

```

# NOTE: The scaffolding code for this part of the assignment
# is adapted from https://github.com/pytorch/examples.
from __future__ import print_function
import argparse
import numpy as np
import torch
import torch.nn as nn
import torch.nn.functional as F
import torch.optim as optim
from torchvision import transforms
import torchvision
from torch.autograd import Variable

# NOTE: This is adapted from the torchvision datasets for
# CIFAR10 and CIFAR100, which can be found at
# https://github.com/pytorch/vision/blob/master/torchvision/datasets/cifar.py
# This version allows a validation dataset to be created.
from __future__ import print_function
from PIL import Image
import os
import os.path
import errno
import numpy as np
import sys
if sys.version_info[0] == 2:
    import cPickle as pickle
else:
    import pickle

import torch.utils.data as data
from torchvision.datasets.utils import download_url, check_integrity
from google.colab import drive
drive.mount('/content/gdrive')

class CIFAR10(data.Dataset):
    """`CIFAR10` <https://www.cs.toronto.edu/~kriz/cifar.html>`_ Dataset.
    Args:
        root (string): Root directory of dataset where directory
            ``cifar-10-batches-py`` exists.
        train (bool, optional): If True, creates dataset from training set, otherwise
            creates from test set.
        transform (callable, optional): A function/transform that takes in an PIL im
            and returns a transformed version. E.g, ``transforms.RandomCrop``
        target_transform (callable, optional): A function/transform that takes in the
            target and transforms it.
        download (bool, optional): If true, downloads the dataset from the internet a
            puts it in root directory. If dataset is already downloaded, it is not
            downloaded again.
    """
    base_folder = 'cifar-10-batches-py'

```

```

url = "http://www.cs.toronto.edu/~kriz/cifar-10-python.tar.gz"
filename = "cifar-10-python.tar.gz"
tgz_md5 = 'c58f30108f718f92721af3b95e74349a'
# validation examples will come from here
train_list = [
    ['data_batch_1', 'c99cafc152244af753f735de768cd75f'],
    ['data_batch_2', 'd4bba439e000b95fd0a9bffe97cbabec'],
    ['data_batch_3', '54ebc095f3ab1f0389bbae665268c751'],
    ['data_batch_4', '634d18415352ddfa80567beed471001a'],
    ['data_batch_5', '482c414d41f54cd18b22e5b47cb7c3cb'],
]

test_list = [
    ['test_batch', '40351d587109b95175f43aff81a1287e'],
]

def __init__(self, root, split='train',
              transform=None, target_transform=None,
              download=False, val_samples=1000):
    self.root = os.path.expanduser(root)
    self.transform = transform
    self.target_transform = target_transform
    self.split = split # train, val, or test

    if download:
        self.download()

    if not self._check_integrity():
        raise RuntimeError('Dataset not found or corrupted.' +
                           ' You can use download=True to download it')

    # now load the picked numpy arrays
    if self.split in ['train', 'val']:
        self.train_data = []
        self.train_labels = []
        for fentry in self.train_list:
            f = fentry[0]
            file = os.path.join(self.root, self.base_folder, f)
            fo = open(file, 'rb')
            if sys.version_info[0] == 2:
                entry = pickle.load(fo)
            else:
                entry = pickle.load(fo, encoding='latin1')
            self.train_data.append(entry['data'])
            if 'labels' in entry:
                self.train_labels += entry['labels']
            else:
                self.train_labels += entry['fine_labels']
            fo.close()

        self.train_data = np.concatenate(self.train_data)
        self.train_data = self.train_data.reshape((50000, 3, 32, 32))

```

```

self.train_data = self.train_data.transpose((0, 2, 3, 1)) # convert to H
self.val_data = self.train_data[-val_samples:]
self.val_labels = self.train_labels[-val_samples:]
self.train_data = self.train_data[:-val_samples]
self.train_labels = self.train_labels[:-val_samples]
elif self.split == 'test':
    f = self.test_list[0][0]
    file = os.path.join(self.root, self.base_folder, f)
    fo = open(file, 'rb')
    if sys.version_info[0] == 2:
        entry = pickle.load(fo)
    else:
        entry = pickle.load(fo, encoding='latin1')
    self.test_data = entry['data']
    if 'labels' in entry:
        self.test_labels = entry['labels']
    else:
        self.test_labels = entry['fine_labels']
    fo.close()
    self.test_data = self.test_data.reshape((10000, 3, 32, 32))
    self.test_data = self.test_data.transpose((0, 2, 3, 1)) # convert to HWC
else:
    raise Exception('Unkown split {}'.format(self.split))

def __getitem__(self, index):
    """
    Args:
        index (int): Index
    Returns:
        tuple: (image, target) where target is index of the target class.
    """
    if self.split == 'train':
        img, target = self.train_data[index], self.train_labels[index]
    elif self.split == 'val':
        img, target = self.val_data[index], self.val_labels[index]
    elif self.split == 'test':
        img, target = self.test_data[index], self.test_labels[index]

    # doing this so that it is consistent with all other datasets
    # to return a PIL Image
    img = Image.fromarray(img)

    if self.transform is not None:
        img = self.transform(img)

    if self.target_transform is not None:
        target = self.target_transform(target)

    return img, target

def __len__(self):

```

```

    if self.split == 'train':
        return len(self.train_data)
    elif self.split == 'val':
        return len(self.val_data)
    elif self.split == 'test':
        return len(self.test_data)

def _check_integrity(self):
    root = self.root
    for fentry in (self.train_list + self.test_list):
        filename, md5 = fentry[0], fentry[1]
        fpath = os.path.join(root, self.base_folder, filename)
        if not check_integrity(fpath, md5):
            return False
    return True

def download(self):
    import tarfile

    if self._check_integrity():
        print('Files already downloaded and verified')
        return

    root = self.root
    download_url(self.url, root, self.filename, self.tgz_md5)

    # extract file
    cwd = os.getcwd()
    tar = tarfile.open(os.path.join(root, self.filename), "r:gz")
    os.chdir(root)
    tar.extractall()
    tar.close()
    os.chdir(cwd)

import matplotlib.pyplot as plt

# You should implement these (softmax.py, twolayernn.py, convnet.py)
description='CIFAR-10 Example'
# Hyperparameters
lr = 0.0001
momentum = 0.9
weight_decay = 0.1
batch_size = 256
epochs = 10
model = 'mymodel'
no_cuda=False
seed=1
test_batch_size=1000
log_interval=10
cifar10_dir = 'data'
cuda = True

```

```

torch.manual_seed(seed)
if cuda:
    torch.cuda.manual_seed(seed)

# Load CIFAR10 using torch data paradigm
kwargs = {'num_workers': 1, 'pin_memory': True} if cuda else {}
# CIFAR10 meta data
n_classes = 10
im_size = (3, 32, 32)
# Subtract the mean color and divide by standard deviation. The mean image
# from part 1 of this homework was essentially a big gray blob, so
# subtracting the same color for all pixels doesn't make much difference.
# mean color of training images
cifar10_mean_color = [0.49131522, 0.48209435, 0.44646862]
# std dev of color across training images
cifar10_std_color = [0.01897398, 0.03039277, 0.03872553]
transform = transforms.Compose([
    transforms.Resize(size=(224, 224)),
    transforms.RandomHorizontalFlip(),
    transforms.RandomCrop(size=(224,224), padding=4),
    transforms.ToTensor(),
    transforms.Normalize(cifar10_mean_color, cifar10_std_color),
])

# Datasets
train_dataset = CIFAR10(cifar10_dir, split='train', download=True,
                        transform=transform)
val_dataset = CIFAR10(cifar10_dir, split='val', download=True,
                      transform=transform)
test_dataset = CIFAR10(cifar10_dir, split='test', download=True,
                       transform=transform)

# DataLoaders
train_loader = torch.utils.data.DataLoader(train_dataset,
                                           batch_size=batch_size, shuffle=True, **kwargs)
val_loader = torch.utils.data.DataLoader(val_dataset,
                                         batch_size=batch_size, shuffle=True, **kwargs)
test_loader = torch.utils.data.DataLoader(test_dataset,
                                          batch_size=batch_size, shuffle=True, **kwargs)

# def imshow(inp, title=None):
#     inp = inp.numpy().transpose((1, 2, 0))
#     mean = np.array(cifar10_mean_color)
#     std = np.array(cifar10_std_color)
#     inp = std * inp + mean
#     inp = np.clip(inp, 0, 1)
#     plt.imshow(inp)
#     if title is not None:
#         plt.title(title)
#     plt.pause(0.001) # Pause a bit so that plots are updated

if model == 'mymodel':
    # model = models.mymodel.MyModel(im_size, args.hidden_dim, args.hidden_dim2,
    #                                args.kernel_size, args.kernel_size2, n_classes)

```

```

# args.kernel_size, args.kernel_size2, n_classes)
model = torchvision.models.resnet152(pretrained=True)
for param in model.parameters():
    param.requires_grad = False
print(model)
# Modify the last layer
# ReLU(inplace=True)
# (5): Dropout(p=0.5, inplace=False)
model.fc = torch.nn.Linear(2048, 10)

print(model)
else:
    raise Exception('Unknown model {}'.format(model))
# cross-entropy loss function
criterion = F.cross_entropy
model.cuda()

#####
# TODO: Initialize an optimizer from the torch.optim package using the
# appropriate hyperparameters found in args. This only requires one line.
#####
optimizer = optim.Adam(model.parameters(), lr=lr, weight_decay=weight_decay)
#####
#                                     END OF YOUR CODE                                     #
#####

def train(epoch):
    """
    Train the model for one epoch.
    """
    # Some models use slightly different forward passes and train and test
    # time (e.g., any model with Dropout). This puts the model in train mode
    # (as opposed to eval mode) so it knows which one to use.
    model.train()
    max_val_loss = 1000
    # train loop
    for batch_idx, batch in enumerate(train_loader):
        # prepare data
        images, targets = Variable(batch[0]), Variable(batch[1])
        print(images.shape)
        # imshow(images[0])
        if cuda:
            images, targets = images.cuda(), targets.cuda()
        #####
        # TODO: Update the parameters in model using the optimizer from above.
        # This only requires a couple lines of code.
        #####
        loss = criterion(model(images), targets)
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()
        #####

```

```

#                                     END OF YOUR CODE                                     #
#####
if batch_idx % log_interval == 0:
    val_loss, val_acc = evaluate('val', n_batches=4)
    train_loss = loss.data
    examples_this_epoch = batch_idx * len(images)
    epoch_progress = 100. * batch_idx / len(train_loader)
    print('Train Epoch: {} [{}/{}] ({:.0f}%)\\t'
          'Train Loss: {:.6f}\\tVal Loss: {:.6f}\\tVal Acc: {}'.format(
              epoch, examples_this_epoch, len(train_loader.dataset),
              epoch_progress, train_loss, val_loss, val_acc))
    if val_loss < max_val_loss:
        model_save_name = 'transferLearning.pt'
        path = F"/content/gdrive/My Drive/transferLearning.pt"
        torch.save(model, path)
        max_val_loss = val_loss
        print("Saved model this time!")

def evaluate(split, verbose=False, n_batches=None):
    """
    Compute loss on val or test data.
    """
    model.eval()
    loss = 0
    correct = 0
    n_examples = 0
    if split == 'val':
        loader = val_loader
    elif split == 'test':
        loader = test_loader
    for batch_i, batch in enumerate(loader):
        data, target = batch
        if cuda:
            data, target = data.cuda(), target.cuda()
        data, target = Variable(data, volatile=True), Variable(target)
        output = model(data)
        loss += criterion(output, target, size_average=False).data
        # predict the argmax of the log-probabilities
        pred = output.data.max(1, keepdim=True)[1]
        correct += pred.eq(target.data.view_as(pred)).cpu().sum()
        n_examples += pred.size(0)
        if n_batches and (batch_i >= n_batches):
            break

    loss /= n_examples
    acc = 100. * correct / n_examples
    if verbose:
        print('\\n{} set: Average loss: {:.4f}, Accuracy: {}/{} ({:.0f}%)\\n'.format(
            split, loss, correct, n_examples, acc))
    return loss, acc

```

```
# train the model one epoch at a time

for epoch in range(1, epochs + 1):
    train(epoch)
    evaluate('test', verbose=True)

# Save the model (architecture and weights)

# Later you can call torch.load(file) to re-load the trained model into python
# See http://pytorch.org/docs/master/notes/serialization.html for more details
```



Go to this URL in a browser: https://accounts.google.com/o/oauth2/auth?client_id=

Enter your authorization code:

.....

Mounted at /content/gdrive

Downloading <http://www.cs.toronto.edu/~kriz/cifar-10-python.tar.gz> to data/cifar-

170500096it [00:06, 26924132.12it/s]

Files already downloaded and verified

Files already downloaded and verified

Downloading: "<https://download.pytorch.org/models/resnet152-b121ed2d.pth>" to /root

100% 230M/230M [00:10<00:00, 23.1MB/s]

ResNet(

(conv1): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=

(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stat

(relu): ReLU(inplace=True)

(maxpool): MaxPool2d(kernel_size=3, stride=2, padding=1, dilation=1, ceil_mode=

(layer1): Sequential(

(0): Bottleneck(

(conv1): Conv2d(64, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)

(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running

(conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),

(bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running

(conv3): Conv2d(64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)

(bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running

(relu): ReLU(inplace=True)

(downsample): Sequential(

(0): Conv2d(64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)

(1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running

)

)

(1): Bottleneck(

(conv1): Conv2d(256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)

(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running

(conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),

(bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running

(conv3): Conv2d(64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)

(bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running

(relu): ReLU(inplace=True)

)

(2): Bottleneck(

(conv1): Conv2d(256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)

(bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running

(conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),

(bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running

(conv3): Conv2d(64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)

(bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running

(relu): ReLU(inplace=True)

)

)

(layer2): Sequential(

(0): Bottleneck(

(conv1): Conv2d(256, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)

(bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running

(conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1)

```

(conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
(bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(rel): ReLU(inplace=True)
(downsample): Sequential(
  (0): Conv2d(256, 512, kernel_size=(1, 1), stride=(2, 2), bias=False)
  (1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
)
)
(1): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(2): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(3): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(4): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(5): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(6): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)

```

```

(conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
(bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(rel): ReLU(inplace=True)
)
(7): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
)
(layer3): Sequential(
  (0): Bottleneck(
    (conv1): Conv2d(512, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1))
    (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu): ReLU(inplace=True)
    (downsample): Sequential(
      (0): Conv2d(512, 1024, kernel_size=(1, 1), stride=(2, 2), bias=False)
      (1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (1): Bottleneck(
    (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu): ReLU(inplace=True)
  )
  (2): Bottleneck(
    (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu): ReLU(inplace=True)
  )
  (3): Bottleneck(
    (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu): ReLU(inplace=True)
  )
  (4): Bottleneck(

```

```

    (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
    (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu): ReLU(inplace=True)
)
(5): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(6): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(7): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(8): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(9): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(10): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```

```

        (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(inplace=True)
    )
(11): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(12): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(13): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(14): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(15): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(16): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```

```

)
(17): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(18): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(19): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(20): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(21): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(22): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(23): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```

```

(conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
(bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(rel): ReLU(inplace=True)
)
(24): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(25): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(26): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(27): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(28): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(29): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```

```

        (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
        (relu): ReLU(inplace=True)
    )
(30): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(31): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(32): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(33): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(34): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(35): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
)
)

```



```

        (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
        (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running
        (conv3): Conv2d(64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running
        (relu): ReLU(inplace=True)
    )
(2): Bottleneck(
  (conv1): Conv2d(256, 64, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running
  (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
  (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running
  (conv3): Conv2d(64, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running
  (relu): ReLU(inplace=True)
)
)
(layer2): Sequential(
  (0): Bottleneck(
    (conv1): Conv2d(256, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running
    (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1),
    (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running
    (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running
    (relu): ReLU(inplace=True)
    (downsample): Sequential(
      (0): Conv2d(256, 512, kernel_size=(1, 1), stride=(2, 2), bias=False)
      (1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running
    )
  )
  (1): Bottleneck(
    (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running
    (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
    (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running
    (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running
    (relu): ReLU(inplace=True)
  )
  (2): Bottleneck(
    (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running
    (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
    (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running
    (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running
    (relu): ReLU(inplace=True)
  )
  (3): Bottleneck(
    (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running
    (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
    (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running
    (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running
    (relu): ReLU(inplace=True)
  )
)

```

```

(4): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(5): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(6): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(7): Bottleneck(
  (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(128, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
)
(layer3): Sequential(
  (0): Bottleneck(
    (conv1): Conv2d(512, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
    (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (relu): ReLU(inplace=True)
    (downsample): Sequential(
      (0): Conv2d(512, 1024, kernel_size=(1, 1), stride=(2, 2), bias=False)
      (1): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (1): Bottleneck(
    (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
    (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  )
)

```

```

    (relu): ReLU(inplace=True)
)
(2): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(3): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(4): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(5): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(6): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(7): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(8): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```

```

(conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
(bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(rel): ReLU(inplace=True)
)
(9): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(10): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(11): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(12): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(13): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(14): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```

```

(conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
(rel): ReLU(inplace=True)
)
(15): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (relu): ReLU(inplace=True)
)
(16): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (relu): ReLU(inplace=True)
)
(17): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (relu): ReLU(inplace=True)
)
(18): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (relu): ReLU(inplace=True)
)
(19): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (relu): ReLU(inplace=True)
)
(20): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_mean=True)
  (relu): ReLU(inplace=True)
)
)

```

```

(21): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(22): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(23): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(24): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(25): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(26): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(27): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```

```

(conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
(bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
(bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(rel): ReLU(inplace=True)
)
(28): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(29): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(30): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(31): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(32): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)
(33): Bottleneck(
  (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
)

```



```

        (relu): ReLU(inplace=True)
    )
    (34): Bottleneck(
      (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
    )
    (35): Bottleneck(
      (conv1): Conv2d(1024, 256, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(256, 1024, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
    )
  )
  (layer4): Sequential(
    (0): Bottleneck(
      (conv1): Conv2d(1024, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(512, 2048, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (downsample): Sequential(
        (0): Conv2d(1024, 2048, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (1): Bottleneck(
      (conv1): Conv2d(2048, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(512, 2048, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
    )
    (2): Bottleneck(
      (conv1): Conv2d(2048, 512, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (conv3): Conv2d(512, 2048, kernel_size=(1, 1), stride=(1, 1), bias=False)
      (bn3): BatchNorm2d(2048, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
    )
  )
  (avgpool): AdaptiveAvgPool2d(output_size=(1, 1))
  (fc): Linear(in_features=2048, out_features=10, bias=True)
)
torch.Size([256, 3, 224, 224])

```

```

/usr/local/lib/python3.6/dist-packages/ipykernel_launcher.py:341: UserWarning: v
/usr/local/lib/python3.6/dist-packages/torch/nn/_reduction.py:43: UserWarning: s
  warnings.warn(warning.format(ret))
Train Epoch: 1 [0/49000 (0%)]    Train Loss: 2.368309    Val Loss: 2.380458    \
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 1 [2560/49000 (5%)]    Train Loss: 2.291346    Val Loss: 2.28630
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 1 [5120/49000 (10%)]    Train Loss: 2.225854    Val Loss: 2.22459
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 1 [7680/49000 (16%)]    Train Loss: 2.191221    Val Loss: 2.19331
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 1 [10240/49000 (21%)]    Train Loss: 2.170921    Val Loss: 2.16461
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [12800/49000 (26%)]

Train Loss: 2.158970

Val Loss: 2.12509

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [15360/49000 (31%)]

Train Loss: 2.112028

Val Loss: 2.11731

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [17920/49000 (36%)]

Train Loss: 2.142910

Val Loss: 2.08831

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [20480/49000 (42%)]

Train Loss: 2.083063

Val Loss: 2.05581

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [23040/49000 (47%)]

Train Loss: 2.026353

Val Loss: 2.04831

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [25600/49000 (52%)]

Train Loss: 2.074434

Val Loss: 2.01419

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [28160/49000 (57%)]

Train Loss: 2.004257

Val Loss: 2.00240

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [30720/49000 (62%)]

Train Loss: 1.993350

Val Loss: 1.99864

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [33280/49000 (68%)]

Train Loss: 1.985238

Val Loss: 1.96971

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 1 [35840/49000 (73%)]

Train Loss: 1.881185

Val Loss: 1.95784

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 1 [38400/49000 (78%)]

Train Loss: 1.940634

Val Loss: 1.93651

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 1 [40960/49000 (83%)]

Train Loss: 1.960648

Val Loss: 1.92324

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 1 [43520/49000 (89%)]

Train Loss: 1.910916

Val Loss: 1.90949

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 1 [46080/49000 (94%)]

Train Loss: 1.876941

Val Loss: 1.89831

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([128, 3, 224, 224])
Train Epoch: 1 [48640/49000 (99%)]      Train Loss: 1.894295      Val Loss: 1.88310
Saved model this time!
torch.Size([104, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [0/49000 (0%)]      Train Loss: 2.264203      Val Loss: 1.868223
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [2560/49000 (5%)]      Train Loss: 1.754707      Val Loss: 1.7538
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [5120/49000 (10%)]      Train Loss: 1.718418      Val Loss: 1.7133
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [7680/49000 (16%)]      Train Loss: 1.744086      Val Loss: 1.6954
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [10240/49000 (21%)]      Train Loss: 1.659988      Val Loss: 1.6659
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [12800/49000 (26%)]    Train Loss: 1.633620    Val Loss: 1.62050
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [15360/49000 (31%)]    Train Loss: 1.580879    Val Loss: 1.60210
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [17920/49000 (36%)]    Train Loss: 1.543830    Val Loss: 1.58540
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [20480/49000 (42%)]    Train Loss: 1.535140    Val Loss: 1.55210
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [23040/49000 (47%)]    Train Loss: 1.539677    Val Loss: 1.53570
Saved model this time!
torch.Size([256, 3, 224, 224])
```

```

.....\....., .., ....., .....
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [25600/49000 (52%)]      Train Loss: 1.514537      Val Loss: 1.5275
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [28160/49000 (57%)]      Train Loss: 1.472236      Val Loss: 1.4940
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [30720/49000 (62%)]      Train Loss: 1.480706      Val Loss: 1.4867
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [33280/49000 (68%)]      Train Loss: 1.478842      Val Loss: 1.4578
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [35840/49000 (73%)]      Train Loss: 1.405777      Val Loss: 1.4522

```



```

Train Epoch: 2 [35840/49000 (73%)]      Train Loss: 1.495777      Val Loss: 1.4553.
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [38400/49000 (78%)]      Train Loss: 1.439593      Val Loss: 1.43089
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [40960/49000 (83%)]      Train Loss: 1.453828      Val Loss: 1.41929
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [43520/49000 (89%)]      Train Loss: 1.430559      Val Loss: 1.40139
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 2 [46080/49000 (94%)]      Train Loss: 1.419873      Val Loss: 1.40189
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
Train Epoch: 2 [48640/49000 (99%)]      Train Loss: 1.345147      Val Loss: 1.36536
Saved model this time!
torch.Size([104, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 3 [0/49000 (0%)]      Train Loss: 1.710355      Val Loss: 1.200107
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 3 [2560/49000 (5%)]      Train Loss: 1.179942      Val Loss: 1.15667
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 3 [5120/49000 (10%)]      Train Loss: 1.180391      Val Loss: 1.14087
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 3 [7680/49000 (16%)]      Train Loss: 1.110277      Val Loss: 1.11997
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 3 [10240/49000 (21%)]      Train Loss: 1.192789      Val Loss: 1.11339
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [12800/49000 (26%)]

Train Loss: 1.145720

Val Loss: 1.09531

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [15360/49000 (31%)]

Train Loss: 1.088721

Val Loss: 1.07946

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [17920/49000 (36%)]

Train Loss: 1.023585

Val Loss: 1.07291

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [20480/49000 (42%)]

Train Loss: 1.083450

Val Loss: 1.05946

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [23040/49000 (47%)]

Train Loss: 1.088979

Val Loss: 1.02561

Saved model this time!

```
torch.Size([256, 3, 224, 224])
```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 3 [25600/49000 (52%)]

Train Loss: 1.013652

Val Loss: 1.03509

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 3 [28160/49000 (57%)]

Train Loss: 1.063842

Val Loss: 1.04013

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 3 [30720/49000 (62%)]

Train Loss: 1.063924

Val Loss: 1.01389

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 3 [33280/49000 (68%)]

Train Loss: 0.989091

Val Loss: 1.01253

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 3 [35840/49000 (73%)]

Train Loss: 1.070152

Val Loss: 0.98714

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [38400/49000 (78%)]

Train Loss: 0.992967

Val Loss: 0.98725

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [40960/49000 (83%)]

Train Loss: 1.004160

Val Loss: 0.97180

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [43520/49000 (89%)]

Train Loss: 0.943324

Val Loss: 0.98284

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [46080/49000 (94%)]

Train Loss: 0.966776

Val Loss: 0.96691

Saved model this time!

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

Train Epoch: 3 [48640/49000 (99%)]

Train Loss: 0.923507

Val Loss: 0.96580

Saved model this time!

```
Saved model this time!
```

```
torch.Size([104, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
Train Epoch: 4 [0/49000 (0%)] Train Loss: 1.451052 Val Loss: 0.854531
```

```
Saved model this time!
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
Train Epoch: 4 [2560/49000 (5%)] Train Loss: 0.826237 Val Loss: 0.84437
```

```
Saved model this time!
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
Train Epoch: 4 [5120/49000 (10%)] Train Loss: 0.785215 Val Loss: 0.83426
```

```
Saved model this time!
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
Train Epoch: 4 [7680/49000 (16%)] Train Loss: 0.893330 Val Loss: 0.81376
```

```
Saved model this time!
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
Train Epoch: 4 [10240/49000 (21%)] Train Loss: 0.800676 Val Loss: 0.81176
```

```
Saved model this time!
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [12800/49000 (26%)]      Train Loss: 0.818147      Val Loss: 0.8023!
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [15360/49000 (31%)]      Train Loss: 0.721139      Val Loss: 0.8035!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [17920/49000 (36%)]      Train Loss: 0.843299      Val Loss: 0.8051!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [20480/49000 (42%)]      Train Loss: 0.828044      Val Loss: 0.7760!
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [23040/49000 (47%)]      Train Loss: 0.773511      Val Loss: 0.7937!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [25600/49000 (52%)]      Train Loss: 0.799816      Val Loss: 0.79550
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [28160/49000 (57%)]      Train Loss: 0.800147      Val Loss: 0.78260
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [30720/49000 (62%)]      Train Loss: 0.804919      Val Loss: 0.76830
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [33280/49000 (68%)]      Train Loss: 0.748627      Val Loss: 0.77730
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [35840/49000 (73%)]      Train Loss: 0.736430      Val Loss: 0.76030
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```



```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [38400/49000 (78%)]      Train Loss: 0.780497      Val Loss: 0.76576
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [40960/49000 (83%)]      Train Loss: 0.780525      Val Loss: 0.75196
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [43520/49000 (89%)]      Train Loss: 0.789780      Val Loss: 0.74356
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [46080/49000 (94%)]      Train Loss: 0.755024      Val Loss: 0.74866
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 4 [48640/49000 (99%)]      Train Loss: 0.786358      Val Loss: 0.74076
Saved model this time!
torch.Size([104, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [0/49000 (0%)]      Train Loss: 1.286258      Val Loss: 0.671899
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [2560/49000 (5%)]      Train Loss: 0.704010      Val Loss: 0.68759
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [5120/49000 (10%)]    Train Loss: 0.684448      Val Loss: 0.68493
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [7680/49000 (16%)]    Train Loss: 0.670135      Val Loss: 0.67608
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [10240/49000 (21%)]   Train Loss: 0.694678      Val Loss: 0.66630
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [12800/49000 (26%)]   Train Loss: 0.667496      Val Loss: 0.67908
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [15360/49000 (31%)]      Train Loss: 0.718447      Val Loss: 0.66069
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [17920/49000 (36%)]      Train Loss: 0.660147      Val Loss: 0.66081
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [20480/49000 (42%)]      Train Loss: 0.713746      Val Loss: 0.65698
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [23040/49000 (47%)]      Train Loss: 0.634169      Val Loss: 0.65150
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [25600/49000 (52%)]      Train Loss: 0.673367      Val Loss: 0.66026
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [28160/49000 (57%)]      Train Loss: 0.605540      Val Loss: 0.6448!
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [30720/49000 (62%)]      Train Loss: 0.685648      Val Loss: 0.6496!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [33280/49000 (68%)]      Train Loss: 0.720873      Val Loss: 0.6503!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [35840/49000 (73%)]      Train Loss: 0.626212      Val Loss: 0.6527!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [38400/49000 (78%)]      Train Loss: 0.599460      Val Loss: 0.6251!
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [40960/49000 (83%)]      Train Loss: 0.600873      Val Loss: 0.6464!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [43520/49000 (89%)]      Train Loss: 0.662426      Val Loss: 0.6429!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [46080/49000 (94%)]      Train Loss: 0.622890      Val Loss: 0.6373!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 5 [48640/49000 (99%)]      Train Loss: 0.679195      Val Loss: 0.6410!
torch.Size([104, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [0/49000 (0%)]      Train Loss: 1.220950      Val Loss: 0.592006      \
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [2560/49000 (5%)]      Train Loss: 0.618742      Val Loss: 0.6148!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [5120/49000 (10%)]      Train Loss: 0.610966      Val Loss: 0.61406
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [7680/49000 (16%)]      Train Loss: 0.638846      Val Loss: 0.58721
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [10240/49000 (21%)]      Train Loss: 0.475878      Val Loss: 0.60216
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [12800/49000 (26%)]      Train Loss: 0.501741      Val Loss: 0.58196
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [15360/49000 (31%)]      Train Loss: 0.610225      Val Loss: 0.58920
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [17920/49000 (36%)]      Train Loss: 0.620539      Val Loss: 0.58637
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [20480/49000 (42%)]      Train Loss: 0.625561      Val Loss: 0.57916
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [23040/49000 (47%)]      Train Loss: 0.574829      Val Loss: 0.59628
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [25600/49000 (52%)]      Train Loss: 0.596519      Val Loss: 0.57720
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [28160/49000 (57%)]      Train Loss: 0.610399      Val Loss: 0.57920
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [30720/49000 (62%)]      Train Loss: 0.523238      Val Loss: 0.58257
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [33280/49000 (68%)]      Train Loss: 0.611538      Val Loss: 0.59904
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [35840/49000 (73%)]      Train Loss: 0.536688      Val Loss: 0.58461
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [38400/49000 (78%)]      Train Loss: 0.528233      Val Loss: 0.57357
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [40960/49000 (83%)]      Train Loss: 0.694261      Val Loss: 0.58777
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```



```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [43520/49000 (89%)]      Train Loss: 0.507718      Val Loss: 0.57864
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [46080/49000 (94%)]      Train Loss: 0.633168      Val Loss: 0.58401
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 6 [48640/49000 (99%)]      Train Loss: 0.534841      Val Loss: 0.57694
torch.Size([104, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [0/49000 (0%)]      Train Loss: 1.168081      Val Loss: 0.562900
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [2560/49000 (5%)]      Train Loss: 0.509286      Val Loss: 0.57624
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [5120/49000 (10%)]      Train Loss: 0.563745      Val Loss: 0.57234
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 7 [7680/49000 (16%)]

Train Loss: 0.609990

Val Loss: 0.56451

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 7 [10240/49000 (21%)]

Train Loss: 0.595339

Val Loss: 0.56381

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 7 [12800/49000 (26%)]

Train Loss: 0.555606

Val Loss: 0.55570

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 7 [15360/49000 (31%)]

Train Loss: 0.531984

Val Loss: 0.55861

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 7 [17920/49000 (36%)]

Train Loss: 0.619223

Val Loss: 0.55791

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [20480/49000 (42%)]      Train Loss: 0.518251      Val Loss: 0.5594!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [23040/49000 (47%)]      Train Loss: 0.515852      Val Loss: 0.5614!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [25600/49000 (52%)]      Train Loss: 0.551437      Val Loss: 0.5532!
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [28160/49000 (57%)]      Train Loss: 0.575369      Val Loss: 0.5530!
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [30720/49000 (62%)]      Train Loss: 0.571790      Val Loss: 0.5636!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [33280/49000 (68%)]      Train Loss: 0.585769      Val Loss: 0.5629
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [35840/49000 (73%)]      Train Loss: 0.615460      Val Loss: 0.5557
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [38400/49000 (78%)]      Train Loss: 0.524053      Val Loss: 0.5522
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [40960/49000 (83%)]      Train Loss: 0.620406      Val Loss: 0.5511
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [43520/49000 (89%)]      Train Loss: 0.556451      Val Loss: 0.5558
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [46080/49000 (94%)]      Train Loss: 0.543320      Val Loss: 0.54631
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 7 [48640/49000 (99%)]      Train Loss: 0.552232      Val Loss: 0.55240
torch.Size([104, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 8 [0/49000 (0%)]      Train Loss: 1.160033      Val Loss: 0.550710
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 8 [2560/49000 (5%)]      Train Loss: 0.578554      Val Loss: 0.55691
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 8 [5120/49000 (10%)]      Train Loss: 0.554731      Val Loss: 0.53861
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 8 [7680/49000 (16%)]      Train Loss: 0.531410      Val Loss: 0.53831
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [10240/49000 (21%)]

Train Loss: 0.572418

Val Loss: 0.53405

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [12800/49000 (26%)]

Train Loss: 0.526520

Val Loss: 0.53846

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [15360/49000 (31%)]

Train Loss: 0.440627

Val Loss: 0.53731

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [17920/49000 (36%)]

Train Loss: 0.531224

Val Loss: 0.53876

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [20480/49000 (42%)]

Train Loss: 0.545627

Val Loss: 0.56376

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [23040/49000 (47%)]

Train Loss: 0.521682

Val Loss: 0.53179

Saved model this time!

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [25600/49000 (52%)]

Train Loss: 0.598017

Val Loss: 0.54031

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [28160/49000 (57%)]

Train Loss: 0.511356

Val Loss: 0.55501

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [30720/49000 (62%)]

Train Loss: 0.539674

Val Loss: 0.54551

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 8 [33280/49000 (68%)]

Train Loss: 0.579209

Val Loss: 0.54891

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 8 [35840/49000 (73%)]      Train Loss: 0.570621      Val Loss: 0.53321
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 8 [38400/49000 (78%)]      Train Loss: 0.563569      Val Loss: 0.53621
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 8 [40960/49000 (83%)]      Train Loss: 0.606820      Val Loss: 0.53481
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 8 [43520/49000 (89%)]      Train Loss: 0.567834      Val Loss: 0.53421
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 8 [46080/49000 (94%)]      Train Loss: 0.460869      Val Loss: 0.54081
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```



```

torch.Size([256, 3, 224, 224])
Train Epoch: 8 [48640/49000 (99%)]      Train Loss: 0.557109      Val Loss: 0.5298
Saved model this time!
torch.Size([104, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [0/49000 (0%)]      Train Loss: 1.024197      Val Loss: 0.540059
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [2560/49000 (5%)]      Train Loss: 0.573959      Val Loss: 0.5425
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [5120/49000 (10%)]      Train Loss: 0.588238      Val Loss: 0.5282
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [7680/49000 (16%)]      Train Loss: 0.543430      Val Loss: 0.5420
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [10240/49000 (21%)]      Train Loss: 0.545229      Val Loss: 0.5430
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 9 [12800/49000 (26%)]

Train Loss: 0.584119

Val Loss: 0.53936

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 9 [15360/49000 (31%)]

Train Loss: 0.560484

Val Loss: 0.54621

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 9 [17920/49000 (36%)]

Train Loss: 0.567873

Val Loss: 0.53394

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 9 [20480/49000 (42%)]

Train Loss: 0.570251

Val Loss: 0.53194

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

Train Epoch: 9 [23040/49000 (47%)]

Train Loss: 0.517306

Val Loss: 0.52936

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [25600/49000 (52%)]      Train Loss: 0.591225      Val Loss: 0.52487
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [28160/49000 (57%)]      Train Loss: 0.517387      Val Loss: 0.51867
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [30720/49000 (62%)]      Train Loss: 0.570756      Val Loss: 0.53147
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [33280/49000 (68%)]      Train Loss: 0.525327      Val Loss: 0.52697
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [35840/49000 (73%)]      Train Loss: 0.556059      Val Loss: 0.52977
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

```

torch.Size([256, 3, 224, 224])
Train Epoch: 9 [38400/49000 (78%)]      Train Loss: 0.467461      Val Loss: 0.53601
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [40960/49000 (83%)]      Train Loss: 0.574416      Val Loss: 0.52181
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [43520/49000 (89%)]      Train Loss: 0.516058      Val Loss: 0.53231
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [46080/49000 (94%)]      Train Loss: 0.546336      Val Loss: 0.52501
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 9 [48640/49000 (99%)]      Train Loss: 0.550028      Val Loss: 0.53361
torch.Size([104, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [0/49000 (0%)]      Train Loss: 0.964872      Val Loss: 0.542323
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [2560/49000 (5%)]      Train Loss: 0.526727      Val Loss: 0.54011
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [5120/49000 (10%)]      Train Loss: 0.501159      Val Loss: 0.54141
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [7680/49000 (16%)]      Train Loss: 0.560064      Val Loss: 0.53771
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [10240/49000 (21%)]      Train Loss: 0.537552      Val Loss: 0.52091
Saved model this time!
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [12800/49000 (26%)]      Train Loss: 0.536036      Val Loss: 0.53201
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [15360/49000 (31%)]    Train Loss: 0.552141    Val Loss: 0.52564
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [17920/49000 (36%)]    Train Loss: 0.514192    Val Loss: 0.54114
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [20480/49000 (42%)]    Train Loss: 0.568552    Val Loss: 0.54271
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [23040/49000 (47%)]    Train Loss: 0.523196    Val Loss: 0.53821
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [25600/49000 (52%)]    Train Loss: 0.595694    Val Loss: 0.53521
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
```

```

Train Epoch: 10 [28160/49000 (57%)]    Train Loss: 0.527919    Val Loss: 0.52445
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [30720/49000 (62%)]    Train Loss: 0.488037    Val Loss: 0.52175
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [33280/49000 (68%)]    Train Loss: 0.449341    Val Loss: 0.52975
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [35840/49000 (73%)]    Train Loss: 0.530765    Val Loss: 0.53025
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [38400/49000 (78%)]    Train Loss: 0.553016    Val Loss: 0.53905
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [40960/49000 (83%)]    Train Loss: 0.533984    Val Loss: 0.53285
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])

```

```

torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [43520/49000 (89%)]      Train Loss: 0.583151      Val Loss: 0.53158
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [46080/49000 (94%)]      Train Loss: 0.548435      Val Loss: 0.53709
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
torch.Size([256, 3, 224, 224])
Train Epoch: 10 [48640/49000 (99%)]      Train Loss: 0.649393      Val Loss: 0.52628
torch.Size([104, 3, 224, 224])

```

test set: Average loss: 0.5452, Accuracy: 8566/10000 (86%)

(tensor(0.5452, device='cuda:0'), tensor(85.6600))

