wk6_hw5_p1

November 7, 2024

[1]:
```
from models import *
from utilities import Utilities as helper
from utilities import AverageMeter
from data_loader import trainloader
from data_loader import testloader
```

```
=> Building model…
Files already downloaded and verified
Files already downloaded and verified
```

[2]:
```
model_name = "resnet20_quan"
resnet20_model_quan = resnet20_quant().cuda()
```

[ ]:
```
lr = 7e-3
weight_decay = 1e-4
epochs = 100
# pre_best_prec = 85.470

criterion = nn.CrossEntropyLoss().cuda()

res20_optimizer = torch.optim.SGD(resnet20_model_quan.parameters(), lr=lr,␣
 ↪momentum=0.9, weight_decay=weight_decay)
pre_best_prec = helper.train_model(resnet20_model_quan, model_name=model_name,␣
 ↪optimizer=res20_optimizer, trainloader=trainloader, testloader=testloader,␣
 ↪criterion=criterion, epochs=epochs, pre_best_prec=pre_best_prec)
```

[8]:
```
helper.test_model(model=resnet20_model_quan, model_name=model_name,␣
 ↪testloader=testloader)
```

```
Test: [0/79]    Time 0.190 (0.190)      Loss 0.2105 (0.2105)    Prec 93.750%
(93.750%)
 * Prec 90.380%
```

[4]:
```
w_bit = 4
quan_con_layer = next(resnet20_model_quan.layer1[0].modules()).conv1
weight_q = quan_con_layer.weight_q # quantized value is stored during the␣
 ↪training
```

```
w_alpha = quan_con_layer.weight_quant.wgt_alpha.data.item()  # alpha is defined in your model already. bring it out here
w_delta = w_alpha/(2**(w_bit-1)-1)   # delta can be calculated by using alpha and w_bit
weight_int = weight_q/w_delta # w_int can be calculated by weight_q and w_delta
print(weight_int) # you should see clean integer numbers


# conv.weight  = torch.nn.Parameter(torch.randn(n_ch_out,n_ch_in,k_size,k_size))
# ############################


# weight_quant = weight_quantize_fn(w_bit= w_bits)  ## define quant function
# weight_quant.wgt_alpha = torch.tensor(w_alpha)
# w_quant      = weight_quant(conv.weight)
# w_delta      = w_alpha/(2**(w_bits-1)-1)
# w_int        = w_quant/w_delta
```

```
tensor([[[[-0.0000,  0.0000,  0.0000],
          [ 0.0000,  1.0000,  1.0000],
          [-1.0000, -0.0000,  1.0000]],

         [[ 2.0000, -0.0000, -1.0000],
          [-1.0000, -4.0000, -5.0000],
          [ 1.0000, -3.0000, -5.0000]],

         [[ 2.0000,  2.0000,  2.0000],
          [ 3.0000,  3.0000,  2.0000],
          [ 2.0000,  3.0000,  2.0000]],

         …,

         [[-0.0000, -4.0000, -1.0000],
          [-2.0000, -5.0000, -0.0000],
          [-1.0000, -3.0000,  1.0000]],

         [[ 1.0000, -3.0000, -3.0000],
          [ 2.0000, -3.0000, -4.0000],
          [ 3.0000, -1.0000, -2.0000]],

         [[ 3.0000,  2.0000,  1.0000],
          [ 3.0000,  3.0000,  4.0000],
          [ 0.0000,  1.0000,  2.0000]]],


        [[[-5.0000, -7.0000, -5.0000],
          [ 1.0000,  2.0000,  0.0000],
```

```
  [ 2.0000,  6.0000,  4.0000]],

 [[ 1.0000,  1.0000,  2.0000],
  [-0.0000, -4.0000, -2.0000],
  [ 1.0000, -1.0000, -3.0000]],

 [[ 1.0000,  2.0000, -0.0000],
  [ 2.0000,  2.0000, -1.0000],
  [-0.0000,  1.0000, -3.0000]],

 …,

 [[-0.0000,  0.0000,  6.0000],
  [-3.0000, -2.0000,  6.0000],
  [-5.0000, -4.0000,  4.0000]],

 [[-0.0000, -2.0000,  0.0000],
  [ 2.0000, -4.0000, -0.0000],
  [ 2.0000, -4.0000,  1.0000]],

 [[ 1.0000, -0.0000,  2.0000],
  [-2.0000, -1.0000,  3.0000],
  [ 2.0000,  5.0000,  3.0000]]],


[[[ 0.0000,  2.0000,  4.0000],
  [ 2.0000,  5.0000,  6.0000],
  [ 1.0000, -0.0000,  3.0000]],

 [[ 3.0000, -1.0000, -4.0000],
  [-3.0000, -2.0000, -1.0000],
  [-2.0000, -2.0000,  2.0000]],

 [[ 3.0000,  3.0000,  2.0000],
  [ 3.0000,  1.0000,  0.0000],
  [ 2.0000, -2.0000, -0.0000]],

 …,

 [[-4.0000, -1.0000, -1.0000],
  [-0.0000, -1.0000, -3.0000],
  [ 2.0000,  3.0000, -7.0000]],

 [[-5.0000, -1.0000, -3.0000],
  [-6.0000, -3.0000,  1.0000],
  [-7.0000, -2.0000,  4.0000]],

 [[-1.0000, -3.0000,  0.0000],
```

```
  [ 0.0000,  0.0000, -3.0000],
  [ 4.0000,  4.0000,  0.0000]]],


…,


[[[-2.0000, -1.0000, -2.0000],
  [-2.0000, -0.0000,  0.0000],
  [-1.0000,  1.0000,  3.0000]],

 [[ 4.0000,  1.0000, -0.0000],
  [ 7.0000, -2.0000,  0.0000],
  [ 7.0000,  2.0000, -0.0000]],

 [[-2.0000, -2.0000,  2.0000],
  [ 0.0000, -0.0000,  1.0000],
  [ 1.0000,  1.0000, -0.0000]],

 …,

 [[ 4.0000,  4.0000, -2.0000],
  [ 4.0000,  1.0000, -5.0000],
  [-1.0000, -2.0000, -1.0000]],

 [[-4.0000,  2.0000,  5.0000],
  [-6.0000, -1.0000,  4.0000],
  [-3.0000, -0.0000, -2.0000]],

 [[-1.0000,  1.0000,  0.0000],
  [ 0.0000, -0.0000, -3.0000],
  [ 1.0000,  0.0000, -1.0000]]],


[[[-3.0000, -0.0000,  3.0000],
  [-5.0000, -2.0000,  3.0000],
  [-5.0000, -1.0000,  3.0000]],

 [[ 7.0000,  2.0000, -2.0000],
  [ 7.0000,  1.0000, -3.0000],
  [ 3.0000, -4.0000, -4.0000]],

 [[ 1.0000,  1.0000, -1.0000],
  [ 0.0000,  1.0000, -1.0000],
  [-0.0000, -0.0000, -1.0000]],

 …,
```

```
          [[-2.0000, -1.0000,  1.0000],
           [-0.0000, -1.0000, -1.0000],
           [-0.0000,  0.0000, -3.0000]],

          [[ 1.0000, -1.0000, -0.0000],
           [-2.0000, -1.0000,  1.0000],
           [-3.0000, -3.0000,  1.0000]],

          [[ 0.0000,  1.0000,  0.0000],
           [-0.0000,  1.0000,  3.0000],
           [-2.0000, -2.0000,  0.0000]]],


         [[[-1.0000,  1.0000,  1.0000],
           [ 1.0000,  2.0000,  2.0000],
           [ 1.0000,  1.0000,  1.0000]],

          [[-4.0000, -2.0000, -2.0000],
           [-0.0000,  1.0000,  2.0000],
           [ 4.0000,  7.0000,  3.0000]],

          [[ 0.0000,  1.0000,  1.0000],
           [ 1.0000,  1.0000,  0.0000],
           [ 0.0000, -1.0000, -0.0000]],

          …,

          [[-2.0000, -1.0000, -0.0000],
           [-1.0000,  1.0000,  1.0000],
           [ 0.0000,  1.0000,  1.0000]],

          [[-0.0000,  0.0000,  1.0000],
           [-0.0000, -2.0000, -0.0000],
           [ 0.0000, -3.0000, -1.0000]],

          [[-3.0000, -1.0000, -0.0000],
           [ 0.0000,  1.0000,  0.0000],
           [-0.0000,  2.0000,  2.0000]]]], device='cuda:0',
        grad_fn=<DivBackward0>)
```

[5]:
```
from saveOutput import SaveOutput

resnet_save_out = SaveOutput()
blocks_res = SaveOutput.hook(resnet20_model_quan.layer1.modules(), QuantConv2d,␣
 ↪resnet_save_out)
```

```
prehooked
prehooked
```

```
prehooked
prehooked
prehooked
prehooked
```

[6]:
```python
def act_quantization(b):

    def uniform_quant(x, b=3):
        xdiv = x.mul(2 ** b - 1)
        xhard = xdiv.round().div(2 ** b - 1)
        return xhard

    class uq(torch.autograd.Function):   # here single underscore means this class is for internal use

        def forward(ctx, input, alpha):
            input_d = input/alpha
            input_c = input_d.clamp(max=1)  # Mingu edited for Alexnet
            input_q = uniform_quant(input_c, b)
            ctx.save_for_backward(input, input_q)
            input_q_out = input_q.mul(alpha)
            return input_q_out

    return uq().apply
```

[9]:
```python
x = resnet_save_out.outputs[0][0]# input of the 2nd conv layer
x_bit = 4

x_alpha  = 3.284
x_delta = x_alpha/(2**x_bit-1)

act_quant_fn = act_quantization(x_bit) # define the quantization function
x_q = act_quant_fn(x, x_alpha)         # create the quantized value for x

x_int = x_q/x_delta
print(x_int) # you should see clean integer numbers
```

```
tensor([[[[ 2.0000,  4.0000,  4.0000,  …,  5.0000,  5.0000,  3.0000],
          [ 2.0000,  4.0000,  4.0000,  …,  4.0000,  4.0000,  3.0000],
          [ 3.0000,  4.0000,  4.0000,  …,  3.0000,  4.0000,  3.0000],
          …,
          [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
          [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
          [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000]],

         [[ 4.0000,  1.0000,  1.0000,  …,  0.0000,  0.0000,  1.0000],
          [10.0000,  7.0000,  8.0000,  …,  7.0000,  6.0000,  4.0000],
          [ 9.0000,  7.0000,  7.0000,  …,  8.0000,  6.0000,  4.0000],
```

```
 …,
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  2.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  1.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000]],

 [[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 …,
  [ 3.0000,  5.0000,  5.0000,  …,  5.0000,  3.0000,  5.0000],
  [ 3.0000,  5.0000,  4.0000,  …,  4.0000,  4.0000,  4.0000],
  [ 3.0000,  4.0000,  4.0000,  …,  4.0000,  4.0000,  3.0000]],

…,

 [[ 3.0000,  1.0000,  1.0000,  …,  2.0000,  3.0000,  0.0000],
  [ 6.0000,  3.0000,  2.0000,  …,  4.0000,  4.0000,  0.0000],
  [ 6.0000,  3.0000,  2.0000,  …,  4.0000,  4.0000,  0.0000],
 …,
  [ 5.0000,  6.0000,  3.0000,  …,  8.0000,  4.0000,  0.0000],
  [ 4.0000,  6.0000,  3.0000,  …,  6.0000,  4.0000,  0.0000],
  [ 2.0000,  1.0000,  1.0000,  …,  2.0000,  0.0000,  0.0000]],

 [[ 1.0000,  2.0000,  2.0000,  …,  1.0000,  1.0000,  4.0000],
  [ 2.0000,  3.0000,  4.0000,  …,  2.0000,  2.0000,  6.0000],
  [ 1.0000,  3.0000,  4.0000,  …,  2.0000,  2.0000,  6.0000],
 …,
  [ 0.0000,  0.0000,  2.0000,  …,  0.0000,  2.0000,  9.0000],
  [ 0.0000,  0.0000,  1.0000,  …,  0.0000,  2.0000,  8.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  5.0000]],

 [[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 …,
  [ 0.0000,  1.0000,  1.0000,  …,  2.0000,  3.0000,  0.0000],
  [ 0.0000,  1.0000,  2.0000,  …,  2.0000,  2.0000,  1.0000],
  [ 0.0000,  0.0000,  1.0000,  …,  2.0000,  1.0000,  1.0000]]],


[[[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 …,
  [ 0.0000,  3.0000,  4.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  2.0000,  3.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000]],
```

```
[[12.0000,  9.0000,  9.0000,  …, 10.0000,  9.0000,  4.0000],
 [ 5.0000,  5.0000,  5.0000,  …,  5.0000,  5.0000,  5.0000],
 [ 5.0000,  4.0000,  4.0000,  …,  4.0000,  4.0000,  5.0000],
 …,
 [ 1.0000,  2.0000,  3.0000,  …,  2.0000,  2.0000,  3.0000],
 [ 1.0000,  2.0000,  3.0000,  …,  2.0000,  2.0000,  3.0000],
 [ 1.0000,  3.0000,  4.0000,  …,  0.0000,  0.0000,  0.0000]],

[[ 6.0000,  6.0000,  6.0000,  …,  6.0000,  6.0000,  4.0000],
 [ 8.0000,  7.0000,  7.0000,  …,  7.0000,  7.0000,  5.0000],
 [ 8.0000,  7.0000,  7.0000,  …,  7.0000,  7.0000,  5.0000],
 …,
 [ 0.0000,  0.0000,  0.0000,  …,  5.0000,  6.0000,  5.0000],
 [ 0.0000,  0.0000,  0.0000,  …,  5.0000,  6.0000,  5.0000],
 [ 0.0000,  0.0000,  0.0000,  …,  3.0000,  4.0000,  3.0000]],

…,

[[ 0.0000,  3.0000,  3.0000,  …,  3.0000,  3.0000, 12.0000],
 [ 0.0000,  3.0000,  3.0000,  …,  3.0000,  3.0000, 14.0000],
 [ 0.0000,  3.0000,  3.0000,  …,  3.0000,  3.0000, 14.0000],
 …,
 [11.0000, 11.0000,  7.0000,  …,  0.0000,  0.0000, 10.0000],
 [11.0000,  9.0000,  5.0000,  …,  0.0000,  1.0000, 10.0000],
 [ 9.0000,  7.0000,  4.0000,  …,  0.0000,  0.0000,  8.0000]],

[[11.0000,  2.0000,  3.0000,  …,  3.0000,  2.0000,  0.0000],
 [12.0000,  2.0000,  2.0000,  …,  2.0000,  2.0000,  0.0000],
 [12.0000,  2.0000,  2.0000,  …,  2.0000,  2.0000,  0.0000],
 …,
 [ 0.0000,  0.0000,  0.0000,  …,  5.0000,  4.0000,  0.0000],
 [ 0.0000,  0.0000,  1.0000,  …,  5.0000,  4.0000,  0.0000],
 [ 0.0000,  0.0000,  1.0000,  …,  5.0000,  4.0000,  0.0000]],

[[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 …,
 [ 2.0000,  5.0000,  6.0000,  …,  0.0000,  0.0000,  0.0000],
 [ 2.0000,  4.0000,  4.0000,  …,  0.0000,  0.0000,  0.0000],
 [ 1.0000,  3.0000,  3.0000,  …,  0.0000,  0.0000,  0.0000]]],


[[[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  …,
  [ 4.0000,  2.0000,  2.0000,  …,  4.0000,  6.0000,  6.0000],
```

```
 [ 3.0000,  2.0000,  2.0000,  …,  5.0000,  6.0000,  5.0000],
 [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000]],

[[ 8.0000,  8.0000,  7.0000,  …, 10.0000, 10.0000,  4.0000],
 [ 0.0000,  1.0000,  1.0000,  …,  5.0000,  5.0000,  4.0000],
 [ 0.0000,  1.0000,  1.0000,  …,  4.0000,  5.0000,  4.0000],
 …,
 [ 1.0000,  2.0000,  2.0000,  …,  2.0000,  1.0000,  3.0000],
 [ 1.0000,  3.0000,  3.0000,  …,  2.0000,  3.0000,  3.0000],
 [ 0.0000,  2.0000,  2.0000,  …,  4.0000,  5.0000,  5.0000]],

[[ 5.0000,  6.0000,  5.0000,  …,  6.0000,  6.0000,  5.0000],
 [ 7.0000,  7.0000,  6.0000,  …,  7.0000,  8.0000,  6.0000],
 [ 7.0000,  7.0000,  6.0000,  …,  7.0000,  8.0000,  6.0000],
 …,
 [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 [ 0.0000,  1.0000,  1.0000,  …,  0.0000,  0.0000,  0.0000]],

…,

[[ 0.0000,  6.0000,  7.0000,  …,  3.0000,  2.0000, 13.0000],
 [ 0.0000,  5.0000,  7.0000,  …,  2.0000,  1.0000, 15.0000],
 [ 0.0000,  4.0000,  7.0000,  …,  2.0000,  1.0000, 15.0000],
 …,
 [13.0000,  3.0000,  3.0000,  …,  5.0000,  7.0000,  0.0000],
 [14.0000,  3.0000,  3.0000,  …,  7.0000,  5.0000,  0.0000],
 [12.0000,  1.0000,  2.0000,  …,  6.0000,  3.0000,  0.0000]],

[[ 8.0000,  1.0000,  1.0000,  …,  3.0000,  3.0000,  0.0000],
 [ 9.0000,  0.0000,  0.0000,  …,  3.0000,  3.0000,  0.0000],
 [ 9.0000,  0.0000,  0.0000,  …,  3.0000,  3.0000,  0.0000],
 …,
 [ 0.0000,  2.0000,  3.0000,  …,  0.0000,  0.0000, 15.0000],
 [ 0.0000,  3.0000,  3.0000,  …,  0.0000,  2.0000, 15.0000],
 [ 0.0000,  1.0000,  2.0000,  …,  0.0000,  1.0000, 11.0000]],

[[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 …,
 [ 8.0000,  4.0000,  4.0000,  …,  5.0000,  7.0000,  3.0000],
 [ 9.0000,  4.0000,  4.0000,  …,  6.0000,  6.0000,  3.0000],
 [ 7.0000,  5.0000,  5.0000,  …, 10.0000,  9.0000,  5.0000]]],


…,
```

```
[[[ 0.0000,  0.0000,  0.0000,  …,  4.0000,  2.0000,  1.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  3.0000,  1.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  3.0000,  1.0000,  0.0000],
  …,
  [ 0.0000,  0.0000,  0.0000,  …,  2.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  2.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000]],

 [[ 7.0000,  3.0000,  4.0000,  …,  0.0000,  0.0000,  1.0000],
  [12.0000,  7.0000,  8.0000,  …, 13.0000,  6.0000,  1.0000],
  [12.0000,  7.0000,  8.0000,  …, 13.0000,  7.0000,  3.0000],
  …,
  [13.0000, 11.0000,  9.0000,  …,  1.0000,  0.0000,  2.0000],
  [10.0000, 10.0000, 10.0000,  …,  5.0000,  4.0000,  3.0000],
  [12.0000, 14.0000, 14.0000,  …,  4.0000,  4.0000,  3.0000]],

 [[ 1.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 1.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 1.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  …,
  [ 2.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  1.0000],
  [ 1.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  1.0000],
  [ 1.0000,  0.0000,  0.0000,  …,  1.0000,  2.0000,  1.0000]],

 …,

 [[ 0.0000,  3.0000,  2.0000,  …,  7.0000,  6.0000,  0.0000],
  [ 0.0000,  5.0000,  5.0000,  …, 11.0000,  8.0000,  0.0000],
  [ 0.0000,  6.0000,  4.0000,  …, 11.0000,  9.0000,  0.0000],
  …,
  [ 0.0000,  2.0000,  6.0000,  …,  5.0000,  7.0000,  1.0000],
  [ 0.0000,  1.0000,  4.0000,  …,  0.0000,  1.0000,  1.0000],
  [ 2.0000,  3.0000,  5.0000,  …,  0.0000,  0.0000,  0.0000]],

 [[ 4.0000,  0.0000,  1.0000,  …,  0.0000,  0.0000,  5.0000],
  [ 5.0000,  1.0000,  1.0000,  …,  0.0000,  0.0000,  6.0000],
  [ 5.0000,  0.0000,  1.0000,  …,  0.0000,  0.0000,  7.0000],
  …,
  [ 9.0000,  5.0000,  1.0000,  …,  3.0000,  1.0000,  5.0000],
  [ 7.0000,  6.0000,  2.0000,  …,  7.0000,  5.0000,  5.0000],
  [ 5.0000,  6.0000,  3.0000,  …,  7.0000,  5.0000,  4.0000]],

 [[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  2.0000,  1.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  2.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  1.0000,  2.0000,  0.0000],
  …,
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
```

```
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  2.0000,  0.0000,  0.0000]]],


[[[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  …,
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000]],

 [[ 7.0000,  7.0000,  7.0000,  …,  7.0000,  7.0000,  5.0000],
  [ 0.0000,  1.0000,  1.0000,  …,  1.0000,  1.0000,  2.0000],
  [ 0.0000,  1.0000,  1.0000,  …,  1.0000,  1.0000,  2.0000],
  …,
  [ 0.0000,  2.0000,  2.0000,  …,  5.0000,  4.0000,  4.0000],
  [ 1.0000,  2.0000,  2.0000,  …,  4.0000,  4.0000,  4.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  2.0000,  2.0000,  3.0000]],

 [[ 4.0000,  5.0000,  5.0000,  …,  5.0000,  5.0000,  4.0000],
  [ 5.0000,  6.0000,  6.0000,  …,  6.0000,  6.0000,  5.0000],
  [ 5.0000,  6.0000,  6.0000,  …,  6.0000,  6.0000,  5.0000],
  …,
  [ 3.0000,  4.0000,  4.0000,  …,  1.0000,  1.0000,  2.0000],
  [ 3.0000,  4.0000,  4.0000,  …,  1.0000,  1.0000,  2.0000],
  [ 3.0000,  3.0000,  3.0000,  …,  2.0000,  2.0000,  2.0000]],

 …,

 [[ 0.0000,  4.0000,  4.0000,  …,  4.0000,  4.0000,  7.0000],
  [ 0.0000,  3.0000,  2.0000,  …,  3.0000,  3.0000,  8.0000],
  [ 0.0000,  2.0000,  2.0000,  …,  3.0000,  3.0000,  8.0000],
  …,
  [ 0.0000,  3.0000,  2.0000,  …,  3.0000,  3.0000,  1.0000],
  [ 1.0000,  3.0000,  2.0000,  …,  3.0000,  2.0000,  0.0000],
  [ 0.0000,  1.0000,  1.0000,  …,  1.0000,  1.0000,  0.0000]],

 [[ 7.0000,  3.0000,  3.0000,  …,  3.0000,  3.0000,  0.0000],
  [ 7.0000,  2.0000,  2.0000,  …,  2.0000,  2.0000,  0.0000],
  [ 7.0000,  2.0000,  2.0000,  …,  2.0000,  2.0000,  0.0000],
  …,
  [ 4.0000,  2.0000,  3.0000,  …,  3.0000,  3.0000,  6.0000],
  [ 4.0000,  3.0000,  3.0000,  …,  3.0000,  3.0000,  6.0000],
  [ 3.0000,  2.0000,  2.0000,  …,  2.0000,  2.0000,  4.0000]],

 [[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
```

```
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 …,
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000]]],


[[[ 2.0000,  3.0000,  3.0000,  …,  6.0000,  7.0000,  6.0000],
  [ 2.0000,  2.0000,  2.0000,  …,  5.0000,  6.0000,  5.0000],
  [ 2.0000,  3.0000,  2.0000,  …,  5.0000,  6.0000,  5.0000],
 …,
  [ 0.0000,  0.0000,  0.0000,  …,  6.0000,  6.0000,  5.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  6.0000,  6.0000,  5.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000]],

 [[ 0.0000,  1.0000,  1.0000,  …,  0.0000,  0.0000,  5.0000],
  [ 5.0000,  6.0000,  6.0000,  …,  3.0000,  3.0000,  3.0000],
  [ 4.0000,  4.0000,  5.0000,  …,  3.0000,  3.0000,  3.0000],
 …,
  [ 6.0000,  5.0000,  5.0000,  …,  3.0000,  3.0000,  3.0000],
  [ 6.0000,  7.0000,  8.0000,  …,  3.0000,  3.0000,  3.0000],
  [ 8.0000,  5.0000,  5.0000,  …,  4.0000,  5.0000,  6.0000]],

 [[ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 0.0000,  0.0000,  0.0000,  …,  0.0000,  0.0000,  0.0000],
 …,
  [ 3.0000,  3.0000,  3.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 3.0000,  2.0000,  2.0000,  …,  0.0000,  0.0000,  0.0000],
  [ 2.0000,  1.0000,  1.0000,  …,  0.0000,  0.0000,  0.0000]],

 …,

 [[ 9.0000,  3.0000,  3.0000,  …,  5.0000,  4.0000,  0.0000],
  [11.0000,  4.0000,  3.0000,  …,  5.0000,  4.0000,  0.0000],
  [11.0000,  5.0000,  2.0000,  …,  6.0000,  4.0000,  0.0000],
 …,
  [ 0.0000,  3.0000,  3.0000,  …,  6.0000,  4.0000,  0.0000],
  [ 0.0000,  2.0000,  2.0000,  …,  5.0000,  3.0000,  0.0000],
  [ 1.0000,  4.0000,  4.0000,  …,  4.0000,  3.0000,  0.0000]],

 [[ 0.0000,  2.0000,  2.0000,  …,  1.0000,  2.0000, 12.0000],
  [ 0.0000,  2.0000,  4.0000,  …,  1.0000,  2.0000, 15.0000],
  [ 0.0000,  2.0000,  4.0000,  …,  0.0000,  2.0000, 15.0000],
 …,
  [ 6.0000,  3.0000,  2.0000,  …,  1.0000,  3.0000, 15.0000],
  [ 5.0000,  3.0000,  3.0000,  …,  1.0000,  3.0000, 15.0000],
  [ 4.0000,  3.0000,  2.0000,  …,  0.0000,  1.0000, 11.0000]],
```

```
          [[ 6.0000,  3.0000,  4.0000,  …,  9.0000,  9.0000,  3.0000],
           [ 4.0000,  2.0000,  1.0000,  …,  7.0000,  7.0000,  4.0000],
           [ 4.0000,  3.0000,  2.0000,  …,  7.0000,  7.0000,  4.0000],
           …,
           [ 0.0000,  0.0000,  0.0000,  …,  7.0000,  7.0000,  4.0000],
           [ 0.0000,  0.0000,  0.0000,  …,  7.0000,  7.0000,  4.0000],
           [ 0.0000,  0.0000,  0.0000,  …, 10.0000, 10.0000,  6.0000]]]],
       device='cuda:0')
```

[10]:
```
conv_quan = nn.Conv2d(16, 16, kernel_size=(3, 3), padding=(1, 1), bias=False)
conv_quan.weight = torch.nn.Parameter(weight_int)

psum_int = conv_quan(x_int)
print("psum int:", psum_int)
psum_recovered = psum_int*w_delta*x_delta
```

```
psum int: tensor([[[[ 3.6000e+01,  1.5300e+02,  1.7800e+02,  …,  1.7400e+02,
            1.1400e+02,  5.1000e+01],
          [-8.8000e+01,  1.6000e+01,  3.4000e+01,  …,  1.5000e+01,
           -9.0000e+00, -1.6000e+01],
          [-1.1100e+02,  7.0000e+00, -5.0000e+00,  …,  2.0000e+00,
            1.3000e+01,  2.0000e+01],
          …,
          [ 2.4800e+02,  2.1200e+02,  2.1300e+02,  …,  1.6900e+02,
            1.3900e+02, -3.3000e+01],
          [ 1.7400e+02,  1.8600e+02,  1.8900e+02,  …,  1.5000e+02,
            1.2800e+02, -3.8000e+01],
          [ 1.0600e+02,  1.2200e+02,  1.2800e+02,  …,  1.4100e+02,
            9.7000e+01,  2.6000e+01]],

         [[ 1.3500e+02,  2.2400e+02,  2.6300e+02,  …,  2.4900e+02,
            2.8600e+02,  2.7000e+01],
          [-3.5000e+01,  2.5000e+01,  4.4000e+01,  …,  9.0000e+00,
            1.3000e+01, -1.0600e+02],
          [ 2.7000e+01,  6.5000e+01,  6.5000e+01,  …,  1.0000e+01,
            3.5000e+01, -7.3000e+01],
          …,
          [ 1.1400e+02, -9.3000e+01, -5.2000e+01,  …,  5.8000e+01,
           -2.1500e+02, -8.2000e+01],
          [ 1.6200e+02,  3.5000e+01,  5.0000e+00,  …,  3.0000e+00,
            1.5900e+02, -1.8200e+02],
          [ 1.6400e+02,  2.4600e+02,  1.9900e+02,  …,  1.5300e+02,
            4.2700e+02, -1.5100e+02]],

         [[ 1.6600e+02,  1.4400e+02,  4.1000e+01,  …,  9.2000e+01,
            1.4600e+02,  3.0000e+01],
          [ 9.2000e+01,  5.1000e+01, -4.7000e+01,  …, -7.3000e+01,
```

```
   1.6000e+01, -1.1800e+02],
 [ 7.9000e+01,  4.0000e+00, -3.8000e+01,  …, -8.2000e+01,
  -4.0000e+00, -7.8000e+01],
 …,
 [-2.1800e+02,  3.3000e+01,  1.4200e+02,  …,  8.8000e+01,
  -4.0000e+00,  1.8300e+02],
 [ 4.0000e+00,  1.6000e+01,  1.5500e+02,  …,  2.3900e+02,
  -7.8000e+01,  2.3600e+02],
 [ 1.2400e+02,  1.3000e+02,  1.4300e+02,  …,  1.3400e+02,
   1.9000e+01,  1.0200e+02]],

…,

[[ 2.6800e+02,  2.7600e+02,  2.5900e+02,  …,  2.2900e+02,
   2.3600e+02,  1.3200e+02],
 [ 4.8000e+01,  1.0400e+02,  9.8000e+01,  …,  1.4100e+02,
   2.0100e+02,  9.5000e+01],
 [ 1.7400e+02,  2.5100e+02,  2.5500e+02,  …,  2.8700e+02,
   2.4400e+02,  2.0900e+02],
 …,
 [ 5.6000e+01,  4.0200e+02,  1.9600e+02,  …,  1.4900e+02,
   2.9900e+02, -2.0000e+00],
 [ 2.1000e+02,  3.7900e+02,  2.9400e+02,  …,  3.8500e+02,
   2.1100e+02,  1.4100e+02],
 [ 2.0200e+02,  2.5400e+02,  1.5300e+02,  …,  2.4600e+02,
   2.7500e+02,  2.8000e+01]],

[[-1.7600e+02, -3.2300e+02, -3.9600e+02,  …, -3.3900e+02,
  -3.7400e+02, -3.6800e+02],
 [-1.1200e+02, -3.0300e+02, -4.3700e+02,  …, -3.3200e+02,
  -4.1500e+02, -4.4800e+02],
 [-9.5000e+01, -2.4800e+02, -3.8400e+02,  …, -2.2800e+02,
  -3.2500e+02, -4.1600e+02],
 …,
 [-3.6700e+02, -2.0100e+02, -2.1500e+02,  …, -6.6000e+01,
  -1.3900e+02, -5.0000e+00],
 [-3.6300e+02, -2.2400e+02, -2.7700e+02,  …, -1.4800e+02,
  -1.5300e+02, -6.5000e+01],
 [-3.8200e+02, -3.2500e+02, -3.6200e+02,  …, -2.8200e+02,
  -2.6400e+02, -1.4100e+02]],

[[-7.4000e+01, -1.8900e+02, -2.1100e+02,  …, -1.6900e+02,
  -1.7500e+02, -3.5000e+01],
 [-1.7300e+02, -3.7800e+02, -3.9100e+02,  …, -3.1900e+02,
  -3.0700e+02, -2.3200e+02],
 [-1.2400e+02, -3.5900e+02, -3.5100e+02,  …, -2.8400e+02,
  -2.8900e+02, -1.9500e+02],
 …,
```

```
  [-1.2800e+02, -1.0100e+02, -7.6000e+01,  …, -1.1200e+02,
   -1.2800e+02, -7.1000e+01],
  [-7.1000e+01, -1.3000e+02, -9.5000e+01,  …, -3.8000e+01,
   -2.9500e+02, -1.6000e+01],
  [ 7.0000e+00, -6.1000e+01, -5.3000e+01,  …, -2.7000e+01,
   -1.0800e+02,  6.0000e+00]]],


[[[ 4.6000e+01,  1.1000e+01, -6.1000e+01,  …, -5.5000e+01,
   -1.4000e+01, -6.2000e+01],
  [ 2.6000e+01,  7.1000e+01, -2.4000e+01,  …, -2.2000e+01,
   -2.4000e+01, -9.7000e+01],
  [ 6.9000e+01,  1.3300e+02,  5.2000e+01,  …,  4.8000e+01,
    2.9000e+01, -6.4000e+01],
  …,
  [ 9.0000e+00, -5.2000e+01,  3.4000e+01,  …,  1.1700e+02,
    1.3600e+02,  5.0000e+00],
  [-1.2000e+01, -5.8000e+01,  1.6000e+01,  …,  1.8700e+02,
    2.1800e+02,  7.9000e+01],
  [-2.3000e+01, -3.5000e+01,  8.5831e-06,  …, -3.1000e+01,
   -8.0000e+00, -8.7000e+01]],

 [[-1.2900e+02, -2.4400e+02, -1.7100e+02,  …, -1.8000e+02,
   -1.8100e+02, -1.6200e+02],
  [-6.3000e+01, -3.0500e+02, -1.1000e+02,  …, -1.1600e+02,
   -9.6000e+01, -8.5000e+01],
  [-3.5000e+01, -2.8400e+02, -7.6000e+01,  …, -7.7000e+01,
   -5.4000e+01, -8.3000e+01],
  …,
  [ 2.1500e+02, -3.0000e+00,  5.5000e+01,  …, -9.2000e+01,
   -3.7000e+01, -8.4000e+01],
  [ 1.9900e+02, -6.2000e+01,  1.2700e+02,  …, -1.0100e+02,
   -3.1000e+01, -1.1100e+02],
  [ 1.5200e+02,  1.9000e+01,  1.3800e+02,  …, -1.3000e+01,
   -9.0000e+00,  5.2000e+01]],

 [[-3.7600e+02, -1.8200e+02, -1.5800e+02,  …, -1.6400e+02,
   -2.4900e+02, -2.1400e+02],
  [-3.2700e+02, -9.4000e+01, -5.5000e+01,  …, -5.5000e+01,
   -1.2400e+02, -1.0400e+02],
  [-2.8800e+02, -6.6000e+01, -1.9000e+01,  …, -8.0000e+00,
   -1.0400e+02, -8.9000e+01],
  …,
  [-8.1000e+01,  1.4700e+02, -2.4000e+01,  …, -1.4000e+02,
   -1.9200e+02, -5.7000e+01],
  [-3.8000e+01,  7.2000e+01, -5.7000e+01,  …, -1.9600e+02,
   -2.7100e+02, -1.4100e+02],
  [-9.4000e+01, -1.3600e+02, -1.8300e+02,  …,  2.3500e+02,
```

```
    2.1700e+02,  1.1900e+02]],

 …,

 [[ 1.1400e+02,  5.0900e+02,  4.0900e+02,  …,  3.9700e+02,
    3.4900e+02,  9.2000e+01],
  [ 6.7000e+01,  4.0800e+02,  3.5000e+02,  …,  3.4300e+02,
    2.5000e+02,  2.2600e+02],
  [ 1.0000e+02,  3.5000e+02,  3.2800e+02,  …,  3.2200e+02,
    2.0900e+02,  1.8800e+02],
  …,
  [ 1.5000e+01,  3.0300e+02,  1.4900e+02,  …,  2.4900e+02,
    1.4300e+02,  8.2000e+01],
  [ 1.4000e+02,  2.8000e+02,  1.7400e+02,  …,  2.3200e+02,
    1.4900e+02,  1.1300e+02],
  [ 1.3300e+02,  2.1100e+02,  1.2500e+02,  …,  8.3000e+01,
    1.8000e+01,  2.6000e+01]],

 [[-2.9800e+02, -2.6000e+02, -2.2400e+02,  …, -2.3200e+02,
   -1.7000e+02, -2.0000e+01],
  [-4.8600e+02, -4.4800e+02, -4.2200e+02,  …, -4.2700e+02,
   -3.5800e+02, -1.2000e+02],
  [-3.8100e+02, -3.4000e+02, -3.0300e+02,  …, -3.0100e+02,
   -2.6600e+02, -4.7000e+01],
  …,
  [-9.1000e+01, -1.4200e+02, -2.5600e+02,  …, -2.3400e+02,
   -2.3900e+02, -7.4000e+01],
  [-5.2000e+01, -1.2500e+02, -2.2200e+02,  …, -2.2600e+02,
   -1.9700e+02, -7.9000e+01],
  [-2.7000e+01, -8.9000e+01, -1.2300e+02,  …, -3.4200e+02,
   -2.8600e+02, -1.0600e+02]],

 [[-2.4900e+02, -1.4800e+02, -1.5100e+02,  …, -1.5900e+02,
   -1.0900e+02, -1.0000e+02],
  [-2.6800e+02, -2.6100e+02, -2.6700e+02,  …, -2.7100e+02,
   -1.7400e+02, -1.2400e+02],
  [-1.9000e+02, -1.5600e+02, -1.7500e+02,  …, -1.7600e+02,
   -8.7000e+01, -7.9000e+01],
  …,
  [-8.0000e+01, -9.0000e+01, -1.7300e+02,  …, -2.1400e+02,
   -1.5400e+02, -7.1000e+01],
  [-9.1000e+01, -1.9200e+02, -2.6400e+02,  …, -2.3600e+02,
   -1.5500e+02, -8.8000e+01],
  [-7.4000e+01, -1.8200e+02, -2.3000e+02,  …, -1.4200e+02,
   -8.5000e+01, -1.0100e+02]]],


[[[ 1.2100e+02,  6.5000e+01, -6.5000e+01,  …, -1.0900e+02,
```

```
   6.0000e+00, -5.8000e+01],
 [ 1.6900e+02,  1.7000e+02,  2.3000e+01,  …, -7.2000e+01,
   3.7000e+01, -7.7000e+01],
 [ 1.8500e+02,  1.8300e+02,  2.4000e+01,  …, -1.9000e+01,
   7.5000e+01, -4.9000e+01],
 …,
 [-6.7000e+01,  4.6000e+01,  5.5000e+01,  …,  6.1000e+01,
  -9.9000e+01, -1.3000e+02],
 [-6.3000e+01,  2.3000e+01,  3.9000e+01,  …, -1.1000e+01,
  -9.7000e+01, -1.5400e+02],
 [ 7.0000e+00,  6.7000e+01,  8.5000e+01,  …,  6.7000e+01,
  -1.6000e+01, -4.3000e+01]],

[[-1.8000e+02, -2.2800e+02, -2.9100e+02,  …, -2.3900e+02,
  -1.3900e+02, -1.5700e+02],
 [-1.2000e+01, -1.0100e+02, -1.3600e+02,  …, -2.2100e+02,
  -1.9000e+01, -9.2000e+01],
 [ 2.0000e+00, -1.3300e+02, -1.1100e+02,  …, -2.0700e+02,
   6.6000e+01, -1.1600e+02],
 …,
 [-5.7000e+01,  1.1900e+02, -6.0000e+00,  …,  1.9400e+02,
   9.8000e+01, -3.3300e+02],
 [-4.8000e+01,  1.8300e+02,  8.3000e+01,  …,  1.8200e+02,
   1.0700e+02, -2.6900e+02],
 [-6.0000e+00,  2.5200e+02,  1.5900e+02,  …,  5.4000e+01,
   1.5000e+02, -2.6600e+02]],

[[-4.0100e+02, -2.9600e+02, -4.8000e+01,  …, -8.9000e+01,
  -3.3800e+02, -2.3100e+02],
 [-3.0400e+02, -1.2200e+02,  7.3000e+01,  …,  4.2000e+01,
  -2.3600e+02, -9.0000e+01],
 [-2.3500e+02, -6.6000e+01,  1.2100e+02,  …,  1.2700e+02,
  -2.0300e+02, -7.1000e+01],
 …,
 [ 1.6000e+02, -6.2000e+01, -4.2000e+01,  …,  5.7000e+01,
   1.5600e+02,  8.0000e+01],
 [ 2.2400e+02,  2.9000e+01,  6.0000e+00,  …,  2.7200e+02,
   1.3700e+02,  6.5000e+01],
 [-6.8000e+01, -1.3300e+02, -1.3300e+02,  …, -1.0100e+02,
  -2.2500e+02, -1.1300e+02]],

…,

[[ 5.5000e+01,  2.6600e+02,  3.0700e+02,  …,  4.1200e+02,
   3.1100e+02,  1.1200e+02],
 [ 8.7000e+01,  3.3200e+02,  3.4500e+02,  …,  3.6600e+02,
   1.7800e+02,  2.5600e+02],
 [ 1.0000e+02,  2.5200e+02,  3.1400e+02,  …,  3.8900e+02,
```

```
    9.4000e+01,  2.2800e+02],
  …,
  [ 1.2100e+02,  1.0300e+02,  1.4300e+02,  …,  2.4000e+01,
    4.0300e+02,  1.2200e+02],
  [ 2.0900e+02,  1.4300e+02,  1.6000e+02,  …,  1.8800e+02,
    3.8000e+02,  1.0200e+02],
  [ 1.8300e+02,  1.8100e+02,  1.1500e+02,  …,  1.6700e+02,
    3.0000e+02,  9.8000e+01]],

 [[-1.7600e+02, -5.1000e+01, -5.8000e+01,  …, -2.3500e+02,
   -1.7100e+02, -6.0000e+00],
  [-4.1400e+02, -2.9400e+02, -2.9200e+02,  …, -4.4300e+02,
   -3.6800e+02, -1.0300e+02],
  [-3.8700e+02, -3.1500e+02, -2.8200e+02,  …, -3.2800e+02,
   -2.7500e+02, -2.6000e+01],
  …,
  [-2.1500e+02, -2.7300e+02, -2.2200e+02,  …, -1.7300e+02,
   -2.5200e+02, -2.6300e+02],
  [-1.7800e+02, -2.7500e+02, -2.1000e+02,  …, -2.3400e+02,
   -2.8900e+02, -2.7300e+02],
  [-1.3500e+02, -2.3700e+02, -1.7900e+02,  …, -1.4400e+02,
   -1.9200e+02, -1.7900e+02]],

 [[-1.7900e+02, -1.4300e+02, -1.1300e+02,  …, -1.2800e+02,
   -1.1800e+02, -9.9000e+01],
  [-2.0600e+02, -1.4800e+02, -1.9300e+02,  …, -2.5300e+02,
   -1.9200e+02, -1.2500e+02],
  [-1.5100e+02, -7.9000e+01, -1.4000e+02,  …, -1.5100e+02,
   -1.0700e+02, -5.7000e+01],
  …,
  [ 4.7000e+01, -1.3900e+02, -5.9000e+01,  …, -2.9000e+01,
   -1.4000e+01,  7.0000e+00],
  [-1.7000e+01, -2.4000e+02, -1.7400e+02,  …, -1.9000e+01,
   -1.6000e+02, -9.3000e+01],
  [-4.1000e+01, -2.0100e+02, -1.5600e+02,  …, -1.0800e+02,
   -2.2700e+02, -8.9000e+01]]],


…,


[[[ 8.0000e+01,  1.0400e+02,  2.3000e+01,  …,  6.4000e+01,
    4.3000e+01, -3.3000e+01],
  [-2.7000e+01, -3.0000e+00, -9.7000e+01,  …, -1.2800e+02,
   -1.1600e+02, -7.9000e+01],
  [-3.3000e+01,  2.0000e+00, -1.0700e+02,  …, -2.0100e+02,
   -1.1500e+02, -8.2000e+01],
  …,
```

```
 [-1.8100e+02,  3.2000e+01,  3.8000e+01,  …, -1.3000e+01,
   2.0000e+01, -4.8000e+01],
 [-2.3700e+02, -1.0100e+02, -3.8000e+01,  …, -1.7600e+02,
  -4.2000e+01, -5.7000e+01],
 [-2.4800e+02, -2.4200e+02, -2.3200e+02,  …, -8.7000e+01,
  -6.0000e+01, -3.9000e+01]],

[[ 1.1800e+02,  1.6000e+01,  3.0000e+01,  …,  1.1400e+02,
   1.8600e+02, -2.0900e+02],
 [ 7.0000e+00, -1.0600e+02, -3.8000e+01,  …, -8.1000e+01,
   4.6000e+01, -2.9100e+02],
 [ 8.3000e+01, -6.6000e+01, -1.9000e+01,  …, -9.0000e+00,
  -6.8000e+01, -2.4400e+02],
 …,
 [-9.4000e+01, -9.8000e+01,  7.1000e+01,  …,  1.8000e+01,
  -6.3000e+01, -1.0400e+02],
 [-1.4600e+02, -2.0100e+02,  1.9000e+01,  …, -1.3000e+01,
  -5.2000e+01, -1.4100e+02],
 [-1.0900e+02, -1.8800e+02,  2.7000e+01,  …, -4.3000e+01,
   1.6600e+02, -1.3300e+02]],

[[-1.5800e+02, -5.3000e+01, -6.8000e+01,  …,  3.9000e+01,
  -9.6000e+01,  7.8000e+01],
 [-1.3500e+02, -3.4000e+01, -8.3000e+01,  …, -1.2700e+02,
  -1.1600e+02, -1.0000e+02],
 [-1.7600e+02, -4.9000e+01, -1.3000e+02,  …, -9.6000e+01,
  -7.0000e+01, -2.2000e+01],
 …,
 [-5.9000e+01, -1.6100e+02, -3.0400e+02,  …, -2.7900e+02,
  -1.8700e+02, -7.0000e+00],
 [-9.2000e+01, -1.2500e+02, -3.1300e+02,  …,  4.1000e+01,
  -3.2200e+02, -5.7000e+01],
 [-7.0000e+00,  2.9000e+01, -1.8000e+01,  …, -1.3300e+02,
  -1.5600e+02, -2.2000e+01]],

…,

[[ 1.8800e+02,  4.6300e+02,  3.4200e+02,  …,  3.0800e+02,
   3.1600e+02,  3.6300e+02],
 [-9.4000e+01,  1.4600e+02,  1.7900e+02,  …, -9.0000e+00,
   5.8600e+02,  4.4600e+02],
 [-1.2000e+01,  3.0500e+02,  3.2700e+02,  …,  6.8000e+01,
   6.2400e+02,  4.9900e+02],
 …,
 [ 8.6000e+01,  6.0000e+01,  1.7300e+02,  …, -4.6000e+01,
   4.2400e+02,  1.5800e+02],
 [ 9.0000e+01,  5.8000e+01,  8.1000e+01,  …,  9.9000e+01,
   1.0900e+02,  2.0400e+02],
```

```
 [-2.5000e+01, -6.6000e+01, -6.2000e+01,  …,  9.8000e+01,
   7.5000e+01,  7.3000e+01]],

[[-3.3900e+02, -3.7900e+02, -3.3300e+02,  …, -6.0300e+02,
  -2.7000e+02, -5.5000e+01],
 [-4.1800e+02, -4.5000e+02, -4.5100e+02,  …, -8.3000e+02,
  -3.9900e+02, -6.2000e+01],
 [-3.2000e+02, -3.1200e+02, -3.1700e+02,  …, -6.3300e+02,
  -3.1600e+02, -5.8000e+01],
 …,
 [-1.5400e+02, -2.1900e+02, -3.4900e+02,  …, -3.9200e+02,
  -4.0800e+02, -1.5600e+02],
 [-1.8800e+02, -2.1700e+02, -3.3500e+02,  …, -3.6300e+02,
  -3.0400e+02, -1.5800e+02],
 [-1.5000e+02, -1.3400e+02, -1.5300e+02,  …, -2.7000e+02,
  -1.9800e+02, -1.0100e+02]],

[[-1.5800e+02, -1.2400e+02, -1.3600e+02,  …,  4.0000e+01,
  -7.8000e+01,  4.7000e+01],
 [-1.8600e+02, -3.0400e+02, -2.5300e+02,  …, -1.4000e+02,
  -1.8700e+02, -9.5000e+01],
 [-1.4600e+02, -2.5100e+02, -2.5200e+02,  …, -1.9400e+02,
  -1.6100e+02, -9.2000e+01],
 …,
 [-1.4100e+02, -2.9200e+02, -2.7200e+02,  …, -2.8300e+02,
  -1.2100e+02, -6.4000e+01],
 [-4.4000e+01, -1.2600e+02, -1.4300e+02,  …, -1.2900e+02,
  -2.6500e+02, -8.1000e+01],
 [-4.8000e+01, -1.5700e+02, -2.1700e+02,  …, -8.4000e+01,
  -1.8200e+02, -6.4000e+01]]],


[[[ 1.1300e+02,  7.4000e+01,  5.5000e+01,  …,  5.5000e+01,
    5.8000e+01, -1.3000e+01],
  [ 1.6200e+02,  1.6200e+02,  1.5500e+02,  …,  1.4400e+02,
    1.3000e+02,  4.0000e+00],
  [ 1.8700e+02,  1.9000e+02,  1.8400e+02,  …,  1.4700e+02,
    1.1300e+02, -2.4000e+01],
  …,
  [ 9.9000e+01,  7.2000e+01,  7.9000e+01,  …, -7.3000e+01,
   -1.0600e+02, -5.5000e+01],
  [ 1.1200e+02,  8.5000e+01,  9.5000e+01,  …, -8.3000e+01,
   -8.2000e+01, -1.0300e+02],
  [ 8.0000e+00, -3.0000e+00,  7.0000e+00,  …, -3.2000e+01,
   -4.4000e+01, -3.8000e+01]],

 [[-1.2300e+02, -2.4900e+02, -2.2600e+02,  …, -2.3600e+02,
   -2.6200e+02, -1.8100e+02],
```

```
 [-9.0000e+00, -1.8100e+02, -9.8000e+01,  …, -1.0000e+02,
  -9.9000e+01, -1.3100e+02],
 [-7.0000e+00, -2.0500e+02, -6.6000e+01,  …, -1.0600e+02,
  -6.9000e+01, -1.4300e+02],
 …,
 [ 5.0000e+00, -1.4800e+02, -8.3000e+01,  …, -9.2000e+01,
  -1.1700e+02, -1.9000e+02],
 [ 5.0000e+00, -1.0900e+02, -3.5000e+01,  …,  1.3000e+01,
  -9.7000e+01, -1.0600e+02],
 [ 6.7000e+01,  4.4000e+01,  8.7000e+01,  …,  1.2300e+02,
   1.3900e+02, -5.3000e+01]],

[[-2.8800e+02, -1.7400e+02, -1.4700e+02,  …, -1.3700e+02,
  -1.9100e+02, -8.9000e+01],
 [-2.2700e+02, -5.8000e+01, -4.8000e+01,  …, -1.9000e+01,
  -9.3000e+01,  5.2000e+01],
 [-1.8900e+02, -2.1000e+01, -6.0000e+00,  …, -1.8120e-05,
  -8.3000e+01,  4.2000e+01],
 …,
 [-9.4000e+01, -3.2000e+01, -5.6000e+01,  …, -1.1100e+02,
  -6.0000e+01, -5.9000e+01],
 [-1.1700e+02, -7.5000e+01, -6.2000e+01,  …, -3.7000e+01,
  -6.0000e+01,  1.8000e+01],
 [ 9.8000e+01,  1.5200e+02,  1.6400e+02,  …,  4.0000e+00,
  -2.3000e+01,  7.8000e+01]],

…,

[[ 5.8000e+01,  2.5100e+02,  1.9800e+02,  …,  1.9900e+02,
   1.7600e+02,  6.0000e+00],
 [ 1.1600e+02,  3.0000e+02,  2.4000e+02,  …,  2.3900e+02,
   2.2300e+02,  1.3400e+02],
 [ 1.3500e+02,  2.8100e+02,  1.9500e+02,  …,  2.2300e+02,
   2.0400e+02,  9.2000e+01],
 …,
 [ 1.0000e+02,  2.3000e+02,  1.9500e+02,  …,  1.1400e+02,
   2.7200e+02,  1.1000e+02],
 [ 1.5500e+02,  2.5900e+02,  2.0700e+02,  …,  1.8100e+02,
   3.2100e+02,  9.1000e+01],
 [ 8.5000e+01,  1.4800e+02,  1.1600e+02,  …,  9.0000e+01,
   1.5000e+02,  9.8000e+01]],

[[-1.2500e+02, -6.3000e+01, -4.4000e+01,  …, -3.7000e+01,
  -1.3000e+01,  7.3000e+01],
 [-2.9800e+02, -2.3800e+02, -2.1700e+02,  …, -2.1000e+02,
  -1.8600e+02,  9.9997e-01],
 [-2.6600e+02, -2.4300e+02, -2.1800e+02,  …, -2.0200e+02,
  -1.8300e+02, -6.0000e+00],
```

```
  …,
   [-2.8200e+02, -2.4900e+02, -2.1900e+02,  …, -2.0900e+02,
    -1.8900e+02, -3.3000e+01],
   [-2.6200e+02, -2.2100e+02, -2.1300e+02,  …, -2.1100e+02,
    -1.7700e+02, -5.0000e+00],
   [-2.9300e+02, -3.0400e+02, -2.9600e+02,  …, -2.4500e+02,
    -2.3100e+02, -8.4000e+01]],

  [[-2.0400e+02, -1.7100e+02, -1.8100e+02,  …, -1.8300e+02,
    -1.5700e+02, -9.9000e+01],
   [-2.0200e+02, -2.0500e+02, -2.2900e+02,  …, -2.1500e+02,
    -1.9700e+02, -8.1000e+01],
   [-1.7400e+02, -1.3700e+02, -1.6800e+02,  …, -1.2300e+02,
    -1.2800e+02, -3.4000e+01],
   …,
   [-1.0400e+02, -9.6000e+01, -1.2700e+02,  …, -9.6000e+01,
    -6.6000e+01, -3.3000e+01],
   [-1.0800e+02, -1.4600e+02, -1.5900e+02,  …, -1.0900e+02,
    -1.3900e+02, -7.4000e+01],
   [-5.6000e+01, -8.2000e+01, -9.8000e+01,  …, -7.6000e+01,
    -1.1700e+02, -3.0000e+00]]],


[[[-4.5000e+01, -1.1000e+01,  1.5000e+01,  …,  7.6000e+01,
    -3.6000e+01, -1.0900e+02],
   [-1.1900e+02, -1.4100e+02, -8.8000e+01,  …, -2.0000e+00,
    -1.2200e+02, -1.5100e+02],
   [-1.5200e+02, -1.0300e+02, -1.0100e+02,  …,  2.7000e+01,
    -7.1000e+01, -1.1800e+02],
   …,
   [-1.3200e+02, -6.9000e+01, -4.5000e+01,  …,  4.2000e+01,
    -5.9000e+01, -1.2300e+02],
   [-8.6000e+01, -4.5000e+01, -1.0200e+02,  …,  3.1000e+01,
    -7.2000e+01, -1.4600e+02],
   [-1.3100e+02, -1.5700e+02, -1.5900e+02,  …,  9.6000e+01,
     9.0000e+00, -3.3000e+01]],

  [[-1.0000e+01,  1.1300e+02,  3.0000e+01,  …,  1.3900e+02,
     2.3100e+02, -1.4800e+02],
   [-7.3000e+01,  5.7000e+01, -1.0200e+02,  …,  3.8000e+01,
     1.3000e+02, -3.6400e+02],
   [ 4.5000e+01,  2.6000e+01, -2.5000e+01,  …,  8.5000e+01,
     1.4400e+02, -3.0900e+02],
   …,
   [ 5.0000e+00, -1.8700e+02,  4.5000e+01,  …,  6.4000e+01,
     1.6800e+02, -3.1100e+02],
   [-2.5000e+01, -2.1100e+02, -8.7000e+01,  …,  1.0800e+02,
     2.0000e+02, -3.0000e+02],
```

```
 [-5.6000e+01,  7.0000e+00, -3.2000e+01,  …,  1.0400e+02,
   1.6000e+02, -3.0600e+02]],

[[ 1.5400e+02,  1.1000e+01,  8.0000e+00,  …,  1.4700e+02,
   1.8900e+02,  2.1500e+02],
 [-2.3000e+01, -3.9000e+01, -1.5400e+02,  …, -2.9000e+01,
  -5.9000e+01, -2.0000e+00],
 [ 6.7000e+01, -9.0000e+00, -2.6000e+01,  …,  4.6000e+01,
   1.1000e+01,  4.0000e+01],
 …,
 [-3.9000e+01, -2.4000e+01, -1.6600e+02,  …,  4.8000e+01,
  -3.7000e+01,  3.9000e+01],
 [-1.1300e+02, -9.4000e+01, -3.9000e+01,  …,  1.5700e+02,
   5.0000e+01,  7.5000e+01],
 [ 5.2000e+01,  5.3000e+01,  1.5000e+02,  …, -1.7100e+02,
  -2.4800e+02, -1.1600e+02]],

…,

[[ 1.9800e+02,  7.0000e+00,  1.4800e+02,  …, -1.0000e+01,
   7.9000e+01,  5.2000e+01],
 [ 1.0500e+02,  2.4900e+02,  2.0500e+02,  …,  1.9900e+02,
   3.5700e+02,  1.0200e+02],
 [ 8.1000e+01,  2.5900e+02,  1.9300e+02,  …,  1.9000e+02,
   3.6100e+02,  1.4300e+02],
 …,
 [ 3.4000e+01,  2.0100e+02,  2.5100e+02,  …,  2.0900e+02,
   3.5200e+02,  1.3900e+02],
 [ 9.1000e+01,  2.3800e+02,  2.0200e+02,  …,  2.1800e+02,
   3.6000e+02,  1.1900e+02],
 [-1.0600e+02, -6.1000e+01,  5.6000e+01,  …,  1.6900e+02,
   3.0700e+02,  8.7000e+01]],

[[-1.1700e+02, -1.5300e+02, -1.4100e+02,  …, -6.4000e+01,
  -1.6800e+02, -2.1500e+02],
 [-2.0500e+02, -2.5800e+02, -2.4700e+02,  …, -1.8100e+02,
  -2.8600e+02, -2.5900e+02],
 [-2.2000e+02, -2.8000e+02, -2.9800e+02,  …, -1.6300e+02,
  -2.2600e+02, -2.1600e+02],
 …,
 [-3.9300e+02, -3.1700e+02, -3.2700e+02,  …, -1.9200e+02,
  -2.8100e+02, -2.2900e+02],
 [-3.5900e+02, -2.9800e+02, -2.5200e+02,  …, -2.4200e+02,
  -3.0600e+02, -2.5100e+02],
 [-2.7400e+02, -2.4800e+02, -2.5800e+02,  …, -1.3700e+02,
  -1.9300e+02, -1.5500e+02]],

[[ 1.4600e+02, -2.2000e+01,  1.8000e+01,  …,  9.4000e+01,
```

```
          1.2000e+01,  1.1700e+02],
         [ 9.9999e-01, -1.2100e+02, -1.0600e+02,  …, -2.1935e-05,
          -9.1000e+01, -9.0000e+00],
         [-2.9000e+01, -1.3400e+02, -1.3900e+02,  …, -2.5000e+01,
          -1.0700e+02, -2.0000e+01],
         …,
         [-7.1000e+01, -8.2000e+01, -9.9000e+01,  …, -1.8000e+01,
          -1.0500e+02, -1.1000e+01],
         [ 2.0000e+00, -9.8000e+01, -1.3700e+02,  …, -1.1900e+02,
          -2.1000e+02, -6.7000e+01],
         [-5.3000e+01, -1.7800e+02, -8.4000e+01,  …, -1.7900e+02,
          -2.2100e+02, -8.6000e+01]]]], device='cuda:0',
       grad_fn=<ConvolutionBackward0>)
```

[11]:
```
conv_unquan = nn.Conv2d(16, 16, kernel_size=(3, 3), padding=(1, 1), bias=False)
conv_unquan.weight = torch.nn.Parameter(weight_q)
out = conv_unquan(x)
```

[12]:
```
diff = abs(out - psum_recovered)
print("gap", diff)
```

```
gap tensor([[[[5.1659e-01, 6.7050e-01, 9.4139e-01,  …, 1.8837e-01,
          5.8331e-01, 2.4019e-01],
         [1.1429e-01, 2.0802e-01, 3.8295e-01,  …, 6.7232e-01,
          2.7058e-01, 7.9369e-01],
         [1.9863e-01, 3.1256e-02, 2.1388e-01,  …, 1.1955e+00,
          4.6978e-01, 5.2290e-01],
         …,
         [3.4516e-01, 6.7371e-01, 8.7989e-01,  …, 6.0625e-01,
          2.7849e-01, 5.2161e-01],
         [4.2660e-01, 4.7936e-01, 1.2529e-01,  …, 2.2255e-01,
          4.5503e-01, 2.8423e+00],
         [1.8825e-02, 4.2435e-01, 2.9205e-01,  …, 7.0882e-02,
          8.7108e-01, 9.2278e-01]],

        [[4.3692e-01, 2.9545e-01, 1.6597e-01,  …, 6.1315e-01,
          2.3375e-01, 4.9136e-01],
         [9.9156e-01, 8.3958e-02, 4.3507e-01,  …, 3.2550e-01,
          3.4743e-01, 7.3152e-01],
         [1.1248e+00, 2.6149e-02, 6.1545e-01,  …, 1.2539e+00,
          1.1430e+00, 4.0929e-01],
         …,
         [4.0891e-01, 1.4092e-01, 3.0194e-01,  …, 1.7048e-01,
          1.3846e+00, 4.6885e-01],
         [6.2307e-01, 4.1691e-01, 5.9603e-01,  …, 7.8116e-03,
          1.4114e-01, 7.5591e-01],
         [6.8830e-01, 3.2772e-01, 1.8181e-01,  …, 3.1025e-01,
          2.3153e+00, 4.3510e+00]],
```

```
[[6.5044e-01, 3.8426e-01, 7.0635e-01,  …, 5.4763e-01,
  1.5085e-01, 8.8686e-01],
 [6.1921e-02, 2.7636e-01, 7.0539e-01,  …, 1.4966e+00,
  7.3114e-01, 1.1154e+00],
 [5.9208e-01, 6.1310e-02, 3.9279e-01,  …, 1.9616e-01,
  1.4218e+00, 5.4240e-02],
 …,
 [1.3114e+00, 5.4916e-02, 1.3638e+00,  …, 9.5369e-01,
  1.2015e+00, 3.4888e-01],
 [1.6831e-01, 5.8005e-01, 3.5946e-01,  …, 1.6365e+00,
  1.2923e+00, 4.4609e+00],
 [9.2791e-02, 3.5758e-01, 2.9470e-01,  …, 5.8231e-01,
  1.4531e+00, 2.5712e+00]],

…,

[[5.1051e-01, 7.3021e-01, 7.1130e-01,  …, 8.5436e-01,
  2.3287e-01, 2.9173e-01],
 [8.1345e-01, 1.5405e-01, 1.3020e+00,  …, 3.7302e-01,
  1.3013e+00, 5.2149e-01],
 [6.1609e-01, 3.8033e-01, 4.3506e-01,  …, 8.5455e-01,
  8.0384e-01, 2.0256e+00],
 …,
 [2.2570e-01, 1.0249e-01, 4.5536e-01,  …, 7.8557e-01,
  1.8171e+00, 1.2723e-01],
 [9.2548e-01, 1.3273e-01, 6.9197e-01,  …, 7.0657e-01,
  2.7231e+00, 2.0390e+00],
 [1.4506e-01, 4.6022e-01, 8.6923e-01,  …, 5.1645e-02,
  1.6757e+00, 2.6752e+00]],

[[2.6193e-01, 4.4637e-01, 6.2078e-01,  …, 7.8300e-01,
  5.9817e-01, 1.7756e-01],
 [6.7313e-01, 3.2362e-01, 5.0590e-01,  …, 6.9812e-01,
  6.5816e-01, 9.0003e-01],
 [1.3038e+00, 5.3690e-01, 4.0259e-01,  …, 1.0052e+00,
  1.0588e+00, 1.1014e+00],
 …,
 [6.8077e-01, 6.3858e-01, 4.2283e-01,  …, 1.3990e-01,
  5.8650e-01, 7.0901e-02],
 [2.6490e-01, 7.3754e-01, 7.1752e-01,  …, 4.8044e-01,
  1.3717e+00, 3.2455e-01],
 [5.0750e-01, 7.3895e-01, 1.0732e-01,  …, 3.9750e-01,
  1.3505e+00, 1.1915e+00]],

[[6.5624e-02, 9.2188e-01, 1.3843e+00,  …, 4.1897e-01,
  2.0055e-01, 1.1315e-01],
 [8.8878e-01, 4.7188e-01, 2.4605e-01,  …, 3.1597e-02,
```

```
   1.0266e+00, 2.1669e-01],
  [1.2950e-01, 8.9456e-01, 1.6690e+00,  …, 3.7082e-01,
   8.5307e-01, 8.2910e-02],
  …,
  [1.0402e+00, 1.4301e-01, 3.3263e-01,  …, 6.3674e-01,
   5.8925e-01, 5.2960e-02],
  [2.6188e-01, 5.1748e-01, 7.9167e-01,  …, 7.6774e-01,
   2.3503e+00, 1.0842e+00],
  [3.1382e-01, 2.3601e-01, 2.2143e-01,  …, 6.8289e-01,
   2.4043e+00, 2.5512e+00]]],


[[[7.1422e-01, 1.0893e+00, 9.1012e-01,  …, 9.5333e-01,
   2.9235e-01, 7.7260e-01],
  [1.4352e+00, 1.4003e+00, 8.3932e-01,  …, 3.2938e-01,
   1.5351e+00, 1.2381e+00],
  [1.0823e+00, 1.3482e+00, 1.7006e+00,  …, 9.7087e-01,
   7.1628e-01, 4.3097e-01],
  …,
  [8.5279e-01, 4.8853e-02, 2.3849e-01,  …, 5.8642e-01,
   2.6699e-01, 8.3883e-01],
  [5.4029e-01, 1.2458e-01, 1.9779e-01,  …, 8.8673e-01,
   2.0042e-02, 3.5401e-02],
  [7.9748e-02, 1.5573e-01, 2.9578e-01,  …, 1.5449e-01,
   1.2682e-01, 5.3056e-01]],

 [[4.9957e-01, 1.2317e+00, 1.1083e+00,  …, 4.8502e-01,
   9.5919e-01, 8.7331e-02],
  [3.0365e-01, 4.8407e-01, 1.4292e+00,  …, 1.0616e+00,
   4.7224e-01, 1.5782e-01],
  [4.1390e-01, 1.7513e+00, 2.2341e+00,  …, 1.9624e+00,
   2.3221e-01, 4.7570e-01],
  …,
  [5.5390e-01, 6.4365e-01, 5.0425e-01,  …, 8.0701e-01,
   1.3551e+00, 3.7633e-01],
  [3.3967e-01, 6.2067e-01, 9.9094e-01,  …, 3.1626e-01,
   1.7610e+00, 4.1331e-01],
  [7.4691e-01, 1.3813e+00, 7.1667e-01,  …, 6.6062e-01,
   1.5544e-01, 2.0377e-01]],

 [[4.3506e-01, 2.1358e-01, 5.3362e-01,  …, 4.6522e-01,
   7.4230e-01, 9.3204e-01],
  [5.5837e-01, 7.8211e-01, 3.7810e-01,  …, 1.0148e+00,
   1.2312e+00, 2.5045e-01],
  [2.9183e-01, 4.8367e-01, 4.9037e-01,  …, 5.4125e-01,
   6.8059e-01, 1.3597e-01],
  …,
  [1.3228e-01, 1.8700e+00, 2.1632e-01,  …, 2.9420e-01,
```

```
  7.9248e-01, 9.2892e-01],
 [1.9126e-03, 1.0682e+00, 9.9787e-01,  …, 8.4015e-02,
  6.5120e-01, 9.6057e-01],
 [2.8137e-01, 1.5378e-01, 7.6610e-01,  …, 8.0122e-02,
  2.2542e-01, 3.0358e-01]],

…,

[[4.7874e-01, 5.3178e-01, 1.3125e+00,  …, 8.6409e-02,
  6.4833e-02, 1.2171e+00],
 [9.8118e-01, 3.2781e-01, 3.0718e-01,  …, 2.9080e-01,
  6.2510e-01, 4.5569e-01],
 [9.8755e-02, 3.4972e-01, 5.2496e-01,  …, 9.2529e-01,
  2.8798e-01, 7.0913e-01],
 …,
 [6.8364e-01, 5.4691e-02, 9.9605e-01,  …, 2.3620e+00,
  1.1136e-01, 1.1311e-01],
 [6.5629e-01, 5.9406e-02, 6.3399e-01,  …, 2.2969e-01,
  1.1743e-01, 2.4218e-01],
 [4.3736e-01, 1.7181e-01, 7.9208e-01,  …, 3.7290e-01,
  8.7631e-02, 5.7144e-03]],

[[1.9190e+00, 2.0986e-01, 7.6046e-03,  …, 5.0693e-01,
  1.3474e+00, 1.4392e+00],
 [1.6935e+00, 8.4301e-02, 1.4668e-01,  …, 4.9329e-01,
  1.2644e+00, 1.9078e+00],
 [1.3667e+00, 2.4447e-01, 1.8546e-01,  …, 1.3215e-01,
  1.7513e+00, 1.0142e+00],
 …,
 [6.6121e-01, 7.7456e-01, 1.2871e+00,  …, 6.4737e-01,
  7.7936e-01, 2.8126e-01],
 [4.8398e-01, 6.4823e-02, 1.4951e-01,  …, 1.0778e+00,
  1.3805e+00, 5.9098e-01],
 [3.7778e-01, 1.1533e-01, 7.0140e-01,  …, 3.0766e-01,
  2.3712e-01, 5.0341e-01]],

[[1.2780e-01, 8.7207e-01, 2.2043e-01,  …, 5.2260e-01,
  5.1268e-01, 8.4286e-01],
 [1.9743e-01, 1.0478e+00, 1.0779e+00,  …, 1.3493e+00,
  6.8558e-01, 2.4875e-01],
 [2.9597e-01, 7.2315e-01, 7.1529e-01,  …, 8.3954e-01,
  2.0444e-01, 1.0093e-01],
 …,
 [5.8755e-01, 2.7815e-01, 6.3272e-02,  …, 1.1181e-01,
  4.3815e-01, 2.7683e-01],
 [1.5765e-01, 7.7670e-02, 3.7846e-02,  …, 4.8082e-01,
  9.4096e-02, 7.1131e-01],
 [1.1777e-01, 2.9844e-01, 2.4902e-02,  …, 8.6706e-01,
```

```
   8.2183e-03, 1.4134e-01]]],


[[[9.3536e-03, 7.5484e-01, 1.6831e+00,  …, 2.6163e-01,
   2.2871e-01, 8.5072e-01],
  [1.8227e-01, 5.9093e-01, 1.6851e+00,  …, 8.4413e-02,
   6.9900e-01, 6.2005e-01],
  [1.1192e-01, 5.8227e-01, 6.0832e-01,  …, 4.3424e-01,
   1.4339e+00, 2.1323e-01],
  …,
  [5.5533e-01, 5.1177e-01, 3.2301e-03,  …, 1.7420e-01,
   1.1909e-01, 1.8280e-01],
  [2.1088e-01, 5.3852e-01, 7.3583e-03,  …, 1.0794e-01,
   5.7110e-01, 2.6842e+00],
  [6.7084e-02, 1.1150e-01, 2.9496e-01,  …, 2.9496e-01,
   4.6769e-01, 6.3962e-01]],

 [[2.1358e-01, 1.3378e-01, 2.3621e-01,  …, 2.4303e-01,
   2.4550e+00, 6.0863e-01],
  [2.3099e-02, 2.6130e-01, 2.7963e-01,  …, 1.5399e+00,
   8.5153e-01, 2.0255e-01],
  [7.0975e-01, 3.3376e-01, 2.8999e-01,  …, 2.9934e-01,
   2.4381e-01, 4.4826e-01],
  …,
  [1.4248e+00, 1.5739e+00, 9.9992e-01,  …, 4.8437e-01,
   4.9456e-01, 6.4116e-01],
  [1.0660e+00, 2.9313e-01, 5.9831e-02,  …, 8.1492e-01,
   5.7221e-01, 6.1424e-01],
  [7.3154e-01, 6.0855e-01, 4.0148e-01,  …, 3.7539e-01,
   2.4045e+00, 3.0970e+00]],

 [[6.8147e-01, 4.6738e-01, 1.1093e-01,  …, 1.0741e-01,
   1.2027e-01, 1.9728e+00],
  [1.5512e-01, 3.5833e-01, 8.6752e-02,  …, 5.4605e-01,
   1.4875e+00, 1.3460e+00],
  [4.4765e-02, 2.1988e-01, 4.5166e-01,  …, 2.0266e-01,
   2.0031e+00, 4.9546e-02],
  …,
  [6.0548e-01, 2.3865e-01, 3.3606e-01,  …, 5.5032e-03,
   1.0533e+00, 1.2163e+00],
  [6.4500e-01, 1.1529e-01, 9.2574e-01,  …, 3.0358e-01,
   1.7585e+00, 2.7139e+00],
  [1.5501e-01, 6.0627e-01, 2.2340e-01,  …, 4.4656e-01,
   1.3363e+00, 2.1126e+00]],

 …,

 [[2.4757e-01, 4.9651e-01, 4.0913e-01,  …, 1.2255e+00,
```

```
   1.8307e+00, 2.1085e+00],
  [4.7750e-01, 2.6064e-01, 5.5314e-01,  …, 1.6565e+00,
   2.9979e+00, 6.1846e-01],
  [4.1115e-01, 2.9179e-01, 3.3119e-01,  …, 9.6369e-02,
   6.2894e-02, 5.9061e-01],
  …,
  [3.5040e-01, 2.1095e-01, 9.0273e-01,  …, 7.7388e-01,
   5.9146e-01, 2.3413e-01],
  [4.9130e-01, 7.2641e-01, 6.3877e-01,  …, 1.2115e-01,
   1.6120e+00, 1.8719e+00],
  [2.3336e-01, 5.0779e-01, 7.9710e-01,  …, 2.7323e-03,
   7.4173e-01, 2.3018e+00]],

 [[4.1741e-02, 1.0679e-01, 7.4069e-01,  …, 1.4275e+00,
   1.5176e+00, 1.2874e+00],
  [3.7144e-02, 2.9123e-01, 7.5951e-01,  …, 1.4677e+00,
   2.5349e-01, 1.1680e+00],
  [6.1058e-01, 5.3751e-01, 5.9012e-01,  …, 5.9052e-01,
   1.1053e-01, 5.8617e-01],
  …,
  [3.3440e-02, 9.6680e-02, 4.7940e-01,  …, 4.4891e-01,
   2.4837e-01, 4.6416e-01],
  [3.2004e-02, 1.5283e-01, 2.1059e-01,  …, 1.4315e+00,
   6.0218e-01, 2.3680e-01],
  [3.2053e-03, 2.1796e-01, 3.4664e-01,  …, 4.9255e-01,
   1.4877e+00, 1.0811e+00]],

 [[4.1559e-01, 6.0890e-01, 6.6211e-01,  …, 7.3751e-01,
   1.2086e+00, 9.0765e-01],
  [1.2682e-01, 2.0588e-01, 6.4935e-01,  …, 5.7220e-06,
   7.9467e-01, 7.9249e-02],
  [5.7714e-02, 2.4816e-01, 1.1416e+00,  …, 1.1693e-01,
   7.1911e-01, 4.1650e-01],
  …,
  [1.3602e-01, 5.1762e-01, 7.9645e-01,  …, 1.5344e-01,
   9.4127e-01, 8.2332e-01],
  [5.6824e-01, 3.0608e-01, 7.5279e-01,  …, 6.6287e-01,
   1.5762e+00, 8.0451e-01],
  [7.2410e-02, 8.0302e-01, 5.0165e-01,  …, 1.2304e-01,
   1.8188e+00, 2.1388e+00]]],


…,


[[[7.7597e-01, 8.1151e-01, 1.0030e+00,  …, 1.3423e+00,
   7.6398e-01, 8.0596e-01],
  [3.2583e-01, 2.9940e-01, 1.4195e-01,  …, 6.3629e-01,
```

```
  1.4588e+00, 2.9305e-01],
 [9.8107e-02, 6.2864e-02, 6.3910e-02,  …, 1.0820e-01,
  4.7898e-01, 2.4467e-01],
 …,
 [1.5250e-01, 1.8132e-02, 3.2378e-01,  …, 1.8960e-01,
  8.4446e-02, 3.3268e-01],
 [1.5353e-01, 6.9099e-01, 8.2695e-01,  …, 1.0804e-01,
  2.6333e-01, 4.6474e-02],
 [1.3523e-02, 1.0514e-01, 5.3501e-02,  …, 4.0139e-01,
  2.7512e-01, 3.5347e-01]],

[[3.5026e-01, 6.2291e-01, 1.1239e+00,  …, 1.0537e+00,
  1.0431e+00, 1.8918e-01],
 [4.1856e-01, 8.5581e-01, 6.0070e-01,  …, 1.8108e+00,
  3.4162e-01, 3.5596e-01],
 [2.8511e-01, 3.2844e-01, 8.0759e-01,  …, 1.0526e+00,
  7.1462e-01, 1.1559e+00],
 …,
 [7.2795e-01, 2.3688e-01, 1.2510e+00,  …, 3.5936e-01,
  8.9996e-02, 3.2906e-01],
 [2.6214e-01, 2.1992e-03, 8.7128e-01,  …, 3.8836e-01,
  8.6500e-01, 3.4199e-01],
 [5.4339e-01, 3.0875e-01, 3.3695e-01,  …, 3.7951e-02,
  3.6069e-01, 1.2634e-01]],

[[6.5255e-01, 9.9868e-01, 6.3024e-01,  …, 1.9855e-01,
  1.1210e+00, 1.7709e-01],
 [8.4859e-01, 4.5186e-01, 2.8531e-01,  …, 6.8722e-01,
  6.9430e-01, 4.1083e-01],
 [4.3023e-01, 4.3287e-01, 8.4441e-02,  …, 7.9416e-01,
  1.2386e+00, 1.6187e-01],
 …,
 [3.7518e-01, 5.3709e-01, 4.0063e-01,  …, 8.2295e-01,
  2.2749e-01, 8.6588e-01],
 [7.4877e-01, 3.7356e-01, 1.3047e-01,  …, 3.6147e-01,
  4.3230e-01, 1.6630e-01],
 [1.7507e-01, 5.5820e-01, 5.6699e-01,  …, 2.2982e-01,
  4.7905e-01, 9.1853e-01]],

…,

[[5.5435e-01, 4.0559e-01, 6.7348e-01,  …, 4.2261e-01,
  1.4926e+00, 1.5766e+00],
 [2.0337e-01, 4.3576e-01, 1.0580e+00,  …, 7.2168e-01,
  1.2745e-02, 6.1867e-01],
 [3.6593e-01, 1.0510e+00, 5.4760e-01,  …, 1.4008e+00,
  1.0482e+00, 3.7970e-01],
 …,
```

```
  [6.9535e-01, 7.0039e-01, 1.2136e+00,  …, 3.5171e+00,
   2.7640e+00, 8.9031e-02],
  [5.9003e-01, 3.3299e-01, 3.8809e-01,  …, 7.9789e-01,
   2.1750e+00, 9.8267e-02],
  [2.2468e-01, 4.9907e-01, 3.5373e-01,  …, 5.6237e-01,
   2.1776e-01, 5.0416e-01]],

 [[9.9923e-01, 4.0092e-01, 3.6274e-01,  …, 1.7395e+00,
   1.7336e+00, 2.8354e-01],
  [9.6809e-01, 1.0378e+00, 8.2633e-01,  …, 1.7452e+00,
   1.2006e+00, 9.2803e-01],
  [1.6110e-01, 1.2375e+00, 4.7428e-01,  …, 2.3277e+00,
   2.2163e-01, 8.9202e-01],
  …,
  [1.1883e+00, 4.9108e-01, 6.8352e-02,  …, 2.6734e+00,
   2.0025e-01, 1.2258e-01],
  [1.1730e+00, 3.7766e-01, 1.9441e-01,  …, 1.1010e+00,
   2.3384e-01, 6.3822e-01],
  [5.5073e-01, 7.7394e-02, 3.6212e-01,  …, 9.8837e-01,
   4.5055e-01, 1.0689e-02]],

 [[6.7192e-01, 1.0829e-01, 1.2238e+00,  …, 2.2515e-01,
   1.8012e-01, 3.2558e-01],
  [6.0306e-01, 1.0819e-01, 7.3367e-01,  …, 2.1937e+00,
   9.7543e-01, 8.2138e-01],
  [8.4143e-01, 5.4906e-01, 3.6080e-01,  …, 4.3387e-02,
   9.7939e-01, 3.8959e-01],
  …,
  [2.7475e-01, 5.0211e-02, 4.4767e-02,  …, 2.1858e+00,
   6.4629e-01, 6.7901e-01],
  [7.1576e-03, 1.9705e-01, 4.3873e-01,  …, 1.4878e-01,
   2.3465e-01, 1.7028e-02],
  [1.0703e-01, 6.8974e-01, 6.5920e-02,  …, 1.5518e-01,
   4.4826e-01, 4.7308e-01]]],


[[[1.9254e-01, 1.7569e-01, 1.9008e-01,  …, 6.1076e-02,
   5.8899e-01, 1.7055e-01],
  [4.4335e-01, 5.9725e-01, 1.3813e+00,  …, 1.1012e+00,
   5.5826e-01, 1.5504e-01],
  [7.1018e-01, 1.4152e+00, 2.4714e+00,  …, 1.4282e+00,
   2.8034e-01, 2.4241e-01],
  …,
  [6.9496e-01, 9.4920e-01, 8.3428e-01,  …, 5.0989e-01,
   4.4878e-01, 4.0266e-01],
  [9.9981e-01, 1.7712e+00, 1.0630e+00,  …, 1.2002e+00,
   4.6423e-01, 6.3462e-01],
  [5.5988e-01, 1.3088e-01, 5.0949e-01,  …, 7.3250e-01,
```

```
  7.1479e-01, 1.0686e-01]],

[[1.1025e-01, 1.1393e+00, 5.0911e-01,  …, 1.0520e+00,
  3.5898e-01, 8.0094e-01],
 [8.8066e-01, 2.5819e-01, 5.7744e-01,  …, 5.8923e-01,
  7.6536e-01, 8.3737e-01],
 [2.7898e-01, 2.5683e-01, 1.5094e+00,  …, 5.1270e-01,
  8.0220e-01, 2.4233e-01],
 …,
 [4.9428e-01, 1.1531e+00, 3.2582e-01,  …, 1.0172e+00,
  4.2978e-01, 4.5582e-01],
 [5.6555e-01, 7.9285e-01, 5.2128e-01,  …, 4.1682e-01,
  2.0910e+00, 2.2084e-02],
 [2.4311e-01, 1.7314e-01, 2.3430e-02,  …, 9.9606e-01,
  3.1066e-02, 1.3214e-01]],

[[2.0714e-02, 6.6891e-01, 1.9496e-01,  …, 7.1726e-01,
  2.6706e-01, 1.5684e-01],
 [5.2769e-01, 2.8303e-01, 1.1035e+00,  …, 4.7423e-01,
  6.1988e-01, 5.1870e-01],
 [3.8456e-02, 1.4502e+00, 2.7164e-01,  …, 7.2076e-01,
  7.8669e-02, 1.1692e+00],
 …,
 [8.0651e-01, 8.4177e-01, 2.4754e-01,  …, 8.2753e-01,
  1.1479e-01, 4.3390e-02],
 [2.2597e-01, 3.0148e-01, 2.0603e-01,  …, 1.0321e+00,
  5.6610e-01, 6.0964e-01],
 [2.3418e-01, 3.9351e-02, 4.5794e-01,  …, 8.2545e-01,
  1.3286e-01, 9.5628e-01]],

…,

[[3.4669e-01, 1.4873e+00, 6.7883e-01,  …, 5.6197e-01,
  5.6243e-01, 1.1205e-01],
 [1.6704e-01, 4.9178e-01, 3.0365e-01,  …, 3.7274e-01,
  1.2047e-02, 1.4136e-01],
 [2.8274e-01, 5.4311e-01, 1.8153e-01,  …, 1.0165e+00,
  7.1060e-01, 1.0929e+00],
 …,
 [1.5402e+00, 5.5328e-01, 1.6375e+00,  …, 4.8938e-02,
  1.2014e+00, 1.1984e+00],
 [3.3405e-02, 4.5252e-01, 8.5541e-01,  …, 1.7093e+00,
  1.7316e+00, 4.8010e-01],
 [1.4940e-01, 8.5993e-01, 5.4261e-01,  …, 6.7912e-01,
  1.1971e-01, 9.4567e-01]],

[[2.5347e-01, 6.0595e-02, 4.4737e-01,  …, 8.7302e-02,
  2.8097e-01, 9.6012e-02],
```

```
  [2.7322e-01, 3.4159e-01, 1.2731e+00,  …, 1.2521e+00,
   5.6762e-01, 4.5584e-01],
  [2.3018e-01, 4.2518e-01, 3.4496e-01,  …, 1.8317e+00,
   2.0160e-01, 3.4485e-02],
  …,
  [9.6515e-01, 1.1581e+00, 3.8223e-01,  …, 1.7821e+00,
   6.1824e-01, 1.0603e+00],
  [7.8257e-01, 5.2805e-01, 6.2919e-01,  …, 6.1807e-01,
   1.6522e+00, 5.3031e-01],
  [2.9273e-01, 3.7149e-01, 1.9806e-01,  …, 6.3072e-02,
   8.2351e-01, 6.2167e-01]],

 [[1.2799e-01, 1.3699e-01, 4.8353e-02,  …, 3.7296e-01,
   3.4602e-01, 1.0427e-02],
  [3.6821e-02, 2.8292e-01, 2.5972e-01,  …, 5.8220e-01,
   3.9719e-01, 5.2955e-01],
  [5.7286e-01, 4.2874e-01, 1.4200e-01,  …, 9.5248e-01,
   3.7645e-01, 5.2656e-01],
  …,
  [1.0322e-01, 1.2778e+00, 6.1798e-01,  …, 1.0710e+00,
   1.5513e-01, 1.3280e-01],
  [2.5123e-01, 2.3682e-01, 5.7850e-03,  …, 9.1203e-01,
   6.5298e-01, 3.0293e-01],
  [5.4786e-01, 5.0929e-01, 8.5984e-01,  …, 7.1861e-01,
   2.9022e-01, 1.1490e+00]]],


[[[1.2611e-01, 6.2123e-01, 3.2231e-01,  …, 7.3926e-02,
   7.9763e-01, 1.4499e-01],
  [3.4400e-01, 4.3915e-01, 1.6150e-01,  …, 5.2547e-01,
   1.1846e+00, 1.3241e+00],
  [5.6351e-01, 1.6474e-01, 4.4202e-01,  …, 1.0653e+00,
   5.4377e-01, 8.3003e-01],
  …,
  [3.8215e-01, 6.3168e-01, 6.6646e-02,  …, 8.0384e-02,
   4.9775e-01, 2.1897e-02],
  [7.4169e-01, 3.8226e-01, 3.6579e-01,  …, 1.4251e-01,
   4.5280e-01, 2.9601e+00],
  [3.0721e-01, 1.1851e-01, 1.7147e-01,  …, 4.0875e-01,
   9.0655e-01, 7.6038e-01]],

 [[3.0092e-03, 7.2797e-01, 3.5301e-01,  …, 7.9228e-01,
   2.9997e+00, 3.8532e+00],
  [3.8966e-01, 5.1220e-01, 3.5895e-01,  …, 4.2475e-01,
   2.0123e+00, 4.0704e+00],
  [5.3417e-01, 3.1895e-01, 5.2775e-01,  …, 1.2519e+00,
   4.1218e-01, 1.4692e+00],
  …,
```

```
 [6.3243e-01, 5.3181e-01, 8.6811e-01,  …, 1.0824e+00,
  1.9368e-01, 1.3300e+00],
 [9.1411e-01, 8.4782e-01, 3.0086e-01,  …, 3.7915e-01,
  5.3559e-01, 9.4519e-02],
 [8.4466e-02, 5.6615e-01, 3.8145e-01,  …, 6.8351e-01,
  3.4610e+00, 2.3065e+00]],

[[3.5540e-01, 1.8324e-01, 8.0090e-01,  …, 7.2805e-01,
  7.9644e-01, 1.5628e+00],
 [2.9950e-02, 4.1141e-01, 1.2257e+00,  …, 1.0380e+00,
  8.9306e-01, 2.1895e+00],
 [9.1862e-01, 5.2428e-01, 8.6294e-02,  …, 5.2746e-01,
  6.7423e-01, 1.7111e-01],
 …,
 [3.6799e-01, 1.0288e+00, 1.3892e-01,  …, 3.1431e-01,
  1.1667e+00, 6.9645e-01],
 [4.7767e-01, 8.3368e-02, 2.4920e-01,  …, 1.7948e-01,
  5.5734e-01, 3.3789e+00],
 [1.7721e-01, 3.0067e-01, 4.6431e-01,  …, 3.0685e-01,
  6.8080e-01, 3.3923e+00]],

…,

[[1.9621e-01, 3.2009e-01, 1.1132e+00,  …, 5.7941e-01,
  4.2009e-01, 2.3432e+00],
 [3.9641e-01, 3.4582e-01, 4.6390e-01,  …, 1.9269e+00,
  1.0147e+00, 4.4679e+00],
 [2.2361e-01, 8.7543e-01, 9.0261e-01,  …, 2.1440e+00,
  2.8856e-02, 6.6558e-01],
 …,
 [3.2711e-01, 1.1300e+00, 2.5863e-01,  …, 9.7422e-01,
  5.9200e-01, 2.1076e-02],
 [7.9700e-01, 1.0287e+00, 8.3844e-01,  …, 4.4240e-01,
  4.8548e+00, 1.6951e+00],
 [9.7963e-01, 6.0355e-01, 1.0041e+00,  …, 4.2072e-02,
  1.8367e+00, 3.5230e+00]],

[[2.4897e-01, 1.4671e+00, 4.1045e-01,  …, 4.7430e-01,
  1.5965e+00, 2.0864e+00],
 [2.5631e-02, 1.0900e+00, 7.8481e-01,  …, 9.2041e-01,
  4.9843e-01, 1.1175e-01],
 [1.2807e+00, 4.2488e-02, 1.6632e+00,  …, 3.9236e-01,
  5.7086e-01, 1.9590e+00],
 …,
 [1.2601e+00, 5.3115e-01, 1.3203e+00,  …, 8.1354e-02,
  1.0595e+00, 1.4882e+00],
 [1.4702e+00, 1.2530e+00, 4.3078e-01,  …, 1.5914e-01,
  5.0841e-01, 4.6161e-01],
```

```
          [9.3919e-01, 7.8232e-01, 3.7109e-02,  …, 4.9173e-01,
           1.8802e+00, 2.4092e+00]],

         [[2.5532e-01, 3.7306e-01, 2.8211e-01,  …, 7.1081e-02,
           2.7513e+00, 2.7442e+00],
          [2.3849e-01, 3.7920e-01, 1.5709e-01,  …, 8.5134e-01,
           7.5342e-01, 1.9515e+00],
          [8.3821e-01, 8.4165e-01, 6.9706e-02,  …, 8.6512e-01,
           6.0488e-02, 1.7450e-01],
          …,
          [1.6983e-01, 1.5249e-01, 5.3876e-01,  …, 1.5062e-01,
           2.1238e-01, 7.1967e-01],
          [3.1201e-01, 5.1463e-01, 4.5135e-01,  …, 2.2796e-02,
           1.1476e+00, 5.2258e-01],
          [4.9031e-01, 3.9788e-01, 2.1007e-01,  …, 2.8001e-01,
           2.3411e+00, 2.8592e+00]]]], device='cuda:0', grad_fn=<AbsBackward0>)
```

[13]:
```
diff.mean()
```

[13]:
```
tensor(0.7316, device='cuda:0', grad_fn=<MeanBackward0>)
```

[16]:
```
resnet20_2bit_quan = resnet20_quant(w_quan_bits=2).cuda()
pre_best_prec = 0
quan_model_name = "2bit_quan_resnet"
```

[ ]:
```
lr = 7e-2
weight_decay = 1e-4
epochs = 100
# pre_best_prec = 85.470

criterion = nn.CrossEntropyLoss().cuda()

res20_2bit_optimizer = torch.optim.SGD(resnet20_2bit_quan.parameters(), lr=lr,␣
 ↪momentum=0.9, weight_decay=weight_decay)
pre_best_prec = helper.train_model(resnet20_2bit_quan,␣
 ↪model_name=quan_model_name, optimizer=res20_2bit_optimizer,␣
 ↪trainloader=trainloader, testloader=testloader, criterion=criterion,␣
 ↪epochs=epochs, pre_best_prec=pre_best_prec)
```

[17]:
```
helper.test_model(model=resnet20_2bit_quan, model_name=quan_model_name,␣
 ↪testloader=testloader)
```

```
Test: [0/79]    Time 0.233 (0.233)      Loss 0.3729 (0.3729)    Prec 87.500%
(87.500%)
 * Prec 83.960%
```

[ ]: