

```
pip install torchviz
```

```
Collecting torchviz
```

```
  Downloading torchviz-0.0.2.tar.gz (4.9 kB)
```

```
  Preparing metadata (setup.py) ... ent already satisfied: torch in  
/usr/local/lib/python3.10/dist-packages (from torchviz) (2.1.0+cu118)
```

```
Requirement already satisfied: graphviz in
```

```
/usr/local/lib/python3.10/dist-packages (from torchviz) (0.20.1)
```

```
Requirement already satisfied: filelock in
```

```
/usr/local/lib/python3.10/dist-packages (from torch->torchviz)  
(3.13.1)
```

```
Requirement already satisfied: typing-extensions in
```

```
/usr/local/lib/python3.10/dist-packages (from torch->torchviz) (4.5.0)
```

```
Requirement already satisfied: sympy in
```

```
/usr/local/lib/python3.10/dist-packages (from torch->torchviz) (1.12)
```

```
Requirement already satisfied: networkx in
```

```
/usr/local/lib/python3.10/dist-packages (from torch->torchviz) (3.2.1)
```

```
Requirement already satisfied: jinja2 in
```

```
/usr/local/lib/python3.10/dist-packages (from torch->torchviz) (3.1.2)
```

```
Requirement already satisfied: fsspec in
```

```
/usr/local/lib/python3.10/dist-packages (from torch->torchviz)  
(2023.6.0)
```

```
Requirement already satisfied: triton==2.1.0 in
```

```
/usr/local/lib/python3.10/dist-packages (from torch->torchviz) (2.1.0)
```

```
Requirement already satisfied: MarkupSafe>=2.0 in
```

```
/usr/local/lib/python3.10/dist-packages (from jinja2->torch->torchviz)  
(2.1.3)
```

```
Requirement already satisfied: mpmath>=0.19 in
```

```
/usr/local/lib/python3.10/dist-packages (from sympy->torch->torchviz)  
(1.3.0)
```

```
Building wheels for collected packages: torchviz
```

```
  Building wheel for torchviz (setup.py) ... e=torchviz-0.0.2-py3-  
none-any.whl size=4133
```

```
sha256=b8d750eb8290d20c86e90cb55bc60b83bc41974fa622440a0e960df35c3acd4  
4
```

```
  Stored in directory:
```

```
/root/.cache/pip/wheels/4c/97/88/a02973217949e0db0c9f4346d154085f4725f  
99c4f15a87094
```

```
Successfully built torchviz
```

```
Installing collected packages: torchviz
```

```
Successfully installed torchviz-0.0.2
```

```
import multiprocessing as mp  
import time  
from typing import Any, Generator, Literal  
import os  
import torch  
import torchvision  
import torchvision.datasets  
import torchvision.transforms
```

```

import torch.utils.data
import torch.nn

from torchvision import models as torch_model
import seaborn as sns
from torchviz import make_dot
from IPython.display import display
import time
from torchsummary import summary
from matplotlib import pyplot as plot

device = torch.device("cuda:0" if torch.cuda.is_available() else
"cpu")
device

device(type='cuda', index=0)

CLASSES = ['plane', 'car', 'bird', 'cat', 'deer', 'dog', 'frog',
'horse', 'ship', 'truck']
N_CLASSES = len(CLASSES)

def show_images(images, title):
    num_showed_imgs_x = 10
    num_showed_imgs_y = 10

    figsize = (10, 10)
    fig, axes = plot.subplots(num_showed_imgs_y, num_showed_imgs_x,
figsize = figsize)
    fig.suptitle(title)
    plot.setp(plot.gcf().get_axes(), xticks = [], yticks = [])
    dataiter = iter(images)
    for i, ax in enumerate(axes.flat):
        img = images[i][0].numpy().transpose(1, 2, 0)
        ax.imshow(img)
        ax.text(1, 1, CLASSES[images[i][1]], bbox=dict(fill=False,
edgecolor='red', linewidth=2))

def download_data(transforms, batch_size):
    dir_name = os.getcwd()

    train_dataset = torchvision.datasets.CIFAR10(
        root = dir_name, train = True, download = True,
        transform = transforms
    )
    test_dataset = torchvision.datasets.CIFAR10(
        root = dir_name, train = False, download = True,
        transform = transforms
    )

    print('Number of train samples: {}'.format(len(train_dataset)))
    show_images(train_dataset, 'Train samples')

```

```

print('Number of test samples: {}'.format(len(test_dataset)))
show_images(test_dataset, 'Test samples')

train_data_loader = torch.utils.data.DataLoader(
    train_dataset, batch_size = batch_size, shuffle = True
)
test_data_loader = torch.utils.data.DataLoader(
    test_dataset, batch_size = batch_size, shuffle = False
)
return train_data_loader, test_data_loader

num_epochs = 20

def get_accuracy(data_loader, model):
    tp = 0
    n = 0
    with torch.no_grad():
        for images, labels in data_loader:
            labels = labels.to(device)
            images = images.to(device)
            outputs = model(images)
            _, predicted = torch.max(outputs.data, 1)
            n += labels.size(0)
            tp += (predicted == labels).sum()
    return tp / n

def train_loop(model, train_dataloader, optimizer, loss_function):
    start_all = time.time()
    for epoch in range(num_epochs):
        start = time.time()
        for images, labels in train_dataloader:
            images = images.to(device)
            labels = labels.to(device)

            outputs = model(images)
            loss = loss_function(outputs.type(torch.float32),
torch.nn.functional.one_hot(labels,
num_classes=10).type(torch.float32))

            optimizer.zero_grad()
            loss.backward()
            optimizer.step()
        end = time.time()
        print('Epoch[{}]: accuracy = {}, time = {}'.format(epoch,
get_accuracy(train_dataloader, model), (end - start)))
        end_all = time.time()
        print('train time = {}'.format((end_all - start_all)))
    # yield model

```

```
def test_model(model, test_dataloader):
    model.eval()
    print('Test accuracy: {}'.format(get_accuracy(test_dataloader,
model)))
    # yield model
```

Modifications:

```
def last_layer(in_features):
    return torch.nn.Linear(in_features, 10)

def last_sequential_layer(in_features):
    return torch.nn.Sequential(
        torch.nn.Linear(in_features, in_features//2),
        torch.nn.ReLU(),
        torch.nn.Linear(in_features//2, 10),
    )

def init_weights(model):
    print(model)
    if type(model) == torch.nn.Linear:
        torch.nn.init.xavier_uniform_(model.weight)
    print(model.weight)
```

1. ResNet18
1. Single layer

```
model = torch_model.resnet18(torch_model.ResNet18_Weights.DEFAULT)

/usr/local/lib/python3.10/dist-packages/torchvision/models/_utils.py:135: UserWarning: Using 'weights' as positional parameter(s)
is deprecated since 0.13 and may be removed in the future. Please use
keyword parameter(s) instead.
  warnings.warn(
```

```
# Информация об архитектуре и визуализация сети
print("Модель \"CNN\":")
summary(model.to(device), input_size=(3, 32, 32))
```

```
print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters())))
```

Модель "CNN":

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 64, 16, 16]	9,408
BatchNorm2d-2	[-1, 64, 16, 16]	128
ReLU-3	[-1, 64, 16, 16]	0

MaxPool2d-4	[-1, 64, 8, 8]	0
Conv2d-5	[-1, 64, 8, 8]	36,864
BatchNorm2d-6	[-1, 64, 8, 8]	128
ReLU-7	[-1, 64, 8, 8]	0
Conv2d-8	[-1, 64, 8, 8]	36,864
BatchNorm2d-9	[-1, 64, 8, 8]	128
ReLU-10	[-1, 64, 8, 8]	0
BasicBlock-11	[-1, 64, 8, 8]	0
Conv2d-12	[-1, 64, 8, 8]	36,864
BatchNorm2d-13	[-1, 64, 8, 8]	128
ReLU-14	[-1, 64, 8, 8]	0
Conv2d-15	[-1, 64, 8, 8]	36,864
BatchNorm2d-16	[-1, 64, 8, 8]	128
ReLU-17	[-1, 64, 8, 8]	0
BasicBlock-18	[-1, 64, 8, 8]	0
Conv2d-19	[-1, 128, 4, 4]	73,728
BatchNorm2d-20	[-1, 128, 4, 4]	256
ReLU-21	[-1, 128, 4, 4]	0
Conv2d-22	[-1, 128, 4, 4]	147,456
BatchNorm2d-23	[-1, 128, 4, 4]	256
Conv2d-24	[-1, 128, 4, 4]	8,192
BatchNorm2d-25	[-1, 128, 4, 4]	256
ReLU-26	[-1, 128, 4, 4]	0
BasicBlock-27	[-1, 128, 4, 4]	0
Conv2d-28	[-1, 128, 4, 4]	147,456
BatchNorm2d-29	[-1, 128, 4, 4]	256
ReLU-30	[-1, 128, 4, 4]	0
Conv2d-31	[-1, 128, 4, 4]	147,456
BatchNorm2d-32	[-1, 128, 4, 4]	256
ReLU-33	[-1, 128, 4, 4]	0
BasicBlock-34	[-1, 128, 4, 4]	0
Conv2d-35	[-1, 256, 2, 2]	294,912
BatchNorm2d-36	[-1, 256, 2, 2]	512
ReLU-37	[-1, 256, 2, 2]	0
Conv2d-38	[-1, 256, 2, 2]	589,824
BatchNorm2d-39	[-1, 256, 2, 2]	512
Conv2d-40	[-1, 256, 2, 2]	32,768
BatchNorm2d-41	[-1, 256, 2, 2]	512
ReLU-42	[-1, 256, 2, 2]	0
BasicBlock-43	[-1, 256, 2, 2]	0
Conv2d-44	[-1, 256, 2, 2]	589,824
BatchNorm2d-45	[-1, 256, 2, 2]	512
ReLU-46	[-1, 256, 2, 2]	0
Conv2d-47	[-1, 256, 2, 2]	589,824
BatchNorm2d-48	[-1, 256, 2, 2]	512
ReLU-49	[-1, 256, 2, 2]	0
BasicBlock-50	[-1, 256, 2, 2]	0
Conv2d-51	[-1, 512, 1, 1]	1,179,648
BatchNorm2d-52	[-1, 512, 1, 1]	1,024

ReLU-53	[-1, 512, 1, 1]	0
Conv2d-54	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-55	[-1, 512, 1, 1]	1,024
Conv2d-56	[-1, 512, 1, 1]	131,072
BatchNorm2d-57	[-1, 512, 1, 1]	1,024
ReLU-58	[-1, 512, 1, 1]	0
BasicBlock-59	[-1, 512, 1, 1]	0
Conv2d-60	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-61	[-1, 512, 1, 1]	1,024
ReLU-62	[-1, 512, 1, 1]	0
Conv2d-63	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-64	[-1, 512, 1, 1]	1,024
ReLU-65	[-1, 512, 1, 1]	0
BasicBlock-66	[-1, 512, 1, 1]	0
AdaptiveAvgPool2d-67	[-1, 512, 1, 1]	0
Linear-68	[-1, 1000]	513,000

```

=====
Total params: 11,689,512
Trainable params: 11,689,512
Non-trainable params: 0
-----

```

```

Input size (MB): 0.01
Forward/backward pass size (MB): 1.29
Params size (MB): 44.59
Estimated Total Size (MB): 45.90
-----

```

Визуализация модели:

[illegible]

[illegible]

[illegible]

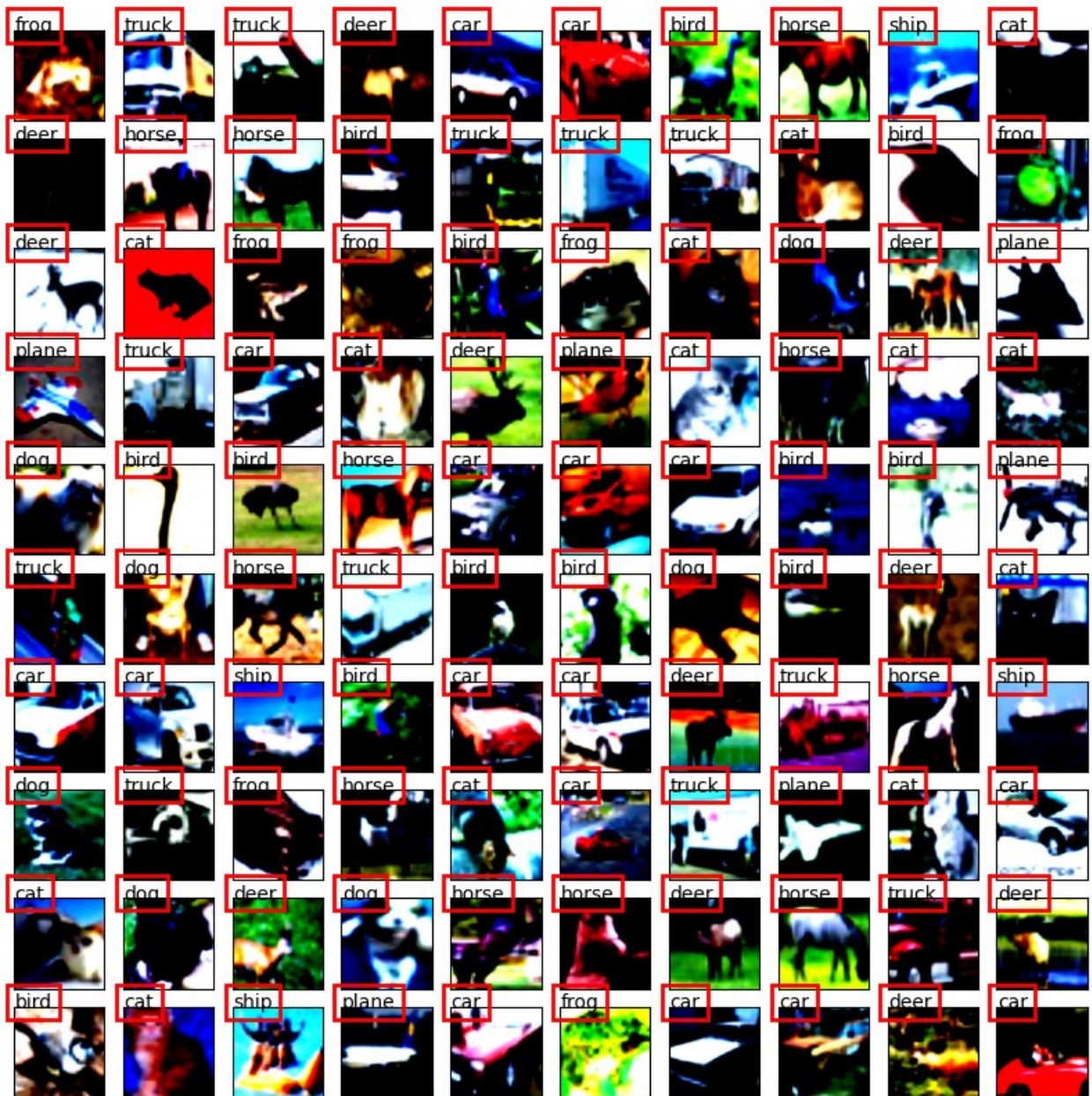
[illegible]

[illegible]

[illegible]

[illegible]

Train samples



Test samples



```
learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.fc.parameters(), lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)

Epoch[0]: accuracy = 0.7552199959754944, time = 134.36429405212402
Epoch[1]: accuracy = 0.7705399990081787, time = 133.38727378845215
```



```
Epoch[2]: accuracy = 0.7748799920082092, time = 134.6206178665161
Epoch[3]: accuracy = 0.7821799516677856, time = 133.35160636901855
Epoch[4]: accuracy = 0.7877799868583679, time = 132.4030101299286
Epoch[5]: accuracy = 0.7893799543380737, time = 130.85477805137634
Epoch[6]: accuracy = 0.7927199602127075, time = 132.90481090545654
Epoch[7]: accuracy = 0.7902599573135376, time = 133.37981700897217
Epoch[8]: accuracy = 0.792739987373352, time = 138.0912709236145
Epoch[9]: accuracy = 0.7948399782180786, time = 140.2856011390686
Epoch[10]: accuracy = 0.7905199527740479, time = 133.56772112846375
Epoch[11]: accuracy = 0.7960399985313416, time = 132.52500438690186
Epoch[12]: accuracy = 0.7979399561882019, time = 131.02950859069824
Epoch[13]: accuracy = 0.799519956111908, time = 132.790180683136
Epoch[14]: accuracy = 0.800279974937439, time = 132.8993923664093
Epoch[15]: accuracy = 0.800059974193573, time = 133.1646978855133
Epoch[16]: accuracy = 0.8001199960708618, time = 133.71345162391663
Epoch[17]: accuracy = 0.7939800024032593, time = 133.75645112991333
Epoch[18]: accuracy = 0.8032400012016296, time = 132.00247764587402
Epoch[19]: accuracy = 0.8007599711418152, time = 130.65841007232666
train time = 5328.962197542191
```

```
test_model(model, test_dataloader)
```

```
Test accuracy: 0.7849000096321106
```

```
print(model)
```

```
ResNet(
  (conv1): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2),
padding=(3, 3), bias=False)
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (relu): ReLU(inplace=True)
  (maxpool): MaxPool2d(kernel_size=3, stride=2, padding=1, dilation=1,
ceil_mode=False)
  (layer1): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
    (1): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
```

```

track_running_stats=True)
    (relu): ReLU(inplace=True)
    (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
    )
    (layer2): Sequential(
    (0): BasicBlock(
        (conv1): Conv2d(64, 128, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (downsample): Sequential(
            (0): Conv2d(64, 128, kernel_size=(1, 1), stride=(2, 2),
bias=False)
            (1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        )
    )
    (1): BasicBlock(
        (conv1): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
    )
    (layer3): Sequential(
    (0): BasicBlock(
        (conv1): Conv2d(128, 256, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
    )

```

```

        (downsample): Sequential(
          (0): Conv2d(128, 256, kernel_size=(1, 1), stride=(2, 2),
bias=False)
          (1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        )
      )
      (1): BasicBlock(
        (conv1): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (layer4): Sequential(
      (0): BasicBlock(
        (conv1): Conv2d(256, 512, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (downsample): Sequential(
          (0): Conv2d(256, 512, kernel_size=(1, 1), stride=(2, 2),
bias=False)
          (1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        )
      )
      (1): BasicBlock(
        (conv1): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
  )
)

```

```
(avgpool): AdaptiveAvgPool2d(output_size=(1, 1))
(fc): Linear(in_features=512, out_features=10, bias=True)
)
```

1. Sequential layer

```
model = torch_model.resnet18(torch_model.ResNet18_Weights.DEFAULT)

/usr/local/lib/python3.10/dist-packages/torchvision/models/_utils.py:135: UserWarning: Using 'weights' as positional parameter(s) is deprecated since 0.13 and may be removed in the future. Please use keyword parameter(s) instead.
  warnings.warn(

# Информация об архитектуре и визуализация сети
print("Модель \"CNN\":")
summary(model.to(device), input_size=(3, 32, 32))

print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters()))))
```

Модель "CNN":

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 64, 16, 16]	9,408
BatchNorm2d-2	[-1, 64, 16, 16]	128
ReLU-3	[-1, 64, 16, 16]	0
MaxPool2d-4	[-1, 64, 8, 8]	0
Conv2d-5	[-1, 64, 8, 8]	36,864
BatchNorm2d-6	[-1, 64, 8, 8]	128
ReLU-7	[-1, 64, 8, 8]	0
Conv2d-8	[-1, 64, 8, 8]	36,864
BatchNorm2d-9	[-1, 64, 8, 8]	128
ReLU-10	[-1, 64, 8, 8]	0
BasicBlock-11	[-1, 64, 8, 8]	0
Conv2d-12	[-1, 64, 8, 8]	36,864
BatchNorm2d-13	[-1, 64, 8, 8]	128
ReLU-14	[-1, 64, 8, 8]	0
Conv2d-15	[-1, 64, 8, 8]	36,864
BatchNorm2d-16	[-1, 64, 8, 8]	128
ReLU-17	[-1, 64, 8, 8]	0
BasicBlock-18	[-1, 64, 8, 8]	0
Conv2d-19	[-1, 128, 4, 4]	73,728
BatchNorm2d-20	[-1, 128, 4, 4]	256
ReLU-21	[-1, 128, 4, 4]	0
Conv2d-22	[-1, 128, 4, 4]	147,456
BatchNorm2d-23	[-1, 128, 4, 4]	256
Conv2d-24	[-1, 128, 4, 4]	8,192

BatchNorm2d-25	[-1, 128, 4, 4]	256
ReLU-26	[-1, 128, 4, 4]	0
BasicBlock-27	[-1, 128, 4, 4]	0
Conv2d-28	[-1, 128, 4, 4]	147,456
BatchNorm2d-29	[-1, 128, 4, 4]	256
ReLU-30	[-1, 128, 4, 4]	0
Conv2d-31	[-1, 128, 4, 4]	147,456
BatchNorm2d-32	[-1, 128, 4, 4]	256
ReLU-33	[-1, 128, 4, 4]	0
BasicBlock-34	[-1, 128, 4, 4]	0
Conv2d-35	[-1, 256, 2, 2]	294,912
BatchNorm2d-36	[-1, 256, 2, 2]	512
ReLU-37	[-1, 256, 2, 2]	0
Conv2d-38	[-1, 256, 2, 2]	589,824
BatchNorm2d-39	[-1, 256, 2, 2]	512
Conv2d-40	[-1, 256, 2, 2]	32,768
BatchNorm2d-41	[-1, 256, 2, 2]	512
ReLU-42	[-1, 256, 2, 2]	0
BasicBlock-43	[-1, 256, 2, 2]	0
Conv2d-44	[-1, 256, 2, 2]	589,824
BatchNorm2d-45	[-1, 256, 2, 2]	512
ReLU-46	[-1, 256, 2, 2]	0
Conv2d-47	[-1, 256, 2, 2]	589,824
BatchNorm2d-48	[-1, 256, 2, 2]	512
ReLU-49	[-1, 256, 2, 2]	0
BasicBlock-50	[-1, 256, 2, 2]	0
Conv2d-51	[-1, 512, 1, 1]	1,179,648
BatchNorm2d-52	[-1, 512, 1, 1]	1,024
ReLU-53	[-1, 512, 1, 1]	0
Conv2d-54	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-55	[-1, 512, 1, 1]	1,024
Conv2d-56	[-1, 512, 1, 1]	131,072
BatchNorm2d-57	[-1, 512, 1, 1]	1,024
ReLU-58	[-1, 512, 1, 1]	0
BasicBlock-59	[-1, 512, 1, 1]	0
Conv2d-60	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-61	[-1, 512, 1, 1]	1,024
ReLU-62	[-1, 512, 1, 1]	0
Conv2d-63	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-64	[-1, 512, 1, 1]	1,024
ReLU-65	[-1, 512, 1, 1]	0
BasicBlock-66	[-1, 512, 1, 1]	0
AdaptiveAvgPool2d-67	[-1, 512, 1, 1]	0
Linear-68	[-1, 1000]	513,000

```

=====
Total params: 11,689,512
Trainable params: 11,689,512
Non-trainable params: 0
-----

```

Input size (MB): 0.01
Forward/backward pass size (MB): 1.29
Params size (MB): 44.59
Estimated Total Size (MB): 45.90

Визуализация модели:


```

for param in model.parameters():
    param.requires_grad = False
model.fc = last_sequential_layer(512)
model = model.to(device)

print(model)

ResNet(
  (conv1): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2),
padding=(3, 3), bias=False)
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (relu): ReLU(inplace=True)
  (maxpool): MaxPool2d(kernel_size=3, stride=2, padding=1, dilation=1,
ceil_mode=False)
  (layer1): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
    (1): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
  )
  (layer2): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(64, 128, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
  )
)

```

```

        (downsample): Sequential(
          (0): Conv2d(64, 128, kernel_size=(1, 1), stride=(2, 2),
bias=False)
          (1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        )
      )
      (1): BasicBlock(
        (conv1): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
    (layer3): Sequential(
      (0): BasicBlock(
        (conv1): Conv2d(128, 256, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (downsample): Sequential(
          (0): Conv2d(128, 256, kernel_size=(1, 1), stride=(2, 2),
bias=False)
          (1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        )
      )
      (1): BasicBlock(
        (conv1): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      )
    )
  )
)

```

```

(layer4): Sequential(
  (0): BasicBlock(
    (conv1): Conv2d(256, 512, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1), bias=False)
    (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (relu): ReLU(inplace=True)
    (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (downsample): Sequential(
      (0): Conv2d(256, 512, kernel_size=(1, 1), stride=(2, 2),
bias=False)
      (1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
  )
  (1): BasicBlock(
    (conv1): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (relu): ReLU(inplace=True)
    (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  )
)
(avgpool): AdaptiveAvgPool2d(output_size=(1, 1))
(fc): Sequential(
  (0): Linear(in_features=512, out_features=256, bias=True)
  (1): ReLU()
  (2): Linear(in_features=256, out_features=10, bias=True)
)
)

```

```

transforms = torch_model.ResNet18_Weights.DEFAULT.transforms()
batch_size = 128
train_dataloader, test_dataloader = download_data(transforms,
batch_size)

```

```

Files already downloaded and verified
Files already downloaded and verified
Number of train samples: 50000

```

```

WARNING:matplotlib.image:Clipping input data to the valid range for
imshow with RGB data ([0..1] for floats or [0..255] for integers).
WARNING:matplotlib.image:Clipping input data to the valid range for

```

[illegible]

[illegible]

[illegible]

[illegible]

WARNING:matplotlib.image:Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

```
Number of test samples: 10000
```

[illegible]

[illegible]

[illegible]

[illegible]

WARNING:matplotlib.image:Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).
WARNING:matplotlib.image:Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).
WARNING:matplotlib.image:Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).
WARNING:matplotlib.image:Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..255] for integers).

Train samples



Test samples



```
learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.fc.parameters(), lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)

Epoch[0]: accuracy = 0.7584599852561951, time = 133.47190737724304
Epoch[1]: accuracy = 0.7836199998855591, time = 134.56931257247925
```

```
Epoch[2]: accuracy = 0.7912200093269348, time = 131.76579809188843
Epoch[3]: accuracy = 0.7971199750900269, time = 132.62603068351746
Epoch[4]: accuracy = 0.7963399887084961, time = 132.72239804267883
Epoch[5]: accuracy = 0.8087999820709229, time = 132.0856158733368
Epoch[6]: accuracy = 0.817799985408783, time = 133.8889524936676
Epoch[7]: accuracy = 0.8195199966430664, time = 133.67858791351318
Epoch[8]: accuracy = 0.8260399699211121, time = 134.11023926734924
Epoch[9]: accuracy = 0.8386799693107605, time = 134.24782919883728
Epoch[10]: accuracy = 0.8505199551582336, time = 133.9661843776703
Epoch[11]: accuracy = 0.8570999503135681, time = 133.6306209564209
Epoch[12]: accuracy = 0.8524199724197388, time = 136.24574303627014
Epoch[13]: accuracy = 0.8709399700164795, time = 133.86288475990295
Epoch[14]: accuracy = 0.8695600032806396, time = 132.84558701515198
Epoch[15]: accuracy = 0.870959997177124, time = 134.21223258972168
Epoch[16]: accuracy = 0.8806599974632263, time = 133.45714831352234
Epoch[17]: accuracy = 0.878879964351654, time = 133.20917320251465
Epoch[18]: accuracy = 0.8874199986457825, time = 133.56253910064697
Epoch[19]: accuracy = 0.8967399597167969, time = 134.24799370765686
train time = 5326.537646770477
```

```
test_model(model, test_dataloader)
```

```
Test accuracy: 0.7937999963760376
```

1. Weights

```
model = torch_model.resnet18(torch_model.ResNet18_Weights.DEFAULT)
```

```
/usr/local/lib/python3.10/dist-packages/torchvision/models/_utils.py:135: UserWarning: Using 'weights' as positional parameter(s) is deprecated since 0.13 and may be removed in the future. Please use keyword parameter(s) instead.
```

```
warnings.warn(
```

```
Downloading: "https://download.pytorch.org/models/resnet18-f37072fd.pth" to /root/.cache/torch/hub/checkpoints/resnet18-f37072fd.pth
```

```
100%|██████████| 44.7M/44.7M [00:00<00:00, 85.8MB/s]
```

```
# Информация об архитектуре и визуализация сети
```

```
print("Модель \"CNN\":")
```

```
summary(model.to(device), input_size=(3, 32, 32))
```

```
print("Визуализация модели:")
```

```
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters())))
```

```
Модель "CNN":
```

Layer (type)	Output Shape	Param #
Conv2d-1	[-1, 64, 16, 16]	9,408

BatchNorm2d-2	[-1, 64, 16, 16]	128
ReLU-3	[-1, 64, 16, 16]	0
MaxPool2d-4	[-1, 64, 8, 8]	0
Conv2d-5	[-1, 64, 8, 8]	36,864
BatchNorm2d-6	[-1, 64, 8, 8]	128
ReLU-7	[-1, 64, 8, 8]	0
Conv2d-8	[-1, 64, 8, 8]	36,864
BatchNorm2d-9	[-1, 64, 8, 8]	128
ReLU-10	[-1, 64, 8, 8]	0
BasicBlock-11	[-1, 64, 8, 8]	0
Conv2d-12	[-1, 64, 8, 8]	36,864
BatchNorm2d-13	[-1, 64, 8, 8]	128
ReLU-14	[-1, 64, 8, 8]	0
Conv2d-15	[-1, 64, 8, 8]	36,864
BatchNorm2d-16	[-1, 64, 8, 8]	128
ReLU-17	[-1, 64, 8, 8]	0
BasicBlock-18	[-1, 64, 8, 8]	0
Conv2d-19	[-1, 128, 4, 4]	73,728
BatchNorm2d-20	[-1, 128, 4, 4]	256
ReLU-21	[-1, 128, 4, 4]	0
Conv2d-22	[-1, 128, 4, 4]	147,456
BatchNorm2d-23	[-1, 128, 4, 4]	256
Conv2d-24	[-1, 128, 4, 4]	8,192
BatchNorm2d-25	[-1, 128, 4, 4]	256
ReLU-26	[-1, 128, 4, 4]	0
BasicBlock-27	[-1, 128, 4, 4]	0
Conv2d-28	[-1, 128, 4, 4]	147,456
BatchNorm2d-29	[-1, 128, 4, 4]	256
ReLU-30	[-1, 128, 4, 4]	0
Conv2d-31	[-1, 128, 4, 4]	147,456
BatchNorm2d-32	[-1, 128, 4, 4]	256
ReLU-33	[-1, 128, 4, 4]	0
BasicBlock-34	[-1, 128, 4, 4]	0
Conv2d-35	[-1, 256, 2, 2]	294,912
BatchNorm2d-36	[-1, 256, 2, 2]	512
ReLU-37	[-1, 256, 2, 2]	0
Conv2d-38	[-1, 256, 2, 2]	589,824
BatchNorm2d-39	[-1, 256, 2, 2]	512
Conv2d-40	[-1, 256, 2, 2]	32,768
BatchNorm2d-41	[-1, 256, 2, 2]	512
ReLU-42	[-1, 256, 2, 2]	0
BasicBlock-43	[-1, 256, 2, 2]	0
Conv2d-44	[-1, 256, 2, 2]	589,824
BatchNorm2d-45	[-1, 256, 2, 2]	512
ReLU-46	[-1, 256, 2, 2]	0
Conv2d-47	[-1, 256, 2, 2]	589,824
BatchNorm2d-48	[-1, 256, 2, 2]	512
ReLU-49	[-1, 256, 2, 2]	0
BasicBlock-50	[-1, 256, 2, 2]	0

Conv2d-51	[-1, 512, 1, 1]	1,179,648
BatchNorm2d-52	[-1, 512, 1, 1]	1,024
ReLU-53	[-1, 512, 1, 1]	0
Conv2d-54	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-55	[-1, 512, 1, 1]	1,024
Conv2d-56	[-1, 512, 1, 1]	131,072
BatchNorm2d-57	[-1, 512, 1, 1]	1,024
ReLU-58	[-1, 512, 1, 1]	0
BasicBlock-59	[-1, 512, 1, 1]	0
Conv2d-60	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-61	[-1, 512, 1, 1]	1,024
ReLU-62	[-1, 512, 1, 1]	0
Conv2d-63	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-64	[-1, 512, 1, 1]	1,024
ReLU-65	[-1, 512, 1, 1]	0
BasicBlock-66	[-1, 512, 1, 1]	0
AdaptiveAvgPool2d-67	[-1, 512, 1, 1]	0
Linear-68	[-1, 1000]	513,000

=====

Total params: 11,689,512

Trainable params: 11,689,512

Non-trainable params: 0

Input size (MB): 0.01

Forward/backward pass size (MB): 1.29

Params size (MB): 44.59

Estimated Total Size (MB): 45.90

Визуализация модели:

[illegible]

[illegible]

[illegible]

[illegible]

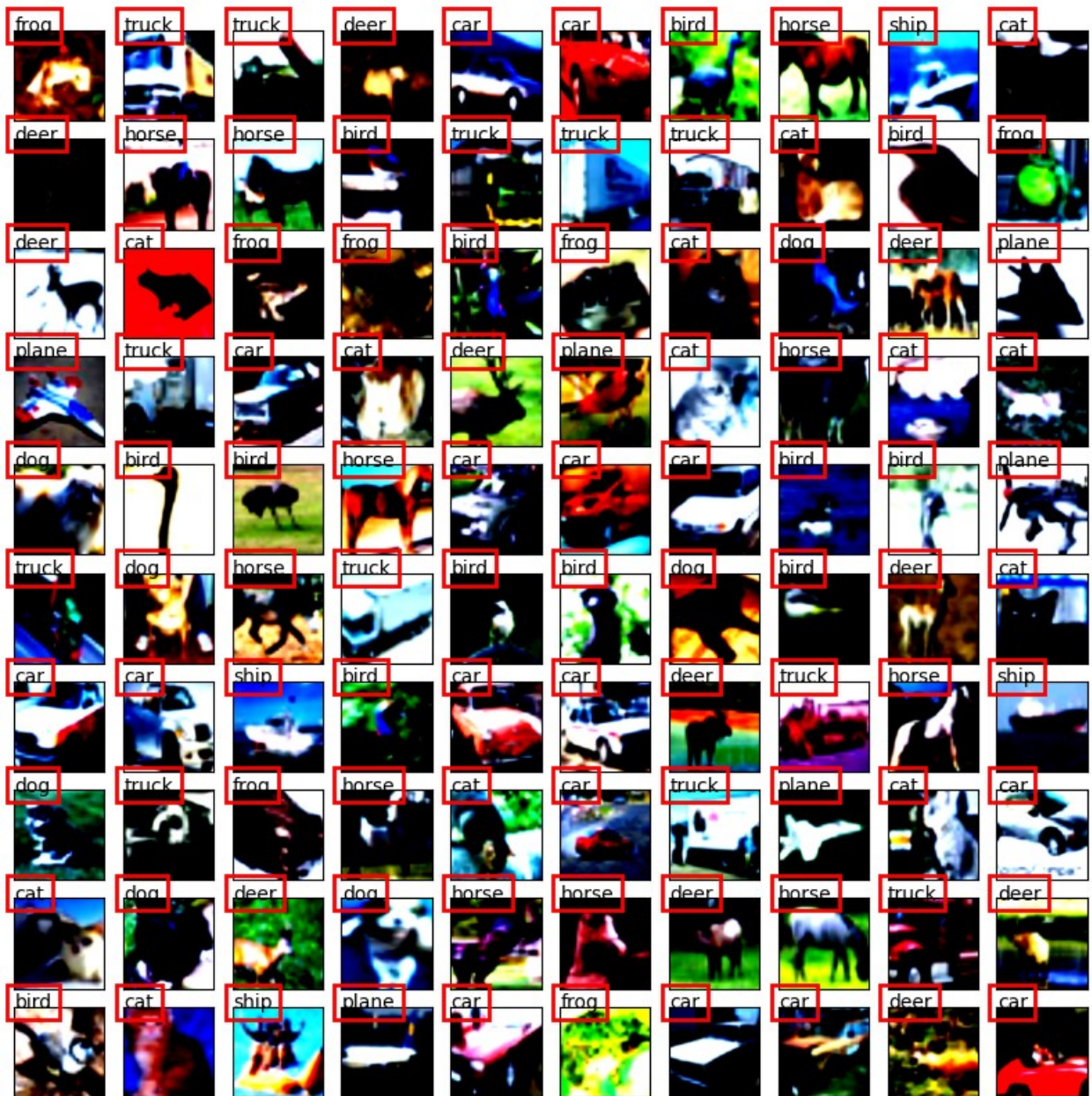
[illegible]

[illegible]

[illegible]

[illegible]

Train samples



Test samples



```
learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.fc.parameters(), lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)

Epoch[0]: accuracy = 0.7274599671363831, time = 134.01403093338013
Epoch[1]: accuracy = 0.7578799724578857, time = 134.17937111854553
```

```
Epoch[2]: accuracy = 0.7758399844169617, time = 136.39211130142212
Epoch[3]: accuracy = 0.7821199893951416, time = 133.69815516471863
Epoch[4]: accuracy = 0.7752799987792969, time = 136.0906012058258
Epoch[5]: accuracy = 0.7862199544906616, time = 136.01660108566284
Epoch[6]: accuracy = 0.7900999784469604, time = 133.74976420402527
Epoch[7]: accuracy = 0.7939800024032593, time = 134.50200629234314
Epoch[8]: accuracy = 0.7942599654197693, time = 134.31048798561096
Epoch[9]: accuracy = 0.7878999710083008, time = 135.3505744934082
Epoch[10]: accuracy = 0.7988199591636658, time = 135.93762755393982
Epoch[11]: accuracy = 0.79367995262146, time = 134.0378932952881
Epoch[12]: accuracy = 0.7895799875259399, time = 130.2157006263733
Epoch[13]: accuracy = 0.7952799797058105, time = 128.0634801387787
Epoch[14]: accuracy = 0.7955399751663208, time = 129.41853523254395
Epoch[15]: accuracy = 0.7999799847602844, time = 128.6872215270996
Epoch[16]: accuracy = 0.7971599698066711, time = 127.93745112419128
Epoch[17]: accuracy = 0.794160008430481, time = 129.3034646511078
Epoch[18]: accuracy = 0.7957800030708313, time = 129.42212867736816
Epoch[19]: accuracy = 0.7978799939155579, time = 129.11809086799622
train time = 5302.444882154465
```

```
test_model(model, test_dataloader)
```

```
Test accuracy: 0.7870999574661255
```

3*. Weights + Sequential

```
model = torch_model.resnet18(torch_model.ResNet18_Weights.DEFAULT)
/usr/local/lib/python3.10/dist-packages/torchvision/models/_
_utils.py:135: UserWarning: Using 'weights' as positional parameter(s)
is deprecated since 0.13 and may be removed in the future. Please use
keyword parameter(s) instead.
  warnings.warn(
Downloading: "https://download.pytorch.org/models/resnet18-
f37072fd.pth" to /root/.cache/torch/hub/checkpoints/resnet18-
f37072fd.pth
100%|██████████| 44.7M/44.7M [00:00<00:00, 142MB/s]
```

```
# Информация об архитектуре и визуализация сети
print("Модель \"CNN\":")
summary(model.to(device), input_size=(3, 32, 32))
```

```
print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters())))
```

```
Модель "CNN":
```

```
-----
Layer (type)                Output Shape                Param #
=====
```

Conv2d-1	[-1, 64, 16, 16]	9,408
BatchNorm2d-2	[-1, 64, 16, 16]	128
ReLU-3	[-1, 64, 16, 16]	0
MaxPool2d-4	[-1, 64, 8, 8]	0
Conv2d-5	[-1, 64, 8, 8]	36,864
BatchNorm2d-6	[-1, 64, 8, 8]	128
ReLU-7	[-1, 64, 8, 8]	0
Conv2d-8	[-1, 64, 8, 8]	36,864
BatchNorm2d-9	[-1, 64, 8, 8]	128
ReLU-10	[-1, 64, 8, 8]	0
BasicBlock-11	[-1, 64, 8, 8]	0
Conv2d-12	[-1, 64, 8, 8]	36,864
BatchNorm2d-13	[-1, 64, 8, 8]	128
ReLU-14	[-1, 64, 8, 8]	0
Conv2d-15	[-1, 64, 8, 8]	36,864
BatchNorm2d-16	[-1, 64, 8, 8]	128
ReLU-17	[-1, 64, 8, 8]	0
BasicBlock-18	[-1, 64, 8, 8]	0
Conv2d-19	[-1, 128, 4, 4]	73,728
BatchNorm2d-20	[-1, 128, 4, 4]	256
ReLU-21	[-1, 128, 4, 4]	0
Conv2d-22	[-1, 128, 4, 4]	147,456
BatchNorm2d-23	[-1, 128, 4, 4]	256
Conv2d-24	[-1, 128, 4, 4]	8,192
BatchNorm2d-25	[-1, 128, 4, 4]	256
ReLU-26	[-1, 128, 4, 4]	0
BasicBlock-27	[-1, 128, 4, 4]	0
Conv2d-28	[-1, 128, 4, 4]	147,456
BatchNorm2d-29	[-1, 128, 4, 4]	256
ReLU-30	[-1, 128, 4, 4]	0
Conv2d-31	[-1, 128, 4, 4]	147,456
BatchNorm2d-32	[-1, 128, 4, 4]	256
ReLU-33	[-1, 128, 4, 4]	0
BasicBlock-34	[-1, 128, 4, 4]	0
Conv2d-35	[-1, 256, 2, 2]	294,912
BatchNorm2d-36	[-1, 256, 2, 2]	512
ReLU-37	[-1, 256, 2, 2]	0
Conv2d-38	[-1, 256, 2, 2]	589,824
BatchNorm2d-39	[-1, 256, 2, 2]	512
Conv2d-40	[-1, 256, 2, 2]	32,768
BatchNorm2d-41	[-1, 256, 2, 2]	512
ReLU-42	[-1, 256, 2, 2]	0
BasicBlock-43	[-1, 256, 2, 2]	0
Conv2d-44	[-1, 256, 2, 2]	589,824
BatchNorm2d-45	[-1, 256, 2, 2]	512
ReLU-46	[-1, 256, 2, 2]	0
Conv2d-47	[-1, 256, 2, 2]	589,824
BatchNorm2d-48	[-1, 256, 2, 2]	512
ReLU-49	[-1, 256, 2, 2]	0

BasicBlock-50	[-1, 256, 2, 2]	0
Conv2d-51	[-1, 512, 1, 1]	1,179,648
BatchNorm2d-52	[-1, 512, 1, 1]	1,024
ReLU-53	[-1, 512, 1, 1]	0
Conv2d-54	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-55	[-1, 512, 1, 1]	1,024
Conv2d-56	[-1, 512, 1, 1]	131,072
BatchNorm2d-57	[-1, 512, 1, 1]	1,024
ReLU-58	[-1, 512, 1, 1]	0
BasicBlock-59	[-1, 512, 1, 1]	0
Conv2d-60	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-61	[-1, 512, 1, 1]	1,024
ReLU-62	[-1, 512, 1, 1]	0
Conv2d-63	[-1, 512, 1, 1]	2,359,296
BatchNorm2d-64	[-1, 512, 1, 1]	1,024
ReLU-65	[-1, 512, 1, 1]	0
BasicBlock-66	[-1, 512, 1, 1]	0
AdaptiveAvgPool2d-67	[-1, 512, 1, 1]	0
Linear-68	[-1, 1000]	513,000

```

=====
Total params: 11,689,512
Trainable params: 11,689,512
Non-trainable params: 0

```

```

-----
Input size (MB): 0.01
Forward/backward pass size (MB): 1.29
Params size (MB): 44.59
Estimated Total Size (MB): 45.90
-----

```

Визуализация модели:


```

for param in model.parameters():
    param.requires_grad = False
model.fc = last_sequential_layer(512)
model.fc.apply(init_weights)
model = model.to(device)

Linear(in_features=512, out_features=256, bias=True)
Parameter containing:
tensor([[ 0.0720, -0.0701, -0.0681, ..., 0.0198, 0.0107, -0.0812],
        [-0.0230, 0.0588, 0.0868, ..., 0.0526, -0.0306, -0.0151],
        [ 0.0261, -0.0610, -0.0360, ..., 0.0334, -0.0109, 0.0788],
        ...,
        [-0.0444, -0.0580, -0.0670, ..., 0.0031, 0.0498, -0.0756],
        [-0.0387, 0.0448, -0.0629, ..., -0.0607, 0.0527, 0.0145],
        [-0.0198, -0.0583, 0.0566, ..., 0.0535, 0.0018, -0.0780]],
        requires_grad=True)
ReLU()
Linear(in_features=256, out_features=10, bias=True)
Parameter containing:
tensor([[ 0.0628, 0.1211, -0.1041, ..., -0.0775, -0.1007, 0.1224],
        [ 0.0815, -0.0602, -0.1055, ..., 0.0397, -0.0264, 0.0204],
        [ 0.0301, -0.0678, 0.0719, ..., -0.1130, 0.1392, -0.1355],
        ...,
        [-0.1405, 0.0262, 0.1393, ..., 0.1376, -0.0812, 0.0704],
        [ 0.0861, 0.0051, -0.0794, ..., -0.0544, 0.0112, 0.1229],
        [ 0.1047, -0.0255, 0.1043, ..., -0.0666, -0.1122, 0.1334]],
        requires_grad=True)
Sequential(
  (0): Linear(in_features=512, out_features=256, bias=True)
  (1): ReLU()
  (2): Linear(in_features=256, out_features=10, bias=True)
)

print(model)

ResNet(
  (conv1): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2),
padding=(3, 3), bias=False)
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  (relu): ReLU(inplace=True)
  (maxpool): MaxPool2d(kernel_size=3, stride=2, padding=1, dilation=1,
ceil_mode=False)
  (layer1): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (relu): ReLU(inplace=True)

```

```

        (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
    (1): BasicBlock(
        (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
)
(layer2): Sequential(
  (0): BasicBlock(
    (conv1): Conv2d(64, 128, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1), bias=False)
    (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (relu): ReLU(inplace=True)
    (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (downsample): Sequential(
      (0): Conv2d(64, 128, kernel_size=(1, 1), stride=(2, 2),
bias=False)
      (1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
  )
  (1): BasicBlock(
    (conv1): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (relu): ReLU(inplace=True)
    (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
  )
)
(layer3): Sequential(
  (0): BasicBlock(

```

```

        (conv1): Conv2d(128, 256, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1), bias=False)
        (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (downsample): Sequential(
          (0): Conv2d(128, 256, kernel_size=(1, 1), stride=(2, 2),
bias=False)
          (1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        )
      )
    (1): BasicBlock(
      (conv1): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
  )
(layer4): Sequential(
  (0): BasicBlock(
    (conv1): Conv2d(256, 512, kernel_size=(3, 3), stride=(2, 2),
padding=(1, 1), bias=False)
    (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (relu): ReLU(inplace=True)
    (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (downsample): Sequential(
      (0): Conv2d(256, 512, kernel_size=(1, 1), stride=(2, 2),
bias=False)
      (1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
  )
  (1): BasicBlock(
    (conv1): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)

```


[illegible]

[illegible]

[illegible]

[illegible]

Number of test samples: 10000

[illegible]

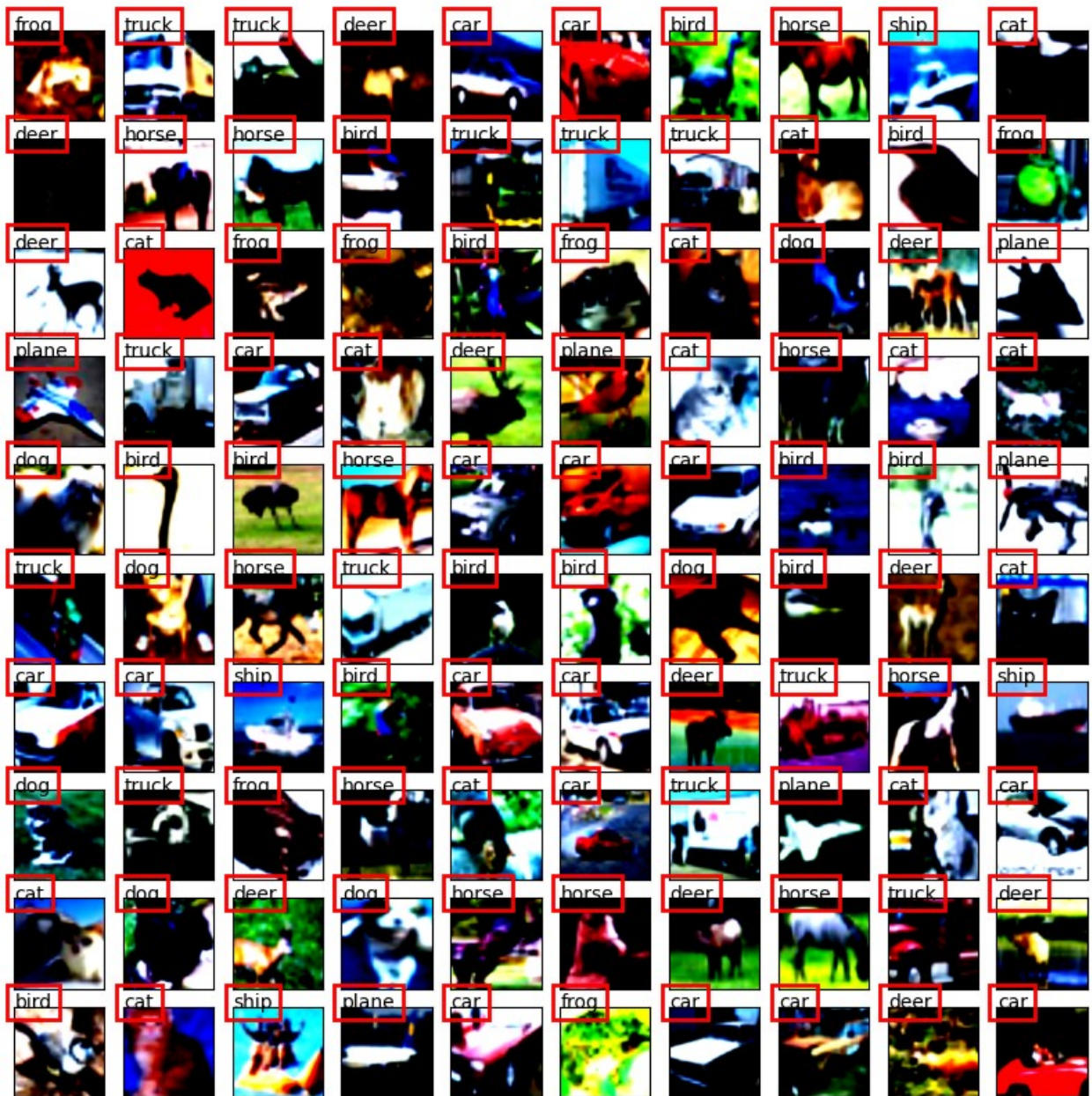
[illegible]

[illegible]

[illegible]

[illegible]

Train samples



Test samples



```
learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.fc.parameters(), lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)

Epoch[0]: accuracy = 0.7594799995422363, time = 133.73124742507935
Epoch[1]: accuracy = 0.7796799540519714, time = 133.07196974754333
```

```
Epoch[2]: accuracy = 0.7849000096321106, time = 131.48094820976257
Epoch[3]: accuracy = 0.7990399599075317, time = 133.5985472202301
Epoch[4]: accuracy = 0.8082799911499023, time = 133.9399676322937
Epoch[5]: accuracy = 0.8157399892807007, time = 131.6291744709015
Epoch[6]: accuracy = 0.8211399912834167, time = 130.24699783325195
Epoch[7]: accuracy = 0.8364599943161011, time = 133.8600754737854
Epoch[8]: accuracy = 0.8285399675369263, time = 133.16687679290771
Epoch[9]: accuracy = 0.8524399995803833, time = 131.25341796875
Epoch[10]: accuracy = 0.8587200045585632, time = 133.86997056007385
Epoch[11]: accuracy = 0.8580799698829651, time = 133.42539286613464
Epoch[12]: accuracy = 0.8598399758338928, time = 133.3869035243988
Epoch[13]: accuracy = 0.8741199970245361, time = 134.33729028701782
Epoch[14]: accuracy = 0.8752399682998657, time = 132.73814916610718
Epoch[15]: accuracy = 0.8773799538612366, time = 132.74400520324707
Epoch[16]: accuracy = 0.8805599808692932, time = 134.41021251678467
Epoch[17]: accuracy = 0.8913599848747253, time = 135.32344675064087
Epoch[18]: accuracy = 0.8877599835395813, time = 133.90785479545593
Epoch[19]: accuracy = 0.8972399830818176, time = 132.90964651107788
train time = 5322.203461408615
```

```
test_model(model, test_dataloader)
```

```
Test accuracy: 0.7958999872207642
```

Densenet121

1. Single layer

```
model =
torch_model.densenet121(torch_model.DenseNet121_Weights.DEFAULT)

/usr/local/lib/python3.10/dist-packages/torchvision/models/
_utils.py:135: UserWarning: Using 'weights' as positional parameter(s)
is deprecated since 0.13 and may be removed in the future. Please use
keyword parameter(s) instead.
  warnings.warn(
Downloading: "https://download.pytorch.org/models/densenet121-
a639ec97.pth" to /root/.cache/torch/hub/checkpoints/densenet121-
a639ec97.pth
100%|██████████| 30.8M/30.8M [00:00<00:00, 132MB/s]

print(model)

DenseNet(
  (features): Sequential(
    (conv0): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2),
padding=(3, 3), bias=False)
    (norm0): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (relu0): ReLU(inplace=True)
    (pool0): MaxPool2d(kernel_size=3, stride=2, padding=1, dilation=1,
```



```

ceil_mode=False)
    (denseblock1): _DenseBlock(
      (denselayer1): _DenseLayer(
        (norm1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(64, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      )
      (denselayer2): _DenseLayer(
        (norm1): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(96, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      )
      (denselayer3): _DenseLayer(
        (norm1): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(128, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      )
      (denselayer4): _DenseLayer(
        (norm1): BatchNorm2d(160, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(160, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)

```

```

    )
    (denselayer5): _DenseLayer(
      (norm1): BatchNorm2d(192, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(192, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer6): _DenseLayer(
      (norm1): BatchNorm2d(224, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(224, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
  )
  (transition1): _Transition(
    (norm): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (relu): ReLU(inplace=True)
    (conv): Conv2d(256, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (pool): AvgPool2d(kernel_size=2, stride=2, padding=0)
  )
  (denseblock2): _DenseBlock(
    (denselayer1): _DenseLayer(
      (norm1): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(128, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer2): _DenseLayer(
      (norm1): BatchNorm2d(160, eps=1e-05, momentum=0.1,

```

```

    affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(160, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer3): _DenseLayer(
        (norm1): BatchNorm2d(192, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(192, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer4): _DenseLayer(
        (norm1): BatchNorm2d(224, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(224, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer5): _DenseLayer(
        (norm1): BatchNorm2d(256, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(256, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer6): _DenseLayer(
        (norm1): BatchNorm2d(288, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)

```

```

        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(288, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer7): _DenseLayer(
        (norm1): BatchNorm2d(320, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(320, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer8): _DenseLayer(
        (norm1): BatchNorm2d(352, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(352, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer9): _DenseLayer(
        (norm1): BatchNorm2d(384, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(384, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer10): _DenseLayer(
        (norm1): BatchNorm2d(416, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)

```

```

        (conv1): Conv2d(416, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer11): _DenseLayer(
        (norm1): BatchNorm2d(448, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(448, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer12): _DenseLayer(
        (norm1): BatchNorm2d(480, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(480, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    )
    (transition2): _Transition(
        (norm): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv): Conv2d(512, 256, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (pool): AvgPool2d(kernel_size=2, stride=2, padding=0)
    )
    (denseblock3): _DenseBlock(
        (denselayer1): _DenseLayer(
            (norm1): BatchNorm2d(256, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)
            (conv1): Conv2d(256, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
            (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,

```

```

    affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer2): _DenseLayer(
        (norm1): BatchNorm2d(288, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(288, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer3): _DenseLayer(
        (norm1): BatchNorm2d(320, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(320, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer4): _DenseLayer(
        (norm1): BatchNorm2d(352, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(352, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer5): _DenseLayer(
        (norm1): BatchNorm2d(384, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(384, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)

```

```

        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer6): _DenseLayer(
        (norm1): BatchNorm2d(416, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(416, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer7): _DenseLayer(
        (norm1): BatchNorm2d(448, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(448, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer8): _DenseLayer(
        (norm1): BatchNorm2d(480, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(480, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer9): _DenseLayer(
        (norm1): BatchNorm2d(512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)

```

```

        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer10): _DenseLayer(
        (norm1): BatchNorm2d(544, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(544, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer11): _DenseLayer(
        (norm1): BatchNorm2d(576, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(576, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer12): _DenseLayer(
        (norm1): BatchNorm2d(608, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(608, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer13): _DenseLayer(
        (norm1): BatchNorm2d(640, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(640, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),

```



```

padding=(1, 1), bias=False)
    )
    (denselayer14): _DenseLayer(
        (norm1): BatchNorm2d(672, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(672, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer15): _DenseLayer(
        (norm1): BatchNorm2d(704, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(704, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer16): _DenseLayer(
        (norm1): BatchNorm2d(736, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(736, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer17): _DenseLayer(
        (norm1): BatchNorm2d(768, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(768, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)

```

```

    )
    (denselayer18): _DenseLayer(
      (norm1): BatchNorm2d(800, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(800, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer19): _DenseLayer(
      (norm1): BatchNorm2d(832, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(832, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer20): _DenseLayer(
      (norm1): BatchNorm2d(864, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(864, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer21): _DenseLayer(
      (norm1): BatchNorm2d(896, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(896, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer22): _DenseLayer(

```

```

        (norm1): BatchNorm2d(928, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(928, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer23): _DenseLayer(
        (norm1): BatchNorm2d(960, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(960, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer24): _DenseLayer(
        (norm1): BatchNorm2d(992, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(992, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    )
    (transition3): _Transition(
        (norm): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv): Conv2d(1024, 512, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (pool): AvgPool2d(kernel_size=2, stride=2, padding=0)
    )
    (denseblock4): _DenseBlock(
        (denselayer1): _DenseLayer(
            (norm1): BatchNorm2d(512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)

```

```

        (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer2): _DenseLayer(
        (norm1): BatchNorm2d(544, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(544, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer3): _DenseLayer(
        (norm1): BatchNorm2d(576, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(576, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer4): _DenseLayer(
        (norm1): BatchNorm2d(608, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(608, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer5): _DenseLayer(
        (norm1): BatchNorm2d(640, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(640, 128, kernel_size=(1, 1), stride=(1, 1),

```

```

bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer6): _DenseLayer(
    (norm1): BatchNorm2d(672, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(672, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer7): _DenseLayer(
    (norm1): BatchNorm2d(704, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(704, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer8): _DenseLayer(
    (norm1): BatchNorm2d(736, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(736, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer9): _DenseLayer(
    (norm1): BatchNorm2d(768, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(768, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)

```

```

        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer10): _DenseLayer(
        (norm1): BatchNorm2d(800, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(800, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer11): _DenseLayer(
        (norm1): BatchNorm2d(832, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(832, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer12): _DenseLayer(
        (norm1): BatchNorm2d(864, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(864, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer13): _DenseLayer(
        (norm1): BatchNorm2d(896, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(896, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,

```

```

    affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer14): _DenseLayer(
        (norm1): BatchNorm2d(928, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(928, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer15): _DenseLayer(
        (norm1): BatchNorm2d(960, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(960, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer16): _DenseLayer(
        (norm1): BatchNorm2d(992, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(992, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    )
    (norm5): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    )
    (classifier): Linear(in_features=1024, out_features=1000, bias=True)
)

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

```

```

model.classifier = last_layer(1024)
model = model.to(device)

print(model)

DenseNet(
  (features): Sequential(
    (conv0): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2),
padding=(3, 3), bias=False)
    (norm0): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (relu0): ReLU(inplace=True)
    (pool0): MaxPool2d(kernel_size=3, stride=2, padding=1, dilation=1,
ceil_mode=False)
    (denseblock1): _DenseBlock(
      (denselayer1): _DenseLayer(
        (norm1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(64, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      )
      (denselayer2): _DenseLayer(
        (norm1): BatchNorm2d(96, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(96, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      )
      (denselayer3): _DenseLayer(
        (norm1): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(128, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
      )
    )
  )
)

```



```

    )
    (denselayer4): _DenseLayer(
      (norm1): BatchNorm2d(160, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(160, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer5): _DenseLayer(
      (norm1): BatchNorm2d(192, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(192, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer6): _DenseLayer(
      (norm1): BatchNorm2d(224, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(224, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
  )
  (transition1): _Transition(
    (norm): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (relu): ReLU(inplace=True)
    (conv): Conv2d(256, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (pool): AvgPool2d(kernel_size=2, stride=2, padding=0)
  )
  (denseblock2): _DenseBlock(
    (denselayer1): _DenseLayer(
      (norm1): BatchNorm2d(128, eps=1e-05, momentum=0.1,

```

```

affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(128, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
)
    (denselayer2): _DenseLayer(
    (norm1): BatchNorm2d(160, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(160, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
)
    (denselayer3): _DenseLayer(
    (norm1): BatchNorm2d(192, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(192, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
)
    (denselayer4): _DenseLayer(
    (norm1): BatchNorm2d(224, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(224, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
)
    (denselayer5): _DenseLayer(
    (norm1): BatchNorm2d(256, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)

```

```

        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(256, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer6): _DenseLayer(
        (norm1): BatchNorm2d(288, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(288, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer7): _DenseLayer(
        (norm1): BatchNorm2d(320, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(320, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer8): _DenseLayer(
        (norm1): BatchNorm2d(352, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(352, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer9): _DenseLayer(
        (norm1): BatchNorm2d(384, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)

```

```

        (conv1): Conv2d(384, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer10): _DenseLayer(
        (norm1): BatchNorm2d(416, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(416, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer11): _DenseLayer(
        (norm1): BatchNorm2d(448, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(448, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer12): _DenseLayer(
        (norm1): BatchNorm2d(480, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(480, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    )
    (transition2): _Transition(
        (norm): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)

```

```

        (conv): Conv2d(512, 256, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (pool): AvgPool2d(kernel_size=2, stride=2, padding=0)
    )
    (denseblock3): _DenseBlock(
        (denselayer1): _DenseLayer(
            (norm1): BatchNorm2d(256, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)
            (conv1): Conv2d(256, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
            (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu2): ReLU(inplace=True)
            (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
        (denselayer2): _DenseLayer(
            (norm1): BatchNorm2d(288, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)
            (conv1): Conv2d(288, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
            (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu2): ReLU(inplace=True)
            (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
        (denselayer3): _DenseLayer(
            (norm1): BatchNorm2d(320, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)
            (conv1): Conv2d(320, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
            (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu2): ReLU(inplace=True)
            (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
        (denselayer4): _DenseLayer(
            (norm1): BatchNorm2d(352, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)
            (conv1): Conv2d(352, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
            (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)

```

```

        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer5): _DenseLayer(
        (norm1): BatchNorm2d(384, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(384, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer6): _DenseLayer(
        (norm1): BatchNorm2d(416, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(416, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer7): _DenseLayer(
        (norm1): BatchNorm2d(448, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(448, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer8): _DenseLayer(
        (norm1): BatchNorm2d(480, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(480, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)

```

```

        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer9): _DenseLayer(
        (norm1): BatchNorm2d(512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer10): _DenseLayer(
        (norm1): BatchNorm2d(544, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(544, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer11): _DenseLayer(
        (norm1): BatchNorm2d(576, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(576, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer12): _DenseLayer(
        (norm1): BatchNorm2d(608, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(608, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),

```

```

padding=(1, 1), bias=False)
    )
    (denselayer13): _DenseLayer(
        (norm1): BatchNorm2d(640, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(640, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer14): _DenseLayer(
        (norm1): BatchNorm2d(672, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(672, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer15): _DenseLayer(
        (norm1): BatchNorm2d(704, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(704, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer16): _DenseLayer(
        (norm1): BatchNorm2d(736, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(736, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)

```



```

    )
    (denselayer17): _DenseLayer(
      (norm1): BatchNorm2d(768, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(768, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer18): _DenseLayer(
      (norm1): BatchNorm2d(800, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(800, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer19): _DenseLayer(
      (norm1): BatchNorm2d(832, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(832, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer20): _DenseLayer(
      (norm1): BatchNorm2d(864, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu1): ReLU(inplace=True)
      (conv1): Conv2d(864, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
      (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
      (relu2): ReLU(inplace=True)
      (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )

```

```

        (denselayer21): _DenseLayer(
          (norm1): BatchNorm2d(896, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (relu1): ReLU(inplace=True)
          (conv1): Conv2d(896, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
          (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (relu2): ReLU(inplace=True)
          (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
        (denselayer22): _DenseLayer(
          (norm1): BatchNorm2d(928, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (relu1): ReLU(inplace=True)
          (conv1): Conv2d(928, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
          (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (relu2): ReLU(inplace=True)
          (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
        (denselayer23): _DenseLayer(
          (norm1): BatchNorm2d(960, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (relu1): ReLU(inplace=True)
          (conv1): Conv2d(960, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
          (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (relu2): ReLU(inplace=True)
          (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
        (denselayer24): _DenseLayer(
          (norm1): BatchNorm2d(992, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (relu1): ReLU(inplace=True)
          (conv1): Conv2d(992, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
          (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
          (relu2): ReLU(inplace=True)
          (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
      )
    (transition3): _Transition(

```

```

        (norm): BatchNorm2d(1024, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
        (relu): ReLU(inplace=True)
        (conv): Conv2d(1024, 512, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (pool): AvgPool2d(kernel_size=2, stride=2, padding=0)
    )
    (denseblock4): _DenseBlock(
        (denselayer1): _DenseLayer(
            (norm1): BatchNorm2d(512, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)
            (conv1): Conv2d(512, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
            (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu2): ReLU(inplace=True)
            (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
        (denselayer2): _DenseLayer(
            (norm1): BatchNorm2d(544, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)
            (conv1): Conv2d(544, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
            (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu2): ReLU(inplace=True)
            (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
        (denselayer3): _DenseLayer(
            (norm1): BatchNorm2d(576, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)
            (conv1): Conv2d(576, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
            (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu2): ReLU(inplace=True)
            (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
        )
        (denselayer4): _DenseLayer(
            (norm1): BatchNorm2d(608, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
            (relu1): ReLU(inplace=True)
            (conv1): Conv2d(608, 128, kernel_size=(1, 1), stride=(1, 1),

```

```

bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer5): _DenseLayer(
    (norm1): BatchNorm2d(640, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(640, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer6): _DenseLayer(
    (norm1): BatchNorm2d(672, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(672, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer7): _DenseLayer(
    (norm1): BatchNorm2d(704, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(704, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer8): _DenseLayer(
    (norm1): BatchNorm2d(736, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(736, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)

```

```

        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer9): _DenseLayer(
        (norm1): BatchNorm2d(768, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(768, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer10): _DenseLayer(
        (norm1): BatchNorm2d(800, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(800, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer11): _DenseLayer(
        (norm1): BatchNorm2d(832, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(832, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu2): ReLU(inplace=True)
        (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer12): _DenseLayer(
        (norm1): BatchNorm2d(864, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
        (relu1): ReLU(inplace=True)
        (conv1): Conv2d(864, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
        (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,

```

```

affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer13): _DenseLayer(
    (norm1): BatchNorm2d(896, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(896, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer14): _DenseLayer(
    (norm1): BatchNorm2d(928, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(928, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer15): _DenseLayer(
    (norm1): BatchNorm2d(960, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(960, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu2): ReLU(inplace=True)
    (conv2): Conv2d(128, 32, kernel_size=(3, 3), stride=(1, 1),
padding=(1, 1), bias=False)
    )
    (denselayer16): _DenseLayer(
    (norm1): BatchNorm2d(992, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)
    (relu1): ReLU(inplace=True)
    (conv1): Conv2d(992, 128, kernel_size=(1, 1), stride=(1, 1),
bias=False)
    (norm2): BatchNorm2d(128, eps=1e-05, momentum=0.1,
affine=True, track_running_stats=True)

```


[illegible]

[illegible]

[illegible]

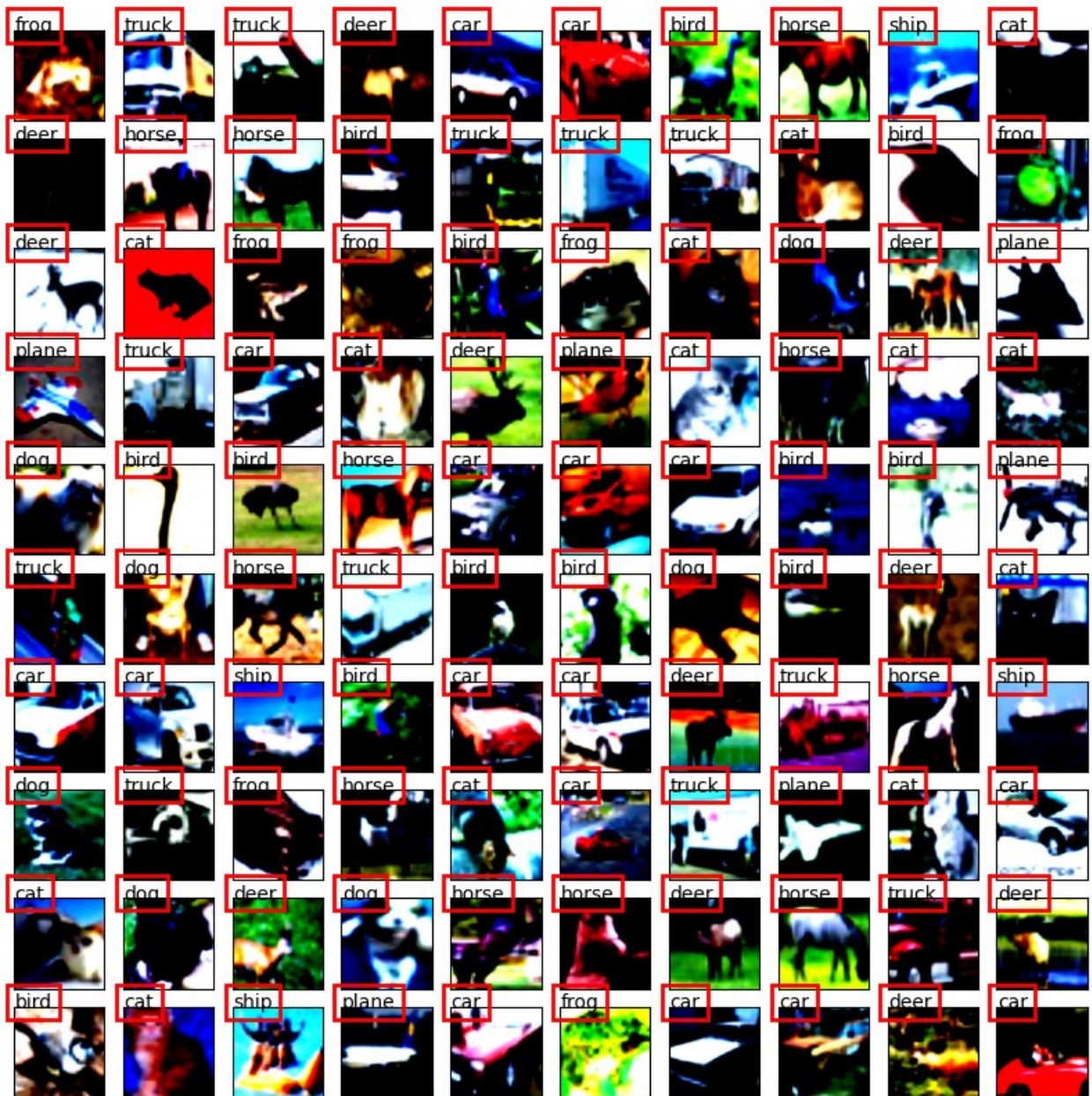
[illegible]

[illegible]

[illegible]

[illegible]

Train samples



Test samples



```
learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.classifier.parameters(),
                               lr=learning_rate)
model = model.to(device)

train_loop(model, train_dataloader, optimizer, loss_function)
```

```
Epoch[0]: accuracy = 0.7708999514579773, time = 172.45439887046814
Epoch[1]: accuracy = 0.7917199730873108, time = 173.8807408809662
Epoch[2]: accuracy = 0.8006599545478821, time = 173.64650082588196
Epoch[3]: accuracy = 0.796999990940094, time = 173.86425352096558
Epoch[4]: accuracy = 0.8100199699401855, time = 173.65966701507568
Epoch[5]: accuracy = 0.8076399564743042, time = 174.50546216964722
Epoch[6]: accuracy = 0.8104000091552734, time = 173.82365036010742
Epoch[7]: accuracy = 0.8085799813270569, time = 173.53456020355225
Epoch[8]: accuracy = 0.811739981174469, time = 173.1445758342743
Epoch[9]: accuracy = 0.811199963092804, time = 174.12666630744934
Epoch[10]: accuracy = 0.8151999711990356, time = 173.08181023597717
Epoch[11]: accuracy = 0.8173199892044067, time = 173.460134267807
Epoch[12]: accuracy = 0.8162199854850769, time = 173.07563638687134
Epoch[13]: accuracy = 0.814300000667572, time = 173.42166662216187
Epoch[14]: accuracy = 0.8192600011825562, time = 174.47900819778442
Epoch[15]: accuracy = 0.8169800043106079, time = 171.43979167938232
Epoch[16]: accuracy = 0.8141999840736389, time = 171.9295723438263
Epoch[17]: accuracy = 0.8133400082588196, time = 172.0942211151123
Epoch[18]: accuracy = 0.8139399886131287, time = 172.2222945690155
Epoch[19]: accuracy = 0.8186999559402466, time = 172.34216904640198
train time = 6931.575216054916
```

```
test_model(model, test_dataloader)
```

```
Test accuracy: 0.8026999831199646
```

1. Sequential layer

```
model =
torch_model.densenet121(torch_model.DenseNet121_Weights.DEFAULT)

# Информация об архитектуре и визуализация сети
print("Модель \"CNN\":")
summary(model.to(device), input_size=(3, 32, 32))

print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters()))))

for param in model.parameters():
    param.requires_grad = False
model.classifier = last_sequential_layer(1024)
model = model.to(device)

print(model)

transforms = torch_model.DenseNet121_Weights.DEFAULT.transforms()
batch_size = 128
train_dataloader, test_dataloader = download_data(transforms,
batch_size)
```

```

learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.classifier.parameters(),
lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)
test_model(model, test_dataloader)

```

1. Weights

```

model =
torch_model.densenet121(torch_model.DenseNet121_Weights.DEFAULT)

# Информация об архитектуре и визуализация сети
print("Модель \ "CNN\ ":")
summary(model.to(device), input_size=(3, 32, 32))

print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters()))))

for param in model.parameters():
    param.requires_grad = False
model.classifier = last_layer(1024)
model.classifier.apply(init_weights)
model = model.to(device)

print(model)

transforms = torch_model.DenseNet121_Weights.DEFAULT.transforms()
batch_size = 128
train_dataloader, test_dataloader = download_data(transforms,
batch_size)

learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.fc.parameters(), lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)
test_model(model, test_dataloader)

```

Resnext50_32x4d

1. Single layer

```

model =
torch_model.resnext50_32x4d(torch_model.ResNeXt50_32X4D_Weights.DEFAULT)

```

```

# Информация об архитектуре и визуализация сети
print("Модель \ "CNN\ ":")
summary(model.to(device), input_size=(3, 32, 32))

print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters()))))

for param in model.parameters():
    param.requires_grad = False
model.fc = last_layer(2048)
model = model.to(device)

print(model)

transforms = torch_model.ResNeXt50_32X4D_Weights.DEFAULT.transforms()
batch_size = 128
train_dataloader, test_dataloader = download_data(transforms,
batch_size)

learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.fc.parameters(), lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)

test_model(model, test_dataloader)

```

1. Sequential layer

```

model =
torch_model.resnext50_32x4d(torch_model.ResNeXt50_32X4D_Weights.DEFAULT)

# Информация об архитектуре и визуализация сети
print("Модель \ "CNN\ ":")
summary(model.to(device), input_size=(3, 32, 32))

print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters()))))

for param in model.parameters():
    param.requires_grad = False
model.fc = last_sequential_layer(2048)
model = model.to(device)

print(model)

transforms = torch_model.ResNeXt50_32X4D_Weights.DEFAULT.transforms()
batch_size = 128

```

```

train_dataloader, test_dataloader = download_data(transforms,
batch_size)

learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.fc.parameters(), lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)

test_model(model, test_dataloader)

```

1. Weights

```

model =
torch_model.resnext50_32x4d(torch_model.ResNeXt50_32X4D_Weights.DEFAULT)

# Информация об архитектуре и визуализация сети
print("Модель \ "CNN\ ":")
summary(model.to(device), input_size=(3, 32, 32))

print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters()))))

for param in model.parameters():
    param.requires_grad = False
model.fc = last_layer(2048)
model.fc.apply(init_weights)
model = model.to(device)

print(model)

transforms = torch_model.ResNeXt50_32X4D_Weights.DEFAULT.transforms()
batch_size = 128
train_dataloader, test_dataloader = download_data(transforms,
batch_size)

learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.fc.parameters(), lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)

test_model(model, test_dataloader)

```

ViT_B_16

1. Single layer

```

model = torch_model.vit_b_16(torch_model.ViT_B_16_Weights.DEFAULT)

```



```

# Информация об архитектуре и визуализация сети
print("Модель \ "CNN\ ":")
summary(model.to(device), input_size=(3, 32, 32))

print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters()))))

for param in model.parameters():
    param.requires_grad = False
model.heads = last_layer(768)
model = model.to(device)

print(model)

transforms = torch_model.ViT_B_16_Weights.DEFAULT.transforms()
batch_size = 128
train_dataloader, test_dataloader = download_data(transforms,
batch_size)

learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.heads.parameters(),
lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)
test_model(model, test_dataloader)

```

1. Sequential layer

```

model = torch_model.vit_b_16(torch_model.ViT_B_16_Weights.DEFAULT)

# Информация об архитектуре и визуализация сети
print("Модель \ "CNN\ ":")
summary(model.to(device), input_size=(3, 32, 32))

print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters()))))

for param in model.parameters():
    param.requires_grad = False
model.heads = last_sequential_layer(768)
model = model.to(device)

print(model)

transforms = torch_model.ViT_B_16_Weights.DEFAULT.transforms()
batch_size = 128
train_dataloader, test_dataloader = download_data(transforms,
batch_size)

```

```

learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.heads.parameters(),
lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)

test_model(model, test_dataloader)

```

1. Weights

```

model = torch_model.vit_b_16(torch_model.ViT_B_16_Weights.DEFAULT)

# Информация об архитектуре и визуализация сети
print("Модель \ "CNN\ ":")
summary(model.to(device), input_size=(3, 32, 32))

print("Визуализация модели:")
display(make_dot(model(torch.randn(16, 3, 32, 32).to(device)),
params=dict(model.named_parameters()))))

for param in model.parameters():
    param.requires_grad = False
model.heads = last_layer(768)
model.heads.apply(init_weights)
model = model.to(device)

print(model)

transforms = torch_model.ViT_B_16_Weights.DEFAULT.transforms()
batch_size = 128
train_dataloader, test_dataloader = download_data(transforms,
batch_size)

learning_rate = 0.001
loss_function = torch.nn.CrossEntropyLoss()
optimizer = torch.optim.Adam(model.heads.parameters(),
lr=learning_rate)

train_loop(model, train_dataloader, optimizer, loss_function)

test_model(model, test_dataloader)

```