time: 0.03s
Iteration 5571, Loss: 0.0201, Time: 0.03s
Iteration 5572, Loss: 0.0157, Time: 0.03s
Iteration 5573, Loss: 0.0167, Time: 0.03s
Iteration 5574, Loss: 0.0072, Time: 0.03s
Iteration 5575, Loss: 0.0144, Time: 0.03s
Iteration 5576, Loss: 0.0100, Time: 0.03s
Iteration 5577, Loss: 0.0146, Time: 0.03s
Iteration 5578, Loss: 0.0143, Time: 0.03s
Iteration 5579, Loss: 0.0147, Time: 0.03s
Iteration 5580, Loss: 0.0146, Time: 0.03s
Iteration 5581, Loss: 0.0163, Time: 0.03s
Iteration 5582, Loss: 0.0145, Time: 0.03s
Iteration 5583, Loss: 0.0160, Time: 0.03s
Iteration 5584, Loss: 0.0153, Time: 0.03s
Iteration 5585, Loss: 0.0108, Time: 0.03s
Iteration 5586, Loss: 0.0194, Time: 0.03s
Iteration 5587, Loss: 0.0151, Time: 0.03s
Iteration 5588, Loss: 0.0110, Time: 0.03s
Iteration 5589, Loss: 0.0122, Time: 0.04s
Iteration 5590, Loss: 0.0191, Time: 0.04s
Iteration 5591, Loss: 0.0117, Time: 0.04s
Iteration 5592, Loss: 0.0108, Time: 0.04s
Iteration 5593, Loss: 0.0073, Time: 0.04s
Iteration 5594, Loss: 0.0161, Time: 0.04s
Iteration 5595, Loss: 0.0186, Time: 0.04s
Iteration 5596, Loss: 0.0087, Time: 0.04s
Iteration 5597, Loss: 0.0148, Time: 0.04s
Iteration 5598, Loss: 0.0080, Time: 0.03s
Iteration 5599, Loss: 0.0125, Time: 0.04s
Iteration 5600, Loss: 0.0112, Time: 0.03s
```

```
Iteration 5600, Loss: 0.0112, Time: 0.04s
Test Loss: 0.0289
Iteration 5601, Loss: 0.0128, Time: 0.04s
Iteration 5602, Loss: 0.0154, Time: 0.04s
Iteration 5603, Loss: 0.0137, Time: 0.04s
Iteration 5604, Loss: 0.0202, Time: 0.03s
Iteration 5605, Loss: 0.0155, Time: 0.03s
Iteration 5606, Loss: 0.0152, Time: 0.03s
Iteration 5607, Loss: 0.0174, Time: 0.03s
Iteration 5608, Loss: 0.0180, Time: 0.03s
Iteration 5609, Loss: 0.0129, Time: 0.03s
Iteration 5610, Loss: 0.0147, Time: 0.03s
Iteration 5611, Loss: 0.0136, Time: 0.03s
Iteration 5612, Loss: 0.0139, Time: 0.03s
Iteration 5613, Loss: 0.0140, Time: 0.03s
Iteration 5614, Loss: 0.0119, Time: 0.03s
Iteration 5615, Loss: 0.0097, Time: 0.03s
Iteration 5616, Loss: 0.0146, Time: 0.03s
Iteration 5617, Loss: 0.0139, Time: 0.03s
Iteration 5618, Loss: 0.0142, Time: 0.03s
Iteration 5619, Loss: 0.0218, Time: 0.03s
Iteration 5620, Loss: 0.0141, Time: 0.03s
Iteration 5621, Loss: 0.0150, Time: 0.03s
Iteration 5622, Loss: 0.0152, Time: 0.03s
Iteration 5623, Loss: 0.0141, Time: 0.03s
Iteration 5624, Loss: 0.0156, Time: 0.03s
Iteration 5625, Loss: 0.0137, Time: 0.03s
Iteration 5626, Loss: 0.0134, Time: 0.03s
Iteration 5627, Loss: 0.0145, Time: 0.03s
Iteration 5628, Loss: 0.0146, Time: 0.03s
Iteration 5629, Loss: 0.0150, Time: 0.03s
Iteration 5630, Loss: 0.0148, Time: 0.03s
Iteration 5631, Loss: 0.0122, Time: 0.03s
Iteration 5632, Loss: 0.0108, Time: 0.03s
Iteration 5633, Loss: 0.0132, Time: 0.03s
Iteration 5634, Loss: 0.0104, Time: 0.03s
Iteration 5635, Loss: 0.0097, Time: 0.03s
Iteration 5636, Loss: 0.0175, Time: 0.03s
Iteration 5637, Loss: 0.0153, Time: 0.03s
Iteration 5638, Loss: 0.0121, Time: 0.03s
Iteration 5639, Loss: 0.0077, Time: 0.03s
Iteration 5640, Loss: 0.0103, Time: 0.03s
Iteration 5641, Loss: 0.0111, Time: 0.03s
Iteration 5642, Loss: 0.0095, Time: 0.03s
Iteration 5643, Loss: 0.0135, Time: 0.03s
Iteration 5644, Loss: 0.0137, Time: 0.03s
Iteration 5645, Loss: 0.0100, Time: 0.03s
Iteration 5646, Loss: 0.0172, Time: 0.03s
```

```
Iteration 5647, Loss: 0.0107, Time: 0.03s
Iteration 5648, Loss: 0.0135, Time: 0.03s
Iteration 5649, Loss: 0.0110, Time: 0.03s
Iteration 5650, Loss: 0.0137, Time: 0.03s
Iteration 5651, Loss: 0.0135, Time: 0.03s
Iteration 5652, Loss: 0.0179, Time: 0.03s
Iteration 5653, Loss: 0.0157, Time: 0.03s
Iteration 5654, Loss: 0.0125, Time: 0.03s
Iteration 5655, Loss: 0.0157, Time: 0.03s
Iteration 5656, Loss: 0.0098, Time: 0.03s
Iteration 5657, Loss: 0.0124, Time: 0.03s
Iteration 5658, Loss: 0.0155, Time: 0.03s
Iteration 5659, Loss: 0.0184, Time: 0.03s
Iteration 5660, Loss: 0.0106, Time: 0.03s
Iteration 5661, Loss: 0.0155, Time: 0.03s
Iteration 5662, Loss: 0.0101, Time: 0.03s
Iteration 5663, Loss: 0.0106, Time: 0.03s
Iteration 5664, Loss: 0.0156, Time: 0.03s
Iteration 5665, Loss: 0.0151, Time: 0.03s
Iteration 5666, Loss: 0.0226, Time: 0.03s
Iteration 5667, Loss: 0.0125, Time: 0.03s
Iteration 5668, Loss: 0.0124, Time: 0.03s
Iteration 5669, Loss: 0.0142, Time: 0.03s
Iteration 5670, Loss: 0.0070, Time: 0.03s
Iteration 5671, Loss: 0.0097, Time: 0.03s
Iteration 5672, Loss: 0.0097, Time: 0.03s
Iteration 5673, Loss: 0.0106, Time: 0.03s
Iteration 5674, Loss: 0.0106, Time: 0.03s
Iteration 5675, Loss: 0.0156, Time: 0.03s
Iteration 5676, Loss: 0.0123, Time: 0.03s
Iteration 5677, Loss: 0.0127, Time: 0.04s
Iteration 5678, Loss: 0.0126, Time: 0.03s
Iteration 5679, Loss: 0.0205, Time: 0.03s
Iteration 5680, Loss: 0.0190, Time: 0.03s
Iteration 5681, Loss: 0.0145, Time: 0.03s
Iteration 5682, Loss: 0.0097, Time: 0.03s
Iteration 5683, Loss: 0.0134, Time: 0.03s
Iteration 5684, Loss: 0.0128, Time: 0.03s
Iteration 5685, Loss: 0.0088, Time: 0.03s
Iteration 5686, Loss: 0.0124, Time: 0.03s
Iteration 5687, Loss: 0.0149, Time: 0.03s
Iteration 5688, Loss: 0.0197, Time: 0.03s
Iteration 5689, Loss: 0.0091, Time: 0.03s
Iteration 5690, Loss: 0.0173, Time: 0.03s
Iteration 5691, Loss: 0.0142, Time: 0.03s
Iteration 5692, Loss: 0.0187, Time: 0.03s
Iteration 5693, Loss: 0.0116, Time: 0.03s
Iteration 5694, Loss: 0.0163, Time: 0.03s
```

```
Iteration 5695, Loss: 0.0119, Time: 0.03s
Iteration 5696, Loss: 0.0108, Time: 0.03s
Iteration 5697, Loss: 0.0127, Time: 0.03s
Iteration 5698, Loss: 0.0137, Time: 0.03s
Iteration 5699, Loss: 0.0093, Time: 0.03s
Iteration 5700, Loss: 0.0189, Time: 0.03s
Iteration 5700, Loss: 0.0189, Time: 0.03s
Test Loss: 0.0446
Iteration 5701, Loss: 0.0141, Time: 0.03s
Iteration 5702, Loss: 0.0116, Time: 0.03s
Iteration 5703, Loss: 0.0140, Time: 0.03s
Iteration 5704, Loss: 0.0131, Time: 0.03s
Iteration 5705, Loss: 0.0181, Time: 0.03s
Iteration 5706, Loss: 0.0147, Time: 0.03s
Iteration 5707, Loss: 0.0132, Time: 0.03s
Iteration 5708, Loss: 0.0108, Time: 0.03s
Iteration 5709, Loss: 0.0123, Time: 0.03s
Iteration 5710, Loss: 0.0164, Time: 0.03s
Iteration 5711, Loss: 0.0150, Time: 0.03s
Iteration 5712, Loss: 0.0104, Time: 0.03s
Iteration 5713, Loss: 0.0129, Time: 0.03s
Iteration 5714, Loss: 0.0115, Time: 0.03s
Iteration 5715, Loss: 0.0143, Time: 0.03s
Iteration 5716, Loss: 0.0084, Time: 0.03s
Iteration 5717, Loss: 0.0157, Time: 0.03s
Iteration 5718, Loss: 0.0157, Time: 0.03s
Iteration 5719, Loss: 0.0150, Time: 0.03s
Iteration 5720, Loss: 0.0089, Time: 0.03s
Iteration 5721, Loss: 0.0133, Time: 0.03s
Iteration 5722, Loss: 0.0105, Time: 0.03s
Iteration 5723, Loss: 0.0084, Time: 0.03s
Iteration 5724, Loss: 0.0176, Time: 0.03s
Iteration 5725, Loss: 0.0131, Time: 0.03s
Iteration 5726, Loss: 0.0125, Time: 0.03s
Iteration 5727, Loss: 0.0137, Time: 0.04s
Iteration 5728, Loss: 0.0137, Time: 0.03s
Iteration 5729, Loss: 0.0186, Time: 0.03s
Iteration 5730, Loss: 0.0136, Time: 0.03s
Iteration 5731, Loss: 0.0158, Time: 0.03s
Iteration 5732, Loss: 0.0112, Time: 0.03s
Iteration 5733, Loss: 0.0123, Time: 0.03s
Iteration 5734, Loss: 0.0119, Time: 0.03s
Iteration 5735, Loss: 0.0096, Time: 0.03s
Iteration 5736, Loss: 0.0110, Time: 0.03s
Iteration 5737, Loss: 0.0155, Time: 0.03s
Iteration 5738, Loss: 0.0141, Time: 0.03s
Iteration 5739, Loss: 0.0111, Time: 0.03s
Iteration 5740, Loss: 0.0103, Time: 0.03s
```

```
Iteration 5741, Loss: 0.0119, Time: 0.03s
Iteration 5742, Loss: 0.0144, Time: 0.03s
Iteration 5743, Loss: 0.0141, Time: 0.03s
Iteration 5744, Loss: 0.0158, Time: 0.03s
Iteration 5745, Loss: 0.0159, Time: 0.03s
Iteration 5746, Loss: 0.0125, Time: 0.03s
Iteration 5747, Loss: 0.0116, Time: 0.03s
Iteration 5748, Loss: 0.0175, Time: 0.03s
Iteration 5749, Loss: 0.0102, Time: 0.03s
Iteration 5750, Loss: 0.0176, Time: 0.03s
Iteration 5751, Loss: 0.0119, Time: 0.03s
Iteration 5752, Loss: 0.0130, Time: 0.03s
Iteration 5753, Loss: 0.0128, Time: 0.03s
Iteration 5754, Loss: 0.0100, Time: 0.03s
Iteration 5755, Loss: 0.0150, Time: 0.03s
Iteration 5756, Loss: 0.0054, Time: 0.03s
Iteration 5757, Loss: 0.0167, Time: 0.03s
Iteration 5758, Loss: 0.0091, Time: 0.03s
Iteration 5759, Loss: 0.0122, Time: 0.03s
Iteration 5760, Loss: 0.0106, Time: 0.03s
Iteration 5761, Loss: 0.0143, Time: 0.03s
Iteration 5762, Loss: 0.0074, Time: 0.03s
Iteration 5763, Loss: 0.0141, Time: 0.03s
Iteration 5764, Loss: 0.0183, Time: 0.03s
Iteration 5765, Loss: 0.0131, Time: 0.03s
Iteration 5766, Loss: 0.0096, Time: 0.03s
Iteration 5767, Loss: 0.0123, Time: 0.03s
Iteration 5768, Loss: 0.0123, Time: 0.03s
Iteration 5769, Loss: 0.0058, Time: 0.04s
Iteration 5770, Loss: 0.0164, Time: 0.03s
Iteration 5771, Loss: 0.0134, Time: 0.03s
Iteration 5772, Loss: 0.0118, Time: 0.03s
Iteration 5773, Loss: 0.0110, Time: 0.03s
Iteration 5774, Loss: 0.0130, Time: 0.03s
Iteration 5775, Loss: 0.0178, Time: 0.03s
Iteration 5776, Loss: 0.0153, Time: 0.03s
Iteration 5777, Loss: 0.0208, Time: 0.03s
Iteration 5778, Loss: 0.0113, Time: 0.03s
Iteration 5779, Loss: 0.0153, Time: 0.03s
Iteration 5780, Loss: 0.0198, Time: 0.03s
Iteration 5781, Loss: 0.0176, Time: 0.03s
Iteration 5782, Loss: 0.0141, Time: 0.03s
Iteration 5783, Loss: 0.0162, Time: 0.03s
Iteration 5784, Loss: 0.0074, Time: 0.03s
Iteration 5785, Loss: 0.0141, Time: 0.03s
Iteration 5786, Loss: 0.0187, Time: 0.03s
Iteration 5787, Loss: 0.0156, Time: 0.03s
Iteration 5788, Loss: 0.0130, Time: 0.03s
```

```
Iteration 5789, Loss: 0.0201, Time: 0.03s
Iteration 5790, Loss: 0.0129, Time: 0.03s
Iteration 5791, Loss: 0.0151, Time: 0.03s
Iteration 5792, Loss: 0.0167, Time: 0.03s
Iteration 5793, Loss: 0.0173, Time: 0.03s
Iteration 5794, Loss: 0.0116, Time: 0.03s
Iteration 5795, Loss: 0.0119, Time: 0.03s
Iteration 5796, Loss: 0.0154, Time: 0.03s
Iteration 5797, Loss: 0.0126, Time: 0.03s
Iteration 5798, Loss: 0.0145, Time: 0.03s
Iteration 5799, Loss: 0.0179, Time: 0.03s
Iteration 5800, Loss: 0.0113, Time: 0.03s
Iteration 5800, Loss: 0.0113, Time: 0.03s
Test Loss: 0.0555
Iteration 5801, Loss: 0.0147, Time: 0.03s
Iteration 5802, Loss: 0.0156, Time: 0.03s
Iteration 5803, Loss: 0.0117, Time: 0.03s
Iteration 5804, Loss: 0.0165, Time: 0.03s
Iteration 5805, Loss: 0.0150, Time: 0.03s
Iteration 5806, Loss: 0.0121, Time: 0.03s
Iteration 5807, Loss: 0.0127, Time: 0.03s
Iteration 5808, Loss: 0.0145, Time: 0.03s
Iteration 5809, Loss: 0.0142, Time: 0.03s
Iteration 5810, Loss: 0.0145, Time: 0.03s
Iteration 5811, Loss: 0.0158, Time: 0.03s
Iteration 5812, Loss: 0.0120, Time: 0.03s
Iteration 5813, Loss: 0.0160, Time: 0.03s
Iteration 5814, Loss: 0.0095, Time: 0.03s
Iteration 5815, Loss: 0.0125, Time: 0.03s
Iteration 5816, Loss: 0.0135, Time: 0.03s
Iteration 5817, Loss: 0.0149, Time: 0.03s
Iteration 5818, Loss: 0.0158, Time: 0.03s
Iteration 5819, Loss: 0.0101, Time: 0.03s
Iteration 5820, Loss: 0.0124, Time: 0.03s
Iteration 5821, Loss: 0.0085, Time: 0.03s
Iteration 5822, Loss: 0.0131, Time: 0.03s
Iteration 5823, Loss: 0.0137, Time: 0.03s
Iteration 5824, Loss: 0.0162, Time: 0.03s
Iteration 5825, Loss: 0.0112, Time: 0.03s
Iteration 5826, Loss: 0.0107, Time: 0.03s
Iteration 5827, Loss: 0.0118, Time: 0.03s
Iteration 5828, Loss: 0.0098, Time: 0.03s
Iteration 5829, Loss: 0.0115, Time: 0.03s
Iteration 5830, Loss: 0.0116, Time: 0.03s
Iteration 5831, Loss: 0.0159, Time: 0.03s
Iteration 5832, Loss: 0.0132, Time: 0.03s
Iteration 5833, Loss: 0.0203, Time: 0.03s
Iteration 5834, Loss: 0.0157, Time: 0.03s
```

```
Iteration 5835, Loss: 0.0092, Time: 0.03s
Iteration 5836, Loss: 0.0185, Time: 0.03s
Iteration 5837, Loss: 0.0144, Time: 0.03s
Iteration 5838, Loss: 0.0155, Time: 0.03s
Iteration 5839, Loss: 0.0081, Time: 0.03s
Iteration 5840, Loss: 0.0146, Time: 0.03s
Iteration 5841, Loss: 0.0130, Time: 0.03s
Iteration 5842, Loss: 0.0128, Time: 0.03s
Iteration 5843, Loss: 0.0184, Time: 0.03s
Iteration 5844, Loss: 0.0197, Time: 0.03s
Iteration 5845, Loss: 0.0126, Time: 0.03s
Iteration 5846, Loss: 0.0116, Time: 0.03s
Iteration 5847, Loss: 0.0109, Time: 0.03s
Iteration 5848, Loss: 0.0174, Time: 0.03s
Iteration 5849, Loss: 0.0107, Time: 0.03s
Iteration 5850, Loss: 0.0157, Time: 0.03s
Iteration 5851, Loss: 0.0119, Time: 0.03s
Iteration 5852, Loss: 0.0156, Time: 0.03s
Iteration 5853, Loss: 0.0112, Time: 0.03s
Iteration 5854, Loss: 0.0142, Time: 0.03s
Iteration 5855, Loss: 0.0131, Time: 0.03s
Iteration 5856, Loss: 0.0146, Time: 0.03s
Iteration 5857, Loss: 0.0122, Time: 0.03s
Iteration 5858, Loss: 0.0111, Time: 0.03s
Iteration 5859, Loss: 0.0121, Time: 0.03s
Iteration 5860, Loss: 0.0148, Time: 0.03s
Iteration 5861, Loss: 0.0123, Time: 0.03s
Iteration 5862, Loss: 0.0087, Time: 0.03s
Iteration 5863, Loss: 0.0132, Time: 0.03s
Iteration 5864, Loss: 0.0106, Time: 0.03s
Iteration 5865, Loss: 0.0105, Time: 0.03s
Iteration 5866, Loss: 0.0087, Time: 0.03s
Iteration 5867, Loss: 0.0137, Time: 0.03s
Iteration 5868, Loss: 0.0121, Time: 0.03s
Iteration 5869, Loss: 0.0124, Time: 0.03s
Iteration 5870, Loss: 0.0123, Time: 0.03s
Iteration 5871, Loss: 0.0125, Time: 0.03s
Iteration 5872, Loss: 0.0143, Time: 0.03s
Iteration 5873, Loss: 0.0109, Time: 0.03s
Iteration 5874, Loss: 0.0099, Time: 0.03s
Iteration 5875, Loss: 0.0097, Time: 0.03s
Iteration 5876, Loss: 0.0109, Time: 0.03s
Iteration 5877, Loss: 0.0123, Time: 0.03s
Iteration 5878, Loss: 0.0112, Time: 0.03s
Iteration 5879, Loss: 0.0112, Time: 0.03s
Iteration 5880, Loss: 0.0098, Time: 0.03s
Iteration 5881, Loss: 0.0082, Time: 0.03s
Iteration 5882, Loss: 0.0126, Time: 0.03s
```

```
Iteration 5883, Loss: 0.0074, Time: 0.03s
Iteration 5884, Loss: 0.0117, Time: 0.03s
Iteration 5885, Loss: 0.0114, Time: 0.03s
Iteration 5886, Loss: 0.0086, Time: 0.03s
Iteration 5887, Loss: 0.0167, Time: 0.03s
Iteration 5888, Loss: 0.0130, Time: 0.03s
Iteration 5889, Loss: 0.0160, Time: 0.03s
Iteration 5890, Loss: 0.0170, Time: 0.03s
Iteration 5891, Loss: 0.0116, Time: 0.03s
Iteration 5892, Loss: 0.0113, Time: 0.03s
Iteration 5893, Loss: 0.0115, Time: 0.03s
Iteration 5894, Loss: 0.0097, Time: 0.03s
Iteration 5895, Loss: 0.0149, Time: 0.03s
Iteration 5896, Loss: 0.0129, Time: 0.03s
Iteration 5897, Loss: 0.0128, Time: 0.03s
Iteration 5898, Loss: 0.0077, Time: 0.03s
Iteration 5899, Loss: 0.0111, Time: 0.03s
Iteration 5900, Loss: 0.0174, Time: 0.03s
Iteration 5900, Loss: 0.0174, Time: 0.03s
Test Loss: 0.0373
Iteration 5901, Loss: 0.0129, Time: 0.03s
Iteration 5902, Loss: 0.0142, Time: 0.03s
Iteration 5903, Loss: 0.0102, Time: 0.03s
Iteration 5904, Loss: 0.0096, Time: 0.03s
Iteration 5905, Loss: 0.0086, Time: 0.03s
Iteration 5906, Loss: 0.0095, Time: 0.03s
Iteration 5907, Loss: 0.0084, Time: 0.03s
Iteration 5908, Loss: 0.0077, Time: 0.03s
Iteration 5909, Loss: 0.0197, Time: 0.03s
Iteration 5910, Loss: 0.0118, Time: 0.03s
Iteration 5911, Loss: 0.0115, Time: 0.03s
Iteration 5912, Loss: 0.0128, Time: 0.03s
Iteration 5913, Loss: 0.0130, Time: 0.03s
Iteration 5914, Loss: 0.0135, Time: 0.03s
Iteration 5915, Loss: 0.0141, Time: 0.03s
Iteration 5916, Loss: 0.0141, Time: 0.03s
Iteration 5917, Loss: 0.0117, Time: 0.03s
Iteration 5918, Loss: 0.0068, Time: 0.03s
Iteration 5919, Loss: 0.0110, Time: 0.03s
Iteration 5920, Loss: 0.0097, Time: 0.03s
Iteration 5921, Loss: 0.0180, Time: 0.03s
Iteration 5922, Loss: 0.0105, Time: 0.03s
Iteration 5923, Loss: 0.0103, Time: 0.03s
Iteration 5924, Loss: 0.0142, Time: 0.03s
Iteration 5925, Loss: 0.0130, Time: 0.03s
Iteration 5926, Loss: 0.0082, Time: 0.03s
Iteration 5927, Loss: 0.0119, Time: 0.03s
Iteration 5928, Loss: 0.0149, Time: 0.03s
```

```
Iteration 5929, Loss: 0.0148, Time: 0.03s
Iteration 5930, Loss: 0.0130, Time: 0.03s
Iteration 5931, Loss: 0.0145, Time: 0.03s
Iteration 5932, Loss: 0.0138, Time: 0.03s
Iteration 5933, Loss: 0.0099, Time: 0.03s
Iteration 5934, Loss: 0.0132, Time: 0.03s
Iteration 5935, Loss: 0.0107, Time: 0.03s
Iteration 5936, Loss: 0.0134, Time: 0.03s
Iteration 5937, Loss: 0.0110, Time: 0.03s
Iteration 5938, Loss: 0.0154, Time: 0.03s
Iteration 5939, Loss: 0.0094, Time: 0.03s
Iteration 5940, Loss: 0.0135, Time: 0.03s
Iteration 5941, Loss: 0.0141, Time: 0.03s
Iteration 5942, Loss: 0.0122, Time: 0.03s
Iteration 5943, Loss: 0.0148, Time: 0.03s
Iteration 5944, Loss: 0.0081, Time: 0.03s
Iteration 5945, Loss: 0.0143, Time: 0.03s
Iteration 5946, Loss: 0.0113, Time: 0.03s
Iteration 5947, Loss: 0.0169, Time: 0.03s
Iteration 5948, Loss: 0.0105, Time: 0.03s
Iteration 5949, Loss: 0.0135, Time: 0.03s
Iteration 5950, Loss: 0.0125, Time: 0.03s
Iteration 5951, Loss: 0.0101, Time: 0.03s
Iteration 5952, Loss: 0.0148, Time: 0.03s
Iteration 5953, Loss: 0.0159, Time: 0.03s
Iteration 5954, Loss: 0.0130, Time: 0.04s
Iteration 5955, Loss: 0.0093, Time: 0.03s
Iteration 5956, Loss: 0.0096, Time: 0.04s
Iteration 5957, Loss: 0.0095, Time: 0.04s
Iteration 5958, Loss: 0.0106, Time: 0.03s
Iteration 5959, Loss: 0.0093, Time: 0.03s
Iteration 5960, Loss: 0.0143, Time: 0.04s
Iteration 5961, Loss: 0.0132, Time: 0.04s
Iteration 5962, Loss: 0.0227, Time: 0.04s
Iteration 5963, Loss: 0.0119, Time: 0.03s
Iteration 5964, Loss: 0.0091, Time: 0.03s
Iteration 5965, Loss: 0.0069, Time: 0.03s
Iteration 5966, Loss: 0.0147, Time: 0.03s
Iteration 5967, Loss: 0.0128, Time: 0.03s
Iteration 5968, Loss: 0.0166, Time: 0.03s
Iteration 5969, Loss: 0.0127, Time: 0.03s
Iteration 5970, Loss: 0.0145, Time: 0.04s
Iteration 5971, Loss: 0.0145, Time: 0.03s
Iteration 5972, Loss: 0.0111, Time: 0.03s
Iteration 5973, Loss: 0.0161, Time: 0.03s
Iteration 5974, Loss: 0.0148, Time: 0.03s
Iteration 5975, Loss: 0.0163, Time: 0.03s
Iteration 5976, Loss: 0.0122, Time: 0.03s
```

```
Iteration 5977, Loss: 0.0146, Time: 0.03s
Iteration 5978, Loss: 0.0112, Time: 0.03s
Iteration 5979, Loss: 0.0151, Time: 0.03s
Iteration 5980, Loss: 0.0150, Time: 0.03s
Iteration 5981, Loss: 0.0121, Time: 0.03s
Iteration 5982, Loss: 0.0157, Time: 0.03s
Iteration 5983, Loss: 0.0135, Time: 0.03s
Iteration 5984, Loss: 0.0130, Time: 0.03s
Iteration 5985, Loss: 0.0114, Time: 0.03s
Iteration 5986, Loss: 0.0122, Time: 0.03s
Iteration 5987, Loss: 0.0091, Time: 0.03s
Iteration 5988, Loss: 0.0116, Time: 0.03s
Iteration 5989, Loss: 0.0156, Time: 0.03s
Iteration 5990, Loss: 0.0104, Time: 0.03s
Iteration 5991, Loss: 0.0171, Time: 0.03s
Iteration 5992, Loss: 0.0141, Time: 0.03s
Iteration 5993, Loss: 0.0211, Time: 0.03s
Iteration 5994, Loss: 0.0111, Time: 0.03s
Iteration 5995, Loss: 0.0102, Time: 0.03s
Iteration 5996, Loss: 0.0141, Time: 0.03s
Iteration 5997, Loss: 0.0124, Time: 0.03s
Iteration 5998, Loss: 0.0108, Time: 0.03s
Iteration 5999, Loss: 0.0117, Time: 0.03s
Iteration 6000, Loss: 0.0139, Time: 0.03s
Iteration 6000, Loss: 0.0139, Time: 0.03s
Test Loss: 0.0784
Iteration 6001, Loss: 0.0127, Time: 0.03s
Iteration 6002, Loss: 0.0260, Time: 0.03s
Iteration 6003, Loss: 0.0101, Time: 0.03s
Iteration 6004, Loss: 0.0126, Time: 0.03s
Iteration 6005, Loss: 0.0111, Time: 0.03s
Iteration 6006, Loss: 0.0123, Time: 0.03s
Iteration 6007, Loss: 0.0171, Time: 0.03s
Iteration 6008, Loss: 0.0100, Time: 0.03s
Iteration 6009, Loss: 0.0151, Time: 0.03s
Iteration 6010, Loss: 0.0190, Time: 0.03s
Iteration 6011, Loss: 0.0147, Time: 0.03s
Iteration 6012, Loss: 0.0182, Time: 0.03s
Iteration 6013, Loss: 0.0152, Time: 0.03s
Iteration 6014, Loss: 0.0196, Time: 0.03s
Iteration 6015, Loss: 0.0138, Time: 0.03s
Iteration 6016, Loss: 0.0098, Time: 0.03s
Iteration 6017, Loss: 0.0108, Time: 0.03s
Iteration 6018, Loss: 0.0084, Time: 0.03s
Iteration 6019, Loss: 0.0128, Time: 0.03s
Iteration 6020, Loss: 0.0137, Time: 0.03s
Iteration 6021, Loss: 0.0186, Time: 0.03s
Iteration 6022, Loss: 0.0210, Time: 0.03s
```

```
Iteration 6023, Loss: 0.0164, Time: 0.03s
Iteration 6024, Loss: 0.0086, Time: 0.03s
Iteration 6025, Loss: 0.0113, Time: 0.03s
Iteration 6026, Loss: 0.0160, Time: 0.03s
Iteration 6027, Loss: 0.0097, Time: 0.03s
Iteration 6028, Loss: 0.0145, Time: 0.03s
Iteration 6029, Loss: 0.0164, Time: 0.03s
Iteration 6030, Loss: 0.0081, Time: 0.03s
Iteration 6031, Loss: 0.0120, Time: 0.03s
Iteration 6032, Loss: 0.0159, Time: 0.03s
Iteration 6033, Loss: 0.0135, Time: 0.03s
Iteration 6034, Loss: 0.0120, Time: 0.03s
Iteration 6035, Loss: 0.0118, Time: 0.03s
Iteration 6036, Loss: 0.0146, Time: 0.03s
Iteration 6037, Loss: 0.0164, Time: 0.03s
Iteration 6038, Loss: 0.0156, Time: 0.03s
Iteration 6039, Loss: 0.0097, Time: 0.03s
Iteration 6040, Loss: 0.0081, Time: 0.03s
Iteration 6041, Loss: 0.0119, Time: 0.03s
Iteration 6042, Loss: 0.0166, Time: 0.03s
Iteration 6043, Loss: 0.0098, Time: 0.03s
Iteration 6044, Loss: 0.0075, Time: 0.03s
Iteration 6045, Loss: 0.0150, Time: 0.03s
Iteration 6046, Loss: 0.0131, Time: 0.03s
Iteration 6047, Loss: 0.0110, Time: 0.03s
Iteration 6048, Loss: 0.0076, Time: 0.03s
Iteration 6049, Loss: 0.0123, Time: 0.03s
Iteration 6050, Loss: 0.0149, Time: 0.03s
Iteration 6051, Loss: 0.0181, Time: 0.03s
Iteration 6052, Loss: 0.0129, Time: 0.03s
Iteration 6053, Loss: 0.0141, Time: 0.03s
Iteration 6054, Loss: 0.0155, Time: 0.03s
Iteration 6055, Loss: 0.0112, Time: 0.03s
Iteration 6056, Loss: 0.0115, Time: 0.03s
Iteration 6057, Loss: 0.0093, Time: 0.03s
Iteration 6058, Loss: 0.0064, Time: 0.03s
Iteration 6059, Loss: 0.0136, Time: 0.03s
Iteration 6060, Loss: 0.0091, Time: 0.03s
Iteration 6061, Loss: 0.0175, Time: 0.03s
Iteration 6062, Loss: 0.0216, Time: 0.03s
Iteration 6063, Loss: 0.0090, Time: 0.03s
Iteration 6064, Loss: 0.0130, Time: 0.03s
Iteration 6065, Loss: 0.0117, Time: 0.03s
Iteration 6066, Loss: 0.0122, Time: 0.03s
Iteration 6067, Loss: 0.0125, Time: 0.03s
Iteration 6068, Loss: 0.0130, Time: 0.03s
Iteration 6069, Loss: 0.0157, Time: 0.03s
Iteration 6070, Loss: 0.0144, Time: 0.03s
```

```
Iteration 6071, Loss: 0.0138, Time: 0.03s
Iteration 6072, Loss: 0.0108, Time: 0.03s
Iteration 6073, Loss: 0.0178, Time: 0.03s
Iteration 6074, Loss: 0.0114, Time: 0.03s
Iteration 6075, Loss: 0.0158, Time: 0.03s
Iteration 6076, Loss: 0.0102, Time: 0.03s
Iteration 6077, Loss: 0.0148, Time: 0.03s
Iteration 6078, Loss: 0.0079, Time: 0.03s
Iteration 6079, Loss: 0.0118, Time: 0.03s
Iteration 6080, Loss: 0.0135, Time: 0.03s
Iteration 6081, Loss: 0.0109, Time: 0.03s
Iteration 6082, Loss: 0.0110, Time: 0.03s
Iteration 6083, Loss: 0.0097, Time: 0.03s
Iteration 6084, Loss: 0.0121, Time: 0.03s
Iteration 6085, Loss: 0.0168, Time: 0.03s
Iteration 6086, Loss: 0.0072, Time: 0.03s
Iteration 6087, Loss: 0.0137, Time: 0.03s
Iteration 6088, Loss: 0.0134, Time: 0.03s
Iteration 6089, Loss: 0.0117, Time: 0.03s
Iteration 6090, Loss: 0.0146, Time: 0.03s
Iteration 6091, Loss: 0.0115, Time: 0.03s
Iteration 6092, Loss: 0.0132, Time: 0.03s
Iteration 6093, Loss: 0.0059, Time: 0.03s
Iteration 6094, Loss: 0.0141, Time: 0.03s
Iteration 6095, Loss: 0.0153, Time: 0.03s
Iteration 6096, Loss: 0.0114, Time: 0.03s
Iteration 6097, Loss: 0.0112, Time: 0.03s
Iteration 6098, Loss: 0.0155, Time: 0.03s
Iteration 6099, Loss: 0.0100, Time: 0.03s
Iteration 6100, Loss: 0.0150, Time: 0.03s
Iteration 6100, Loss: 0.0150, Time: 0.03s
Test Loss: 0.0268
Iteration 6101, Loss: 0.0079, Time: 0.03s
Iteration 6102, Loss: 0.0133, Time: 0.03s
Iteration 6103, Loss: 0.0151, Time: 0.03s
Iteration 6104, Loss: 0.0119, Time: 0.03s
Iteration 6105, Loss: 0.0137, Time: 0.03s
Iteration 6106, Loss: 0.0137, Time: 0.03s
Iteration 6107, Loss: 0.0113, Time: 0.03s
Iteration 6108, Loss: 0.0207, Time: 0.03s
Iteration 6109, Loss: 0.0089, Time: 0.03s
Iteration 6110, Loss: 0.0067, Time: 0.03s
Iteration 6111, Loss: 0.0123, Time: 0.03s
Iteration 6112, Loss: 0.0162, Time: 0.03s
Iteration 6113, Loss: 0.0083, Time: 0.03s
Iteration 6114, Loss: 0.0119, Time: 0.03s
Iteration 6115, Loss: 0.0108, Time: 0.03s
Iteration 6116, Loss: 0.0209, Time: 0.03s
```

```
Iteration 6117, Loss: 0.0165, Time: 0.03s
Iteration 6118, Loss: 0.0120, Time: 0.03s
Iteration 6119, Loss: 0.0160, Time: 0.03s
Iteration 6120, Loss: 0.0140, Time: 0.03s
Iteration 6121, Loss: 0.0145, Time: 0.03s
Iteration 6122, Loss: 0.0141, Time: 0.03s
Iteration 6123, Loss: 0.0134, Time: 0.03s
Iteration 6124, Loss: 0.0091, Time: 0.03s
Iteration 6125, Loss: 0.0103, Time: 0.03s
Iteration 6126, Loss: 0.0125, Time: 0.03s
Iteration 6127, Loss: 0.0123, Time: 0.03s
Iteration 6128, Loss: 0.0144, Time: 0.03s
Iteration 6129, Loss: 0.0099, Time: 0.03s
Iteration 6130, Loss: 0.0130, Time: 0.03s
Iteration 6131, Loss: 0.0108, Time: 0.03s
Iteration 6132, Loss: 0.0111, Time: 0.03s
Iteration 6133, Loss: 0.0090, Time: 0.03s
Iteration 6134, Loss: 0.0129, Time: 0.03s
Iteration 6135, Loss: 0.0162, Time: 0.03s
Iteration 6136, Loss: 0.0155, Time: 0.03s
Iteration 6137, Loss: 0.0141, Time: 0.03s
Iteration 6138, Loss: 0.0116, Time: 0.03s
Iteration 6139, Loss: 0.0122, Time: 0.03s
Iteration 6140, Loss: 0.0146, Time: 0.03s
Iteration 6141, Loss: 0.0200, Time: 0.03s
Iteration 6142, Loss: 0.0162, Time: 0.03s
Iteration 6143, Loss: 0.0126, Time: 0.03s
Iteration 6144, Loss: 0.0113, Time: 0.03s
Iteration 6145, Loss: 0.0101, Time: 0.03s
Iteration 6146, Loss: 0.0124, Time: 0.03s
Iteration 6147, Loss: 0.0117, Time: 0.03s
Iteration 6148, Loss: 0.0112, Time: 0.03s
Iteration 6149, Loss: 0.0119, Time: 0.03s
Iteration 6150, Loss: 0.0110, Time: 0.03s
Iteration 6151, Loss: 0.0091, Time: 0.03s
Iteration 6152, Loss: 0.0146, Time: 0.03s
Iteration 6153, Loss: 0.0091, Time: 0.03s
Iteration 6154, Loss: 0.0063, Time: 0.03s
Iteration 6155, Loss: 0.0069, Time: 0.03s
Iteration 6156, Loss: 0.0137, Time: 0.03s
Iteration 6157, Loss: 0.0206, Time: 0.03s
Iteration 6158, Loss: 0.0108, Time: 0.03s
Iteration 6159, Loss: 0.0090, Time: 0.03s
Iteration 6160, Loss: 0.0174, Time: 0.03s
Iteration 6161, Loss: 0.0127, Time: 0.03s
Iteration 6162, Loss: 0.0137, Time: 0.03s
Iteration 6163, Loss: 0.0076, Time: 0.03s
Iteration 6164, Loss: 0.0093, Time: 0.03s
```

```
Iteration 6165, Loss: 0.0104, Time: 0.03s
Iteration 6166, Loss: 0.0077, Time: 0.03s
Iteration 6167, Loss: 0.0102, Time: 0.03s
Iteration 6168, Loss: 0.0153, Time: 0.03s
Iteration 6169, Loss: 0.0121, Time: 0.03s
Iteration 6170, Loss: 0.0181, Time: 0.03s
Iteration 6171, Loss: 0.0090, Time: 0.03s
Iteration 6172, Loss: 0.0089, Time: 0.03s
Iteration 6173, Loss: 0.0087, Time: 0.03s
Iteration 6174, Loss: 0.0099, Time: 0.04s
Iteration 6175, Loss: 0.0130, Time: 0.03s
Iteration 6176, Loss: 0.0110, Time: 0.03s
Iteration 6177, Loss: 0.0091, Time: 0.03s
Iteration 6178, Loss: 0.0105, Time: 0.03s
Iteration 6179, Loss: 0.0076, Time: 0.03s
Iteration 6180, Loss: 0.0106, Time: 0.04s
Iteration 6181, Loss: 0.0105, Time: 0.03s
Iteration 6182, Loss: 0.0147, Time: 0.03s
Iteration 6183, Loss: 0.0124, Time: 0.04s
Iteration 6184, Loss: 0.0153, Time: 0.04s
Iteration 6185, Loss: 0.0091, Time: 0.03s
Iteration 6186, Loss: 0.0153, Time: 0.03s
Iteration 6187, Loss: 0.0121, Time: 0.04s
Iteration 6188, Loss: 0.0094, Time: 0.03s
Iteration 6189, Loss: 0.0097, Time: 0.03s
Iteration 6190, Loss: 0.0065, Time: 0.03s
Iteration 6191, Loss: 0.0126, Time: 0.03s
Iteration 6192, Loss: 0.0078, Time: 0.03s
Iteration 6193, Loss: 0.0164, Time: 0.03s
Iteration 6194, Loss: 0.0126, Time: 0.03s
Iteration 6195, Loss: 0.0145, Time: 0.03s
Iteration 6196, Loss: 0.0136, Time: 0.03s
Iteration 6197, Loss: 0.0097, Time: 0.03s
Iteration 6198, Loss: 0.0134, Time: 0.03s
Iteration 6199, Loss: 0.0139, Time: 0.03s
Iteration 6200, Loss: 0.0136, Time: 0.03s
Iteration 6200, Loss: 0.0136, Time: 0.03s
Test Loss: 0.0425
Iteration 6201, Loss: 0.0104, Time: 0.03s
Iteration 6202, Loss: 0.0099, Time: 0.03s
Iteration 6203, Loss: 0.0141, Time: 0.03s
Iteration 6204, Loss: 0.0161, Time: 0.03s
Iteration 6205, Loss: 0.0130, Time: 0.03s
Iteration 6206, Loss: 0.0138, Time: 0.03s
Iteration 6207, Loss: 0.0130, Time: 0.03s
Iteration 6208, Loss: 0.0069, Time: 0.03s
Iteration 6209, Loss: 0.0113, Time: 0.03s
Iteration 6210, Loss: 0.0151, Time: 0.03s
```

```
Iteration 6211, Loss: 0.0136, Time: 0.03s
Iteration 6212, Loss: 0.0162, Time: 0.03s
Iteration 6213, Loss: 0.0074, Time: 0.03s
Iteration 6214, Loss: 0.0058, Time: 0.03s
Iteration 6215, Loss: 0.0089, Time: 0.03s
Iteration 6216, Loss: 0.0110, Time: 0.03s
Iteration 6217, Loss: 0.0166, Time: 0.03s
Iteration 6218, Loss: 0.0126, Time: 0.03s
Iteration 6219, Loss: 0.0134, Time: 0.03s
Iteration 6220, Loss: 0.0124, Time: 0.03s
Iteration 6221, Loss: 0.0092, Time: 0.03s
Iteration 6222, Loss: 0.0105, Time: 0.03s
Iteration 6223, Loss: 0.0122, Time: 0.03s
Iteration 6224, Loss: 0.0120, Time: 0.03s
Iteration 6225, Loss: 0.0087, Time: 0.03s
Iteration 6226, Loss: 0.0087, Time: 0.03s
Iteration 6227, Loss: 0.0125, Time: 0.03s
Iteration 6228, Loss: 0.0146, Time: 0.03s
Iteration 6229, Loss: 0.0117, Time: 0.03s
Iteration 6230, Loss: 0.0147, Time: 0.03s
Iteration 6231, Loss: 0.0129, Time: 0.03s
Iteration 6232, Loss: 0.0096, Time: 0.03s
Iteration 6233, Loss: 0.0182, Time: 0.03s
Iteration 6234, Loss: 0.0099, Time: 0.03s
Iteration 6235, Loss: 0.0179, Time: 0.03s
Iteration 6236, Loss: 0.0071, Time: 0.03s
Iteration 6237, Loss: 0.0083, Time: 0.03s
Iteration 6238, Loss: 0.0139, Time: 0.03s
Iteration 6239, Loss: 0.0110, Time: 0.03s
Iteration 6240, Loss: 0.0106, Time: 0.03s
Iteration 6241, Loss: 0.0089, Time: 0.03s
Iteration 6242, Loss: 0.0090, Time: -0.11s
Iteration 6243, Loss: 0.0099, Time: 0.03s
Iteration 6244, Loss: 0.0128, Time: 0.03s
Iteration 6245, Loss: 0.0117, Time: 0.04s
Iteration 6246, Loss: 0.0100, Time: 0.04s
Iteration 6247, Loss: 0.0151, Time: 0.04s
Iteration 6248, Loss: 0.0109, Time: 0.03s
Iteration 6249, Loss: 0.0151, Time: 0.04s
Iteration 6250, Loss: 0.0129, Time: 0.04s
Iteration 6251, Loss: 0.0106, Time: 0.03s
Iteration 6252, Loss: 0.0130, Time: 0.03s
Iteration 6253, Loss: 0.0115, Time: 0.03s
Iteration 6254, Loss: 0.0097, Time: 0.03s
Iteration 6255, Loss: 0.0124, Time: 0.04s
Iteration 6256, Loss: 0.0107, Time: 0.04s
Iteration 6257, Loss: 0.0159, Time: 0.03s
Iteration 6258, Loss: 0.0120, Time: 0.03s
```

```
Iteration 6259, Loss: 0.0152, Time: 0.03s
Iteration 6260, Loss: 0.0123, Time: 0.03s
Iteration 6261, Loss: 0.0108, Time: 0.03s
Iteration 6262, Loss: 0.0117, Time: 0.03s
Iteration 6263, Loss: 0.0138, Time: 0.03s
Iteration 6264, Loss: 0.0096, Time: 0.03s
Iteration 6265, Loss: 0.0209, Time: 0.03s
Iteration 6266, Loss: 0.0065, Time: 0.03s
Iteration 6267, Loss: 0.0127, Time: 0.03s
Iteration 6268, Loss: 0.0195, Time: 0.03s
Iteration 6269, Loss: 0.0111, Time: 0.03s
Iteration 6270, Loss: 0.0227, Time: 0.03s
Iteration 6271, Loss: 0.0150, Time: 0.03s
Iteration 6272, Loss: 0.0163, Time: 0.03s
Iteration 6273, Loss: 0.0090, Time: 0.03s
Iteration 6274, Loss: 0.0182, Time: 0.03s
Iteration 6275, Loss: 0.0129, Time: 0.03s
Iteration 6276, Loss: 0.0104, Time: 0.03s
Iteration 6277, Loss: 0.0143, Time: 0.03s
Iteration 6278, Loss: 0.0118, Time: 0.03s
Iteration 6279, Loss: 0.0067, Time: 0.03s
Iteration 6280, Loss: 0.0098, Time: 0.03s
Iteration 6281, Loss: 0.0176, Time: 0.03s
Iteration 6282, Loss: 0.0095, Time: 0.03s
Iteration 6283, Loss: 0.0153, Time: 0.03s
Iteration 6284, Loss: 0.0100, Time: 0.03s
Iteration 6285, Loss: 0.0082, Time: 0.03s
Iteration 6286, Loss: 0.0097, Time: 0.03s
Iteration 6287, Loss: 0.0139, Time: 0.03s
Iteration 6288, Loss: 0.0132, Time: 0.03s
Iteration 6289, Loss: 0.0168, Time: 0.03s
Iteration 6290, Loss: 0.0071, Time: 0.03s
Iteration 6291, Loss: 0.0086, Time: 0.03s
Iteration 6292, Loss: 0.0154, Time: 0.03s
Iteration 6293, Loss: 0.0186, Time: 0.03s
Iteration 6294, Loss: 0.0120, Time: 0.03s
Iteration 6295, Loss: 0.0103, Time: 0.03s
Iteration 6296, Loss: 0.0135, Time: 0.03s
Iteration 6297, Loss: 0.0199, Time: 0.03s
Iteration 6298, Loss: 0.0134, Time: 0.03s
Iteration 6299, Loss: 0.0142, Time: 0.03s
Iteration 6300, Loss: 0.0103, Time: 0.03s
Iteration 6300, Loss: 0.0103, Time: 0.03s
Test Loss: 0.0253
Iteration 6301, Loss: 0.0144, Time: 0.03s
Iteration 6302, Loss: 0.0120, Time: 0.03s
Iteration 6303, Loss: 0.0087, Time: 0.03s
Iteration 6304, Loss: 0.0142, Time: 0.03s
```

```
Iteration 6305, Loss: 0.0131, Time: 0.03s
Iteration 6306, Loss: 0.0094, Time: 0.03s
Iteration 6307, Loss: 0.0080, Time: 0.03s
Iteration 6308, Loss: 0.0127, Time: 0.03s
Iteration 6309, Loss: 0.0094, Time: 0.03s
Iteration 6310, Loss: 0.0135, Time: 0.03s
Iteration 6311, Loss: 0.0144, Time: 0.03s
Iteration 6312, Loss: 0.0122, Time: 0.03s
Iteration 6313, Loss: 0.0146, Time: 0.03s
Iteration 6314, Loss: 0.0068, Time: 0.03s
Iteration 6315, Loss: 0.0132, Time: 0.03s
Iteration 6316, Loss: 0.0091, Time: 0.03s
Iteration 6317, Loss: 0.0108, Time: 0.03s
Iteration 6318, Loss: 0.0141, Time: 0.03s
Iteration 6319, Loss: 0.0127, Time: 0.03s
Iteration 6320, Loss: 0.0158, Time: 0.03s
Iteration 6321, Loss: 0.0103, Time: 0.03s
Iteration 6322, Loss: 0.0094, Time: 0.03s
Iteration 6323, Loss: 0.0134, Time: 0.03s
Iteration 6324, Loss: 0.0093, Time: 0.03s
Iteration 6325, Loss: 0.0119, Time: 0.03s
Iteration 6326, Loss: 0.0151, Time: 0.03s
Iteration 6327, Loss: 0.0105, Time: 0.03s
Iteration 6328, Loss: 0.0112, Time: 0.03s
Iteration 6329, Loss: 0.0104, Time: 0.03s
Iteration 6330, Loss: 0.0180, Time: 0.03s
Iteration 6331, Loss: 0.0177, Time: 0.03s
Iteration 6332, Loss: 0.0120, Time: 0.03s
Iteration 6333, Loss: 0.0113, Time: 0.03s
Iteration 6334, Loss: 0.0123, Time: 0.03s
Iteration 6335, Loss: 0.0131, Time: 0.03s
Iteration 6336, Loss: 0.0115, Time: 0.03s
Iteration 6337, Loss: 0.0094, Time: 0.03s
Iteration 6338, Loss: 0.0143, Time: 0.03s
Iteration 6339, Loss: 0.0123, Time: 0.03s
Iteration 6340, Loss: 0.0153, Time: 0.03s
Iteration 6341, Loss: 0.0073, Time: 0.03s
Iteration 6342, Loss: 0.0154, Time: 0.03s
Iteration 6343, Loss: 0.0082, Time: 0.03s
Iteration 6344, Loss: 0.0090, Time: 0.03s
Iteration 6345, Loss: 0.0078, Time: 0.03s
Iteration 6346, Loss: 0.0177, Time: 0.03s
Iteration 6347, Loss: 0.0124, Time: 0.03s
Iteration 6348, Loss: 0.0087, Time: 0.03s
Iteration 6349, Loss: 0.0118, Time: 0.03s
Iteration 6350, Loss: 0.0087, Time: 0.03s
Iteration 6351, Loss: 0.0117, Time: 0.03s
Iteration 6352, Loss: 0.0108, Time: 0.03s
```

```
Iteration 6353, Loss: 0.0055, Time: 0.03s
Iteration 6354, Loss: 0.0104, Time: 0.03s
Iteration 6355, Loss: 0.0161, Time: 0.03s
Iteration 6356, Loss: 0.0087, Time: 0.03s
Iteration 6357, Loss: 0.0140, Time: 0.03s
Iteration 6358, Loss: 0.0175, Time: 0.03s
Iteration 6359, Loss: 0.0132, Time: 0.03s
Iteration 6360, Loss: 0.0145, Time: 0.03s
Iteration 6361, Loss: 0.0130, Time: 0.03s
Iteration 6362, Loss: 0.0118, Time: 0.03s
Iteration 6363, Loss: 0.0151, Time: 0.03s
Iteration 6364, Loss: 0.0056, Time: 0.03s
Iteration 6365, Loss: 0.0091, Time: 0.03s
Iteration 6366, Loss: 0.0148, Time: 0.04s
Iteration 6367, Loss: 0.0120, Time: 0.03s
Iteration 6368, Loss: 0.0164, Time: 0.03s
Iteration 6369, Loss: 0.0098, Time: 0.03s
Iteration 6370, Loss: 0.0097, Time: 0.03s
Iteration 6371, Loss: 0.0112, Time: 0.03s
Iteration 6372, Loss: 0.0102, Time: 0.03s
Iteration 6373, Loss: 0.0113, Time: 0.03s
Iteration 6374, Loss: 0.0106, Time: 0.03s
Iteration 6375, Loss: 0.0119, Time: 0.03s
Iteration 6376, Loss: 0.0158, Time: 0.03s
Iteration 6377, Loss: 0.0124, Time: 0.03s
Iteration 6378, Loss: 0.0080, Time: 0.03s
Iteration 6379, Loss: 0.0111, Time: 0.03s
Iteration 6380, Loss: 0.0126, Time: 0.03s
Iteration 6381, Loss: 0.0161, Time: 0.03s
Iteration 6382, Loss: 0.0167, Time: 0.03s
Iteration 6383, Loss: 0.0129, Time: 0.03s
Iteration 6384, Loss: 0.0096, Time: 0.03s
Iteration 6385, Loss: 0.0160, Time: 0.03s
Iteration 6386, Loss: 0.0096, Time: 0.03s
Iteration 6387, Loss: 0.0151, Time: 0.03s
Iteration 6388, Loss: 0.0120, Time: 0.03s
Iteration 6389, Loss: 0.0075, Time: 0.03s
Iteration 6390, Loss: 0.0110, Time: 0.03s
Iteration 6391, Loss: 0.0114, Time: 0.03s
Iteration 6392, Loss: 0.0194, Time: 0.03s
Iteration 6393, Loss: 0.0113, Time: 0.03s
Iteration 6394, Loss: 0.0193, Time: 0.03s
Iteration 6395, Loss: 0.0090, Time: 0.03s
Iteration 6396, Loss: 0.0136, Time: 0.03s
Iteration 6397, Loss: 0.0170, Time: 0.03s
Iteration 6398, Loss: 0.0071, Time: 0.03s
Iteration 6399, Loss: 0.0140, Time: 0.03s
Iteration 6400, Loss: 0.0122, Time: 0.03s
```

```
Iteration 6400, Loss: 0.0122, Time: 0.03s
Test Loss: 0.0352
Iteration 6401, Loss: 0.0096, Time: 0.03s
Iteration 6402, Loss: 0.0071, Time: 0.03s
Iteration 6403, Loss: 0.0199, Time: 0.03s
Iteration 6404, Loss: 0.0116, Time: 0.03s
Iteration 6405, Loss: 0.0119, Time: 0.03s
Iteration 6406, Loss: 0.0139, Time: 0.03s
Iteration 6407, Loss: 0.0110, Time: 0.03s
Iteration 6408, Loss: 0.0130, Time: 0.03s
Iteration 6409, Loss: 0.0080, Time: 0.03s
Iteration 6410, Loss: 0.0080, Time: 0.03s
Iteration 6411, Loss: 0.0168, Time: 0.03s
Iteration 6412, Loss: 0.0051, Time: 0.03s
Iteration 6413, Loss: 0.0135, Time: 0.03s
Iteration 6414, Loss: 0.0157, Time: 0.03s
Iteration 6415, Loss: 0.0120, Time: 0.03s
Iteration 6416, Loss: 0.0110, Time: 0.03s
Iteration 6417, Loss: 0.0103, Time: 0.03s
Iteration 6418, Loss: 0.0139, Time: 0.03s
Iteration 6419, Loss: 0.0111, Time: 0.03s
Iteration 6420, Loss: 0.0151, Time: 0.03s
Iteration 6421, Loss: 0.0132, Time: 0.03s
Iteration 6422, Loss: 0.0154, Time: 0.03s
Iteration 6423, Loss: 0.0149, Time: 0.03s
Iteration 6424, Loss: 0.0136, Time: 0.03s
Iteration 6425, Loss: 0.0101, Time: 0.03s
Iteration 6426, Loss: 0.0101, Time: 0.03s
Iteration 6427, Loss: 0.0097, Time: 0.03s
Iteration 6428, Loss: 0.0123, Time: 0.03s
Iteration 6429, Loss: 0.0144, Time: 0.03s
Iteration 6430, Loss: 0.0110, Time: 0.03s
Iteration 6431, Loss: 0.0104, Time: 0.03s
Iteration 6432, Loss: 0.0158, Time: 0.03s
Iteration 6433, Loss: 0.0111, Time: 0.03s
Iteration 6434, Loss: 0.0123, Time: 0.03s
Iteration 6435, Loss: 0.0173, Time: 0.03s
Iteration 6436, Loss: 0.0164, Time: 0.03s
Iteration 6437, Loss: 0.0129, Time: 0.03s
Iteration 6438, Loss: 0.0120, Time: 0.03s
Iteration 6439, Loss: 0.0108, Time: 0.03s
Iteration 6440, Loss: 0.0077, Time: 0.03s
Iteration 6441, Loss: 0.0127, Time: 0.03s
Iteration 6442, Loss: 0.0116, Time: 0.03s
Iteration 6443, Loss: 0.0141, Time: 0.03s
Iteration 6444, Loss: 0.0180, Time: 0.03s
Iteration 6445, Loss: 0.0109, Time: 0.03s
Iteration 6446, Loss: 0.0170, Time: 0.03s
```

```
Iteration 6447, Loss: 0.0110, Time: 0.03s
Iteration 6448, Loss: 0.0121, Time: 0.03s
Iteration 6449, Loss: 0.0138, Time: 0.03s
Iteration 6450, Loss: 0.0116, Time: 0.03s
Iteration 6451, Loss: 0.0124, Time: 0.03s
Iteration 6452, Loss: 0.0113, Time: 0.03s
Iteration 6453, Loss: 0.0102, Time: 0.03s
Iteration 6454, Loss: 0.0144, Time: 0.03s
Iteration 6455, Loss: 0.0098, Time: 0.03s
Iteration 6456, Loss: 0.0134, Time: 0.03s
Iteration 6457, Loss: 0.0130, Time: 0.03s
Iteration 6458, Loss: 0.0092, Time: 0.03s
Iteration 6459, Loss: 0.0122, Time: 0.03s
Iteration 6460, Loss: 0.0096, Time: 0.03s
Iteration 6461, Loss: 0.0151, Time: 0.03s
Iteration 6462, Loss: 0.0123, Time: 0.03s
Iteration 6463, Loss: 0.0184, Time: 0.03s
Iteration 6464, Loss: 0.0107, Time: 0.03s
Iteration 6465, Loss: 0.0072, Time: 0.03s
Iteration 6466, Loss: 0.0120, Time: 0.03s
Iteration 6467, Loss: 0.0085, Time: 0.03s
Iteration 6468, Loss: 0.0155, Time: 0.03s
Iteration 6469, Loss: 0.0091, Time: 0.03s
Iteration 6470, Loss: 0.0132, Time: 0.03s
Iteration 6471, Loss: 0.0150, Time: 0.03s
Iteration 6472, Loss: 0.0104, Time: 0.03s
Iteration 6473, Loss: 0.0119, Time: 0.03s
Iteration 6474, Loss: 0.0094, Time: 0.03s
Iteration 6475, Loss: 0.0122, Time: 0.03s
Iteration 6476, Loss: 0.0093, Time: 0.03s
Iteration 6477, Loss: 0.0134, Time: 0.03s
Iteration 6478, Loss: 0.0192, Time: 0.03s
Iteration 6479, Loss: 0.0111, Time: 0.03s
Iteration 6480, Loss: 0.0123, Time: 0.03s
Iteration 6481, Loss: 0.0112, Time: 0.03s
Iteration 6482, Loss: 0.0099, Time: 0.03s
Iteration 6483, Loss: 0.0108, Time: 0.03s
Iteration 6484, Loss: 0.0091, Time: 0.03s
Iteration 6485, Loss: 0.0100, Time: 0.03s
Iteration 6486, Loss: 0.0111, Time: 0.03s
Iteration 6487, Loss: 0.0083, Time: 0.03s
Iteration 6488, Loss: 0.0124, Time: 0.03s
Iteration 6489, Loss: 0.0153, Time: 0.03s
Iteration 6490, Loss: 0.0134, Time: 0.03s
Iteration 6491, Loss: 0.0122, Time: 0.03s
Iteration 6492, Loss: 0.0093, Time: 0.03s
Iteration 6493, Loss: 0.0085, Time: 0.03s
Iteration 6494, Loss: 0.0093, Time: 0.03s
```

```
Iteration 6495, Loss: 0.0172, Time: 0.03s
Iteration 6496, Loss: 0.0127, Time: 0.03s
Iteration 6497, Loss: 0.0130, Time: 0.03s
Iteration 6498, Loss: 0.0084, Time: 0.03s
Iteration 6499, Loss: 0.0145, Time: 0.03s
Iteration 6500, Loss: 0.0116, Time: 0.03s
Iteration 6500, Loss: 0.0116, Time: 0.03s
Test Loss: 0.0392
Iteration 6501, Loss: 0.0109, Time: 0.03s
Iteration 6502, Loss: 0.0116, Time: 0.03s
Iteration 6503, Loss: 0.0103, Time: 0.03s
Iteration 6504, Loss: 0.0100, Time: 0.03s
Iteration 6505, Loss: 0.0060, Time: 0.03s
Iteration 6506, Loss: 0.0110, Time: 0.03s
Iteration 6507, Loss: 0.0139, Time: 0.03s
Iteration 6508, Loss: 0.0111, Time: 0.03s
Iteration 6509, Loss: 0.0116, Time: 0.03s
Iteration 6510, Loss: 0.0118, Time: 0.03s
Iteration 6511, Loss: 0.0138, Time: 0.03s
Iteration 6512, Loss: 0.0094, Time: 0.03s
Iteration 6513, Loss: 0.0053, Time: 0.03s
Iteration 6514, Loss: 0.0104, Time: 0.03s
Iteration 6515, Loss: 0.0093, Time: 0.03s
Iteration 6516, Loss: 0.0099, Time: 0.03s
Iteration 6517, Loss: 0.0119, Time: 0.03s
Iteration 6518, Loss: 0.0130, Time: 0.03s
Iteration 6519, Loss: 0.0092, Time: 0.03s
Iteration 6520, Loss: 0.0068, Time: 0.04s
Iteration 6521, Loss: 0.0169, Time: 0.03s
Iteration 6522, Loss: 0.0120, Time: 0.03s
Iteration 6523, Loss: 0.0167, Time: 0.03s
Iteration 6524, Loss: 0.0087, Time: 0.03s
Iteration 6525, Loss: 0.0133, Time: 0.03s
Iteration 6526, Loss: 0.0121, Time: 0.03s
Iteration 6527, Loss: 0.0117, Time: 0.03s
Iteration 6528, Loss: 0.0075, Time: 0.03s
Iteration 6529, Loss: 0.0136, Time: 0.03s
Iteration 6530, Loss: 0.0098, Time: 0.03s
Iteration 6531, Loss: 0.0080, Time: 0.03s
Iteration 6532, Loss: 0.0125, Time: 0.03s
Iteration 6533, Loss: 0.0167, Time: 0.03s
Iteration 6534, Loss: 0.0107, Time: 0.03s
Iteration 6535, Loss: 0.0161, Time: 0.03s
Iteration 6536, Loss: 0.0085, Time: 0.03s
Iteration 6537, Loss: 0.0089, Time: 0.03s
Iteration 6538, Loss: 0.0124, Time: 0.03s
Iteration 6539, Loss: 0.0135, Time: 0.03s
Iteration 6540, Loss: 0.0088, Time: 0.03s
```

```
Iteration 6541, Loss: 0.0126, Time: 0.03s
Iteration 6542, Loss: 0.0143, Time: 0.03s
Iteration 6543, Loss: 0.0078, Time: 0.03s
Iteration 6544, Loss: 0.0082, Time: 0.03s
Iteration 6545, Loss: 0.0130, Time: 0.03s
Iteration 6546, Loss: 0.0117, Time: 0.03s
Iteration 6547, Loss: 0.0092, Time: 0.03s
Iteration 6548, Loss: 0.0138, Time: 0.03s
Iteration 6549, Loss: 0.0162, Time: 0.03s
Iteration 6550, Loss: 0.0104, Time: 0.03s
Iteration 6551, Loss: 0.0160, Time: 0.03s
Iteration 6552, Loss: 0.0129, Time: 0.03s
Iteration 6553, Loss: 0.0095, Time: 0.03s
Iteration 6554, Loss: 0.0098, Time: 0.03s
Iteration 6555, Loss: 0.0131, Time: 0.03s
Iteration 6556, Loss: 0.0082, Time: 0.03s
Iteration 6557, Loss: 0.0155, Time: 0.03s
Iteration 6558, Loss: 0.0132, Time: 0.03s
Iteration 6559, Loss: 0.0123, Time: 0.03s
Iteration 6560, Loss: 0.0123, Time: 0.03s
Iteration 6561, Loss: 0.0146, Time: 0.03s
Iteration 6562, Loss: 0.0156, Time: 0.03s
Iteration 6563, Loss: 0.0118, Time: 0.03s
Iteration 6564, Loss: 0.0158, Time: 0.03s
Iteration 6565, Loss: 0.0100, Time: 0.03s
Iteration 6566, Loss: 0.0111, Time: 0.03s
Iteration 6567, Loss: 0.0121, Time: 0.03s
Iteration 6568, Loss: 0.0069, Time: 0.03s
Iteration 6569, Loss: 0.0108, Time: 0.03s
Iteration 6570, Loss: 0.0093, Time: 0.03s
Iteration 6571, Loss: 0.0066, Time: 0.03s
Iteration 6572, Loss: 0.0128, Time: 0.03s
Iteration 6573, Loss: 0.0122, Time: 0.03s
Iteration 6574, Loss: 0.0098, Time: 0.03s
Iteration 6575, Loss: 0.0145, Time: 0.03s
Iteration 6576, Loss: 0.0109, Time: 0.03s
Iteration 6577, Loss: 0.0155, Time: 0.03s
Iteration 6578, Loss: 0.0194, Time: 0.03s
Iteration 6579, Loss: 0.0141, Time: 0.03s
Iteration 6580, Loss: 0.0161, Time: 0.03s
Iteration 6581, Loss: 0.0101, Time: 0.03s
Iteration 6582, Loss: 0.0142, Time: 0.03s
Iteration 6583, Loss: 0.0157, Time: 0.03s
Iteration 6584, Loss: 0.0130, Time: 0.03s
Iteration 6585, Loss: 0.0089, Time: 0.03s
Iteration 6586, Loss: 0.0143, Time: 0.03s
Iteration 6587, Loss: 0.0088, Time: 0.03s
Iteration 6588, Loss: 0.0153, Time: 0.03s
```

```
Iteration 6589, Loss: 0.0082, Time: 0.03s
Iteration 6590, Loss: 0.0117, Time: 0.03s
Iteration 6591, Loss: 0.0093, Time: 0.03s
Iteration 6592, Loss: 0.0172, Time: 0.03s
Iteration 6593, Loss: 0.0061, Time: 0.03s
Iteration 6594, Loss: 0.0088, Time: 0.03s
Iteration 6595, Loss: 0.0140, Time: 0.03s
Iteration 6596, Loss: 0.0145, Time: 0.03s
Iteration 6597, Loss: 0.0090, Time: 0.03s
Iteration 6598, Loss: 0.0099, Time: 0.03s
Iteration 6599, Loss: 0.0093, Time: 0.03s
Iteration 6600, Loss: 0.0175, Time: 0.03s
Iteration 6600, Loss: 0.0175, Time: 0.03s
Test Loss: 0.0195
Iteration 6601, Loss: 0.0147, Time: 0.03s
Iteration 6602, Loss: 0.0142, Time: 0.03s
Iteration 6603, Loss: 0.0141, Time: 0.03s
Iteration 6604, Loss: 0.0116, Time: 0.03s
Iteration 6605, Loss: 0.0113, Time: 0.03s
Iteration 6606, Loss: 0.0129, Time: 0.03s
Iteration 6607, Loss: 0.0142, Time: 0.03s
Iteration 6608, Loss: 0.0121, Time: 0.03s
Iteration 6609, Loss: 0.0127, Time: 0.03s
Iteration 6610, Loss: 0.0112, Time: 0.03s
Iteration 6611, Loss: 0.0123, Time: 0.03s
Iteration 6612, Loss: 0.0116, Time: 0.03s
Iteration 6613, Loss: 0.0163, Time: 0.03s
Iteration 6614, Loss: 0.0131, Time: 0.03s
Iteration 6615, Loss: 0.0098, Time: 0.03s
Iteration 6616, Loss: 0.0099, Time: 0.03s
Iteration 6617, Loss: 0.0099, Time: 0.03s
Iteration 6618, Loss: 0.0156, Time: 0.03s
Iteration 6619, Loss: 0.0142, Time: 0.03s
Iteration 6620, Loss: 0.0130, Time: 0.03s
Iteration 6621, Loss: 0.0104, Time: 0.03s
Iteration 6622, Loss: 0.0149, Time: 0.03s
Iteration 6623, Loss: 0.0089, Time: 0.03s
Iteration 6624, Loss: 0.0105, Time: 0.04s
Iteration 6625, Loss: 0.0165, Time: 0.03s
Iteration 6626, Loss: 0.0081, Time: 0.03s
Iteration 6627, Loss: 0.0156, Time: 0.04s
Iteration 6628, Loss: 0.0094, Time: 0.04s
Iteration 6629, Loss: 0.0126, Time: 0.04s
Iteration 6630, Loss: 0.0099, Time: 0.03s
Iteration 6631, Loss: 0.0143, Time: 0.04s
Iteration 6632, Loss: 0.0126, Time: 0.03s
Iteration 6633, Loss: 0.0148, Time: 0.03s
Iteration 6634, Loss: 0.0158, Time: 0.04s
```

```
Iteration 6635, Loss: 0.0124, Time: 0.04s
Iteration 6636, Loss: 0.0130, Time: 0.04s
Iteration 6637, Loss: 0.0136, Time: 0.03s
Iteration 6638, Loss: 0.0109, Time: 0.03s
Iteration 6639, Loss: 0.0170, Time: 0.03s
Iteration 6640, Loss: 0.0079, Time: 0.03s
Iteration 6641, Loss: 0.0084, Time: 0.03s
Iteration 6642, Loss: 0.0095, Time: 0.03s
Iteration 6643, Loss: 0.0102, Time: 0.03s
Iteration 6644, Loss: 0.0113, Time: 0.03s
Iteration 6645, Loss: 0.0096, Time: 0.03s
Iteration 6646, Loss: 0.0169, Time: 0.03s
Iteration 6647, Loss: 0.0170, Time: 0.03s
Iteration 6648, Loss: 0.0149, Time: 0.03s
Iteration 6649, Loss: 0.0118, Time: 0.03s
Iteration 6650, Loss: 0.0093, Time: 0.03s
Iteration 6651, Loss: 0.0105, Time: 0.03s
Iteration 6652, Loss: 0.0084, Time: 0.03s
Iteration 6653, Loss: 0.0061, Time: 0.03s
Iteration 6654, Loss: 0.0186, Time: 0.03s
Iteration 6655, Loss: 0.0117, Time: 0.03s
Iteration 6656, Loss: 0.0155, Time: 0.03s
Iteration 6657, Loss: 0.0118, Time: 0.03s
Iteration 6658, Loss: 0.0109, Time: 0.03s
Iteration 6659, Loss: 0.0114, Time: 0.03s
Iteration 6660, Loss: 0.0110, Time: 0.03s
Iteration 6661, Loss: 0.0181, Time: 0.03s
Iteration 6662, Loss: 0.0133, Time: 0.03s
Iteration 6663, Loss: 0.0107, Time: 0.03s
Iteration 6664, Loss: 0.0111, Time: 0.03s
Iteration 6665, Loss: 0.0070, Time: 0.03s
Iteration 6666, Loss: 0.0174, Time: 0.03s
Iteration 6667, Loss: 0.0131, Time: 0.03s
Iteration 6668, Loss: 0.0112, Time: 0.03s
Iteration 6669, Loss: 0.0124, Time: 0.03s
Iteration 6670, Loss: 0.0115, Time: 0.03s
Iteration 6671, Loss: 0.0114, Time: 0.03s
Iteration 6672, Loss: 0.0134, Time: 0.03s
Iteration 6673, Loss: 0.0131, Time: 0.03s
Iteration 6674, Loss: 0.0163, Time: 0.03s
Iteration 6675, Loss: 0.0113, Time: 0.03s
Iteration 6676, Loss: 0.0118, Time: 0.03s
Iteration 6677, Loss: 0.0063, Time: 0.03s
Iteration 6678, Loss: 0.0152, Time: 0.03s
Iteration 6679, Loss: 0.0133, Time: 0.03s
Iteration 6680, Loss: 0.0083, Time: 0.03s
Iteration 6681, Loss: 0.0104, Time: 0.03s
Iteration 6682, Loss: 0.0104, Time: 0.03s
```

```
Iteration 6683, Loss: 0.0219, Time: 0.03s
Iteration 6684, Loss: 0.0164, Time: 0.03s
Iteration 6685, Loss: 0.0154, Time: 0.03s
Iteration 6686, Loss: 0.0121, Time: 0.03s
Iteration 6687, Loss: 0.0107, Time: 0.03s
Iteration 6688, Loss: 0.0153, Time: 0.03s
Iteration 6689, Loss: 0.0105, Time: 0.03s
Iteration 6690, Loss: 0.0159, Time: 0.03s
Iteration 6691, Loss: 0.0104, Time: 0.03s
Iteration 6692, Loss: 0.0102, Time: 0.03s
Iteration 6693, Loss: 0.0155, Time: 0.03s
Iteration 6694, Loss: 0.0081, Time: 0.03s
Iteration 6695, Loss: 0.0128, Time: 0.03s
Iteration 6696, Loss: 0.0061, Time: 0.03s
Iteration 6697, Loss: 0.0124, Time: 0.03s
Iteration 6698, Loss: 0.0137, Time: 0.03s
Iteration 6699, Loss: 0.0130, Time: 0.03s
Iteration 6700, Loss: 0.0108, Time: 0.03s
Iteration 6700, Loss: 0.0108, Time: 0.03s
Test Loss: 0.0401
Iteration 6701, Loss: 0.0134, Time: 0.03s
Iteration 6702, Loss: 0.0118, Time: 0.03s
Iteration 6703, Loss: 0.0129, Time: 0.03s
Iteration 6704, Loss: 0.0150, Time: 0.03s
Iteration 6705, Loss: 0.0132, Time: 0.03s
Iteration 6706, Loss: 0.0110, Time: 0.03s
Iteration 6707, Loss: 0.0112, Time: 0.03s
Iteration 6708, Loss: 0.0173, Time: 0.03s
Iteration 6709, Loss: 0.0116, Time: 0.03s
Iteration 6710, Loss: 0.0119, Time: 0.03s
Iteration 6711, Loss: 0.0149, Time: 0.03s
Iteration 6712, Loss: 0.0125, Time: 0.03s
Iteration 6713, Loss: 0.0107, Time: 0.03s
Iteration 6714, Loss: 0.0104, Time: 0.03s
Iteration 6715, Loss: 0.0152, Time: 0.03s
Iteration 6716, Loss: 0.0138, Time: 0.03s
Iteration 6717, Loss: 0.0155, Time: 0.03s
Iteration 6718, Loss: 0.0108, Time: 0.03s
Iteration 6719, Loss: 0.0129, Time: 0.03s
Iteration 6720, Loss: 0.0131, Time: 0.03s
Iteration 6721, Loss: 0.0162, Time: 0.03s
Iteration 6722, Loss: 0.0086, Time: 0.03s
Iteration 6723, Loss: 0.0132, Time: 0.03s
Iteration 6724, Loss: 0.0114, Time: 0.03s
Iteration 6725, Loss: 0.0134, Time: 0.03s
Iteration 6726, Loss: 0.0148, Time: 0.03s
Iteration 6727, Loss: 0.0133, Time: 0.03s
Iteration 6728, Loss: 0.0141, Time: 0.03s
```

```
Iteration 6729, Loss: 0.0104, Time: 0.03s
Iteration 6730, Loss: 0.0104, Time: 0.03s
Iteration 6731, Loss: 0.0125, Time: 0.03s
Iteration 6732, Loss: 0.0126, Time: 0.03s
Iteration 6733, Loss: 0.0123, Time: 0.03s
Iteration 6734, Loss: 0.0099, Time: 0.03s
Iteration 6735, Loss: 0.0094, Time: 0.03s
Iteration 6736, Loss: 0.0125, Time: 0.03s
Iteration 6737, Loss: 0.0139, Time: 0.03s
Iteration 6738, Loss: 0.0106, Time: 0.03s
Iteration 6739, Loss: 0.0122, Time: 0.03s
Iteration 6740, Loss: 0.0102, Time: 0.03s
Iteration 6741, Loss: 0.0125, Time: 0.03s
Iteration 6742, Loss: 0.0069, Time: 0.03s
Iteration 6743, Loss: 0.0097, Time: 0.03s
Iteration 6744, Loss: 0.0230, Time: 0.03s
Iteration 6745, Loss: 0.0137, Time: 0.03s
Iteration 6746, Loss: 0.0201, Time: 0.03s
Iteration 6747, Loss: 0.0120, Time: 0.03s
Iteration 6748, Loss: 0.0093, Time: 0.03s
Iteration 6749, Loss: 0.0128, Time: 0.03s
Iteration 6750, Loss: 0.0100, Time: 0.03s
Iteration 6751, Loss: 0.0061, Time: 0.03s
Iteration 6752, Loss: 0.0142, Time: 0.03s
Iteration 6753, Loss: 0.0079, Time: 0.03s
Iteration 6754, Loss: 0.0094, Time: 0.03s
Iteration 6755, Loss: 0.0137, Time: 0.03s
Iteration 6756, Loss: 0.0097, Time: 0.03s
Iteration 6757, Loss: 0.0120, Time: 0.03s
Iteration 6758, Loss: 0.0132, Time: 0.03s
Iteration 6759, Loss: 0.0110, Time: 0.03s
Iteration 6760, Loss: 0.0148, Time: 0.03s
Iteration 6761, Loss: 0.0103, Time: 0.04s
Iteration 6762, Loss: 0.0154, Time: 0.03s
Iteration 6763, Loss: 0.0072, Time: 0.03s
Iteration 6764, Loss: 0.0131, Time: 0.03s
Iteration 6765, Loss: 0.0148, Time: 0.03s
Iteration 6766, Loss: 0.0126, Time: 0.03s
Iteration 6767, Loss: 0.0147, Time: 0.03s
Iteration 6768, Loss: 0.0101, Time: 0.03s
Iteration 6769, Loss: 0.0096, Time: 0.03s
Iteration 6770, Loss: 0.0130, Time: 0.04s
Iteration 6771, Loss: 0.0107, Time: 0.03s
Iteration 6772, Loss: 0.0107, Time: 0.03s
Iteration 6773, Loss: 0.0122, Time: 0.04s
Iteration 6774, Loss: 0.0118, Time: 0.03s
Iteration 6775, Loss: 0.0136, Time: 0.03s
Iteration 6776, Loss: 0.0118, Time: 0.03s
```

```
Iteration 6777, Loss: 0.0086, Time: 0.03s
Iteration 6778, Loss: 0.0100, Time: 0.03s
Iteration 6779, Loss: 0.0095, Time: 0.04s
Iteration 6780, Loss: 0.0116, Time: 0.03s
Iteration 6781, Loss: 0.0051, Time: 0.03s
Iteration 6782, Loss: 0.0080, Time: 0.03s
Iteration 6783, Loss: 0.0203, Time: 0.03s
Iteration 6784, Loss: 0.0086, Time: 0.03s
Iteration 6785, Loss: 0.0101, Time: 0.03s
Iteration 6786, Loss: 0.0102, Time: 0.03s
Iteration 6787, Loss: 0.0115, Time: 0.03s
Iteration 6788, Loss: 0.0114, Time: 0.03s
Iteration 6789, Loss: 0.0079, Time: 0.03s
Iteration 6790, Loss: 0.0166, Time: 0.03s
Iteration 6791, Loss: 0.0185, Time: 0.03s
Iteration 6792, Loss: 0.0155, Time: 0.03s
Iteration 6793, Loss: 0.0102, Time: 0.03s
Iteration 6794, Loss: 0.0119, Time: 0.03s
Iteration 6795, Loss: 0.0128, Time: 0.03s
Iteration 6796, Loss: 0.0138, Time: 0.03s
Iteration 6797, Loss: 0.0149, Time: 0.03s
Iteration 6798, Loss: 0.0107, Time: 0.03s
Iteration 6799, Loss: 0.0103, Time: 0.03s
Iteration 6800, Loss: 0.0216, Time: 0.03s
Iteration 6800, Loss: 0.0216, Time: 0.03s
Test Loss: 0.0262
Iteration 6801, Loss: 0.0148, Time: 0.03s
Iteration 6802, Loss: 0.0105, Time: 0.03s
Iteration 6803, Loss: 0.0096, Time: 0.04s
Iteration 6804, Loss: 0.0171, Time: 0.03s
Iteration 6805, Loss: 0.0124, Time: 0.03s
Iteration 6806, Loss: 0.0083, Time: 0.03s
Iteration 6807, Loss: 0.0091, Time: 0.03s
Iteration 6808, Loss: 0.0182, Time: 0.03s
Iteration 6809, Loss: 0.0131, Time: 0.04s
Iteration 6810, Loss: 0.0119, Time: 0.03s
Iteration 6811, Loss: 0.0185, Time: 0.04s
Iteration 6812, Loss: 0.0084, Time: 0.03s
Iteration 6813, Loss: 0.0093, Time: 0.03s
Iteration 6814, Loss: 0.0137, Time: 0.03s
Iteration 6815, Loss: 0.0171, Time: 0.03s
Iteration 6816, Loss: 0.0169, Time: 0.04s
Iteration 6817, Loss: 0.0190, Time: 0.03s
Iteration 6818, Loss: 0.0108, Time: 0.03s
Iteration 6819, Loss: 0.0127, Time: 0.03s
Iteration 6820, Loss: 0.0127, Time: 0.03s
Iteration 6821, Loss: 0.0080, Time: 0.03s
Iteration 6822, Loss: 0.0162, Time: 0.03s
```

```
Iteration 6823, Loss: 0.0157, Time: 0.03s
Iteration 6824, Loss: 0.0173, Time: 0.04s
Iteration 6825, Loss: 0.0146, Time: 0.04s
Iteration 6826, Loss: 0.0139, Time: 0.03s
Iteration 6827, Loss: 0.0095, Time: 0.03s
Iteration 6828, Loss: 0.0137, Time: 0.03s
Iteration 6829, Loss: 0.0078, Time: 0.03s
Iteration 6830, Loss: 0.0141, Time: 0.03s
Iteration 6831, Loss: 0.0095, Time: 0.03s
Iteration 6832, Loss: 0.0106, Time: 0.03s
Iteration 6833, Loss: 0.0120, Time: 0.03s
Iteration 6834, Loss: 0.0096, Time: 0.03s
Iteration 6835, Loss: 0.0160, Time: 0.03s
Iteration 6836, Loss: 0.0180, Time: 0.03s
Iteration 6837, Loss: 0.0105, Time: 0.03s
Iteration 6838, Loss: 0.0094, Time: 0.03s
Iteration 6839, Loss: 0.0147, Time: 0.03s
Iteration 6840, Loss: 0.0117, Time: 0.03s
Iteration 6841, Loss: 0.0129, Time: 0.03s
Iteration 6842, Loss: 0.0148, Time: 0.03s
Iteration 6843, Loss: 0.0100, Time: 0.03s
Iteration 6844, Loss: 0.0102, Time: 0.03s
Iteration 6845, Loss: 0.0109, Time: 0.04s
Iteration 6846, Loss: 0.0120, Time: 0.03s
Iteration 6847, Loss: 0.0147, Time: 0.04s
Iteration 6848, Loss: 0.0083, Time: 0.04s
Iteration 6849, Loss: 0.0084, Time: 0.04s
Iteration 6850, Loss: 0.0077, Time: 0.04s
Iteration 6851, Loss: 0.0140, Time: 0.04s
Iteration 6852, Loss: 0.0098, Time: 0.03s
Iteration 6853, Loss: 0.0194, Time: 0.03s
Iteration 6854, Loss: 0.0107, Time: 0.03s
Iteration 6855, Loss: 0.0110, Time: 0.03s
Iteration 6856, Loss: 0.0102, Time: 0.03s
Iteration 6857, Loss: 0.0125, Time: 0.03s
Iteration 6858, Loss: 0.0095, Time: 0.03s
Iteration 6859, Loss: 0.0148, Time: 0.03s
Iteration 6860, Loss: 0.0142, Time: 0.03s
Iteration 6861, Loss: 0.0089, Time: 0.03s
Iteration 6862, Loss: 0.0151, Time: 0.03s
Iteration 6863, Loss: 0.0096, Time: 0.03s
Iteration 6864, Loss: 0.0140, Time: 0.03s
Iteration 6865, Loss: 0.0088, Time: 0.03s
Iteration 6866, Loss: 0.0112, Time: 0.03s
Iteration 6867, Loss: 0.0125, Time: 0.03s
Iteration 6868, Loss: 0.0142, Time: 0.03s
Iteration 6869, Loss: 0.0102, Time: 0.03s
Iteration 6870, Loss: 0.0108, Time: 0.03s
```

```
Iteration 6871, Loss: 0.0128, Time: 0.03s
Iteration 6872, Loss: 0.0106, Time: 0.03s
Iteration 6873, Loss: 0.0148, Time: 0.03s
Iteration 6874, Loss: 0.0123, Time: 0.03s
Iteration 6875, Loss: 0.0100, Time: 0.03s
Iteration 6876, Loss: 0.0131, Time: 0.03s
Iteration 6877, Loss: 0.0119, Time: 0.03s
Iteration 6878, Loss: 0.0083, Time: 0.03s
Iteration 6879, Loss: 0.0114, Time: 0.03s
Iteration 6880, Loss: 0.0129, Time: 0.04s
Iteration 6881, Loss: 0.0074, Time: 0.04s
Iteration 6882, Loss: 0.0144, Time: 0.04s
Iteration 6883, Loss: 0.0110, Time: 0.03s
Iteration 6884, Loss: 0.0140, Time: 0.04s
Iteration 6885, Loss: 0.0123, Time: 0.04s
Iteration 6886, Loss: 0.0115, Time: 0.03s
Iteration 6887, Loss: 0.0078, Time: 0.03s
Iteration 6888, Loss: 0.0110, Time: 0.03s
Iteration 6889, Loss: 0.0093, Time: 0.03s
Iteration 6890, Loss: 0.0120, Time: 0.03s
Iteration 6891, Loss: 0.0109, Time: 0.03s
Iteration 6892, Loss: 0.0132, Time: 0.03s
Iteration 6893, Loss: 0.0131, Time: 0.03s
Iteration 6894, Loss: 0.0100, Time: 0.03s
Iteration 6895, Loss: 0.0077, Time: 0.03s
Iteration 6896, Loss: 0.0111, Time: 0.03s
Iteration 6897, Loss: 0.0085, Time: 0.03s
Iteration 6898, Loss: 0.0153, Time: 0.03s
Iteration 6899, Loss: 0.0102, Time: 0.03s
Iteration 6900, Loss: 0.0115, Time: 0.03s
Iteration 6900, Loss: 0.0115, Time: 0.03s
Test Loss: 0.0466
Iteration 6901, Loss: 0.0085, Time: 0.03s
Iteration 6902, Loss: 0.0148, Time: 0.03s
Iteration 6903, Loss: 0.0096, Time: 0.03s
Iteration 6904, Loss: 0.0144, Time: 0.03s
Iteration 6905, Loss: 0.0186, Time: 0.03s
Iteration 6906, Loss: 0.0081, Time: 0.03s
Iteration 6907, Loss: 0.0135, Time: 0.03s
Iteration 6908, Loss: 0.0108, Time: 0.03s
Iteration 6909, Loss: 0.0132, Time: 0.03s
Iteration 6910, Loss: 0.0100, Time: 0.03s
Iteration 6911, Loss: 0.0163, Time: 0.03s
Iteration 6912, Loss: 0.0125, Time: 0.03s
Iteration 6913, Loss: 0.0092, Time: 0.03s
Iteration 6914, Loss: 0.0122, Time: 0.03s
Iteration 6915, Loss: 0.0167, Time: 0.03s
Iteration 6916, Loss: 0.0164, Time: 0.03s
```

```
Iteration 6917, Loss: 0.0074, Time: 0.03s
Iteration 6918, Loss: 0.0081, Time: 0.03s
Iteration 6919, Loss: 0.0116, Time: 0.03s
Iteration 6920, Loss: 0.0167, Time: 0.03s
Iteration 6921, Loss: 0.0076, Time: 0.03s
Iteration 6922, Loss: 0.0103, Time: 0.03s
Iteration 6923, Loss: 0.0177, Time: 0.03s
Iteration 6924, Loss: 0.0124, Time: 0.03s
Iteration 6925, Loss: 0.0125, Time: 0.03s
Iteration 6926, Loss: 0.0122, Time: 0.03s
Iteration 6927, Loss: 0.0109, Time: 0.03s
Iteration 6928, Loss: 0.0111, Time: 0.03s
Iteration 6929, Loss: 0.0127, Time: 0.03s
Iteration 6930, Loss: 0.0125, Time: 0.03s
Iteration 6931, Loss: 0.0093, Time: 0.03s
Iteration 6932, Loss: 0.0114, Time: 0.03s
Iteration 6933, Loss: 0.0114, Time: 0.04s
Iteration 6934, Loss: 0.0106, Time: 0.03s
Iteration 6935, Loss: 0.0127, Time: 0.03s
Iteration 6936, Loss: 0.0144, Time: 0.03s
Iteration 6937, Loss: 0.0064, Time: 0.03s
Iteration 6938, Loss: 0.0101, Time: 0.03s
Iteration 6939, Loss: 0.0093, Time: 0.03s
Iteration 6940, Loss: 0.0077, Time: 0.03s
Iteration 6941, Loss: 0.0084, Time: 0.03s
Iteration 6942, Loss: 0.0127, Time: 0.03s
Iteration 6943, Loss: 0.0128, Time: 0.03s
Iteration 6944, Loss: 0.0144, Time: 0.03s
Iteration 6945, Loss: 0.0108, Time: 0.04s
Iteration 6946, Loss: 0.0151, Time: 0.04s
Iteration 6947, Loss: 0.0111, Time: 0.04s
Iteration 6948, Loss: 0.0131, Time: 0.04s
Iteration 6949, Loss: 0.0109, Time: 0.03s
Iteration 6950, Loss: 0.0138, Time: 0.04s
Iteration 6951, Loss: 0.0073, Time: 0.03s
Iteration 6952, Loss: 0.0089, Time: 0.03s
Iteration 6953, Loss: 0.0118, Time: 0.03s
Iteration 6954, Loss: 0.0132, Time: 0.04s
Iteration 6955, Loss: 0.0123, Time: 0.03s
Iteration 6956, Loss: 0.0125, Time: 0.03s
Iteration 6957, Loss: 0.0083, Time: 0.03s
Iteration 6958, Loss: 0.0063, Time: 0.03s
Iteration 6959, Loss: 0.0089, Time: 0.03s
Iteration 6960, Loss: 0.0140, Time: 0.03s
Iteration 6961, Loss: 0.0104, Time: 0.03s
Iteration 6962, Loss: 0.0085, Time: 0.03s
Iteration 6963, Loss: 0.0115, Time: 0.03s
Iteration 6964, Loss: 0.0123, Time: 0.03s
```

```
Iteration 6965, Loss: 0.0104, Time: 0.03s
Iteration 6966, Loss: 0.0139, Time: 0.03s
Iteration 6967, Loss: 0.0086, Time: 0.03s
Iteration 6968, Loss: 0.0086, Time: 0.03s
Iteration 6969, Loss: 0.0168, Time: 0.03s
Iteration 6970, Loss: 0.0146, Time: 0.03s
Iteration 6971, Loss: 0.0127, Time: 0.03s
Iteration 6972, Loss: 0.0095, Time: 0.03s
Iteration 6973, Loss: 0.0131, Time: 0.03s
Iteration 6974, Loss: 0.0141, Time: 0.03s
Iteration 6975, Loss: 0.0138, Time: 0.03s
Iteration 6976, Loss: 0.0070, Time: 0.03s
Iteration 6977, Loss: 0.0143, Time: 0.03s
Iteration 6978, Loss: 0.0167, Time: 0.03s
Iteration 6979, Loss: 0.0079, Time: 0.03s
Iteration 6980, Loss: 0.0132, Time: 0.03s
Iteration 6981, Loss: 0.0138, Time: 0.03s
Iteration 6982, Loss: 0.0084, Time: 0.03s
Iteration 6983, Loss: 0.0108, Time: 0.03s
Iteration 6984, Loss: 0.0061, Time: 0.03s
Iteration 6985, Loss: 0.0125, Time: 0.03s
Iteration 6986, Loss: 0.0107, Time: 0.03s
Iteration 6987, Loss: 0.0104, Time: 0.03s
Iteration 6988, Loss: 0.0113, Time: 0.03s
Iteration 6989, Loss: 0.0090, Time: 0.03s
Iteration 6990, Loss: 0.0098, Time: 0.03s
Iteration 6991, Loss: 0.0120, Time: 0.03s
Iteration 6992, Loss: 0.0088, Time: 0.03s
Iteration 6993, Loss: 0.0097, Time: 0.03s
Iteration 6994, Loss: 0.0143, Time: 0.03s
Iteration 6995, Loss: 0.0111, Time: 0.03s
Iteration 6996, Loss: 0.0126, Time: 0.03s
Iteration 6997, Loss: 0.0119, Time: 0.03s
Iteration 6998, Loss: 0.0092, Time: 0.03s
Iteration 6999, Loss: 0.0128, Time: 0.03s
Iteration 7000, Loss: 0.0164, Time: 0.03s
Iteration 7000, Loss: 0.0164, Time: 0.03s
Test Loss: 0.0182
Iteration 7001, Loss: 0.0088, Time: 0.03s
Iteration 7002, Loss: 0.0130, Time: 0.03s
Iteration 7003, Loss: 0.0128, Time: 0.03s
Iteration 7004, Loss: 0.0150, Time: 0.03s
Iteration 7005, Loss: 0.0077, Time: 0.03s
Iteration 7006, Loss: 0.0104, Time: 0.03s
Iteration 7007, Loss: 0.0084, Time: 0.03s
Iteration 7008, Loss: 0.0092, Time: 0.03s
Iteration 7009, Loss: 0.0149, Time: 0.03s
Iteration 7010, Loss: 0.0071, Time: 0.03s
```

```
Iteration 7011, Loss: 0.0129, Time: 0.03s
Iteration 7012, Loss: 0.0132, Time: 0.03s
Iteration 7013, Loss: 0.0121, Time: 0.03s
Iteration 7014, Loss: 0.0139, Time: 0.03s
Iteration 7015, Loss: 0.0151, Time: 0.03s
Iteration 7016, Loss: 0.0115, Time: 0.03s
Iteration 7017, Loss: 0.0089, Time: 0.03s
Iteration 7018, Loss: 0.0152, Time: 0.03s
Iteration 7019, Loss: 0.0105, Time: 0.03s
Iteration 7020, Loss: 0.0063, Time: 0.03s
Iteration 7021, Loss: 0.0094, Time: 0.03s
Iteration 7022, Loss: 0.0111, Time: 0.03s
Iteration 7023, Loss: 0.0093, Time: 0.03s
Iteration 7024, Loss: 0.0081, Time: 0.03s
Iteration 7025, Loss: 0.0130, Time: 0.03s
Iteration 7026, Loss: 0.0061, Time: 0.03s
Iteration 7027, Loss: 0.0120, Time: 0.03s
Iteration 7028, Loss: 0.0097, Time: 0.03s
Iteration 7029, Loss: 0.0094, Time: 0.03s
Iteration 7030, Loss: 0.0143, Time: 0.03s
Iteration 7031, Loss: 0.0119, Time: 0.03s
Iteration 7032, Loss: 0.0063, Time: 0.03s
Iteration 7033, Loss: 0.0124, Time: 0.03s
Iteration 7034, Loss: 0.0121, Time: 0.03s
Iteration 7035, Loss: 0.0121, Time: 0.03s
Iteration 7036, Loss: 0.0118, Time: 0.03s
Iteration 7037, Loss: 0.0153, Time: 0.03s
Iteration 7038, Loss: 0.0136, Time: 0.03s
Iteration 7039, Loss: 0.0093, Time: 0.03s
Iteration 7040, Loss: 0.0073, Time: 0.03s
Iteration 7041, Loss: 0.0102, Time: 0.03s
Iteration 7042, Loss: 0.0133, Time: 0.03s
Iteration 7043, Loss: 0.0081, Time: 0.03s
Iteration 7044, Loss: 0.0150, Time: 0.03s
Iteration 7045, Loss: 0.0105, Time: 0.03s
Iteration 7046, Loss: 0.0155, Time: 0.03s
Iteration 7047, Loss: 0.0118, Time: 0.03s
Iteration 7048, Loss: 0.0059, Time: 0.03s
Iteration 7049, Loss: 0.0132, Time: 0.03s
Iteration 7050, Loss: 0.0120, Time: 0.03s
Iteration 7051, Loss: 0.0138, Time: 0.03s
Iteration 7052, Loss: 0.0092, Time: 0.03s
Iteration 7053, Loss: 0.0141, Time: 0.03s
Iteration 7054, Loss: 0.0121, Time: 0.03s
Iteration 7055, Loss: 0.0097, Time: 0.03s
Iteration 7056, Loss: 0.0107, Time: 0.03s
Iteration 7057, Loss: 0.0106, Time: 0.03s
Iteration 7058, Loss: 0.0111, Time: 0.03s
```

```
Iteration 7059, Loss: 0.0134, Time: 0.03s
Iteration 7060, Loss: 0.0167, Time: 0.03s
Iteration 7061, Loss: 0.0130, Time: 0.03s
Iteration 7062, Loss: 0.0133, Time: 0.03s
Iteration 7063, Loss: 0.0159, Time: 0.03s
Iteration 7064, Loss: 0.0126, Time: 0.03s
Iteration 7065, Loss: 0.0103, Time: 0.03s
Iteration 7066, Loss: 0.0088, Time: 0.03s
Iteration 7067, Loss: 0.0100, Time: 0.03s
Iteration 7068, Loss: 0.0162, Time: 0.03s
Iteration 7069, Loss: 0.0088, Time: 0.03s
Iteration 7070, Loss: 0.0143, Time: 0.03s
Iteration 7071, Loss: 0.0130, Time: 0.03s
Iteration 7072, Loss: 0.0127, Time: 0.03s
Iteration 7073, Loss: 0.0179, Time: 0.03s
Iteration 7074, Loss: 0.0103, Time: 0.03s
Iteration 7075, Loss: 0.0133, Time: 0.03s
Iteration 7076, Loss: 0.0101, Time: 0.03s
Iteration 7077, Loss: 0.0049, Time: 0.03s
Iteration 7078, Loss: 0.0135, Time: 0.03s
Iteration 7079, Loss: 0.0104, Time: 0.03s
Iteration 7080, Loss: 0.0099, Time: 0.03s
Iteration 7081, Loss: 0.0073, Time: 0.03s
Iteration 7082, Loss: 0.0149, Time: 0.03s
Iteration 7083, Loss: 0.0131, Time: 0.03s
Iteration 7084, Loss: 0.0100, Time: 0.03s
Iteration 7085, Loss: 0.0167, Time: 0.03s
Iteration 7086, Loss: 0.0146, Time: 0.03s
Iteration 7087, Loss: 0.0129, Time: 0.03s
Iteration 7088, Loss: 0.0075, Time: 0.03s
Iteration 7089, Loss: 0.0176, Time: 0.03s
Iteration 7090, Loss: 0.0126, Time: 0.03s
Iteration 7091, Loss: 0.0137, Time: 0.03s
Iteration 7092, Loss: 0.0139, Time: 0.03s
Iteration 7093, Loss: 0.0118, Time: 0.03s
Iteration 7094, Loss: 0.0113, Time: 0.03s
Iteration 7095, Loss: 0.0138, Time: 0.03s
Iteration 7096, Loss: 0.0136, Time: 0.03s
Iteration 7097, Loss: 0.0118, Time: 0.03s
Iteration 7098, Loss: 0.0099, Time: 0.03s
Iteration 7099, Loss: 0.0110, Time: 0.03s
Iteration 7100, Loss: 0.0104, Time: 0.03s
Iteration 7100, Loss: 0.0104, Time: 0.03s
Test Loss: 0.0382
Iteration 7101, Loss: 0.0115, Time: 0.03s
Iteration 7102, Loss: 0.0122, Time: 0.03s
Iteration 7103, Loss: 0.0109, Time: 0.03s
Iteration 7104, Loss: 0.0128, Time: 0.03s
```

```
Iteration 7105, Loss: 0.0112, Time: 0.03s
Iteration 7106, Loss: 0.0139, Time: 0.03s
Iteration 7107, Loss: 0.0084, Time: 0.03s
Iteration 7108, Loss: 0.0083, Time: 0.03s
Iteration 7109, Loss: 0.0145, Time: 0.03s
Iteration 7110, Loss: 0.0106, Time: 0.03s
Iteration 7111, Loss: 0.0137, Time: 0.03s
Iteration 7112, Loss: 0.0100, Time: 0.04s
Iteration 7113, Loss: 0.0140, Time: 0.03s
Iteration 7114, Loss: 0.0122, Time: 0.03s
Iteration 7115, Loss: 0.0121, Time: 0.03s
Iteration 7116, Loss: 0.0102, Time: 0.03s
Iteration 7117, Loss: 0.0073, Time: 0.04s
Iteration 7118, Loss: 0.0159, Time: 0.03s
Iteration 7119, Loss: 0.0173, Time: 0.03s
Iteration 7120, Loss: 0.0108, Time: 0.03s
Iteration 7121, Loss: 0.0097, Time: 0.03s
Iteration 7122, Loss: 0.0110, Time: 0.03s
Iteration 7123, Loss: 0.0086, Time: 0.03s
Iteration 7124, Loss: 0.0106, Time: 0.03s
Iteration 7125, Loss: 0.0137, Time: 0.03s
Iteration 7126, Loss: 0.0094, Time: 0.03s
Iteration 7127, Loss: 0.0113, Time: 0.03s
Iteration 7128, Loss: 0.0124, Time: 0.03s
Iteration 7129, Loss: 0.0141, Time: 0.03s
Iteration 7130, Loss: 0.0142, Time: 0.03s
Iteration 7131, Loss: 0.0130, Time: 0.03s
Iteration 7132, Loss: 0.0097, Time: 0.03s
Iteration 7133, Loss: 0.0135, Time: 0.03s
Iteration 7134, Loss: 0.0149, Time: 0.03s
Iteration 7135, Loss: 0.0120, Time: 0.03s
Iteration 7136, Loss: 0.0085, Time: 0.03s
Iteration 7137, Loss: 0.0142, Time: 0.03s
Iteration 7138, Loss: 0.0099, Time: 0.03s
Iteration 7139, Loss: 0.0139, Time: 0.03s
Iteration 7140, Loss: 0.0119, Time: 0.03s
Iteration 7141, Loss: 0.0126, Time: 0.03s
Iteration 7142, Loss: 0.0121, Time: 0.03s
Iteration 7143, Loss: 0.0164, Time: 0.03s
Iteration 7144, Loss: 0.0121, Time: 0.03s
Iteration 7145, Loss: 0.0124, Time: 0.03s
Iteration 7146, Loss: 0.0146, Time: 0.03s
Iteration 7147, Loss: 0.0121, Time: 0.03s
Iteration 7148, Loss: 0.0111, Time: 0.03s
Iteration 7149, Loss: 0.0195, Time: 0.03s
Iteration 7150, Loss: 0.0122, Time: 0.03s
Iteration 7151, Loss: 0.0139, Time: 0.03s
Iteration 7152, Loss: 0.0076, Time: 0.03s
```

```
Iteration 7153, Loss: 0.0128, Time: 0.03s
Iteration 7154, Loss: 0.0085, Time: 0.03s
Iteration 7155, Loss: 0.0090, Time: 0.03s
Iteration 7156, Loss: 0.0195, Time: 0.03s
Iteration 7157, Loss: 0.0113, Time: 0.03s
Iteration 7158, Loss: 0.0100, Time: 0.03s
Iteration 7159, Loss: 0.0114, Time: 0.03s
Iteration 7160, Loss: 0.0126, Time: 0.03s
Iteration 7161, Loss: 0.0137, Time: 0.03s
Iteration 7162, Loss: 0.0107, Time: 0.03s
Iteration 7163, Loss: 0.0121, Time: 0.03s
Iteration 7164, Loss: 0.0090, Time: 0.03s
Iteration 7165, Loss: 0.0109, Time: 0.03s
Iteration 7166, Loss: 0.0103, Time: 0.03s
Iteration 7167, Loss: 0.0142, Time: 0.03s
Iteration 7168, Loss: 0.0086, Time: 0.03s
Iteration 7169, Loss: 0.0194, Time: 0.03s
Iteration 7170, Loss: 0.0103, Time: 0.03s
Iteration 7171, Loss: 0.0097, Time: 0.03s
Iteration 7172, Loss: 0.0121, Time: 0.03s
Iteration 7173, Loss: 0.0126, Time: 0.03s
Iteration 7174, Loss: 0.0117, Time: 0.03s
Iteration 7175, Loss: 0.0113, Time: 0.03s
Iteration 7176, Loss: 0.0109, Time: 0.03s
Iteration 7177, Loss: 0.0085, Time: 0.03s
Iteration 7178, Loss: 0.0102, Time: 0.03s
Iteration 7179, Loss: 0.0107, Time: 0.03s
Iteration 7180, Loss: 0.0079, Time: 0.03s
Iteration 7181, Loss: 0.0150, Time: 0.03s
Iteration 7182, Loss: 0.0159, Time: 0.03s
Iteration 7183, Loss: 0.0076, Time: 0.03s
Iteration 7184, Loss: 0.0139, Time: 0.03s
Iteration 7185, Loss: 0.0112, Time: 0.03s
Iteration 7186, Loss: 0.0137, Time: 0.03s
Iteration 7187, Loss: 0.0092, Time: 0.03s
Iteration 7188, Loss: 0.0182, Time: 0.03s
Iteration 7189, Loss: 0.0113, Time: 0.03s
Iteration 7190, Loss: 0.0093, Time: 0.03s
Iteration 7191, Loss: 0.0097, Time: 0.03s
Iteration 7192, Loss: 0.0101, Time: 0.03s
Iteration 7193, Loss: 0.0233, Time: 0.03s
Iteration 7194, Loss: 0.0117, Time: 0.03s
Iteration 7195, Loss: 0.0088, Time: 0.03s
Iteration 7196, Loss: 0.0084, Time: 0.03s
Iteration 7197, Loss: 0.0071, Time: 0.03s
Iteration 7198, Loss: 0.0064, Time: 0.03s
Iteration 7199, Loss: 0.0130, Time: 0.03s
Iteration 7200, Loss: 0.0113, Time: 0.03s
```

```
Iteration 7200, Loss: 0.0113, Time: 0.03s
Test Loss: 0.0688
Iteration 7201, Loss: 0.0088, Time: 0.03s
Iteration 7202, Loss: 0.0111, Time: 0.03s
Iteration 7203, Loss: 0.0130, Time: 0.03s
Iteration 7204, Loss: 0.0112, Time: 0.03s
Iteration 7205, Loss: 0.0139, Time: 0.03s
Iteration 7206, Loss: 0.0136, Time: 0.03s
Iteration 7207, Loss: 0.0065, Time: 0.03s
Iteration 7208, Loss: 0.0067, Time: 0.03s
Iteration 7209, Loss: 0.0117, Time: 0.03s
Iteration 7210, Loss: 0.0079, Time: 0.03s
Iteration 7211, Loss: 0.0109, Time: 0.03s
Iteration 7212, Loss: 0.0068, Time: 0.03s
Iteration 7213, Loss: 0.0090, Time: 0.03s
Iteration 7214, Loss: 0.0151, Time: 0.03s
Iteration 7215, Loss: 0.0125, Time: 0.03s
Iteration 7216, Loss: 0.0064, Time: 0.03s
Iteration 7217, Loss: 0.0098, Time: 0.03s
Iteration 7218, Loss: 0.0076, Time: 0.03s
Iteration 7219, Loss: 0.0105, Time: 0.03s
Iteration 7220, Loss: 0.0145, Time: 0.03s
Iteration 7221, Loss: 0.0086, Time: 0.03s
Iteration 7222, Loss: 0.0154, Time: 0.03s
Iteration 7223, Loss: 0.0146, Time: 0.03s
Iteration 7224, Loss: 0.0137, Time: 0.03s
Iteration 7225, Loss: 0.0105, Time: 0.03s
Iteration 7226, Loss: 0.0080, Time: 0.03s
Iteration 7227, Loss: 0.0115, Time: 0.03s
Iteration 7228, Loss: 0.0133, Time: 0.03s
Iteration 7229, Loss: 0.0081, Time: 0.03s
Iteration 7230, Loss: 0.0071, Time: 0.03s
Iteration 7231, Loss: 0.0068, Time: 0.03s
Iteration 7232, Loss: 0.0114, Time: 0.03s
Iteration 7233, Loss: 0.0104, Time: 0.03s
Iteration 7234, Loss: 0.0097, Time: 0.03s
Iteration 7235, Loss: 0.0144, Time: 0.03s
Iteration 7236, Loss: 0.0119, Time: 0.03s
Iteration 7237, Loss: 0.0082, Time: 0.03s
Iteration 7238, Loss: 0.0156, Time: 0.03s
Iteration 7239, Loss: 0.0147, Time: 0.03s
Iteration 7240, Loss: 0.0124, Time: -0.12s
Iteration 7241, Loss: 0.0068, Time: 0.03s
Iteration 7242, Loss: 0.0080, Time: 0.03s
Iteration 7243, Loss: 0.0048, Time: 0.03s
Iteration 7244, Loss: 0.0124, Time: 0.03s
Iteration 7245, Loss: 0.0147, Time: 0.03s
Iteration 7246, Loss: 0.0143, Time: 0.03s
```

```
Iteration 7247, Loss: 0.0178, Time: 0.03s
Iteration 7248, Loss: 0.0090, Time: 0.03s
Iteration 7249, Loss: 0.0109, Time: 0.03s
Iteration 7250, Loss: 0.0103, Time: 0.03s
Iteration 7251, Loss: 0.0151, Time: 0.03s
Iteration 7252, Loss: 0.0139, Time: 0.03s
Iteration 7253, Loss: 0.0122, Time: 0.03s
Iteration 7254, Loss: 0.0137, Time: 0.03s
Iteration 7255, Loss: 0.0089, Time: 0.03s
Iteration 7256, Loss: 0.0131, Time: 0.03s
Iteration 7257, Loss: 0.0137, Time: 0.03s
Iteration 7258, Loss: 0.0141, Time: 0.03s
Iteration 7259, Loss: 0.0141, Time: 0.03s
Iteration 7260, Loss: 0.0097, Time: 0.03s
Iteration 7261, Loss: 0.0112, Time: 0.03s
Iteration 7262, Loss: 0.0100, Time: 0.03s
Iteration 7263, Loss: 0.0120, Time: 0.03s
Iteration 7264, Loss: 0.0121, Time: 0.03s
Iteration 7265, Loss: 0.0075, Time: 0.03s
Iteration 7266, Loss: 0.0143, Time: 0.03s
Iteration 7267, Loss: 0.0110, Time: 0.03s
Iteration 7268, Loss: 0.0154, Time: 0.03s
Iteration 7269, Loss: 0.0103, Time: 0.03s
Iteration 7270, Loss: 0.0082, Time: 0.03s
Iteration 7271, Loss: 0.0104, Time: 0.03s
Iteration 7272, Loss: 0.0121, Time: 0.03s
Iteration 7273, Loss: 0.0106, Time: 0.03s
Iteration 7274, Loss: 0.0088, Time: 0.03s
Iteration 7275, Loss: 0.0094, Time: 0.03s
Iteration 7276, Loss: 0.0111, Time: 0.03s
Iteration 7277, Loss: 0.0107, Time: 0.03s
Iteration 7278, Loss: 0.0130, Time: 0.03s
Iteration 7279, Loss: 0.0090, Time: 0.03s
Iteration 7280, Loss: 0.0092, Time: 0.03s
Iteration 7281, Loss: 0.0082, Time: 0.03s
Iteration 7282, Loss: 0.0128, Time: 0.03s
Iteration 7283, Loss: 0.0091, Time: 0.03s
Iteration 7284, Loss: 0.0096, Time: 0.03s
Iteration 7285, Loss: 0.0074, Time: 0.03s
Iteration 7286, Loss: 0.0056, Time: 0.03s
Iteration 7287, Loss: 0.0074, Time: 0.03s
Iteration 7288, Loss: 0.0120, Time: 0.03s
Iteration 7289, Loss: 0.0120, Time: 0.03s
Iteration 7290, Loss: 0.0117, Time: 0.03s
Iteration 7291, Loss: 0.0112, Time: 0.03s
Iteration 7292, Loss: 0.0100, Time: 0.03s
Iteration 7293, Loss: 0.0107, Time: 0.03s
Iteration 7294, Loss: 0.0083, Time: 0.03s
```

```
Iteration 7295, Loss: 0.0094, Time: 0.03s
Iteration 7296, Loss: 0.0140, Time: 0.03s
Iteration 7297, Loss: 0.0097, Time: 0.03s
Iteration 7298, Loss: 0.0112, Time: 0.03s
Iteration 7299, Loss: 0.0175, Time: 0.03s
Iteration 7300, Loss: 0.0084, Time: 0.03s
Iteration 7300, Loss: 0.0084, Time: 0.03s
Test Loss: 0.0268
Iteration 7301, Loss: 0.0110, Time: 0.03s
Iteration 7302, Loss: 0.0103, Time: 0.03s
Iteration 7303, Loss: 0.0148, Time: 0.03s
Iteration 7304, Loss: 0.0117, Time: 0.03s
Iteration 7305, Loss: 0.0111, Time: 0.03s
Iteration 7306, Loss: 0.0128, Time: 0.03s
Iteration 7307, Loss: 0.0111, Time: 0.03s
Iteration 7308, Loss: 0.0066, Time: 0.03s
Iteration 7309, Loss: 0.0099, Time: 0.03s
Iteration 7310, Loss: 0.0097, Time: 0.03s
Iteration 7311, Loss: 0.0081, Time: 0.03s
Iteration 7312, Loss: 0.0070, Time: 0.03s
Iteration 7313, Loss: 0.0083, Time: 0.03s
Iteration 7314, Loss: 0.0170, Time: 0.03s
Iteration 7315, Loss: 0.0062, Time: 0.03s
Iteration 7316, Loss: 0.0137, Time: 0.03s
Iteration 7317, Loss: 0.0096, Time: 0.03s
Iteration 7318, Loss: 0.0093, Time: 0.03s
Iteration 7319, Loss: 0.0195, Time: 0.03s
Iteration 7320, Loss: 0.0082, Time: 0.03s
Iteration 7321, Loss: 0.0142, Time: 0.03s
Iteration 7322, Loss: 0.0093, Time: 0.03s
Iteration 7323, Loss: 0.0071, Time: 0.03s
Iteration 7324, Loss: 0.0084, Time: 0.03s
Iteration 7325, Loss: 0.0103, Time: 0.03s
Iteration 7326, Loss: 0.0112, Time: 0.03s
Iteration 7327, Loss: 0.0095, Time: 0.03s
Iteration 7328, Loss: 0.0126, Time: 0.03s
Iteration 7329, Loss: 0.0187, Time: 0.03s
Iteration 7330, Loss: 0.0110, Time: 0.03s
Iteration 7331, Loss: 0.0105, Time: 0.03s
Iteration 7332, Loss: 0.0111, Time: 0.03s
Iteration 7333, Loss: 0.0091, Time: 0.03s
Iteration 7334, Loss: 0.0112, Time: 0.03s
Iteration 7335, Loss: 0.0097, Time: 0.03s
Iteration 7336, Loss: 0.0109, Time: 0.03s
Iteration 7337, Loss: 0.0125, Time: 0.03s
Iteration 7338, Loss: 0.0134, Time: 0.03s
Iteration 7339, Loss: 0.0080, Time: 0.03s
Iteration 7340, Loss: 0.0147, Time: 0.03s
```

```
Iteration 7341, Loss: 0.0101, Time: 0.03s
Iteration 7342, Loss: 0.0107, Time: 0.03s
Iteration 7343, Loss: 0.0124, Time: 0.03s
Iteration 7344, Loss: 0.0132, Time: 0.03s
Iteration 7345, Loss: 0.0116, Time: 0.03s
Iteration 7346, Loss: 0.0064, Time: 0.03s
Iteration 7347, Loss: 0.0093, Time: 0.03s
Iteration 7348, Loss: 0.0130, Time: 0.03s
Iteration 7349, Loss: 0.0151, Time: 0.03s
Iteration 7350, Loss: 0.0124, Time: 0.03s
Iteration 7351, Loss: 0.0082, Time: 0.03s
Iteration 7352, Loss: 0.0121, Time: 0.03s
Iteration 7353, Loss: 0.0077, Time: 0.03s
Iteration 7354, Loss: 0.0145, Time: 0.03s
Iteration 7355, Loss: 0.0137, Time: 0.03s
Iteration 7356, Loss: 0.0159, Time: 0.03s
Iteration 7357, Loss: 0.0120, Time: 0.03s
Iteration 7358, Loss: 0.0135, Time: 0.03s
Iteration 7359, Loss: 0.0106, Time: 0.03s
Iteration 7360, Loss: 0.0096, Time: 0.03s
Iteration 7361, Loss: 0.0152, Time: 0.03s
Iteration 7362, Loss: 0.0122, Time: 0.03s
Iteration 7363, Loss: 0.0067, Time: 0.03s
Iteration 7364, Loss: 0.0040, Time: 0.03s
Iteration 7365, Loss: 0.0100, Time: 0.03s
Iteration 7366, Loss: 0.0093, Time: 0.03s
Iteration 7367, Loss: 0.0119, Time: 0.03s
Iteration 7368, Loss: 0.0075, Time: 0.03s
Iteration 7369, Loss: 0.0097, Time: 0.03s
Iteration 7370, Loss: 0.0132, Time: 0.03s
Iteration 7371, Loss: 0.0106, Time: 0.03s
Iteration 7372, Loss: 0.0101, Time: 0.03s
Iteration 7373, Loss: 0.0096, Time: 0.03s
Iteration 7374, Loss: 0.0123, Time: 0.03s
Iteration 7375, Loss: 0.0085, Time: 0.03s
Iteration 7376, Loss: 0.0097, Time: 0.03s
Iteration 7377, Loss: 0.0077, Time: 0.03s
Iteration 7378, Loss: 0.0073, Time: 0.03s
Iteration 7379, Loss: 0.0059, Time: 0.03s
Iteration 7380, Loss: 0.0055, Time: 0.03s
Iteration 7381, Loss: 0.0090, Time: 0.03s
Iteration 7382, Loss: 0.0159, Time: 0.03s
Iteration 7383, Loss: 0.0092, Time: 0.03s
Iteration 7384, Loss: 0.0133, Time: 0.03s
Iteration 7385, Loss: 0.0103, Time: 0.03s
Iteration 7386, Loss: 0.0091, Time: 0.03s
Iteration 7387, Loss: 0.0116, Time: 0.03s
Iteration 7388, Loss: 0.0133, Time: 0.03s
```

```
Iteration 7389, Loss: 0.0139, Time: 0.03s
Iteration 7390, Loss: 0.0125, Time: 0.03s
Iteration 7391, Loss: 0.0118, Time: 0.03s
Iteration 7392, Loss: 0.0096, Time: 0.03s
Iteration 7393, Loss: 0.0102, Time: 0.03s
Iteration 7394, Loss: 0.0113, Time: 0.03s
Iteration 7395, Loss: 0.0104, Time: 0.03s
Iteration 7396, Loss: 0.0125, Time: 0.03s
Iteration 7397, Loss: 0.0105, Time: 0.03s
Iteration 7398, Loss: 0.0124, Time: 0.03s
Iteration 7399, Loss: 0.0090, Time: 0.03s
Iteration 7400, Loss: 0.0134, Time: 0.03s
Iteration 7400, Loss: 0.0134, Time: 0.03s
Test Loss: 0.0235
Iteration 7401, Loss: 0.0082, Time: 0.03s
Iteration 7402, Loss: 0.0120, Time: 0.03s
Iteration 7403, Loss: 0.0137, Time: 0.03s
Iteration 7404, Loss: 0.0127, Time: 0.03s
Iteration 7405, Loss: 0.0069, Time: 0.03s
Iteration 7406, Loss: 0.0085, Time: 0.03s
Iteration 7407, Loss: 0.0128, Time: 0.03s
Iteration 7408, Loss: 0.0138, Time: 0.03s
Iteration 7409, Loss: 0.0084, Time: 0.03s
Iteration 7410, Loss: 0.0127, Time: 0.03s
Iteration 7411, Loss: 0.0097, Time: 0.03s
Iteration 7412, Loss: 0.0104, Time: 0.03s
Iteration 7413, Loss: 0.0079, Time: 0.03s
Iteration 7414, Loss: 0.0134, Time: 0.03s
Iteration 7415, Loss: 0.0061, Time: 0.03s
Iteration 7416, Loss: 0.0125, Time: 0.03s
Iteration 7417, Loss: 0.0086, Time: 0.03s
Iteration 7418, Loss: 0.0165, Time: 0.03s
Iteration 7419, Loss: 0.0093, Time: 0.03s
Iteration 7420, Loss: 0.0090, Time: 0.03s
Iteration 7421, Loss: 0.0070, Time: 0.03s
Iteration 7422, Loss: 0.0105, Time: 0.03s
Iteration 7423, Loss: 0.0115, Time: 0.03s
Iteration 7424, Loss: 0.0107, Time: 0.03s
Iteration 7425, Loss: 0.0052, Time: 0.03s
Iteration 7426, Loss: 0.0111, Time: 0.03s
Iteration 7427, Loss: 0.0054, Time: 0.03s
Iteration 7428, Loss: 0.0124, Time: 0.03s
Iteration 7429, Loss: 0.0138, Time: 0.03s
Iteration 7430, Loss: 0.0092, Time: 0.03s
Iteration 7431, Loss: 0.0133, Time: 0.03s
Iteration 7432, Loss: 0.0078, Time: 0.03s
Iteration 7433, Loss: 0.0082, Time: 0.03s
Iteration 7434, Loss: 0.0127, Time: 0.03s
```

```
Iteration 7435, Loss: 0.0090, Time: 0.03s
Iteration 7436, Loss: 0.0151, Time: 0.03s
Iteration 7437, Loss: 0.0091, Time: 0.03s
Iteration 7438, Loss: 0.0097, Time: 0.03s
Iteration 7439, Loss: 0.0129, Time: 0.03s
Iteration 7440, Loss: 0.0100, Time: 0.03s
Iteration 7441, Loss: 0.0088, Time: 0.03s
Iteration 7442, Loss: 0.0141, Time: 0.03s
Iteration 7443, Loss: 0.0133, Time: 0.03s
Iteration 7444, Loss: 0.0123, Time: 0.03s
Iteration 7445, Loss: 0.0105, Time: 0.03s
Iteration 7446, Loss: 0.0094, Time: 0.03s
Iteration 7447, Loss: 0.0152, Time: 0.03s
Iteration 7448, Loss: 0.0099, Time: 0.03s
Iteration 7449, Loss: 0.0107, Time: 0.03s
Iteration 7450, Loss: 0.0109, Time: 0.03s
Iteration 7451, Loss: 0.0123, Time: 0.03s
Iteration 7452, Loss: 0.0159, Time: 0.03s
Iteration 7453, Loss: 0.0128, Time: 0.03s
Iteration 7454, Loss: 0.0138, Time: 0.03s
Iteration 7455, Loss: 0.0094, Time: 0.03s
Iteration 7456, Loss: 0.0154, Time: 0.03s
Iteration 7457, Loss: 0.0120, Time: 0.03s
Iteration 7458, Loss: 0.0124, Time: 0.03s
Iteration 7459, Loss: 0.0112, Time: 0.03s
Iteration 7460, Loss: 0.0136, Time: 0.03s
Iteration 7461, Loss: 0.0117, Time: 0.03s
Iteration 7462, Loss: 0.0066, Time: 0.03s
Iteration 7463, Loss: 0.0110, Time: 0.03s
Iteration 7464, Loss: 0.0190, Time: 0.03s
Iteration 7465, Loss: 0.0133, Time: 0.03s
Iteration 7466, Loss: 0.0098, Time: 0.03s
Iteration 7467, Loss: 0.0110, Time: 0.03s
Iteration 7468, Loss: 0.0121, Time: 0.03s
Iteration 7469, Loss: 0.0063, Time: 0.03s
Iteration 7470, Loss: 0.0175, Time: 0.03s
Iteration 7471, Loss: 0.0135, Time: 0.03s
Iteration 7472, Loss: 0.0117, Time: 0.03s
Iteration 7473, Loss: 0.0151, Time: 0.03s
Iteration 7474, Loss: 0.0106, Time: 0.03s
Iteration 7475, Loss: 0.0146, Time: 0.03s
Iteration 7476, Loss: 0.0096, Time: 0.03s
Iteration 7477, Loss: 0.0109, Time: 0.03s
Iteration 7478, Loss: 0.0080, Time: 0.03s
Iteration 7479, Loss: 0.0096, Time: 0.03s
Iteration 7480, Loss: 0.0091, Time: 0.03s
Iteration 7481, Loss: 0.0100, Time: 0.03s
Iteration 7482, Loss: 0.0130, Time: 0.03s
```

```
Iteration 7483, Loss: 0.0082, Time: 0.03s
Iteration 7484, Loss: 0.0109, Time: 0.03s
Iteration 7485, Loss: 0.0116, Time: 0.03s
Iteration 7486, Loss: 0.0067, Time: 0.03s
Iteration 7487, Loss: 0.0112, Time: 0.03s
Iteration 7488, Loss: 0.0119, Time: 0.03s
Iteration 7489, Loss: 0.0084, Time: 0.03s
Iteration 7490, Loss: 0.0128, Time: 0.03s
Iteration 7491, Loss: 0.0117, Time: 0.03s
Iteration 7492, Loss: 0.0082, Time: 0.03s
Iteration 7493, Loss: 0.0104, Time: 0.03s
Iteration 7494, Loss: 0.0102, Time: 0.03s
Iteration 7495, Loss: 0.0100, Time: 0.03s
Iteration 7496, Loss: 0.0131, Time: 0.03s
Iteration 7497, Loss: 0.0107, Time: 0.03s
Iteration 7498, Loss: 0.0108, Time: 0.03s
Iteration 7499, Loss: 0.0114, Time: 0.03s
Iteration 7500, Loss: 0.0099, Time: 0.03s
Iteration 7500, Loss: 0.0099, Time: 0.03s
Test Loss: 0.0492
Iteration 7501, Loss: 0.0100, Time: 0.03s
Iteration 7502, Loss: 0.0142, Time: 0.03s
Iteration 7503, Loss: 0.0116, Time: 0.03s
Iteration 7504, Loss: 0.0085, Time: 0.03s
Iteration 7505, Loss: 0.0078, Time: 0.03s
Iteration 7506, Loss: 0.0146, Time: 0.03s
Iteration 7507, Loss: 0.0111, Time: 0.03s
Iteration 7508, Loss: 0.0124, Time: 0.03s
Iteration 7509, Loss: 0.0141, Time: 0.03s
Iteration 7510, Loss: 0.0142, Time: 0.03s
Iteration 7511, Loss: 0.0076, Time: 0.03s
Iteration 7512, Loss: 0.0217, Time: 0.03s
Iteration 7513, Loss: 0.0077, Time: 0.03s
Iteration 7514, Loss: 0.0096, Time: 0.03s
Iteration 7515, Loss: 0.0091, Time: 0.03s
Iteration 7516, Loss: 0.0098, Time: 0.03s
Iteration 7517, Loss: 0.0096, Time: 0.03s
Iteration 7518, Loss: 0.0161, Time: 0.03s
Iteration 7519, Loss: 0.0104, Time: 0.03s
Iteration 7520, Loss: 0.0109, Time: 0.03s
Iteration 7521, Loss: 0.0120, Time: 0.03s
Iteration 7522, Loss: 0.0052, Time: 0.03s
Iteration 7523, Loss: 0.0128, Time: 0.03s
Iteration 7524, Loss: 0.0069, Time: 0.03s
Iteration 7525, Loss: 0.0170, Time: 0.03s
Iteration 7526, Loss: 0.0094, Time: 0.03s
Iteration 7527, Loss: 0.0073, Time: 0.03s
Iteration 7528, Loss: 0.0104, Time: 0.03s
```

```
Iteration 7529, Loss: 0.0124, Time: 0.03s
Iteration 7530, Loss: 0.0115, Time: 0.03s
Iteration 7531, Loss: 0.0163, Time: 0.03s
Iteration 7532, Loss: 0.0136, Time: 0.03s
Iteration 7533, Loss: 0.0088, Time: 0.03s
Iteration 7534, Loss: 0.0092, Time: 0.03s
Iteration 7535, Loss: 0.0087, Time: 0.03s
Iteration 7536, Loss: 0.0077, Time: 0.03s
Iteration 7537, Loss: 0.0085, Time: 0.03s
Iteration 7538, Loss: 0.0147, Time: 0.03s
Iteration 7539, Loss: 0.0110, Time: 0.03s
Iteration 7540, Loss: 0.0135, Time: 0.03s
Iteration 7541, Loss: 0.0114, Time: 0.03s
Iteration 7542, Loss: 0.0099, Time: 0.03s
Iteration 7543, Loss: 0.0072, Time: 0.03s
Iteration 7544, Loss: 0.0107, Time: 0.03s
Iteration 7545, Loss: 0.0105, Time: 0.03s
Iteration 7546, Loss: 0.0098, Time: 0.03s
Iteration 7547, Loss: 0.0091, Time: 0.03s
Iteration 7548, Loss: 0.0098, Time: 0.03s
Iteration 7549, Loss: 0.0150, Time: 0.03s
Iteration 7550, Loss: 0.0104, Time: 0.03s
Iteration 7551, Loss: 0.0149, Time: 0.03s
Iteration 7552, Loss: 0.0100, Time: 0.03s
Iteration 7553, Loss: 0.0129, Time: 0.03s
Iteration 7554, Loss: 0.0127, Time: 0.03s
Iteration 7555, Loss: 0.0078, Time: 0.03s
Iteration 7556, Loss: 0.0120, Time: 0.03s
Iteration 7557, Loss: 0.0113, Time: 0.03s
Iteration 7558, Loss: 0.0156, Time: 0.03s
Iteration 7559, Loss: 0.0104, Time: 0.03s
Iteration 7560, Loss: 0.0084, Time: 0.03s
Iteration 7561, Loss: 0.0185, Time: 0.03s
Iteration 7562, Loss: 0.0096, Time: 0.03s
Iteration 7563, Loss: 0.0105, Time: 0.03s
Iteration 7564, Loss: 0.0103, Time: 0.03s
Iteration 7565, Loss: 0.0095, Time: 0.03s
Iteration 7566, Loss: 0.0107, Time: 0.03s
Iteration 7567, Loss: 0.0120, Time: 0.03s
Iteration 7568, Loss: 0.0098, Time: 0.03s
Iteration 7569, Loss: 0.0105, Time: 0.03s
Iteration 7570, Loss: 0.0112, Time: 0.03s
Iteration 7571, Loss: 0.0146, Time: 0.03s
Iteration 7572, Loss: 0.0129, Time: 0.03s
Iteration 7573, Loss: 0.0144, Time: 0.03s
Iteration 7574, Loss: 0.0098, Time: 0.03s
Iteration 7575, Loss: 0.0115, Time: 0.03s
Iteration 7576, Loss: 0.0094, Time: 0.03s
```

```
Iteration 7577, Loss: 0.0071, Time: 0.03s
Iteration 7578, Loss: 0.0121, Time: 0.03s
Iteration 7579, Loss: 0.0082, Time: 0.03s
Iteration 7580, Loss: 0.0120, Time: 0.03s
Iteration 7581, Loss: 0.0128, Time: 0.03s
Iteration 7582, Loss: 0.0064, Time: 0.03s
Iteration 7583, Loss: 0.0133, Time: 0.03s
Iteration 7584, Loss: 0.0124, Time: 0.03s
Iteration 7585, Loss: 0.0116, Time: 0.03s
Iteration 7586, Loss: 0.0099, Time: 0.03s
Iteration 7587, Loss: 0.0112, Time: 0.03s
Iteration 7588, Loss: 0.0085, Time: 0.03s
Iteration 7589, Loss: 0.0085, Time: 0.03s
Iteration 7590, Loss: 0.0063, Time: 0.03s
Iteration 7591, Loss: 0.0109, Time: 0.03s
Iteration 7592, Loss: 0.0105, Time: 0.03s
Iteration 7593, Loss: 0.0113, Time: 0.03s
Iteration 7594, Loss: 0.0121, Time: 0.03s
Iteration 7595, Loss: 0.0133, Time: 0.03s
Iteration 7596, Loss: 0.0070, Time: 0.03s
Iteration 7597, Loss: 0.0150, Time: 0.03s
Iteration 7598, Loss: 0.0122, Time: 0.03s
Iteration 7599, Loss: 0.0093, Time: 0.03s
Iteration 7600, Loss: 0.0121, Time: 0.03s
Iteration 7600, Loss: 0.0121, Time: 0.03s
Test Loss: 0.0303
Iteration 7601, Loss: 0.0095, Time: 0.03s
Iteration 7602, Loss: 0.0097, Time: 0.03s
Iteration 7603, Loss: 0.0097, Time: 0.03s
Iteration 7604, Loss: 0.0089, Time: 0.03s
Iteration 7605, Loss: 0.0124, Time: 0.03s
Iteration 7606, Loss: 0.0126, Time: 0.03s
Iteration 7607, Loss: 0.0101, Time: 0.03s
Iteration 7608, Loss: 0.0085, Time: 0.03s
Iteration 7609, Loss: 0.0097, Time: 0.03s
Iteration 7610, Loss: 0.0158, Time: 0.03s
Iteration 7611, Loss: 0.0141, Time: 0.03s
Iteration 7612, Loss: 0.0102, Time: 0.03s
Iteration 7613, Loss: 0.0132, Time: 0.03s
Iteration 7614, Loss: 0.0131, Time: 0.03s
Iteration 7615, Loss: 0.0096, Time: 0.03s
Iteration 7616, Loss: 0.0131, Time: 0.03s
Iteration 7617, Loss: 0.0148, Time: 0.03s
Iteration 7618, Loss: 0.0086, Time: 0.03s
Iteration 7619, Loss: 0.0079, Time: 0.03s
Iteration 7620, Loss: 0.0142, Time: 0.03s
Iteration 7621, Loss: 0.0098, Time: 0.03s
Iteration 7622, Loss: 0.0090, Time: 0.03s
```

```
Iteration 7623, Loss: 0.0110, Time: 0.03s
Iteration 7624, Loss: 0.0083, Time: 0.03s
Iteration 7625, Loss: 0.0099, Time: 0.03s
Iteration 7626, Loss: 0.0108, Time: 0.03s
Iteration 7627, Loss: 0.0074, Time: 0.03s
Iteration 7628, Loss: 0.0133, Time: 0.03s
Iteration 7629, Loss: 0.0080, Time: 0.03s
Iteration 7630, Loss: 0.0098, Time: 0.03s
Iteration 7631, Loss: 0.0111, Time: 0.03s
Iteration 7632, Loss: 0.0117, Time: 0.03s
Iteration 7633, Loss: 0.0090, Time: 0.03s
Iteration 7634, Loss: 0.0072, Time: 0.03s
Iteration 7635, Loss: 0.0074, Time: 0.03s
Iteration 7636, Loss: 0.0107, Time: 0.03s
Iteration 7637, Loss: 0.0093, Time: 0.03s
Iteration 7638, Loss: 0.0124, Time: 0.03s
Iteration 7639, Loss: 0.0144, Time: 0.03s
Iteration 7640, Loss: 0.0089, Time: 0.03s
Iteration 7641, Loss: 0.0090, Time: 0.03s
Iteration 7642, Loss: 0.0101, Time: 0.03s
Iteration 7643, Loss: 0.0105, Time: 0.03s
Iteration 7644, Loss: 0.0124, Time: 0.03s
Iteration 7645, Loss: 0.0059, Time: 0.03s
Iteration 7646, Loss: 0.0113, Time: 0.03s
Iteration 7647, Loss: 0.0114, Time: 0.03s
Iteration 7648, Loss: 0.0100, Time: 0.03s
Iteration 7649, Loss: 0.0083, Time: 0.03s
Iteration 7650, Loss: 0.0085, Time: 0.03s
Iteration 7651, Loss: 0.0104, Time: 0.03s
Iteration 7652, Loss: 0.0116, Time: 0.03s
Iteration 7653, Loss: 0.0121, Time: 0.03s
Iteration 7654, Loss: 0.0124, Time: 0.03s
Iteration 7655, Loss: 0.0073, Time: 0.03s
Iteration 7656, Loss: 0.0085, Time: 0.03s
Iteration 7657, Loss: 0.0051, Time: 0.03s
Iteration 7658, Loss: 0.0117, Time: 0.03s
Iteration 7659, Loss: 0.0098, Time: 0.03s
Iteration 7660, Loss: 0.0153, Time: 0.03s
Iteration 7661, Loss: 0.0126, Time: 0.03s
Iteration 7662, Loss: 0.0138, Time: 0.03s
Iteration 7663, Loss: 0.0075, Time: 0.03s
Iteration 7664, Loss: 0.0091, Time: 0.03s
Iteration 7665, Loss: 0.0079, Time: 0.03s
Iteration 7666, Loss: 0.0134, Time: 0.03s
Iteration 7667, Loss: 0.0092, Time: 0.03s
Iteration 7668, Loss: 0.0134, Time: 0.03s
Iteration 7669, Loss: 0.0058, Time: 0.03s
Iteration 7670, Loss: 0.0124, Time: 0.03s
```

```
Iteration 7671, Loss: 0.0097, Time: 0.03s
Iteration 7672, Loss: 0.0144, Time: 0.03s
Iteration 7673, Loss: 0.0126, Time: 0.03s
Iteration 7674, Loss: 0.0077, Time: 0.03s
Iteration 7675, Loss: 0.0143, Time: 0.03s
Iteration 7676, Loss: 0.0131, Time: 0.03s
Iteration 7677, Loss: 0.0134, Time: 0.03s
Iteration 7678, Loss: 0.0130, Time: 0.03s
Iteration 7679, Loss: 0.0121, Time: 0.03s
Iteration 7680, Loss: 0.0141, Time: 0.03s
Iteration 7681, Loss: 0.0096, Time: 0.03s
Iteration 7682, Loss: 0.0090, Time: 0.03s
Iteration 7683, Loss: 0.0106, Time: 0.03s
Iteration 7684, Loss: 0.0083, Time: 0.03s
Iteration 7685, Loss: 0.0090, Time: 0.03s
Iteration 7686, Loss: 0.0120, Time: 0.03s
Iteration 7687, Loss: 0.0103, Time: 0.03s
Iteration 7688, Loss: 0.0105, Time: 0.03s
Iteration 7689, Loss: 0.0158, Time: 0.03s
Iteration 7690, Loss: 0.0074, Time: 0.03s
Iteration 7691, Loss: 0.0090, Time: 0.03s
Iteration 7692, Loss: 0.0083, Time: 0.03s
Iteration 7693, Loss: 0.0147, Time: 0.03s
Iteration 7694, Loss: 0.0116, Time: 0.03s
Iteration 7695, Loss: 0.0102, Time: 0.03s
Iteration 7696, Loss: 0.0129, Time: 0.03s
Iteration 7697, Loss: 0.0150, Time: 0.03s
Iteration 7698, Loss: 0.0059, Time: 0.03s
Iteration 7699, Loss: 0.0127, Time: 0.03s
Iteration 7700, Loss: 0.0110, Time: 0.03s
Iteration 7700, Loss: 0.0110, Time: 0.03s
Test Loss: 0.0500
Iteration 7701, Loss: 0.0123, Time: 0.03s
Iteration 7702, Loss: 0.0084, Time: 0.03s
Iteration 7703, Loss: 0.0085, Time: 0.03s
Iteration 7704, Loss: 0.0064, Time: 0.03s
Iteration 7705, Loss: 0.0100, Time: 0.03s
Iteration 7706, Loss: 0.0088, Time: 0.03s
Iteration 7707, Loss: 0.0110, Time: 0.03s
Iteration 7708, Loss: 0.0086, Time: 0.03s
Iteration 7709, Loss: 0.0099, Time: 0.03s
Iteration 7710, Loss: 0.0114, Time: 0.03s
Iteration 7711, Loss: 0.0097, Time: 0.03s
Iteration 7712, Loss: 0.0157, Time: 0.03s
Iteration 7713, Loss: 0.0097, Time: 0.03s
Iteration 7714, Loss: 0.0122, Time: 0.03s
Iteration 7715, Loss: 0.0084, Time: 0.03s
Iteration 7716, Loss: 0.0075, Time: 0.03s
```

```
Iteration 7717, Loss: 0.0133, Time: 0.03s
Iteration 7718, Loss: 0.0082, Time: 0.03s
Iteration 7719, Loss: 0.0163, Time: 0.03s
Iteration 7720, Loss: 0.0070, Time: 0.03s
Iteration 7721, Loss: 0.0095, Time: 0.03s
Iteration 7722, Loss: 0.0071, Time: 0.03s
Iteration 7723, Loss: 0.0109, Time: 0.03s
Iteration 7724, Loss: 0.0098, Time: 0.03s
Iteration 7725, Loss: 0.0084, Time: 0.03s
Iteration 7726, Loss: 0.0113, Time: 0.03s
Iteration 7727, Loss: 0.0133, Time: 0.03s
Iteration 7728, Loss: 0.0117, Time: 0.03s
Iteration 7729, Loss: 0.0120, Time: 0.03s
Iteration 7730, Loss: 0.0097, Time: 0.03s
Iteration 7731, Loss: 0.0075, Time: 0.03s
Iteration 7732, Loss: 0.0091, Time: 0.03s
Iteration 7733, Loss: 0.0093, Time: 0.04s
Iteration 7734, Loss: 0.0108, Time: 0.04s
Iteration 7735, Loss: 0.0147, Time: 0.03s
Iteration 7736, Loss: 0.0094, Time: 0.04s
Iteration 7737, Loss: 0.0154, Time: 0.04s
Iteration 7738, Loss: 0.0127, Time: 0.04s
Iteration 7739, Loss: 0.0099, Time: 0.04s
Iteration 7740, Loss: 0.0156, Time: 0.03s
Iteration 7741, Loss: 0.0092, Time: 0.04s
Iteration 7742, Loss: 0.0099, Time: 0.04s
Iteration 7743, Loss: 0.0077, Time: 0.04s
Iteration 7744, Loss: 0.0124, Time: 0.03s
Iteration 7745, Loss: 0.0125, Time: 0.03s
Iteration 7746, Loss: 0.0118, Time: 0.03s
Iteration 7747, Loss: 0.0160, Time: 0.03s
Iteration 7748, Loss: 0.0090, Time: 0.03s
Iteration 7749, Loss: 0.0087, Time: 0.03s
Iteration 7750, Loss: 0.0109, Time: 0.03s
Iteration 7751, Loss: 0.0068, Time: 0.03s
Iteration 7752, Loss: 0.0156, Time: 0.03s
Iteration 7753, Loss: 0.0109, Time: 0.03s
Iteration 7754, Loss: 0.0092, Time: 0.03s
Iteration 7755, Loss: 0.0086, Time: 0.03s
Iteration 7756, Loss: 0.0087, Time: 0.03s
Iteration 7757, Loss: 0.0108, Time: 0.03s
Iteration 7758, Loss: 0.0108, Time: 0.03s
Iteration 7759, Loss: 0.0129, Time: 0.03s
Iteration 7760, Loss: 0.0103, Time: 0.03s
Iteration 7761, Loss: 0.0123, Time: 0.03s
Iteration 7762, Loss: 0.0105, Time: 0.03s
Iteration 7763, Loss: 0.0088, Time: 0.03s
Iteration 7764, Loss: 0.0082, Time: 0.03s
```

```
Iteration 7765, Loss: 0.0115, Time: 0.03s
Iteration 7766, Loss: 0.0094, Time: 0.03s
Iteration 7767, Loss: 0.0091, Time: 0.03s
Iteration 7768, Loss: 0.0145, Time: 0.03s
Iteration 7769, Loss: 0.0115, Time: 0.03s
Iteration 7770, Loss: 0.0139, Time: 0.03s
Iteration 7771, Loss: 0.0093, Time: 0.03s
Iteration 7772, Loss: 0.0121, Time: 0.03s
Iteration 7773, Loss: 0.0093, Time: 0.03s
Iteration 7774, Loss: 0.0128, Time: 0.03s
Iteration 7775, Loss: 0.0144, Time: 0.03s
Iteration 7776, Loss: 0.0113, Time: 0.03s
Iteration 7777, Loss: 0.0114, Time: 0.03s
Iteration 7778, Loss: 0.0058, Time: 0.04s
Iteration 7779, Loss: 0.0136, Time: 0.03s
Iteration 7780, Loss: 0.0125, Time: 0.03s
Iteration 7781, Loss: 0.0105, Time: 0.03s
Iteration 7782, Loss: 0.0081, Time: 0.03s
Iteration 7783, Loss: 0.0065, Time: 0.03s
Iteration 7784, Loss: 0.0080, Time: 0.03s
Iteration 7785, Loss: 0.0100, Time: 0.03s
Iteration 7786, Loss: 0.0107, Time: 0.03s
Iteration 7787, Loss: 0.0100, Time: 0.03s
Iteration 7788, Loss: 0.0100, Time: 0.03s
Iteration 7789, Loss: 0.0113, Time: 0.03s
Iteration 7790, Loss: 0.0092, Time: 0.03s
Iteration 7791, Loss: 0.0105, Time: 0.03s
Iteration 7792, Loss: 0.0076, Time: 0.03s
Iteration 7793, Loss: 0.0084, Time: 0.03s
Iteration 7794, Loss: 0.0080, Time: 0.03s
Iteration 7795, Loss: 0.0103, Time: 0.03s
Iteration 7796, Loss: 0.0162, Time: 0.03s
Iteration 7797, Loss: 0.0074, Time: 0.03s
Iteration 7798, Loss: 0.0064, Time: 0.03s
Iteration 7799, Loss: 0.0083, Time: 0.03s
Iteration 7800, Loss: 0.0116, Time: 0.04s
Iteration 7800, Loss: 0.0116, Time: 0.04s
Test Loss: 0.0377
Iteration 7801, Loss: 0.0094, Time: 0.03s
Iteration 7802, Loss: 0.0130, Time: 0.03s
Iteration 7803, Loss: 0.0106, Time: 0.03s
Iteration 7804, Loss: 0.0137, Time: 0.03s
Iteration 7805, Loss: 0.0109, Time: 0.03s
Iteration 7806, Loss: 0.0132, Time: 0.03s
Iteration 7807, Loss: 0.0099, Time: 0.03s
Iteration 7808, Loss: 0.0104, Time: 0.03s
Iteration 7809, Loss: 0.0088, Time: 0.03s
Iteration 7810, Loss: 0.0109, Time: 0.03s
```

```
Iteration 7811, Loss: 0.0102, Time: 0.03s
Iteration 7812, Loss: 0.0117, Time: 0.03s
Iteration 7813, Loss: 0.0119, Time: 0.03s
Iteration 7814, Loss: 0.0089, Time: 0.03s
Iteration 7815, Loss: 0.0085, Time: 0.03s
Iteration 7816, Loss: 0.0133, Time: 0.03s
Iteration 7817, Loss: 0.0140, Time: 0.03s
Iteration 7818, Loss: 0.0147, Time: 0.03s
Iteration 7819, Loss: 0.0054, Time: 0.03s
Iteration 7820, Loss: 0.0084, Time: 0.03s
Iteration 7821, Loss: 0.0090, Time: 0.03s
Iteration 7822, Loss: 0.0080, Time: 0.03s
Iteration 7823, Loss: 0.0091, Time: 0.03s
Iteration 7824, Loss: 0.0069, Time: 0.03s
Iteration 7825, Loss: 0.0068, Time: 0.03s
Iteration 7826, Loss: 0.0147, Time: 0.04s
Iteration 7827, Loss: 0.0106, Time: 0.03s
Iteration 7828, Loss: 0.0092, Time: 0.03s
Iteration 7829, Loss: 0.0137, Time: 0.03s
Iteration 7830, Loss: 0.0083, Time: 0.03s
Iteration 7831, Loss: 0.0124, Time: 0.03s
Iteration 7832, Loss: 0.0076, Time: 0.03s
Iteration 7833, Loss: 0.0103, Time: 0.03s
Iteration 7834, Loss: 0.0091, Time: 0.03s
Iteration 7835, Loss: 0.0123, Time: 0.03s
Iteration 7836, Loss: 0.0125, Time: 0.03s
Iteration 7837, Loss: 0.0124, Time: 0.03s
Iteration 7838, Loss: 0.0087, Time: 0.03s
Iteration 7839, Loss: 0.0109, Time: 0.03s
Iteration 7840, Loss: 0.0099, Time: 0.03s
Iteration 7841, Loss: 0.0072, Time: 0.03s
Iteration 7842, Loss: 0.0064, Time: 0.03s
Iteration 7843, Loss: 0.0110, Time: 0.03s
Iteration 7844, Loss: 0.0065, Time: 0.03s
Iteration 7845, Loss: 0.0087, Time: 0.03s
Iteration 7846, Loss: 0.0122, Time: 0.03s
Iteration 7847, Loss: 0.0174, Time: 0.03s
Iteration 7848, Loss: 0.0094, Time: 0.03s
Iteration 7849, Loss: 0.0103, Time: 0.03s
Iteration 7850, Loss: 0.0092, Time: 0.03s
Iteration 7851, Loss: 0.0095, Time: 0.03s
Iteration 7852, Loss: 0.0052, Time: 0.03s
Iteration 7853, Loss: 0.0067, Time: 0.03s
Iteration 7854, Loss: 0.0085, Time: 0.03s
Iteration 7855, Loss: 0.0121, Time: 0.03s
Iteration 7856, Loss: 0.0110, Time: 0.03s
Iteration 7857, Loss: 0.0087, Time: 0.03s
Iteration 7858, Loss: 0.0064, Time: 0.03s
```

```
Iteration 7859, Loss: 0.0094, Time: 0.03s
Iteration 7860, Loss: 0.0053, Time: 0.03s
Iteration 7861, Loss: 0.0149, Time: 0.03s
Iteration 7862, Loss: 0.0099, Time: 0.03s
Iteration 7863, Loss: 0.0111, Time: 0.03s
Iteration 7864, Loss: 0.0102, Time: 0.03s
Iteration 7865, Loss: 0.0097, Time: 0.03s
Iteration 7866, Loss: 0.0109, Time: 0.03s
Iteration 7867, Loss: 0.0102, Time: 0.03s
Iteration 7868, Loss: 0.0138, Time: 0.03s
Iteration 7869, Loss: 0.0116, Time: 0.03s
Iteration 7870, Loss: 0.0092, Time: 0.03s
Iteration 7871, Loss: 0.0065, Time: 0.03s
Iteration 7872, Loss: 0.0154, Time: 0.03s
Iteration 7873, Loss: 0.0145, Time: 0.03s
Iteration 7874, Loss: 0.0077, Time: 0.03s
Iteration 7875, Loss: 0.0104, Time: 0.03s
Iteration 7876, Loss: 0.0115, Time: 0.03s
Iteration 7877, Loss: 0.0114, Time: 0.03s
Iteration 7878, Loss: 0.0096, Time: 0.04s
Iteration 7879, Loss: 0.0132, Time: 0.03s
Iteration 7880, Loss: 0.0135, Time: 0.03s
Iteration 7881, Loss: 0.0081, Time: 0.03s
Iteration 7882, Loss: 0.0106, Time: 0.03s
Iteration 7883, Loss: 0.0120, Time: 0.03s
Iteration 7884, Loss: 0.0116, Time: 0.03s
Iteration 7885, Loss: 0.0096, Time: 0.03s
Iteration 7886, Loss: 0.0092, Time: 0.03s
Iteration 7887, Loss: 0.0090, Time: 0.03s
Iteration 7888, Loss: 0.0107, Time: 0.03s
Iteration 7889, Loss: 0.0130, Time: 0.03s
Iteration 7890, Loss: 0.0107, Time: 0.03s
Iteration 7891, Loss: 0.0107, Time: 0.03s
Iteration 7892, Loss: 0.0097, Time: 0.03s
Iteration 7893, Loss: 0.0110, Time: 0.03s
Iteration 7894, Loss: 0.0152, Time: 0.03s
Iteration 7895, Loss: 0.0109, Time: 0.03s
Iteration 7896, Loss: 0.0109, Time: 0.03s
Iteration 7897, Loss: 0.0101, Time: 0.03s
Iteration 7898, Loss: 0.0108, Time: 0.03s
Iteration 7899, Loss: 0.0062, Time: 0.03s
Iteration 7900, Loss: 0.0115, Time: 0.03s
Iteration 7900, Loss: 0.0115, Time: 0.03s
Test Loss: 0.0722
Iteration 7901, Loss: 0.0140, Time: 0.03s
Iteration 7902, Loss: 0.0092, Time: 0.03s
Iteration 7903, Loss: 0.0118, Time: 0.03s
Iteration 7904, Loss: 0.0113, Time: 0.03s
```

```
Iteration 7905, Loss: 0.0112, Time: 0.03s
Iteration 7906, Loss: 0.0092, Time: 0.03s
Iteration 7907, Loss: 0.0096, Time: 0.03s
Iteration 7908, Loss: 0.0115, Time: 0.03s
Iteration 7909, Loss: 0.0147, Time: 0.03s
Iteration 7910, Loss: 0.0076, Time: 0.03s
Iteration 7911, Loss: 0.0074, Time: 0.03s
Iteration 7912, Loss: 0.0152, Time: 0.03s
Iteration 7913, Loss: 0.0066, Time: 0.03s
Iteration 7914, Loss: 0.0074, Time: 0.03s
Iteration 7915, Loss: 0.0094, Time: 0.03s
Iteration 7916, Loss: 0.0067, Time: 0.03s
Iteration 7917, Loss: 0.0155, Time: 0.03s
Iteration 7918, Loss: 0.0028, Time: 0.03s
Iteration 7919, Loss: 0.0112, Time: 0.03s
Iteration 7920, Loss: 0.0132, Time: 0.03s
Iteration 7921, Loss: 0.0143, Time: 0.03s
Iteration 7922, Loss: 0.0106, Time: 0.03s
Iteration 7923, Loss: 0.0094, Time: 0.03s
Iteration 7924, Loss: 0.0097, Time: 0.03s
Iteration 7925, Loss: 0.0087, Time: 0.03s
Iteration 7926, Loss: 0.0080, Time: 0.03s
Iteration 7927, Loss: 0.0098, Time: 0.03s
Iteration 7928, Loss: 0.0135, Time: 0.03s
Iteration 7929, Loss: 0.0139, Time: 0.03s
Iteration 7930, Loss: 0.0106, Time: 0.03s
Iteration 7931, Loss: 0.0093, Time: 0.03s
Iteration 7932, Loss: 0.0084, Time: 0.03s
Iteration 7933, Loss: 0.0098, Time: 0.03s
Iteration 7934, Loss: 0.0118, Time: 0.03s
Iteration 7935, Loss: 0.0084, Time: 0.03s
Iteration 7936, Loss: 0.0079, Time: 0.03s
Iteration 7937, Loss: 0.0093, Time: 0.03s
Iteration 7938, Loss: 0.0109, Time: 0.03s
Iteration 7939, Loss: 0.0146, Time: 0.03s
Iteration 7940, Loss: 0.0084, Time: 0.03s
Iteration 7941, Loss: 0.0094, Time: 0.03s
Iteration 7942, Loss: 0.0122, Time: 0.03s
Iteration 7943, Loss: 0.0151, Time: 0.03s
Iteration 7944, Loss: 0.0099, Time: 0.03s
Iteration 7945, Loss: 0.0096, Time: 0.03s
Iteration 7946, Loss: 0.0089, Time: 0.03s
Iteration 7947, Loss: 0.0105, Time: 0.03s
Iteration 7948, Loss: 0.0103, Time: 0.03s
Iteration 7949, Loss: 0.0133, Time: 0.03s
Iteration 7950, Loss: 0.0084, Time: 0.03s
Iteration 7951, Loss: 0.0100, Time: 0.03s
Iteration 7952, Loss: 0.0090, Time: 0.03s
```

```
Iteration 7953, Loss: 0.0096, Time: 0.03s
Iteration 7954, Loss: 0.0168, Time: 0.03s
Iteration 7955, Loss: 0.0115, Time: 0.03s
Iteration 7956, Loss: 0.0140, Time: 0.03s
Iteration 7957, Loss: 0.0101, Time: 0.03s
Iteration 7958, Loss: 0.0132, Time: 0.03s
Iteration 7959, Loss: 0.0051, Time: 0.03s
Iteration 7960, Loss: 0.0080, Time: 0.03s
Iteration 7961, Loss: 0.0139, Time: 0.03s
Iteration 7962, Loss: 0.0112, Time: 0.03s
Iteration 7963, Loss: 0.0117, Time: 0.03s
Iteration 7964, Loss: 0.0120, Time: 0.03s
Iteration 7965, Loss: 0.0151, Time: 0.03s
Iteration 7966, Loss: 0.0084, Time: 0.03s
Iteration 7967, Loss: 0.0084, Time: 0.03s
Iteration 7968, Loss: 0.0137, Time: 0.03s
Iteration 7969, Loss: 0.0080, Time: 0.03s
Iteration 7970, Loss: 0.0110, Time: 0.03s
Iteration 7971, Loss: 0.0094, Time: 0.03s
Iteration 7972, Loss: 0.0075, Time: 0.03s
Iteration 7973, Loss: 0.0069, Time: 0.03s
Iteration 7974, Loss: 0.0130, Time: 0.03s
Iteration 7975, Loss: 0.0064, Time: 0.03s
Iteration 7976, Loss: 0.0098, Time: 0.03s
Iteration 7977, Loss: 0.0186, Time: 0.03s
Iteration 7978, Loss: 0.0108, Time: 0.03s
Iteration 7979, Loss: 0.0086, Time: 0.03s
Iteration 7980, Loss: 0.0151, Time: 0.03s
Iteration 7981, Loss: 0.0073, Time: 0.03s
Iteration 7982, Loss: 0.0047, Time: 0.03s
Iteration 7983, Loss: 0.0104, Time: 0.03s
Iteration 7984, Loss: 0.0074, Time: 0.03s
Iteration 7985, Loss: 0.0115, Time: 0.03s
Iteration 7986, Loss: 0.0105, Time: 0.03s
Iteration 7987, Loss: 0.0116, Time: 0.03s
Iteration 7988, Loss: 0.0123, Time: 0.03s
Iteration 7989, Loss: 0.0107, Time: 0.03s
Iteration 7990, Loss: 0.0152, Time: 0.03s
Iteration 7991, Loss: 0.0102, Time: 0.03s
Iteration 7992, Loss: 0.0089, Time: 0.03s
Iteration 7993, Loss: 0.0104, Time: 0.03s
Iteration 7994, Loss: 0.0095, Time: 0.03s
Iteration 7995, Loss: 0.0118, Time: 0.03s
Iteration 7996, Loss: 0.0080, Time: 0.03s
Iteration 7997, Loss: 0.0129, Time: 0.03s
Iteration 7998, Loss: 0.0076, Time: 0.03s
Iteration 7999, Loss: 0.0122, Time: 0.03s
Iteration 8000, Loss: 0.0114, Time: 0.03s
```

```
Iteration 8000, Loss: 0.0114, Time: 0.03s
Test Loss: 0.0897
Iteration 8001, Loss: 0.0096, Time: 0.03s
Iteration 8002, Loss: 0.0105, Time: 0.03s
Iteration 8003, Loss: 0.0119, Time: 0.03s
Iteration 8004, Loss: 0.0098, Time: 0.03s
Iteration 8005, Loss: 0.0101, Time: 0.03s
Iteration 8006, Loss: 0.0107, Time: 0.03s
Iteration 8007, Loss: 0.0105, Time: 0.03s
Iteration 8008, Loss: 0.0157, Time: 0.03s
Iteration 8009, Loss: 0.0098, Time: 0.03s
Iteration 8010, Loss: 0.0085, Time: 0.03s
Iteration 8011, Loss: 0.0118, Time: 0.03s
Iteration 8012, Loss: 0.0114, Time: 0.03s
Iteration 8013, Loss: 0.0126, Time: 0.03s
Iteration 8014, Loss: 0.0085, Time: 0.03s
Iteration 8015, Loss: 0.0111, Time: 0.03s
Iteration 8016, Loss: 0.0121, Time: 0.03s
Iteration 8017, Loss: 0.0108, Time: 0.03s
Iteration 8018, Loss: 0.0081, Time: 0.03s
Iteration 8019, Loss: 0.0105, Time: 0.03s
Iteration 8020, Loss: 0.0090, Time: 0.03s
Iteration 8021, Loss: 0.0069, Time: 0.03s
Iteration 8022, Loss: 0.0085, Time: 0.03s
Iteration 8023, Loss: 0.0121, Time: 0.03s
Iteration 8024, Loss: 0.0140, Time: 0.03s
Iteration 8025, Loss: 0.0120, Time: 0.03s
Iteration 8026, Loss: 0.0095, Time: 0.03s
Iteration 8027, Loss: 0.0108, Time: 0.03s
Iteration 8028, Loss: 0.0079, Time: 0.03s
Iteration 8029, Loss: 0.0123, Time: 0.03s
Iteration 8030, Loss: 0.0129, Time: 0.03s
Iteration 8031, Loss: 0.0086, Time: 0.03s
Iteration 8032, Loss: 0.0114, Time: 0.03s
Iteration 8033, Loss: 0.0121, Time: 0.03s
Iteration 8034, Loss: 0.0054, Time: 0.03s
Iteration 8035, Loss: 0.0169, Time: 0.03s
Iteration 8036, Loss: 0.0125, Time: 0.03s
Iteration 8037, Loss: 0.0100, Time: 0.03s
Iteration 8038, Loss: 0.0072, Time: 0.03s
Iteration 8039, Loss: 0.0078, Time: 0.03s
Iteration 8040, Loss: 0.0083, Time: 0.03s
Iteration 8041, Loss: 0.0088, Time: 0.03s
Iteration 8042, Loss: 0.0072, Time: 0.03s
Iteration 8043, Loss: 0.0136, Time: 0.03s
Iteration 8044, Loss: 0.0138, Time: 0.03s
Iteration 8045, Loss: 0.0065, Time: 0.03s
Iteration 8046, Loss: 0.0143, Time: 0.03s
```

```
Iteration 8047, Loss: 0.0156, Time: 0.03s
Iteration 8048, Loss: 0.0060, Time: 0.03s
Iteration 8049, Loss: 0.0108, Time: 0.03s
Iteration 8050, Loss: 0.0114, Time: 0.03s
Iteration 8051, Loss: 0.0078, Time: 0.03s
Iteration 8052, Loss: 0.0103, Time: 0.03s
Iteration 8053, Loss: 0.0113, Time: 0.03s
Iteration 8054, Loss: 0.0112, Time: 0.03s
Iteration 8055, Loss: 0.0125, Time: 0.03s
Iteration 8056, Loss: 0.0121, Time: 0.03s
Iteration 8057, Loss: 0.0060, Time: 0.03s
Iteration 8058, Loss: 0.0134, Time: 0.03s
Iteration 8059, Loss: 0.0100, Time: 0.03s
Iteration 8060, Loss: 0.0132, Time: 0.03s
Iteration 8061, Loss: 0.0150, Time: 0.03s
Iteration 8062, Loss: 0.0118, Time: 0.03s
Iteration 8063, Loss: 0.0126, Time: 0.03s
Iteration 8064, Loss: 0.0104, Time: 0.03s
Iteration 8065, Loss: 0.0071, Time: 0.03s
Iteration 8066, Loss: 0.0098, Time: 0.03s
Iteration 8067, Loss: 0.0116, Time: 0.03s
Iteration 8068, Loss: 0.0087, Time: 0.03s
Iteration 8069, Loss: 0.0125, Time: 0.03s
Iteration 8070, Loss: 0.0067, Time: 0.03s
Iteration 8071, Loss: 0.0099, Time: 0.03s
Iteration 8072, Loss: 0.0047, Time: 0.03s
Iteration 8073, Loss: 0.0111, Time: 0.03s
Iteration 8074, Loss: 0.0075, Time: 0.03s
Iteration 8075, Loss: 0.0092, Time: 0.03s
Iteration 8076, Loss: 0.0121, Time: 0.03s
Iteration 8077, Loss: 0.0114, Time: 0.03s
Iteration 8078, Loss: 0.0082, Time: 0.03s
Iteration 8079, Loss: 0.0068, Time: 0.03s
Iteration 8080, Loss: 0.0124, Time: 0.03s
Iteration 8081, Loss: 0.0119, Time: 0.03s
Iteration 8082, Loss: 0.0122, Time: 0.03s
Iteration 8083, Loss: 0.0046, Time: 0.03s
Iteration 8084, Loss: 0.0084, Time: 0.03s
Iteration 8085, Loss: 0.0105, Time: 0.03s
Iteration 8086, Loss: 0.0078, Time: 0.03s
Iteration 8087, Loss: 0.0110, Time: 0.03s
Iteration 8088, Loss: 0.0117, Time: 0.03s
Iteration 8089, Loss: 0.0145, Time: 0.03s
Iteration 8090, Loss: 0.0136, Time: 0.03s
Iteration 8091, Loss: 0.0113, Time: 0.03s
Iteration 8092, Loss: 0.0093, Time: 0.03s
Iteration 8093, Loss: 0.0059, Time: 0.03s
Iteration 8094, Loss: 0.0074, Time: 0.03s
```

```
Iteration 8095, Loss: 0.0087, Time: 0.03s
Iteration 8096, Loss: 0.0062, Time: 0.03s
Iteration 8097, Loss: 0.0140, Time: 0.03s
Iteration 8098, Loss: 0.0062, Time: 0.03s
Iteration 8099, Loss: 0.0105, Time: 0.03s
Iteration 8100, Loss: 0.0086, Time: 0.03s
Iteration 8100, Loss: 0.0086, Time: 0.03s
Test Loss: 0.0427
Iteration 8101, Loss: 0.0077, Time: 0.03s
Iteration 8102, Loss: 0.0134, Time: 0.03s
Iteration 8103, Loss: 0.0063, Time: 0.03s
Iteration 8104, Loss: 0.0119, Time: 0.03s
Iteration 8105, Loss: 0.0108, Time: 0.03s
Iteration 8106, Loss: 0.0107, Time: 0.03s
Iteration 8107, Loss: 0.0103, Time: 0.03s
Iteration 8108, Loss: 0.0078, Time: 0.03s
Iteration 8109, Loss: 0.0150, Time: 0.03s
Iteration 8110, Loss: 0.0138, Time: 0.03s
Iteration 8111, Loss: 0.0099, Time: 0.03s
Iteration 8112, Loss: 0.0110, Time: 0.03s
Iteration 8113, Loss: 0.0091, Time: 0.03s
Iteration 8114, Loss: 0.0100, Time: 0.03s
Iteration 8115, Loss: 0.0125, Time: 0.03s
Iteration 8116, Loss: 0.0139, Time: 0.03s
Iteration 8117, Loss: 0.0126, Time: 0.03s
Iteration 8118, Loss: 0.0109, Time: 0.03s
Iteration 8119, Loss: 0.0069, Time: 0.03s
Iteration 8120, Loss: 0.0210, Time: 0.03s
Iteration 8121, Loss: 0.0105, Time: 0.03s
Iteration 8122, Loss: 0.0112, Time: 0.03s
Iteration 8123, Loss: 0.0121, Time: 0.03s
Iteration 8124, Loss: 0.0114, Time: 0.03s
Iteration 8125, Loss: 0.0137, Time: 0.03s
Iteration 8126, Loss: 0.0123, Time: 0.03s
Iteration 8127, Loss: 0.0133, Time: 0.03s
Iteration 8128, Loss: 0.0084, Time: 0.03s
Iteration 8129, Loss: 0.0104, Time: 0.03s
Iteration 8130, Loss: 0.0117, Time: 0.03s
Iteration 8131, Loss: 0.0155, Time: 0.03s
Iteration 8132, Loss: 0.0090, Time: 0.03s
Iteration 8133, Loss: 0.0090, Time: 0.03s
Iteration 8134, Loss: 0.0140, Time: 0.03s
Iteration 8135, Loss: 0.0117, Time: 0.03s
Iteration 8136, Loss: 0.0103, Time: 0.03s
Iteration 8137, Loss: 0.0123, Time: 0.03s
Iteration 8138, Loss: 0.0071, Time: 0.03s
Iteration 8139, Loss: 0.0148, Time: 0.03s
Iteration 8140, Loss: 0.0127, Time: 0.03s
```

```
Iteration 8141, Loss: 0.0090, Time: 0.03s
Iteration 8142, Loss: 0.0152, Time: 0.03s
Iteration 8143, Loss: 0.0078, Time: 0.03s
Iteration 8144, Loss: 0.0154, Time: 0.03s
Iteration 8145, Loss: 0.0096, Time: 0.03s
Iteration 8146, Loss: 0.0087, Time: 0.03s
Iteration 8147, Loss: 0.0177, Time: 0.03s
Iteration 8148, Loss: 0.0140, Time: 0.03s
Iteration 8149, Loss: 0.0100, Time: 0.03s
Iteration 8150, Loss: 0.0133, Time: 0.03s
Iteration 8151, Loss: 0.0100, Time: 0.03s
Iteration 8152, Loss: 0.0111, Time: 0.03s
Iteration 8153, Loss: 0.0079, Time: 0.03s
Iteration 8154, Loss: 0.0079, Time: 0.03s
Iteration 8155, Loss: 0.0117, Time: 0.03s
Iteration 8156, Loss: 0.0093, Time: 0.03s
Iteration 8157, Loss: 0.0134, Time: 0.03s
Iteration 8158, Loss: 0.0117, Time: 0.03s
Iteration 8159, Loss: 0.0111, Time: 0.03s
Iteration 8160, Loss: 0.0124, Time: 0.03s
Iteration 8161, Loss: 0.0083, Time: 0.03s
Iteration 8162, Loss: 0.0161, Time: 0.03s
Iteration 8163, Loss: 0.0079, Time: 0.03s
Iteration 8164, Loss: 0.0088, Time: 0.03s
Iteration 8165, Loss: 0.0096, Time: 0.03s
Iteration 8166, Loss: 0.0093, Time: 0.03s
Iteration 8167, Loss: 0.0226, Time: 0.03s
Iteration 8168, Loss: 0.0139, Time: 0.03s
Iteration 8169, Loss: 0.0044, Time: 0.03s
Iteration 8170, Loss: 0.0078, Time: 0.03s
Iteration 8171, Loss: 0.0118, Time: 0.03s
Iteration 8172, Loss: 0.0094, Time: 0.03s
Iteration 8173, Loss: 0.0069, Time: 0.03s
Iteration 8174, Loss: 0.0140, Time: 0.03s
Iteration 8175, Loss: 0.0102, Time: 0.03s
Iteration 8176, Loss: 0.0066, Time: 0.03s
Iteration 8177, Loss: 0.0094, Time: 0.03s
Iteration 8178, Loss: 0.0071, Time: 0.03s
Iteration 8179, Loss: 0.0103, Time: 0.03s
Iteration 8180, Loss: 0.0133, Time: 0.03s
Iteration 8181, Loss: 0.0083, Time: 0.03s
Iteration 8182, Loss: 0.0086, Time: 0.03s
Iteration 8183, Loss: 0.0115, Time: 0.03s
Iteration 8184, Loss: 0.0087, Time: 0.03s
Iteration 8185, Loss: 0.0088, Time: 0.03s
Iteration 8186, Loss: 0.0111, Time: 0.03s
Iteration 8187, Loss: 0.0108, Time: 0.03s
Iteration 8188, Loss: 0.0126, Time: 0.03s
```

```
Iteration 8189, Loss: 0.0128, Time: 0.03s
Iteration 8190, Loss: 0.0103, Time: 0.03s
Iteration 8191, Loss: 0.0096, Time: 0.03s
Iteration 8192, Loss: 0.0080, Time: 0.03s
Iteration 8193, Loss: 0.0074, Time: 0.03s
Iteration 8194, Loss: 0.0147, Time: 0.03s
Iteration 8195, Loss: 0.0085, Time: 0.03s
Iteration 8196, Loss: 0.0111, Time: 0.03s
Iteration 8197, Loss: 0.0086, Time: 0.03s
Iteration 8198, Loss: 0.0099, Time: 0.03s
Iteration 8199, Loss: 0.0090, Time: 0.03s
Iteration 8200, Loss: 0.0112, Time: 0.03s
Iteration 8200, Loss: 0.0112, Time: 0.03s
Test Loss: 0.0352
Iteration 8201, Loss: 0.0155, Time: 0.03s
Iteration 8202, Loss: 0.0119, Time: 0.03s
Iteration 8203, Loss: 0.0094, Time: 0.03s
Iteration 8204, Loss: 0.0077, Time: 0.03s
Iteration 8205, Loss: 0.0100, Time: 0.03s
Iteration 8206, Loss: 0.0114, Time: 0.03s
Iteration 8207, Loss: 0.0119, Time: 0.03s
Iteration 8208, Loss: 0.0137, Time: 0.03s
Iteration 8209, Loss: 0.0097, Time: 0.03s
Iteration 8210, Loss: 0.0138, Time: 0.03s
Iteration 8211, Loss: 0.0085, Time: 0.04s
Iteration 8212, Loss: 0.0126, Time: 0.03s
Iteration 8213, Loss: 0.0076, Time: 0.03s
Iteration 8214, Loss: 0.0084, Time: 0.03s
Iteration 8215, Loss: 0.0108, Time: 0.03s
Iteration 8216, Loss: 0.0110, Time: 0.03s
Iteration 8217, Loss: 0.0095, Time: 0.03s
Iteration 8218, Loss: 0.0098, Time: 0.03s
Iteration 8219, Loss: 0.0141, Time: 0.03s
Iteration 8220, Loss: 0.0113, Time: 0.03s
Iteration 8221, Loss: 0.0115, Time: 0.03s
Iteration 8222, Loss: 0.0090, Time: 0.03s
Iteration 8223, Loss: 0.0089, Time: 0.03s
Iteration 8224, Loss: 0.0081, Time: 0.03s
Iteration 8225, Loss: 0.0099, Time: 0.03s
Iteration 8226, Loss: 0.0127, Time: 0.03s
Iteration 8227, Loss: 0.0067, Time: 0.03s
Iteration 8228, Loss: 0.0099, Time: 0.03s
Iteration 8229, Loss: 0.0142, Time: 0.03s
Iteration 8230, Loss: 0.0077, Time: 0.03s
Iteration 8231, Loss: 0.0086, Time: 0.03s
Iteration 8232, Loss: 0.0105, Time: 0.03s
Iteration 8233, Loss: 0.0119, Time: 0.03s
Iteration 8234, Loss: 0.0033, Time: 0.03s
```

```
Iteration 8235, Loss: 0.0117, Time: 0.03s
Iteration 8236, Loss: 0.0093, Time: 0.03s
Iteration 8237, Loss: 0.0104, Time: 0.03s
Iteration 8238, Loss: 0.0144, Time: 0.03s
Iteration 8239, Loss: 0.0072, Time: 0.03s
Iteration 8240, Loss: 0.0110, Time: 0.03s
Iteration 8241, Loss: 0.0114, Time: 0.03s
Iteration 8242, Loss: 0.0104, Time: 0.03s
Iteration 8243, Loss: 0.0098, Time: 0.03s
Iteration 8244, Loss: 0.0131, Time: 0.03s
Iteration 8245, Loss: 0.0115, Time: 0.03s
Iteration 8246, Loss: 0.0097, Time: 0.03s
Iteration 8247, Loss: 0.0116, Time: 0.03s
Iteration 8248, Loss: 0.0113, Time: 0.03s
Iteration 8249, Loss: 0.0065, Time: 0.03s
Iteration 8250, Loss: 0.0086, Time: 0.03s
Iteration 8251, Loss: 0.0099, Time: 0.03s
Iteration 8252, Loss: 0.0139, Time: 0.03s
Iteration 8253, Loss: 0.0100, Time: 0.03s
Iteration 8254, Loss: 0.0066, Time: 0.03s
Iteration 8255, Loss: 0.0104, Time: 0.03s
Iteration 8256, Loss: 0.0127, Time: 0.03s
Iteration 8257, Loss: 0.0135, Time: 0.03s
Iteration 8258, Loss: 0.0109, Time: 0.03s
Iteration 8259, Loss: 0.0082, Time: 0.03s
Iteration 8260, Loss: 0.0090, Time: 0.03s
Iteration 8261, Loss: 0.0136, Time: 0.03s
Iteration 8262, Loss: 0.0138, Time: 0.03s
Iteration 8263, Loss: 0.0103, Time: 0.03s
Iteration 8264, Loss: 0.0105, Time: 0.03s
Iteration 8265, Loss: 0.0110, Time: 0.03s
Iteration 8266, Loss: 0.0082, Time: 0.03s
Iteration 8267, Loss: 0.0108, Time: 0.03s
Iteration 8268, Loss: 0.0035, Time: 0.03s
Iteration 8269, Loss: 0.0111, Time: -0.12s
Iteration 8270, Loss: 0.0118, Time: 0.03s
Iteration 8271, Loss: 0.0058, Time: 0.03s
Iteration 8272, Loss: 0.0101, Time: 0.03s
Iteration 8273, Loss: 0.0126, Time: 0.03s
Iteration 8274, Loss: 0.0109, Time: 0.03s
Iteration 8275, Loss: 0.0147, Time: 0.03s
Iteration 8276, Loss: 0.0081, Time: 0.03s
Iteration 8277, Loss: 0.0120, Time: 0.03s
Iteration 8278, Loss: 0.0093, Time: 0.03s
Iteration 8279, Loss: 0.0120, Time: 0.03s
Iteration 8280, Loss: 0.0143, Time: 0.03s
Iteration 8281, Loss: 0.0072, Time: 0.03s
Iteration 8282, Loss: 0.0103, Time: 0.03s
```

```
Iteration 8283, Loss: 0.0113, Time: 0.03s
Iteration 8284, Loss: 0.0073, Time: 0.03s
Iteration 8285, Loss: 0.0112, Time: 0.03s
Iteration 8286, Loss: 0.0090, Time: 0.03s
Iteration 8287, Loss: 0.0078, Time: 0.03s
Iteration 8288, Loss: 0.0072, Time: 0.03s
Iteration 8289, Loss: 0.0073, Time: 0.03s
Iteration 8290, Loss: 0.0107, Time: 0.03s
Iteration 8291, Loss: 0.0103, Time: 0.03s
Iteration 8292, Loss: 0.0087, Time: 0.03s
Iteration 8293, Loss: 0.0104, Time: 0.03s
Iteration 8294, Loss: 0.0137, Time: 0.03s
Iteration 8295, Loss: 0.0076, Time: 0.03s
Iteration 8296, Loss: 0.0051, Time: 0.03s
Iteration 8297, Loss: 0.0120, Time: 0.03s
Iteration 8298, Loss: 0.0080, Time: 0.03s
Iteration 8299, Loss: 0.0082, Time: 0.03s
Iteration 8300, Loss: 0.0079, Time: 0.03s
Iteration 8300, Loss: 0.0079, Time: 0.03s
Test Loss: 0.0355
Iteration 8301, Loss: 0.0048, Time: 0.03s
Iteration 8302, Loss: 0.0038, Time: 0.03s
Iteration 8303, Loss: 0.0134, Time: 0.03s
Iteration 8304, Loss: 0.0120, Time: 0.03s
Iteration 8305, Loss: 0.0136, Time: 0.03s
Iteration 8306, Loss: 0.0080, Time: 0.03s
Iteration 8307, Loss: 0.0123, Time: 0.03s
Iteration 8308, Loss: 0.0098, Time: 0.03s
Iteration 8309, Loss: 0.0084, Time: 0.03s
Iteration 8310, Loss: 0.0130, Time: 0.03s
Iteration 8311, Loss: 0.0123, Time: 0.03s
Iteration 8312, Loss: 0.0112, Time: 0.03s
Iteration 8313, Loss: 0.0121, Time: 0.03s
Iteration 8314, Loss: 0.0134, Time: 0.03s
Iteration 8315, Loss: 0.0053, Time: 0.03s
Iteration 8316, Loss: 0.0080, Time: 0.03s
Iteration 8317, Loss: 0.0093, Time: 0.03s
Iteration 8318, Loss: 0.0128, Time: 0.03s
Iteration 8319, Loss: 0.0080, Time: 0.03s
Iteration 8320, Loss: 0.0125, Time: 0.03s
Iteration 8321, Loss: 0.0097, Time: 0.03s
Iteration 8322, Loss: 0.0130, Time: 0.03s
Iteration 8323, Loss: 0.0106, Time: 0.03s
Iteration 8324, Loss: 0.0120, Time: 0.03s
Iteration 8325, Loss: 0.0111, Time: 0.03s
Iteration 8326, Loss: 0.0113, Time: 0.03s
Iteration 8327, Loss: 0.0074, Time: 0.03s
Iteration 8328, Loss: 0.0082, Time: 0.03s
```

```
Iteration 8329, Loss: 0.0097, Time: 0.03s
Iteration 8330, Loss: 0.0082, Time: 0.03s
Iteration 8331, Loss: 0.0110, Time: 0.03s
Iteration 8332, Loss: 0.0085, Time: 0.03s
Iteration 8333, Loss: 0.0086, Time: 0.03s
Iteration 8334, Loss: 0.0089, Time: 0.03s
Iteration 8335, Loss: 0.0110, Time: 0.03s
Iteration 8336, Loss: 0.0074, Time: 0.03s
Iteration 8337, Loss: 0.0088, Time: 0.03s
Iteration 8338, Loss: 0.0113, Time: 0.03s
Iteration 8339, Loss: 0.0075, Time: 0.03s
Iteration 8340, Loss: 0.0105, Time: 0.03s
Iteration 8341, Loss: 0.0125, Time: 0.03s
Iteration 8342, Loss: 0.0135, Time: 0.03s
Iteration 8343, Loss: 0.0068, Time: 0.03s
Iteration 8344, Loss: 0.0088, Time: 0.03s
Iteration 8345, Loss: 0.0137, Time: 0.03s
Iteration 8346, Loss: 0.0085, Time: 0.03s
Iteration 8347, Loss: 0.0086, Time: 0.03s
Iteration 8348, Loss: 0.0073, Time: 0.03s
Iteration 8349, Loss: 0.0107, Time: 0.03s
Iteration 8350, Loss: 0.0094, Time: 0.03s
Iteration 8351, Loss: 0.0068, Time: 0.03s
Iteration 8352, Loss: 0.0107, Time: 0.03s
Iteration 8353, Loss: 0.0089, Time: 0.03s
Iteration 8354, Loss: 0.0116, Time: 0.03s
Iteration 8355, Loss: 0.0091, Time: 0.03s
Iteration 8356, Loss: 0.0082, Time: 0.03s
Iteration 8357, Loss: 0.0097, Time: 0.03s
Iteration 8358, Loss: 0.0131, Time: 0.03s
Iteration 8359, Loss: 0.0148, Time: 0.03s
Iteration 8360, Loss: 0.0124, Time: 0.03s
Iteration 8361, Loss: 0.0091, Time: 0.03s
Iteration 8362, Loss: 0.0110, Time: 0.03s
Iteration 8363, Loss: 0.0128, Time: 0.03s
Iteration 8364, Loss: 0.0100, Time: 0.03s
Iteration 8365, Loss: 0.0123, Time: 0.03s
Iteration 8366, Loss: 0.0108, Time: 0.03s
Iteration 8367, Loss: 0.0095, Time: 0.03s
Iteration 8368, Loss: 0.0129, Time: 0.03s
Iteration 8369, Loss: 0.0067, Time: 0.03s
Iteration 8370, Loss: 0.0124, Time: 0.03s
Iteration 8371, Loss: 0.0178, Time: 0.03s
Iteration 8372, Loss: 0.0119, Time: 0.03s
Iteration 8373, Loss: 0.0106, Time: 0.03s
Iteration 8374, Loss: 0.0093, Time: 0.03s
Iteration 8375, Loss: 0.0111, Time: 0.03s
Iteration 8376, Loss: 0.0086, Time: 0.03s
```

```
Iteration 8377, Loss: 0.0086, Time: 0.03s
Iteration 8378, Loss: 0.0069, Time: 0.03s
Iteration 8379, Loss: 0.0116, Time: 0.03s
Iteration 8380, Loss: 0.0076, Time: 0.03s
Iteration 8381, Loss: 0.0093, Time: 0.03s
Iteration 8382, Loss: 0.0130, Time: 0.03s
Iteration 8383, Loss: 0.0082, Time: 0.03s
Iteration 8384, Loss: 0.0099, Time: 0.03s
Iteration 8385, Loss: 0.0125, Time: 0.03s
Iteration 8386, Loss: 0.0108, Time: 0.03s
Iteration 8387, Loss: 0.0056, Time: 0.03s
Iteration 8388, Loss: 0.0132, Time: 0.03s
Iteration 8389, Loss: 0.0079, Time: 0.03s
Iteration 8390, Loss: 0.0083, Time: 0.03s
Iteration 8391, Loss: 0.0116, Time: 0.03s
Iteration 8392, Loss: 0.0128, Time: 0.03s
Iteration 8393, Loss: 0.0075, Time: 0.03s
Iteration 8394, Loss: 0.0097, Time: 0.03s
Iteration 8395, Loss: 0.0086, Time: 0.03s
Iteration 8396, Loss: 0.0095, Time: 0.03s
Iteration 8397, Loss: 0.0065, Time: 0.03s
Iteration 8398, Loss: 0.0083, Time: 0.03s
Iteration 8399, Loss: 0.0080, Time: 0.03s
Iteration 8400, Loss: 0.0128, Time: 0.03s
Iteration 8400, Loss: 0.0128, Time: 0.03s
Test Loss: 0.0601
Iteration 8401, Loss: 0.0075, Time: 0.03s
Iteration 8402, Loss: 0.0105, Time: 0.03s
Iteration 8403, Loss: 0.0174, Time: 0.03s
Iteration 8404, Loss: 0.0099, Time: 0.03s
Iteration 8405, Loss: 0.0102, Time: 0.03s
Iteration 8406, Loss: 0.0092, Time: 0.03s
Iteration 8407, Loss: 0.0119, Time: 0.03s
Iteration 8408, Loss: 0.0061, Time: 0.03s
Iteration 8409, Loss: 0.0127, Time: 0.03s
Iteration 8410, Loss: 0.0060, Time: 0.03s
Iteration 8411, Loss: 0.0123, Time: 0.03s
Iteration 8412, Loss: 0.0086, Time: 0.03s
Iteration 8413, Loss: 0.0108, Time: 0.03s
Iteration 8414, Loss: 0.0102, Time: 0.03s
Iteration 8415, Loss: 0.0088, Time: 0.03s
Iteration 8416, Loss: 0.0103, Time: 0.03s
Iteration 8417, Loss: 0.0119, Time: 0.03s
Iteration 8418, Loss: 0.0095, Time: 0.03s
Iteration 8419, Loss: 0.0090, Time: 0.03s
Iteration 8420, Loss: 0.0095, Time: 0.03s
Iteration 8421, Loss: 0.0101, Time: 0.03s
Iteration 8422, Loss: 0.0094, Time: 0.03s
```

```
Iteration 8423, Loss: 0.0125, Time: 0.03s
Iteration 8424, Loss: 0.0102, Time: 0.03s
Iteration 8425, Loss: 0.0118, Time: 0.03s
Iteration 8426, Loss: 0.0107, Time: 0.03s
Iteration 8427, Loss: 0.0081, Time: 0.03s
Iteration 8428, Loss: 0.0067, Time: 0.03s
Iteration 8429, Loss: 0.0065, Time: 0.03s
Iteration 8430, Loss: 0.0077, Time: 0.03s
Iteration 8431, Loss: 0.0105, Time: 0.03s
Iteration 8432, Loss: 0.0070, Time: 0.03s
Iteration 8433, Loss: 0.0135, Time: 0.03s
Iteration 8434, Loss: 0.0066, Time: 0.03s
Iteration 8435, Loss: 0.0128, Time: 0.03s
Iteration 8436, Loss: 0.0115, Time: 0.03s
Iteration 8437, Loss: 0.0092, Time: 0.03s
Iteration 8438, Loss: 0.0056, Time: 0.03s
Iteration 8439, Loss: 0.0083, Time: 0.03s
Iteration 8440, Loss: 0.0091, Time: 0.03s
Iteration 8441, Loss: 0.0067, Time: 0.03s
Iteration 8442, Loss: 0.0133, Time: 0.03s
Iteration 8443, Loss: 0.0060, Time: 0.03s
Iteration 8444, Loss: 0.0112, Time: 0.03s
Iteration 8445, Loss: 0.0089, Time: 0.03s
Iteration 8446, Loss: 0.0076, Time: 0.03s
Iteration 8447, Loss: 0.0081, Time: 0.03s
Iteration 8448, Loss: 0.0106, Time: 0.03s
Iteration 8449, Loss: 0.0128, Time: 0.03s
Iteration 8450, Loss: 0.0143, Time: 0.03s
Iteration 8451, Loss: 0.0109, Time: 0.03s
Iteration 8452, Loss: 0.0066, Time: 0.03s
Iteration 8453, Loss: 0.0149, Time: 0.03s
Iteration 8454, Loss: 0.0097, Time: 0.03s
Iteration 8455, Loss: 0.0085, Time: 0.03s
Iteration 8456, Loss: 0.0128, Time: 0.03s
Iteration 8457, Loss: 0.0102, Time: 0.03s
Iteration 8458, Loss: 0.0086, Time: 0.03s
Iteration 8459, Loss: 0.0162, Time: 0.03s
Iteration 8460, Loss: 0.0102, Time: 0.03s
Iteration 8461, Loss: 0.0125, Time: 0.03s
Iteration 8462, Loss: 0.0099, Time: 0.03s
Iteration 8463, Loss: 0.0105, Time: 0.03s
Iteration 8464, Loss: 0.0140, Time: 0.03s
Iteration 8465, Loss: 0.0116, Time: 0.03s
Iteration 8466, Loss: 0.0083, Time: 0.03s
Iteration 8467, Loss: 0.0081, Time: 0.03s
Iteration 8468, Loss: 0.0083, Time: 0.03s
Iteration 8469, Loss: 0.0115, Time: 0.03s
Iteration 8470, Loss: 0.0070, Time: 0.03s
```

```
Iteration 8471, Loss: 0.0141, Time: 0.03s
Iteration 8472, Loss: 0.0090, Time: 0.03s
Iteration 8473, Loss: 0.0080, Time: 0.03s
Iteration 8474, Loss: 0.0093, Time: 0.03s
Iteration 8475, Loss: 0.0125, Time: 0.03s
Iteration 8476, Loss: 0.0135, Time: 0.03s
Iteration 8477, Loss: 0.0083, Time: 0.03s
Iteration 8478, Loss: 0.0047, Time: 0.03s
Iteration 8479, Loss: 0.0099, Time: 0.03s
Iteration 8480, Loss: 0.0036, Time: 0.03s
Iteration 8481, Loss: 0.0122, Time: 0.03s
Iteration 8482, Loss: 0.0139, Time: 0.03s
Iteration 8483, Loss: 0.0082, Time: 0.03s
Iteration 8484, Loss: 0.0163, Time: 0.03s
Iteration 8485, Loss: 0.0088, Time: 0.03s
Iteration 8486, Loss: 0.0113, Time: 0.03s
Iteration 8487, Loss: 0.0119, Time: 0.03s
Iteration 8488, Loss: 0.0113, Time: 0.03s
Iteration 8489, Loss: 0.0081, Time: 0.03s
Iteration 8490, Loss: 0.0089, Time: 0.03s
Iteration 8491, Loss: 0.0130, Time: 0.03s
Iteration 8492, Loss: 0.0087, Time: 0.03s
Iteration 8493, Loss: 0.0107, Time: 0.03s
Iteration 8494, Loss: 0.0103, Time: 0.03s
Iteration 8495, Loss: 0.0041, Time: 0.03s
Iteration 8496, Loss: 0.0091, Time: 0.03s
Iteration 8497, Loss: 0.0099, Time: 0.03s
Iteration 8498, Loss: 0.0058, Time: 0.03s
Iteration 8499, Loss: 0.0111, Time: 0.03s
Iteration 8500, Loss: 0.0111, Time: 0.03s
Iteration 8500, Loss: 0.0111, Time: 0.03s
Test Loss: 0.0571
Iteration 8501, Loss: 0.0140, Time: 0.03s
Iteration 8502, Loss: 0.0101, Time: 0.03s
Iteration 8503, Loss: 0.0114, Time: 0.03s
Iteration 8504, Loss: 0.0119, Time: 0.03s
Iteration 8505, Loss: 0.0102, Time: 0.03s
Iteration 8506, Loss: 0.0059, Time: 0.03s
Iteration 8507, Loss: 0.0064, Time: 0.03s
Iteration 8508, Loss: 0.0101, Time: 0.03s
Iteration 8509, Loss: 0.0133, Time: 0.03s
Iteration 8510, Loss: 0.0088, Time: 0.03s
Iteration 8511, Loss: 0.0081, Time: 0.03s
Iteration 8512, Loss: 0.0090, Time: 0.03s
Iteration 8513, Loss: 0.0058, Time: 0.03s
Iteration 8514, Loss: 0.0111, Time: 0.03s
Iteration 8515, Loss: 0.0049, Time: 0.03s
Iteration 8516, Loss: 0.0135, Time: 0.03s
```

```
Iteration 8517, Loss: 0.0088, Time: 0.03s
Iteration 8518, Loss: 0.0082, Time: 0.03s
Iteration 8519, Loss: 0.0079, Time: 0.03s
Iteration 8520, Loss: 0.0114, Time: 0.03s
Iteration 8521, Loss: 0.0081, Time: 0.03s
Iteration 8522, Loss: 0.0056, Time: 0.03s
Iteration 8523, Loss: 0.0107, Time: 0.03s
Iteration 8524, Loss: 0.0090, Time: 0.03s
Iteration 8525, Loss: 0.0096, Time: 0.03s
Iteration 8526, Loss: 0.0164, Time: 0.03s
Iteration 8527, Loss: 0.0063, Time: 0.03s
Iteration 8528, Loss: 0.0093, Time: 0.03s
Iteration 8529, Loss: 0.0126, Time: 0.03s
Iteration 8530, Loss: 0.0067, Time: 0.03s
Iteration 8531, Loss: 0.0114, Time: 0.03s
Iteration 8532, Loss: 0.0161, Time: 0.03s
Iteration 8533, Loss: 0.0078, Time: 0.03s
Iteration 8534, Loss: 0.0120, Time: 0.03s
Iteration 8535, Loss: 0.0137, Time: 0.03s
Iteration 8536, Loss: 0.0098, Time: 0.03s
Iteration 8537, Loss: 0.0126, Time: 0.03s
Iteration 8538, Loss: 0.0072, Time: 0.03s
Iteration 8539, Loss: 0.0082, Time: 0.03s
Iteration 8540, Loss: 0.0131, Time: 0.03s
Iteration 8541, Loss: 0.0065, Time: 0.03s
Iteration 8542, Loss: 0.0096, Time: 0.03s
Iteration 8543, Loss: 0.0085, Time: 0.03s
Iteration 8544, Loss: 0.0103, Time: 0.03s
Iteration 8545, Loss: 0.0136, Time: 0.03s
Iteration 8546, Loss: 0.0110, Time: 0.03s
Iteration 8547, Loss: 0.0128, Time: 0.03s
Iteration 8548, Loss: 0.0109, Time: 0.03s
Iteration 8549, Loss: 0.0124, Time: 0.03s
Iteration 8550, Loss: 0.0095, Time: 0.03s
Iteration 8551, Loss: 0.0118, Time: 0.03s
Iteration 8552, Loss: 0.0102, Time: 0.03s
Iteration 8553, Loss: 0.0094, Time: 0.03s
Iteration 8554, Loss: 0.0027, Time: 0.03s
Iteration 8555, Loss: 0.0123, Time: 0.03s
Iteration 8556, Loss: 0.0116, Time: 0.03s
Iteration 8557, Loss: 0.0120, Time: 0.03s
Iteration 8558, Loss: 0.0118, Time: 0.03s
Iteration 8559, Loss: 0.0071, Time: 0.03s
Iteration 8560, Loss: 0.0101, Time: 0.03s
Iteration 8561, Loss: 0.0127, Time: 0.03s
Iteration 8562, Loss: 0.0094, Time: 0.03s
Iteration 8563, Loss: 0.0071, Time: 0.03s
Iteration 8564, Loss: 0.0126, Time: 0.03s
```

```
Iteration 8565, Loss: 0.0129, Time: 0.03s
Iteration 8566, Loss: 0.0111, Time: 0.03s
Iteration 8567, Loss: 0.0090, Time: 0.03s
Iteration 8568, Loss: 0.0097, Time: 0.03s
Iteration 8569, Loss: 0.0095, Time: 0.03s
Iteration 8570, Loss: 0.0101, Time: 0.03s
Iteration 8571, Loss: 0.0104, Time: 0.03s
Iteration 8572, Loss: 0.0116, Time: 0.03s
Iteration 8573, Loss: 0.0123, Time: 0.03s
Iteration 8574, Loss: 0.0097, Time: 0.03s
Iteration 8575, Loss: 0.0094, Time: 0.03s
Iteration 8576, Loss: 0.0123, Time: 0.03s
Iteration 8577, Loss: 0.0092, Time: 0.03s
Iteration 8578, Loss: 0.0112, Time: 0.03s
Iteration 8579, Loss: 0.0105, Time: 0.03s
Iteration 8580, Loss: 0.0088, Time: 0.03s
Iteration 8581, Loss: 0.0072, Time: 0.03s
Iteration 8582, Loss: 0.0104, Time: 0.03s
Iteration 8583, Loss: 0.0117, Time: 0.03s
Iteration 8584, Loss: 0.0079, Time: 0.03s
Iteration 8585, Loss: 0.0107, Time: 0.03s
Iteration 8586, Loss: 0.0087, Time: 0.03s
Iteration 8587, Loss: 0.0088, Time: 0.03s
Iteration 8588, Loss: 0.0071, Time: 0.03s
Iteration 8589, Loss: 0.0088, Time: 0.03s
Iteration 8590, Loss: 0.0066, Time: 0.03s
Iteration 8591, Loss: 0.0052, Time: 0.03s
Iteration 8592, Loss: 0.0094, Time: 0.03s
Iteration 8593, Loss: 0.0093, Time: 0.03s
Iteration 8594, Loss: 0.0106, Time: 0.03s
Iteration 8595, Loss: 0.0096, Time: 0.03s
Iteration 8596, Loss: 0.0086, Time: 0.03s
Iteration 8597, Loss: 0.0142, Time: 0.03s
Iteration 8598, Loss: 0.0101, Time: 0.03s
Iteration 8599, Loss: 0.0061, Time: 0.03s
Iteration 8600, Loss: 0.0107, Time: 0.03s
Iteration 8600, Loss: 0.0107, Time: 0.03s
Test Loss: 0.0594
Iteration 8601, Loss: 0.0115, Time: 0.03s
Iteration 8602, Loss: 0.0067, Time: 0.03s
Iteration 8603, Loss: 0.0104, Time: 0.03s
Iteration 8604, Loss: 0.0069, Time: 0.03s
Iteration 8605, Loss: 0.0115, Time: 0.03s
Iteration 8606, Loss: 0.0163, Time: 0.03s
Iteration 8607, Loss: 0.0050, Time: 0.03s
Iteration 8608, Loss: 0.0103, Time: 0.03s
Iteration 8609, Loss: 0.0137, Time: 0.03s
Iteration 8610, Loss: 0.0129, Time: 0.03s
```

```
Iteration 8611, Loss: 0.0074, Time: 0.03s
Iteration 8612, Loss: 0.0096, Time: 0.03s
Iteration 8613, Loss: 0.0058, Time: 0.03s
Iteration 8614, Loss: 0.0102, Time: 0.03s
Iteration 8615, Loss: 0.0088, Time: 0.03s
Iteration 8616, Loss: 0.0122, Time: 0.03s
Iteration 8617, Loss: 0.0099, Time: 0.03s
Iteration 8618, Loss: 0.0118, Time: 0.03s
Iteration 8619, Loss: 0.0107, Time: 0.03s
Iteration 8620, Loss: 0.0123, Time: 0.03s
Iteration 8621, Loss: 0.0144, Time: 0.03s
Iteration 8622, Loss: 0.0085, Time: 0.03s
Iteration 8623, Loss: 0.0090, Time: 0.03s
Iteration 8624, Loss: 0.0092, Time: 0.03s
Iteration 8625, Loss: 0.0072, Time: 0.03s
Iteration 8626, Loss: 0.0117, Time: 0.03s
Iteration 8627, Loss: 0.0062, Time: 0.03s
Iteration 8628, Loss: 0.0082, Time: 0.03s
Iteration 8629, Loss: 0.0103, Time: 0.03s
Iteration 8630, Loss: 0.0075, Time: 0.03s
Iteration 8631, Loss: 0.0058, Time: 0.03s
Iteration 8632, Loss: 0.0085, Time: 0.03s
Iteration 8633, Loss: 0.0103, Time: 0.03s
Iteration 8634, Loss: 0.0084, Time: 0.03s
Iteration 8635, Loss: 0.0091, Time: 0.03s
Iteration 8636, Loss: 0.0058, Time: 0.03s
Iteration 8637, Loss: 0.0076, Time: 0.03s
Iteration 8638, Loss: 0.0094, Time: 0.03s
Iteration 8639, Loss: 0.0103, Time: 0.03s
Iteration 8640, Loss: 0.0083, Time: 0.03s
Iteration 8641, Loss: 0.0091, Time: 0.03s
Iteration 8642, Loss: 0.0159, Time: 0.03s
Iteration 8643, Loss: 0.0075, Time: 0.03s
Iteration 8644, Loss: 0.0068, Time: 0.03s
Iteration 8645, Loss: 0.0069, Time: 0.03s
Iteration 8646, Loss: 0.0071, Time: 0.03s
Iteration 8647, Loss: 0.0097, Time: 0.03s
Iteration 8648, Loss: 0.0175, Time: 0.03s
Iteration 8649, Loss: 0.0092, Time: 0.03s
Iteration 8650, Loss: 0.0108, Time: 0.03s
Iteration 8651, Loss: 0.0080, Time: 0.03s
Iteration 8652, Loss: 0.0135, Time: 0.03s
Iteration 8653, Loss: 0.0063, Time: 0.03s
Iteration 8654, Loss: 0.0077, Time: 0.03s
Iteration 8655, Loss: 0.0114, Time: 0.03s
Iteration 8656, Loss: 0.0120, Time: 0.03s
Iteration 8657, Loss: 0.0104, Time: 0.03s
Iteration 8658, Loss: 0.0094, Time: 0.03s
```

```
Iteration 8659, Loss: 0.0109, Time: 0.03s
Iteration 8660, Loss: 0.0093, Time: 0.03s
Iteration 8661, Loss: 0.0107, Time: 0.03s
Iteration 8662, Loss: 0.0054, Time: 0.03s
Iteration 8663, Loss: 0.0089, Time: 0.03s
Iteration 8664, Loss: 0.0114, Time: 0.03s
Iteration 8665, Loss: 0.0086, Time: 0.03s
Iteration 8666, Loss: 0.0135, Time: 0.03s
Iteration 8667, Loss: 0.0107, Time: 0.03s
Iteration 8668, Loss: 0.0103, Time: 0.03s
Iteration 8669, Loss: 0.0066, Time: 0.03s
Iteration 8670, Loss: 0.0126, Time: 0.03s
Iteration 8671, Loss: 0.0060, Time: 0.03s
Iteration 8672, Loss: 0.0071, Time: 0.03s
Iteration 8673, Loss: 0.0074, Time: 0.03s
Iteration 8674, Loss: 0.0114, Time: 0.03s
Iteration 8675, Loss: 0.0106, Time: 0.03s
Iteration 8676, Loss: 0.0131, Time: 0.03s
Iteration 8677, Loss: 0.0097, Time: 0.03s
Iteration 8678, Loss: 0.0100, Time: 0.03s
Iteration 8679, Loss: 0.0068, Time: 0.03s
Iteration 8680, Loss: 0.0167, Time: 0.03s
Iteration 8681, Loss: 0.0110, Time: 0.03s
Iteration 8682, Loss: 0.0119, Time: 0.03s
Iteration 8683, Loss: 0.0084, Time: 0.03s
Iteration 8684, Loss: 0.0128, Time: 0.03s
Iteration 8685, Loss: 0.0116, Time: 0.03s
Iteration 8686, Loss: 0.0099, Time: 0.03s
Iteration 8687, Loss: 0.0080, Time: 0.03s
Iteration 8688, Loss: 0.0119, Time: 0.03s
Iteration 8689, Loss: 0.0086, Time: 0.03s
Iteration 8690, Loss: 0.0126, Time: 0.03s
Iteration 8691, Loss: 0.0081, Time: 0.03s
Iteration 8692, Loss: 0.0119, Time: 0.03s
Iteration 8693, Loss: 0.0100, Time: 0.03s
Iteration 8694, Loss: 0.0081, Time: 0.03s
Iteration 8695, Loss: 0.0135, Time: 0.03s
Iteration 8696, Loss: 0.0101, Time: 0.03s
Iteration 8697, Loss: 0.0082, Time: 0.03s
Iteration 8698, Loss: 0.0094, Time: 0.03s
Iteration 8699, Loss: 0.0096, Time: 0.03s
Iteration 8700, Loss: 0.0075, Time: 0.03s
Iteration 8700, Loss: 0.0075, Time: 0.03s
Test Loss: 0.0483
Iteration 8701, Loss: 0.0103, Time: 0.03s
Iteration 8702, Loss: 0.0114, Time: 0.03s
Iteration 8703, Loss: 0.0099, Time: 0.03s
Iteration 8704, Loss: 0.0096, Time: 0.03s
```

```
Iteration 8705, Loss: 0.0124, Time: 0.03s
Iteration 8706, Loss: 0.0097, Time: 0.03s
Iteration 8707, Loss: 0.0147, Time: 0.03s
Iteration 8708, Loss: 0.0090, Time: 0.03s
Iteration 8709, Loss: 0.0132, Time: 0.03s
Iteration 8710, Loss: 0.0116, Time: 0.03s
Iteration 8711, Loss: 0.0099, Time: 0.03s
Iteration 8712, Loss: 0.0073, Time: 0.03s
Iteration 8713, Loss: 0.0112, Time: 0.03s
Iteration 8714, Loss: 0.0106, Time: 0.03s
Iteration 8715, Loss: 0.0097, Time: 0.03s
Iteration 8716, Loss: 0.0120, Time: 0.04s
Iteration 8717, Loss: 0.0105, Time: 0.03s
Iteration 8718, Loss: 0.0126, Time: 0.03s
Iteration 8719, Loss: 0.0097, Time: 0.03s
Iteration 8720, Loss: 0.0105, Time: 0.03s
Iteration 8721, Loss: 0.0091, Time: 0.03s
Iteration 8722, Loss: 0.0108, Time: 0.03s
Iteration 8723, Loss: 0.0091, Time: 0.03s
Iteration 8724, Loss: 0.0091, Time: 0.03s
Iteration 8725, Loss: 0.0075, Time: 0.03s
Iteration 8726, Loss: 0.0129, Time: 0.03s
Iteration 8727, Loss: 0.0078, Time: 0.03s
Iteration 8728, Loss: 0.0079, Time: 0.03s
Iteration 8729, Loss: 0.0078, Time: 0.03s
Iteration 8730, Loss: 0.0124, Time: 0.03s
Iteration 8731, Loss: 0.0087, Time: 0.03s
Iteration 8732, Loss: 0.0085, Time: 0.03s
Iteration 8733, Loss: 0.0083, Time: 0.03s
Iteration 8734, Loss: 0.0074, Time: 0.03s
Iteration 8735, Loss: 0.0094, Time: 0.03s
Iteration 8736, Loss: 0.0104, Time: 0.03s
Iteration 8737, Loss: 0.0104, Time: 0.03s
Iteration 8738, Loss: 0.0107, Time: 0.03s
Iteration 8739, Loss: 0.0093, Time: 0.03s
Iteration 8740, Loss: 0.0125, Time: 0.03s
Iteration 8741, Loss: 0.0078, Time: 0.03s
Iteration 8742, Loss: 0.0075, Time: 0.03s
Iteration 8743, Loss: 0.0107, Time: 0.03s
Iteration 8744, Loss: 0.0110, Time: 0.03s
Iteration 8745, Loss: 0.0068, Time: 0.03s
Iteration 8746, Loss: 0.0056, Time: 0.03s
Iteration 8747, Loss: 0.0068, Time: 0.03s
Iteration 8748, Loss: 0.0126, Time: 0.03s
Iteration 8749, Loss: 0.0042, Time: 0.03s
Iteration 8750, Loss: 0.0094, Time: 0.03s
Iteration 8751, Loss: 0.0076, Time: 0.03s
Iteration 8752, Loss: 0.0090, Time: 0.03s
```

```
Iteration 8753, Loss: 0.0127, Time: 0.03s
Iteration 8754, Loss: 0.0076, Time: 0.03s
Iteration 8755, Loss: 0.0138, Time: 0.03s
Iteration 8756, Loss: 0.0103, Time: 0.03s
Iteration 8757, Loss: 0.0073, Time: 0.03s
Iteration 8758, Loss: 0.0121, Time: 0.03s
Iteration 8759, Loss: 0.0106, Time: 0.03s
Iteration 8760, Loss: 0.0155, Time: 0.03s
Iteration 8761, Loss: 0.0099, Time: 0.03s
Iteration 8762, Loss: 0.0076, Time: 0.03s
Iteration 8763, Loss: 0.0081, Time: 0.03s
Iteration 8764, Loss: 0.0082, Time: 0.03s
Iteration 8765, Loss: 0.0117, Time: 0.03s
Iteration 8766, Loss: 0.0100, Time: 0.03s
Iteration 8767, Loss: 0.0099, Time: 0.03s
Iteration 8768, Loss: 0.0099, Time: 0.03s
Iteration 8769, Loss: 0.0142, Time: 0.03s
Iteration 8770, Loss: 0.0093, Time: 0.03s
Iteration 8771, Loss: 0.0066, Time: 0.03s
Iteration 8772, Loss: 0.0072, Time: 0.03s
Iteration 8773, Loss: 0.0093, Time: 0.03s
Iteration 8774, Loss: 0.0103, Time: 0.03s
Iteration 8775, Loss: 0.0110, Time: 0.03s
Iteration 8776, Loss: 0.0154, Time: 0.03s
Iteration 8777, Loss: 0.0120, Time: 0.03s
Iteration 8778, Loss: 0.0110, Time: 0.03s
Iteration 8779, Loss: 0.0085, Time: 0.03s
Iteration 8780, Loss: 0.0084, Time: 0.03s
Iteration 8781, Loss: 0.0101, Time: 0.03s
Iteration 8782, Loss: 0.0089, Time: 0.03s
Iteration 8783, Loss: 0.0141, Time: 0.03s
Iteration 8784, Loss: 0.0067, Time: 0.03s
Iteration 8785, Loss: 0.0070, Time: 0.03s
Iteration 8786, Loss: 0.0124, Time: 0.03s
Iteration 8787, Loss: 0.0089, Time: 0.03s
Iteration 8788, Loss: 0.0055, Time: 0.03s
Iteration 8789, Loss: 0.0138, Time: 0.03s
Iteration 8790, Loss: 0.0140, Time: 0.03s
Iteration 8791, Loss: 0.0100, Time: 0.03s
Iteration 8792, Loss: 0.0105, Time: 0.03s
Iteration 8793, Loss: 0.0147, Time: 0.03s
Iteration 8794, Loss: 0.0109, Time: 0.03s
Iteration 8795, Loss: 0.0093, Time: 0.03s
Iteration 8796, Loss: 0.0108, Time: 0.03s
Iteration 8797, Loss: 0.0101, Time: 0.03s
Iteration 8798, Loss: 0.0098, Time: 0.03s
Iteration 8799, Loss: 0.0169, Time: 0.03s
Iteration 8800, Loss: 0.0121, Time: 0.03s
```

```
Iteration 8800, Loss: 0.0121, Time: 0.03s
Test Loss: 0.0562
Iteration 8801, Loss: 0.0080, Time: 0.03s
Iteration 8802, Loss: 0.0094, Time: 0.03s
Iteration 8803, Loss: 0.0091, Time: 0.03s
Iteration 8804, Loss: 0.0084, Time: 0.03s
Iteration 8805, Loss: 0.0154, Time: 0.03s
Iteration 8806, Loss: 0.0115, Time: 0.03s
Iteration 8807, Loss: 0.0120, Time: 0.03s
Iteration 8808, Loss: 0.0094, Time: 0.03s
Iteration 8809, Loss: 0.0084, Time: 0.03s
Iteration 8810, Loss: 0.0104, Time: 0.03s
Iteration 8811, Loss: 0.0078, Time: 0.03s
Iteration 8812, Loss: 0.0099, Time: 0.03s
Iteration 8813, Loss: 0.0113, Time: 0.03s
Iteration 8814, Loss: 0.0125, Time: 0.03s
Iteration 8815, Loss: 0.0078, Time: 0.03s
Iteration 8816, Loss: 0.0084, Time: 0.03s
Iteration 8817, Loss: 0.0056, Time: 0.03s
Iteration 8818, Loss: 0.0063, Time: 0.03s
Iteration 8819, Loss: 0.0115, Time: 0.03s
Iteration 8820, Loss: 0.0093, Time: 0.03s
Iteration 8821, Loss: 0.0116, Time: 0.03s
Iteration 8822, Loss: 0.0076, Time: 0.03s
Iteration 8823, Loss: 0.0080, Time: 0.03s
Iteration 8824, Loss: 0.0044, Time: 0.03s
Iteration 8825, Loss: 0.0092, Time: 0.03s
Iteration 8826, Loss: 0.0119, Time: 0.03s
Iteration 8827, Loss: 0.0088, Time: 0.03s
Iteration 8828, Loss: 0.0090, Time: 0.03s
Iteration 8829, Loss: 0.0101, Time: 0.03s
Iteration 8830, Loss: 0.0098, Time: 0.03s
Iteration 8831, Loss: 0.0076, Time: 0.03s
Iteration 8832, Loss: 0.0113, Time: 0.03s
Iteration 8833, Loss: 0.0124, Time: 0.03s
Iteration 8834, Loss: 0.0089, Time: 0.03s
Iteration 8835, Loss: 0.0094, Time: 0.03s
Iteration 8836, Loss: 0.0077, Time: 0.03s
Iteration 8837, Loss: 0.0110, Time: 0.03s
Iteration 8838, Loss: 0.0112, Time: 0.03s
Iteration 8839, Loss: 0.0099, Time: 0.03s
Iteration 8840, Loss: 0.0093, Time: 0.03s
Iteration 8841, Loss: 0.0079, Time: 0.03s
Iteration 8842, Loss: 0.0132, Time: 0.03s
Iteration 8843, Loss: 0.0032, Time: 0.03s
Iteration 8844, Loss: 0.0039, Time: 0.03s
Iteration 8845, Loss: 0.0107, Time: 0.03s
Iteration 8846, Loss: 0.0083, Time: 0.03s
```

```
Iteration 8847, Loss: 0.0064, Time: 0.03s
Iteration 8848, Loss: 0.0118, Time: 0.03s
Iteration 8849, Loss: 0.0109, Time: 0.03s
Iteration 8850, Loss: 0.0092, Time: 0.03s
Iteration 8851, Loss: 0.0093, Time: 0.03s
Iteration 8852, Loss: 0.0089, Time: 0.03s
Iteration 8853, Loss: 0.0146, Time: 0.03s
Iteration 8854, Loss: 0.0086, Time: 0.03s
Iteration 8855, Loss: 0.0133, Time: 0.03s
Iteration 8856, Loss: 0.0146, Time: 0.03s
Iteration 8857, Loss: 0.0085, Time: 0.04s
Iteration 8858, Loss: 0.0050, Time: 0.03s
Iteration 8859, Loss: 0.0089, Time: 0.03s
Iteration 8860, Loss: 0.0113, Time: 0.03s
Iteration 8861, Loss: 0.0172, Time: 0.03s
Iteration 8862, Loss: 0.0097, Time: 0.03s
Iteration 8863, Loss: 0.0125, Time: 0.03s
Iteration 8864, Loss: 0.0043, Time: 0.03s
Iteration 8865, Loss: 0.0068, Time: 0.03s
Iteration 8866, Loss: 0.0100, Time: 0.03s
Iteration 8867, Loss: 0.0063, Time: 0.03s
Iteration 8868, Loss: 0.0094, Time: 0.03s
Iteration 8869, Loss: 0.0105, Time: 0.03s
Iteration 8870, Loss: 0.0111, Time: 0.03s
Iteration 8871, Loss: 0.0144, Time: 0.03s
Iteration 8872, Loss: 0.0110, Time: 0.03s
Iteration 8873, Loss: 0.0112, Time: 0.03s
Iteration 8874, Loss: 0.0125, Time: 0.03s
Iteration 8875, Loss: 0.0119, Time: 0.03s
Iteration 8876, Loss: 0.0078, Time: 0.03s
Iteration 8877, Loss: 0.0076, Time: 0.03s
Iteration 8878, Loss: 0.0102, Time: 0.03s
Iteration 8879, Loss: 0.0105, Time: 0.03s
Iteration 8880, Loss: 0.0073, Time: 0.03s
Iteration 8881, Loss: 0.0067, Time: 0.03s
Iteration 8882, Loss: 0.0100, Time: 0.03s
Iteration 8883, Loss: 0.0140, Time: 0.03s
Iteration 8884, Loss: 0.0119, Time: 0.03s
Iteration 8885, Loss: 0.0141, Time: 0.04s
Iteration 8886, Loss: 0.0087, Time: 0.05s
Iteration 8887, Loss: 0.0088, Time: 0.05s
Iteration 8888, Loss: 0.0104, Time: 0.03s
Iteration 8889, Loss: 0.0082, Time: 0.03s
Iteration 8890, Loss: 0.0052, Time: 0.03s
Iteration 8891, Loss: 0.0086, Time: 0.03s
Iteration 8892, Loss: 0.0090, Time: 0.03s
Iteration 8893, Loss: 0.0079, Time: 0.04s
Iteration 8894, Loss: 0.0140, Time: 0.03s
```

```
Iteration 8895, Loss: 0.0099, Time: 0.03s
Iteration 8896, Loss: 0.0082, Time: 0.03s
Iteration 8897, Loss: 0.0117, Time: 0.03s
Iteration 8898, Loss: 0.0101, Time: 0.03s
Iteration 8899, Loss: 0.0094, Time: 0.03s
Iteration 8900, Loss: 0.0074, Time: 0.03s
Iteration 8900, Loss: 0.0074, Time: 0.03s
Test Loss: 0.0419
Iteration 8901, Loss: 0.0115, Time: 0.03s
Iteration 8902, Loss: 0.0094, Time: 0.03s
Iteration 8903, Loss: 0.0133, Time: 0.03s
Iteration 8904, Loss: 0.0126, Time: 0.03s
Iteration 8905, Loss: 0.0094, Time: 0.03s
Iteration 8906, Loss: 0.0113, Time: 0.03s
Iteration 8907, Loss: 0.0082, Time: 0.03s
Iteration 8908, Loss: 0.0173, Time: 0.03s
Iteration 8909, Loss: 0.0105, Time: 0.03s
Iteration 8910, Loss: 0.0107, Time: 0.03s
Iteration 8911, Loss: 0.0084, Time: 0.03s
Iteration 8912, Loss: 0.0113, Time: 0.03s
Iteration 8913, Loss: 0.0107, Time: 0.03s
Iteration 8914, Loss: 0.0085, Time: 0.03s
Iteration 8915, Loss: 0.0082, Time: 0.03s
Iteration 8916, Loss: 0.0125, Time: 0.03s
Iteration 8917, Loss: 0.0066, Time: 0.03s
Iteration 8918, Loss: 0.0115, Time: 0.03s
Iteration 8919, Loss: 0.0113, Time: 0.03s
Iteration 8920, Loss: 0.0084, Time: 0.03s
Iteration 8921, Loss: 0.0129, Time: 0.03s
Iteration 8922, Loss: 0.0133, Time: 0.03s
Iteration 8923, Loss: 0.0096, Time: 0.03s
Iteration 8924, Loss: 0.0105, Time: 0.03s
Iteration 8925, Loss: 0.0066, Time: 0.03s
Iteration 8926, Loss: 0.0106, Time: 0.04s
Iteration 8927, Loss: 0.0117, Time: 0.03s
Iteration 8928, Loss: 0.0110, Time: 0.03s
Iteration 8929, Loss: 0.0105, Time: 0.03s
Iteration 8930, Loss: 0.0082, Time: 0.03s
Iteration 8931, Loss: 0.0068, Time: 0.03s
Iteration 8932, Loss: 0.0107, Time: 0.03s
Iteration 8933, Loss: 0.0127, Time: 0.03s
Iteration 8934, Loss: 0.0083, Time: 0.03s
Iteration 8935, Loss: 0.0115, Time: 0.03s
Iteration 8936, Loss: 0.0102, Time: 0.03s
Iteration 8937, Loss: 0.0099, Time: 0.03s
Iteration 8938, Loss: 0.0150, Time: 0.03s
Iteration 8939, Loss: 0.0078, Time: 0.03s
Iteration 8940, Loss: 0.0090, Time: 0.04s
```

```
Iteration 8941, Loss: 0.0108, Time: 0.03s
Iteration 8942, Loss: 0.0132, Time: 0.03s
Iteration 8943, Loss: 0.0098, Time: 0.03s
Iteration 8944, Loss: 0.0160, Time: 0.03s
Iteration 8945, Loss: 0.0117, Time: 0.03s
Iteration 8946, Loss: 0.0088, Time: 0.03s
Iteration 8947, Loss: 0.0080, Time: 0.03s
Iteration 8948, Loss: 0.0140, Time: 0.03s
Iteration 8949, Loss: 0.0107, Time: 0.03s
Iteration 8950, Loss: 0.0107, Time: 0.03s
Iteration 8951, Loss: 0.0085, Time: 0.03s
Iteration 8952, Loss: 0.0124, Time: 0.03s
Iteration 8953, Loss: 0.0099, Time: 0.03s
Iteration 8954, Loss: 0.0079, Time: 0.03s
Iteration 8955, Loss: 0.0082, Time: 0.03s
Iteration 8956, Loss: 0.0090, Time: 0.03s
Iteration 8957, Loss: 0.0090, Time: 0.03s
Iteration 8958, Loss: 0.0127, Time: 0.03s
Iteration 8959, Loss: 0.0091, Time: 0.03s
Iteration 8960, Loss: 0.0079, Time: 0.03s
Iteration 8961, Loss: 0.0073, Time: 0.03s
Iteration 8962, Loss: 0.0100, Time: 0.03s
Iteration 8963, Loss: 0.0073, Time: 0.03s
Iteration 8964, Loss: 0.0044, Time: 0.03s
Iteration 8965, Loss: 0.0054, Time: 0.03s
Iteration 8966, Loss: 0.0082, Time: 0.03s
Iteration 8967, Loss: 0.0083, Time: 0.03s
Iteration 8968, Loss: 0.0093, Time: 0.03s
Iteration 8969, Loss: 0.0058, Time: 0.03s
Iteration 8970, Loss: 0.0091, Time: 0.03s
Iteration 8971, Loss: 0.0100, Time: 0.03s
Iteration 8972, Loss: 0.0116, Time: 0.03s
Iteration 8973, Loss: 0.0117, Time: 0.03s
Iteration 8974, Loss: 0.0117, Time: 0.03s
Iteration 8975, Loss: 0.0121, Time: 0.03s
Iteration 8976, Loss: 0.0126, Time: 0.03s
Iteration 8977, Loss: 0.0066, Time: 0.03s
Iteration 8978, Loss: 0.0107, Time: 0.03s
Iteration 8979, Loss: 0.0077, Time: 0.03s
Iteration 8980, Loss: 0.0123, Time: 0.03s
Iteration 8981, Loss: 0.0076, Time: 0.03s
Iteration 8982, Loss: 0.0103, Time: 0.03s
Iteration 8983, Loss: 0.0072, Time: 0.03s
Iteration 8984, Loss: 0.0083, Time: 0.03s
Iteration 8985, Loss: 0.0083, Time: 0.03s
Iteration 8986, Loss: 0.0090, Time: 0.03s
Iteration 8987, Loss: 0.0117, Time: 0.03s
Iteration 8988, Loss: 0.0081, Time: 0.03s
```

```
Iteration 8989, Loss: 0.0067, Time: 0.03s
Iteration 8990, Loss: 0.0065, Time: 0.03s
Iteration 8991, Loss: 0.0072, Time: 0.03s
Iteration 8992, Loss: 0.0070, Time: 0.03s
Iteration 8993, Loss: 0.0121, Time: 0.03s
Iteration 8994, Loss: 0.0106, Time: 0.03s
Iteration 8995, Loss: 0.0100, Time: 0.03s
Iteration 8996, Loss: 0.0124, Time: 0.03s
Iteration 8997, Loss: 0.0111, Time: 0.03s
Iteration 8998, Loss: 0.0124, Time: 0.03s
Iteration 8999, Loss: 0.0111, Time: 0.03s
Iteration 9000, Loss: 0.0065, Time: 0.03s
Iteration 9000, Loss: 0.0065, Time: 0.03s
Test Loss: 0.0330
Iteration 9001, Loss: 0.0089, Time: 0.03s
Iteration 9002, Loss: 0.0118, Time: 0.03s
Iteration 9003, Loss: 0.0093, Time: 0.03s
Iteration 9004, Loss: 0.0137, Time: 0.03s
Iteration 9005, Loss: 0.0059, Time: 0.03s
Iteration 9006, Loss: 0.0080, Time: 0.03s
Iteration 9007, Loss: 0.0106, Time: 0.03s
Iteration 9008, Loss: 0.0101, Time: 0.03s
Iteration 9009, Loss: 0.0144, Time: 0.03s
Iteration 9010, Loss: 0.0064, Time: 0.03s
Iteration 9011, Loss: 0.0079, Time: 0.03s
Iteration 9012, Loss: 0.0172, Time: 0.03s
Iteration 9013, Loss: 0.0114, Time: 0.03s
Iteration 9014, Loss: 0.0074, Time: 0.03s
Iteration 9015, Loss: 0.0068, Time: 0.03s
Iteration 9016, Loss: 0.0103, Time: 0.03s
Iteration 9017, Loss: 0.0078, Time: 0.03s
Iteration 9018, Loss: 0.0096, Time: 0.03s
Iteration 9019, Loss: 0.0073, Time: 0.03s
Iteration 9020, Loss: 0.0052, Time: 0.03s
Iteration 9021, Loss: 0.0131, Time: 0.03s
Iteration 9022, Loss: 0.0092, Time: 0.03s
Iteration 9023, Loss: 0.0128, Time: 0.03s
Iteration 9024, Loss: 0.0090, Time: 0.03s
Iteration 9025, Loss: 0.0083, Time: 0.03s
Iteration 9026, Loss: 0.0049, Time: 0.03s
Iteration 9027, Loss: 0.0072, Time: 0.04s
Iteration 9028, Loss: 0.0096, Time: 0.03s
Iteration 9029, Loss: 0.0141, Time: 0.03s
Iteration 9030, Loss: 0.0204, Time: 0.03s
Iteration 9031, Loss: 0.0086, Time: 0.03s
Iteration 9032, Loss: 0.0076, Time: 0.03s
Iteration 9033, Loss: 0.0108, Time: 0.03s
Iteration 9034, Loss: 0.0067, Time: 0.03s
```

```
Iteration 9035, Loss: 0.0097, Time: 0.03s
Iteration 9036, Loss: 0.0103, Time: 0.03s
Iteration 9037, Loss: 0.0088, Time: 0.03s
Iteration 9038, Loss: 0.0129, Time: 0.03s
Iteration 9039, Loss: 0.0089, Time: 0.03s
Iteration 9040, Loss: 0.0142, Time: 0.03s
Iteration 9041, Loss: 0.0108, Time: 0.03s
Iteration 9042, Loss: 0.0086, Time: 0.03s
Iteration 9043, Loss: 0.0081, Time: 0.03s
Iteration 9044, Loss: 0.0098, Time: 0.03s
Iteration 9045, Loss: 0.0121, Time: 0.03s
Iteration 9046, Loss: 0.0071, Time: 0.03s
Iteration 9047, Loss: 0.0115, Time: 0.03s
Iteration 9048, Loss: 0.0095, Time: 0.03s
Iteration 9049, Loss: 0.0073, Time: 0.03s
Iteration 9050, Loss: 0.0074, Time: 0.03s
Iteration 9051, Loss: 0.0091, Time: 0.03s
Iteration 9052, Loss: 0.0134, Time: 0.03s
Iteration 9053, Loss: 0.0143, Time: 0.03s
Iteration 9054, Loss: 0.0103, Time: 0.03s
Iteration 9055, Loss: 0.0090, Time: 0.03s
Iteration 9056, Loss: 0.0108, Time: 0.03s
Iteration 9057, Loss: 0.0086, Time: 0.03s
Iteration 9058, Loss: 0.0023, Time: 0.03s
Iteration 9059, Loss: 0.0180, Time: 0.03s
Iteration 9060, Loss: 0.0121, Time: 0.03s
Iteration 9061, Loss: 0.0080, Time: 0.03s
Iteration 9062, Loss: 0.0072, Time: 0.03s
Iteration 9063, Loss: 0.0040, Time: 0.03s
Iteration 9064, Loss: 0.0075, Time: 0.03s
Iteration 9065, Loss: 0.0103, Time: 0.03s
Iteration 9066, Loss: 0.0066, Time: 0.03s
Iteration 9067, Loss: 0.0072, Time: 0.03s
Iteration 9068, Loss: 0.0120, Time: 0.03s
Iteration 9069, Loss: 0.0103, Time: 0.03s
Iteration 9070, Loss: 0.0055, Time: 0.03s
Iteration 9071, Loss: 0.0079, Time: 0.03s
Iteration 9072, Loss: 0.0087, Time: 0.03s
Iteration 9073, Loss: 0.0075, Time: 0.03s
Iteration 9074, Loss: 0.0094, Time: 0.03s
Iteration 9075, Loss: 0.0112, Time: 0.03s
Iteration 9076, Loss: 0.0130, Time: 0.03s
Iteration 9077, Loss: 0.0103, Time: 0.03s
Iteration 9078, Loss: 0.0127, Time: 0.03s
Iteration 9079, Loss: 0.0113, Time: 0.03s
Iteration 9080, Loss: 0.0134, Time: 0.03s
Iteration 9081, Loss: 0.0102, Time: 0.03s
Iteration 9082, Loss: 0.0084, Time: 0.03s
```

```
Iteration 9083, Loss: 0.0078, Time: 0.03s
Iteration 9084, Loss: 0.0040, Time: 0.03s
Iteration 9085, Loss: 0.0078, Time: 0.03s
Iteration 9086, Loss: 0.0088, Time: 0.03s
Iteration 9087, Loss: 0.0059, Time: 0.03s
Iteration 9088, Loss: 0.0110, Time: 0.03s
Iteration 9089, Loss: 0.0067, Time: 0.03s
Iteration 9090, Loss: 0.0116, Time: 0.03s
Iteration 9091, Loss: 0.0085, Time: 0.03s
Iteration 9092, Loss: 0.0122, Time: 0.03s
Iteration 9093, Loss: 0.0109, Time: 0.03s
Iteration 9094, Loss: 0.0130, Time: 0.03s
Iteration 9095, Loss: 0.0163, Time: 0.03s
Iteration 9096, Loss: 0.0054, Time: 0.03s
Iteration 9097, Loss: 0.0087, Time: 0.03s
Iteration 9098, Loss: 0.0105, Time: 0.03s
Iteration 9099, Loss: 0.0095, Time: 0.03s
Iteration 9100, Loss: 0.0176, Time: 0.03s
Iteration 9100, Loss: 0.0176, Time: 0.03s
Test Loss: 0.0281
Iteration 9101, Loss: 0.0102, Time: 0.03s
Iteration 9102, Loss: 0.0092, Time: 0.03s
Iteration 9103, Loss: 0.0058, Time: 0.03s
Iteration 9104, Loss: 0.0147, Time: 0.03s
Iteration 9105, Loss: 0.0093, Time: 0.03s
Iteration 9106, Loss: 0.0064, Time: 0.03s
Iteration 9107, Loss: 0.0076, Time: 0.03s
Iteration 9108, Loss: 0.0110, Time: 0.03s
Iteration 9109, Loss: 0.0158, Time: 0.03s
Iteration 9110, Loss: 0.0081, Time: 0.03s
Iteration 9111, Loss: 0.0151, Time: 0.03s
Iteration 9112, Loss: 0.0091, Time: 0.03s
Iteration 9113, Loss: 0.0092, Time: 0.03s
Iteration 9114, Loss: 0.0120, Time: 0.03s
Iteration 9115, Loss: 0.0102, Time: 0.03s
Iteration 9116, Loss: 0.0119, Time: 0.03s
Iteration 9117, Loss: 0.0060, Time: 0.03s
Iteration 9118, Loss: 0.0111, Time: 0.03s
Iteration 9119, Loss: 0.0106, Time: 0.03s
Iteration 9120, Loss: 0.0110, Time: 0.03s
Iteration 9121, Loss: 0.0120, Time: 0.03s
Iteration 9122, Loss: 0.0117, Time: 0.03s
Iteration 9123, Loss: 0.0056, Time: 0.03s
Iteration 9124, Loss: 0.0073, Time: 0.03s
Iteration 9125, Loss: 0.0113, Time: 0.03s
Iteration 9126, Loss: 0.0088, Time: 0.03s
Iteration 9127, Loss: 0.0067, Time: 0.03s
Iteration 9128, Loss: 0.0092, Time: 0.03s
```

```
Iteration 9129, Loss: 0.0116, Time: 0.03s
Iteration 9130, Loss: 0.0094, Time: 0.03s
Iteration 9131, Loss: 0.0067, Time: 0.03s
Iteration 9132, Loss: 0.0113, Time: 0.03s
Iteration 9133, Loss: 0.0071, Time: 0.03s
Iteration 9134, Loss: 0.0114, Time: 0.03s
Iteration 9135, Loss: 0.0031, Time: 0.03s
Iteration 9136, Loss: 0.0089, Time: 0.03s
Iteration 9137, Loss: 0.0123, Time: 0.03s
Iteration 9138, Loss: 0.0105, Time: 0.03s
Iteration 9139, Loss: 0.0086, Time: 0.03s
Iteration 9140, Loss: 0.0068, Time: 0.03s
Iteration 9141, Loss: 0.0067, Time: 0.03s
Iteration 9142, Loss: 0.0105, Time: 0.03s
Iteration 9143, Loss: 0.0103, Time: 0.03s
Iteration 9144, Loss: 0.0123, Time: 0.03s
Iteration 9145, Loss: 0.0059, Time: 0.03s
Iteration 9146, Loss: 0.0121, Time: 0.03s
Iteration 9147, Loss: 0.0073, Time: 0.03s
Iteration 9148, Loss: 0.0127, Time: 0.03s
Iteration 9149, Loss: 0.0115, Time: 0.03s
Iteration 9150, Loss: 0.0065, Time: 0.03s
Iteration 9151, Loss: 0.0092, Time: 0.03s
Iteration 9152, Loss: 0.0097, Time: 0.03s
Iteration 9153, Loss: 0.0073, Time: 0.03s
Iteration 9154, Loss: 0.0109, Time: 0.03s
Iteration 9155, Loss: 0.0116, Time: 0.03s
Iteration 9156, Loss: 0.0090, Time: 0.03s
Iteration 9157, Loss: 0.0074, Time: 0.03s
Iteration 9158, Loss: 0.0079, Time: 0.03s
Iteration 9159, Loss: 0.0068, Time: 0.03s
Iteration 9160, Loss: 0.0083, Time: 0.03s
Iteration 9161, Loss: 0.0097, Time: 0.03s
Iteration 9162, Loss: 0.0063, Time: 0.03s
Iteration 9163, Loss: 0.0120, Time: 0.03s
Iteration 9164, Loss: 0.0101, Time: 0.03s
Iteration 9165, Loss: 0.0086, Time: 0.03s
Iteration 9166, Loss: 0.0119, Time: 0.03s
Iteration 9167, Loss: 0.0088, Time: 0.03s
Iteration 9168, Loss: 0.0086, Time: 0.03s
Iteration 9169, Loss: 0.0100, Time: 0.03s
Iteration 9170, Loss: 0.0101, Time: 0.03s
Iteration 9171, Loss: 0.0078, Time: 0.03s
Iteration 9172, Loss: 0.0100, Time: 0.03s
Iteration 9173, Loss: 0.0086, Time: 0.03s
Iteration 9174, Loss: 0.0087, Time: 0.03s
Iteration 9175, Loss: 0.0144, Time: 0.03s
Iteration 9176, Loss: 0.0095, Time: 0.03s
```

```
Iteration 9177, Loss: 0.0067, Time: 0.03s
Iteration 9178, Loss: 0.0130, Time: 0.03s
Iteration 9179, Loss: 0.0089, Time: 0.03s
Iteration 9180, Loss: 0.0093, Time: 0.03s
Iteration 9181, Loss: 0.0106, Time: 0.03s
Iteration 9182, Loss: 0.0124, Time: 0.03s
Iteration 9183, Loss: 0.0124, Time: 0.03s
Iteration 9184, Loss: 0.0081, Time: 0.03s
Iteration 9185, Loss: 0.0084, Time: 0.03s
Iteration 9186, Loss: 0.0129, Time: 0.03s
Iteration 9187, Loss: 0.0064, Time: 0.03s
Iteration 9188, Loss: 0.0105, Time: 0.03s
Iteration 9189, Loss: 0.0080, Time: 0.03s
Iteration 9190, Loss: 0.0101, Time: 0.03s
Iteration 9191, Loss: 0.0081, Time: 0.03s
Iteration 9192, Loss: 0.0131, Time: 0.03s
Iteration 9193, Loss: 0.0092, Time: 0.03s
Iteration 9194, Loss: 0.0095, Time: 0.03s
Iteration 9195, Loss: 0.0108, Time: 0.03s
Iteration 9196, Loss: 0.0089, Time: 0.03s
Iteration 9197, Loss: 0.0064, Time: 0.03s
Iteration 9198, Loss: 0.0088, Time: 0.03s
Iteration 9199, Loss: 0.0094, Time: 0.03s
Iteration 9200, Loss: 0.0036, Time: 0.03s
Iteration 9200, Loss: 0.0036, Time: 0.03s
Test Loss: 0.0643
Iteration 9201, Loss: 0.0131, Time: 0.03s
Iteration 9202, Loss: 0.0092, Time: 0.03s
Iteration 9203, Loss: 0.0093, Time: 0.03s
Iteration 9204, Loss: 0.0125, Time: 0.03s
Iteration 9205, Loss: 0.0099, Time: 0.03s
Iteration 9206, Loss: 0.0113, Time: 0.03s
Iteration 9207, Loss: 0.0105, Time: 0.03s
Iteration 9208, Loss: 0.0070, Time: 0.03s
Iteration 9209, Loss: 0.0084, Time: 0.03s
Iteration 9210, Loss: 0.0101, Time: 0.03s
Iteration 9211, Loss: 0.0107, Time: 0.03s
Iteration 9212, Loss: 0.0066, Time: 0.03s
Iteration 9213, Loss: 0.0117, Time: 0.03s
Iteration 9214, Loss: 0.0070, Time: 0.03s
Iteration 9215, Loss: 0.0069, Time: 0.03s
Iteration 9216, Loss: 0.0093, Time: 0.03s
Iteration 9217, Loss: 0.0105, Time: 0.03s
Iteration 9218, Loss: 0.0070, Time: 0.03s
Iteration 9219, Loss: 0.0099, Time: 0.03s
Iteration 9220, Loss: 0.0087, Time: 0.03s
Iteration 9221, Loss: 0.0092, Time: 0.03s
Iteration 9222, Loss: 0.0064, Time: 0.03s
```

```
Iteration 9223, Loss: 0.0070, Time: 0.03s
Iteration 9224, Loss: 0.0100, Time: 0.03s
Iteration 9225, Loss: 0.0093, Time: 0.03s
Iteration 9226, Loss: 0.0091, Time: 0.03s
Iteration 9227, Loss: 0.0098, Time: 0.03s
Iteration 9228, Loss: 0.0084, Time: 0.03s
Iteration 9229, Loss: 0.0127, Time: 0.03s
Iteration 9230, Loss: 0.0062, Time: 0.03s
Iteration 9231, Loss: 0.0149, Time: 0.03s
Iteration 9232, Loss: 0.0108, Time: 0.03s
Iteration 9233, Loss: 0.0041, Time: 0.03s
Iteration 9234, Loss: 0.0111, Time: 0.03s
Iteration 9235, Loss: 0.0083, Time: 0.03s
Iteration 9236, Loss: 0.0097, Time: 0.03s
Iteration 9237, Loss: 0.0157, Time: 0.03s
Iteration 9238, Loss: 0.0094, Time: 0.03s
Iteration 9239, Loss: 0.0073, Time: 0.03s
Iteration 9240, Loss: 0.0072, Time: 0.03s
Iteration 9241, Loss: 0.0104, Time: 0.03s
Iteration 9242, Loss: 0.0119, Time: 0.03s
Iteration 9243, Loss: 0.0076, Time: 0.03s
Iteration 9244, Loss: 0.0100, Time: 0.03s
Iteration 9245, Loss: 0.0085, Time: 0.03s
Iteration 9246, Loss: 0.0081, Time: 0.03s
Iteration 9247, Loss: 0.0099, Time: 0.03s
Iteration 9248, Loss: 0.0072, Time: 0.03s
Iteration 9249, Loss: 0.0069, Time: 0.03s
Iteration 9250, Loss: 0.0096, Time: 0.03s
Iteration 9251, Loss: 0.0057, Time: 0.03s
Iteration 9252, Loss: 0.0104, Time: 0.03s
Iteration 9253, Loss: 0.0086, Time: 0.03s
Iteration 9254, Loss: 0.0101, Time: 0.03s
Iteration 9255, Loss: 0.0140, Time: 0.03s
Iteration 9256, Loss: 0.0068, Time: 0.03s
Iteration 9257, Loss: 0.0063, Time: 0.03s
Iteration 9258, Loss: 0.0066, Time: 0.03s
Iteration 9259, Loss: 0.0150, Time: 0.03s
Iteration 9260, Loss: 0.0078, Time: 0.03s
Iteration 9261, Loss: 0.0106, Time: 0.03s
Iteration 9262, Loss: 0.0097, Time: 0.03s
Iteration 9263, Loss: 0.0093, Time: 0.03s
Iteration 9264, Loss: 0.0105, Time: 0.03s
Iteration 9265, Loss: 0.0061, Time: 0.03s
Iteration 9266, Loss: 0.0078, Time: 0.03s
Iteration 9267, Loss: 0.0095, Time: 0.03s
Iteration 9268, Loss: 0.0059, Time: 0.03s
Iteration 9269, Loss: 0.0116, Time: 0.03s
Iteration 9270, Loss: 0.0112, Time: 0.03s
```

```
Iteration 9271, Loss: 0.0125, Time: 0.03s
Iteration 9272, Loss: 0.0088, Time: 0.03s
Iteration 9273, Loss: 0.0092, Time: 0.03s
Iteration 9274, Loss: 0.0123, Time: 0.03s
Iteration 9275, Loss: 0.0052, Time: 0.03s
Iteration 9276, Loss: 0.0138, Time: 0.03s
Iteration 9277, Loss: 0.0079, Time: 0.03s
Iteration 9278, Loss: 0.0060, Time: 0.03s
Iteration 9279, Loss: 0.0096, Time: 0.03s
Iteration 9280, Loss: 0.0074, Time: 0.03s
Iteration 9281, Loss: 0.0073, Time: 0.03s
Iteration 9282, Loss: 0.0063, Time: 0.03s
Iteration 9283, Loss: 0.0085, Time: 0.03s
Iteration 9284, Loss: 0.0104, Time: 0.03s
Iteration 9285, Loss: 0.0051, Time: 0.03s
Iteration 9286, Loss: 0.0074, Time: 0.03s
Iteration 9287, Loss: 0.0103, Time: 0.03s
Iteration 9288, Loss: 0.0074, Time: 0.03s
Iteration 9289, Loss: 0.0075, Time: 0.03s
Iteration 9290, Loss: 0.0079, Time: 0.03s
Iteration 9291, Loss: 0.0075, Time: 0.03s
Iteration 9292, Loss: 0.0099, Time: 0.03s
Iteration 9293, Loss: 0.0074, Time: 0.03s
Iteration 9294, Loss: 0.0084, Time: 0.03s
Iteration 9295, Loss: 0.0106, Time: 0.03s
Iteration 9296, Loss: 0.0112, Time: 0.03s
Iteration 9297, Loss: 0.0081, Time: -0.12s
Iteration 9298, Loss: 0.0118, Time: 0.03s
Iteration 9299, Loss: 0.0086, Time: 0.03s
Iteration 9300, Loss: 0.0125, Time: 0.03s
Iteration 9300, Loss: 0.0125, Time: 0.03s
Test Loss: 0.0402
Iteration 9301, Loss: 0.0091, Time: 0.03s
Iteration 9302, Loss: 0.0076, Time: 0.03s
Iteration 9303, Loss: 0.0079, Time: 0.03s
Iteration 9304, Loss: 0.0078, Time: 0.03s
Iteration 9305, Loss: 0.0088, Time: 0.03s
Iteration 9306, Loss: 0.0100, Time: 0.03s
Iteration 9307, Loss: 0.0163, Time: 0.03s
Iteration 9308, Loss: 0.0084, Time: 0.03s
Iteration 9309, Loss: 0.0101, Time: 0.03s
Iteration 9310, Loss: 0.0119, Time: 0.03s
Iteration 9311, Loss: 0.0126, Time: 0.03s
Iteration 9312, Loss: 0.0118, Time: 0.03s
Iteration 9313, Loss: 0.0103, Time: 0.03s
Iteration 9314, Loss: 0.0168, Time: 0.03s
Iteration 9315, Loss: 0.0132, Time: 0.03s
Iteration 9316, Loss: 0.0050, Time: 0.03s
```

```
Iteration 9317, Loss: 0.0093, Time: 0.03s
Iteration 9318, Loss: 0.0090, Time: 0.03s
Iteration 9319, Loss: 0.0093, Time: 0.03s
Iteration 9320, Loss: 0.0111, Time: 0.03s
Iteration 9321, Loss: 0.0099, Time: 0.03s
Iteration 9322, Loss: 0.0083, Time: 0.03s
Iteration 9323, Loss: 0.0095, Time: 0.03s
Iteration 9324, Loss: 0.0099, Time: 0.03s
Iteration 9325, Loss: 0.0130, Time: 0.03s
Iteration 9326, Loss: 0.0142, Time: 0.03s
Iteration 9327, Loss: 0.0064, Time: 0.03s
Iteration 9328, Loss: 0.0094, Time: 0.03s
Iteration 9329, Loss: 0.0098, Time: 0.03s
Iteration 9330, Loss: 0.0096, Time: 0.03s
Iteration 9331, Loss: 0.0089, Time: 0.03s
Iteration 9332, Loss: 0.0141, Time: 0.03s
Iteration 9333, Loss: 0.0100, Time: 0.03s
Iteration 9334, Loss: 0.0066, Time: 0.03s
Iteration 9335, Loss: 0.0147, Time: 0.03s
Iteration 9336, Loss: 0.0070, Time: 0.03s
Iteration 9337, Loss: 0.0058, Time: 0.03s
Iteration 9338, Loss: 0.0096, Time: 0.03s
Iteration 9339, Loss: 0.0125, Time: 0.03s
Iteration 9340, Loss: 0.0101, Time: 0.03s
Iteration 9341, Loss: 0.0130, Time: 0.03s
Iteration 9342, Loss: 0.0090, Time: 0.03s
Iteration 9343, Loss: 0.0058, Time: 0.03s
Iteration 9344, Loss: 0.0066, Time: 0.03s
Iteration 9345, Loss: 0.0087, Time: 0.03s
Iteration 9346, Loss: 0.0128, Time: 0.03s
Iteration 9347, Loss: 0.0080, Time: 0.03s
Iteration 9348, Loss: 0.0080, Time: 0.03s
Iteration 9349, Loss: 0.0134, Time: 0.03s
Iteration 9350, Loss: 0.0088, Time: 0.03s
Iteration 9351, Loss: 0.0100, Time: 0.03s
Iteration 9352, Loss: 0.0110, Time: 0.04s
Iteration 9353, Loss: 0.0105, Time: 0.03s
Iteration 9354, Loss: 0.0075, Time: 0.03s
Iteration 9355, Loss: 0.0053, Time: 0.03s
Iteration 9356, Loss: 0.0093, Time: 0.04s
Iteration 9357, Loss: 0.0112, Time: 0.03s
Iteration 9358, Loss: 0.0086, Time: 0.03s
Iteration 9359, Loss: 0.0088, Time: 0.03s
Iteration 9360, Loss: 0.0068, Time: 0.03s
Iteration 9361, Loss: 0.0134, Time: 0.03s
Iteration 9362, Loss: 0.0111, Time: 0.03s
Iteration 9363, Loss: 0.0062, Time: 0.03s
Iteration 9364, Loss: 0.0083, Time: 0.03s
```

```
Iteration 9365, Loss: 0.0091, Time: 0.03s
Iteration 9366, Loss: 0.0075, Time: 0.03s
Iteration 9367, Loss: 0.0077, Time: 0.03s
Iteration 9368, Loss: 0.0109, Time: 0.03s
Iteration 9369, Loss: 0.0136, Time: 0.03s
Iteration 9370, Loss: 0.0088, Time: 0.03s
Iteration 9371, Loss: 0.0095, Time: 0.03s
Iteration 9372, Loss: 0.0127, Time: 0.03s
Iteration 9373, Loss: 0.0133, Time: 0.03s
Iteration 9374, Loss: 0.0099, Time: 0.03s
Iteration 9375, Loss: 0.0112, Time: 0.03s
Iteration 9376, Loss: 0.0134, Time: 0.03s
Iteration 9377, Loss: 0.0108, Time: 0.03s
Iteration 9378, Loss: 0.0086, Time: 0.03s
Iteration 9379, Loss: 0.0070, Time: 0.03s
Iteration 9380, Loss: 0.0090, Time: 0.03s
Iteration 9381, Loss: 0.0122, Time: 0.03s
Iteration 9382, Loss: 0.0135, Time: 0.03s
Iteration 9383, Loss: 0.0097, Time: 0.03s
Iteration 9384, Loss: 0.0089, Time: 0.03s
Iteration 9385, Loss: 0.0117, Time: 0.03s
Iteration 9386, Loss: 0.0084, Time: 0.03s
Iteration 9387, Loss: 0.0066, Time: 0.03s
Iteration 9388, Loss: 0.0132, Time: 0.03s
Iteration 9389, Loss: 0.0077, Time: 0.03s
Iteration 9390, Loss: 0.0094, Time: 0.03s
Iteration 9391, Loss: 0.0107, Time: 0.03s
Iteration 9392, Loss: 0.0091, Time: 0.03s
Iteration 9393, Loss: 0.0059, Time: 0.03s
Iteration 9394, Loss: 0.0095, Time: 0.03s
Iteration 9395, Loss: 0.0058, Time: 0.03s
Iteration 9396, Loss: 0.0067, Time: 0.03s
Iteration 9397, Loss: 0.0098, Time: 0.03s
Iteration 9398, Loss: 0.0107, Time: 0.03s
Iteration 9399, Loss: 0.0086, Time: 0.03s
Iteration 9400, Loss: 0.0099, Time: 0.03s
Iteration 9400, Loss: 0.0099, Time: 0.03s
Test Loss: 0.0519
Iteration 9401, Loss: 0.0098, Time: 0.03s
Iteration 9402, Loss: 0.0079, Time: 0.03s
Iteration 9403, Loss: 0.0110, Time: 0.03s
Iteration 9404, Loss: 0.0088, Time: 0.03s
Iteration 9405, Loss: 0.0071, Time: 0.03s
Iteration 9406, Loss: 0.0104, Time: 0.03s
Iteration 9407, Loss: 0.0090, Time: 0.03s
Iteration 9408, Loss: 0.0101, Time: 0.03s
Iteration 9409, Loss: 0.0067, Time: 0.03s
Iteration 9410, Loss: 0.0052, Time: 0.03s
```

```
Iteration 9411, Loss: 0.0138, Time: 0.03s
Iteration 9412, Loss: 0.0137, Time: 0.03s
Iteration 9413, Loss: 0.0103, Time: 0.03s
Iteration 9414, Loss: 0.0117, Time: 0.03s
Iteration 9415, Loss: 0.0071, Time: 0.03s
Iteration 9416, Loss: 0.0111, Time: 0.03s
Iteration 9417, Loss: 0.0116, Time: 0.03s
Iteration 9418, Loss: 0.0088, Time: 0.03s
Iteration 9419, Loss: 0.0092, Time: 0.03s
Iteration 9420, Loss: 0.0130, Time: 0.03s
Iteration 9421, Loss: 0.0084, Time: 0.03s
Iteration 9422, Loss: 0.0074, Time: 0.03s
Iteration 9423, Loss: 0.0076, Time: 0.03s
Iteration 9424, Loss: 0.0109, Time: 0.03s
Iteration 9425, Loss: 0.0124, Time: 0.03s
Iteration 9426, Loss: 0.0122, Time: 0.03s
Iteration 9427, Loss: 0.0106, Time: 0.03s
Iteration 9428, Loss: 0.0118, Time: 0.03s
Iteration 9429, Loss: 0.0128, Time: 0.03s
Iteration 9430, Loss: 0.0146, Time: 0.03s
Iteration 9431, Loss: 0.0080, Time: 0.03s
Iteration 9432, Loss: 0.0121, Time: 0.03s
Iteration 9433, Loss: 0.0059, Time: 0.03s
Iteration 9434, Loss: 0.0078, Time: 0.03s
Iteration 9435, Loss: 0.0117, Time: 0.03s
Iteration 9436, Loss: 0.0080, Time: 0.03s
Iteration 9437, Loss: 0.0138, Time: 0.03s
Iteration 9438, Loss: 0.0076, Time: 0.03s
Iteration 9439, Loss: 0.0103, Time: 0.03s
Iteration 9440, Loss: 0.0087, Time: 0.03s
Iteration 9441, Loss: 0.0105, Time: 0.03s
Iteration 9442, Loss: 0.0082, Time: 0.03s
Iteration 9443, Loss: 0.0115, Time: 0.03s
Iteration 9444, Loss: 0.0065, Time: 0.03s
Iteration 9445, Loss: 0.0065, Time: 0.03s
Iteration 9446, Loss: 0.0100, Time: 0.03s
Iteration 9447, Loss: 0.0099, Time: 0.03s
Iteration 9448, Loss: 0.0100, Time: 0.03s
Iteration 9449, Loss: 0.0115, Time: 0.03s
Iteration 9450, Loss: 0.0089, Time: 0.03s
Iteration 9451, Loss: 0.0108, Time: 0.03s
Iteration 9452, Loss: 0.0090, Time: 0.03s
Iteration 9453, Loss: 0.0061, Time: 0.04s
Iteration 9454, Loss: 0.0068, Time: 0.03s
Iteration 9455, Loss: 0.0111, Time: 0.03s
Iteration 9456, Loss: 0.0128, Time: 0.03s
Iteration 9457, Loss: 0.0051, Time: 0.03s
Iteration 9458, Loss: 0.0078, Time: 0.03s
```

```
Iteration 9459, Loss: 0.0080, Time: 0.03s
Iteration 9460, Loss: 0.0097, Time: 0.03s
Iteration 9461, Loss: 0.0069, Time: 0.03s
Iteration 9462, Loss: 0.0136, Time: 0.03s
Iteration 9463, Loss: 0.0098, Time: 0.03s
Iteration 9464, Loss: 0.0112, Time: 0.03s
Iteration 9465, Loss: 0.0134, Time: 0.03s
Iteration 9466, Loss: 0.0106, Time: 0.03s
Iteration 9467, Loss: 0.0101, Time: 0.03s
Iteration 9468, Loss: 0.0088, Time: 0.03s
Iteration 9469, Loss: 0.0107, Time: 0.03s
Iteration 9470, Loss: 0.0092, Time: 0.03s
Iteration 9471, Loss: 0.0082, Time: 0.03s
Iteration 9472, Loss: 0.0130, Time: 0.03s
Iteration 9473, Loss: 0.0130, Time: 0.03s
Iteration 9474, Loss: 0.0106, Time: 0.03s
Iteration 9475, Loss: 0.0130, Time: 0.03s
Iteration 9476, Loss: 0.0092, Time: 0.03s
Iteration 9477, Loss: 0.0176, Time: 0.03s
Iteration 9478, Loss: 0.0114, Time: 0.03s
Iteration 9479, Loss: 0.0102, Time: 0.03s
Iteration 9480, Loss: 0.0099, Time: 0.03s
Iteration 9481, Loss: 0.0094, Time: 0.03s
Iteration 9482, Loss: 0.0080, Time: 0.03s
Iteration 9483, Loss: 0.0109, Time: 0.03s
Iteration 9484, Loss: 0.0112, Time: 0.03s
Iteration 9485, Loss: 0.0112, Time: 0.03s
Iteration 9486, Loss: 0.0122, Time: 0.03s
Iteration 9487, Loss: 0.0087, Time: 0.03s
Iteration 9488, Loss: 0.0037, Time: 0.03s
Iteration 9489, Loss: 0.0083, Time: 0.03s
Iteration 9490, Loss: 0.0098, Time: 0.03s
Iteration 9491, Loss: 0.0070, Time: 0.03s
Iteration 9492, Loss: 0.0105, Time: 0.03s
Iteration 9493, Loss: 0.0090, Time: 0.03s
Iteration 9494, Loss: 0.0066, Time: 0.03s
Iteration 9495, Loss: 0.0088, Time: 0.03s
Iteration 9496, Loss: 0.0105, Time: 0.03s
Iteration 9497, Loss: 0.0123, Time: 0.03s
Iteration 9498, Loss: 0.0069, Time: 0.03s
Iteration 9499, Loss: 0.0123, Time: 0.03s
Iteration 9500, Loss: 0.0109, Time: 0.03s
Iteration 9500, Loss: 0.0109, Time: 0.03s
Test Loss: 0.0600
Iteration 9501, Loss: 0.0079, Time: 0.03s
Iteration 9502, Loss: 0.0108, Time: 0.03s
Iteration 9503, Loss: 0.0082, Time: 0.03s
Iteration 9504, Loss: 0.0066, Time: 0.03s
```

```
Iteration 9505, Loss: 0.0068, Time: 0.03s
Iteration 9506, Loss: 0.0084, Time: 0.03s
Iteration 9507, Loss: 0.0099, Time: 0.03s
Iteration 9508, Loss: 0.0112, Time: 0.03s
Iteration 9509, Loss: 0.0049, Time: 0.03s
Iteration 9510, Loss: 0.0068, Time: 0.03s
Iteration 9511, Loss: 0.0137, Time: 0.03s
Iteration 9512, Loss: 0.0059, Time: 0.03s
Iteration 9513, Loss: 0.0081, Time: 0.03s
Iteration 9514, Loss: 0.0117, Time: 0.03s
Iteration 9515, Loss: 0.0064, Time: 0.03s
Iteration 9516, Loss: 0.0106, Time: 0.03s
Iteration 9517, Loss: 0.0108, Time: 0.03s
Iteration 9518, Loss: 0.0107, Time: 0.03s
Iteration 9519, Loss: 0.0070, Time: 0.03s
Iteration 9520, Loss: 0.0102, Time: 0.03s
Iteration 9521, Loss: 0.0071, Time: 0.03s
Iteration 9522, Loss: 0.0082, Time: 0.03s
Iteration 9523, Loss: 0.0051, Time: 0.03s
Iteration 9524, Loss: 0.0055, Time: 0.03s
Iteration 9525, Loss: 0.0102, Time: 0.03s
Iteration 9526, Loss: 0.0117, Time: 0.03s
Iteration 9527, Loss: 0.0094, Time: 0.03s
Iteration 9528, Loss: 0.0086, Time: 0.03s
Iteration 9529, Loss: 0.0088, Time: 0.03s
Iteration 9530, Loss: 0.0105, Time: 0.03s
Iteration 9531, Loss: 0.0128, Time: 0.03s
Iteration 9532, Loss: 0.0071, Time: 0.03s
Iteration 9533, Loss: 0.0085, Time: 0.03s
Iteration 9534, Loss: 0.0144, Time: 0.03s
Iteration 9535, Loss: 0.0092, Time: 0.03s
Iteration 9536, Loss: 0.0086, Time: 0.03s
Iteration 9537, Loss: 0.0064, Time: 0.03s
Iteration 9538, Loss: 0.0092, Time: 0.03s
Iteration 9539, Loss: 0.0125, Time: 0.03s
Iteration 9540, Loss: 0.0096, Time: 0.03s
Iteration 9541, Loss: 0.0071, Time: 0.03s
Iteration 9542, Loss: 0.0114, Time: 0.03s
Iteration 9543, Loss: 0.0116, Time: 0.03s
Iteration 9544, Loss: 0.0101, Time: 0.03s
Iteration 9545, Loss: 0.0094, Time: 0.03s
Iteration 9546, Loss: 0.0098, Time: 0.03s
Iteration 9547, Loss: 0.0113, Time: 0.03s
Iteration 9548, Loss: 0.0107, Time: 0.03s
Iteration 9549, Loss: 0.0100, Time: 0.03s
Iteration 9550, Loss: 0.0138, Time: 0.03s
Iteration 9551, Loss: 0.0100, Time: 0.03s
Iteration 9552, Loss: 0.0076, Time: 0.03s
```

```
Iteration 9553, Loss: 0.0115, Time: 0.03s
Iteration 9554, Loss: 0.0053, Time: 0.03s
Iteration 9555, Loss: 0.0092, Time: 0.03s
Iteration 9556, Loss: 0.0108, Time: 0.03s
Iteration 9557, Loss: 0.0096, Time: 0.03s
Iteration 9558, Loss: 0.0078, Time: 0.03s
Iteration 9559, Loss: 0.0071, Time: 0.03s
Iteration 9560, Loss: 0.0068, Time: 0.03s
Iteration 9561, Loss: 0.0088, Time: 0.03s
Iteration 9562, Loss: 0.0068, Time: 0.03s
Iteration 9563, Loss: 0.0119, Time: 0.03s
Iteration 9564, Loss: 0.0114, Time: 0.03s
Iteration 9565, Loss: 0.0097, Time: 0.03s
Iteration 9566, Loss: 0.0085, Time: 0.03s
Iteration 9567, Loss: 0.0126, Time: 0.03s
Iteration 9568, Loss: 0.0067, Time: 0.03s
Iteration 9569, Loss: 0.0119, Time: 0.03s
Iteration 9570, Loss: 0.0116, Time: 0.03s
Iteration 9571, Loss: 0.0062, Time: 0.03s
Iteration 9572, Loss: 0.0102, Time: 0.03s
Iteration 9573, Loss: 0.0070, Time: 0.03s
Iteration 9574, Loss: 0.0159, Time: 0.03s
Iteration 9575, Loss: 0.0093, Time: 0.03s
Iteration 9576, Loss: 0.0081, Time: 0.03s
Iteration 9577, Loss: 0.0091, Time: 0.03s
Iteration 9578, Loss: 0.0056, Time: 0.03s
Iteration 9579, Loss: 0.0065, Time: 0.03s
Iteration 9580, Loss: 0.0083, Time: 0.03s
Iteration 9581, Loss: 0.0091, Time: 0.03s
Iteration 9582, Loss: 0.0100, Time: 0.03s
Iteration 9583, Loss: 0.0137, Time: 0.03s
Iteration 9584, Loss: 0.0171, Time: 0.03s
Iteration 9585, Loss: 0.0090, Time: 0.03s
Iteration 9586, Loss: 0.0069, Time: 0.03s
Iteration 9587, Loss: 0.0118, Time: 0.03s
Iteration 9588, Loss: 0.0106, Time: 0.03s
Iteration 9589, Loss: 0.0086, Time: 0.03s
Iteration 9590, Loss: 0.0074, Time: 0.03s
Iteration 9591, Loss: 0.0123, Time: 0.03s
Iteration 9592, Loss: 0.0104, Time: 0.03s
Iteration 9593, Loss: 0.0153, Time: 0.03s
Iteration 9594, Loss: 0.0098, Time: 0.03s
Iteration 9595, Loss: 0.0153, Time: 0.03s
Iteration 9596, Loss: 0.0124, Time: 0.03s
Iteration 9597, Loss: 0.0131, Time: 0.03s
Iteration 9598, Loss: 0.0152, Time: 0.03s
Iteration 9599, Loss: 0.0087, Time: 0.03s
Iteration 9600, Loss: 0.0085, Time: 0.03s
```

```
Iteration 9600, Loss: 0.0085, Time: 0.03s
Test Loss: 0.0517
Iteration 9601, Loss: 0.0077, Time: 0.03s
Iteration 9602, Loss: 0.0105, Time: 0.03s
Iteration 9603, Loss: 0.0126, Time: 0.03s
Iteration 9604, Loss: 0.0055, Time: 0.03s
Iteration 9605, Loss: 0.0117, Time: 0.03s
Iteration 9606, Loss: 0.0073, Time: 0.03s
Iteration 9607, Loss: 0.0078, Time: 0.03s
Iteration 9608, Loss: 0.0057, Time: 0.03s
Iteration 9609, Loss: 0.0115, Time: 0.03s
Iteration 9610, Loss: 0.0121, Time: 0.03s
Iteration 9611, Loss: 0.0121, Time: 0.03s
Iteration 9612, Loss: 0.0128, Time: 0.03s
Iteration 9613, Loss: 0.0075, Time: 0.03s
Iteration 9614, Loss: 0.0116, Time: 0.03s
Iteration 9615, Loss: 0.0107, Time: 0.03s
Iteration 9616, Loss: 0.0102, Time: 0.03s
Iteration 9617, Loss: 0.0103, Time: 0.03s
Iteration 9618, Loss: 0.0057, Time: 0.03s
Iteration 9619, Loss: 0.0137, Time: 0.03s
Iteration 9620, Loss: 0.0074, Time: 0.03s
Iteration 9621, Loss: 0.0110, Time: 0.03s
Iteration 9622, Loss: 0.0080, Time: 0.03s
Iteration 9623, Loss: 0.0097, Time: 0.03s
Iteration 9624, Loss: 0.0102, Time: 0.03s
Iteration 9625, Loss: 0.0101, Time: 0.03s
Iteration 9626, Loss: 0.0068, Time: 0.03s
Iteration 9627, Loss: 0.0054, Time: 0.03s
Iteration 9628, Loss: 0.0131, Time: 0.03s
Iteration 9629, Loss: 0.0064, Time: 0.03s
Iteration 9630, Loss: 0.0097, Time: 0.03s
Iteration 9631, Loss: 0.0058, Time: 0.03s
Iteration 9632, Loss: 0.0065, Time: 0.03s
Iteration 9633, Loss: 0.0077, Time: 0.03s
Iteration 9634, Loss: 0.0109, Time: 0.03s
Iteration 9635, Loss: 0.0118, Time: 0.03s
Iteration 9636, Loss: 0.0083, Time: 0.03s
Iteration 9637, Loss: 0.0086, Time: 0.03s
Iteration 9638, Loss: 0.0078, Time: 0.03s
Iteration 9639, Loss: 0.0099, Time: 0.03s
Iteration 9640, Loss: 0.0116, Time: 0.03s
Iteration 9641, Loss: 0.0093, Time: 0.03s
Iteration 9642, Loss: 0.0080, Time: 0.03s
Iteration 9643, Loss: 0.0088, Time: 0.03s
Iteration 9644, Loss: 0.0065, Time: 0.03s
Iteration 9645, Loss: 0.0073, Time: 0.03s
Iteration 9646, Loss: 0.0141, Time: 0.03s
```

```
Iteration 9647, Loss: 0.0079, Time: 0.03s
Iteration 9648, Loss: 0.0111, Time: 0.03s
Iteration 9649, Loss: 0.0103, Time: 0.03s
Iteration 9650, Loss: 0.0078, Time: 0.03s
Iteration 9651, Loss: 0.0114, Time: 0.03s
Iteration 9652, Loss: 0.0103, Time: 0.03s
Iteration 9653, Loss: 0.0066, Time: 0.03s
Iteration 9654, Loss: 0.0040, Time: 0.03s
Iteration 9655, Loss: 0.0073, Time: 0.03s
Iteration 9656, Loss: 0.0062, Time: 0.03s
Iteration 9657, Loss: 0.0075, Time: 0.03s
Iteration 9658, Loss: 0.0081, Time: 0.03s
Iteration 9659, Loss: 0.0070, Time: 0.03s
Iteration 9660, Loss: 0.0074, Time: 0.03s
Iteration 9661, Loss: 0.0081, Time: 0.03s
Iteration 9662, Loss: 0.0046, Time: 0.03s
Iteration 9663, Loss: 0.0082, Time: 0.03s
Iteration 9664, Loss: 0.0085, Time: 0.03s
Iteration 9665, Loss: 0.0092, Time: 0.03s
Iteration 9666, Loss: 0.0106, Time: 0.03s
Iteration 9667, Loss: 0.0041, Time: 0.03s
Iteration 9668, Loss: 0.0082, Time: 0.03s
Iteration 9669, Loss: 0.0081, Time: 0.03s
Iteration 9670, Loss: 0.0102, Time: 0.03s
Iteration 9671, Loss: 0.0072, Time: 0.03s
Iteration 9672, Loss: 0.0106, Time: 0.03s
Iteration 9673, Loss: 0.0081, Time: 0.03s
Iteration 9674, Loss: 0.0062, Time: 0.03s
Iteration 9675, Loss: 0.0095, Time: 0.03s
Iteration 9676, Loss: 0.0085, Time: 0.03s
Iteration 9677, Loss: 0.0074, Time: 0.03s
Iteration 9678, Loss: 0.0059, Time: 0.03s
Iteration 9679, Loss: 0.0083, Time: 0.03s
Iteration 9680, Loss: 0.0069, Time: 0.03s
Iteration 9681, Loss: 0.0135, Time: 0.03s
Iteration 9682, Loss: 0.0048, Time: 0.03s
Iteration 9683, Loss: 0.0084, Time: 0.03s
Iteration 9684, Loss: 0.0118, Time: 0.03s
Iteration 9685, Loss: 0.0087, Time: 0.03s
Iteration 9686, Loss: 0.0076, Time: 0.03s
Iteration 9687, Loss: 0.0088, Time: 0.03s
Iteration 9688, Loss: 0.0076, Time: 0.03s
Iteration 9689, Loss: 0.0092, Time: 0.03s
Iteration 9690, Loss: 0.0083, Time: 0.03s
Iteration 9691, Loss: 0.0070, Time: 0.03s
Iteration 9692, Loss: 0.0115, Time: 0.03s
Iteration 9693, Loss: 0.0084, Time: 0.03s
Iteration 9694, Loss: 0.0091, Time: 0.03s
```

```
Iteration 9695, Loss: 0.0095, Time: 0.03s
Iteration 9696, Loss: 0.0074, Time: 0.03s
Iteration 9697, Loss: 0.0060, Time: 0.03s
Iteration 9698, Loss: 0.0122, Time: 0.03s
Iteration 9699, Loss: 0.0095, Time: 0.03s
Iteration 9700, Loss: 0.0120, Time: 0.03s
Iteration 9700, Loss: 0.0120, Time: 0.03s
Test Loss: 0.0408
Iteration 9701, Loss: 0.0036, Time: 0.03s
Iteration 9702, Loss: 0.0084, Time: 0.03s
Iteration 9703, Loss: 0.0111, Time: 0.03s
Iteration 9704, Loss: 0.0137, Time: 0.03s
Iteration 9705, Loss: 0.0135, Time: 0.03s
Iteration 9706, Loss: 0.0071, Time: 0.03s
Iteration 9707, Loss: 0.0103, Time: 0.03s
Iteration 9708, Loss: 0.0132, Time: 0.03s
Iteration 9709, Loss: 0.0075, Time: 0.03s
Iteration 9710, Loss: 0.0042, Time: 0.03s
Iteration 9711, Loss: 0.0118, Time: 0.03s
Iteration 9712, Loss: 0.0095, Time: 0.03s
Iteration 9713, Loss: 0.0072, Time: 0.03s
Iteration 9714, Loss: 0.0144, Time: 0.03s
Iteration 9715, Loss: 0.0133, Time: 0.03s
Iteration 9716, Loss: 0.0074, Time: 0.03s
Iteration 9717, Loss: 0.0099, Time: 0.03s
Iteration 9718, Loss: 0.0079, Time: 0.03s
Iteration 9719, Loss: 0.0078, Time: 0.03s
Iteration 9720, Loss: 0.0093, Time: 0.03s
Iteration 9721, Loss: 0.0085, Time: 0.03s
Iteration 9722, Loss: 0.0091, Time: 0.03s
Iteration 9723, Loss: 0.0110, Time: 0.03s
Iteration 9724, Loss: 0.0085, Time: 0.03s
Iteration 9725, Loss: 0.0137, Time: 0.03s
Iteration 9726, Loss: 0.0084, Time: 0.03s
Iteration 9727, Loss: 0.0128, Time: 0.03s
Iteration 9728, Loss: 0.0092, Time: 0.03s
Iteration 9729, Loss: 0.0138, Time: 0.03s
Iteration 9730, Loss: 0.0068, Time: 0.03s
Iteration 9731, Loss: 0.0109, Time: 0.03s
Iteration 9732, Loss: 0.0061, Time: 0.03s
Iteration 9733, Loss: 0.0128, Time: 0.03s
Iteration 9734, Loss: 0.0124, Time: 0.03s
Iteration 9735, Loss: 0.0153, Time: 0.03s
Iteration 9736, Loss: 0.0073, Time: 0.03s
Iteration 9737, Loss: 0.0072, Time: 0.03s
Iteration 9738, Loss: 0.0074, Time: 0.03s
Iteration 9739, Loss: 0.0110, Time: 0.03s
Iteration 9740, Loss: 0.0089, Time: 0.03s
```

```
Iteration 9741, Loss: 0.0075, Time: 0.03s
Iteration 9742, Loss: 0.0103, Time: 0.03s
Iteration 9743, Loss: 0.0047, Time: 0.03s
Iteration 9744, Loss: 0.0075, Time: 0.03s
Iteration 9745, Loss: 0.0095, Time: 0.03s
Iteration 9746, Loss: 0.0124, Time: 0.03s
Iteration 9747, Loss: 0.0087, Time: 0.03s
Iteration 9748, Loss: 0.0109, Time: 0.03s
Iteration 9749, Loss: 0.0087, Time: 0.03s
Iteration 9750, Loss: 0.0128, Time: 0.03s
Iteration 9751, Loss: 0.0071, Time: 0.03s
Iteration 9752, Loss: 0.0073, Time: 0.03s
Iteration 9753, Loss: 0.0095, Time: 0.03s
Iteration 9754, Loss: 0.0137, Time: 0.03s
Iteration 9755, Loss: 0.0063, Time: 0.03s
Iteration 9756, Loss: 0.0109, Time: 0.03s
Iteration 9757, Loss: 0.0090, Time: 0.03s
Iteration 9758, Loss: 0.0113, Time: 0.03s
Iteration 9759, Loss: 0.0104, Time: 0.03s
Iteration 9760, Loss: 0.0096, Time: 0.03s
Iteration 9761, Loss: 0.0055, Time: 0.03s
Iteration 9762, Loss: 0.0076, Time: 0.03s
Iteration 9763, Loss: 0.0051, Time: 0.03s
Iteration 9764, Loss: 0.0101, Time: 0.03s
Iteration 9765, Loss: 0.0149, Time: 0.03s
Iteration 9766, Loss: 0.0087, Time: 0.03s
Iteration 9767, Loss: 0.0112, Time: 0.03s
Iteration 9768, Loss: 0.0045, Time: 0.03s
Iteration 9769, Loss: 0.0098, Time: 0.03s
Iteration 9770, Loss: 0.0093, Time: 0.03s
Iteration 9771, Loss: 0.0106, Time: 0.03s
Iteration 9772, Loss: 0.0050, Time: 0.03s
Iteration 9773, Loss: 0.0102, Time: 0.03s
Iteration 9774, Loss: 0.0053, Time: 0.03s
Iteration 9775, Loss: 0.0097, Time: 0.03s
Iteration 9776, Loss: 0.0059, Time: 0.03s
Iteration 9777, Loss: 0.0044, Time: 0.03s
Iteration 9778, Loss: 0.0056, Time: 0.03s
Iteration 9779, Loss: 0.0127, Time: 0.03s
Iteration 9780, Loss: 0.0111, Time: 0.03s
Iteration 9781, Loss: 0.0083, Time: 0.03s
Iteration 9782, Loss: 0.0064, Time: 0.03s
Iteration 9783, Loss: 0.0083, Time: 0.03s
Iteration 9784, Loss: 0.0101, Time: 0.03s
Iteration 9785, Loss: 0.0108, Time: 0.03s
Iteration 9786, Loss: 0.0087, Time: 0.03s
Iteration 9787, Loss: 0.0065, Time: 0.03s
Iteration 9788, Loss: 0.0069, Time: 0.03s
```

```
Iteration 9789, Loss: 0.0086, Time: 0.03s
Iteration 9790, Loss: 0.0067, Time: 0.03s
Iteration 9791, Loss: 0.0116, Time: 0.03s
Iteration 9792, Loss: 0.0082, Time: 0.03s
Iteration 9793, Loss: 0.0106, Time: 0.03s
Iteration 9794, Loss: 0.0107, Time: 0.03s
Iteration 9795, Loss: 0.0073, Time: 0.03s
Iteration 9796, Loss: 0.0105, Time: 0.03s
Iteration 9797, Loss: 0.0079, Time: 0.03s
Iteration 9798, Loss: 0.0096, Time: 0.03s
Iteration 9799, Loss: 0.0129, Time: 0.03s
Iteration 9800, Loss: 0.0115, Time: 0.03s
Iteration 9800, Loss: 0.0115, Time: 0.03s
Test Loss: 0.0688
Iteration 9801, Loss: 0.0103, Time: 0.03s
Iteration 9802, Loss: 0.0103, Time: 0.03s
Iteration 9803, Loss: 0.0111, Time: 0.03s
Iteration 9804, Loss: 0.0074, Time: 0.03s
Iteration 9805, Loss: 0.0105, Time: 0.03s
Iteration 9806, Loss: 0.0113, Time: 0.03s
Iteration 9807, Loss: 0.0064, Time: 0.03s
Iteration 9808, Loss: 0.0072, Time: 0.03s
Iteration 9809, Loss: 0.0085, Time: 0.03s
Iteration 9810, Loss: 0.0111, Time: 0.03s
Iteration 9811, Loss: 0.0110, Time: 0.03s
Iteration 9812, Loss: 0.0090, Time: 0.03s
Iteration 9813, Loss: 0.0070, Time: 0.03s
Iteration 9814, Loss: 0.0108, Time: 0.03s
Iteration 9815, Loss: 0.0079, Time: 0.03s
Iteration 9816, Loss: 0.0120, Time: 0.03s
Iteration 9817, Loss: 0.0079, Time: 0.03s
Iteration 9818, Loss: 0.0105, Time: 0.03s
Iteration 9819, Loss: 0.0086, Time: 0.03s
Iteration 9820, Loss: 0.0112, Time: 0.03s
Iteration 9821, Loss: 0.0078, Time: 0.03s
Iteration 9822, Loss: 0.0126, Time: 0.03s
Iteration 9823, Loss: 0.0071, Time: 0.03s
Iteration 9824, Loss: 0.0075, Time: 0.03s
Iteration 9825, Loss: 0.0095, Time: 0.03s
Iteration 9826, Loss: 0.0118, Time: 0.03s
Iteration 9827, Loss: 0.0112, Time: 0.03s
Iteration 9828, Loss: 0.0097, Time: 0.03s
Iteration 9829, Loss: 0.0079, Time: 0.03s
Iteration 9830, Loss: 0.0118, Time: 0.03s
Iteration 9831, Loss: 0.0085, Time: 0.03s
Iteration 9832, Loss: 0.0087, Time: 0.03s
Iteration 9833, Loss: 0.0131, Time: 0.03s
Iteration 9834, Loss: 0.0111, Time: 0.03s
```

```
Iteration 9835, Loss: 0.0173, Time: 0.03s
Iteration 9836, Loss: 0.0053, Time: 0.03s
Iteration 9837, Loss: 0.0078, Time: 0.03s
Iteration 9838, Loss: 0.0079, Time: 0.03s
Iteration 9839, Loss: 0.0103, Time: 0.03s
Iteration 9840, Loss: 0.0136, Time: 0.03s
Iteration 9841, Loss: 0.0107, Time: 0.03s
Iteration 9842, Loss: 0.0062, Time: 0.03s
Iteration 9843, Loss: 0.0099, Time: 0.03s
Iteration 9844, Loss: 0.0131, Time: 0.03s
Iteration 9845, Loss: 0.0093, Time: 0.03s
Iteration 9846, Loss: 0.0095, Time: 0.03s
Iteration 9847, Loss: 0.0101, Time: 0.03s
Iteration 9848, Loss: 0.0061, Time: 0.03s
Iteration 9849, Loss: 0.0088, Time: 0.03s
Iteration 9850, Loss: 0.0084, Time: 0.03s
Iteration 9851, Loss: 0.0120, Time: 0.03s
Iteration 9852, Loss: 0.0099, Time: 0.03s
Iteration 9853, Loss: 0.0099, Time: 0.03s
Iteration 9854, Loss: 0.0084, Time: 0.03s
Iteration 9855, Loss: 0.0078, Time: 0.03s
Iteration 9856, Loss: 0.0149, Time: 0.03s
Iteration 9857, Loss: 0.0066, Time: 0.04s
Iteration 9858, Loss: 0.0070, Time: 0.03s
Iteration 9859, Loss: 0.0118, Time: 0.03s
Iteration 9860, Loss: 0.0108, Time: 0.03s
Iteration 9861, Loss: 0.0062, Time: 0.03s
Iteration 9862, Loss: 0.0085, Time: 0.03s
Iteration 9863, Loss: 0.0085, Time: 0.03s
Iteration 9864, Loss: 0.0117, Time: 0.03s
Iteration 9865, Loss: 0.0055, Time: 0.03s
Iteration 9866, Loss: 0.0074, Time: 0.03s
Iteration 9867, Loss: 0.0090, Time: 0.03s
Iteration 9868, Loss: 0.0122, Time: 0.03s
Iteration 9869, Loss: 0.0063, Time: 0.03s
Iteration 9870, Loss: 0.0098, Time: 0.03s
Iteration 9871, Loss: 0.0096, Time: 0.03s
Iteration 9872, Loss: 0.0065, Time: 0.03s
Iteration 9873, Loss: 0.0074, Time: 0.03s
Iteration 9874, Loss: 0.0112, Time: 0.03s
Iteration 9875, Loss: 0.0117, Time: 0.03s
Iteration 9876, Loss: 0.0096, Time: 0.03s
Iteration 9877, Loss: 0.0048, Time: 0.03s
Iteration 9878, Loss: 0.0079, Time: 0.03s
Iteration 9879, Loss: 0.0093, Time: 0.03s
Iteration 9880, Loss: 0.0058, Time: 0.03s
Iteration 9881, Loss: 0.0081, Time: 0.03s
Iteration 9882, Loss: 0.0095, Time: 0.03s
```

```
Iteration 9883, Loss: 0.0077, Time: 0.03s
Iteration 9884, Loss: 0.0078, Time: 0.03s
Iteration 9885, Loss: 0.0063, Time: 0.03s
Iteration 9886, Loss: 0.0106, Time: 0.03s
Iteration 9887, Loss: 0.0078, Time: 0.03s
Iteration 9888, Loss: 0.0087, Time: 0.03s
Iteration 9889, Loss: 0.0090, Time: 0.03s
Iteration 9890, Loss: 0.0088, Time: 0.03s
Iteration 9891, Loss: 0.0082, Time: 0.03s
Iteration 9892, Loss: 0.0093, Time: 0.03s
Iteration 9893, Loss: 0.0123, Time: 0.03s
Iteration 9894, Loss: 0.0084, Time: 0.03s
Iteration 9895, Loss: 0.0110, Time: 0.03s
Iteration 9896, Loss: 0.0066, Time: 0.03s
Iteration 9897, Loss: 0.0080, Time: 0.03s
Iteration 9898, Loss: 0.0077, Time: 0.03s
Iteration 9899, Loss: 0.0061, Time: 0.03s
Iteration 9900, Loss: 0.0083, Time: 0.03s
Iteration 9900, Loss: 0.0083, Time: 0.03s
Test Loss: 0.0550
Iteration 9901, Loss: 0.0105, Time: 0.03s
Iteration 9902, Loss: 0.0115, Time: 0.03s
Iteration 9903, Loss: 0.0098, Time: 0.03s
Iteration 9904, Loss: 0.0076, Time: 0.03s
Iteration 9905, Loss: 0.0105, Time: 0.03s
Iteration 9906, Loss: 0.0129, Time: 0.03s
Iteration 9907, Loss: 0.0070, Time: 0.03s
Iteration 9908, Loss: 0.0083, Time: 0.03s
Iteration 9909, Loss: 0.0067, Time: 0.03s
Iteration 9910, Loss: 0.0049, Time: 0.03s
Iteration 9911, Loss: 0.0104, Time: 0.03s
Iteration 9912, Loss: 0.0051, Time: 0.03s
Iteration 9913, Loss: 0.0080, Time: 0.03s
Iteration 9914, Loss: 0.0064, Time: 0.03s
Iteration 9915, Loss: 0.0124, Time: 0.03s
Iteration 9916, Loss: 0.0089, Time: 0.03s
Iteration 9917, Loss: 0.0078, Time: 0.03s
Iteration 9918, Loss: 0.0056, Time: 0.03s
Iteration 9919, Loss: 0.0101, Time: 0.03s
Iteration 9920, Loss: 0.0065, Time: 0.03s
Iteration 9921, Loss: 0.0096, Time: 0.03s
Iteration 9922, Loss: 0.0112, Time: 0.03s
Iteration 9923, Loss: 0.0109, Time: 0.03s
Iteration 9924, Loss: 0.0125, Time: 0.03s
Iteration 9925, Loss: 0.0131, Time: 0.03s
Iteration 9926, Loss: 0.0060, Time: 0.03s
Iteration 9927, Loss: 0.0093, Time: 0.03s
Iteration 9928, Loss: 0.0097, Time: 0.03s
```

```
Iteration 9929, Loss: 0.0096, Time: 0.03s
Iteration 9930, Loss: 0.0056, Time: 0.03s
Iteration 9931, Loss: 0.0073, Time: 0.03s
Iteration 9932, Loss: 0.0087, Time: 0.03s
Iteration 9933, Loss: 0.0036, Time: 0.03s
Iteration 9934, Loss: 0.0065, Time: 0.03s
Iteration 9935, Loss: 0.0088, Time: 0.03s
Iteration 9936, Loss: 0.0091, Time: 0.03s
Iteration 9937, Loss: 0.0076, Time: 0.03s
Iteration 9938, Loss: 0.0092, Time: 0.03s
Iteration 9939, Loss: 0.0096, Time: 0.03s
Iteration 9940, Loss: 0.0061, Time: 0.03s
Iteration 9941, Loss: 0.0065, Time: 0.03s
Iteration 9942, Loss: 0.0134, Time: 0.03s
Iteration 9943, Loss: 0.0089, Time: 0.03s
Iteration 9944, Loss: 0.0084, Time: 0.03s
Iteration 9945, Loss: 0.0060, Time: 0.03s
Iteration 9946, Loss: 0.0105, Time: 0.03s
Iteration 9947, Loss: 0.0125, Time: 0.03s
Iteration 9948, Loss: 0.0071, Time: 0.03s
Iteration 9949, Loss: 0.0100, Time: 0.03s
Iteration 9950, Loss: 0.0102, Time: 0.03s
Iteration 9951, Loss: 0.0085, Time: 0.03s
Iteration 9952, Loss: 0.0110, Time: 0.03s
Iteration 9953, Loss: 0.0044, Time: 0.03s
Iteration 9954, Loss: 0.0084, Time: 0.03s
Iteration 9955, Loss: 0.0080, Time: 0.03s
Iteration 9956, Loss: 0.0129, Time: 0.03s
Iteration 9957, Loss: 0.0064, Time: 0.03s
Iteration 9958, Loss: 0.0117, Time: 0.03s
Iteration 9959, Loss: 0.0093, Time: 0.03s
Iteration 9960, Loss: 0.0107, Time: 0.03s
Iteration 9961, Loss: 0.0126, Time: 0.03s
Iteration 9962, Loss: 0.0100, Time: 0.03s
Iteration 9963, Loss: 0.0100, Time: 0.03s
Iteration 9964, Loss: 0.0090, Time: 0.03s
Iteration 9965, Loss: 0.0100, Time: 0.03s
Iteration 9966, Loss: 0.0091, Time: 0.03s
Iteration 9967, Loss: 0.0096, Time: 0.03s
Iteration 9968, Loss: 0.0075, Time: 0.03s
Iteration 9969, Loss: 0.0051, Time: 0.03s
Iteration 9970, Loss: 0.0073, Time: 0.03s
Iteration 9971, Loss: 0.0070, Time: 0.03s
Iteration 9972, Loss: 0.0072, Time: 0.03s
Iteration 9973, Loss: 0.0160, Time: 0.03s
Iteration 9974, Loss: 0.0113, Time: 0.03s
Iteration 9975, Loss: 0.0067, Time: 0.03s
Iteration 9976, Loss: 0.0117, Time: 0.03s
```

```
Iteration 9977, Loss: 0.0086, Time: 0.03s
Iteration 9978, Loss: 0.0073, Time: 0.03s
Iteration 9979, Loss: 0.0120, Time: 0.03s
Iteration 9980, Loss: 0.0084, Time: 0.03s
Iteration 9981, Loss: 0.0059, Time: 0.03s
Iteration 9982, Loss: 0.0057, Time: 0.03s
Iteration 9983, Loss: 0.0136, Time: 0.03s
Iteration 9984, Loss: 0.0082, Time: 0.03s
Iteration 9985, Loss: 0.0132, Time: 0.03s
Iteration 9986, Loss: 0.0080, Time: 0.03s
Iteration 9987, Loss: 0.0108, Time: 0.03s
Iteration 9988, Loss: 0.0059, Time: 0.03s
Iteration 9989, Loss: 0.0086, Time: 0.03s
Iteration 9990, Loss: 0.0073, Time: 0.03s
Iteration 9991, Loss: 0.0082, Time: 0.03s
Iteration 9992, Loss: 0.0116, Time: 0.03s
Iteration 9993, Loss: 0.0093, Time: 0.03s
Iteration 9994, Loss: 0.0115, Time: 0.03s
Iteration 9995, Loss: 0.0072, Time: 0.03s
Iteration 9996, Loss: 0.0093, Time: 0.03s
Iteration 9997, Loss: 0.0088, Time: 0.03s
Iteration 9998, Loss: 0.0051, Time: 0.03s
Iteration 9999, Loss: 0.0098, Time: 0.03s
Iteration 10000, Loss: 0.0086, Time: 0.03s
Iteration 10000, Loss: 0.0086, Time: 0.03s
Test Loss: 0.0602
Training took 382.370s in total.
```

1.9   5. Deploy the trained model to a random set of 4 test images and visualise the automated segmentation.

You can show the images as a 4 x 3 panel. Each row shows one example, with the 3 columns being the test image, automated segmentation and ground truth segmentation.

```
[9]: ### Insert your code ###
     # Visualise the segmentation results
     # Load the test images and labels
     test_image_path = 'Task01_BrainTumour_2D/test_images'
     test_label_path = 'Task01_BrainTumour_2D/test_labels'
     test_image_names = sorted(os.listdir(test_image_path))
     test_label_names = sorted(os.listdir(test_label_path))

     # Randomly select 4 test images
     random_test_indices = random.sample(range(len(test_image_names)), 4)

     fig, axes = plt.subplots(4, 3, figsize=(15, 20))

     for i, idx in enumerate(random_test_indices):
```

```python
    test_image = imageio.imread(os.path.join(test_image_path,
 ↪test_image_names[idx]))
    test_label = imageio.imread(os.path.join(test_label_path,
 ↪test_label_names[idx]))

    # Preprocess the test image
    test_image_norm = normalise_intensity(test_image)
    test_image_tensor = torch.from_numpy(test_image_norm).unsqueeze(0).
 ↪unsqueeze(0).to(device, dtype=torch.float32)

    # Get the model prediction
    model.eval()
    with torch.no_grad():
        pred = model(test_image_tensor)
        pred = torch.argmax(pred, dim=1).squeeze().cpu().numpy()

    # Plot the test image
    axes[i, 0].imshow(test_image, cmap=image_cmap)
    axes[i, 0].set_title(f'Test Image {idx}')
    axes[i, 0].axis('off')

    # Plot the predicted segmentation
    axes[i, 1].imshow(pred, cmap=label_cmap)
    axes[i, 1].set_title(f'Predicted Segmentation {idx}')
    axes[i, 1].axis('off')

    # Plot the ground truth segmentation
    axes[i, 2].imshow(test_label, cmap=label_cmap)
    axes[i, 2].set_title(f'Ground Truth {idx}')
    axes[i, 2].axis('off')

plt.tight_layout()
plt.show()
### End of your code ###
```
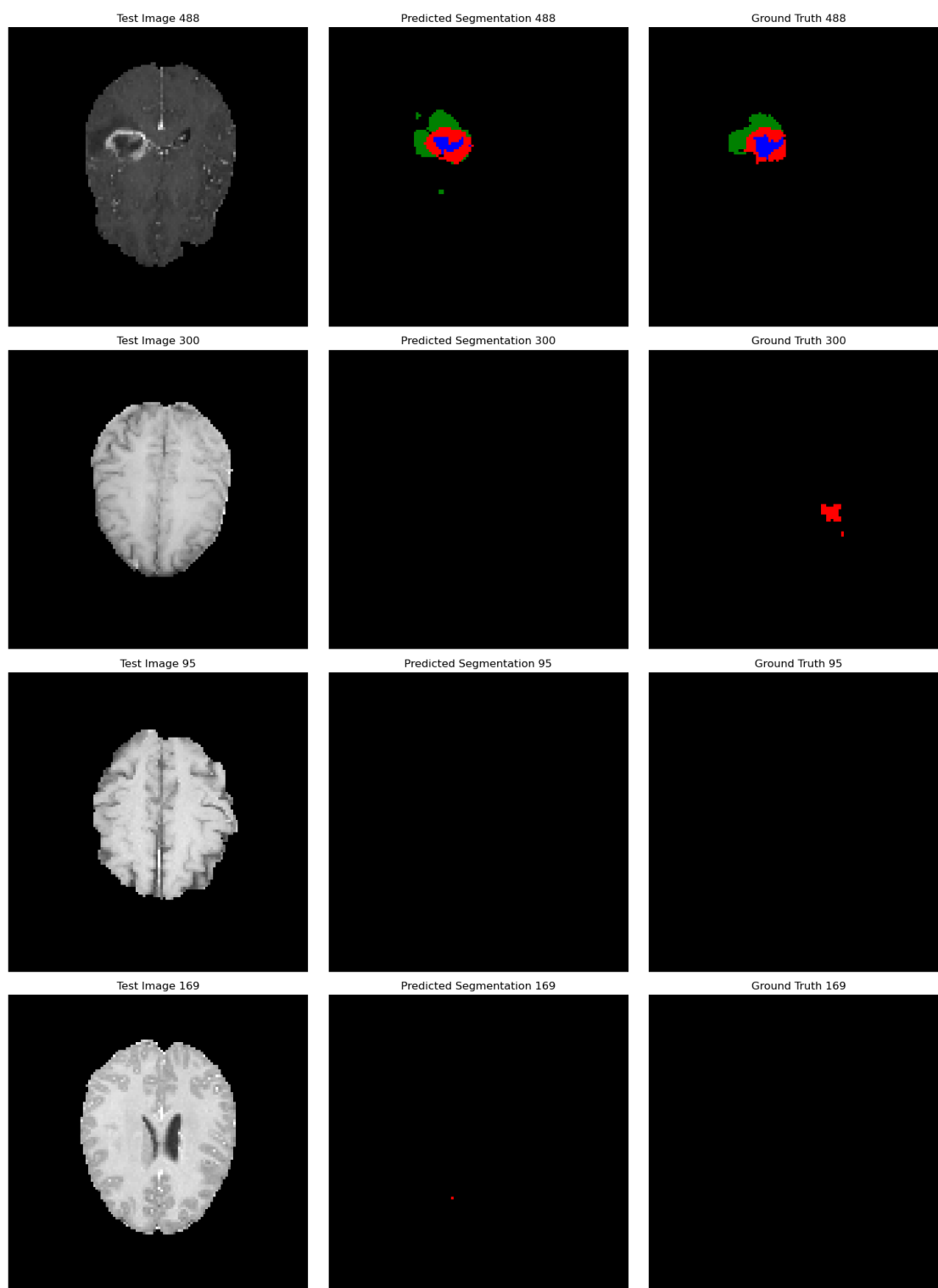
/tmp/ipykernel_8866/2287680980.py:15: DeprecationWarning: Starting with ImageIO
v3 the behavior of this function will switch to that of iio.v3.imread. To keep
the current behavior (and make this warning disappear) use `import imageio.v2 as
imageio` or call `imageio.v2.imread` directly.
  test_image = imageio.imread(os.path.join(test_image_path,
test_image_names[idx]))
/tmp/ipykernel_8866/2287680980.py:16: DeprecationWarning: Starting with ImageIO
v3 the behavior of this function will switch to that of iio.v3.imread. To keep
the current behavior (and make this warning disappear) use `import imageio.v2 as
imageio` or call `imageio.v2.imread` directly.
  test_label = imageio.imread(os.path.join(test_label_path,
test_label_names[idx]))

Test Image 488 — Predicted Segmentation 488 — Ground Truth 488

Test Image 300 — Predicted Segmentation 300 — Ground Truth 300

Test Image 95 — Predicted Segmentation 95 — Ground Truth 95

Test Image 169 — Predicted Segmentation 169 — Ground Truth 169

1.10 6. Discussion. Does your trained model work well? How would you improve this model so it can be deployed to the real clinic?

```
# Discussion: Does your trained model work well? How would you improve this
↪model so it can be deployed to the real clinic?

Yes, the trained model shows promising results in segmenting brain tumors from
↪MRI images. The model's performance, as indicated by the training and test
↪losses, suggests that it has learned to identify and segment different tumor
↪regions effectively. However, there is still room for improvement to enhance
↪its accuracy and robustness for clinical deployment.

Improvements:

1. Data Augmentation:
Increase the diversity of the training data by applying data augmentation
↪techniques such as rotation, scaling, flipping, and adding noise. This can
↪help the model generalize better to unseen data.

2. Hyperparameter Tuning:
Experiment with different hyperparameters such as learning rate, batch size,
↪and the number of filters in each convolutional layer to find the optimal
↪configuration for the model.

3. Model Architecture:
Explore more advanced architectures such as U-Net++ or Attention U-Net, which
↪have shown improved performance in medical image segmentation tasks.

4. Cross-validation:
Use cross-validation to ensure that the model's performance is consistent
↪across different subsets of the data and to avoid overfitting.

5. Clinical Validation:
Collaborate with medical professionals to validate the model's performance on
↪real clinical data and gather feedback for further improvements.

6. Robustness and Reliability:
Ensure the model is robust to variations in image quality, different MRI
↪machines, and patient demographics. Conduct extensive testing to ensure the
↪model's reliability in different clinical scenarios.

By implementing these improvements, the model can be made more accurate,
↪reliable, and suitable for deployment in a real clinical setting, ultimately
↪aiding in the diagnosis and treatment of brain tumors.
```