Pneumonia_Detection

November 11, 2024

1 Pneumonia Detection from Chest X-Ray Images

Dataset available on kaggle: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

1.1 Install and load modules

[1]:
```
# !pip install timm torchsummary tqdm numpy matplotlib torch torchvision kagglehub
```

[2]:
```
!pip install -q torchmetrics
```

```
                                        891.4/891.4 kB
18.4 MB/s eta 0:00:00
```

[3]:
```
import torch
import numpy as np
import matplotlib.pyplot as plt
import os
from tqdm.notebook import tqdm
import torchmetrics
```

1.2 Download the dataset

[4]:
```
import kagglehub

# Download latest version
path = kagglehub.dataset_download("paultimothymooney/chest-xray-pneumonia")

print("Path to dataset files:", path)
```

```
Downloading from
https://www.kaggle.com/api/v1/datasets/download/paultimothymooney/chest-xray-
pneumonia?dataset_version_number=2…

100%|          | 2.29G/2.29G [00:23<00:00, 104MB/s]

Extracting files…
```

```
Path to dataset files: /root/.cache/kagglehub/datasets/paultimothymooney/chest-
xray-pneumonia/versions/2
```

[5]:

```
!pwd
```

```
/content
```

[6]:

```
! cp -r /root/.cache/kagglehub/datasets/paultimothymooney/chest-xray-pneumonia/
 ↪versions/2 /content/
```

1.3 Define problem parameters

[7]:

```
class config:

    epochs = 1                                     # No. of epochs of training␣
 ↪the model
    lr_custom = 0.000005                           # Learning rate for custom␣
 ↪model
    lr_pretrained = 0.00001                        # Learning rate for␣
 ↪pretrained model
    batch_size = 128                               # Batch Size For Dataset

    model_name = 'tf_efficientnet_b7.ap_in1k'      # Model name (We are going␣
 ↪to import model from timm)

    img_size = 224

    in_mean = [0.485, 0.456, 0.406]  # ImageNet dataset mean
    in_std = [0.229, 0.224, 0.225]   # ImageNet dataset std

    pn_mean = [0.4752]  # Pneumonia dataset mean
    pn_std = [0.2234]   # Pneumonia dataset std

    # Going to be use for loading dataset
    data_dir = '/content/2/chest_xray/chest_xray'   # Data Directory
    test_dir = 'test'                               # Test folder name in data␣
 ↪directory
    train_dir = 'train'                             # Train folder name in data␣
 ↪directory
    valid_dir = 'val'                               # Valid folder name in data␣
 ↪directory
```

1.4 Choose available device

[8]:
```
device = torch.device('cuda' if torch.cuda.is_available() else 'cpu')

print(f"device: {device}")

if torch.cuda.is_available():
    !nvidia-smi
```

```
device: cuda
Mon Nov 11 05:27:06 2024
+--------------------------------------------------------------------------------
--------+
| NVIDIA-SMI 535.104.05             Driver Version: 535.104.05   CUDA Version:
12.2     |
|-----------------------------------------+----------------------+--------------
--------+
| GPU  Name                 Persistence-M | Bus-Id        Disp.A | Volatile
Uncorr. ECC |
| Fan  Temp   Perf          Pwr:Usage/Cap |         Memory-Usage | GPU-Util
Compute M. |
|                                         |                      |
MIG M. |
|=========================================+======================+==============
========|
|   0  Tesla T4                       Off | 00000000:00:04.0 Off |
0 |
| N/A   51C    P8              10W /  70W |      3MiB / 15360MiB |      0%
Default |
|                                         |                      |
N/A |
+-----------------------------------------+----------------------+--------------
--------+

+--------------------------------------------------------------------------------
--------+
| Processes:
|
|  GPU   GI   CI        PID   Type   Process name                            GPU
Memory |
|        ID   ID
Usage      |
|================================================================================
========|
|  No running processes found
|
+--------------------------------------------------------------------------------
--------+
```

1.5 Image transformation pipelines

[9]:

```
from torchvision import transforms as T, datasets
from torch.utils.data import Dataset
```

[10]:

```
class Gray2RGB:
    def __call__(self, image):
        return image.repeat(3, 1, 1) # Repeat the single channel across 3␣
 ↪channels to convert to RGB

train_transform_custom = T.Compose([
                            T.Resize(size=(config.img_size, config.img_size)),
                            T.RandomHorizontalFlip(p=0.5),
                            T.RandomRotation(degrees=(-20, 20)),
                            T.ToTensor(),
                            T.Normalize(config.pn_mean, config.pn_std),
                            Gray2RGB(),
                        ])

valid_transform_custom = T.Compose([
                            T.Resize(size=(config.img_size, config.img_size)),
                            T.ToTensor(),
                            T.Normalize(config.pn_mean, config.pn_std),
                            Gray2RGB(),
                        ])

test_transform_custom = T.Compose([
                            T.Resize(size=(config.img_size, config.img_size)),
                            T.ToTensor(),
                            T.Normalize(config.pn_mean, config.pn_std),
                            Gray2RGB(),
                        ])
```

[11]:

```
train_transform_pretrained = T.Compose([
                            T.Resize(size=(config.img_size, config.img_size)),
                            T.RandomHorizontalFlip(p=0.5),
                            T.RandomRotation(degrees=(-20, 20)),
                            T.ToTensor(),
                            Gray2RGB(),
                            T.Normalize(config.in_mean, config.in_std)
                        ])

valid_transform_pretrained = T.Compose([
                            T.Resize(size=(config.img_size, config.img_size)),
                            T.ToTensor(),
                            Gray2RGB(),
                            T.Normalize(config.in_mean, config.in_std)
                        ])

test_transform_pretrained = T.Compose([
                            T.Resize(size=(config.img_size, config.img_size)),
                            T.ToTensor(),
                            Gray2RGB(),
                            T.Normalize(config.in_mean, config.in_std)
                        ])
```

1.6 Pre-load datasets

[12]:

```
train_path = os.path.join(config.data_dir, config.train_dir)
valid_path = os.path.join(config.data_dir, config.valid_dir)
test_path = os.path.join(config.data_dir, config.test_dir)
```

[13]:

```
trainset_temp = datasets.ImageFolder(train_path, transform=None)
validset_temp = datasets.ImageFolder(valid_path, transform=None)
testset_temp = datasets.ImageFolder(test_path, transform=None)
```

[14]:

```
print(f"Train set size : {len(trainset_temp)}")
print(f"Valid set size : {len(validset_temp)}")
print(f"Test set size : {len(testset_temp)}")
```

```
Train set size : 5216
Valid set size : 16
Test set size : 624
```

[15]:

```
# Loading the datasets into the memory to boost performance

train_img_list = []
train_label_list = []
for img, label in trainset_temp:
    img = img.convert("L")
    train_img_list.append(img)
    train_label_list.append(label)

valid_img_list = []
valid_label_list = []
for img, label in validset_temp:
    img = img.convert("L")
    valid_img_list.append(img)
    valid_label_list.append(label)

test_img_list = []
test_label_list = []
for img, label in testset_temp:
    img = img.convert("L")
    test_img_list.append(img)
    test_label_list.append(label)
```

1.7 Create custom dataset

[16]:

```
class CustomDataset(Dataset):

    def __init__(self, img_list, label_list, transform=None):
        super().__init__()
        self.img_list = img_list
        self.label_list = label_list
        self.transform = transform

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

    def __getitem__(self, index):
        image = self.img_list[index]
        label = self.label_list[index]

        if self.transform:
            image = self.transform(image)

        return image, label
```

[17]:

```
trainset_custom = CustomDataset(train_img_list, train_label_list,␣
 ↪train_transform_custom)
trainset_pretrained = CustomDataset(train_img_list, train_label_list,␣
 ↪train_transform_pretrained)

validset_custom = CustomDataset(valid_img_list, valid_label_list,␣
 ↪valid_transform_custom)
validset_pretrained = CustomDataset(valid_img_list, valid_label_list,␣
 ↪valid_transform_pretrained)

testset_custom = CustomDataset(test_img_list, test_label_list,␣
 ↪test_transform_custom)
testset_pretrained = CustomDataset(test_img_list, test_label_list,␣
 ↪test_transform_pretrained)
```

1.8 Count samples in each class

[18]:

```
from collections import Counter

# Function for counting the number of samples in each class
def count_classes(dataset):
    label_counts = Counter()
    for _, label in dataset:
        label_counts[label] += 1
    return label_counts
```

[19]:

```
label_counts_train = count_classes(trainset_custom)

print(label_counts_train)  # Counter({1: 3875, 0: 1341})
```

```
Counter({1: 3875, 0: 1341})
```

[20]:

```
label_counts_valid = count_classes(validset_custom)

print(label_counts_valid)
```

```
Counter({0: 8, 1: 8})
```

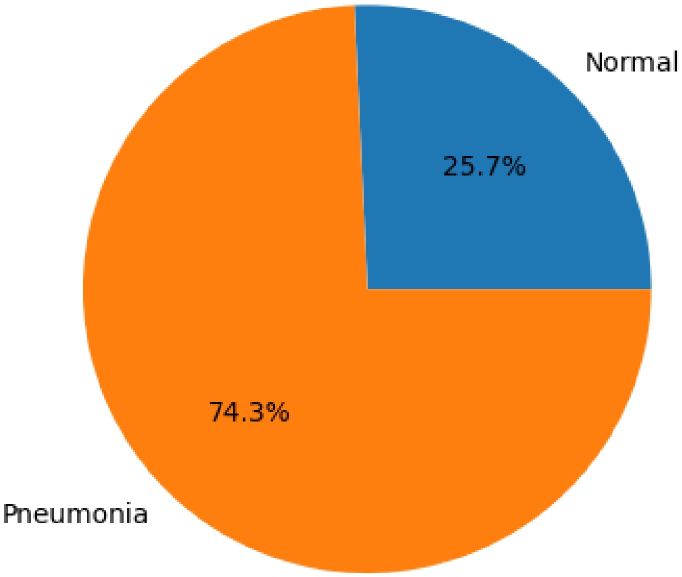
[21]:

```
label_counts_test = count_classes(testset_custom)

print(label_counts_test)
```

```
Counter({1: 390, 0: 234})
```

[22]:

```
sizes = [label_counts_train[0], label_counts_train[1]]
labels = ['Normal', 'Pneumonia']

plt.pie(sizes, labels=labels, autopct='%1.1f%%');
```



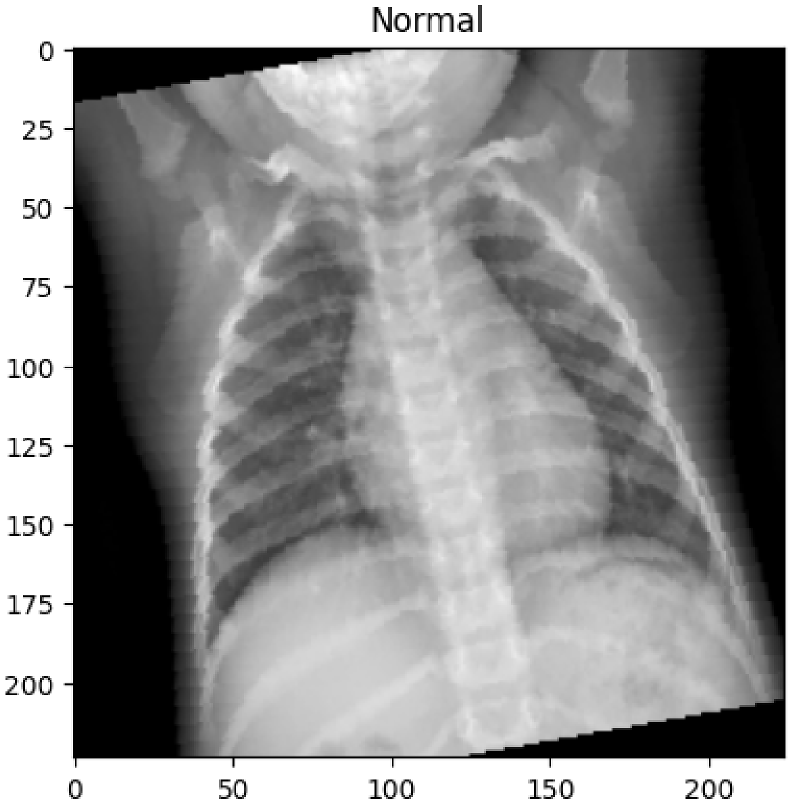
1.9 Calculate the weight for each class

[23]:
```
# calculating weight for each class
from sklearn.utils.class_weight import compute_class_weight

# Get class labels from dataset
y_labels = [label for _, label in trainset_custom]
class_weights = compute_class_weight(class_weight='balanced', classes=np.
 ↪unique(y_labels), y=y_labels)

# Convert to torch tensor
class_weights = torch.tensor(class_weights, dtype=torch.float)

print(f'class_weights: {class_weights}')
```

```
class_weights: tensor([1.9448, 0.6730])
```

[24]:
```
def show_image(image, label='-', mean=None, std=None):

    image = image.permute(1, 2, 0)

    if mean and std:
        mean = torch.FloatTensor(mean)
        std = torch.FloatTensor(std)
        image = image*std + mean
        image = torch.clip(image, 0, 1)
        image = image.numpy()

    plt.imshow(image, cmap="gray")
    plt.title(label)
```

```
[25]: sample_img, sample_label = trainset_custom[2]

      show_image(sample_img, labels[sample_label], config.pn_mean, config.pn_std)
```



1.10 Load dataset into batches

```
[26]: from torch.utils.data import DataLoader
      from torchvision.utils import make_grid
```

```
[27]: train_loader_custom = DataLoader(trainset_custom, batch_size=config.batch_size,␣
        ↪shuffle=True)
```

```
valid_loader_custom = DataLoader(validset_custom, batch_size=config.batch_size,␣
 ↪shuffle=False)
test_loader_custom = DataLoader(testset_custom, batch_size=config.batch_size,␣
 ↪shuffle=False)
```

[28]:

```
train_loader_pretrained = DataLoader(trainset_pretrained, batch_size=config.
 ↪batch_size, shuffle=True)
valid_loader_pretrained = DataLoader(validset_pretrained, batch_size=config.
 ↪batch_size, shuffle=False)
test_loader_pretrained = DataLoader(testset_pretrained, batch_size=config.
 ↪batch_size, shuffle=False)
```

[29]:

```
print(f"No. of batches in train loader : {len(train_loader_custom)}")
print(f"No. of batches in valid loader : {len(valid_loader_custom)}")
print(f"No. of batches in test loader : {len(test_loader_custom)}")
```

```
No. of batches in train loader : 41
No. of batches in valid loader : 1
No. of batches in test loader : 5
```

[30]:

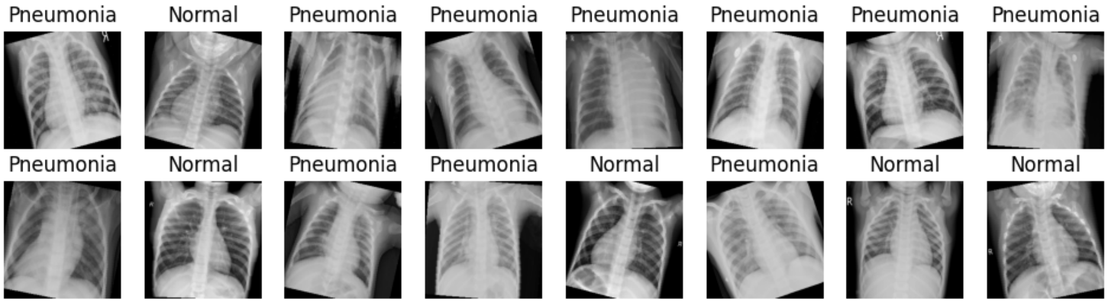
```
train_iter = iter(train_loader_custom)
sample_batch, sample_batch_label = next(train_iter)

print(f'sample_batch.shape: {sample_batch.shape}')
```

```
sample_batch.shape: torch.Size([128, 3, 224, 224])
```

1.11 Show samples of dataset

[31]:

```
fig, ax = plt.subplots(2, 8, figsize=(12,3))

sample_batch, sample_batch_label = next(iter(train_loader_custom))
for n in range(2):
    for m in range(8):
        image = sample_batch[n*8+m].cpu().permute(1, 2, 0)
        mean = torch.FloatTensor(config.pn_mean)
        std = torch.FloatTensor(config.pn_std)
        image = image*std + mean
        image = torch.clip(image, 0, 1)
        image = image.numpy()

        ax[n,m].imshow(image, cmap="gray")
        ax[n,m].grid(False)
        ax[n,m].axis(False)
        ax[n,m].tick_params(labelbottom=False, labelleft=False)
        ax[n,m].set_title(f'{labels[sample_batch_label[n*8+m]]}')
```



1.12 Design custom CNN model

[32]:

```
import torch.nn as nn
import torch.nn.functional as F
```

[33]:

```
# Calculating the size of the final layer before feeding into the linear layer
X = torch.zeros(1, 3, config.img_size, config.img_size)

conv1 = nn.Sequential(
                nn.Conv2d(3, 8, 3, stride=1, padding=1),
                nn.BatchNorm2d(8),
                nn.ReLU(inplace=True),
                nn.MaxPool2d(2,2))

conv2 = nn.Sequential(
                nn.Conv2d(8, 16, 3, stride=1, padding=1),
                nn.BatchNorm2d(16),
                nn.ReLU(inplace=True),
                nn.MaxPool2d(2,2))

conv3 = nn.Sequential(
                nn.Conv2d(16, 32, 3, stride=1, padding=1),
                nn.BatchNorm2d(32),
                nn.ReLU(inplace=True),
                nn.MaxPool2d(2,2))

conv4 = nn.Sequential(
                nn.Conv2d(32, 64, 3, stride=1, padding=1),
                nn.BatchNorm2d(64),
                nn.ReLU(inplace=True),
                nn.MaxPool2d(2,2))

conv5 = nn.Sequential(
                nn.Conv2d(64, 128, 3, stride=1, padding=1),
                nn.BatchNorm2d(128),
                nn.ReLU(inplace=True),
                nn.MaxPool2d(2,2))

X = conv1(X)
X = conv2(X)
X = conv3(X)
X = conv4(X)
X = conv5(X)

X.shape # [1, 128, 7, 7]
```

[33]:

```
torch.Size([1, 128, 7, 7])
```

[34]:

```
class ConvolutionalNetwork(nn.Module):
    def __init__(self):
        super().__init__()

        self.conv1 = nn.Sequential(
                        nn.Conv2d(3, 8, 3, stride=1, padding=1),
                        nn.ReLU(inplace=True),
                        nn.BatchNorm2d(8),
                        nn.MaxPool2d(2,2))

        self.conv2 = nn.Sequential(
                        nn.Conv2d(8, 16, 3, stride=1, padding=1),
                        nn.ReLU(inplace=True),
                        nn.BatchNorm2d(16),
                        nn.MaxPool2d(2,2))

        self.conv3 = nn.Sequential(
                        nn.Conv2d(16, 32, 3, stride=1, padding=1),
                        nn.ReLU(inplace=True),
                        nn.BatchNorm2d(32),
                        nn.MaxPool2d(2,2))

        self.conv4 = nn.Sequential(
                        nn.Conv2d(32, 64, 3, stride=1, padding=1),
                        nn.ReLU(inplace=True),
                        nn.BatchNorm2d(64),
                        nn.MaxPool2d(2,2))

        self.conv5 = nn.Sequential(
                        nn.Conv2d(64, 128, 3, stride=1, padding=1),
                        nn.ReLU(inplace=True),
                        nn.BatchNorm2d(128),
                        nn.MaxPool2d(2,2))
        self.fc = nn.Sequential(
                    nn.Linear(128*7*7, 512),
                    nn.ReLU(inplace=True),
                    nn.BatchNorm1d(512),
                    nn.Dropout(0.5),
                    nn.Linear(512, 2))

    def forward(self, x):
        x = self.conv1(x)
        x = self.conv2(x)
        x = self.conv3(x)
        x = self.conv4(x)
        x = self.conv5(x)
        x = x.view(x.shape[0], -1)
        x = self.fc(x)
        return x
```

[35]:

```
torch.manual_seed(42)
torch.cuda.manual_seed(42)
cnn_model = ConvolutionalNetwork()

cnn_model.to(device)
```

[35]:

```
ConvolutionalNetwork(
  (conv1): Sequential(
    (0): Conv2d(3, 8, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU(inplace=True)
    (2): BatchNorm2d(8, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  )
  (conv2): Sequential(
    (0): Conv2d(8, 16, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU(inplace=True)
    (2): BatchNorm2d(16, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  )
  (conv3): Sequential(
    (0): Conv2d(16, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU(inplace=True)
    (2): BatchNorm2d(32, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
```

```
  )
  (conv4): Sequential(
    (0): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU(inplace=True)
    (2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  )
  (conv5): Sequential(
    (0): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
    (1): ReLU(inplace=True)
    (2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (3): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  )
  (fc): Sequential(
    (0): Linear(in_features=6272, out_features=512, bias=True)
    (1): ReLU(inplace=True)
    (2): BatchNorm1d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (3): Dropout(p=0.5, inplace=False)
    (4): Linear(in_features=512, out_features=2, bias=True)
  )
)
```

[36]:

```python
from torchsummary import summary

summary(cnn_model, input_size=(3, config.img_size, config.img_size))
```

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
            Conv2d-1          [-1, 8, 224, 224]             224
              ReLU-2          [-1, 8, 224, 224]               0
       BatchNorm2d-3          [-1, 8, 224, 224]              16
         MaxPool2d-4          [-1, 8, 112, 112]               0
            Conv2d-5         [-1, 16, 112, 112]           1,168
              ReLU-6         [-1, 16, 112, 112]               0
       BatchNorm2d-7         [-1, 16, 112, 112]              32
         MaxPool2d-8           [-1, 16, 56, 56]               0
            Conv2d-9           [-1, 32, 56, 56]           4,640
             ReLU-10           [-1, 32, 56, 56]               0
      BatchNorm2d-11           [-1, 32, 56, 56]              64
        MaxPool2d-12           [-1, 32, 28, 28]               0
           Conv2d-13           [-1, 64, 28, 28]          18,496
           ReLU-14          [-1, 64, 28, 28]               0
      BatchNorm2d-15          [-1, 64, 28, 28]             128
        MaxPool2d-16          [-1, 64, 14, 14]               0
           Conv2d-17         [-1, 128, 14, 14]          73,856
             ReLU-18         [-1, 128, 14, 14]               0
      BatchNorm2d-19         [-1, 128, 14, 14]             256
        MaxPool2d-20           [-1, 128, 7, 7]               0
           Linear-21                  [-1, 512]       3,211,776
             ReLU-22                  [-1, 512]               0
      BatchNorm1d-23                  [-1, 512]           1,024
          Dropout-24                  [-1, 512]               0
           Linear-25                    [-1, 2]           1,026
================================================================
Total params: 3,312,706
Trainable params: 3,312,706
Non-trainable params: 0
----------------------------------------------------------------
Input size (MB): 0.57
Forward/backward pass size (MB): 19.30
Params size (MB): 12.64
Estimated Total Size (MB): 32.51
----------------------------------------------------------------
```

1.13 Define pretrained models for fine-tuning

[37]:

```python
from torch import nn
import torch.nn.functional as F
import timm

# Models that we will be using:
# tf_efficientnet_b4.ap_in1k
# tf_efficientnet_b7.ap_in1k
# resnet50.tv_in1k
# resnet101.tv_in1k

torch.manual_seed(42)
torch.cuda.manual_seed(42)

config.model_name = 'tf_efficientnet_b7.ap_in1k'

model = timm.create_model(config.model_name, pretrained=True)
model
```

```
/usr/local/lib/python3.10/dist-packages/huggingface_hub/utils/_token.py:89:
UserWarning:
The secret `HF_TOKEN` does not exist in your Colab secrets.
To authenticate with the Hugging Face Hub, create a token in your settings tab
(https://huggingface.co/settings/tokens), set it as secret in your Google Colab
```

```
and restart your session.
You will be able to reuse this secret in all of your notebooks.
Please note that authentication is recommended but still optional to access
public models or datasets.
  warnings.warn(

model.safetensors:   0%|          | 0.00/267M [00:00<?, ?B/s]
```

[37]:

```
EfficientNet(
  (conv_stem): Conv2dSame(3, 64, kernel_size=(3, 3), stride=(2, 2), bias=False)
  (bn1): BatchNormAct2d(
    64, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
    (drop): Identity()
    (act): SiLU(inplace=True)
  )
  (blocks): Sequential(
    (0): Sequential(
      (0): DepthwiseSeparableConv(
        (conv_dw): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), groups=64, bias=False)
        (bn1): BatchNormAct2d(
          64, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pw): Conv2d(64, 32, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn2): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): DepthwiseSeparableConv(
        (conv_dw): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), groups=32, bias=False)
        (bn1): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
    (aa): Identity()
    (se): SqueezeExcite(
     (conv_reduce): Conv2d(32, 8, kernel_size=(1, 1), stride=(1, 1))
     (act1): SiLU(inplace=True)
     (conv_expand): Conv2d(8, 32, kernel_size=(1, 1), stride=(1, 1))
     (gate): Sigmoid()
    )
    (conv_pw): Conv2d(32, 32, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn2): BatchNormAct2d(
     32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
     (drop): Identity()
     (act): Identity()
    )
    (drop_path): Identity()
   )
   (2): DepthwiseSeparableConv(
    (conv_dw): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), groups=32, bias=False)
    (bn1): BatchNormAct2d(
     32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
     (drop): Identity()
     (act): SiLU(inplace=True)
    )
    (aa): Identity()
    (se): SqueezeExcite(
     (conv_reduce): Conv2d(32, 8, kernel_size=(1, 1), stride=(1, 1))
     (act1): SiLU(inplace=True)
     (conv_expand): Conv2d(8, 32, kernel_size=(1, 1), stride=(1, 1))
     (gate): Sigmoid()
    )
    (conv_pw): Conv2d(32, 32, kernel_size=(1, 1), stride=(1, 1), bias=False)
    (bn2): BatchNormAct2d(
     32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
     (drop): Identity()
     (act): Identity()
    )
    (drop_path): Identity()
   )
   (3): DepthwiseSeparableConv(
    (conv_dw): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), groups=32, bias=False)
    (bn1): BatchNormAct2d(
     32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
     (drop): Identity()
     (act): SiLU(inplace=True)
    )
    (aa): Identity()
```

```
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(32, 8, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(8, 32, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pw): Conv2d(32, 32, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn2): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (1): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(32, 192, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          192, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2dSame(192, 192, kernel_size=(3, 3), stride=(2, 2),
groups=192, bias=False)
        (bn2): BatchNormAct2d(
          192, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(192, 8, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(8, 192, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(192, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (5): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (6): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (2): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2dSame(288, 288, kernel_size=(5, 5), stride=(2, 2),
groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (5): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (6): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (3): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2dSame(480, 480, kernel_size=(3, 3), stride=(2, 2),
groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (5): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (6): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (7): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (8): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
```

```
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (9): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (4): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
```

```
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (5): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (6): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (7): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (8): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (9): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (5): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
```

```
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2dSame(1344, 1344, kernel_size=(5, 5), stride=(2, 2),
groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
```

```
     (gate): Sigmoid()
    )
    (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (bn3): BatchNormAct2d(
     384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
     (drop): Identity()
     (act): Identity()
    )
    (drop_path): Identity()
   )
   (5): InvertedResidual(
    (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (bn1): BatchNormAct2d(
     2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
     (drop): Identity()
     (act): SiLU(inplace=True)
    )
    (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
    (bn2): BatchNormAct2d(
     2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
     (drop): Identity()
     (act): SiLU(inplace=True)
    )
    (aa): Identity()
    (se): SqueezeExcite(
     (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
     (act1): SiLU(inplace=True)
     (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
     (gate): Sigmoid()
    )
    (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (bn3): BatchNormAct2d(
     384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
     (drop): Identity()
     (act): Identity()
    )
    (drop_path): Identity()
   )
   (6): InvertedResidual(
    (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (bn1): BatchNormAct2d(
     2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
```

```
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (7): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (8): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (9): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (10): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (11): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (12): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (6): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 640, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          640, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(640, 3840, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(3840, 3840, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=3840, bias=False)
        (bn2): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(3840, 160, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(160, 3840, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(3840, 640, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          640, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(640, 3840, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(3840, 3840, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=3840, bias=False)
        (bn2): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(3840, 160, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(160, 3840, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(3840, 640, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          640, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(640, 3840, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(3840, 3840, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=3840, bias=False)
        (bn2): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(3840, 160, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(160, 3840, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(3840, 640, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          640, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
  )
  (conv_head): Conv2d(640, 2560, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn2): BatchNormAct2d(
    2560, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
    (drop): Identity()
    (act): SiLU(inplace=True)
  )
  (global_pool): SelectAdaptivePool2d(pool_type=avg,
flatten=Flatten(start_dim=1, end_dim=-1))
  (classifier): Linear(in_features=2560, out_features=1000, bias=True)
)
```

[38]:

```
for param in model.parameters():
    param.requires_grad = False

# Number of in- and out- features for each model:
# 1792 -> 768  -> 256 -> 2   EfficientNet_b4
# 2560 -> 1024 -> 256 -> 2   EfficientNet_b7
# 2048 -> 1024 -> 256 -> 2   ResNet50
# 2048 -> 1024 -> 256 -> 2   ResNet101

# model.classifier for efficientnet family
# model.fc for ResNet family
model.classifier = nn.Sequential(
    nn.Linear(in_features=2560, out_features=1024),
    nn.ReLU(),
    nn.Dropout(p=0.4),
    nn.Linear(in_features=1024, out_features=256),
    nn.ReLU(),
    nn.Linear(in_features=256, out_features=2),
)

model.to(device)
```

[38]:

```
EfficientNet(
  (conv_stem): Conv2dSame(3, 64, kernel_size=(3, 3), stride=(2, 2), bias=False)
  (bn1): BatchNormAct2d(
    64, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
    (drop): Identity()
    (act): SiLU(inplace=True)
  )
  (blocks): Sequential(
    (0): Sequential(
      (0): DepthwiseSeparableConv(
        (conv_dw): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), groups=64, bias=False)
        (bn1): BatchNormAct2d(
          64, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(64, 16, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(16, 64, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pw): Conv2d(64, 32, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn2): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): DepthwiseSeparableConv(
        (conv_dw): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), groups=32, bias=False)
        (bn1): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(32, 8, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(8, 32, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pw): Conv2d(32, 32, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn2): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): DepthwiseSeparableConv(
        (conv_dw): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), groups=32, bias=False)
        (bn1): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(32, 8, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(8, 32, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pw): Conv2d(32, 32, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn2): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): DepthwiseSeparableConv(
        (conv_dw): Conv2d(32, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1,
1), groups=32, bias=False)
        (bn1): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(32, 8, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(8, 32, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pw): Conv2d(32, 32, kernel_size=(1, 1), stride=(1, 1), bias=False)
        (bn2): BatchNormAct2d(
          32, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (1): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(32, 192, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          192, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2dSame(192, 192, kernel_size=(3, 3), stride=(2, 2),
groups=192, bias=False)
        (bn2): BatchNormAct2d(
          192, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(192, 8, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(8, 192, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(192, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
```

```
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (5): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (6): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(288, 288, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=288, bias=False)
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 48, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          48, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (2): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(48, 288, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2dSame(288, 288, kernel_size=(5, 5), stride=(2, 2),
groups=288, bias=False)
```

```
        (bn2): BatchNormAct2d(
          288, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(288, 12, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(12, 288, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(288, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (5): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (6): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(480, 480, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 80, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          80, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (3): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(80, 480, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2dSame(480, 480, kernel_size=(3, 3), stride=(2, 2),
groups=480, bias=False)
        (bn2): BatchNormAct2d(
          480, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(480, 20, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(20, 480, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(480, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (5): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (6): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (7): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (8): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (9): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 160, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          160, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (4): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(160, 960, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(960, 960, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=960, bias=False)
        (bn2): BatchNormAct2d(
          960, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(960, 40, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(40, 960, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(960, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (5): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (6): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (7): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (8): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (9): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(1344, 1344, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 224, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          224, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (5): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(224, 1344, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2dSame(1344, 1344, kernel_size=(5, 5), stride=(2, 2),
groups=1344, bias=False)
        (bn2): BatchNormAct2d(
          1344, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(1344, 56, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(56, 1344, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(1344, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (4): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (5): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (6): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (7): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (8): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (9): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (10): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (11): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (12): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(5, 5), stride=(1, 1),
padding=(2, 2), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 384, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          384, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
    (6): Sequential(
      (0): InvertedResidual(
        (conv_pw): Conv2d(384, 2304, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(2304, 2304, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=2304, bias=False)
        (bn2): BatchNormAct2d(
          2304, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(2304, 96, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(96, 2304, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(2304, 640, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          640, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (1): InvertedResidual(
        (conv_pw): Conv2d(640, 3840, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(3840, 3840, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=3840, bias=False)
        (bn2): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(3840, 160, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(160, 3840, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(3840, 640, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          640, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (2): InvertedResidual(
        (conv_pw): Conv2d(640, 3840, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(3840, 3840, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=3840, bias=False)
        (bn2): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(3840, 160, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(160, 3840, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(3840, 640, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          640, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
      (3): InvertedResidual(
        (conv_pw): Conv2d(640, 3840, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn1): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (conv_dw): Conv2d(3840, 3840, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), groups=3840, bias=False)
        (bn2): BatchNormAct2d(
          3840, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): SiLU(inplace=True)
        )
        (aa): Identity()
        (se): SqueezeExcite(
          (conv_reduce): Conv2d(3840, 160, kernel_size=(1, 1), stride=(1, 1))
          (act1): SiLU(inplace=True)
          (conv_expand): Conv2d(160, 3840, kernel_size=(1, 1), stride=(1, 1))
          (gate): Sigmoid()
        )
        (conv_pwl): Conv2d(3840, 640, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (bn3): BatchNormAct2d(
          640, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
          (drop): Identity()
          (act): Identity()
        )
        (drop_path): Identity()
      )
    )
```

```
  )
  (conv_head): Conv2d(640, 2560, kernel_size=(1, 1), stride=(1, 1), bias=False)
  (bn2): BatchNormAct2d(
    2560, eps=0.001, momentum=0.1, affine=True, track_running_stats=True
    (drop): Identity()
    (act): SiLU(inplace=True)
  )
  (global_pool): SelectAdaptivePool2d(pool_type=avg,
flatten=Flatten(start_dim=1, end_dim=-1))
  (classifier): Sequential(
    (0): Linear(in_features=2560, out_features=1024, bias=True)
    (1): ReLU()
    (2): Dropout(p=0.4, inplace=False)
    (3): Linear(in_features=1024, out_features=256, bias=True)
    (4): ReLU()
    (5): Linear(in_features=256, out_features=2, bias=True)
  )
)
```

[39]:

```python
from torchsummary import summary

summary(model, input_size=(3, config.img_size, config.img_size))
```

```
----------------------------------------------------------------
        Layer (type)               Output Shape         Param #
================================================================
        Conv2dSame-1         [-1, 64, 112, 112]           1,728
          Identity-2         [-1, 64, 112, 112]               0
              SiLU-3         [-1, 64, 112, 112]               0
    BatchNormAct2d-4         [-1, 64, 112, 112]             128
            Conv2d-5         [-1, 64, 112, 112]             576
          Identity-6         [-1, 64, 112, 112]               0
              SiLU-7         [-1, 64, 112, 112]               0
    BatchNormAct2d-8         [-1, 64, 112, 112]             128
          Identity-9         [-1, 64, 112, 112]               0
           Conv2d-10             [-1, 16, 1, 1]           1,040
             SiLU-11             [-1, 16, 1, 1]               0
           Conv2d-12             [-1, 64, 1, 1]           1,088
          Sigmoid-13             [-1, 64, 1, 1]               0
    SqueezeExcite-14         [-1, 64, 112, 112]               0
           Conv2d-15         [-1, 32, 112, 112]           2,048
         Identity-16         [-1, 32, 112, 112]               0
         Identity-17         [-1, 32, 112, 112]               0
   BatchNormAct2d-18         [-1, 32, 112, 112]              64
DepthwiseSeparableConv-19         [-1, 32, 112, 112]               0
           Conv2d-20         [-1, 32, 112, 112]             288
         Identity-21         [-1, 32, 112, 112]               0
```

```
             SiLU-22         [-1, 32, 112, 112]               0
   BatchNormAct2d-23         [-1, 32, 112, 112]              64
         Identity-24         [-1, 32, 112, 112]               0
           Conv2d-25              [-1, 8, 1, 1]             264
             SiLU-26              [-1, 8, 1, 1]               0
           Conv2d-27             [-1, 32, 1, 1]             288
          Sigmoid-28             [-1, 32, 1, 1]               0
    SqueezeExcite-29         [-1, 32, 112, 112]               0
           Conv2d-30         [-1, 32, 112, 112]           1,024
         Identity-31         [-1, 32, 112, 112]               0
         Identity-32         [-1, 32, 112, 112]               0
   BatchNormAct2d-33         [-1, 32, 112, 112]              64
         Identity-34         [-1, 32, 112, 112]               0
DepthwiseSeparableConv-35         [-1, 32, 112, 112]               0
           Conv2d-36         [-1, 32, 112, 112]             288
         Identity-37         [-1, 32, 112, 112]               0
             SiLU-38         [-1, 32, 112, 112]               0
   BatchNormAct2d-39         [-1, 32, 112, 112]              64
         Identity-40         [-1, 32, 112, 112]               0
           Conv2d-41              [-1, 8, 1, 1]             264
             SiLU-42              [-1, 8, 1, 1]               0
           Conv2d-43             [-1, 32, 1, 1]             288
          Sigmoid-44             [-1, 32, 1, 1]               0
    SqueezeExcite-45         [-1, 32, 112, 112]               0
           Conv2d-46         [-1, 32, 112, 112]           1,024
         Identity-47         [-1, 32, 112, 112]               0
         Identity-48         [-1, 32, 112, 112]               0
   BatchNormAct2d-49         [-1, 32, 112, 112]              64
         Identity-50         [-1, 32, 112, 112]               0
DepthwiseSeparableConv-51         [-1, 32, 112, 112]               0
           Conv2d-52         [-1, 32, 112, 112]             288
         Identity-53         [-1, 32, 112, 112]               0
             SiLU-54         [-1, 32, 112, 112]               0
   BatchNormAct2d-55         [-1, 32, 112, 112]              64
         Identity-56         [-1, 32, 112, 112]               0
           Conv2d-57              [-1, 8, 1, 1]             264
             SiLU-58              [-1, 8, 1, 1]               0
           Conv2d-59             [-1, 32, 1, 1]             288
          Sigmoid-60             [-1, 32, 1, 1]               0
    SqueezeExcite-61         [-1, 32, 112, 112]               0
           Conv2d-62         [-1, 32, 112, 112]           1,024
         Identity-63         [-1, 32, 112, 112]               0
         Identity-64         [-1, 32, 112, 112]               0
   BatchNormAct2d-65         [-1, 32, 112, 112]              64
         Identity-66         [-1, 32, 112, 112]               0
DepthwiseSeparableConv-67         [-1, 32, 112, 112]               0
           Conv2d-68        [-1, 192, 112, 112]           6,144
         Identity-69        [-1, 192, 112, 112]               0
```

```
             SiLU-70        [-1, 192, 112, 112]               0
   BatchNormAct2d-71        [-1, 192, 112, 112]             384
       Conv2dSame-72          [-1, 192, 56, 56]           1,728
         Identity-73          [-1, 192, 56, 56]               0
             SiLU-74          [-1, 192, 56, 56]               0
   BatchNormAct2d-75          [-1, 192, 56, 56]             384
         Identity-76          [-1, 192, 56, 56]               0
           Conv2d-77              [-1, 8, 1, 1]           1,544
             SiLU-78              [-1, 8, 1, 1]               0
           Conv2d-79            [-1, 192, 1, 1]           1,728
          Sigmoid-80            [-1, 192, 1, 1]               0
    SqueezeExcite-81          [-1, 192, 56, 56]               0
           Conv2d-82           [-1, 48, 56, 56]           9,216
         Identity-83           [-1, 48, 56, 56]               0
         Identity-84           [-1, 48, 56, 56]               0
   BatchNormAct2d-85           [-1, 48, 56, 56]              96
 InvertedResidual-86           [-1, 48, 56, 56]               0
           Conv2d-87          [-1, 288, 56, 56]          13,824
         Identity-88          [-1, 288, 56, 56]               0
             SiLU-89          [-1, 288, 56, 56]               0
   BatchNormAct2d-90          [-1, 288, 56, 56]             576
           Conv2d-91          [-1, 288, 56, 56]           2,592
         Identity-92          [-1, 288, 56, 56]               0
             SiLU-93          [-1, 288, 56, 56]               0
   BatchNormAct2d-94          [-1, 288, 56, 56]             576
         Identity-95          [-1, 288, 56, 56]               0
           Conv2d-96             [-1, 12, 1, 1]           3,468
             SiLU-97             [-1, 12, 1, 1]               0
           Conv2d-98            [-1, 288, 1, 1]           3,744
          Sigmoid-99            [-1, 288, 1, 1]               0
   SqueezeExcite-100          [-1, 288, 56, 56]               0
          Conv2d-101           [-1, 48, 56, 56]          13,824
        Identity-102           [-1, 48, 56, 56]               0
        Identity-103           [-1, 48, 56, 56]               0
  BatchNormAct2d-104           [-1, 48, 56, 56]              96
        Identity-105           [-1, 48, 56, 56]               0
InvertedResidual-106           [-1, 48, 56, 56]               0
          Conv2d-107          [-1, 288, 56, 56]          13,824
        Identity-108          [-1, 288, 56, 56]               0
            SiLU-109          [-1, 288, 56, 56]               0
  BatchNormAct2d-110          [-1, 288, 56, 56]             576
          Conv2d-111          [-1, 288, 56, 56]           2,592
        Identity-112          [-1, 288, 56, 56]               0
            SiLU-113          [-1, 288, 56, 56]               0
  BatchNormAct2d-114          [-1, 288, 56, 56]             576
        Identity-115          [-1, 288, 56, 56]               0
          Conv2d-116             [-1, 12, 1, 1]           3,468
            SiLU-117             [-1, 12, 1, 1]               0
          Conv2d-118          [-1, 288, 1, 1]           3,744
         Sigmoid-119          [-1, 288, 1, 1]               0
   SqueezeExcite-120        [-1, 288, 56, 56]               0
          Conv2d-121         [-1, 48, 56, 56]          13,824
        Identity-122         [-1, 48, 56, 56]               0
        Identity-123         [-1, 48, 56, 56]               0
  BatchNormAct2d-124         [-1, 48, 56, 56]              96
        Identity-125         [-1, 48, 56, 56]               0
InvertedResidual-126         [-1, 48, 56, 56]               0
          Conv2d-127        [-1, 288, 56, 56]          13,824
        Identity-128        [-1, 288, 56, 56]               0
            SiLU-129        [-1, 288, 56, 56]               0
  BatchNormAct2d-130        [-1, 288, 56, 56]             576
          Conv2d-131        [-1, 288, 56, 56]           2,592
        Identity-132        [-1, 288, 56, 56]               0
            SiLU-133        [-1, 288, 56, 56]               0
  BatchNormAct2d-134        [-1, 288, 56, 56]             576
        Identity-135        [-1, 288, 56, 56]               0
          Conv2d-136           [-1, 12, 1, 1]           3,468
            SiLU-137           [-1, 12, 1, 1]               0
          Conv2d-138          [-1, 288, 1, 1]           3,744
         Sigmoid-139          [-1, 288, 1, 1]               0
   SqueezeExcite-140        [-1, 288, 56, 56]               0
          Conv2d-141         [-1, 48, 56, 56]          13,824
        Identity-142         [-1, 48, 56, 56]               0
        Identity-143         [-1, 48, 56, 56]               0
  BatchNormAct2d-144         [-1, 48, 56, 56]              96
        Identity-145         [-1, 48, 56, 56]               0
InvertedResidual-146         [-1, 48, 56, 56]               0
          Conv2d-147        [-1, 288, 56, 56]          13,824
        Identity-148        [-1, 288, 56, 56]               0
            SiLU-149        [-1, 288, 56, 56]               0
  BatchNormAct2d-150        [-1, 288, 56, 56]             576
          Conv2d-151        [-1, 288, 56, 56]           2,592
        Identity-152        [-1, 288, 56, 56]               0
            SiLU-153        [-1, 288, 56, 56]               0
  BatchNormAct2d-154        [-1, 288, 56, 56]             576
        Identity-155        [-1, 288, 56, 56]               0
          Conv2d-156           [-1, 12, 1, 1]           3,468
            SiLU-157           [-1, 12, 1, 1]               0
          Conv2d-158          [-1, 288, 1, 1]           3,744
         Sigmoid-159          [-1, 288, 1, 1]               0
   SqueezeExcite-160        [-1, 288, 56, 56]               0
          Conv2d-161         [-1, 48, 56, 56]          13,824
        Identity-162         [-1, 48, 56, 56]               0
        Identity-163         [-1, 48, 56, 56]               0
  BatchNormAct2d-164         [-1, 48, 56, 56]              96
        Identity-165         [-1, 48, 56, 56]               0
```

```
InvertedResidual-166           [-1, 48, 56, 56]               0
          Conv2d-167          [-1, 288, 56, 56]          13,824
        Identity-168          [-1, 288, 56, 56]               0
            SiLU-169          [-1, 288, 56, 56]               0
  BatchNormAct2d-170          [-1, 288, 56, 56]             576
          Conv2d-171          [-1, 288, 56, 56]           2,592
        Identity-172          [-1, 288, 56, 56]               0
            SiLU-173          [-1, 288, 56, 56]               0
  BatchNormAct2d-174          [-1, 288, 56, 56]             576
        Identity-175          [-1, 288, 56, 56]               0
          Conv2d-176             [-1, 12, 1, 1]           3,468
            SiLU-177             [-1, 12, 1, 1]               0
          Conv2d-178            [-1, 288, 1, 1]           3,744
         Sigmoid-179            [-1, 288, 1, 1]               0
   SqueezeExcite-180          [-1, 288, 56, 56]               0
          Conv2d-181           [-1, 48, 56, 56]          13,824
        Identity-182           [-1, 48, 56, 56]               0
        Identity-183           [-1, 48, 56, 56]               0
  BatchNormAct2d-184           [-1, 48, 56, 56]              96
        Identity-185           [-1, 48, 56, 56]               0
InvertedResidual-186           [-1, 48, 56, 56]               0
          Conv2d-187          [-1, 288, 56, 56]          13,824
        Identity-188          [-1, 288, 56, 56]               0
            SiLU-189          [-1, 288, 56, 56]               0
  BatchNormAct2d-190          [-1, 288, 56, 56]             576
          Conv2d-191          [-1, 288, 56, 56]           2,592
        Identity-192          [-1, 288, 56, 56]               0
            SiLU-193          [-1, 288, 56, 56]               0
  BatchNormAct2d-194          [-1, 288, 56, 56]             576
        Identity-195          [-1, 288, 56, 56]               0
          Conv2d-196             [-1, 12, 1, 1]           3,468
            SiLU-197             [-1, 12, 1, 1]               0
          Conv2d-198            [-1, 288, 1, 1]           3,744
         Sigmoid-199            [-1, 288, 1, 1]               0
   SqueezeExcite-200          [-1, 288, 56, 56]               0
          Conv2d-201           [-1, 48, 56, 56]          13,824
        Identity-202           [-1, 48, 56, 56]               0
        Identity-203           [-1, 48, 56, 56]               0
  BatchNormAct2d-204           [-1, 48, 56, 56]              96
        Identity-205           [-1, 48, 56, 56]               0
InvertedResidual-206           [-1, 48, 56, 56]               0
          Conv2d-207          [-1, 288, 56, 56]          13,824
        Identity-208          [-1, 288, 56, 56]               0
            SiLU-209          [-1, 288, 56, 56]               0
  BatchNormAct2d-210          [-1, 288, 56, 56]             576
      Conv2dSame-211          [-1, 288, 28, 28]           7,200
        Identity-212          [-1, 288, 28, 28]               0
            SiLU-213          [-1, 288, 28, 28]               0
  BatchNormAct2d-214          [-1, 288, 28, 28]             576
        Identity-215          [-1, 288, 28, 28]               0
          Conv2d-216             [-1, 12, 1, 1]           3,468
            SiLU-217             [-1, 12, 1, 1]               0
          Conv2d-218            [-1, 288, 1, 1]           3,744
         Sigmoid-219            [-1, 288, 1, 1]               0
   SqueezeExcite-220          [-1, 288, 28, 28]               0
          Conv2d-221           [-1, 80, 28, 28]          23,040
        Identity-222           [-1, 80, 28, 28]               0
        Identity-223           [-1, 80, 28, 28]               0
  BatchNormAct2d-224           [-1, 80, 28, 28]             160
InvertedResidual-225           [-1, 80, 28, 28]               0
          Conv2d-226          [-1, 480, 28, 28]          38,400
        Identity-227          [-1, 480, 28, 28]               0
            SiLU-228          [-1, 480, 28, 28]               0
  BatchNormAct2d-229          [-1, 480, 28, 28]             960
          Conv2d-230          [-1, 480, 28, 28]          12,000
        Identity-231          [-1, 480, 28, 28]               0
            SiLU-232          [-1, 480, 28, 28]               0
  BatchNormAct2d-233          [-1, 480, 28, 28]             960
        Identity-234          [-1, 480, 28, 28]               0
          Conv2d-235             [-1, 20, 1, 1]           9,620
            SiLU-236             [-1, 20, 1, 1]               0
          Conv2d-237            [-1, 480, 1, 1]          10,080
         Sigmoid-238            [-1, 480, 1, 1]               0
   SqueezeExcite-239          [-1, 480, 28, 28]               0
          Conv2d-240           [-1, 80, 28, 28]          38,400
        Identity-241           [-1, 80, 28, 28]               0
        Identity-242           [-1, 80, 28, 28]               0
  BatchNormAct2d-243           [-1, 80, 28, 28]             160
        Identity-244           [-1, 80, 28, 28]               0
InvertedResidual-245           [-1, 80, 28, 28]               0
          Conv2d-246          [-1, 480, 28, 28]          38,400
        Identity-247          [-1, 480, 28, 28]               0
            SiLU-248          [-1, 480, 28, 28]               0
  BatchNormAct2d-249          [-1, 480, 28, 28]             960
          Conv2d-250          [-1, 480, 28, 28]          12,000
        Identity-251          [-1, 480, 28, 28]               0
            SiLU-252          [-1, 480, 28, 28]               0
  BatchNormAct2d-253          [-1, 480, 28, 28]             960
        Identity-254          [-1, 480, 28, 28]               0
          Conv2d-255             [-1, 20, 1, 1]           9,620
            SiLU-256             [-1, 20, 1, 1]               0
          Conv2d-257            [-1, 480, 1, 1]          10,080
         Sigmoid-258            [-1, 480, 1, 1]               0
   SqueezeExcite-259          [-1, 480, 28, 28]               0
          Conv2d-260           [-1, 80, 28, 28]          38,400
        Identity-261           [-1, 80, 28, 28]               0
        Identity-262          [-1, 80, 28, 28]               0
  BatchNormAct2d-263          [-1, 80, 28, 28]             160
        Identity-264          [-1, 80, 28, 28]               0
InvertedResidual-265          [-1, 80, 28, 28]               0
          Conv2d-266         [-1, 480, 28, 28]          38,400
        Identity-267         [-1, 480, 28, 28]               0
            SiLU-268         [-1, 480, 28, 28]               0
  BatchNormAct2d-269         [-1, 480, 28, 28]             960
          Conv2d-270         [-1, 480, 28, 28]          12,000
        Identity-271         [-1, 480, 28, 28]               0
            SiLU-272         [-1, 480, 28, 28]               0
  BatchNormAct2d-273         [-1, 480, 28, 28]             960
        Identity-274         [-1, 480, 28, 28]               0
          Conv2d-275            [-1, 20, 1, 1]           9,620
            SiLU-276            [-1, 20, 1, 1]               0
          Conv2d-277           [-1, 480, 1, 1]          10,080
         Sigmoid-278           [-1, 480, 1, 1]               0
   SqueezeExcite-279         [-1, 480, 28, 28]               0
          Conv2d-280          [-1, 80, 28, 28]          38,400
        Identity-281          [-1, 80, 28, 28]               0
        Identity-282          [-1, 80, 28, 28]               0
  BatchNormAct2d-283          [-1, 80, 28, 28]             160
        Identity-284          [-1, 80, 28, 28]               0
InvertedResidual-285          [-1, 80, 28, 28]               0
          Conv2d-286         [-1, 480, 28, 28]          38,400
        Identity-287         [-1, 480, 28, 28]               0
            SiLU-288         [-1, 480, 28, 28]               0
  BatchNormAct2d-289         [-1, 480, 28, 28]             960
          Conv2d-290         [-1, 480, 28, 28]          12,000
        Identity-291         [-1, 480, 28, 28]               0
            SiLU-292         [-1, 480, 28, 28]               0
  BatchNormAct2d-293         [-1, 480, 28, 28]             960
        Identity-294         [-1, 480, 28, 28]               0
          Conv2d-295            [-1, 20, 1, 1]           9,620
            SiLU-296            [-1, 20, 1, 1]               0
          Conv2d-297           [-1, 480, 1, 1]          10,080
         Sigmoid-298           [-1, 480, 1, 1]               0
   SqueezeExcite-299         [-1, 480, 28, 28]               0
          Conv2d-300          [-1, 80, 28, 28]          38,400
        Identity-301          [-1, 80, 28, 28]               0
        Identity-302          [-1, 80, 28, 28]               0
  BatchNormAct2d-303          [-1, 80, 28, 28]             160
        Identity-304          [-1, 80, 28, 28]               0
InvertedResidual-305          [-1, 80, 28, 28]               0
          Conv2d-306         [-1, 480, 28, 28]          38,400
        Identity-307         [-1, 480, 28, 28]               0
            SiLU-308         [-1, 480, 28, 28]               0
  BatchNormAct2d-309         [-1, 480, 28, 28]             960
```

```
          Conv2d-310          [-1, 480, 28, 28]          12,000
        Identity-311          [-1, 480, 28, 28]               0
            SiLU-312          [-1, 480, 28, 28]               0
  BatchNormAct2d-313          [-1, 480, 28, 28]             960
        Identity-314          [-1, 480, 28, 28]               0
          Conv2d-315             [-1, 20, 1, 1]           9,620
            SiLU-316             [-1, 20, 1, 1]               0
          Conv2d-317            [-1, 480, 1, 1]          10,080
         Sigmoid-318            [-1, 480, 1, 1]               0
   SqueezeExcite-319          [-1, 480, 28, 28]               0
          Conv2d-320           [-1, 80, 28, 28]          38,400
        Identity-321           [-1, 80, 28, 28]               0
        Identity-322           [-1, 80, 28, 28]               0
  BatchNormAct2d-323           [-1, 80, 28, 28]             160
        Identity-324           [-1, 80, 28, 28]               0
InvertedResidual-325           [-1, 80, 28, 28]               0
          Conv2d-326          [-1, 480, 28, 28]          38,400
        Identity-327          [-1, 480, 28, 28]               0
            SiLU-328          [-1, 480, 28, 28]               0
  BatchNormAct2d-329          [-1, 480, 28, 28]             960
          Conv2d-330          [-1, 480, 28, 28]          12,000
        Identity-331          [-1, 480, 28, 28]               0
            SiLU-332          [-1, 480, 28, 28]               0
  BatchNormAct2d-333          [-1, 480, 28, 28]             960
        Identity-334          [-1, 480, 28, 28]               0
          Conv2d-335             [-1, 20, 1, 1]           9,620
            SiLU-336             [-1, 20, 1, 1]               0
          Conv2d-337            [-1, 480, 1, 1]          10,080
         Sigmoid-338            [-1, 480, 1, 1]               0
   SqueezeExcite-339          [-1, 480, 28, 28]               0
          Conv2d-340           [-1, 80, 28, 28]          38,400
        Identity-341           [-1, 80, 28, 28]               0
        Identity-342           [-1, 80, 28, 28]               0
  BatchNormAct2d-343           [-1, 80, 28, 28]             160
        Identity-344           [-1, 80, 28, 28]               0
InvertedResidual-345           [-1, 80, 28, 28]               0
          Conv2d-346          [-1, 480, 28, 28]          38,400
        Identity-347          [-1, 480, 28, 28]               0
            SiLU-348          [-1, 480, 28, 28]               0
  BatchNormAct2d-349          [-1, 480, 28, 28]             960
      Conv2dSame-350          [-1, 480, 14, 14]           4,320
        Identity-351          [-1, 480, 14, 14]               0
            SiLU-352          [-1, 480, 14, 14]               0
  BatchNormAct2d-353          [-1, 480, 14, 14]             960
        Identity-354          [-1, 480, 14, 14]               0
          Conv2d-355             [-1, 20, 1, 1]           9,620
            SiLU-356             [-1, 20, 1, 1]               0
          Conv2d-357            [-1, 480, 1, 1]          10,080
         Sigmoid-358            [-1, 480, 1, 1]               0
   SqueezeExcite-359          [-1, 480, 14, 14]               0
          Conv2d-360          [-1, 160, 14, 14]          76,800
        Identity-361          [-1, 160, 14, 14]               0
        Identity-362          [-1, 160, 14, 14]               0
  BatchNormAct2d-363          [-1, 160, 14, 14]             320
InvertedResidual-364          [-1, 160, 14, 14]               0
          Conv2d-365          [-1, 960, 14, 14]         153,600
        Identity-366          [-1, 960, 14, 14]               0
            SiLU-367          [-1, 960, 14, 14]               0
  BatchNormAct2d-368          [-1, 960, 14, 14]           1,920
          Conv2d-369          [-1, 960, 14, 14]           8,640
        Identity-370          [-1, 960, 14, 14]               0
            SiLU-371          [-1, 960, 14, 14]               0
  BatchNormAct2d-372          [-1, 960, 14, 14]           1,920
        Identity-373          [-1, 960, 14, 14]               0
          Conv2d-374             [-1, 40, 1, 1]          38,440
            SiLU-375             [-1, 40, 1, 1]               0
          Conv2d-376            [-1, 960, 1, 1]          39,360
         Sigmoid-377            [-1, 960, 1, 1]               0
   SqueezeExcite-378          [-1, 960, 14, 14]               0
          Conv2d-379          [-1, 160, 14, 14]         153,600
        Identity-380          [-1, 160, 14, 14]               0
        Identity-381          [-1, 160, 14, 14]               0
  BatchNormAct2d-382          [-1, 160, 14, 14]             320
        Identity-383          [-1, 160, 14, 14]               0
InvertedResidual-384          [-1, 160, 14, 14]               0
          Conv2d-385          [-1, 960, 14, 14]         153,600
        Identity-386          [-1, 960, 14, 14]               0
            SiLU-387          [-1, 960, 14, 14]               0
  BatchNormAct2d-388          [-1, 960, 14, 14]           1,920
          Conv2d-389          [-1, 960, 14, 14]           8,640
        Identity-390          [-1, 960, 14, 14]               0
            SiLU-391          [-1, 960, 14, 14]               0
  BatchNormAct2d-392          [-1, 960, 14, 14]           1,920
        Identity-393          [-1, 960, 14, 14]               0
          Conv2d-394             [-1, 40, 1, 1]          38,440
            SiLU-395             [-1, 40, 1, 1]               0
          Conv2d-396            [-1, 960, 1, 1]          39,360
         Sigmoid-397            [-1, 960, 1, 1]               0
   SqueezeExcite-398          [-1, 960, 14, 14]               0
          Conv2d-399          [-1, 160, 14, 14]         153,600
        Identity-400          [-1, 160, 14, 14]               0
        Identity-401          [-1, 160, 14, 14]               0
  BatchNormAct2d-402          [-1, 160, 14, 14]             320
        Identity-403          [-1, 160, 14, 14]               0
InvertedResidual-404          [-1, 160, 14, 14]               0
          Conv2d-405          [-1, 960, 14, 14]         153,600
        Identity-406          [-1, 960, 14, 14]               0
            SiLU-407          [-1, 960, 14, 14]               0
  BatchNormAct2d-408          [-1, 960, 14, 14]           1,920
          Conv2d-409          [-1, 960, 14, 14]           8,640
        Identity-410          [-1, 960, 14, 14]               0
            SiLU-411          [-1, 960, 14, 14]               0
  BatchNormAct2d-412          [-1, 960, 14, 14]           1,920
        Identity-413          [-1, 960, 14, 14]               0
          Conv2d-414             [-1, 40, 1, 1]          38,440
            SiLU-415             [-1, 40, 1, 1]               0
          Conv2d-416            [-1, 960, 1, 1]          39,360
         Sigmoid-417            [-1, 960, 1, 1]               0
   SqueezeExcite-418          [-1, 960, 14, 14]               0
          Conv2d-419          [-1, 160, 14, 14]         153,600
        Identity-420          [-1, 160, 14, 14]               0
        Identity-421          [-1, 160, 14, 14]               0
  BatchNormAct2d-422          [-1, 160, 14, 14]             320
        Identity-423          [-1, 160, 14, 14]               0
InvertedResidual-424          [-1, 160, 14, 14]               0
          Conv2d-425          [-1, 960, 14, 14]         153,600
        Identity-426          [-1, 960, 14, 14]               0
            SiLU-427          [-1, 960, 14, 14]               0
  BatchNormAct2d-428          [-1, 960, 14, 14]           1,920
          Conv2d-429          [-1, 960, 14, 14]           8,640
        Identity-430          [-1, 960, 14, 14]               0
            SiLU-431          [-1, 960, 14, 14]               0
  BatchNormAct2d-432          [-1, 960, 14, 14]           1,920
        Identity-433          [-1, 960, 14, 14]               0
          Conv2d-434             [-1, 40, 1, 1]          38,440
            SiLU-435             [-1, 40, 1, 1]               0
          Conv2d-436            [-1, 960, 1, 1]          39,360
         Sigmoid-437            [-1, 960, 1, 1]               0
   SqueezeExcite-438          [-1, 960, 14, 14]               0
          Conv2d-439          [-1, 160, 14, 14]         153,600
        Identity-440          [-1, 160, 14, 14]               0
        Identity-441          [-1, 160, 14, 14]               0
  BatchNormAct2d-442          [-1, 160, 14, 14]             320
        Identity-443          [-1, 160, 14, 14]               0
InvertedResidual-444          [-1, 160, 14, 14]               0
          Conv2d-445          [-1, 960, 14, 14]         153,600
        Identity-446          [-1, 960, 14, 14]               0
            SiLU-447          [-1, 960, 14, 14]               0
  BatchNormAct2d-448          [-1, 960, 14, 14]           1,920
          Conv2d-449          [-1, 960, 14, 14]           8,640
        Identity-450          [-1, 960, 14, 14]               0
            SiLU-451          [-1, 960, 14, 14]               0
  BatchNormAct2d-452          [-1, 960, 14, 14]           1,920
        Identity-453          [-1, 960, 14, 14]               0
```

```
          Conv2d-454             [-1, 40, 1, 1]          38,440
            SiLU-455             [-1, 40, 1, 1]               0
          Conv2d-456            [-1, 960, 1, 1]          39,360
         Sigmoid-457            [-1, 960, 1, 1]               0
   SqueezeExcite-458          [-1, 960, 14, 14]               0
          Conv2d-459          [-1, 160, 14, 14]         153,600
        Identity-460          [-1, 160, 14, 14]               0
        Identity-461          [-1, 160, 14, 14]               0
  BatchNormAct2d-462          [-1, 160, 14, 14]             320
        Identity-463          [-1, 160, 14, 14]               0
InvertedResidual-464          [-1, 160, 14, 14]               0
          Conv2d-465          [-1, 960, 14, 14]         153,600
        Identity-466          [-1, 960, 14, 14]               0
            SiLU-467          [-1, 960, 14, 14]               0
  BatchNormAct2d-468          [-1, 960, 14, 14]           1,920
          Conv2d-469          [-1, 960, 14, 14]           8,640
        Identity-470          [-1, 960, 14, 14]               0
            SiLU-471          [-1, 960, 14, 14]               0
  BatchNormAct2d-472          [-1, 960, 14, 14]           1,920
        Identity-473          [-1, 960, 14, 14]               0
          Conv2d-474             [-1, 40, 1, 1]          38,440
            SiLU-475             [-1, 40, 1, 1]               0
          Conv2d-476            [-1, 960, 1, 1]          39,360
         Sigmoid-477            [-1, 960, 1, 1]               0
   SqueezeExcite-478          [-1, 960, 14, 14]               0
          Conv2d-479          [-1, 160, 14, 14]         153,600
        Identity-480          [-1, 160, 14, 14]               0
        Identity-481          [-1, 160, 14, 14]               0
  BatchNormAct2d-482          [-1, 160, 14, 14]             320
        Identity-483          [-1, 160, 14, 14]               0
InvertedResidual-484          [-1, 160, 14, 14]               0
          Conv2d-485          [-1, 960, 14, 14]         153,600
        Identity-486          [-1, 960, 14, 14]               0
            SiLU-487          [-1, 960, 14, 14]               0
  BatchNormAct2d-488          [-1, 960, 14, 14]           1,920
          Conv2d-489          [-1, 960, 14, 14]           8,640
        Identity-490          [-1, 960, 14, 14]               0
            SiLU-491          [-1, 960, 14, 14]               0
  BatchNormAct2d-492          [-1, 960, 14, 14]           1,920
        Identity-493          [-1, 960, 14, 14]               0
          Conv2d-494             [-1, 40, 1, 1]          38,440
            SiLU-495             [-1, 40, 1, 1]               0
          Conv2d-496            [-1, 960, 1, 1]          39,360
         Sigmoid-497            [-1, 960, 1, 1]               0
   SqueezeExcite-498          [-1, 960, 14, 14]               0
          Conv2d-499          [-1, 160, 14, 14]         153,600
        Identity-500          [-1, 160, 14, 14]               0
        Identity-501          [-1, 160, 14, 14]               0
  BatchNormAct2d-502          [-1, 160, 14, 14]             320
        Identity-503          [-1, 160, 14, 14]               0
InvertedResidual-504          [-1, 160, 14, 14]               0
          Conv2d-505          [-1, 960, 14, 14]         153,600
        Identity-506          [-1, 960, 14, 14]               0
            SiLU-507          [-1, 960, 14, 14]               0
  BatchNormAct2d-508          [-1, 960, 14, 14]           1,920
          Conv2d-509          [-1, 960, 14, 14]           8,640
        Identity-510          [-1, 960, 14, 14]               0
            SiLU-511          [-1, 960, 14, 14]               0
  BatchNormAct2d-512          [-1, 960, 14, 14]           1,920
        Identity-513          [-1, 960, 14, 14]               0
          Conv2d-514             [-1, 40, 1, 1]          38,440
            SiLU-515             [-1, 40, 1, 1]               0
          Conv2d-516            [-1, 960, 1, 1]          39,360
         Sigmoid-517            [-1, 960, 1, 1]               0
   SqueezeExcite-518          [-1, 960, 14, 14]               0
          Conv2d-519          [-1, 160, 14, 14]         153,600
        Identity-520          [-1, 160, 14, 14]               0
        Identity-521          [-1, 160, 14, 14]               0
  BatchNormAct2d-522          [-1, 160, 14, 14]             320
        Identity-523          [-1, 160, 14, 14]               0
InvertedResidual-524          [-1, 160, 14, 14]               0
          Conv2d-525          [-1, 960, 14, 14]         153,600
        Identity-526          [-1, 960, 14, 14]               0
            SiLU-527          [-1, 960, 14, 14]               0
  BatchNormAct2d-528          [-1, 960, 14, 14]           1,920
          Conv2d-529          [-1, 960, 14, 14]           8,640
        Identity-530          [-1, 960, 14, 14]               0
            SiLU-531          [-1, 960, 14, 14]               0
  BatchNormAct2d-532          [-1, 960, 14, 14]           1,920
        Identity-533          [-1, 960, 14, 14]               0
          Conv2d-534             [-1, 40, 1, 1]          38,440
            SiLU-535             [-1, 40, 1, 1]               0
          Conv2d-536            [-1, 960, 1, 1]          39,360
         Sigmoid-537            [-1, 960, 1, 1]               0
   SqueezeExcite-538          [-1, 960, 14, 14]               0
          Conv2d-539          [-1, 160, 14, 14]         153,600
        Identity-540          [-1, 160, 14, 14]               0
        Identity-541          [-1, 160, 14, 14]               0
  BatchNormAct2d-542          [-1, 160, 14, 14]             320
        Identity-543          [-1, 160, 14, 14]               0
InvertedResidual-544          [-1, 160, 14, 14]               0
          Conv2d-545          [-1, 960, 14, 14]         153,600
        Identity-546          [-1, 960, 14, 14]               0
            SiLU-547          [-1, 960, 14, 14]               0
  BatchNormAct2d-548          [-1, 960, 14, 14]           1,920
          Conv2d-549          [-1, 960, 14, 14]          24,000
        Identity-550          [-1, 960, 14, 14]               0
            SiLU-551          [-1, 960, 14, 14]               0
  BatchNormAct2d-552          [-1, 960, 14, 14]           1,920
        Identity-553          [-1, 960, 14, 14]               0
          Conv2d-554             [-1, 40, 1, 1]          38,440
            SiLU-555             [-1, 40, 1, 1]               0
          Conv2d-556            [-1, 960, 1, 1]          39,360
         Sigmoid-557            [-1, 960, 1, 1]               0
   SqueezeExcite-558          [-1, 960, 14, 14]               0
          Conv2d-559          [-1, 224, 14, 14]         215,040
        Identity-560          [-1, 224, 14, 14]               0
        Identity-561          [-1, 224, 14, 14]               0
  BatchNormAct2d-562          [-1, 224, 14, 14]             448
InvertedResidual-563          [-1, 224, 14, 14]               0
          Conv2d-564         [-1, 1344, 14, 14]         301,056
        Identity-565         [-1, 1344, 14, 14]               0
            SiLU-566         [-1, 1344, 14, 14]               0
  BatchNormAct2d-567         [-1, 1344, 14, 14]           2,688
          Conv2d-568         [-1, 1344, 14, 14]          33,600
        Identity-569         [-1, 1344, 14, 14]               0
            SiLU-570         [-1, 1344, 14, 14]               0
  BatchNormAct2d-571         [-1, 1344, 14, 14]           2,688
        Identity-572         [-1, 1344, 14, 14]               0
          Conv2d-573             [-1, 56, 1, 1]          75,320
            SiLU-574             [-1, 56, 1, 1]               0
          Conv2d-575           [-1, 1344, 1, 1]          76,608
         Sigmoid-576           [-1, 1344, 1, 1]               0
   SqueezeExcite-577         [-1, 1344, 14, 14]               0
          Conv2d-578          [-1, 224, 14, 14]         301,056
        Identity-579          [-1, 224, 14, 14]               0
        Identity-580          [-1, 224, 14, 14]               0
  BatchNormAct2d-581          [-1, 224, 14, 14]             448
        Identity-582          [-1, 224, 14, 14]               0
InvertedResidual-583          [-1, 224, 14, 14]               0
          Conv2d-584         [-1, 1344, 14, 14]         301,056
        Identity-585         [-1, 1344, 14, 14]               0
            SiLU-586         [-1, 1344, 14, 14]               0
  BatchNormAct2d-587         [-1, 1344, 14, 14]           2,688
          Conv2d-588         [-1, 1344, 14, 14]          33,600
        Identity-589         [-1, 1344, 14, 14]               0
            SiLU-590         [-1, 1344, 14, 14]               0
  BatchNormAct2d-591         [-1, 1344, 14, 14]           2,688
        Identity-592         [-1, 1344, 14, 14]               0
          Conv2d-593             [-1, 56, 1, 1]          75,320
            SiLU-594             [-1, 56, 1, 1]               0
          Conv2d-595           [-1, 1344, 1, 1]          76,608
         Sigmoid-596           [-1, 1344, 1, 1]               0
   SqueezeExcite-597         [-1, 1344, 14, 14]               0
```

```
          Conv2d-598         [-1, 224, 14, 14]         301,056
        Identity-599         [-1, 224, 14, 14]               0
        Identity-600         [-1, 224, 14, 14]               0
  BatchNormAct2d-601         [-1, 224, 14, 14]             448
        Identity-602         [-1, 224, 14, 14]               0
InvertedResidual-603         [-1, 224, 14, 14]               0
          Conv2d-604        [-1, 1344, 14, 14]         301,056
        Identity-605        [-1, 1344, 14, 14]               0
            SiLU-606        [-1, 1344, 14, 14]               0
  BatchNormAct2d-607        [-1, 1344, 14, 14]           2,688
          Conv2d-608        [-1, 1344, 14, 14]          33,600
        Identity-609        [-1, 1344, 14, 14]               0
            SiLU-610        [-1, 1344, 14, 14]               0
  BatchNormAct2d-611        [-1, 1344, 14, 14]           2,688
        Identity-612        [-1, 1344, 14, 14]               0
          Conv2d-613            [-1, 56, 1, 1]          75,320
            SiLU-614            [-1, 56, 1, 1]               0
          Conv2d-615          [-1, 1344, 1, 1]          76,608
         Sigmoid-616          [-1, 1344, 1, 1]               0
   SqueezeExcite-617        [-1, 1344, 14, 14]               0
          Conv2d-618         [-1, 224, 14, 14]         301,056
        Identity-619         [-1, 224, 14, 14]               0
        Identity-620         [-1, 224, 14, 14]               0
  BatchNormAct2d-621         [-1, 224, 14, 14]             448
        Identity-622         [-1, 224, 14, 14]               0
InvertedResidual-623         [-1, 224, 14, 14]               0
          Conv2d-624        [-1, 1344, 14, 14]         301,056
        Identity-625        [-1, 1344, 14, 14]               0
            SiLU-626        [-1, 1344, 14, 14]               0
  BatchNormAct2d-627        [-1, 1344, 14, 14]           2,688
          Conv2d-628        [-1, 1344, 14, 14]          33,600
        Identity-629        [-1, 1344, 14, 14]               0
            SiLU-630        [-1, 1344, 14, 14]               0
  BatchNormAct2d-631        [-1, 1344, 14, 14]           2,688
        Identity-632        [-1, 1344, 14, 14]               0
          Conv2d-633            [-1, 56, 1, 1]          75,320
            SiLU-634            [-1, 56, 1, 1]               0
          Conv2d-635          [-1, 1344, 1, 1]          76,608
         Sigmoid-636          [-1, 1344, 1, 1]               0
   SqueezeExcite-637        [-1, 1344, 14, 14]               0
          Conv2d-638         [-1, 224, 14, 14]         301,056
        Identity-639         [-1, 224, 14, 14]               0
        Identity-640         [-1, 224, 14, 14]               0
  BatchNormAct2d-641         [-1, 224, 14, 14]             448
        Identity-642         [-1, 224, 14, 14]               0
InvertedResidual-643         [-1, 224, 14, 14]               0
          Conv2d-644        [-1, 1344, 14, 14]         301,056
        Identity-645        [-1, 1344, 14, 14]               0
```

```
            SiLU-646         [-1, 1344, 14, 14]               0
  BatchNormAct2d-647         [-1, 1344, 14, 14]           2,688
          Conv2d-648         [-1, 1344, 14, 14]          33,600
        Identity-649         [-1, 1344, 14, 14]               0
            SiLU-650         [-1, 1344, 14, 14]               0
  BatchNormAct2d-651         [-1, 1344, 14, 14]           2,688
        Identity-652         [-1, 1344, 14, 14]               0
          Conv2d-653             [-1, 56, 1, 1]          75,320
            SiLU-654             [-1, 56, 1, 1]               0
          Conv2d-655           [-1, 1344, 1, 1]          76,608
         Sigmoid-656           [-1, 1344, 1, 1]               0
   SqueezeExcite-657         [-1, 1344, 14, 14]               0
          Conv2d-658          [-1, 224, 14, 14]         301,056
        Identity-659          [-1, 224, 14, 14]               0
        Identity-660          [-1, 224, 14, 14]               0
  BatchNormAct2d-661          [-1, 224, 14, 14]             448
        Identity-662          [-1, 224, 14, 14]               0
InvertedResidual-663          [-1, 224, 14, 14]               0
          Conv2d-664         [-1, 1344, 14, 14]         301,056
        Identity-665         [-1, 1344, 14, 14]               0
            SiLU-666         [-1, 1344, 14, 14]               0
  BatchNormAct2d-667         [-1, 1344, 14, 14]           2,688
          Conv2d-668         [-1, 1344, 14, 14]          33,600
        Identity-669         [-1, 1344, 14, 14]               0
            SiLU-670         [-1, 1344, 14, 14]               0
  BatchNormAct2d-671         [-1, 1344, 14, 14]           2,688
        Identity-672         [-1, 1344, 14, 14]               0
          Conv2d-673             [-1, 56, 1, 1]          75,320
            SiLU-674             [-1, 56, 1, 1]               0
          Conv2d-675           [-1, 1344, 1, 1]          76,608
         Sigmoid-676           [-1, 1344, 1, 1]               0
   SqueezeExcite-677         [-1, 1344, 14, 14]               0
          Conv2d-678          [-1, 224, 14, 14]         301,056
        Identity-679          [-1, 224, 14, 14]               0
        Identity-680          [-1, 224, 14, 14]               0
  BatchNormAct2d-681          [-1, 224, 14, 14]             448
        Identity-682          [-1, 224, 14, 14]               0
InvertedResidual-683          [-1, 224, 14, 14]               0
          Conv2d-684         [-1, 1344, 14, 14]         301,056
        Identity-685         [-1, 1344, 14, 14]               0
            SiLU-686         [-1, 1344, 14, 14]               0
  BatchNormAct2d-687         [-1, 1344, 14, 14]           2,688
          Conv2d-688         [-1, 1344, 14, 14]          33,600
        Identity-689         [-1, 1344, 14, 14]               0
            SiLU-690         [-1, 1344, 14, 14]               0
  BatchNormAct2d-691         [-1, 1344, 14, 14]           2,688
        Identity-692         [-1, 1344, 14, 14]               0
          Conv2d-693             [-1, 56, 1, 1]          75,320
```

```
            SiLU-694             [-1, 56, 1, 1]               0
          Conv2d-695           [-1, 1344, 1, 1]          76,608
         Sigmoid-696           [-1, 1344, 1, 1]               0
   SqueezeExcite-697         [-1, 1344, 14, 14]               0
          Conv2d-698          [-1, 224, 14, 14]         301,056
        Identity-699          [-1, 224, 14, 14]               0
        Identity-700          [-1, 224, 14, 14]               0
  BatchNormAct2d-701          [-1, 224, 14, 14]             448
        Identity-702          [-1, 224, 14, 14]               0
InvertedResidual-703          [-1, 224, 14, 14]               0
          Conv2d-704         [-1, 1344, 14, 14]         301,056
        Identity-705         [-1, 1344, 14, 14]               0
            SiLU-706         [-1, 1344, 14, 14]               0
  BatchNormAct2d-707         [-1, 1344, 14, 14]           2,688
          Conv2d-708         [-1, 1344, 14, 14]          33,600
        Identity-709         [-1, 1344, 14, 14]               0
            SiLU-710         [-1, 1344, 14, 14]               0
  BatchNormAct2d-711         [-1, 1344, 14, 14]           2,688
        Identity-712         [-1, 1344, 14, 14]               0
          Conv2d-713             [-1, 56, 1, 1]          75,320
            SiLU-714             [-1, 56, 1, 1]               0
          Conv2d-715           [-1, 1344, 1, 1]          76,608
         Sigmoid-716           [-1, 1344, 1, 1]               0
   SqueezeExcite-717         [-1, 1344, 14, 14]               0
          Conv2d-718          [-1, 224, 14, 14]         301,056
        Identity-719          [-1, 224, 14, 14]               0
        Identity-720          [-1, 224, 14, 14]               0
  BatchNormAct2d-721          [-1, 224, 14, 14]             448
        Identity-722          [-1, 224, 14, 14]               0
InvertedResidual-723          [-1, 224, 14, 14]               0
          Conv2d-724         [-1, 1344, 14, 14]         301,056
        Identity-725         [-1, 1344, 14, 14]               0
            SiLU-726         [-1, 1344, 14, 14]               0
  BatchNormAct2d-727         [-1, 1344, 14, 14]           2,688
          Conv2d-728         [-1, 1344, 14, 14]          33,600
        Identity-729         [-1, 1344, 14, 14]               0
            SiLU-730         [-1, 1344, 14, 14]               0
  BatchNormAct2d-731         [-1, 1344, 14, 14]           2,688
        Identity-732         [-1, 1344, 14, 14]               0
          Conv2d-733             [-1, 56, 1, 1]          75,320
            SiLU-734             [-1, 56, 1, 1]               0
          Conv2d-735           [-1, 1344, 1, 1]          76,608
         Sigmoid-736           [-1, 1344, 1, 1]               0
   SqueezeExcite-737         [-1, 1344, 14, 14]               0
          Conv2d-738          [-1, 224, 14, 14]         301,056
        Identity-739          [-1, 224, 14, 14]               0
        Identity-740          [-1, 224, 14, 14]               0
  BatchNormAct2d-741          [-1, 224, 14, 14]             448
```

```
        Identity-742         [-1, 224, 14, 14]               0
InvertedResidual-743         [-1, 224, 14, 14]               0
          Conv2d-744        [-1, 1344, 14, 14]         301,056
        Identity-745        [-1, 1344, 14, 14]               0
            SiLU-746        [-1, 1344, 14, 14]               0
  BatchNormAct2d-747        [-1, 1344, 14, 14]           2,688
      Conv2dSame-748          [-1, 1344, 7, 7]          33,600
        Identity-749          [-1, 1344, 7, 7]               0
            SiLU-750          [-1, 1344, 7, 7]               0
  BatchNormAct2d-751          [-1, 1344, 7, 7]           2,688
        Identity-752          [-1, 1344, 7, 7]               0
          Conv2d-753            [-1, 56, 1, 1]          75,320
            SiLU-754            [-1, 56, 1, 1]               0
          Conv2d-755          [-1, 1344, 1, 1]          76,608
         Sigmoid-756          [-1, 1344, 1, 1]               0
   SqueezeExcite-757          [-1, 1344, 7, 7]               0
          Conv2d-758           [-1, 384, 7, 7]         516,096
        Identity-759           [-1, 384, 7, 7]               0
        Identity-760           [-1, 384, 7, 7]               0
  BatchNormAct2d-761           [-1, 384, 7, 7]             768
InvertedResidual-762           [-1, 384, 7, 7]               0
          Conv2d-763          [-1, 2304, 7, 7]         884,736
        Identity-764          [-1, 2304, 7, 7]               0
            SiLU-765          [-1, 2304, 7, 7]               0
  BatchNormAct2d-766          [-1, 2304, 7, 7]           4,608
          Conv2d-767          [-1, 2304, 7, 7]          57,600
        Identity-768          [-1, 2304, 7, 7]               0
            SiLU-769          [-1, 2304, 7, 7]               0
  BatchNormAct2d-770          [-1, 2304, 7, 7]           4,608
        Identity-771          [-1, 2304, 7, 7]               0
          Conv2d-772            [-1, 96, 1, 1]         221,280
            SiLU-773            [-1, 96, 1, 1]               0
          Conv2d-774          [-1, 2304, 1, 1]         223,488
         Sigmoid-775          [-1, 2304, 1, 1]               0
   SqueezeExcite-776          [-1, 2304, 7, 7]               0
          Conv2d-777           [-1, 384, 7, 7]         884,736
        Identity-778           [-1, 384, 7, 7]               0
        Identity-779           [-1, 384, 7, 7]               0
  BatchNormAct2d-780           [-1, 384, 7, 7]             768
        Identity-781           [-1, 384, 7, 7]               0
InvertedResidual-782           [-1, 384, 7, 7]               0
          Conv2d-783          [-1, 2304, 7, 7]         884,736
        Identity-784          [-1, 2304, 7, 7]               0
            SiLU-785          [-1, 2304, 7, 7]               0
  BatchNormAct2d-786          [-1, 2304, 7, 7]           4,608
          Conv2d-787          [-1, 2304, 7, 7]          57,600
        Identity-788          [-1, 2304, 7, 7]               0
            SiLU-789          [-1, 2304, 7, 7]               0
  BatchNormAct2d-790           [-1, 2304, 7, 7]           4,608
        Identity-791           [-1, 2304, 7, 7]               0
          Conv2d-792             [-1, 96, 1, 1]         221,280
            SiLU-793             [-1, 96, 1, 1]               0
          Conv2d-794           [-1, 2304, 1, 1]         223,488
         Sigmoid-795           [-1, 2304, 1, 1]               0
   SqueezeExcite-796           [-1, 2304, 7, 7]               0
          Conv2d-797            [-1, 384, 7, 7]         884,736
        Identity-798            [-1, 384, 7, 7]               0
        Identity-799            [-1, 384, 7, 7]               0
  BatchNormAct2d-800            [-1, 384, 7, 7]             768
        Identity-801            [-1, 384, 7, 7]               0
InvertedResidual-802            [-1, 384, 7, 7]               0
          Conv2d-803           [-1, 2304, 7, 7]         884,736
        Identity-804           [-1, 2304, 7, 7]               0
            SiLU-805           [-1, 2304, 7, 7]               0
  BatchNormAct2d-806           [-1, 2304, 7, 7]           4,608
          Conv2d-807           [-1, 2304, 7, 7]          57,600
        Identity-808           [-1, 2304, 7, 7]               0
            SiLU-809           [-1, 2304, 7, 7]               0
  BatchNormAct2d-810           [-1, 2304, 7, 7]           4,608
        Identity-811           [-1, 2304, 7, 7]               0
          Conv2d-812             [-1, 96, 1, 1]         221,280
            SiLU-813             [-1, 96, 1, 1]               0
          Conv2d-814           [-1, 2304, 1, 1]         223,488
         Sigmoid-815           [-1, 2304, 1, 1]               0
   SqueezeExcite-816           [-1, 2304, 7, 7]               0
          Conv2d-817            [-1, 384, 7, 7]         884,736
        Identity-818            [-1, 384, 7, 7]               0
        Identity-819            [-1, 384, 7, 7]               0
  BatchNormAct2d-820            [-1, 384, 7, 7]             768
        Identity-821            [-1, 384, 7, 7]               0
InvertedResidual-822            [-1, 384, 7, 7]               0
          Conv2d-823           [-1, 2304, 7, 7]         884,736
        Identity-824           [-1, 2304, 7, 7]               0
            SiLU-825           [-1, 2304, 7, 7]               0
  BatchNormAct2d-826           [-1, 2304, 7, 7]           4,608
          Conv2d-827           [-1, 2304, 7, 7]          57,600
        Identity-828           [-1, 2304, 7, 7]               0
            SiLU-829           [-1, 2304, 7, 7]               0
  BatchNormAct2d-830           [-1, 2304, 7, 7]           4,608
        Identity-831           [-1, 2304, 7, 7]               0
          Conv2d-832             [-1, 96, 1, 1]         221,280
            SiLU-833             [-1, 96, 1, 1]               0
          Conv2d-834           [-1, 2304, 1, 1]         223,488
         Sigmoid-835           [-1, 2304, 1, 1]               0
   SqueezeExcite-836           [-1, 2304, 7, 7]               0
          Conv2d-837            [-1, 384, 7, 7]         884,736
        Identity-838            [-1, 384, 7, 7]               0
        Identity-839            [-1, 384, 7, 7]               0
  BatchNormAct2d-840            [-1, 384, 7, 7]             768
        Identity-841            [-1, 384, 7, 7]               0
InvertedResidual-842            [-1, 384, 7, 7]               0
          Conv2d-843           [-1, 2304, 7, 7]         884,736
        Identity-844           [-1, 2304, 7, 7]               0
            SiLU-845           [-1, 2304, 7, 7]               0
  BatchNormAct2d-846           [-1, 2304, 7, 7]           4,608
          Conv2d-847           [-1, 2304, 7, 7]          57,600
        Identity-848           [-1, 2304, 7, 7]               0
            SiLU-849           [-1, 2304, 7, 7]               0
  BatchNormAct2d-850           [-1, 2304, 7, 7]           4,608
        Identity-851           [-1, 2304, 7, 7]               0
          Conv2d-852             [-1, 96, 1, 1]         221,280
            SiLU-853             [-1, 96, 1, 1]               0
          Conv2d-854           [-1, 2304, 1, 1]         223,488
         Sigmoid-855           [-1, 2304, 1, 1]               0
   SqueezeExcite-856           [-1, 2304, 7, 7]               0
          Conv2d-857            [-1, 384, 7, 7]         884,736
        Identity-858            [-1, 384, 7, 7]               0
        Identity-859            [-1, 384, 7, 7]               0
  BatchNormAct2d-860            [-1, 384, 7, 7]             768
        Identity-861            [-1, 384, 7, 7]               0
InvertedResidual-862            [-1, 384, 7, 7]               0
          Conv2d-863           [-1, 2304, 7, 7]         884,736
        Identity-864           [-1, 2304, 7, 7]               0
            SiLU-865           [-1, 2304, 7, 7]               0
  BatchNormAct2d-866           [-1, 2304, 7, 7]           4,608
          Conv2d-867           [-1, 2304, 7, 7]          57,600
        Identity-868           [-1, 2304, 7, 7]               0
            SiLU-869           [-1, 2304, 7, 7]               0
  BatchNormAct2d-870           [-1, 2304, 7, 7]           4,608
        Identity-871           [-1, 2304, 7, 7]               0
          Conv2d-872             [-1, 96, 1, 1]         221,280
            SiLU-873             [-1, 96, 1, 1]               0
          Conv2d-874           [-1, 2304, 1, 1]         223,488
         Sigmoid-875           [-1, 2304, 1, 1]               0
   SqueezeExcite-876           [-1, 2304, 7, 7]               0
          Conv2d-877            [-1, 384, 7, 7]         884,736
        Identity-878            [-1, 384, 7, 7]               0
        Identity-879            [-1, 384, 7, 7]               0
  BatchNormAct2d-880            [-1, 384, 7, 7]             768
        Identity-881            [-1, 384, 7, 7]               0
InvertedResidual-882            [-1, 384, 7, 7]               0
          Conv2d-883           [-1, 2304, 7, 7]         884,736
        Identity-884           [-1, 2304, 7, 7]               0
            SiLU-885           [-1, 2304, 7, 7]               0
  BatchNormAct2d-886           [-1, 2304, 7, 7]           4,608
          Conv2d-887           [-1, 2304, 7, 7]          57,600
        Identity-888           [-1, 2304, 7, 7]               0
            SiLU-889           [-1, 2304, 7, 7]               0
  BatchNormAct2d-890           [-1, 2304, 7, 7]           4,608
        Identity-891           [-1, 2304, 7, 7]               0
          Conv2d-892             [-1, 96, 1, 1]         221,280
            SiLU-893             [-1, 96, 1, 1]               0
          Conv2d-894           [-1, 2304, 1, 1]         223,488
         Sigmoid-895           [-1, 2304, 1, 1]               0
   SqueezeExcite-896           [-1, 2304, 7, 7]               0
          Conv2d-897            [-1, 384, 7, 7]         884,736
        Identity-898            [-1, 384, 7, 7]               0
        Identity-899            [-1, 384, 7, 7]               0
  BatchNormAct2d-900            [-1, 384, 7, 7]             768
        Identity-901            [-1, 384, 7, 7]               0
InvertedResidual-902            [-1, 384, 7, 7]               0
          Conv2d-903           [-1, 2304, 7, 7]         884,736
        Identity-904           [-1, 2304, 7, 7]               0
            SiLU-905           [-1, 2304, 7, 7]               0
  BatchNormAct2d-906           [-1, 2304, 7, 7]           4,608
          Conv2d-907           [-1, 2304, 7, 7]          57,600
        Identity-908           [-1, 2304, 7, 7]               0
            SiLU-909           [-1, 2304, 7, 7]               0
  BatchNormAct2d-910           [-1, 2304, 7, 7]           4,608
        Identity-911           [-1, 2304, 7, 7]               0
          Conv2d-912             [-1, 96, 1, 1]         221,280
            SiLU-913             [-1, 96, 1, 1]               0
          Conv2d-914           [-1, 2304, 1, 1]         223,488
         Sigmoid-915           [-1, 2304, 1, 1]               0
   SqueezeExcite-916           [-1, 2304, 7, 7]               0
          Conv2d-917            [-1, 384, 7, 7]         884,736
        Identity-918            [-1, 384, 7, 7]               0
        Identity-919            [-1, 384, 7, 7]               0
  BatchNormAct2d-920            [-1, 384, 7, 7]             768
        Identity-921            [-1, 384, 7, 7]               0
InvertedResidual-922            [-1, 384, 7, 7]               0
          Conv2d-923           [-1, 2304, 7, 7]         884,736
        Identity-924           [-1, 2304, 7, 7]               0
            SiLU-925           [-1, 2304, 7, 7]               0
  BatchNormAct2d-926           [-1, 2304, 7, 7]           4,608
          Conv2d-927           [-1, 2304, 7, 7]          57,600
        Identity-928           [-1, 2304, 7, 7]               0
            SiLU-929           [-1, 2304, 7, 7]               0
  BatchNormAct2d-930           [-1, 2304, 7, 7]           4,608
        Identity-931           [-1, 2304, 7, 7]               0
          Conv2d-932             [-1, 96, 1, 1]         221,280
            SiLU-933             [-1, 96, 1, 1]               0
          Conv2d-934           [-1, 2304, 1, 1]         223,488
         Sigmoid-935           [-1, 2304, 1, 1]               0
   SqueezeExcite-936           [-1, 2304, 7, 7]               0
          Conv2d-937            [-1, 384, 7, 7]         884,736
        Identity-938            [-1, 384, 7, 7]               0
        Identity-939            [-1, 384, 7, 7]               0
  BatchNormAct2d-940            [-1, 384, 7, 7]             768
        Identity-941            [-1, 384, 7, 7]               0
InvertedResidual-942            [-1, 384, 7, 7]               0
          Conv2d-943           [-1, 2304, 7, 7]         884,736
        Identity-944           [-1, 2304, 7, 7]               0
            SiLU-945           [-1, 2304, 7, 7]               0
  BatchNormAct2d-946           [-1, 2304, 7, 7]           4,608
          Conv2d-947           [-1, 2304, 7, 7]          57,600
        Identity-948           [-1, 2304, 7, 7]               0
            SiLU-949           [-1, 2304, 7, 7]               0
  BatchNormAct2d-950           [-1, 2304, 7, 7]           4,608
        Identity-951           [-1, 2304, 7, 7]               0
          Conv2d-952             [-1, 96, 1, 1]         221,280
            SiLU-953             [-1, 96, 1, 1]               0
          Conv2d-954           [-1, 2304, 1, 1]         223,488
         Sigmoid-955           [-1, 2304, 1, 1]               0
   SqueezeExcite-956           [-1, 2304, 7, 7]               0
          Conv2d-957            [-1, 384, 7, 7]         884,736
        Identity-958            [-1, 384, 7, 7]               0
        Identity-959            [-1, 384, 7, 7]               0
  BatchNormAct2d-960            [-1, 384, 7, 7]             768
        Identity-961            [-1, 384, 7, 7]               0
InvertedResidual-962            [-1, 384, 7, 7]               0
          Conv2d-963           [-1, 2304, 7, 7]         884,736
        Identity-964           [-1, 2304, 7, 7]               0
            SiLU-965           [-1, 2304, 7, 7]               0
  BatchNormAct2d-966           [-1, 2304, 7, 7]           4,608
          Conv2d-967           [-1, 2304, 7, 7]          57,600
        Identity-968           [-1, 2304, 7, 7]               0
            SiLU-969           [-1, 2304, 7, 7]               0
  BatchNormAct2d-970           [-1, 2304, 7, 7]           4,608
        Identity-971           [-1, 2304, 7, 7]               0
          Conv2d-972             [-1, 96, 1, 1]         221,280
            SiLU-973             [-1, 96, 1, 1]               0
          Conv2d-974           [-1, 2304, 1, 1]         223,488
         Sigmoid-975           [-1, 2304, 1, 1]               0
   SqueezeExcite-976           [-1, 2304, 7, 7]               0
          Conv2d-977            [-1, 384, 7, 7]         884,736
        Identity-978            [-1, 384, 7, 7]               0
        Identity-979            [-1, 384, 7, 7]               0
  BatchNormAct2d-980            [-1, 384, 7, 7]             768
        Identity-981            [-1, 384, 7, 7]               0
```

```
InvertedResidual-982            [-1, 384, 7, 7]               0
          Conv2d-983           [-1, 2304, 7, 7]         884,736
        Identity-984           [-1, 2304, 7, 7]               0
            SiLU-985           [-1, 2304, 7, 7]               0
  BatchNormAct2d-986           [-1, 2304, 7, 7]           4,608
          Conv2d-987           [-1, 2304, 7, 7]          57,600
        Identity-988           [-1, 2304, 7, 7]               0
            SiLU-989           [-1, 2304, 7, 7]               0
  BatchNormAct2d-990           [-1, 2304, 7, 7]           4,608
        Identity-991           [-1, 2304, 7, 7]               0
          Conv2d-992             [-1, 96, 1, 1]         221,280
            SiLU-993             [-1, 96, 1, 1]               0
          Conv2d-994           [-1, 2304, 1, 1]         223,488
         Sigmoid-995           [-1, 2304, 1, 1]               0
   SqueezeExcite-996           [-1, 2304, 7, 7]               0
          Conv2d-997            [-1, 384, 7, 7]         884,736
        Identity-998            [-1, 384, 7, 7]               0
        Identity-999            [-1, 384, 7, 7]               0
 BatchNormAct2d-1000            [-1, 384, 7, 7]             768
       Identity-1001            [-1, 384, 7, 7]               0
InvertedResidual-1002            [-1, 384, 7, 7]               0
         Conv2d-1003           [-1, 2304, 7, 7]         884,736
       Identity-1004           [-1, 2304, 7, 7]               0
           SiLU-1005           [-1, 2304, 7, 7]               0
 BatchNormAct2d-1006           [-1, 2304, 7, 7]           4,608
         Conv2d-1007           [-1, 2304, 7, 7]          20,736
       Identity-1008           [-1, 2304, 7, 7]               0
           SiLU-1009           [-1, 2304, 7, 7]               0
 BatchNormAct2d-1010           [-1, 2304, 7, 7]           4,608
       Identity-1011           [-1, 2304, 7, 7]               0
         Conv2d-1012             [-1, 96, 1, 1]         221,280
           SiLU-1013             [-1, 96, 1, 1]               0
         Conv2d-1014           [-1, 2304, 1, 1]         223,488
        Sigmoid-1015           [-1, 2304, 1, 1]               0
  SqueezeExcite-1016           [-1, 2304, 7, 7]               0
         Conv2d-1017            [-1, 640, 7, 7]       1,474,560
       Identity-1018            [-1, 640, 7, 7]               0
       Identity-1019            [-1, 640, 7, 7]               0
 BatchNormAct2d-1020            [-1, 640, 7, 7]           1,280
InvertedResidual-1021            [-1, 640, 7, 7]               0
         Conv2d-1022           [-1, 3840, 7, 7]       2,457,600
       Identity-1023           [-1, 3840, 7, 7]               0
           SiLU-1024           [-1, 3840, 7, 7]               0
 BatchNormAct2d-1025           [-1, 3840, 7, 7]           7,680
         Conv2d-1026           [-1, 3840, 7, 7]          34,560
       Identity-1027           [-1, 3840, 7, 7]               0
           SiLU-1028           [-1, 3840, 7, 7]               0
 BatchNormAct2d-1029           [-1, 3840, 7, 7]           7,680
```

```
       Identity-1030          [-1, 3840, 7, 7]               0
         Conv2d-1031           [-1, 160, 1, 1]         614,560
           SiLU-1032           [-1, 160, 1, 1]               0
         Conv2d-1033          [-1, 3840, 1, 1]         618,240
        Sigmoid-1034          [-1, 3840, 1, 1]               0
  SqueezeExcite-1035          [-1, 3840, 7, 7]               0
         Conv2d-1036           [-1, 640, 7, 7]       2,457,600
       Identity-1037           [-1, 640, 7, 7]               0
       Identity-1038           [-1, 640, 7, 7]               0
 BatchNormAct2d-1039           [-1, 640, 7, 7]           1,280
       Identity-1040           [-1, 640, 7, 7]               0
InvertedResidual-1041           [-1, 640, 7, 7]               0
         Conv2d-1042          [-1, 3840, 7, 7]       2,457,600
       Identity-1043          [-1, 3840, 7, 7]               0
           SiLU-1044          [-1, 3840, 7, 7]               0
 BatchNormAct2d-1045          [-1, 3840, 7, 7]           7,680
         Conv2d-1046          [-1, 3840, 7, 7]          34,560
       Identity-1047          [-1, 3840, 7, 7]               0
           SiLU-1048          [-1, 3840, 7, 7]               0
 BatchNormAct2d-1049          [-1, 3840, 7, 7]           7,680
       Identity-1050          [-1, 3840, 7, 7]               0
         Conv2d-1051           [-1, 160, 1, 1]         614,560
           SiLU-1052           [-1, 160, 1, 1]               0
         Conv2d-1053          [-1, 3840, 1, 1]         618,240
        Sigmoid-1054          [-1, 3840, 1, 1]               0
  SqueezeExcite-1055          [-1, 3840, 7, 7]               0
         Conv2d-1056           [-1, 640, 7, 7]       2,457,600
       Identity-1057           [-1, 640, 7, 7]               0
       Identity-1058           [-1, 640, 7, 7]               0
 BatchNormAct2d-1059           [-1, 640, 7, 7]           1,280
       Identity-1060           [-1, 640, 7, 7]               0
InvertedResidual-1061           [-1, 640, 7, 7]               0
         Conv2d-1062          [-1, 3840, 7, 7]       2,457,600
       Identity-1063          [-1, 3840, 7, 7]               0
           SiLU-1064          [-1, 3840, 7, 7]               0
 BatchNormAct2d-1065          [-1, 3840, 7, 7]           7,680
         Conv2d-1066          [-1, 3840, 7, 7]          34,560
       Identity-1067          [-1, 3840, 7, 7]               0
           SiLU-1068          [-1, 3840, 7, 7]               0
 BatchNormAct2d-1069          [-1, 3840, 7, 7]           7,680
       Identity-1070          [-1, 3840, 7, 7]               0
         Conv2d-1071           [-1, 160, 1, 1]         614,560
           SiLU-1072           [-1, 160, 1, 1]               0
         Conv2d-1073          [-1, 3840, 1, 1]         618,240
        Sigmoid-1074          [-1, 3840, 1, 1]               0
  SqueezeExcite-1075          [-1, 3840, 7, 7]               0
         Conv2d-1076           [-1, 640, 7, 7]       2,457,600
       Identity-1077           [-1, 640, 7, 7]               0
       Identity-1078            [-1, 640, 7, 7]               0
 BatchNormAct2d-1079            [-1, 640, 7, 7]           1,280
       Identity-1080            [-1, 640, 7, 7]               0
InvertedResidual-1081            [-1, 640, 7, 7]               0
         Conv2d-1082           [-1, 2560, 7, 7]       1,638,400
       Identity-1083           [-1, 2560, 7, 7]               0
           SiLU-1084           [-1, 2560, 7, 7]               0
 BatchNormAct2d-1085           [-1, 2560, 7, 7]           5,120
AdaptiveAvgPool2d-1086           [-1, 2560, 1, 1]               0
        Flatten-1087                 [-1, 2560]               0
SelectAdaptivePool2d-1088                 [-1, 2560]               0
         Linear-1089                 [-1, 1024]       2,622,464
           ReLU-1090                 [-1, 1024]               0
        Dropout-1091                 [-1, 1024]               0
         Linear-1092                  [-1, 256]         262,400
           ReLU-1093                  [-1, 256]               0
         Linear-1094                    [-1, 2]             514
================================================================
Total params: 66,672,338
Trainable params: 2,885,378
Non-trainable params: 63,786,960
----------------------------------------------------------------
Input size (MB): 0.57
Forward/backward pass size (MB): 1539.97
Params size (MB): 254.33
Estimated Total Size (MB): 1794.87
----------------------------------------------------------------
```

1.14 Trainer class for training and validation

[40]:

```
# Define a named tuple for metrics
from collections import namedtuple

Metrics = namedtuple('Metrics', ['avg_loss', 'avg_accuracy', 'avg_recall',␣
 ↪'avg_precision', 'avg_f1_score'])
```

[41]:

```
class PneumoniaTrainer():

    def __init__(self, model, optimizer, criterion, device):
        self.model = model
        self.optimizer = optimizer
        self.criterion = criterion
        self.min_valid_loss = 1e10 # a large number
        self.best_valid_accuracy = 0.
        self.best_valid_recall = 0.
        self.best_valid_precision = 0.
        self.best_valid_f1_score = 0.
        self.device = device
        self.num_classes = 2

    def train_batch_loop(self, train_loader):

        self.model.to(self.device)
        self.model.train()

        total_loss = 0.

        metric_accuracy = torchmetrics.classification.
↪Accuracy(task="multiclass", num_classes=self.num_classes,␣
↪average="weighted").to(self.device)
        metric_recall = torchmetrics.classification.Recall(task="multiclass",␣
↪num_classes=self.num_classes, average="weighted").to(self.device)
        metric_precision = torchmetrics.classification.
↪Precision(task="multiclass", num_classes=self.num_classes,␣
↪average="weighted").to(self.device)
        metric_f1_score = torchmetrics.classification.
↪F1Score(task="multiclass", num_classes=self.num_classes, average="weighted").
↪to(self.device)

        for images, labels in tqdm(train_loader):
            images = images.to(self.device)
            labels = labels.to(self.device)

            logits = self.model(images)

            loss = self.criterion(logits, labels)

            self.optimizer.zero_grad()
            loss.backward()
            self.optimizer.step()

            total_loss += loss.item()

            accuracy_batch = metric_accuracy(logits, labels)
            recall_batch = metric_recall(logits, labels)
            precision_batch = metric_precision(logits, labels)
            f1_score_batch = metric_f1_score(logits, labels)


        avg_loss = total_loss / len(train_loader)
        avg_accuracy = metric_accuracy.compute().item()
        avg_recall = metric_recall.compute().item()
        avg_precision = metric_precision.compute().item()
        avg_f1_score = metric_f1_score.compute().item()
        metric_accuracy.reset()
        metric_recall.reset()
        metric_precision.reset()
        metric_f1_score.reset()

        return Metrics(avg_loss=avg_loss,
                       avg_accuracy=avg_accuracy,
                       avg_recall=avg_recall,
                       avg_precision=avg_precision,
                       avg_f1_score=avg_f1_score)

    def valid_batch_loop(self, valid_loader):

        self.model.to(self.device)
        self.model.eval() # turns off drop out and batch normalization layers

        total_loss = 0.

        metric_accuracy = torchmetrics.classification.
↪Accuracy(task="multiclass", num_classes=self.num_classes,␣
↪average="weighted").to(self.device)
        metric_recall = torchmetrics.classification.Recall(task="multiclass",␣
↪num_classes=self.num_classes, average="weighted").to(self.device)
        metric_precision = torchmetrics.classification.
↪Precision(task="multiclass", num_classes=self.num_classes,␣
↪average="weighted").to(self.device)
        metric_f1_score = torchmetrics.classification.
↪F1Score(task="multiclass", num_classes=self.num_classes, average="weighted").
↪to(self.device)

        with torch.no_grad():
            for images, labels in tqdm(valid_loader):
                images = images.to(self.device)
                labels = labels.to(self.device)

                logits = self.model(images)

                loss = self.criterion(logits, labels)

                total_loss += loss.item()

                accuracy_batch = metric_accuracy(logits, labels)
                recall_batch = metric_recall(logits, labels)
                precision_batch = metric_precision(logits, labels)
                f1_score_batch = metric_f1_score(logits, labels)
            avg_loss = total_loss / len(valid_loader)
            avg_accuracy = metric_accuracy.compute().item()
            avg_recall = metric_recall.compute().item()
            avg_precision = metric_precision.compute().item()
            avg_f1_score = metric_f1_score.compute().item()

            metric_accuracy.reset()
            metric_recall.reset()
            metric_precision.reset()
            metric_f1_score.reset()

        return Metrics(avg_loss=avg_loss,
                       avg_accuracy=avg_accuracy,
                       avg_recall=avg_recall,
                       avg_precision=avg_precision,
                       avg_f1_score=avg_f1_score)


    def fit(self, n_epoch, train_loader, valid_loader, save_filename):

        train_loss = []
        train_accuracy = []
        train_recall = []
        train_precision = []
        train_f1_score = []

        valid_loss = []
        valid_accuracy = []
        valid_recall = []
        valid_precision = []
        valid_f1_score = []

        # Early stopping parameters
        patience = 5
        patience_counter = 0

        for i_epoch in range(n_epoch):

            train_metrics = self.train_batch_loop(train_loader)

            valid_metrics = self.valid_batch_loop(valid_loader)

            train_loss.append(train_metrics.avg_loss)
            train_accuracy.append(train_metrics.avg_accuracy)
            train_recall.append(train_metrics.avg_recall)
            train_precision.append(train_metrics.avg_precision)
            train_f1_score.append(train_metrics.avg_f1_score)

            valid_loss.append(valid_metrics.avg_loss)
            valid_accuracy.append(valid_metrics.avg_accuracy)
            valid_recall.append(valid_metrics.avg_recall)
            valid_precision.append(valid_metrics.avg_precision)
            valid_f1_score.append(valid_metrics.avg_f1_score)

            print(f'Epoch: {i_epoch}, Train Loss: {train_metrics.avg_loss:.3f},␣
↪Train Acc.: {train_metrics.avg_accuracy:.3f}, ' +
                  f'Train recall: {train_metrics.avg_recall:.3f}, ' +
                  f'Train precision: {train_metrics.avg_precision:.3f}, Train␣
↪f1_score: {train_metrics.avg_f1_score:.3f}')

            print(f'Epoch: {i_epoch}, Valid Loss: {valid_metrics.avg_loss:.3f},␣
↪Valid Acc.: {valid_metrics.avg_accuracy:.3f}, ' +
                  f'Valid recall: {valid_metrics.avg_recall:.3f}, ' +
                  f'Valid precision: {valid_metrics.avg_precision:.3f}, Valid␣
↪f1_score: {valid_metrics.avg_f1_score:.3f}')

            if valid_metrics.avg_loss < self.min_valid_loss:
                self.min_valid_loss = valid_metrics.avg_loss
                self.best_valid_accuracy = valid_metrics.avg_accuracy
                self.best_valid_recall = valid_metrics.avg_recall
                self.best_valid_precision = valid_metrics.avg_precision
                self.best_valid_f1_score = valid_metrics.avg_f1_score
                patience_counter = 0
                torch.save(self.model.state_dict(), save_filename)
                print(f'Model saved to {save_filename}')
            else:
                patience_counter += 1
                if patience_counter >= patience:
                    print("Early stopping triggered.")
                    break

        print(f'\nBest Valid Loss: {self.min_valid_loss:.3f}')
        print(f'Best Valid Accuracy: {self.best_valid_accuracy:.3f}')
        print(f'Best Valid recall: {self.best_valid_recall:.3f}')
        print(f'Best Valid precision: {self.best_valid_precision:.3f}')
        print(f'Best Valid f1_score: {self.best_valid_f1_score:.3f}')

        train_lists = {'train_loss': train_loss,
                       'train_accuracy': train_accuracy,
                       'train_recall': train_recall,
                       'train_precision': train_precision,
                       'train_f1_score': train_f1_score
                       }
        valid_lists = {'valid_loss': valid_loss,
                       'valid_accuracy': valid_accuracy,
                       'valid_recall': valid_recall,
                       'valid_precision': valid_precision,
                       'valid_f1_score': valid_f1_score
                       }


        return train_lists, valid_lists
```

1.15 Train and fine-tune the models

[42]:
```
cnn_optimizer = torch.optim.Adam(cnn_model.parameters(), lr=config.lr_custom)
pretrained_optimizer = torch.optim.Adam(model.parameters(), lr=config.
  ↪lr_pretrained)
criterion = nn.CrossEntropyLoss(weight=class_weights.to(device))

cnn_trainer = PneumoniaTrainer(cnn_model, cnn_optimizer, criterion, device)
pretrained_trainer = PneumoniaTrainer(model, pretrained_optimizer, criterion,␣
  ↪device)

print('Training custom CNN model:')
n_epoch = 20
train_lists1, valid_lists1 = cnn_trainer.fit(n_epoch, train_loader_custom,␣
  ↪valid_loader_custom, 'pneumonia_cnn_model.pt')

print('Fine-tuning pretrained model:')
n_epoch = 20
save_filename_pretrained = 'pneumonia_pretrained_model_' + config.model_name +␣
  ↪'.pt'
train_lists2, valid_lists2  = pretrained_trainer.fit(n_epoch,␣
  ↪train_loader_pretrained, valid_loader_pretrained, save_filename_pretrained)
```

```
Training custom CNN model:

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 0, Train Loss: 0.573, Train Acc.: 0.677, Train recall: 0.677, Train
precision: 0.779, Train f1_score: 0.698
Epoch: 0, Valid Loss: 1.446, Valid Acc.: 0.500, Valid recall: 0.500, Valid
precision: 0.250, Valid f1_score: 0.333
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]
```

```
Epoch: 1, Train Loss: 0.392, Train Acc.: 0.785, Train recall: 0.785, Train
precision: 0.867, Train f1_score: 0.798
Epoch: 1, Valid Loss: 0.588, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.767, Valid f1_score: 0.746
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 2, Train Loss: 0.353, Train Acc.: 0.811, Train recall: 0.811, Train
precision: 0.879, Train f1_score: 0.823
Epoch: 2, Valid Loss: 0.558, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 3, Train Loss: 0.327, Train Acc.: 0.825, Train recall: 0.825, Train
precision: 0.886, Train f1_score: 0.835
Epoch: 3, Valid Loss: 0.530, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 4, Train Loss: 0.312, Train Acc.: 0.842, Train recall: 0.842, Train
precision: 0.896, Train f1_score: 0.851
Epoch: 4, Valid Loss: 0.514, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 5, Train Loss: 0.288, Train Acc.: 0.851, Train recall: 0.851, Train
precision: 0.899, Train f1_score: 0.859
Epoch: 5, Valid Loss: 0.456, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.767, Valid f1_score: 0.746
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 6, Train Loss: 0.279, Train Acc.: 0.852, Train recall: 0.852, Train
precision: 0.899, Train f1_score: 0.860
Epoch: 6, Valid Loss: 0.430, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.767, Valid f1_score: 0.746
Model saved to pneumonia_cnn_model.pt
```

```
  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 7, Train Loss: 0.268, Train Acc.: 0.867, Train recall: 0.867, Train
precision: 0.907, Train f1_score: 0.874
Epoch: 7, Valid Loss: 0.421, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.767, Valid f1_score: 0.746
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 8, Train Loss: 0.252, Train Acc.: 0.870, Train recall: 0.870, Train
precision: 0.908, Train f1_score: 0.877
Epoch: 8, Valid Loss: 0.401, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.817, Valid f1_score: 0.812
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 9, Train Loss: 0.242, Train Acc.: 0.874, Train recall: 0.874, Train
precision: 0.911, Train f1_score: 0.880
Epoch: 9, Valid Loss: 0.405, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.767, Valid f1_score: 0.746

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 10, Train Loss: 0.235, Train Acc.: 0.880, Train recall: 0.880, Train
precision: 0.913, Train f1_score: 0.886
Epoch: 10, Valid Loss: 0.366, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.817, Valid f1_score: 0.812
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 11, Train Loss: 0.228, Train Acc.: 0.887, Train recall: 0.887, Train
precision: 0.916, Train f1_score: 0.893
Epoch: 11, Valid Loss: 0.328, Valid Acc.: 0.938, Valid recall: 0.938, Valid
precision: 0.944, Valid f1_score: 0.937
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 12, Train Loss: 0.217, Train Acc.: 0.889, Train recall: 0.889, Train
precision: 0.918, Train f1_score: 0.894
```

```
Epoch: 12, Valid Loss: 0.338, Valid Acc.: 0.875, Valid recall: 0.875, Valid
precision: 0.900, Valid f1_score: 0.873

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 13, Train Loss: 0.217, Train Acc.: 0.899, Train recall: 0.899, Train
precision: 0.923, Train f1_score: 0.904
Epoch: 13, Valid Loss: 0.324, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.817, Valid f1_score: 0.812
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 14, Train Loss: 0.209, Train Acc.: 0.901, Train recall: 0.901, Train
precision: 0.924, Train f1_score: 0.905
Epoch: 14, Valid Loss: 0.310, Valid Acc.: 0.938, Valid recall: 0.938, Valid
precision: 0.944, Valid f1_score: 0.937
Model saved to pneumonia_cnn_model.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 15, Train Loss: 0.204, Train Acc.: 0.909, Train recall: 0.909, Train
precision: 0.929, Train f1_score: 0.913
Epoch: 15, Valid Loss: 0.319, Valid Acc.: 0.875, Valid recall: 0.875, Valid
precision: 0.875, Valid f1_score: 0.875

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 16, Train Loss: 0.192, Train Acc.: 0.915, Train recall: 0.915, Train
precision: 0.932, Train f1_score: 0.918
Epoch: 16, Valid Loss: 0.321, Valid Acc.: 0.875, Valid recall: 0.875, Valid
precision: 0.900, Valid f1_score: 0.873

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 17, Train Loss: 0.197, Train Acc.: 0.912, Train recall: 0.912, Train
precision: 0.930, Train f1_score: 0.915
Epoch: 17, Valid Loss: 0.346, Valid Acc.: 0.875, Valid recall: 0.875, Valid
precision: 0.900, Valid f1_score: 0.873

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 18, Train Loss: 0.185, Train Acc.: 0.917, Train recall: 0.917, Train
precision: 0.933, Train f1_score: 0.920
```

```
Epoch: 18, Valid Loss: 0.349, Valid Acc.: 0.875, Valid recall: 0.875, Valid
precision: 0.900, Valid f1_score: 0.873

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 19, Train Loss: 0.178, Train Acc.: 0.925, Train recall: 0.925, Train
precision: 0.938, Train f1_score: 0.928
Epoch: 19, Valid Loss: 0.337, Valid Acc.: 0.875, Valid recall: 0.875, Valid
precision: 0.900, Valid f1_score: 0.873
Early stopping triggered.

Best Valid Loss: 0.310
Best Valid Accuracy: 0.938
Best Valid recall: 0.938
Best Valid precision: 0.944
Best Valid f1_score: 0.937
Fine-tuning pretrained model:

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 0, Train Loss: 0.684, Train Acc.: 0.513, Train recall: 0.513, Train
precision: 0.798, Train f1_score: 0.519
Epoch: 0, Valid Loss: 0.663, Valid Acc.: 0.625, Valid recall: 0.625, Valid
precision: 0.786, Valid f1_score: 0.564
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 1, Train Loss: 0.656, Train Acc.: 0.738, Train recall: 0.738, Train
precision: 0.866, Train f1_score: 0.755
Epoch: 1, Valid Loss: 0.618, Valid Acc.: 0.688, Valid recall: 0.688, Valid
precision: 0.808, Valid f1_score: 0.654
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 2, Train Loss: 0.615, Train Acc.: 0.797, Train recall: 0.797, Train
precision: 0.882, Train f1_score: 0.809
Epoch: 2, Valid Loss: 0.557, Valid Acc.: 0.688, Valid recall: 0.688, Valid
precision: 0.808, Valid f1_score: 0.654
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]
```

```
Epoch: 3, Train Loss: 0.553, Train Acc.: 0.833, Train recall: 0.833, Train
precision: 0.894, Train f1_score: 0.843
Epoch: 3, Valid Loss: 0.490, Valid Acc.: 0.688, Valid recall: 0.688, Valid
precision: 0.808, Valid f1_score: 0.654
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 4, Train Loss: 0.481, Train Acc.: 0.850, Train recall: 0.850, Train
precision: 0.901, Train f1_score: 0.859
Epoch: 4, Valid Loss: 0.431, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.833, Valid f1_score: 0.733
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 5, Train Loss: 0.410, Train Acc.: 0.865, Train recall: 0.865, Train
precision: 0.907, Train f1_score: 0.872
Epoch: 5, Valid Loss: 0.393, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.833, Valid f1_score: 0.733
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 6, Train Loss: 0.348, Train Acc.: 0.876, Train recall: 0.876, Train
precision: 0.911, Train f1_score: 0.882
Epoch: 6, Valid Loss: 0.368, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.833, Valid f1_score: 0.733
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 7, Train Loss: 0.303, Train Acc.: 0.891, Train recall: 0.891, Train
precision: 0.917, Train f1_score: 0.895
Epoch: 7, Valid Loss: 0.356, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.833, Valid f1_score: 0.733
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 8, Train Loss: 0.275, Train Acc.: 0.898, Train recall: 0.898, Train
precision: 0.920, Train f1_score: 0.902
Epoch: 8, Valid Loss: 0.342, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.833, Valid f1_score: 0.733
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt
```

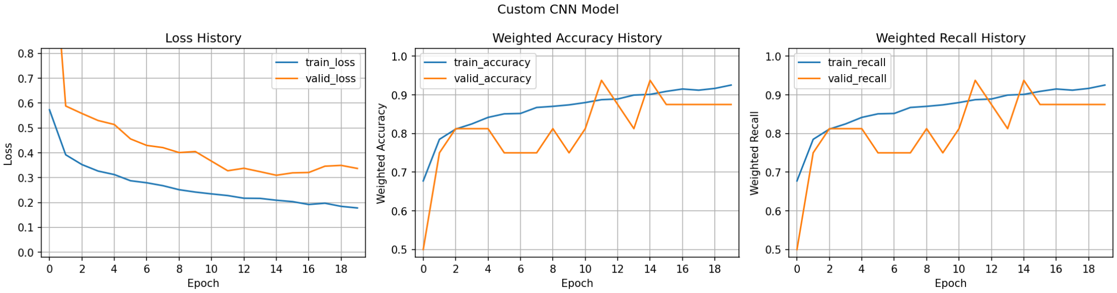
```
  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 9, Train Loss: 0.249, Train Acc.: 0.900, Train recall: 0.900, Train
precision: 0.922, Train f1_score: 0.904
Epoch: 9, Valid Loss: 0.340, Valid Acc.: 0.750, Valid recall: 0.750, Valid
precision: 0.833, Valid f1_score: 0.733
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 10, Train Loss: 0.228, Train Acc.: 0.906, Train recall: 0.906, Train
precision: 0.924, Train f1_score: 0.909
Epoch: 10, Valid Loss: 0.330, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 11, Train Loss: 0.212, Train Acc.: 0.913, Train recall: 0.913, Train
precision: 0.929, Train f1_score: 0.916
Epoch: 11, Valid Loss: 0.328, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 12, Train Loss: 0.200, Train Acc.: 0.922, Train recall: 0.922, Train
precision: 0.936, Train f1_score: 0.925
Epoch: 12, Valid Loss: 0.317, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 13, Train Loss: 0.196, Train Acc.: 0.921, Train recall: 0.921, Train
precision: 0.934, Train f1_score: 0.924
Epoch: 13, Valid Loss: 0.317, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 14, Train Loss: 0.189, Train Acc.: 0.922, Train recall: 0.922, Train
precision: 0.934, Train f1_score: 0.924
Epoch: 14, Valid Loss: 0.315, Valid Acc.: 0.812, Valid recall: 0.812, Valid
```

```
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 15, Train Loss: 0.179, Train Acc.: 0.927, Train recall: 0.927, Train
precision: 0.938, Train f1_score: 0.929
Epoch: 15, Valid Loss: 0.308, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 16, Train Loss: 0.169, Train Acc.: 0.930, Train recall: 0.930, Train
precision: 0.940, Train f1_score: 0.932
Epoch: 16, Valid Loss: 0.310, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 17, Train Loss: 0.161, Train Acc.: 0.937, Train recall: 0.937, Train
precision: 0.946, Train f1_score: 0.939
Epoch: 17, Valid Loss: 0.305, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 18, Train Loss: 0.161, Train Acc.: 0.936, Train recall: 0.936, Train
precision: 0.944, Train f1_score: 0.938
Epoch: 18, Valid Loss: 0.299, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806
Model saved to pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt

  0%|          | 0/41 [00:00<?, ?it/s]

  0%|          | 0/1 [00:00<?, ?it/s]

Epoch: 19, Train Loss: 0.162, Train Acc.: 0.933, Train recall: 0.933, Train
precision: 0.941, Train f1_score: 0.934
Epoch: 19, Valid Loss: 0.309, Valid Acc.: 0.812, Valid recall: 0.812, Valid
precision: 0.864, Valid f1_score: 0.806

Best Valid Loss: 0.299
Best Valid Accuracy: 0.812
Best Valid recall: 0.812
Best Valid precision: 0.864
Best Valid f1_score: 0.806
```

1.16 Plot the histories

[43]:

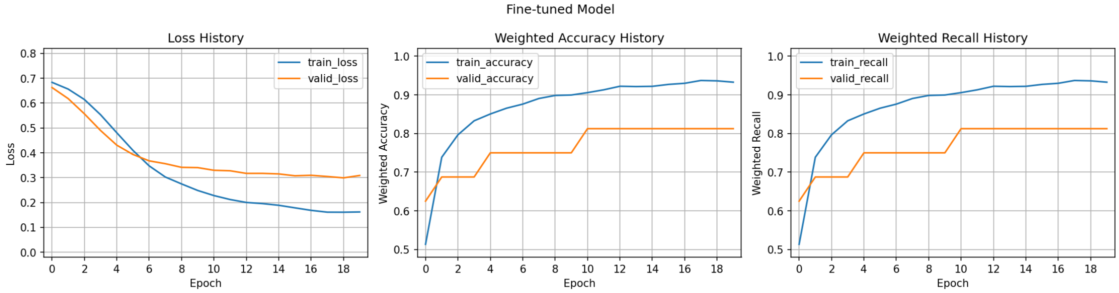
```
fig, ax = plt.subplots(nrows=1, ncols=3, figsize=(15,4), dpi=150)
plt.suptitle("Custom CNN Model")
ax[0].plot(train_lists1['train_loss'], label='train_loss')
ax[0].plot(valid_lists1['valid_loss'], label='valid_loss')
ax[0].set_title('Loss History')
ax[0].set_xlabel('Epoch')
ax[0].set_ylabel('Loss')
ax[0].legend(loc='best')
ax[0].grid(True)
ax[0].set_xlim([-0.5, n_epoch-0.5])
ax[0].set_ylim([-0.02, 0.82])
ax[0].set_xticks(np.arange(0, n_epoch, step=2))

ax[1].plot(train_lists1['train_accuracy'], label='train_accuracy')
ax[1].plot(valid_lists1['valid_accuracy'], label='valid_accuracy')
ax[1].set_title('Weighted Accuracy History')
ax[1].set_xlabel('Epoch')
ax[1].set_ylabel('Weighted Accuracy')
ax[1].legend(loc='best')
ax[1].grid(True)
ax[1].set_xlim([-0.5, n_epoch-0.5])
ax[1].set_ylim([0.48, 1.02])
ax[1].set_xticks(np.arange(0, n_epoch, step=2))

ax[2].plot(train_lists1['train_recall'], label='train_recall')
ax[2].plot(valid_lists1['valid_recall'], label='valid_recall')
ax[2].set_title('Weighted Recall History')
ax[2].set_xlabel('Epoch')
ax[2].set_ylabel('Weighted Recall')
ax[2].legend(loc='best')
ax[2].grid(True)
ax[2].set_xlim([-0.5, n_epoch-0.5])
ax[2].set_ylim([0.48, 1.02])
ax[2].set_xticks(np.arange(0, n_epoch, step=2))

fig.tight_layout();

fig_filename = 'pneumonia_custom_model_' + f'_lr_{config.lr_custom}_bs_{config.
  ↪batch_size}' + '.png'
fig.savefig(fig_filename)
```



[44]:

```
fig, ax = plt.subplots(nrows=1, ncols=3, figsize=(15,4), dpi=150)
plt.suptitle("Fine-tuned Model")
ax[0].plot(train_lists2['train_loss'], label='train_loss')
ax[0].plot(valid_lists2['valid_loss'], label='valid_loss')
ax[0].set_title('Loss History')
ax[0].set_xlabel('Epoch')
ax[0].set_ylabel('Loss')
ax[0].legend(loc='best')
ax[0].grid(True)
ax[0].set_xlim([-0.5, n_epoch-0.5])
ax[0].set_ylim([-0.02, 0.82])
ax[0].set_xticks(np.arange(0, n_epoch, step=2))

ax[1].plot(train_lists2['train_accuracy'], label='train_accuracy')
ax[1].plot(valid_lists2['valid_accuracy'], label='valid_accuracy')
ax[1].set_title('Weighted Accuracy History')
ax[1].set_xlabel('Epoch')
ax[1].set_ylabel('Weighted Accuracy')
ax[1].legend(loc='best')
ax[1].grid(True)
ax[1].set_xlim([-0.5, n_epoch-0.5])
ax[1].set_ylim([0.48, 1.02])
ax[1].set_xticks(np.arange(0, n_epoch, step=2))

ax[2].plot(train_lists2['train_recall'], label='train_recall')
ax[2].plot(valid_lists2['valid_recall'], label='valid_recall')
ax[2].set_title('Weighted Recall History')
ax[2].set_xlabel('Epoch')
ax[2].set_ylabel('Weighted Recall')
ax[2].legend(loc='best')
ax[2].grid(True)
ax[2].set_xlim([-0.5, n_epoch-0.5])
ax[2].set_ylim([0.48, 1.02])
ax[2].set_xticks(np.arange(0, n_epoch, step=2))

fig.tight_layout();
```

```
fig_filename = 'pneumonia_pretrained_model_' + config.model_name +␣
↪f'_lr_{config.lr_pretrained}_bs_{config.batch_size}' + '.png'
fig.savefig(fig_filename)
```



1.17 Evaluate the performance on testset

[45]:

```
status = cnn_model.load_state_dict(torch.load('pneumonia_cnn_model.pt',␣
↪map_location=device, weights_only=True))
print(f"Status: {status}")

model_size = os.path.getsize('pneumonia_cnn_model.pt') / (1024 * 1024)  # in MB
print(f"Custom CNN model size: {model_size:.2f} MB")

test_metrics = cnn_trainer.valid_batch_loop(test_loader_custom)

print(f"Test Loss: {test_metrics.avg_loss:.3f}")
print(f"Test Acc.: {test_metrics.avg_accuracy:.3f}")
print(f"Test Recall: {test_metrics.avg_recall:.3f}")
print(f"Test Precision: {test_metrics.avg_precision:.3f}")
print(f"Test F1-score: {test_metrics.avg_f1_score:.3f}")
```

```
Status: <All keys matched successfully>
Custom CNN model size: 12.66 MB

  0%|          | 0/5 [00:00<?, ?it/s]

Test Loss: 0.317
Test Acc.: 0.894
Test Recall: 0.894
Test Precision: 0.896
Test F1-score: 0.895
```

[46]:

```
save_filename_pretrained = 'pneumonia_pretrained_model_' + config.model_name +␣
↪'.pt'
```

```
status = model.load_state_dict(torch.load('/content/' +␣
 ↪save_filename_pretrained, map_location=device, weights_only=True))
print(f"Status: {status}")
print(f"save_filename_pretrained: {save_filename_pretrained}")

model_size = os.path.getsize(save_filename_pretrained) / (1024 * 1024)  # in MB
print(f"Pretrained model size: {model_size:.2f} MB")

test_metrics_pretrained = pretrained_trainer.
 ↪valid_batch_loop(test_loader_pretrained)

print(f"Test Loss: {test_metrics_pretrained.avg_loss:.3f}")
print(f"Test Acc.: {test_metrics_pretrained.avg_accuracy:.3f}")
print(f"Test Recall: {test_metrics_pretrained.avg_recall:.3f}")
print(f"Test Precision: {test_metrics_pretrained.avg_precision:.3f}")
print(f"Test F1-score: {test_metrics_pretrained.avg_f1_score:.3f}")
```

```
Status: <All keys matched successfully>
save_filename_pretrained:
pneumonia_pretrained_model_tf_efficientnet_b7.ap_in1k.pt
Pretrained model size: 255.98 MB
```

```
  0%|          | 0/5 [00:00<?, ?it/s]
```

```
Test Loss: 0.336
Test Acc.: 0.872
Test Recall: 0.872
Test Precision: 0.876
Test F1-score: 0.873
```

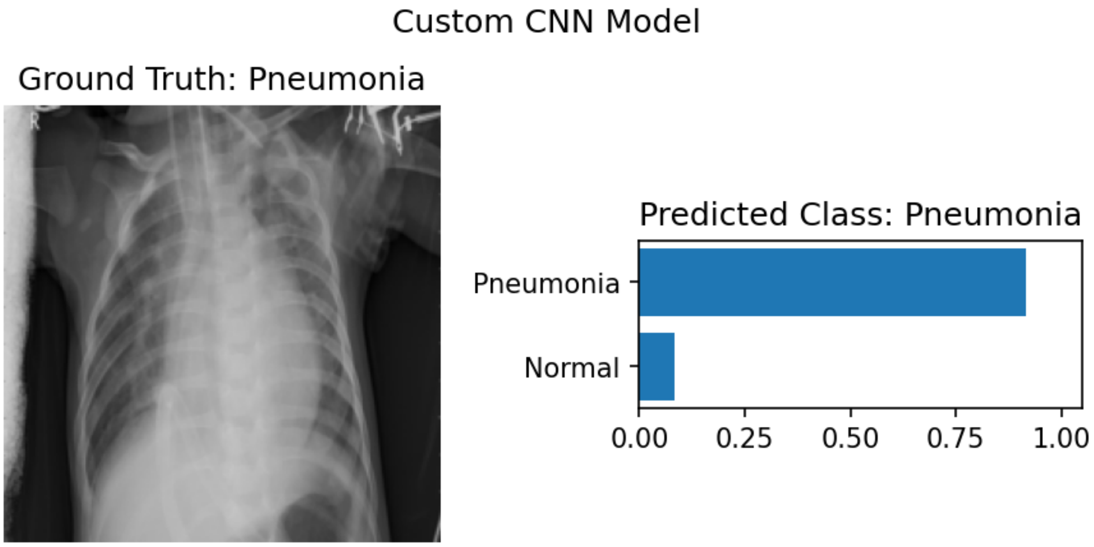
1.18 Show the prediction for a sample image

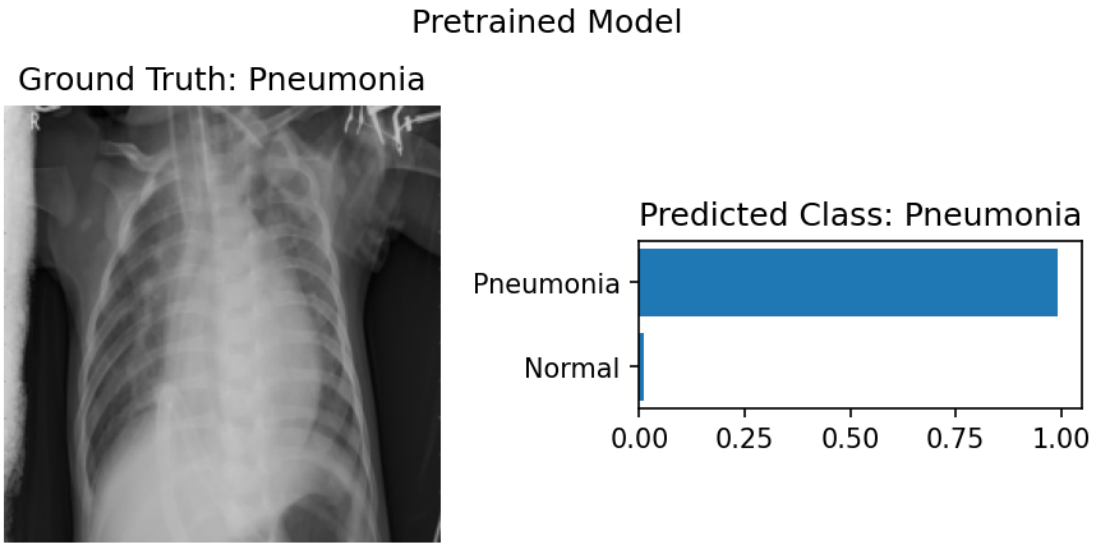
[47]:

```
def show_predicted_class(image, label_prob, label, mean, std, title):

    class_names = ['Normal', 'Pneumonia']

    label_prob = label_prob.cpu().detach().numpy().squeeze(0)

    image = image.to('cpu')
    image = image.permute(1, 2, 0)
    mean = torch.FloatTensor(mean)
    std = torch.FloatTensor(std)
    image = image*std + mean
    img = torch.clip(image, 0, 1)
    img = img.numpy()

    fig, (ax1, ax2) = plt.subplots(figsize=(6,3), ncols=2, dpi=150)
    plt.suptitle(title)
    ax1.imshow(img, cmap="gray")
    ax1.set_title(f'Ground Truth: {class_names[label]}')
    ax1.axis(False)

    ax2.barh(class_names, label_prob)
    ax2.set_aspect(0.2)
    ax2.set_yticks(class_names)
    ax2.set_yticklabels(class_names)
    ax2.set_title(f'Predicted Class: {class_names[label_prob.argmax()]}')
    ax2.set_xlim(0, 1.05)

    plt.tight_layout()
```

[48]:

```
import random

random_int = random.randint(0, len(testset_custom)-1)
# For custom model
image, label = testset_custom[random_int]
image = image.to(device)

logit = cnn_model(image.unsqueeze(0))
label_predict = F.softmax(logit, dim = 1)

show_predicted_class(image, label_predict, label, config.pn_mean, config.
 ↪pn_std, title="Custom CNN Model")

# For fine-tuned model
image, label = testset_pretrained[random_int]
image = image.to(device)

logit = model(image.unsqueeze(0))
label_predict = F.softmax(logit, dim = 1)

show_predicted_class(image, label_predict, label, config.in_mean, config.
 ↪in_std, title="Pretrained Model")
```


Custom CNN Model
Ground Truth: Pneumonia
Predicted Class: Pneumonia
Pneumonia
Normal
0.00
0.25
0.50
0.75
1.00


Pretrained Model
Ground Truth: Pneumonia
Predicted Class: Pneumonia
Pneumonia
Normal
0.00
0.25
0.50
0.75
1.00

[48]:

[48]: