```
In [1]:
```

```python
# This Python 3 environment comes with many helpful analytics libraries installed
# It is defined by the kaggle/python docker image: https://github.com/kaggle/docker-python
# For example, here's several helpful packages to load in

import numpy as np # linear algebra
import pandas as pd # data processing, CSV file I/O (e.g. pd.read_csv)

# Input data files are available in the "../input/" directory.
# For example, running this (by clicking run or pressing Shift+Enter) will list all files
# under the input directory

import os
for dirname, _, filenames in os.walk('/kaggle/input'):
    for filename in filenames:
        pass
        #print(os.path.join(dirname, filename))

# Any results you write to the current directory are saved as output.
```

```
In [ ]:
```

```python
# Сделано в интерфейсе Kaggle
# Лишние блоки кода из пайплайна убраны из-за их бесполезности
# Score 98.405
# Team Name: Владимир_Терентьев_341670400
```

```
In [2]:
```

```python
import torch
import numpy as np

train_on_gpu = torch.cuda.is_available()

if not train_on_gpu:
    print('CUDA is not available.  Training on CPU ...')
else:
    print('CUDA is available!  Training on GPU ...')
```

```
CUDA is available!  Training on GPU ...
```

```
In [3]:
```

```python
import pickle
import numpy as np
import random
from skimage import io

from tqdm import tqdm, tqdm_notebook
from PIL import Image
from pathlib import Path

from torchvision import transforms
from multiprocessing.pool import ThreadPool
from sklearn.preprocessing import LabelEncoder
from torch.utils.data import Dataset, DataLoader
import torch.nn as nn

from matplotlib import colors, pyplot as plt
%matplotlib inline

# в sklearn не все гладко, чтобы в colab удобно выводить картинки
# мы будем игнорировать warnings
import warnings
warnings.filterwarnings(action='ignore', category=DeprecationWarning)
```

```python
# разные режимы датасета
DATA_MODES = ['train', 'val', 'test']
# все изображения будут масштабированы к размеру 224x224 px
RESCALE_SIZE = 224
# работаем на видеокарте
DEVICE = torch.device("cuda")
```

```python
class SimpsonsDataset(Dataset):
    """
    Датасет с картинками, который паралельно подгружает их из папок
    производит скалирование и превращение в торчевые тензоры
    """
    def __init__(self, files, mode):
        super().__init__()
        # список файлов для загрузки
        self.files = sorted(files)
        # режим работы
        self.mode = mode

        if self.mode not in DATA_MODES:
            print(f"{self.mode} is not correct; correct modes: {DATA_MODES}")
            raise NameError

        self.len_ = len(self.files)

        self.label_encoder = LabelEncoder()

        if self.mode != 'test':
            self.labels = [path.parent.name for path in self.files]
            self.label_encoder.fit(self.labels)

            with open('label_encoder.pkl', 'wb') as le_dump_file:
                pickle.dump(self.label_encoder, le_dump_file)

    def __len__(self):
        return self.len_

    def load_sample(self, file):
        image = Image.open(file)
        image.load()
        return image

    def __getitem__(self, index):
        # для преобразования изображений в тензоры PyTorch и нормализации входа
        transform = transforms.Compose([
            transforms.ToTensor(),
            transforms.Normalize([0.485, 0.456, 0.406], [0.229, 0.224, 0.225])
        ])
        x = self.load_sample(self.files[index])
        x = self._prepare_sample(x)
        x = np.array(x / 255, dtype='float32')
        x = transform(x)
        if self.mode == 'test':
            return x
        else:
            label = self.labels[index]
            label_id = self.label_encoder.transform([label])
            y = label_id.item()
            return x, y

    def _prepare_sample(self, image):
        image = image.resize((RESCALE_SIZE, RESCALE_SIZE))
        return np.array(image)
```

```python
def imshow(inp, title=None, plt_ax=plt, default=False):
    """Imshow для тензоров"""
```

```python
        inp = inp.numpy().transpose((1, 2, 0))
        mean = np.array([0.485, 0.456, 0.406])
        std = np.array([0.229, 0.224, 0.225])
        inp = std * inp + mean
        inp = np.clip(inp, 0, 1)
        plt_ax.imshow(inp)
        if title is not None:
            plt_ax.set_title(title)
        plt_ax.grid(False)
```

In [7]:

```python
TRAIN_DIR = Path('/kaggle/input/journey-springfield/train/simpsons_dataset')
TEST_DIR = Path('/kaggle/input/journey-springfield/testset/testset')

train_val_files = sorted(list(TRAIN_DIR.rglob('*.jpg')))
test_files = sorted(list(TEST_DIR.rglob('*.jpg')))
```

In [8]:

```python
from sklearn.model_selection import train_test_split

train_val_labels = [path.parent.name for path in train_val_files]
train_files, val_files = train_test_split(train_val_files, stratify=train_val_labels)
```

In [9]:

```python
val_dataset = SimpsonsDataset(val_files, mode='val')
```

In [10]:

```python
def fit_epoch(model, train_loader, criterion, optimizer):
    running_loss = 0.0
    running_corrects = 0
    processed_data = 0

    for inputs, labels in train_loader:
        inputs = inputs.to(DEVICE)
        labels = labels.to(DEVICE)
        optimizer.zero_grad()

        outputs = model(inputs)
        loss = criterion(outputs, labels)
        loss.backward()
        optimizer.step()
        preds = torch.argmax(outputs, 1)
        running_loss += loss.item() * inputs.size(0)
        running_corrects += torch.sum(preds == labels.data)
        processed_data += inputs.size(0)

    train_loss = running_loss / processed_data
    train_acc = running_corrects.cpu().numpy() / processed_data
    return train_loss, train_acc
```

In [11]:

```python
def eval_epoch(model, val_loader, criterion):
    model.eval()
    running_loss = 0.0
    running_corrects = 0
    processed_size = 0

    for inputs, labels in val_loader:
        inputs = inputs.to(DEVICE)
        labels = labels.to(DEVICE)

        with torch.set_grad_enabled(False):
            outputs = model(inputs)
            loss = criterion(outputs, labels)
            preds = torch.argmax(outputs, 1)
```

```
            running_loss += loss.item() * inputs.size(0)
            running_corrects += torch.sum(preds == labels.data)
            processed_size += inputs.size(0)
    val_loss = running_loss / processed_size
    val_acc = running_corrects.double() / processed_size
    return val_loss, val_acc
```

In [12]:

```python
def train(train_files, val_files, model, epochs, batch_size):
    train_loader = DataLoader(train_dataset, batch_size=batch_size, shuffle=True)
    val_loader = DataLoader(val_dataset, batch_size=batch_size, shuffle=False)

    history = []
    log_template = "\nEpoch {ep:03d} train_loss: {t_loss:0.4f} \
    val_loss {v_loss:0.4f} train_acc {t_acc:0.4f} val_acc {v_acc:0.4f}"

    with tqdm(desc="epoch", total=epochs) as pbar_outer:
        opt = torch.optim.Adam(model.parameters())
        criterion = nn.CrossEntropyLoss()

        for epoch in range(epochs):
            train_loss, train_acc = fit_epoch(model, train_loader, criterion, opt)
            print("loss", train_loss)

            val_loss, val_acc = eval_epoch(model, val_loader, criterion)
            history.append((train_loss, train_acc, val_loss, val_acc))

            pbar_outer.update(1)
            tqdm.write(log_template.format(ep=epoch+1, t_loss=train_loss,\
                                           v_loss=val_loss, t_acc=train_acc, v_acc=val_a
cc))

    return history
```

In [13]:

```python
def predict(model, test_loader):
    with torch.no_grad():
        logits = []

        for inputs in test_loader:
            inputs = inputs.to(DEVICE)
            model.eval()
            outputs = model(inputs).cpu()
            logits.append(outputs)

    probs = nn.functional.softmax(torch.cat(logits), dim=-1).numpy()
    return probs
```

In [14]:

```python
if val_dataset is None:
    val_dataset = SimpsonsDataset(val_files, mode='val')

train_dataset = SimpsonsDataset(train_files, mode='train')
```

In [15]:

```python
# Лучшая модель
!pip install efficientnet_pytorch
```

```
Collecting efficientnet_pytorch
  Downloading efficientnet_pytorch-0.7.1.tar.gz (21 kB)
Requirement already satisfied: torch in /opt/conda/lib/python3.7/site-packages (from effi
cientnet_pytorch) (1.7.0)
Requirement already satisfied: future in /opt/conda/lib/python3.7/site-packages (from tor
ch->efficientnet_pytorch) (0.18.2)
Requirement already satisfied: typing_extensions in /opt/conda/lib/python3.7/site-package
s (from torch->efficientnet_pytorch) (3.7.4.3)
Requirement already satisfied: dataclasses in /opt/conda/lib/python3.7/site-packages (fro
```

```
m torch->efficientnet_pytorch) (0.6)
Requirement already satisfied: numpy in /opt/conda/lib/python3.7/site-packages (from torc
h->efficientnet_pytorch) (1.19.5)
Building wheels for collected packages: efficientnet-pytorch
  Building wheel for efficientnet-pytorch (setup.py) ... - \ done
  Created wheel for efficientnet-pytorch: filename=efficientnet_pytorch-0.7.1-py3-none-an
y.whl size=16446 sha256=8fc6fcde9c8c4be689ea46b2b82bd398870f3d7437aa4b6acd5535fd55be29f5
  Stored in directory: /root/.cache/pip/wheels/0e/cc/b2/49e74588263573ff778da58cc99b9c634
9b496636a7e165be6
Successfully built efficientnet-pytorch
Installing collected packages: efficientnet-pytorch
Successfully installed efficientnet-pytorch-0.7.1
```

In [16]:

```python
from efficientnet_pytorch import EfficientNet
model = EfficientNet.from_pretrained('efficientnet-b3')
```

```
Downloading: "https://github.com/lukemelas/EfficientNet-PyTorch/releases/download/1.0/eff
icientnet-b3-5fb5a3c3.pth" to /root/.cache/torch/hub/checkpoints/efficientnet-b3-5fb5a3c3
.pth
```

```
Loaded pretrained weights for efficientnet-b3
```

In [ ]:

```python
torch.cuda.empty_cache()
# очистка сиды
```

In [18]:

```python
model
```

Out[18]:

```
EfficientNet(
  (_conv_stem): Conv2dStaticSamePadding(
    3, 40, kernel_size=(3, 3), stride=(2, 2), bias=False
    (static_padding): ZeroPad2d(padding=(0, 1, 0, 1), value=0.0)
  )
  (_bn0): BatchNorm2d(40, eps=0.001, momentum=0.010000000000000009, affine=True, track_ru
nning_stats=True)
  (_blocks): ModuleList(
    (0): MBConvBlock(
      (_depthwise_conv): Conv2dStaticSamePadding(
        40, 40, kernel_size=(3, 3), stride=[1, 1], groups=40, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(40, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        40, 10, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        10, 40, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        40, 24, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(24, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (1): MBConvBlock(
      (_depthwise_conv): Conv2dStaticSamePadding(
        24, 24, kernel_size=(3, 3), stride=(1, 1), groups=24, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(24, eps=0.001, momentum=0.010000000000000009, affine=True, trac
```

```
k_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        24, 6, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        6, 24, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        24, 24, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(24, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (2): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        24, 144, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(144, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        144, 144, kernel_size=(3, 3), stride=[2, 2], groups=144, bias=False
        (static_padding): ZeroPad2d(padding=(0, 1, 0, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(144, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        144, 6, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        6, 144, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        144, 32, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(32, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (3): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        32, 192, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        192, 192, kernel_size=(3, 3), stride=(1, 1), groups=192, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        192, 8, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        8, 192, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        192, 32, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(32, eps=0.001, momentum=0.010000000000000009, affine=True, trac
```

```
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (4): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        32, 192, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        192, 192, kernel_size=(3, 3), stride=(1, 1), groups=192, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        192, 8, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        8, 192, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        192, 32, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(32, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (5): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        32, 192, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        192, 192, kernel_size=(5, 5), stride=[2, 2], groups=192, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        192, 8, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        8, 192, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        192, 48, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(48, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (6): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        48, 288, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        288, 288, kernel_size=(5, 5), stride=(1, 1), groups=288, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
```

```
      (_bn1): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        288, 12, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        12, 288, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        288, 48, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(48, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (7): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        48, 288, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        288, 288, kernel_size=(5, 5), stride=(1, 1), groups=288, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        288, 12, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        12, 288, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        288, 48, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(48, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (8): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        48, 288, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        288, 288, kernel_size=(3, 3), stride=[2, 2], groups=288, bias=False
        (static_padding): ZeroPad2d(padding=(0, 1, 0, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        288, 12, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        12, 288, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        288, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
```

```
      (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (9): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        576, 576, kernel_size=(3, 3), stride=(1, 1), groups=576, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        576, 24, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        24, 576, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        576, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (10): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        576, 576, kernel_size=(3, 3), stride=(1, 1), groups=576, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        576, 24, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        24, 576, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        576, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (11): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        576, 576, kernel_size=(3, 3), stride=(1, 1), groups=576, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
```

```
    )
    (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_se_reduce): Conv2dStaticSamePadding(
      576, 24, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_se_expand): Conv2dStaticSamePadding(
      24, 576, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_project_conv): Conv2dStaticSamePadding(
      576, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
    (_swish): MemoryEfficientSwish()
  )
  (12): MBConvBlock(
    (_expand_conv): Conv2dStaticSamePadding(
      96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_depthwise_conv): Conv2dStaticSamePadding(
      576, 576, kernel_size=(3, 3), stride=(1, 1), groups=576, bias=False
      (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
    )
    (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_se_reduce): Conv2dStaticSamePadding(
      576, 24, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_se_expand): Conv2dStaticSamePadding(
      24, 576, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_project_conv): Conv2dStaticSamePadding(
      576, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
    (_swish): MemoryEfficientSwish()
  )
  (13): MBConvBlock(
    (_expand_conv): Conv2dStaticSamePadding(
      96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_depthwise_conv): Conv2dStaticSamePadding(
      576, 576, kernel_size=(5, 5), stride=[1, 1], groups=576, bias=False
      (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
    )
    (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_se_reduce): Conv2dStaticSamePadding(
      576, 24, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_se_expand): Conv2dStaticSamePadding(
      24, 576, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_project_conv): Conv2dStaticSamePadding(
      576, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
```

```
      )
      (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (14): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        816, 816, kernel_size=(5, 5), stride=(1, 1), groups=816, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        816, 34, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        34, 816, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        816, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (15): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        816, 816, kernel_size=(5, 5), stride=(1, 1), groups=816, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        816, 34, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        34, 816, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        816, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (16): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        816, 816, kernel_size=(5, 5), stride=(1, 1), groups=816, bias=False
```

```
      (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
    )
    (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_se_reduce): Conv2dStaticSamePadding(
      816, 34, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_se_expand): Conv2dStaticSamePadding(
      34, 816, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_project_conv): Conv2dStaticSamePadding(
      816, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_swish): MemoryEfficientSwish()
  )
  (17): MBConvBlock(
    (_expand_conv): Conv2dStaticSamePadding(
      136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_depthwise_conv): Conv2dStaticSamePadding(
      816, 816, kernel_size=(5, 5), stride=(1, 1), groups=816, bias=False
      (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
    )
    (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_se_reduce): Conv2dStaticSamePadding(
      816, 34, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_se_expand): Conv2dStaticSamePadding(
      34, 816, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_project_conv): Conv2dStaticSamePadding(
      816, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_swish): MemoryEfficientSwish()
  )
  (18): MBConvBlock(
    (_expand_conv): Conv2dStaticSamePadding(
      136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_depthwise_conv): Conv2dStaticSamePadding(
      816, 816, kernel_size=(5, 5), stride=[2, 2], groups=816, bias=False
      (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
    )
    (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_se_reduce): Conv2dStaticSamePadding(
      816, 34, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_se_expand): Conv2dStaticSamePadding(
      34, 816, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_project_conv): Conv2dStaticSamePadding(
      816, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
```

```
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (19): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (20): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (21): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
```

```
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (22): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (23): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
```

```
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (24): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(3, 3), stride=[1, 1], groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 384, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(384, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (25): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        384, 2304, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(2304, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        2304, 2304, kernel_size=(3, 3), stride=(1, 1), groups=2304, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(2304, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        2304, 96, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        96, 2304, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        2304, 384, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(384, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
  )
  (_conv_head): Conv2dStaticSamePadding(
    384, 1536, kernel_size=(1, 1), stride=(1, 1), bias=False
    (static_padding): Identity()
  )
  (_bn1): BatchNorm2d(1536, eps=0.001, momentum=0.010000000000000009, affine=True, track_
running_stats=True)
```

```
    (_avg_pooling): AdaptiveAvgPool2d(output_size=1)
    (_dropout): Dropout(p=0.3, inplace=False)
    (_fc): Linear(in_features=1536, out_features=1000, bias=True)
    (_swish): MemoryEfficientSwish()
  )
```

In [19]:

```
model._fc=nn.Linear(in_features=1536, out_features=42, bias=True)
```

In [20]:

```
model.cuda()
```

Out[20]:

```
EfficientNet(
  (_conv_stem): Conv2dStaticSamePadding(
    3, 40, kernel_size=(3, 3), stride=(2, 2), bias=False
    (static_padding): ZeroPad2d(padding=(0, 1, 0, 1), value=0.0)
  )
  (_bn0): BatchNorm2d(40, eps=0.001, momentum=0.010000000000000009, affine=True, track_ru
nning_stats=True)
  (_blocks): ModuleList(
    (0): MBConvBlock(
      (_depthwise_conv): Conv2dStaticSamePadding(
        40, 40, kernel_size=(3, 3), stride=[1, 1], groups=40, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(40, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        40, 10, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        10, 40, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        40, 24, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(24, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (1): MBConvBlock(
      (_depthwise_conv): Conv2dStaticSamePadding(
        24, 24, kernel_size=(3, 3), stride=(1, 1), groups=24, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(24, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        24, 6, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        6, 24, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        24, 24, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(24, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (2): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
```

```
        24, 144, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(144, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        144, 144, kernel_size=(3, 3), stride=[2, 2], groups=144, bias=False
        (static_padding): ZeroPad2d(padding=(0, 1, 0, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(144, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        144, 6, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        6, 144, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        144, 32, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(32, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (3): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        32, 192, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        192, 192, kernel_size=(3, 3), stride=(1, 1), groups=192, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        192, 8, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        8, 192, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        192, 32, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(32, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (4): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        32, 192, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        192, 192, kernel_size=(3, 3), stride=(1, 1), groups=192, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        192, 8, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
```

```
    )
    (_se_expand): Conv2dStaticSamePadding(
      8, 192, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_project_conv): Conv2dStaticSamePadding(
      192, 32, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn2): BatchNorm2d(32, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
    (_swish): MemoryEfficientSwish()
  )
  (5): MBConvBlock(
    (_expand_conv): Conv2dStaticSamePadding(
      32, 192, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn0): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_depthwise_conv): Conv2dStaticSamePadding(
      192, 192, kernel_size=(5, 5), stride=[2, 2], groups=192, bias=False
      (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
    )
    (_bn1): BatchNorm2d(192, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_se_reduce): Conv2dStaticSamePadding(
      192, 8, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_se_expand): Conv2dStaticSamePadding(
      8, 192, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_project_conv): Conv2dStaticSamePadding(
      192, 48, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn2): BatchNorm2d(48, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
    (_swish): MemoryEfficientSwish()
  )
  (6): MBConvBlock(
    (_expand_conv): Conv2dStaticSamePadding(
      48, 288, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn0): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_depthwise_conv): Conv2dStaticSamePadding(
      288, 288, kernel_size=(5, 5), stride=(1, 1), groups=288, bias=False
      (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
    )
    (_bn1): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
    (_se_reduce): Conv2dStaticSamePadding(
      288, 12, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_se_expand): Conv2dStaticSamePadding(
      12, 288, kernel_size=(1, 1), stride=(1, 1)
      (static_padding): Identity()
    )
    (_project_conv): Conv2dStaticSamePadding(
      288, 48, kernel_size=(1, 1), stride=(1, 1), bias=False
      (static_padding): Identity()
    )
    (_bn2): BatchNorm2d(48, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
    (_swish): MemoryEfficientSwish()
  )
  (7): MBConvBlock(
```

```
      (_expand_conv): Conv2dStaticSamePadding(
        48, 288, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        288, 288, kernel_size=(5, 5), stride=(1, 1), groups=288, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        288, 12, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        12, 288, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        288, 48, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(48, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (8): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        48, 288, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        288, 288, kernel_size=(3, 3), stride=[2, 2], groups=288, bias=False
        (static_padding): ZeroPad2d(padding=(0, 1, 0, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(288, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        288, 12, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        12, 288, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        288, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (9): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        576, 576, kernel_size=(3, 3), stride=(1, 1), groups=576, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        576, 24, kernel_size=(1, 1), stride=(1, 1)
```

```
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        24, 576, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        576, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (10): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        576, 576, kernel_size=(3, 3), stride=(1, 1), groups=576, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        576, 24, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        24, 576, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        576, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (11): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        576, 576, kernel_size=(3, 3), stride=(1, 1), groups=576, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        576, 24, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        24, 576, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        576, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
```

```
    (12): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        576, 576, kernel_size=(3, 3), stride=(1, 1), groups=576, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        576, 24, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        24, 576, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        576, 96, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(96, eps=0.001, momentum=0.010000000000000009, affine=True, trac
k_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (13): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        96, 576, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        576, 576, kernel_size=(5, 5), stride=[1, 1], groups=576, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(576, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        576, 24, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        24, 576, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        576, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (14): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        816, 816, kernel_size=(5, 5), stride=(1, 1), groups=816, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
```

```
        816, 34, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        34, 816, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        816, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (15): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        816, 816, kernel_size=(5, 5), stride=(1, 1), groups=816, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        816, 34, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        34, 816, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        816, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (16): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        816, 816, kernel_size=(5, 5), stride=(1, 1), groups=816, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        816, 34, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        34, 816, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        816, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
```

```
    )
    (17): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        816, 816, kernel_size=(5, 5), stride=(1, 1), groups=816, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        816, 34, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        34, 816, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        816, 136, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(136, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (18): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        136, 816, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        816, 816, kernel_size=(5, 5), stride=[2, 2], groups=816, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(816, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        816, 34, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        34, 816, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        816, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (19): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
```

```
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (20): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (21): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
```

```
      (_swish): MemoryEfficientSwish()
    )
    (22): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (23): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(5, 5), stride=(1, 1), groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(2, 2, 2, 2), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 232, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(232, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (24): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        232, 1392, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        1392, 1392, kernel_size=(3, 3), stride=[1, 1], groups=1392, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(1392, eps=0.001, momentum=0.010000000000000009, affine=True, tr
```

```
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        1392, 58, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        58, 1392, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        1392, 384, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(384, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
    (25): MBConvBlock(
      (_expand_conv): Conv2dStaticSamePadding(
        384, 2304, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn0): BatchNorm2d(2304, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_depthwise_conv): Conv2dStaticSamePadding(
        2304, 2304, kernel_size=(3, 3), stride=(1, 1), groups=2304, bias=False
        (static_padding): ZeroPad2d(padding=(1, 1, 1, 1), value=0.0)
      )
      (_bn1): BatchNorm2d(2304, eps=0.001, momentum=0.010000000000000009, affine=True, tr
ack_running_stats=True)
      (_se_reduce): Conv2dStaticSamePadding(
        2304, 96, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_se_expand): Conv2dStaticSamePadding(
        96, 2304, kernel_size=(1, 1), stride=(1, 1)
        (static_padding): Identity()
      )
      (_project_conv): Conv2dStaticSamePadding(
        2304, 384, kernel_size=(1, 1), stride=(1, 1), bias=False
        (static_padding): Identity()
      )
      (_bn2): BatchNorm2d(384, eps=0.001, momentum=0.010000000000000009, affine=True, tra
ck_running_stats=True)
      (_swish): MemoryEfficientSwish()
    )
  )
  (_conv_head): Conv2dStaticSamePadding(
    384, 1536, kernel_size=(1, 1), stride=(1, 1), bias=False
    (static_padding): Identity()
  )
  (_bn1): BatchNorm2d(1536, eps=0.001, momentum=0.010000000000000009, affine=True, track_
running_stats=True)
  (_avg_pooling): AdaptiveAvgPool2d(output_size=1)
  (_dropout): Dropout(p=0.3, inplace=False)
  (_fc): Linear(in_features=1536, out_features=42, bias=True)
  (_swish): MemoryEfficientSwish()
)
```

In [21]:

```python
loss_fn=nn.CrossEntropyLoss()
```

In [22]:

```python
opt = torch.optim.Adam(model.parameters(),lr=0.01)
```

In [23]:

```python
exp_lr_scheduler=torch.optim.lr_scheduler.StepLR(opt, step_size=7, gamma=0.1)
```

```
In [24]:
```

```
history=train(train_dataset, val_dataset, model=model, epochs=10, batch_size=64)
```

```
epoch:   0%|          | 0/10 [00:00<?, ?it/s]
```

loss 0.5146420984896469

```
epoch:  10%|█         | 1/10 [07:53<1:10:57, 473.11s/it]
```

Epoch 001 train_loss: 0.5146   val_loss 0.2348 train_acc 0.8799 val_acc 0.9392
loss 0.3829744393986588

```
epoch:  20%|██        | 2/10 [12:39<48:25, 363.23s/it]
```

Epoch 002 train_loss: 0.3830   val_loss 0.2176 train_acc 0.9044 val_acc 0.9421
loss 0.11727185271671693

```
epoch:  30%|███       | 3/10 [17:20<38:01, 325.86s/it]
```

Epoch 003 train_loss: 0.1173   val_loss 0.2170 train_acc 0.9722 val_acc 0.9532
loss 0.09874495174335943

```
epoch:  40%|████      | 4/10 [22:01<30:47, 307.98s/it]
```

Epoch 004 train_loss: 0.0987   val_loss 0.2163 train_acc 0.9729 val_acc 0.9492
loss 0.13552717879605344

```
epoch:  50%|█████     | 5/10 [26:41<24:50, 298.03s/it]
```

Epoch 005 train_loss: 0.1355   val_loss 0.4590 train_acc 0.9659 val_acc 0.9083
loss 0.13686933817192545

```
epoch:  60%|██████    | 6/10 [31:23<19:29, 292.38s/it]
```

Epoch 006 train_loss: 0.1369   val_loss 0.1855 train_acc 0.9646 val_acc 0.9591
loss 0.07739384250458028

```
epoch:  70%|███████   | 7/10 [36:20<14:41, 293.90s/it]
```

Epoch 007 train_loss: 0.0774   val_loss 0.1923 train_acc 0.9808 val_acc 0.9543
loss 0.05691597149921239

```
epoch:  80%|████████  | 8/10 [41:19<09:50, 295.47s/it]
```

Epoch 008 train_loss: 0.0569   val_loss 0.2237 train_acc 0.9839 val_acc 0.9442
loss 0.0909518531781525

```
epoch:  90%|█████████ | 9/10 [46:20<04:57, 297.29s/it]
```

Epoch 009 train_loss: 0.0910   val_loss 0.2265 train_acc 0.9763 val_acc 0.9541
loss 0.1098189698175095

```
epoch: 100%|██████████| 10/10 [51:20<00:00, 308.00s/it]
```

Epoch 010 train_loss: 0.1098   val_loss 0.1894 train_acc 0.9734 val_acc 0.9645

```
In [25]:
```

```
history=train(val_dataset, train_dataset, model=model, epochs=5, batch_size=64)
```

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

loss 0.12040306932737184

```
epoch:  20%|██        | 1/5 [04:59<19:58, 299.63s/it]
```

Epoch 001 train_loss: 0.1204   val_loss 0.2018 train_acc 0.9701 val_acc 0.9515
loss 0.09255204745670668

```
epoch:  40%|████      | 2/5 [10:00<15:01, 300.49s/it]
```

Epoch 002 train_loss: 0.0926   val_loss 0.2809 train_acc 0.9786 val_acc 0.9364
loss 0.08480644948992679

```
epoch:  60%|██████    | 3/5 [15:01<10:01, 300.68s/it]
```

Epoch 003 train_loss: 0.0848   val_loss 0.2355 train_acc 0.9777 val_acc 0.9540

```
Epoch 003 train_loss: 0.0848      val_loss 0.2355 train_acc 0.9777 val_acc 0.9540
loss 0.10535751649322195
```

```
Epoch 004 train_loss: 0.1054      val_loss 0.1949 train_acc 0.9743 val_acc 0.9551
loss 0.07838044400811345
```
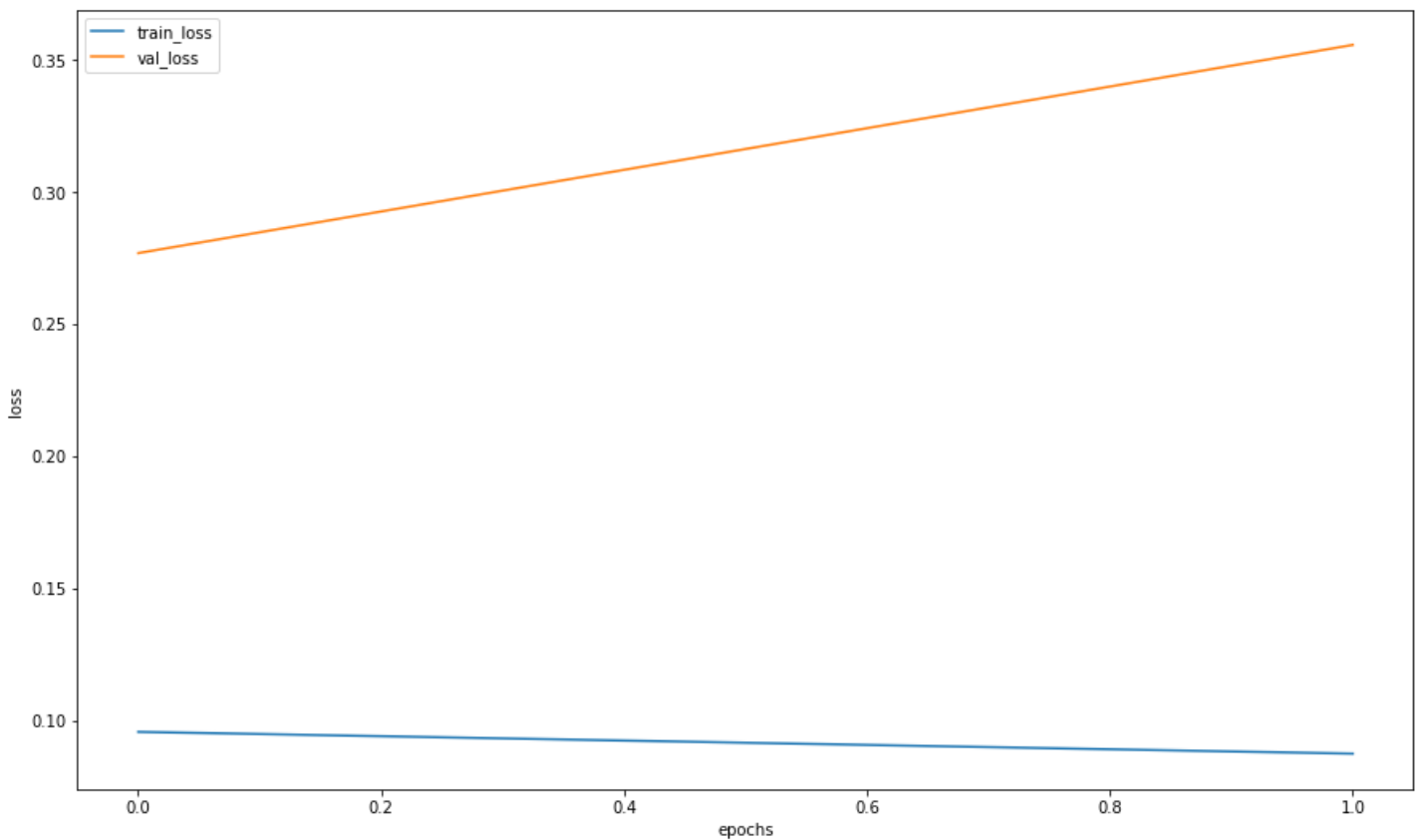
```
Epoch 005 train_loss: 0.0784      val_loss 0.1700 train_acc 0.9808 val_acc 0.9654
```

In [27]:

```python
loss, acc, val_loss, val_acc = zip(*history)
```

In [28]:

```python
plt.figure(figsize=(15, 9))
plt.plot(loss, label="train_loss")
plt.plot(val_loss, label="val_loss")
plt.legend(loc='best')
plt.xlabel("epochs")
plt.ylabel("loss")
plt.show()
```


In [29]:

```python
label_encoder = pickle.load(open("label_encoder.pkl", 'rb'))
test_dataset = SimpsonsDataset(test_files, mode="test")
test_loader = DataLoader(test_dataset, shuffle=False, batch_size=64)
probs = predict(model, test_loader)


preds = label_encoder.inverse_transform(np.argmax(probs, axis=1))
test_filenames = [path.name for path in test_dataset.files]
```

In [30]:

```python
my_submit = pd.DataFrame({'Id': test_filenames, 'Expected': preds})
my_submit.to_csv('/kaggle/working/submit.csv', index=False)
my_submit.head()
```

Out[30]:

	Id	Expected
0	**img0.jpg**	**nelson_muntz**
1	**img1.jpg**	**bart_simpson**
2	**img10.jpg**	**ned_flanders**
3	**img100.jpg**	**chief_wiggum**
4	**img101.jpg**	**apu_nahasapeemapetilon**

In [31]:

```python
import os
os.chdir('/kaggle/working')
```

In [32]:

```python
my_submit.to_csv('submit.csv')
```

In [7]:

```python
from IPython.display import FileLink
FileLink('submit.csv')
```

Out[7]:

submit.csv

In [5]:

```python
# score на Kaggle-98.405%
# Team Name: Владимир_Терентьев_341670400
# выводимый файл имеет score 96.386(потому что делал на Kaggle и при повторном обучении результат изменился,
# хотя код тот же(не ставид seed))
```

In [ ]: