In [ ]:

```
import argparse
import os
import time
import shutil

import torch
import torch.nn as nn
import torch.optim as optim
import torch.nn.functional as F
import torch.backends.cudnn as cudnn


import torchvision
import torchvision.transforms as transforms
import sys

sys.path.insert(1, './misc')
from models import *


global best_prec
use_gpu = torch.cuda.is_available()
print('=> Building model...')
```

```
batch_size = 128
model_name = "final_VGG16_quant"
model = final_VGG16_quant()

print(model)

normalize = transforms.Normalize(mean=[0.491, 0.482, 0.447], std=[0.247, 0.243, 0.262])


train_dataset = torchvision.datasets.CIFAR10(
    root='./data',
    train=True,
    download=True,
    transform=transforms.Compose([
        transforms.RandomCrop(32, padding=4),
        transforms.RandomHorizontalFlip(),
        transforms.ToTensor(),
        normalize,
    ]))
trainloader = torch.utils.data.DataLoader(train_dataset, batch_size=batch_size, shuffle=True, num_workers=2


test_dataset = torchvision.datasets.CIFAR10(
    root='./data',
    train=False,
    download=True,
    transform=transforms.Compose([
        transforms.ToTensor(),
        normalize,
    ]))

testloader = torch.utils.data.DataLoader(test_dataset, batch_size=batch_size, shuffle=False, num_workers=2)


print_freq = 100 # every 100 batches, accuracy printed. Here, each batch includes "batch_size" data points
# CIFAR10 has 50,000 training data, and 10,000 validation data.

def train(trainloader, model, criterion, optimizer, epoch):
    batch_time = AverageMeter()
    data_time = AverageMeter()
    losses = AverageMeter()
    top1 = AverageMeter()

    model.train()

    end = time.time()
    for i, (input, target) in enumerate(trainloader):
        # measure data loading time
        data_time.update(time.time() - end)

        input, target = input.cuda(), target.cuda()

        # compute output
        output = model(input)
```

```
        loss = criterion(output, target)

        # measure accuracy and record loss
        prec = accuracy(output, target)[0]
        losses.update(loss.item(), input.size(0))
        top1.update(prec.item(), input.size(0))

        # compute gradient and do SGD step
        optimizer.zero_grad()
        loss.backward()
        optimizer.step()

        # measure elapsed time
        batch_time.update(time.time() - end)
        end = time.time()


        if i % print_freq == 0:
            print('Epoch: [{0}][{1}/{2}]\t'
                  'Time {batch_time.val:.3f} ({batch_time.avg:.3f})\t'
                  'Data {data_time.val:.3f} ({data_time.avg:.3f})\t'
                  'Loss {loss.val:.4f} ({loss.avg:.4f})\t'
                  'Prec {top1.val:.3f}% ({top1.avg:.3f}%)'.format(
                   epoch, i, len(trainloader), batch_time=batch_time,
                   data_time=data_time, loss=losses, top1=top1))


def validate(val_loader, model, criterion ):
    batch_time = AverageMeter()
    losses = AverageMeter()
    top1 = AverageMeter()

    # switch to evaluate mode
    model.eval()

    end = time.time()
    with torch.no_grad():
        for i, (input, target) in enumerate(val_loader):

            input, target = input.cuda(), target.cuda()

            # compute output
            output = model(input)
            loss = criterion(output, target)

            # measure accuracy and record loss
            prec = accuracy(output, target)[0]
            losses.update(loss.item(), input.size(0))
            top1.update(prec.item(), input.size(0))

            # measure elapsed time
            batch_time.update(time.time() - end)
            end = time.time()

            if i % print_freq == 0:  # This line shows how frequently print out the status. e.g., i%5 => ev
                print('Test: [{0}/{1}]\t'
                  'Time {batch_time.val:.3f} ({batch_time.avg:.3f})\t'
                  'Loss {loss.val:.4f} ({loss.avg:.4f})\t'
                  'Prec {top1.val:.3f}% ({top1.avg:.3f}%)'.format(
                   i, len(val_loader), batch_time=batch_time, loss=losses,
                   top1=top1))

    print(' * Prec {top1.avg:.3f}% '.format(top1=top1))
    return top1.avg


def accuracy(output, target, topk=(1,)):
    """Computes the precision@k for the specified values of k"""
    maxk = max(topk)
    batch_size = target.size(0)

    _, pred = output.topk(maxk, 1, True, True)
    pred = pred.t()
    correct = pred.eq(target.view(1, -1).expand_as(pred))

    res = []
    for k in topk:
        correct_k = correct[:k].view(-1).float().sum(0)
        res.append(correct_k.mul_(100.0 / batch_size))
    return res
```

```python
class AverageMeter(object):
    """Computes and stores the average and current value"""
    def __init__(self):
        self.reset()

    def reset(self):
        self.val = 0
        self.avg = 0
        self.sum = 0
        self.count = 0

    def update(self, val, n=1):
        self.val = val
        self.sum += val * n
        self.count += n
        self.avg = self.sum / self.count


def save_checkpoint(state, is_best, fdir):
    filepath = os.path.join(fdir, 'checkpoint.pth')
    torch.save(state, filepath)
    if is_best:
        shutil.copyfile(filepath, os.path.join(fdir, 'model_best.pth.tar'))
```

```python
def adjust_learning_rate(optimizer, epoch):
    """For resnet, the lr starts from 0.1, and is divided by 10 at 80 and 120 epochs"""
    adjust_list = [150, 225]
    if epoch in adjust_list:
        for param_group in optimizer.param_groups:
            param_group['lr'] = param_group['lr'] * 0.1

#model = nn.DataParallel(model).cuda()
#all_params = checkpoint['state_dict']
#model.load_state_dict(all_params, strict=False)
#criterion = nn.CrossEntropyLoss().cuda()
#validate(testloader, model, criterion)
```

```
=> Building model...
VGG_quant(
  (features): Sequential(
    (0): QuantConv2d(
      3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
      (weight_quant): weight_quantize_fn()
    )
    (1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (2): ReLU(inplace=True)
    (3): QuantConv2d(
      64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
      (weight_quant): weight_quantize_fn()
    )
    (4): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (5): ReLU(inplace=True)
    (6): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    (7): QuantConv2d(
      64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
      (weight_quant): weight_quantize_fn()
    )
    (8): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (9): ReLU(inplace=True)
    (10): QuantConv2d(
      128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
      (weight_quant): weight_quantize_fn()
    )
    (11): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (12): ReLU(inplace=True)
    (13): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
    (14): QuantConv2d(
      128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
      (weight_quant): weight_quantize_fn()
    )
    (15): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (16): ReLU(inplace=True)
    (17): QuantConv2d(
      256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
      (weight_quant): weight_quantize_fn()
    )
    (18): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    (19): ReLU(inplace=True)
    (20): QuantConv2d(
      256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
```

```
  (weight_quant): weight_quantize_fn()
)
(21): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(22): ReLU(inplace=True)
(23): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
(24): QuantConv2d(
  256, 8, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
  (weight_quant): weight_quantize_fn()
)
(25): BatchNorm2d(8, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(26): ReLU(inplace=True)
(27): QuantConv2d(
  8, 8, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
  (weight_quant): weight_quantize_fn()
)
(28): ReLU(inplace=True)
(29): QuantConv2d(
  8, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
  (weight_quant): weight_quantize_fn()
)
(30): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(31): ReLU(inplace=True)
(32): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
(33): QuantConv2d(
  512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
  (weight_quant): weight_quantize_fn()
)
(34): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(35): ReLU(inplace=True)
(36): QuantConv2d(
  512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
  (weight_quant): weight_quantize_fn()
)
(37): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(38): ReLU(inplace=True)
(39): QuantConv2d(
  512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False
  (weight_quant): weight_quantize_fn()
)
(40): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
(41): ReLU(inplace=True)
(42): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1, ceil_mode=False)
(43): AvgPool2d(kernel_size=1, stride=1, padding=0)
  )
  (classifier): Linear(in_features=512, out_features=10, bias=True)
)
Files already downloaded and verified
Files already downloaded and verified
```

In [2]:

```
# This cell won't be given, but students will complete the training

lr = 1e-2
weight_decay = 8e-4
epochs = 150
best_prec = 0

#model = nn.DataParallel(model).cuda()
model.cuda()
criterion = nn.CrossEntropyLoss().cuda()
optimizer = torch.optim.SGD(model.parameters(), lr=lr, momentum=0.9, weight_decay=weight_decay)
#cudnn.benchmark = True

if not os.path.exists('result'):
    os.makedirs('result')
fdir = 'result/'+str(model_name)
if not os.path.exists(fdir):
    os.makedirs(fdir)


untrained = True

if (untrained):
    for epoch in range(0, epochs):
        adjust_learning_rate(optimizer, epoch)

        train(trainloader, model, criterion, optimizer, epoch)

        # evaluate on test set
        print("Validation starts")
        prec = validate(testloader, model, criterion)

        # remember best precision and save checkpoint
```

```
        is_best = prec > best_prec
        best_prec = max(prec,best_prec)
        print('best acc: {:1f}'.format(best_prec))
        save_checkpoint({
            'epoch': epoch + 1,
            'state_dict': model.state_dict(),
            'best_prec': best_prec,
            'optimizer': optimizer.state_dict(),
        }, is_best, fdir)
```

```
Epoch: [0][0/391]       Time 0.743 (0.743)      Data 0.306 (0.306)      Loss 2.5791 (2.5791)    Prec 9.375%
(9.375%)
Epoch: [0][100/391]     Time 0.021 (0.028)      Data 0.001 (0.004)      Loss 1.7758 (2.0248)    Prec 28.90
6% (26.106%)
Epoch: [0][200/391]     Time 0.021 (0.024)      Data 0.001 (0.003)      Loss 1.4695 (1.8586)    Prec 45.31
2% (31.926%)
Epoch: [0][300/391]     Time 0.020 (0.023)      Data 0.001 (0.002)      Loss 1.6487 (1.7344)    Prec 37.50
0% (36.496%)
Validation starts
Test: [0/79]    Time 0.250 (0.250)      Loss 1.3022 (1.3022)    Prec 56.250% (56.250%)
 * Prec 52.950%
best acc: 52.950000
Epoch: [1][0/391]       Time 0.352 (0.352)      Data 0.331 (0.331)      Loss 1.2386 (1.2386)    Prec 48.43
8% (48.438%)
Epoch: [1][100/391]     Time 0.018 (0.025)      Data 0.001 (0.006)      Loss 1.5457 (1.2643)    Prec 42.96
9% (53.968%)
Epoch: [1][200/391]     Time 0.019 (0.023)      Data 0.001 (0.004)      Loss 0.9439 (1.2208)    Prec 64.84
4% (56.017%)
Epoch: [1][300/391]     Time 0.018 (0.022)      Data 0.001 (0.003)      Loss 1.0101 (1.1894)    Prec 65.62
5% (57.389%)
Validation starts
Test: [0/79]    Time 0.191 (0.191)      Loss 0.9986 (0.9986)    Prec 62.500% (62.500%)
 * Prec 60.000%
best acc: 60.000000
Epoch: [2][0/391]       Time 0.322 (0.322)      Data 0.299 (0.299)      Loss 1.0368 (1.0368)    Prec 68.75
0% (68.750%)
Epoch: [2][100/391]     Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.8784 (1.0367)    Prec 65.62
5% (63.057%)
Epoch: [2][200/391]     Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.8969 (1.0008)    Prec 66.40
6% (64.373%)
Epoch: [2][300/391]     Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.8804 (0.9804)    Prec 67.18
8% (65.192%)
Validation starts
Test: [0/79]    Time 0.205 (0.205)      Loss 0.7445 (0.7445)    Prec 69.531% (69.531%)
 * Prec 67.990%
best acc: 67.990000
Epoch: [3][0/391]       Time 0.388 (0.388)      Data 0.366 (0.366)      Loss 1.0043 (1.0043)    Prec 64.06
2% (64.062%)
Epoch: [3][100/391]     Time 0.021 (0.026)      Data 0.001 (0.006)      Loss 0.9652 (0.8660)    Prec 62.50
0% (69.106%)
Epoch: [3][200/391]     Time 0.021 (0.023)      Data 0.002 (0.004)      Loss 0.9788 (0.8654)    Prec 64.84
4% (69.026%)
Epoch: [3][300/391]     Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.8837 (0.8526)    Prec 69.53
1% (69.643%)
Validation starts
Test: [0/79]    Time 0.196 (0.196)      Loss 0.8932 (0.8932)    Prec 70.312% (70.312%)
 * Prec 67.570%
best acc: 67.990000
Epoch: [4][0/391]       Time 0.309 (0.309)      Data 0.285 (0.285)      Loss 0.7523 (0.7523)    Prec 74.21
9% (74.219%)
Epoch: [4][100/391]     Time 0.021 (0.025)      Data 0.002 (0.006)      Loss 0.9098 (0.7965)    Prec 67.18
8% (72.030%)
Epoch: [4][200/391]     Time 0.021 (0.023)      Data 0.002 (0.004)      Loss 0.7407 (0.7751)    Prec 75.00
0% (72.956%)
Epoch: [4][300/391]     Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 1.0573 (0.7693)    Prec 64.84
4% (73.194%)
Validation starts
Test: [0/79]    Time 0.196 (0.196)      Loss 0.6859 (0.6859)    Prec 80.469% (80.469%)
 * Prec 73.370%
best acc: 73.370000
Epoch: [5][0/391]       Time 0.287 (0.287)      Data 0.263 (0.263)      Loss 0.9251 (0.9251)    Prec 60.93
8% (60.938%)
Epoch: [5][100/391]     Time 0.020 (0.025)      Data 0.001 (0.005)      Loss 0.6752 (0.7176)    Prec 72.65
6% (75.085%)
Epoch: [5][200/391]     Time 0.020 (0.023)      Data 0.001 (0.003)      Loss 0.7273 (0.7028)    Prec 74.21
9% (75.704%)
Epoch: [5][300/391]     Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.6467 (0.6997)    Prec 74.21
9% (75.831%)
Validation starts
Test: [0/79]    Time 0.192 (0.192)      Loss 0.6206 (0.6206)    Prec 80.469% (80.469%)
 * Prec 77.090%
best acc: 77.090000
```

```
Epoch: [6][0/391]       Time 0.368 (0.368)      Data 0.342 (0.342)      Loss 0.6353 (0.6353)    Prec 79.68
8% (79.688%)
Epoch: [6][100/391]     Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.5207 (0.6482)    Prec 82.81
2% (77.731%)
Epoch: [6][200/391]     Time 0.021 (0.022)      Data 0.001 (0.003)      Loss 0.8391 (0.6468)    Prec 72.65
6% (77.670%)
Epoch: [6][300/391]     Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.6056 (0.6475)    Prec 80.46
9% (77.655%)
Validation starts
Test: [0/79]    Time 0.201 (0.201)      Loss 0.6375 (0.6375)    Prec 78.906% (78.906%)
 * Prec 76.700%
best acc: 77.090000
Epoch: [7][0/391]       Time 0.328 (0.328)      Data 0.305 (0.305)      Loss 0.5685 (0.5685)    Prec 78.90
6% (78.906%)
Epoch: [7][100/391]     Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.6095 (0.6069)    Prec 78.12
5% (79.146%)
Epoch: [7][200/391]     Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.5858 (0.6023)    Prec 79.68
8% (79.396%)
Epoch: [7][300/391]     Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.6502 (0.6043)    Prec 76.56
2% (79.399%)
Validation starts
Test: [0/79]    Time 0.199 (0.199)      Loss 0.5380 (0.5380)    Prec 80.469% (80.469%)
 * Prec 79.250%
best acc: 79.250000
Epoch: [8][0/391]       Time 0.378 (0.378)      Data 0.355 (0.355)      Loss 0.5200 (0.5200)    Prec 82.03
1% (82.031%)
Epoch: [8][100/391]     Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.6008 (0.5655)    Prec 79.68
8% (80.732%)
Epoch: [8][200/391]     Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.3746 (0.5650)    Prec 90.62
5% (80.741%)
Epoch: [8][300/391]     Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.5080 (0.5628)    Prec 84.37
5% (80.762%)
Validation starts
Test: [0/79]    Time 0.198 (0.198)      Loss 0.8274 (0.8274)    Prec 73.438% (73.438%)
 * Prec 76.040%
best acc: 79.250000
Epoch: [9][0/391]       Time 0.304 (0.304)      Data 0.282 (0.282)      Loss 0.4598 (0.4598)    Prec 81.25
0% (81.250%)
Epoch: [9][100/391]     Time 0.021 (0.023)      Data 0.002 (0.004)      Loss 0.4523 (0.5523)    Prec 82.81
2% (80.716%)
Epoch: [9][200/391]     Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.4381 (0.5446)    Prec 84.37
5% (81.192%)
Epoch: [9][300/391]     Time 0.020 (0.022)      Data 0.001 (0.002)      Loss 0.5210 (0.5442)    Prec 83.59
4% (81.377%)
Validation starts
Test: [0/79]    Time 0.223 (0.223)      Loss 0.5068 (0.5068)    Prec 82.031% (82.031%)
 * Prec 79.610%
best acc: 79.610000
Epoch: [10][0/391]      Time 0.353 (0.353)      Data 0.330 (0.330)      Loss 0.4799 (0.4799)    Prec 85.93
8% (85.938%)
Epoch: [10][100/391]    Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.3815 (0.5167)    Prec 86.71
9% (82.163%)
Epoch: [10][200/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.4889 (0.5112)    Prec 82.81
2% (82.591%)
Epoch: [10][300/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.5007 (0.5121)    Prec 82.81
2% (82.548%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.5801 (0.5801)    Prec 81.250% (81.250%)
 * Prec 78.690%
best acc: 79.610000
Epoch: [11][0/391]      Time 0.308 (0.308)      Data 0.286 (0.286)      Loss 0.4510 (0.4510)    Prec 84.37
5% (84.375%)
Epoch: [11][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.5490 (0.4826)    Prec 83.59
4% (83.594%)
Epoch: [11][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.4095 (0.4808)    Prec 83.59
4% (83.590%)
Epoch: [11][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.5860 (0.4844)    Prec 79.68
8% (83.464%)
Validation starts
Test: [0/79]    Time 0.187 (0.187)      Loss 0.5272 (0.5272)    Prec 79.688% (79.688%)
 * Prec 80.610%
best acc: 80.610000
Epoch: [12][0/391]      Time 0.356 (0.356)      Data 0.333 (0.333)      Loss 0.4696 (0.4696)    Prec 85.93
8% (85.938%)
Epoch: [12][100/391]    Time 0.021 (0.024)      Data 0.001 (0.005)      Loss 0.4436 (0.4600)    Prec 82.03
1% (84.336%)
Epoch: [12][200/391]    Time 0.021 (0.023)      Data 0.002 (0.003)      Loss 0.5268 (0.4610)    Prec 82.03
1% (84.282%)
Epoch: [12][300/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.6446 (0.4634)    Prec 78.90
6% (84.269%)
Validation starts
Test: [0/79]    Time 0.202 (0.202)      Loss 0.5679 (0.5679)    Prec 78.125% (78.125%)
```

```
 * Prec 79.260%
best acc: 80.610000
Epoch: [13][0/391]      Time 0.400 (0.400)      Data 0.374 (0.374)      Loss 0.4644 (0.4644)    Prec 82.03
1% (82.031%)
Epoch: [13][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.5199 (0.4261)    Prec 82.81
2% (85.628%)
Epoch: [13][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.4393 (0.4369)    Prec 82.81
2% (85.277%)
Epoch: [13][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.3786 (0.4380)    Prec 85.15
6% (85.156%)
Validation starts
Test: [0/79]    Time 0.207 (0.207)      Loss 0.5065 (0.5065)    Prec 84.375% (84.375%)
 * Prec 80.910%
best acc: 80.910000
Epoch: [14][0/391]      Time 0.358 (0.358)      Data 0.337 (0.337)      Loss 0.3949 (0.3949)    Prec 85.93
8% (85.938%)
Epoch: [14][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.3513 (0.4139)    Prec 83.59
4% (85.651%)
Epoch: [14][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.6133 (0.4215)    Prec 79.68
8% (85.576%)
Epoch: [14][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.4780 (0.4232)    Prec 84.37
5% (85.631%)
Validation starts
Test: [0/79]    Time 0.203 (0.203)      Loss 0.7413 (0.7413)    Prec 77.344% (77.344%)
 * Prec 78.600%
best acc: 80.910000
Epoch: [15][0/391]      Time 0.322 (0.322)      Data 0.299 (0.299)      Loss 0.4373 (0.4373)    Prec 86.71
9% (86.719%)
Epoch: [15][100/391]    Time 0.020 (0.024)      Data 0.001 (0.004)      Loss 0.4290 (0.3935)    Prec 85.15
6% (86.371%)
Epoch: [15][200/391]    Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.3628 (0.3997)    Prec 89.06
2% (86.276%)
Epoch: [15][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.4227 (0.4023)    Prec 83.59
4% (86.293%)
Validation starts
Test: [0/79]    Time 0.230 (0.230)      Loss 0.3290 (0.3290)    Prec 85.156% (85.156%)
 * Prec 84.790%
best acc: 84.790000
Epoch: [16][0/391]      Time 0.412 (0.412)      Data 0.388 (0.388)      Loss 0.2732 (0.2732)    Prec 89.06
2% (89.062%)
Epoch: [16][100/391]    Time 0.020 (0.025)      Data 0.001 (0.005)      Loss 0.4116 (0.3801)    Prec 87.50
0% (87.067%)
Epoch: [16][200/391]    Time 0.021 (0.023)      Data 0.001 (0.003)      Loss 0.3257 (0.3844)    Prec 91.40
6% (86.831%)
Epoch: [16][300/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.4769 (0.3887)    Prec 82.03
1% (86.675%)
Validation starts
Test: [0/79]    Time 0.195 (0.195)      Loss 0.4817 (0.4817)    Prec 85.156% (85.156%)
 * Prec 84.590%
best acc: 84.790000
Epoch: [17][0/391]      Time 0.276 (0.276)      Data 0.252 (0.252)      Loss 0.5246 (0.5246)    Prec 82.03
1% (82.031%)
Epoch: [17][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.3631 (0.3673)    Prec 87.50
0% (87.477%)
Epoch: [17][200/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.3666 (0.3696)    Prec 87.50
0% (87.364%)
Epoch: [17][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.3906 (0.3767)    Prec 87.50
0% (87.139%)
Validation starts
Test: [0/79]    Time 0.211 (0.211)      Loss 0.3563 (0.3563)    Prec 88.281% (88.281%)
 * Prec 84.000%
best acc: 84.790000
Epoch: [18][0/391]      Time 0.328 (0.328)      Data 0.301 (0.301)      Loss 0.2177 (0.2177)    Prec 93.75
0% (93.750%)
Epoch: [18][100/391]    Time 0.021 (0.024)      Data 0.001 (0.005)      Loss 0.3371 (0.3428)    Prec 89.06
2% (88.227%)
Epoch: [18][200/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.2366 (0.3522)    Prec 91.40
6% (87.990%)
Epoch: [18][300/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.4188 (0.3578)    Prec 85.93
8% (87.926%)
Validation starts
Test: [0/79]    Time 0.196 (0.196)      Loss 0.4840 (0.4840)    Prec 85.156% (85.156%)
 * Prec 82.110%
best acc: 84.790000
Epoch: [19][0/391]      Time 0.313 (0.313)      Data 0.287 (0.287)      Loss 0.3541 (0.3541)    Prec 87.50
0% (87.500%)
Epoch: [19][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.3485 (0.3425)    Prec 87.50
0% (88.390%)
Epoch: [19][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.4093 (0.3470)    Prec 83.59
4% (88.196%)
Epoch: [19][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.4351 (0.3510)    Prec 83.59
4% (88.087%)
```

```
Validation starts
Test: [0/79]    Time 0.308 (0.308)      Loss 0.3683 (0.3683)    Prec 88.281% (88.281%)
 * Prec 85.530%
best acc: 85.530000
Epoch: [20][0/391]      Time 0.366 (0.366)      Data 0.344 (0.344)      Loss 0.2454 (0.2454)    Prec 92.18
8% (92.188%)
Epoch: [20][100/391]    Time 0.018 (0.024)      Data 0.001 (0.005)      Loss 0.2607 (0.3326)    Prec 90.62
5% (88.420%)
Epoch: [20][200/391]    Time 0.018 (0.022)      Data 0.001 (0.003)      Loss 0.4002 (0.3374)    Prec 85.93
8% (88.413%)
Epoch: [20][300/391]    Time 0.019 (0.021)      Data 0.001 (0.002)      Loss 0.2603 (0.3394)    Prec 89.84
4% (88.375%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.3558 (0.3558)    Prec 86.719% (86.719%)
 * Prec 83.470%
best acc: 85.530000
Epoch: [21][0/391]      Time 0.362 (0.362)      Data 0.340 (0.340)      Loss 0.3329 (0.3329)    Prec 90.62
5% (90.625%)
Epoch: [21][100/391]    Time 0.021 (0.024)      Data 0.001 (0.005)      Loss 0.3584 (0.3337)    Prec 89.06
2% (88.645%)
Epoch: [21][200/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.4405 (0.3268)    Prec 86.71
9% (89.008%)
Epoch: [21][300/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.2172 (0.3313)    Prec 93.75
0% (88.831%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.3247 (0.3247)    Prec 89.062% (89.062%)
 * Prec 84.300%
best acc: 85.530000
Epoch: [22][0/391]      Time 0.396 (0.396)      Data 0.371 (0.371)      Loss 0.2414 (0.2414)    Prec 90.62
5% (90.625%)
Epoch: [22][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.3533 (0.2952)    Prec 88.28
1% (89.666%)
Epoch: [22][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2909 (0.3074)    Prec 91.40
6% (89.459%)
Epoch: [22][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.3399 (0.3102)    Prec 90.62
5% (89.395%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.4021 (0.4021)    Prec 85.938% (85.938%)
 * Prec 83.730%
best acc: 85.530000
Epoch: [23][0/391]      Time 0.355 (0.355)      Data 0.330 (0.330)      Loss 0.2452 (0.2452)    Prec 90.62
5% (90.625%)
Epoch: [23][100/391]    Time 0.019 (0.024)      Data 0.002 (0.005)      Loss 0.2264 (0.3073)    Prec 91.40
6% (89.898%)
Epoch: [23][200/391]    Time 0.019 (0.022)      Data 0.001 (0.003)      Loss 0.3041 (0.3105)    Prec 89.84
4% (89.700%)
Epoch: [23][300/391]    Time 0.018 (0.021)      Data 0.001 (0.002)      Loss 0.3099 (0.3119)    Prec 89.84
4% (89.644%)
Validation starts
Test: [0/79]    Time 0.277 (0.277)      Loss 0.4055 (0.4055)    Prec 84.375% (84.375%)
 * Prec 83.760%
best acc: 85.530000
Epoch: [24][0/391]      Time 0.335 (0.335)      Data 0.312 (0.312)      Loss 0.3712 (0.3712)    Prec 85.15
6% (85.156%)
Epoch: [24][100/391]    Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.2435 (0.2926)    Prec 91.40
6% (90.068%)
Epoch: [24][200/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.3556 (0.2957)    Prec 88.28
1% (89.898%)
Epoch: [24][300/391]    Time 0.021 (0.021)      Data 0.002 (0.002)      Loss 0.2025 (0.2976)    Prec 94.53
1% (89.823%)
Validation starts
Test: [0/79]    Time 0.198 (0.198)      Loss 0.4024 (0.4024)    Prec 85.938% (85.938%)
 * Prec 85.730%
best acc: 85.730000
Epoch: [25][0/391]      Time 0.340 (0.340)      Data 0.319 (0.319)      Loss 0.2340 (0.2340)    Prec 91.40
6% (91.406%)
Epoch: [25][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.3043 (0.2727)    Prec 89.84
4% (90.903%)
Epoch: [25][200/391]    Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.3425 (0.2876)    Prec 86.71
9% (90.330%)
Epoch: [25][300/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.2114 (0.2929)    Prec 94.53
1% (90.015%)
Validation starts
Test: [0/79]    Time 0.199 (0.199)      Loss 0.3631 (0.3631)    Prec 89.062% (89.062%)
 * Prec 86.450%
best acc: 86.450000
Epoch: [26][0/391]      Time 0.342 (0.342)      Data 0.320 (0.320)      Loss 0.2756 (0.2756)    Prec 90.62
5% (90.625%)
Epoch: [26][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.2341 (0.2810)    Prec 92.18
8% (90.432%)
Epoch: [26][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.3135 (0.2783)    Prec 89.06
2% (90.594%)
```

```
Epoch: [26][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.2105 (0.2778)    Prec 92.18
8% (90.480%)
Validation starts
Test: [0/79]    Time 0.212 (0.212)      Loss 0.3823 (0.3823)    Prec 88.281% (88.281%)
 * Prec 86.410%
best acc: 86.450000
Epoch: [27][0/391]      Time 0.418 (0.418)      Data 0.393 (0.393)      Loss 0.2108 (0.2108)    Prec 90.62
5% (90.625%)
Epoch: [27][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.3648 (0.2676)    Prec 89.06
2% (90.958%)
Epoch: [27][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.3018 (0.2710)    Prec 90.62
5% (90.780%)
Epoch: [27][300/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2363 (0.2723)    Prec 92.96
9% (90.760%)
Validation starts
Test: [0/79]    Time 0.192 (0.192)      Loss 0.3551 (0.3551)    Prec 87.500% (87.500%)
 * Prec 85.910%
best acc: 86.450000
Epoch: [28][0/391]      Time 0.306 (0.306)      Data 0.282 (0.282)      Loss 0.2251 (0.2251)    Prec 91.40
6% (91.406%)
Epoch: [28][100/391]    Time 0.021 (0.023)      Data 0.001 (0.004)      Loss 0.2626 (0.2453)    Prec 92.18
8% (91.785%)
Epoch: [28][200/391]    Time 0.021 (0.021)      Data 0.001 (0.003)      Loss 0.2671 (0.2519)    Prec 94.53
1% (91.356%)
Epoch: [28][300/391]    Time 0.021 (0.021)      Data 0.001 (0.002)      Loss 0.1848 (0.2534)    Prec 92.18
8% (91.360%)
Validation starts
Test: [0/79]    Time 0.202 (0.202)      Loss 0.3053 (0.3053)    Prec 91.406% (91.406%)
 * Prec 85.250%
best acc: 86.450000
Epoch: [29][0/391]      Time 0.301 (0.301)      Data 0.279 (0.279)      Loss 0.3254 (0.3254)    Prec 88.28
1% (88.281%)
Epoch: [29][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.2258 (0.2395)    Prec 91.40
6% (91.909%)
Epoch: [29][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2934 (0.2517)    Prec 87.50
0% (91.379%)
Epoch: [29][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.2821 (0.2553)    Prec 89.84
4% (91.271%)
Validation starts
Test: [0/79]    Time 0.202 (0.202)      Loss 0.3592 (0.3592)    Prec 88.281% (88.281%)
 * Prec 87.360%
best acc: 87.360000
Epoch: [30][0/391]      Time 0.359 (0.359)      Data 0.336 (0.336)      Loss 0.2583 (0.2583)    Prec 92.18
8% (92.188%)
Epoch: [30][100/391]    Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.3061 (0.2454)    Prec 89.84
4% (91.770%)
Epoch: [30][200/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.2618 (0.2529)    Prec 92.96
9% (91.395%)
Epoch: [30][300/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.3543 (0.2510)    Prec 85.93
8% (91.466%)
Validation starts
Test: [0/79]    Time 0.196 (0.196)      Loss 0.4396 (0.4396)    Prec 88.281% (88.281%)
 * Prec 84.760%
best acc: 87.360000
Epoch: [31][0/391]      Time 0.405 (0.405)      Data 0.378 (0.378)      Loss 0.2094 (0.2094)    Prec 91.40
6% (91.406%)
Epoch: [31][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.1808 (0.2245)    Prec 94.53
1% (92.319%)
Epoch: [31][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.3192 (0.2364)    Prec 85.15
6% (91.892%)
Epoch: [31][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1974 (0.2386)    Prec 93.75
0% (91.783%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.2548 (0.2548)    Prec 91.406% (91.406%)
 * Prec 87.030%
best acc: 87.360000
Epoch: [32][0/391]      Time 0.315 (0.315)      Data 0.291 (0.291)      Loss 0.1527 (0.1527)    Prec 94.53
1% (94.531%)
Epoch: [32][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.2744 (0.2201)    Prec 91.40
6% (92.489%)
Epoch: [32][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.1796 (0.2341)    Prec 93.75
0% (92.009%)
Epoch: [32][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.2441 (0.2383)    Prec 92.18
8% (91.866%)
Validation starts
Test: [0/79]    Time 0.198 (0.198)      Loss 0.2738 (0.2738)    Prec 89.062% (89.062%)
 * Prec 86.330%
best acc: 87.360000
Epoch: [33][0/391]      Time 0.298 (0.298)      Data 0.273 (0.273)      Loss 0.1369 (0.1369)    Prec 97.65
6% (97.656%)
Epoch: [33][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1373 (0.2154)    Prec 93.75
0% (92.597%)
```

```
Epoch: [33][200/391]    Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.1413 (0.2232)    Prec 96.09
4% (92.324%)
Epoch: [33][300/391]    Time 0.021 (0.022)      Data 0.001 (0.002)      Loss 0.1844 (0.2300)    Prec 92.96
9% (92.112%)
Validation starts
Test: [0/79]    Time 0.199 (0.199)      Loss 0.3492 (0.3492)    Prec 90.625% (90.625%)
 * Prec 86.530%
best acc: 87.360000
Epoch: [34][0/391]      Time 0.400 (0.400)      Data 0.376 (0.376)      Loss 0.1864 (0.1864)    Prec 94.53
1% (94.531%)
Epoch: [34][100/391]    Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.2225 (0.2109)    Prec 92.18
8% (92.613%)
Epoch: [34][200/391]    Time 0.021 (0.023)      Data 0.001 (0.003)      Loss 0.1829 (0.2180)    Prec 92.96
9% (92.460%)
Epoch: [34][300/391]    Time 0.021 (0.022)      Data 0.001 (0.003)      Loss 0.1570 (0.2209)    Prec 94.53
1% (92.369%)
Validation starts
Test: [0/79]    Time 0.197 (0.197)      Loss 0.3607 (0.3607)    Prec 90.625% (90.625%)
 * Prec 87.150%
best acc: 87.360000
Epoch: [35][0/391]      Time 0.300 (0.300)      Data 0.279 (0.279)      Loss 0.2901 (0.2901)    Prec 90.62
5% (90.625%)
Epoch: [35][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.2880 (0.1990)    Prec 87.50
0% (93.131%)
Epoch: [35][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.2521 (0.2168)    Prec 90.62
5% (92.471%)
Epoch: [35][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.2608 (0.2175)    Prec 90.62
5% (92.553%)
Validation starts
Test: [0/79]    Time 0.196 (0.196)      Loss 0.2149 (0.2149)    Prec 92.188% (92.188%)
 * Prec 87.640%
best acc: 87.640000
Epoch: [36][0/391]      Time 0.363 (0.363)      Data 0.342 (0.342)      Loss 0.1701 (0.1701)    Prec 92.96
9% (92.969%)
Epoch: [36][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.3164 (0.2185)    Prec 85.93
8% (92.342%)
Epoch: [36][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1947 (0.2113)    Prec 92.96
9% (92.654%)
Epoch: [36][300/391]    Time 0.021 (0.021)      Data 0.001 (0.003)      Loss 0.1284 (0.2131)    Prec 96.87
5% (92.624%)
Validation starts
Test: [0/79]    Time 0.197 (0.197)      Loss 0.4657 (0.4657)    Prec 84.375% (84.375%)
 * Prec 86.250%
best acc: 87.640000
Epoch: [37][0/391]      Time 0.324 (0.324)      Data 0.300 (0.300)      Loss 0.2231 (0.2231)    Prec 91.40
6% (91.406%)
Epoch: [37][100/391]    Time 0.021 (0.024)      Data 0.002 (0.004)      Loss 0.2075 (0.2031)    Prec 93.75
0% (93.100%)
Epoch: [37][200/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.2252 (0.2094)    Prec 92.18
8% (93.000%)
Epoch: [37][300/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.1696 (0.2108)    Prec 94.53
1% (92.901%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.3721 (0.3721)    Prec 84.375% (84.375%)
 * Prec 87.780%
best acc: 87.780000
Epoch: [38][0/391]      Time 0.366 (0.366)      Data 0.344 (0.344)      Loss 0.1839 (0.1839)    Prec 95.31
2% (95.312%)
Epoch: [38][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.2117 (0.1882)    Prec 93.75
0% (93.688%)
Epoch: [38][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1329 (0.1920)    Prec 95.31
2% (93.466%)
Epoch: [38][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.2224 (0.1946)    Prec 93.75
0% (93.355%)
Validation starts
Test: [0/79]    Time 0.216 (0.216)      Loss 0.3605 (0.3605)    Prec 89.062% (89.062%)
 * Prec 87.530%
best acc: 87.780000
Epoch: [39][0/391]      Time 0.377 (0.377)      Data 0.354 (0.354)      Loss 0.1795 (0.1795)    Prec 94.53
1% (94.531%)
Epoch: [39][100/391]    Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.1856 (0.1926)    Prec 93.75
0% (93.325%)
Epoch: [39][200/391]    Time 0.021 (0.023)      Data 0.001 (0.003)      Loss 0.2422 (0.1965)    Prec 92.96
9% (93.295%)
Epoch: [39][300/391]    Time 0.021 (0.022)      Data 0.001 (0.003)      Loss 0.2577 (0.2031)    Prec 93.75
0% (93.036%)
Validation starts
Test: [0/79]    Time 0.200 (0.200)      Loss 0.3342 (0.3342)    Prec 87.500% (87.500%)
 * Prec 86.500%
best acc: 87.780000
Epoch: [40][0/391]      Time 0.347 (0.347)      Data 0.323 (0.323)      Loss 0.2767 (0.2767)    Prec 91.40
6% (91.406%)
```

```
Epoch: [40][100/391]    Time 0.021 (0.024)      Data 0.001 (0.005)      Loss 0.0923 (0.1923)    Prec 96.87
5% (93.209%)
Epoch: [40][200/391]    Time 0.021 (0.022)      Data 0.001 (0.003)      Loss 0.1600 (0.1970)    Prec 94.53
1% (93.218%)
Epoch: [40][300/391]    Time 0.021 (0.022)      Data 0.001 (0.003)      Loss 0.1865 (0.2013)    Prec 94.53
1% (93.106%)
Validation starts
Test: [0/79]    Time 0.188 (0.188)      Loss 0.3392 (0.3392)    Prec 89.844% (89.844%)
 * Prec 85.600%
best acc: 87.780000
Epoch: [41][0/391]      Time 0.311 (0.311)      Data 0.287 (0.287)      Loss 0.2360 (0.2360)    Prec 93.75
0% (93.750%)
Epoch: [41][100/391]    Time 0.021 (0.024)      Data 0.001 (0.004)      Loss 0.1614 (0.1921)    Prec 95.31
2% (93.448%)
Epoch: [41][200/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.1960 (0.1960)    Prec 92.96
9% (93.377%)
Epoch: [41][300/391]    Time 0.021 (0.022)      Data 0.001 (0.002)      Loss 0.2438 (0.1978)    Prec 91.40
6% (93.301%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.2464 (0.2464)    Prec 92.188% (92.188%)
 * Prec 86.070%
best acc: 87.780000
Epoch: [42][0/391]      Time 0.418 (0.418)      Data 0.391 (0.391)      Loss 0.0861 (0.0861)    Prec 96.09
4% (96.094%)
Epoch: [42][100/391]    Time 0.021 (0.024)      Data 0.001 (0.005)      Loss 0.2536 (0.1664)    Prec 90.62
5% (94.253%)
Epoch: [42][200/391]    Time 0.021 (0.022)      Data 0.001 (0.003)      Loss 0.2638 (0.1801)    Prec 90.62
5% (93.781%)
Epoch: [42][300/391]    Time 0.021 (0.021)      Data 0.001 (0.003)      Loss 0.2554 (0.1833)    Prec 90.62
5% (93.740%)
Validation starts
Test: [0/79]    Time 0.200 (0.200)      Loss 0.2722 (0.2722)    Prec 91.406% (91.406%)
 * Prec 88.300%
best acc: 88.300000
Epoch: [43][0/391]      Time 0.369 (0.369)      Data 0.349 (0.349)      Loss 0.1449 (0.1449)    Prec 96.09
4% (96.094%)
Epoch: [43][100/391]    Time 0.020 (0.023)      Data 0.001 (0.005)      Loss 0.2240 (0.1810)    Prec 91.40
6% (93.905%)
Epoch: [43][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1043 (0.1808)    Prec 96.09
4% (94.003%)
Epoch: [43][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1743 (0.1831)    Prec 93.75
0% (93.893%)
Validation starts
Test: [0/79]    Time 0.203 (0.203)      Loss 0.3214 (0.3214)    Prec 89.062% (89.062%)
 * Prec 87.160%
best acc: 88.300000
Epoch: [44][0/391]      Time 0.324 (0.324)      Data 0.303 (0.303)      Loss 0.2263 (0.2263)    Prec 94.53
1% (94.531%)
Epoch: [44][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1207 (0.1798)    Prec 95.31
2% (93.866%)
Epoch: [44][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.1730 (0.1865)    Prec 94.53
1% (93.560%)
Epoch: [44][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1428 (0.1866)    Prec 94.53
1% (93.644%)
Validation starts
Test: [0/79]    Time 0.205 (0.205)      Loss 0.2855 (0.2855)    Prec 89.844% (89.844%)
 * Prec 87.830%
best acc: 88.300000
Epoch: [45][0/391]      Time 0.260 (0.260)      Data 0.212 (0.212)      Loss 0.1075 (0.1075)    Prec 96.09
4% (96.094%)
Epoch: [45][100/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2845 (0.1606)    Prec 91.40
6% (94.616%)
Epoch: [45][200/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1353 (0.1764)    Prec 96.09
4% (93.995%)
Epoch: [45][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1692 (0.1805)    Prec 92.96
9% (93.856%)
Validation starts
Test: [0/79]    Time 0.207 (0.207)      Loss 0.2204 (0.2204)    Prec 93.750% (93.750%)
 * Prec 88.280%
best acc: 88.300000
Epoch: [46][0/391]      Time 0.306 (0.306)      Data 0.284 (0.284)      Loss 0.1723 (0.1723)    Prec 94.53
1% (94.531%)
Epoch: [46][100/391]    Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.2117 (0.1721)    Prec 89.84
4% (94.152%)
Epoch: [46][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1849 (0.1700)    Prec 92.96
9% (94.271%)
Epoch: [46][300/391]    Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.1974 (0.1770)    Prec 92.96
9% (94.067%)
Validation starts
Test: [0/79]    Time 0.205 (0.205)      Loss 0.3572 (0.3572)    Prec 90.625% (90.625%)
 * Prec 86.080%
best acc: 88.300000
```

```
Epoch: [47][0/391]      Time 0.303 (0.303)      Data 0.280 (0.280)      Loss 0.2202 (0.2202)    Prec 89.84
4% (89.844%)
Epoch: [47][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.2170 (0.1720)    Prec 91.40
6% (94.261%)
Epoch: [47][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1349 (0.1759)    Prec 94.53
1% (94.127%)
Epoch: [47][300/391]    Time 0.020 (0.021)      Data 0.002 (0.002)      Loss 0.1823 (0.1815)    Prec 93.75
0% (93.893%)
Validation starts
Test: [0/79]    Time 0.190 (0.190)      Loss 0.3632 (0.3632)    Prec 89.062% (89.062%)
 * Prec 87.830%
best acc: 88.300000
Epoch: [48][0/391]      Time 0.434 (0.434)      Data 0.407 (0.407)      Loss 0.2487 (0.2487)    Prec 93.75
0% (93.750%)
Epoch: [48][100/391]    Time 0.021 (0.025)      Data 0.002 (0.006)      Loss 0.1544 (0.1567)    Prec 96.87
5% (94.879%)
Epoch: [48][200/391]    Time 0.021 (0.023)      Data 0.002 (0.004)      Loss 0.1373 (0.1685)    Prec 95.31
2% (94.422%)
Epoch: [48][300/391]    Time 0.021 (0.022)      Data 0.001 (0.003)      Loss 0.1352 (0.1715)    Prec 96.09
4% (94.295%)
Validation starts
Test: [0/79]    Time 0.200 (0.200)      Loss 0.2746 (0.2746)    Prec 89.062% (89.062%)
 * Prec 88.280%
best acc: 88.300000
Epoch: [49][0/391]      Time 0.325 (0.325)      Data 0.303 (0.303)      Loss 0.2677 (0.2677)    Prec 88.28
1% (88.281%)
Epoch: [49][100/391]    Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.2170 (0.1575)    Prec 92.18
8% (94.570%)
Epoch: [49][200/391]    Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.1336 (0.1620)    Prec 94.53
1% (94.454%)
Epoch: [49][300/391]    Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.2912 (0.1698)    Prec 92.18
8% (94.178%)
Validation starts
Test: [0/79]    Time 0.210 (0.210)      Loss 0.3709 (0.3709)    Prec 86.719% (86.719%)
 * Prec 87.060%
best acc: 88.300000
Epoch: [50][0/391]      Time 0.378 (0.378)      Data 0.353 (0.353)      Loss 0.0738 (0.0738)    Prec 98.43
8% (98.438%)
Epoch: [50][100/391]    Time 0.021 (0.024)      Data 0.001 (0.005)      Loss 0.1064 (0.1616)    Prec 96.87
5% (94.609%)
Epoch: [50][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2526 (0.1672)    Prec 92.96
9% (94.430%)
Epoch: [50][300/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1356 (0.1686)    Prec 95.31
2% (94.396%)
Validation starts
Test: [0/79]    Time 0.209 (0.209)      Loss 0.4208 (0.4208)    Prec 87.500% (87.500%)
 * Prec 86.410%
best acc: 88.300000
Epoch: [51][0/391]      Time 0.315 (0.315)      Data 0.293 (0.293)      Loss 0.2368 (0.2368)    Prec 91.40
6% (91.406%)
Epoch: [51][100/391]    Time 0.020 (0.024)      Data 0.002 (0.004)      Loss 0.1506 (0.1485)    Prec 94.53
1% (94.918%)
Epoch: [51][200/391]    Time 0.020 (0.023)      Data 0.001 (0.003)      Loss 0.2050 (0.1633)    Prec 89.84
4% (94.450%)
Epoch: [51][300/391]    Time 0.021 (0.022)      Data 0.002 (0.002)      Loss 0.1463 (0.1652)    Prec 94.53
1% (94.451%)
Validation starts
Test: [0/79]    Time 0.191 (0.191)      Loss 0.3538 (0.3538)    Prec 89.062% (89.062%)
 * Prec 86.760%
best acc: 88.300000
Epoch: [52][0/391]      Time 0.313 (0.313)      Data 0.293 (0.293)      Loss 0.1312 (0.1312)    Prec 96.87
5% (96.875%)
Epoch: [52][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.2019 (0.1648)    Prec 91.40
6% (94.384%)
Epoch: [52][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1267 (0.1636)    Prec 94.53
1% (94.426%)
Epoch: [52][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.2066 (0.1639)    Prec 93.75
0% (94.498%)
Validation starts
Test: [0/79]    Time 0.291 (0.291)      Loss 0.3539 (0.3539)    Prec 86.719% (86.719%)
 * Prec 87.430%
best acc: 88.300000
Epoch: [53][0/391]      Time 0.349 (0.349)      Data 0.327 (0.327)      Loss 0.1490 (0.1490)    Prec 94.53
1% (94.531%)
Epoch: [53][100/391]    Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.2196 (0.1460)    Prec 92.18
8% (94.833%)
Epoch: [53][200/391]    Time 0.021 (0.023)      Data 0.002 (0.003)      Loss 0.1839 (0.1567)    Prec 93.75
0% (94.632%)
Epoch: [53][300/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1494 (0.1599)    Prec 95.31
2% (94.575%)
Validation starts
Test: [0/79]    Time 0.198 (0.198)      Loss 0.2710 (0.2710)    Prec 91.406% (91.406%)
```

```
 * Prec 87.620%
best acc: 88.300000
Epoch: [54][0/391]      Time 0.303 (0.303)      Data 0.281 (0.281)      Loss 0.1307 (0.1307)    Prec 95.31
2% (95.312%)
Epoch: [54][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1316 (0.1495)    Prec 96.09
4% (94.918%)
Epoch: [54][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2746 (0.1535)    Prec 92.96
9% (94.788%)
Epoch: [54][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1777 (0.1594)    Prec 95.31
2% (94.518%)
Validation starts
Test: [0/79]    Time 0.208 (0.208)      Loss 0.3251 (0.3251)    Prec 89.844% (89.844%)
 * Prec 87.430%
best acc: 88.300000
Epoch: [55][0/391]      Time 0.312 (0.312)      Data 0.290 (0.290)      Loss 0.2038 (0.2038)    Prec 94.53
1% (94.531%)
Epoch: [55][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1505 (0.1465)    Prec 95.31
2% (95.150%)
Epoch: [55][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1566 (0.1552)    Prec 94.53
1% (94.807%)
Epoch: [55][300/391]    Time 0.021 (0.021)      Data 0.001 (0.002)      Loss 0.1442 (0.1558)    Prec 93.75
0% (94.760%)
Validation starts
Test: [0/79]    Time 0.210 (0.210)      Loss 0.2244 (0.2244)    Prec 95.312% (95.312%)
 * Prec 87.790%
best acc: 88.300000
Epoch: [56][0/391]      Time 0.331 (0.331)      Data 0.308 (0.308)      Loss 0.0684 (0.0684)    Prec 96.87
5% (96.875%)
Epoch: [56][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1275 (0.1532)    Prec 96.09
4% (94.771%)
Epoch: [56][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.2502 (0.1539)    Prec 92.18
8% (94.862%)
Epoch: [56][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.2150 (0.1574)    Prec 92.18
8% (94.703%)
Validation starts
Test: [0/79]    Time 0.207 (0.207)      Loss 0.1670 (0.1670)    Prec 95.312% (95.312%)
 * Prec 87.570%
best acc: 88.300000
Epoch: [57][0/391]      Time 0.388 (0.388)      Data 0.321 (0.321)      Loss 0.1889 (0.1889)    Prec 92.18
8% (92.188%)
Epoch: [57][100/391]    Time 0.020 (0.024)      Data 0.001 (0.004)      Loss 0.0885 (0.1401)    Prec 96.87
5% (95.282%)
Epoch: [57][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2494 (0.1481)    Prec 93.75
0% (94.951%)
Epoch: [57][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1000 (0.1519)    Prec 96.87
5% (94.806%)
Validation starts
Test: [0/79]    Time 0.199 (0.199)      Loss 0.4388 (0.4388)    Prec 88.281% (88.281%)
 * Prec 87.840%
best acc: 88.300000
Epoch: [58][0/391]      Time 0.306 (0.306)      Data 0.284 (0.284)      Loss 0.1431 (0.1431)    Prec 92.96
9% (92.969%)
Epoch: [58][100/391]    Time 0.018 (0.023)      Data 0.001 (0.004)      Loss 0.1788 (0.1456)    Prec 92.96
9% (95.088%)
Epoch: [58][200/391]    Time 0.018 (0.022)      Data 0.001 (0.003)      Loss 0.1224 (0.1533)    Prec 96.87
5% (94.904%)
Epoch: [58][300/391]    Time 0.018 (0.021)      Data 0.001 (0.002)      Loss 0.1711 (0.1567)    Prec 92.96
9% (94.760%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.3232 (0.3232)    Prec 89.062% (89.062%)
 * Prec 87.050%
best acc: 88.300000
Epoch: [59][0/391]      Time 0.309 (0.309)      Data 0.286 (0.286)      Loss 0.1079 (0.1079)    Prec 95.31
2% (95.312%)
Epoch: [59][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1660 (0.1289)    Prec 94.53
1% (95.614%)
Epoch: [59][200/391]    Time 0.020 (0.021)      Data 0.002 (0.003)      Loss 0.1406 (0.1408)    Prec 95.31
2% (95.227%)
Epoch: [59][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1608 (0.1476)    Prec 94.53
1% (94.962%)
Validation starts
Test: [0/79]    Time 0.193 (0.193)      Loss 0.3209 (0.3209)    Prec 91.406% (91.406%)
 * Prec 88.050%
best acc: 88.300000
Epoch: [60][0/391]      Time 0.319 (0.319)      Data 0.298 (0.298)      Loss 0.0914 (0.0914)    Prec 96.87
5% (96.875%)
Epoch: [60][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1627 (0.1377)    Prec 94.53
1% (95.243%)
Epoch: [60][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1598 (0.1443)    Prec 96.09
4% (95.079%)
Epoch: [60][300/391]    Time 0.020 (0.021)      Data 0.002 (0.002)      Loss 0.0974 (0.1498)    Prec 96.09
4% (94.908%)
```

```
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.3721 (0.3721)    Prec 87.500% (87.500%)
 * Prec 86.870%
best acc: 88.300000
Epoch: [61][0/391]      Time 0.334 (0.334)      Data 0.312 (0.312)      Loss 0.1163 (0.1163)    Prec 97.65
6% (97.656%)
Epoch: [61][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1530 (0.1458)    Prec 94.53
1% (95.104%)
Epoch: [61][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.1127 (0.1499)    Prec 96.09
4% (94.893%)
Epoch: [61][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.2287 (0.1489)    Prec 93.75
0% (94.921%)
Validation starts
Test: [0/79]    Time 0.212 (0.212)      Loss 0.2620 (0.2620)    Prec 92.969% (92.969%)
 * Prec 88.250%
best acc: 88.300000
Epoch: [62][0/391]      Time 0.318 (0.318)      Data 0.298 (0.298)      Loss 0.1942 (0.1942)    Prec 93.75
0% (93.750%)
Epoch: [62][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1017 (0.1411)    Prec 95.31
2% (95.436%)
Epoch: [62][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.1997 (0.1427)    Prec 94.53
1% (95.274%)
Epoch: [62][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1824 (0.1428)    Prec 95.31
2% (95.271%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.3081 (0.3081)    Prec 86.719% (86.719%)
 * Prec 88.650%
best acc: 88.650000
Epoch: [63][0/391]      Time 0.383 (0.383)      Data 0.361 (0.361)      Loss 0.1677 (0.1677)    Prec 94.53
1% (94.531%)
Epoch: [63][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.0944 (0.1430)    Prec 96.09
4% (95.243%)
Epoch: [63][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1030 (0.1459)    Prec 97.65
6% (95.075%)
Epoch: [63][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1383 (0.1472)    Prec 93.75
0% (95.014%)
Validation starts
Test: [0/79]    Time 0.213 (0.213)      Loss 0.3583 (0.3583)    Prec 89.844% (89.844%)
 * Prec 86.980%
best acc: 88.650000
Epoch: [64][0/391]      Time 0.314 (0.314)      Data 0.293 (0.293)      Loss 0.1099 (0.1099)    Prec 96.87
5% (96.875%)
Epoch: [64][100/391]    Time 0.020 (0.024)      Data 0.001 (0.004)      Loss 0.1608 (0.1528)    Prec 95.31
2% (94.988%)
Epoch: [64][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1919 (0.1451)    Prec 94.53
1% (95.141%)
Epoch: [64][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1700 (0.1460)    Prec 93.75
0% (95.019%)
Validation starts
Test: [0/79]    Time 0.202 (0.202)      Loss 0.3486 (0.3486)    Prec 89.844% (89.844%)
 * Prec 86.910%
best acc: 88.650000
Epoch: [65][0/391]      Time 0.306 (0.306)      Data 0.285 (0.285)      Loss 0.1158 (0.1158)    Prec 94.53
1% (94.531%)
Epoch: [65][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1194 (0.1243)    Prec 95.31
2% (95.808%)
Epoch: [65][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.2213 (0.1367)    Prec 93.75
0% (95.429%)
Epoch: [65][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.2503 (0.1457)    Prec 91.40
6% (95.170%)
Validation starts
Test: [0/79]    Time 0.202 (0.202)      Loss 0.3339 (0.3339)    Prec 90.625% (90.625%)
 * Prec 87.710%
best acc: 88.650000
Epoch: [66][0/391]      Time 0.360 (0.360)      Data 0.334 (0.334)      Loss 0.2296 (0.2296)    Prec 92.96
9% (92.969%)
Epoch: [66][100/391]    Time 0.021 (0.024)      Data 0.001 (0.005)      Loss 0.1225 (0.1214)    Prec 95.31
2% (96.047%)
Epoch: [66][200/391]    Time 0.021 (0.023)      Data 0.001 (0.003)      Loss 0.1528 (0.1321)    Prec 93.75
0% (95.608%)
Epoch: [66][300/391]    Time 0.021 (0.022)      Data 0.001 (0.002)      Loss 0.0885 (0.1364)    Prec 97.65
6% (95.398%)
Validation starts
Test: [0/79]    Time 0.214 (0.214)      Loss 0.3211 (0.3211)    Prec 90.625% (90.625%)
 * Prec 87.320%
best acc: 88.650000
Epoch: [67][0/391]      Time 0.339 (0.339)      Data 0.316 (0.316)      Loss 0.1711 (0.1711)    Prec 94.53
1% (94.531%)
Epoch: [67][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1410 (0.1436)    Prec 94.53
1% (95.011%)
Epoch: [67][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2026 (0.1473)    Prec 89.84
4% (94.908%)
```

```
Epoch: [67][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1710 (0.1461)    Prec 94.53
1% (95.037%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.3136 (0.3136)    Prec 89.844% (89.844%)
 * Prec 88.660%
best acc: 88.660000
Epoch: [68][0/391]      Time 0.483 (0.483)      Data 0.457 (0.457)      Loss 0.0636 (0.0636)    Prec 97.65
6% (97.656%)
Epoch: [68][100/391]    Time 0.020 (0.025)      Data 0.001 (0.006)      Loss 0.0905 (0.1311)    Prec 97.65
6% (95.591%)
Epoch: [68][200/391]    Time 0.020 (0.022)      Data 0.001 (0.004)      Loss 0.0869 (0.1381)    Prec 96.09
4% (95.309%)
Epoch: [68][300/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1355 (0.1403)    Prec 95.31
2% (95.198%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.2870 (0.2870)    Prec 92.969% (92.969%)
 * Prec 88.640%
best acc: 88.660000
Epoch: [69][0/391]      Time 0.424 (0.424)      Data 0.403 (0.403)      Loss 0.1009 (0.1009)    Prec 96.09
4% (96.094%)
Epoch: [69][100/391]    Time 0.021 (0.025)      Data 0.001 (0.005)      Loss 0.1612 (0.1315)    Prec 93.75
0% (95.614%)
Epoch: [69][200/391]    Time 0.021 (0.023)      Data 0.001 (0.003)      Loss 0.1229 (0.1293)    Prec 96.87
5% (95.756%)
Epoch: [69][300/391]    Time 0.021 (0.022)      Data 0.001 (0.003)      Loss 0.2065 (0.1361)    Prec 95.31
2% (95.453%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.2643 (0.2643)    Prec 90.625% (90.625%)
 * Prec 87.910%
best acc: 88.660000
Epoch: [70][0/391]      Time 0.342 (0.342)      Data 0.319 (0.319)      Loss 0.1068 (0.1068)    Prec 96.87
5% (96.875%)
Epoch: [70][100/391]    Time 0.021 (0.024)      Data 0.001 (0.005)      Loss 0.1444 (0.1277)    Prec 96.09
4% (95.521%)
Epoch: [70][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1596 (0.1347)    Prec 92.96
9% (95.379%)
Epoch: [70][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1201 (0.1366)    Prec 94.53
1% (95.331%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.3576 (0.3576)    Prec 86.719% (86.719%)
 * Prec 88.460%
best acc: 88.660000
Epoch: [71][0/391]      Time 0.319 (0.319)      Data 0.298 (0.298)      Loss 0.0852 (0.0852)    Prec 98.43
8% (98.438%)
Epoch: [71][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1804 (0.1321)    Prec 95.31
2% (95.374%)
Epoch: [71][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1293 (0.1393)    Prec 96.09
4% (95.188%)
Epoch: [71][300/391]    Time 0.020 (0.021)      Data 0.002 (0.002)      Loss 0.1269 (0.1387)    Prec 95.31
2% (95.263%)
Validation starts
Test: [0/79]    Time 0.201 (0.201)      Loss 0.2952 (0.2952)    Prec 92.969% (92.969%)
 * Prec 89.210%
best acc: 89.210000
Epoch: [72][0/391]      Time 0.352 (0.352)      Data 0.330 (0.330)      Loss 0.0684 (0.0684)    Prec 99.21
9% (99.219%)
Epoch: [72][100/391]    Time 0.020 (0.023)      Data 0.001 (0.005)      Loss 0.1225 (0.1302)    Prec 96.09
4% (95.583%)
Epoch: [72][200/391]    Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.1613 (0.1310)    Prec 95.31
2% (95.553%)
Epoch: [72][300/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.0961 (0.1320)    Prec 96.09
4% (95.551%)
Validation starts
Test: [0/79]    Time 0.211 (0.211)      Loss 0.2762 (0.2762)    Prec 91.406% (91.406%)
 * Prec 84.460%
best acc: 89.210000
Epoch: [73][0/391]      Time 0.334 (0.334)      Data 0.313 (0.313)      Loss 0.1118 (0.1118)    Prec 96.87
5% (96.875%)
Epoch: [73][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1716 (0.1279)    Prec 94.53
1% (95.692%)
Epoch: [73][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1178 (0.1326)    Prec 96.09
4% (95.550%)
Epoch: [73][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1155 (0.1332)    Prec 96.09
4% (95.551%)
Validation starts
Test: [0/79]    Time 0.200 (0.200)      Loss 0.2680 (0.2680)    Prec 92.969% (92.969%)
 * Prec 89.220%
best acc: 89.220000
Epoch: [74][0/391]      Time 0.377 (0.377)      Data 0.356 (0.356)      Loss 0.1039 (0.1039)    Prec 96.87
5% (96.875%)
Epoch: [74][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.1150 (0.1184)    Prec 97.65
6% (96.086%)
```

```
Epoch: [74][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1013 (0.1259)    Prec 95.31
2% (95.779%)
Epoch: [74][300/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.0652 (0.1340)    Prec 99.21
9% (95.453%)
Validation starts
Test: [0/79]    Time 0.216 (0.216)      Loss 0.3084 (0.3084)    Prec 90.625% (90.625%)
 * Prec 88.430%
best acc: 89.220000
Epoch: [75][0/391]      Time 0.315 (0.315)      Data 0.293 (0.293)      Loss 0.0886 (0.0886)    Prec 98.43
8% (98.438%)
Epoch: [75][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1516 (0.1166)    Prec 95.31
2% (95.970%)
Epoch: [75][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1279 (0.1182)    Prec 95.31
2% (96.024%)
Epoch: [75][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.0816 (0.1246)    Prec 97.65
6% (95.782%)
Validation starts
Test: [0/79]    Time 0.286 (0.286)      Loss 0.1827 (0.1827)    Prec 93.750% (93.750%)
 * Prec 88.580%
best acc: 89.220000
Epoch: [76][0/391]      Time 0.342 (0.342)      Data 0.321 (0.321)      Loss 0.0760 (0.0760)    Prec 96.09
4% (96.094%)
Epoch: [76][100/391]    Time 0.020 (0.026)      Data 0.001 (0.005)      Loss 0.1000 (0.1228)    Prec 96.87
5% (95.777%)
Epoch: [76][200/391]    Time 0.027 (0.023)      Data 0.011 (0.003)      Loss 0.1356 (0.1269)    Prec 95.31
2% (95.744%)
Epoch: [76][300/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1296 (0.1301)    Prec 93.75
0% (95.603%)
Validation starts
Test: [0/79]    Time 0.209 (0.209)      Loss 0.2764 (0.2764)    Prec 92.188% (92.188%)
 * Prec 88.360%
best acc: 89.220000
Epoch: [77][0/391]      Time 0.323 (0.323)      Data 0.297 (0.297)      Loss 0.0655 (0.0655)    Prec 98.43
8% (98.438%)
Epoch: [77][100/391]    Time 0.021 (0.023)      Data 0.001 (0.004)      Loss 0.0816 (0.1248)    Prec 97.65
6% (95.792%)
Epoch: [77][200/391]    Time 0.021 (0.021)      Data 0.001 (0.003)      Loss 0.0972 (0.1288)    Prec 98.43
8% (95.705%)
Epoch: [77][300/391]    Time 0.021 (0.021)      Data 0.001 (0.002)      Loss 0.0670 (0.1307)    Prec 98.43
8% (95.647%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.2305 (0.2305)    Prec 91.406% (91.406%)
 * Prec 88.230%
best acc: 89.220000
Epoch: [78][0/391]      Time 0.326 (0.326)      Data 0.305 (0.305)      Loss 0.1305 (0.1305)    Prec 94.53
1% (94.531%)
Epoch: [78][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1012 (0.1231)    Prec 96.09
4% (95.908%)
Epoch: [78][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2100 (0.1252)    Prec 94.53
1% (95.814%)
Epoch: [78][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1146 (0.1279)    Prec 95.31
2% (95.668%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.3550 (0.3550)    Prec 89.844% (89.844%)
 * Prec 87.500%
best acc: 89.220000
Epoch: [79][0/391]      Time 0.340 (0.340)      Data 0.312 (0.312)      Loss 0.1858 (0.1858)    Prec 93.75
0% (93.750%)
Epoch: [79][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1970 (0.1150)    Prec 92.18
8% (96.078%)
Epoch: [79][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1918 (0.1233)    Prec 94.53
1% (95.794%)
Epoch: [79][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1085 (0.1255)    Prec 93.75
0% (95.746%)
Validation starts
Test: [0/79]    Time 0.210 (0.210)      Loss 0.3572 (0.3572)    Prec 89.844% (89.844%)
 * Prec 88.970%
best acc: 89.220000
Epoch: [80][0/391]      Time 0.356 (0.356)      Data 0.335 (0.335)      Loss 0.1290 (0.1290)    Prec 96.09
4% (96.094%)
Epoch: [80][100/391]    Time 0.020 (0.023)      Data 0.001 (0.005)      Loss 0.1877 (0.1063)    Prec 94.53
1% (96.419%)
Epoch: [80][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.0609 (0.1142)    Prec 98.43
8% (96.195%)
Epoch: [80][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.0576 (0.1196)    Prec 98.43
8% (95.995%)
Validation starts
Test: [0/79]    Time 0.220 (0.220)      Loss 0.4044 (0.4044)    Prec 89.844% (89.844%)
 * Prec 86.580%
best acc: 89.220000
Epoch: [81][0/391]      Time 0.339 (0.339)      Data 0.318 (0.318)      Loss 0.1117 (0.1117)    Prec 95.31
2% (95.312%)
```

```
Epoch: [81][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.0574 (0.1045)    Prec 98.43
8% (96.310%)
Epoch: [81][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.1143 (0.1070)    Prec 94.53
1% (96.276%)
Epoch: [81][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1498 (0.1158)    Prec 92.96
9% (95.998%)
Validation starts
Test: [0/79]    Time 0.209 (0.209)      Loss 0.2730 (0.2730)    Prec 92.969% (92.969%)
 * Prec 88.850%
best acc: 89.220000
Epoch: [82][0/391]      Time 0.323 (0.323)      Data 0.300 (0.300)      Loss 0.1529 (0.1529)    Prec 94.53
1% (94.531%)
Epoch: [82][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.0572 (0.1176)    Prec 98.43
8% (96.148%)
Epoch: [82][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1307 (0.1143)    Prec 96.09
4% (96.199%)
Epoch: [82][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1340 (0.1169)    Prec 95.31
2% (96.081%)
Validation starts
Test: [0/79]    Time 0.203 (0.203)      Loss 0.3128 (0.3128)    Prec 88.281% (88.281%)
 * Prec 88.870%
best acc: 89.220000
Epoch: [83][0/391]      Time 0.326 (0.326)      Data 0.303 (0.303)      Loss 0.2423 (0.2423)    Prec 92.96
9% (92.969%)
Epoch: [83][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1461 (0.1091)    Prec 96.09
4% (96.357%)
Epoch: [83][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2586 (0.1181)    Prec 92.18
8% (96.094%)
Epoch: [83][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1185 (0.1194)    Prec 96.87
5% (95.972%)
Validation starts
Test: [0/79]    Time 0.203 (0.203)      Loss 0.4659 (0.4659)    Prec 89.844% (89.844%)
 * Prec 86.210%
best acc: 89.220000
Epoch: [84][0/391]      Time 0.324 (0.324)      Data 0.301 (0.301)      Loss 0.0379 (0.0379)    Prec 98.43
8% (98.438%)
Epoch: [84][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1159 (0.1148)    Prec 96.87
5% (96.241%)
Epoch: [84][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.1077 (0.1123)    Prec 97.65
6% (96.230%)
Epoch: [84][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1650 (0.1165)    Prec 94.53
1% (96.076%)
Validation starts
Test: [0/79]    Time 0.208 (0.208)      Loss 0.3336 (0.3336)    Prec 89.062% (89.062%)
 * Prec 88.160%
best acc: 89.220000
Epoch: [85][0/391]      Time 0.327 (0.327)      Data 0.304 (0.304)      Loss 0.0513 (0.0513)    Prec 98.43
8% (98.438%)
Epoch: [85][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.0616 (0.1054)    Prec 98.43
8% (96.349%)
Epoch: [85][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1285 (0.1115)    Prec 96.87
5% (96.292%)
Epoch: [85][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1908 (0.1167)    Prec 94.53
1% (96.102%)
Validation starts
Test: [0/79]    Time 0.212 (0.212)      Loss 0.2252 (0.2252)    Prec 94.531% (94.531%)
 * Prec 87.620%
best acc: 89.220000
Epoch: [86][0/391]      Time 0.345 (0.345)      Data 0.322 (0.322)      Loss 0.1102 (0.1102)    Prec 97.65
6% (97.656%)
Epoch: [86][100/391]    Time 0.020 (0.023)      Data 0.001 (0.005)      Loss 0.0871 (0.1075)    Prec 96.09
4% (96.434%)
Epoch: [86][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1490 (0.1131)    Prec 94.53
1% (96.214%)
Epoch: [86][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1868 (0.1179)    Prec 92.18
8% (96.011%)
Validation starts
Test: [0/79]    Time 0.274 (0.274)      Loss 0.2331 (0.2331)    Prec 95.312% (95.312%)
 * Prec 88.490%
best acc: 89.220000
Epoch: [87][0/391]      Time 0.397 (0.397)      Data 0.375 (0.375)      Loss 0.0643 (0.0643)    Prec 97.65
6% (97.656%)
Epoch: [87][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.1212 (0.1007)    Prec 97.65
6% (96.488%)
Epoch: [87][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1635 (0.1071)    Prec 94.53
1% (96.319%)
Epoch: [87][300/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.1504 (0.1117)    Prec 92.96
9% (96.177%)
Validation starts
Test: [0/79]    Time 0.185 (0.185)      Loss 0.4988 (0.4988)    Prec 86.719% (86.719%)
 * Prec 86.570%
best acc: 89.220000
```

```
Epoch: [88][0/391]      Time 0.339 (0.339)      Data 0.315 (0.315)      Loss 0.1276 (0.1276)    Prec 96.87
5% (96.875%)
Epoch: [88][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.0983 (0.1161)    Prec 98.43
8% (96.047%)
Epoch: [88][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.1337 (0.1170)    Prec 94.53
1% (95.954%)
Epoch: [88][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.0929 (0.1204)    Prec 98.43
8% (95.899%)
Validation starts
Test: [0/79]    Time 0.213 (0.213)      Loss 0.2768 (0.2768)    Prec 90.625% (90.625%)
 * Prec 89.610%
best acc: 89.610000
Epoch: [89][0/391]      Time 0.372 (0.372)      Data 0.351 (0.351)      Loss 0.0868 (0.0868)    Prec 97.65
6% (97.656%)
Epoch: [89][100/391]    Time 0.020 (0.023)      Data 0.001 (0.005)      Loss 0.1461 (0.1078)    Prec 95.31
2% (96.357%)
Epoch: [89][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1052 (0.1126)    Prec 97.65
6% (96.191%)
Epoch: [89][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1074 (0.1122)    Prec 96.09
4% (96.203%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.4052 (0.4052)    Prec 87.500% (87.500%)
 * Prec 88.030%
best acc: 89.610000
Epoch: [90][0/391]      Time 0.311 (0.311)      Data 0.290 (0.290)      Loss 0.1331 (0.1331)    Prec 94.53
1% (94.531%)
Epoch: [90][100/391]    Time 0.021 (0.023)      Data 0.001 (0.004)      Loss 0.1492 (0.0978)    Prec 97.65
6% (96.658%)
Epoch: [90][200/391]    Time 0.021 (0.021)      Data 0.001 (0.003)      Loss 0.1018 (0.1121)    Prec 95.31
2% (96.249%)
Epoch: [90][300/391]    Time 0.022 (0.021)      Data 0.001 (0.002)      Loss 0.1621 (0.1146)    Prec 95.31
2% (96.143%)
Validation starts
Test: [0/79]    Time 0.181 (0.181)      Loss 0.3883 (0.3883)    Prec 89.062% (89.062%)
 * Prec 86.940%
best acc: 89.610000
Epoch: [91][0/391]      Time 0.401 (0.401)      Data 0.378 (0.378)      Loss 0.1513 (0.1513)    Prec 96.09
4% (96.094%)
Epoch: [91][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.1172 (0.1000)    Prec 96.87
5% (96.720%)
Epoch: [91][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1996 (0.1101)    Prec 94.53
1% (96.288%)
Epoch: [91][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1256 (0.1165)    Prec 96.09
4% (96.151%)
Validation starts
Test: [0/79]    Time 0.209 (0.209)      Loss 0.2326 (0.2326)    Prec 89.844% (89.844%)
 * Prec 88.500%
best acc: 89.610000
Epoch: [92][0/391]      Time 0.320 (0.320)      Data 0.298 (0.298)      Loss 0.1376 (0.1376)    Prec 94.53
1% (94.531%)
Epoch: [92][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1236 (0.1138)    Prec 96.09
4% (96.078%)
Epoch: [92][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2634 (0.1149)    Prec 92.18
8% (96.043%)
Epoch: [92][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.0564 (0.1168)    Prec 99.21
9% (96.013%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.2617 (0.2617)    Prec 89.844% (89.844%)
 * Prec 88.480%
best acc: 89.610000
Epoch: [93][0/391]      Time 0.328 (0.328)      Data 0.305 (0.305)      Loss 0.1535 (0.1535)    Prec 94.53
1% (94.531%)
Epoch: [93][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.0505 (0.1099)    Prec 98.43
8% (96.202%)
Epoch: [93][200/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.0655 (0.1104)    Prec 98.43
8% (96.203%)
Epoch: [93][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1661 (0.1135)    Prec 94.53
1% (96.120%)
Validation starts
Test: [0/79]    Time 0.198 (0.198)      Loss 0.1903 (0.1903)    Prec 93.750% (93.750%)
 * Prec 88.540%
best acc: 89.610000
Epoch: [94][0/391]      Time 0.336 (0.336)      Data 0.313 (0.313)      Loss 0.1518 (0.1518)    Prec 94.53
1% (94.531%)
Epoch: [94][100/391]    Time 0.022 (0.023)      Data 0.002 (0.005)      Loss 0.0852 (0.1035)    Prec 97.65
6% (96.457%)
Epoch: [94][200/391]    Time 0.022 (0.022)      Data 0.001 (0.003)      Loss 0.0600 (0.1125)    Prec 97.65
6% (96.230%)
Epoch: [94][300/391]    Time 0.021 (0.021)      Data 0.001 (0.003)      Loss 0.1504 (0.1151)    Prec 94.53
1% (96.104%)
Validation starts
Test: [0/79]    Time 0.198 (0.198)      Loss 0.2661 (0.2661)    Prec 89.844% (89.844%)
```

```
 * Prec 87.530%
best acc: 89.610000
Epoch: [95][0/391]      Time 0.425 (0.425)      Data 0.398 (0.398)      Loss 0.1624 (0.1624)    Prec 94.53
1% (94.531%)
Epoch: [95][100/391]    Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.1447 (0.1016)    Prec 95.31
2% (96.581%)
Epoch: [95][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1441 (0.1108)    Prec 93.75
0% (96.249%)
Epoch: [95][300/391]    Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.0711 (0.1137)    Prec 96.87
5% (96.161%)
Validation starts
Test: [0/79]    Time 0.200 (0.200)      Loss 0.3427 (0.3427)    Prec 89.844% (89.844%)
 * Prec 86.550%
best acc: 89.610000
Epoch: [96][0/391]      Time 0.327 (0.327)      Data 0.305 (0.305)      Loss 0.0481 (0.0481)    Prec 97.65
6% (97.656%)
Epoch: [96][100/391]    Time 0.021 (0.024)      Data 0.002 (0.005)      Loss 0.0499 (0.1064)    Prec 98.43
8% (96.434%)
Epoch: [96][200/391]    Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.1273 (0.1148)    Prec 96.87
5% (96.121%)
Epoch: [96][300/391]    Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.1163 (0.1164)    Prec 96.09
4% (96.112%)
Validation starts
Test: [0/79]    Time 0.208 (0.208)      Loss 0.3221 (0.3221)    Prec 89.844% (89.844%)
 * Prec 87.610%
best acc: 89.610000
Epoch: [97][0/391]      Time 0.333 (0.333)      Data 0.311 (0.311)      Loss 0.0563 (0.0563)    Prec 97.65
6% (97.656%)
Epoch: [97][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1572 (0.1067)    Prec 94.53
1% (96.450%)
Epoch: [97][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1645 (0.1090)    Prec 93.75
0% (96.276%)
Epoch: [97][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1463 (0.1126)    Prec 94.53
1% (96.200%)
Validation starts
Test: [0/79]    Time 0.188 (0.188)      Loss 0.3119 (0.3119)    Prec 91.406% (91.406%)
 * Prec 86.970%
best acc: 89.610000
Epoch: [98][0/391]      Time 0.373 (0.373)      Data 0.296 (0.296)      Loss 0.1141 (0.1141)    Prec 96.09
4% (96.094%)
Epoch: [98][100/391]    Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1811 (0.0946)    Prec 95.31
2% (96.890%)
Epoch: [98][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1499 (0.1119)    Prec 93.75
0% (96.226%)
Epoch: [98][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1159 (0.1155)    Prec 94.53
1% (96.060%)
Validation starts
Test: [0/79]    Time 0.208 (0.208)      Loss 0.2666 (0.2666)    Prec 90.625% (90.625%)
 * Prec 88.160%
best acc: 89.610000
Epoch: [99][0/391]      Time 0.378 (0.378)      Data 0.355 (0.355)      Loss 0.0762 (0.0762)    Prec 98.43
8% (98.438%)
Epoch: [99][100/391]    Time 0.020 (0.023)      Data 0.001 (0.005)      Loss 0.1092 (0.1074)    Prec 95.31
2% (96.519%)
Epoch: [99][200/391]    Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.1102 (0.1089)    Prec 96.87
5% (96.362%)
Epoch: [99][300/391]    Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.0827 (0.1150)    Prec 97.65
6% (96.130%)
Validation starts
Test: [0/79]    Time 0.212 (0.212)      Loss 0.2617 (0.2617)    Prec 89.062% (89.062%)
 * Prec 88.400%
best acc: 89.610000
Epoch: [100][0/391]     Time 0.330 (0.330)      Data 0.306 (0.306)      Loss 0.0599 (0.0599)    Prec 96.87
5% (96.875%)
Epoch: [100][100/391]   Time 0.020 (0.024)      Data 0.002 (0.005)      Loss 0.1081 (0.1014)    Prec 96.87
5% (96.744%)
Epoch: [100][200/391]   Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.0911 (0.1011)    Prec 96.87
5% (96.661%)
Epoch: [100][300/391]   Time 0.020 (0.021)      Data 0.001 (0.003)      Loss 0.1149 (0.1075)    Prec 95.31
2% (96.377%)
Validation starts
Test: [0/79]    Time 0.272 (0.272)      Loss 0.3835 (0.3835)    Prec 89.844% (89.844%)
 * Prec 89.090%
best acc: 89.610000
Epoch: [101][0/391]     Time 0.338 (0.338)      Data 0.316 (0.316)      Loss 0.0658 (0.0658)    Prec 97.65
6% (97.656%)
Epoch: [101][100/391]   Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.0607 (0.1047)    Prec 96.87
5% (96.496%)
Epoch: [101][200/391]   Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.0756 (0.1104)    Prec 97.65
6% (96.241%)
Epoch: [101][300/391]   Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.0709 (0.1097)    Prec 95.31
2% (96.281%)
```

```
Validation starts
Test: [0/79]    Time 0.205 (0.205)      Loss 0.4179 (0.4179)    Prec 90.625% (90.625%)
 * Prec 86.550%
best acc: 89.610000
Epoch: [102][0/391]     Time 0.393 (0.393)      Data 0.371 (0.371)      Loss 0.1934 (0.1934)    Prec 95.31
2% (95.312%)
Epoch: [102][100/391]   Time 0.020 (0.024)      Data 0.002 (0.005)      Loss 0.0840 (0.1109)    Prec 97.65
6% (96.341%)
Epoch: [102][200/391]   Time 0.021 (0.022)      Data 0.002 (0.004)      Loss 0.1031 (0.1087)    Prec 96.09
4% (96.432%)
Epoch: [102][300/391]   Time 0.021 (0.022)      Data 0.002 (0.003)      Loss 0.0917 (0.1095)    Prec 97.65
6% (96.410%)
Validation starts
Test: [0/79]    Time 0.198 (0.198)      Loss 0.3667 (0.3667)    Prec 92.188% (92.188%)
 * Prec 88.960%
best acc: 89.610000
Epoch: [103][0/391]     Time 0.299 (0.299)      Data 0.278 (0.278)      Loss 0.1445 (0.1445)    Prec 95.31
2% (95.312%)
Epoch: [103][100/391]   Time 0.020 (0.023)      Data 0.001 (0.004)      Loss 0.1964 (0.1093)    Prec 92.96
9% (96.326%)
Epoch: [103][200/391]   Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2024 (0.1116)    Prec 93.75
0% (96.269%)
Epoch: [103][300/391]   Time 0.022 (0.021)      Data 0.001 (0.002)      Loss 0.1236 (0.1119)    Prec 96.09
4% (96.226%)
Validation starts
Test: [0/79]    Time 0.202 (0.202)      Loss 0.2578 (0.2578)    Prec 89.844% (89.844%)
 * Prec 89.550%
best acc: 89.610000
Epoch: [104][0/391]     Time 0.336 (0.336)      Data 0.315 (0.315)      Loss 0.0964 (0.0964)    Prec 95.31
2% (95.312%)
Epoch: [104][100/391]   Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.1125 (0.1106)    Prec 96.87
5% (96.218%)
Epoch: [104][200/391]   Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.0720 (0.1114)    Prec 97.65
6% (96.164%)
Epoch: [104][300/391]   Time 0.020 (0.021)      Data 0.002 (0.003)      Loss 0.0868 (0.1137)    Prec 96.87
5% (96.117%)
Validation starts
Test: [0/79]    Time 0.209 (0.209)      Loss 0.3468 (0.3468)    Prec 90.625% (90.625%)
 * Prec 85.310%
best acc: 89.610000
Epoch: [105][0/391]     Time 0.377 (0.377)      Data 0.355 (0.355)      Loss 0.1497 (0.1497)    Prec 93.75
0% (93.750%)
Epoch: [105][100/391]   Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.0810 (0.1040)    Prec 97.65
6% (96.225%)
Epoch: [105][200/391]   Time 0.019 (0.022)      Data 0.001 (0.003)      Loss 0.2182 (0.1020)    Prec 92.18
8% (96.374%)
Epoch: [105][300/391]   Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1195 (0.1048)    Prec 94.53
1% (96.356%)
Validation starts
Test: [0/79]    Time 0.207 (0.207)      Loss 0.2860 (0.2860)    Prec 88.281% (88.281%)
 * Prec 89.790%
best acc: 89.790000
Epoch: [106][0/391]     Time 0.361 (0.361)      Data 0.339 (0.339)      Loss 0.0997 (0.0997)    Prec 97.65
6% (97.656%)
Epoch: [106][100/391]   Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.1389 (0.1048)    Prec 94.53
1% (96.403%)
Epoch: [106][200/391]   Time 0.020 (0.022)      Data 0.001 (0.003)      Loss 0.2191 (0.1056)    Prec 94.53
1% (96.459%)
Epoch: [106][300/391]   Time 0.020 (0.021)      Data 0.001 (0.002)      Loss 0.1999 (0.1117)    Prec 90.62
5% (96.247%)
Validation starts
Test: [0/79]    Time 0.211 (0.211)      Loss 0.4581 (0.4581)    Prec 86.719% (86.719%)
 * Prec 87.790%
best acc: 89.790000
Epoch: [107][0/391]     Time 0.309 (0.309)      Data 0.285 (0.285)      Loss 0.1092 (0.1092)    Prec 94.53
1% (94.531%)
Epoch: [107][100/391]   Time 0.020 (0.023)      Data 0.002 (0.004)      Loss 0.0972 (0.1077)    Prec 96.09
4% (96.334%)
Epoch: [107][200/391]   Time 0.020 (0.022)      Data 0.002 (0.003)      Loss 0.0944 (0.1054)    Prec 98.43
8% (96.393%)
Epoch: [107][300/391]   Time 0.020 (0.021)      Data 0.002 (0.003)      Loss 0.1167 (0.1093)    Prec 96.87
5% (96.325%)
Validation starts
Test: [0/79]    Time 0.205 (0.205)      Loss 0.2870 (0.2870)    Prec 89.062% (89.062%)
 * Prec 88.200%
best acc: 89.790000
Epoch: [108][0/391]     Time 0.353 (0.353)      Data 0.331 (0.331)      Loss 0.0381 (0.0381)    Prec 99.21
9% (99.219%)
Epoch: [108][100/391]   Time 0.020 (0.024)      Data 0.001 (0.005)      Loss 0.0293 (0.1000)    Prec 100.00
0% (96.821%)
Epoch: [108][200/391]   Time 0.021 (0.022)      Data 0.001 (0.003)      Loss 0.1870 (0.1027)    Prec 92.96
9% (96.681%)
```

```
Epoch: [108][300/391]   Time 0.020 (0.025)      Data 0.002 (0.003)      Loss 0.1730 (0.1103)    Prec 92.18
8% (96.333%)
Validation starts
Test: [0/79]    Time 0.196 (0.196)      Loss 0.2934 (0.2934)    Prec 91.406% (91.406%)
 * Prec 89.260%
best acc: 89.790000
Epoch: [109][0/391]     Time 0.344 (0.344)      Data 0.318 (0.318)      Loss 0.0796 (0.0796)    Prec 98.43
8% (98.438%)
Epoch: [109][100/391]   Time 0.044 (0.046)      Data 0.001 (0.004)      Loss 0.0547 (0.1048)    Prec 98.43
8% (96.287%)
Epoch: [109][200/391]   Time 0.043 (0.045)      Data 0.001 (0.003)      Loss 0.1254 (0.1006)    Prec 96.09
4% (96.494%)
Epoch: [109][300/391]   Time 0.043 (0.044)      Data 0.001 (0.002)      Loss 0.1123 (0.1059)    Prec 97.65
6% (96.353%)
Validation starts
Test: [0/79]    Time 0.224 (0.224)      Loss 0.3733 (0.3733)    Prec 88.281% (88.281%)
 * Prec 88.750%
best acc: 89.790000
Epoch: [110][0/391]     Time 0.492 (0.492)      Data 0.466 (0.466)      Loss 0.0380 (0.0380)    Prec 99.21
9% (99.219%)
Epoch: [110][100/391]   Time 0.043 (0.048)      Data 0.002 (0.006)      Loss 0.0436 (0.0865)    Prec 97.65
6% (97.099%)
Epoch: [110][200/391]   Time 0.043 (0.045)      Data 0.002 (0.004)      Loss 0.1825 (0.0985)    Prec 92.96
9% (96.646%)
Epoch: [110][300/391]   Time 0.043 (0.045)      Data 0.002 (0.003)      Loss 0.1089 (0.1037)    Prec 95.31
2% (96.480%)
Validation starts
Test: [0/79]    Time 0.238 (0.238)      Loss 0.2704 (0.2704)    Prec 90.625% (90.625%)
 * Prec 88.160%
best acc: 89.790000
Epoch: [111][0/391]     Time 0.558 (0.558)      Data 0.529 (0.529)      Loss 0.0382 (0.0382)    Prec 99.21
9% (99.219%)
Epoch: [111][100/391]   Time 0.044 (0.048)      Data 0.001 (0.006)      Loss 0.0661 (0.1076)    Prec 96.87
5% (96.334%)
Epoch: [111][200/391]   Time 0.046 (0.046)      Data 0.001 (0.004)      Loss 0.1226 (0.1073)    Prec 96.87
5% (96.311%)
Epoch: [111][300/391]   Time 0.040 (0.045)      Data 0.001 (0.003)      Loss 0.1271 (0.1080)    Prec 95.31
2% (96.307%)
Validation starts
Test: [0/79]    Time 0.424 (0.424)      Loss 0.3419 (0.3419)    Prec 87.500% (87.500%)
 * Prec 87.920%
best acc: 89.790000
Epoch: [112][0/391]     Time 0.332 (0.332)      Data 0.308 (0.308)      Loss 0.1359 (0.1359)    Prec 93.75
0% (93.750%)
Epoch: [112][100/391]   Time 0.044 (0.045)      Data 0.001 (0.004)      Loss 0.0372 (0.0961)    Prec 99.21
9% (96.612%)
Epoch: [112][200/391]   Time 0.044 (0.044)      Data 0.001 (0.003)      Loss 0.1147 (0.1012)    Prec 96.09
4% (96.529%)
Epoch: [112][300/391]   Time 0.042 (0.044)      Data 0.001 (0.002)      Loss 0.1187 (0.1033)    Prec 96.09
4% (96.480%)
Validation starts
Test: [0/79]    Time 0.207 (0.207)      Loss 0.3657 (0.3657)    Prec 88.281% (88.281%)
 * Prec 88.010%
best acc: 89.790000
Epoch: [113][0/391]     Time 0.397 (0.397)      Data 0.375 (0.375)      Loss 0.1344 (0.1344)    Prec 96.09
4% (96.094%)
Epoch: [113][100/391]   Time 0.043 (0.040)      Data 0.001 (0.005)      Loss 0.1052 (0.0985)    Prec 96.09
4% (96.597%)
Epoch: [113][200/391]   Time 0.044 (0.042)      Data 0.001 (0.003)      Loss 0.1254 (0.1013)    Prec 98.43
8% (96.525%)
Epoch: [113][300/391]   Time 0.043 (0.042)      Data 0.002 (0.003)      Loss 0.1576 (0.1070)    Prec 94.53
1% (96.296%)
Validation starts
Test: [0/79]    Time 0.301 (0.301)      Loss 0.3133 (0.3133)    Prec 91.406% (91.406%)
 * Prec 88.650%
best acc: 89.790000
Epoch: [114][0/391]     Time 0.444 (0.444)      Data 0.417 (0.417)      Loss 0.1696 (0.1696)    Prec 93.75
0% (93.750%)
Epoch: [114][100/391]   Time 0.043 (0.039)      Data 0.001 (0.006)      Loss 0.1200 (0.0970)    Prec 96.87
5% (96.627%)
Epoch: [114][200/391]   Time 0.043 (0.041)      Data 0.001 (0.004)      Loss 0.1085 (0.0998)    Prec 96.87
5% (96.646%)
Epoch: [114][300/391]   Time 0.046 (0.042)      Data 0.001 (0.003)      Loss 0.1549 (0.1011)    Prec 95.31
2% (96.571%)
Validation starts
Test: [0/79]    Time 0.306 (0.306)      Loss 0.3212 (0.3212)    Prec 89.062% (89.062%)
 * Prec 87.880%
best acc: 89.790000
Epoch: [115][0/391]     Time 0.399 (0.399)      Data 0.369 (0.369)      Loss 0.0384 (0.0384)    Prec 98.43
8% (98.438%)
Epoch: [115][100/391]   Time 0.044 (0.041)      Data 0.001 (0.007)      Loss 0.0567 (0.0911)    Prec 97.65
6% (96.867%)
```

```
Epoch: [115][200/391]   Time 0.044 (0.042)      Data 0.001 (0.004)      Loss 0.0178 (0.0972)    Prec 100.00
0% (96.638%)
Epoch: [115][300/391]   Time 0.044 (0.043)      Data 0.001 (0.003)      Loss 0.1664 (0.1021)    Prec 94.53
1% (96.475%)
Validation starts
Test: [0/79]    Time 0.300 (0.300)      Loss 0.2697 (0.2697)    Prec 89.844% (89.844%)
 * Prec 88.160%
best acc: 89.790000
Epoch: [116][0/391]     Time 0.355 (0.355)      Data 0.327 (0.327)      Loss 0.1411 (0.1411)    Prec 95.31
2% (95.312%)
Epoch: [116][100/391]   Time 0.044 (0.045)      Data 0.001 (0.008)      Loss 0.0877 (0.0978)    Prec 97.65
6% (96.566%)
Epoch: [116][200/391]   Time 0.043 (0.044)      Data 0.001 (0.004)      Loss 0.1571 (0.0937)    Prec 92.96
9% (96.712%)
Epoch: [116][300/391]   Time 0.043 (0.044)      Data 0.001 (0.003)      Loss 0.0977 (0.0992)    Prec 96.87
5% (96.564%)
Validation starts
Test: [0/79]    Time 0.322 (0.322)      Loss 0.5547 (0.5547)    Prec 84.375% (84.375%)
 * Prec 85.470%
best acc: 89.790000
Epoch: [117][0/391]     Time 0.480 (0.480)      Data 0.453 (0.453)      Loss 0.0330 (0.0330)    Prec 100.00
0% (100.000%)
Epoch: [117][100/391]   Time 0.041 (0.042)      Data 0.001 (0.009)      Loss 0.0987 (0.0986)    Prec 96.87
5% (96.658%)
Epoch: [117][200/391]   Time 0.043 (0.043)      Data 0.001 (0.005)      Loss 0.2590 (0.1064)    Prec 89.84
4% (96.381%)
Epoch: [117][300/391]   Time 0.043 (0.043)      Data 0.001 (0.004)      Loss 0.1254 (0.1059)    Prec 96.09
4% (96.405%)
Validation starts
Test: [0/79]    Time 0.321 (0.321)      Loss 0.5207 (0.5207)    Prec 84.375% (84.375%)
 * Prec 86.750%
best acc: 89.790000
Epoch: [118][0/391]     Time 0.393 (0.393)      Data 0.369 (0.369)      Loss 0.0659 (0.0659)    Prec 98.43
8% (98.438%)
Epoch: [118][100/391]   Time 0.020 (0.041)      Data 0.001 (0.005)      Loss 0.1540 (0.0929)    Prec 94.53
1% (96.976%)
Epoch: [118][200/391]   Time 0.044 (0.042)      Data 0.001 (0.004)      Loss 0.0518 (0.0935)    Prec 99.21
9% (96.980%)
Epoch: [118][300/391]   Time 0.044 (0.043)      Data 0.001 (0.003)      Loss 0.0681 (0.0998)    Prec 99.21
9% (96.709%)
Validation starts
Test: [0/79]    Time 0.287 (0.287)      Loss 0.1926 (0.1926)    Prec 92.969% (92.969%)
 * Prec 86.750%
best acc: 89.790000
Epoch: [119][0/391]     Time 0.373 (0.373)      Data 0.308 (0.308)      Loss 0.0584 (0.0584)    Prec 98.43
8% (98.438%)
Epoch: [119][100/391]   Time 0.020 (0.045)      Data 0.001 (0.005)      Loss 0.1268 (0.1014)    Prec 96.09
4% (96.566%)
Epoch: [119][200/391]   Time 0.043 (0.042)      Data 0.001 (0.003)      Loss 0.0353 (0.0987)    Prec 99.21
9% (96.735%)
Epoch: [119][300/391]   Time 0.044 (0.042)      Data 0.001 (0.003)      Loss 0.0600 (0.0969)    Prec 98.43
8% (96.758%)
Validation starts
Test: [0/79]    Time 0.299 (0.299)      Loss 0.3658 (0.3658)    Prec 87.500% (87.500%)
 * Prec 88.380%
best acc: 89.790000
Epoch: [120][0/391]     Time 0.411 (0.411)      Data 0.338 (0.338)      Loss 0.0766 (0.0766)    Prec 97.65
6% (97.656%)
Epoch: [120][100/391]   Time 0.041 (0.047)      Data 0.001 (0.006)      Loss 0.1740 (0.1045)    Prec 93.75
0% (96.419%)
Epoch: [120][200/391]   Time 0.043 (0.043)      Data 0.002 (0.004)      Loss 0.0619 (0.1018)    Prec 97.65
6% (96.529%)
Epoch: [120][300/391]   Time 0.043 (0.043)      Data 0.002 (0.003)      Loss 0.1125 (0.1035)    Prec 94.53
1% (96.527%)
Validation starts
Test: [0/79]    Time 0.211 (0.211)      Loss 0.2833 (0.2833)    Prec 92.969% (92.969%)
 * Prec 88.820%
best acc: 89.790000
Epoch: [121][0/391]     Time 0.317 (0.317)      Data 0.286 (0.286)      Loss 0.0663 (0.0663)    Prec 96.87
5% (96.875%)
Epoch: [121][100/391]   Time 0.037 (0.046)      Data 0.001 (0.004)      Loss 0.1108 (0.0944)    Prec 96.09
4% (96.674%)
Epoch: [121][200/391]   Time 0.044 (0.043)      Data 0.002 (0.003)      Loss 0.0286 (0.1003)    Prec 99.21
9% (96.587%)
Epoch: [121][300/391]   Time 0.043 (0.043)      Data 0.001 (0.003)      Loss 0.0777 (0.1028)    Prec 97.65
6% (96.504%)
Validation starts
Test: [0/79]    Time 0.224 (0.224)      Loss 0.3429 (0.3429)    Prec 90.625% (90.625%)
 * Prec 87.810%
best acc: 89.790000
Epoch: [122][0/391]     Time 0.426 (0.426)      Data 0.351 (0.351)      Loss 0.1146 (0.1146)    Prec 96.87
5% (96.875%)
```

```
Epoch: [122][100/391]   Time 0.043 (0.048)      Data 0.002 (0.006)      Loss 0.0801 (0.0869)    Prec 97.65
6% (97.053%)
Epoch: [122][200/391]   Time 0.043 (0.044)      Data 0.002 (0.004)      Loss 0.0910 (0.0980)    Prec 95.31
2% (96.661%)
Epoch: [122][300/391]   Time 0.046 (0.044)      Data 0.001 (0.003)      Loss 0.1298 (0.1019)    Prec 95.31
2% (96.548%)
Validation starts
Test: [0/79]    Time 0.295 (0.295)      Loss 0.3525 (0.3525)    Prec 90.625% (90.625%)
 * Prec 87.720%
best acc: 89.790000
Epoch: [123][0/391]     Time 0.349 (0.349)      Data 0.317 (0.317)      Loss 0.0447 (0.0447)    Prec 98.43
8% (98.438%)
Epoch: [123][100/391]   Time 0.037 (0.047)      Data 0.001 (0.005)      Loss 0.0991 (0.0969)    Prec 96.87
5% (96.604%)
Epoch: [123][200/391]   Time 0.043 (0.042)      Data 0.002 (0.005)      Loss 0.1537 (0.0995)    Prec 95.31
2% (96.564%)
Epoch: [123][300/391]   Time 0.043 (0.042)      Data 0.002 (0.004)      Loss 0.1108 (0.0994)    Prec 95.31
2% (96.551%)
Validation starts
Test: [0/79]    Time 0.281 (0.281)      Loss 0.3759 (0.3759)    Prec 89.844% (89.844%)
 * Prec 87.790%
best acc: 89.790000
Epoch: [124][0/391]     Time 0.374 (0.374)      Data 0.348 (0.348)      Loss 0.0653 (0.0653)    Prec 98.43
8% (98.438%)
Epoch: [124][100/391]   Time 0.044 (0.047)      Data 0.001 (0.005)      Loss 0.0959 (0.1090)    Prec 96.09
4% (96.349%)
Epoch: [124][200/391]   Time 0.044 (0.043)      Data 0.001 (0.003)      Loss 0.0554 (0.1007)    Prec 98.43
8% (96.560%)
Epoch: [124][300/391]   Time 0.043 (0.043)      Data 0.001 (0.002)      Loss 0.1262 (0.0988)    Prec 96.09
4% (96.639%)
Validation starts
Test: [0/79]    Time 0.219 (0.219)      Loss 0.2874 (0.2874)    Prec 88.281% (88.281%)
 * Prec 88.260%
best acc: 89.790000
Epoch: [125][0/391]     Time 0.383 (0.383)      Data 0.355 (0.355)      Loss 0.0947 (0.0947)    Prec 96.09
4% (96.094%)
Epoch: [125][100/391]   Time 0.043 (0.046)      Data 0.002 (0.005)      Loss 0.0596 (0.0986)    Prec 97.65
6% (96.558%)
Epoch: [125][200/391]   Time 0.043 (0.043)      Data 0.001 (0.004)      Loss 0.0883 (0.1029)    Prec 97.65
6% (96.455%)
Epoch: [125][300/391]   Time 0.043 (0.043)      Data 0.001 (0.003)      Loss 0.0964 (0.1001)    Prec 98.43
8% (96.556%)
Validation starts
Test: [0/79]    Time 0.311 (0.311)      Loss 0.3019 (0.3019)    Prec 91.406% (91.406%)
 * Prec 88.210%
best acc: 89.790000
Epoch: [126][0/391]     Time 0.341 (0.341)      Data 0.316 (0.316)      Loss 0.0729 (0.0729)    Prec 97.65
6% (97.656%)
Epoch: [126][100/391]   Time 0.043 (0.046)      Data 0.001 (0.004)      Loss 0.0387 (0.0893)    Prec 99.21
9% (96.968%)
Epoch: [126][200/391]   Time 0.042 (0.043)      Data 0.002 (0.003)      Loss 0.1057 (0.0935)    Prec 98.43
8% (96.867%)
Epoch: [126][300/391]   Time 0.043 (0.043)      Data 0.001 (0.003)      Loss 0.1099 (0.0968)    Prec 95.31
2% (96.701%)
Validation starts
Test: [0/79]    Time 0.302 (0.302)      Loss 0.3150 (0.3150)    Prec 89.062% (89.062%)
 * Prec 88.530%
best acc: 89.790000
Epoch: [127][0/391]     Time 0.415 (0.415)      Data 0.386 (0.386)      Loss 0.0783 (0.0783)    Prec 96.87
5% (96.875%)
Epoch: [127][100/391]   Time 0.043 (0.047)      Data 0.001 (0.005)      Loss 0.1606 (0.0809)    Prec 92.96
9% (97.192%)
Epoch: [127][200/391]   Time 0.043 (0.042)      Data 0.001 (0.004)      Loss 0.0621 (0.0915)    Prec 99.21
9% (96.856%)
Epoch: [127][300/391]   Time 0.043 (0.043)      Data 0.001 (0.003)      Loss 0.0696 (0.0940)    Prec 97.65
6% (96.795%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.2411 (0.2411)    Prec 92.969% (92.969%)
 * Prec 89.060%
best acc: 89.790000
Epoch: [128][0/391]     Time 0.330 (0.330)      Data 0.301 (0.301)      Loss 0.1418 (0.1418)    Prec 93.75
0% (93.750%)
Epoch: [128][100/391]   Time 0.043 (0.046)      Data 0.001 (0.005)      Loss 0.1499 (0.0861)    Prec 92.96
9% (97.208%)
Epoch: [128][200/391]   Time 0.043 (0.043)      Data 0.002 (0.004)      Loss 0.1667 (0.0955)    Prec 93.75
0% (96.813%)
Epoch: [128][300/391]   Time 0.045 (0.043)      Data 0.001 (0.003)      Loss 0.1323 (0.0975)    Prec 93.75
0% (96.722%)
Validation starts
Test: [0/79]    Time 0.201 (0.201)      Loss 0.3068 (0.3068)    Prec 89.062% (89.062%)
 * Prec 89.230%
best acc: 89.790000
```

```
Epoch: [129][0/391]     Time 0.367 (0.367)      Data 0.340 (0.340)      Loss 0.1697 (0.1697)    Prec 93.75
0% (93.750%)
Epoch: [129][100/391]   Time 0.043 (0.046)      Data 0.002 (0.005)      Loss 0.0842 (0.0903)    Prec 96.87
5% (96.921%)
Epoch: [129][200/391]   Time 0.044 (0.043)      Data 0.001 (0.004)      Loss 0.1182 (0.0957)    Prec 96.87
5% (96.696%)
Epoch: [129][300/391]   Time 0.043 (0.043)      Data 0.001 (0.003)      Loss 0.1237 (0.0969)    Prec 96.87
5% (96.613%)
Validation starts
Test: [0/79]    Time 0.200 (0.200)      Loss 0.1888 (0.1888)    Prec 92.188% (92.188%)
 * Prec 89.950%
best acc: 89.950000
Epoch: [130][0/391]     Time 0.340 (0.340)      Data 0.309 (0.309)      Loss 0.0876 (0.0876)    Prec 96.87
5% (96.875%)
Epoch: [130][100/391]   Time 0.043 (0.046)      Data 0.002 (0.005)      Loss 0.0644 (0.0926)    Prec 97.65
6% (96.674%)
Epoch: [130][200/391]   Time 0.018 (0.041)      Data 0.001 (0.003)      Loss 0.1866 (0.0929)    Prec 93.75
0% (96.739%)
Epoch: [130][300/391]   Time 0.043 (0.042)      Data 0.002 (0.003)      Loss 0.1128 (0.0949)    Prec 96.87
5% (96.706%)
Validation starts
Test: [0/79]    Time 0.208 (0.208)      Loss 0.4803 (0.4803)    Prec 86.719% (86.719%)
 * Prec 86.770%
best acc: 89.950000
Epoch: [131][0/391]     Time 0.402 (0.402)      Data 0.335 (0.335)      Loss 0.0596 (0.0596)    Prec 99.21
9% (99.219%)
Epoch: [131][100/391]   Time 0.043 (0.047)      Data 0.001 (0.005)      Loss 0.1014 (0.0857)    Prec 95.31
2% (97.138%)
Epoch: [131][200/391]   Time 0.018 (0.044)      Data 0.001 (0.003)      Loss 0.0706 (0.0923)    Prec 96.87
5% (96.891%)
Epoch: [131][300/391]   Time 0.043 (0.044)      Data 0.001 (0.003)      Loss 0.0545 (0.0979)    Prec 99.21
9% (96.649%)
Validation starts
Test: [0/79]    Time 0.296 (0.296)      Loss 0.2797 (0.2797)    Prec 90.625% (90.625%)
 * Prec 88.290%
best acc: 89.950000
Epoch: [132][0/391]     Time 0.333 (0.333)      Data 0.309 (0.309)      Loss 0.1077 (0.1077)    Prec 98.43
8% (98.438%)
Epoch: [132][100/391]   Time 0.037 (0.046)      Data 0.002 (0.005)      Loss 0.0594 (0.0970)    Prec 98.43
8% (96.720%)
Epoch: [132][200/391]   Time 0.021 (0.045)      Data 0.001 (0.003)      Loss 0.1344 (0.0988)    Prec 96.87
5% (96.603%)
Epoch: [132][300/391]   Time 0.043 (0.044)      Data 0.001 (0.002)      Loss 0.0713 (0.1002)    Prec 98.43
8% (96.558%)
Validation starts
Test: [0/79]    Time 0.304 (0.304)      Loss 0.2875 (0.2875)    Prec 91.406% (91.406%)
 * Prec 87.910%
best acc: 89.950000
Epoch: [133][0/391]     Time 0.330 (0.330)      Data 0.302 (0.302)      Loss 0.0610 (0.0610)    Prec 97.65
6% (97.656%)
Epoch: [133][100/391]   Time 0.044 (0.046)      Data 0.001 (0.005)      Loss 0.0588 (0.0889)    Prec 99.21
9% (97.130%)
Epoch: [133][200/391]   Time 0.021 (0.044)      Data 0.001 (0.003)      Loss 0.1125 (0.0908)    Prec 96.09
4% (97.023%)
Epoch: [133][300/391]   Time 0.043 (0.044)      Data 0.001 (0.003)      Loss 0.0847 (0.0936)    Prec 96.09
4% (96.859%)
Validation starts
Test: [0/79]    Time 0.226 (0.226)      Loss 0.2765 (0.2765)    Prec 91.406% (91.406%)
 * Prec 89.080%
best acc: 89.950000
Epoch: [134][0/391]     Time 0.406 (0.406)      Data 0.379 (0.379)      Loss 0.2234 (0.2234)    Prec 93.75
0% (93.750%)
Epoch: [134][100/391]   Time 0.045 (0.046)      Data 0.002 (0.005)      Loss 0.0569 (0.0904)    Prec 99.21
9% (97.061%)
Epoch: [134][200/391]   Time 0.040 (0.044)      Data 0.002 (0.003)      Loss 0.0681 (0.0924)    Prec 96.87
5% (96.914%)
Epoch: [134][300/391]   Time 0.043 (0.043)      Data 0.001 (0.003)      Loss 0.1943 (0.0978)    Prec 93.75
0% (96.727%)
Validation starts
Test: [0/79]    Time 0.288 (0.288)      Loss 0.2277 (0.2277)    Prec 92.969% (92.969%)
 * Prec 89.700%
best acc: 89.950000
Epoch: [135][0/391]     Time 0.421 (0.421)      Data 0.389 (0.389)      Loss 0.0600 (0.0600)    Prec 98.43
8% (98.438%)
Epoch: [135][100/391]   Time 0.044 (0.046)      Data 0.001 (0.005)      Loss 0.1197 (0.0881)    Prec 96.09
4% (96.976%)
Epoch: [135][200/391]   Time 0.037 (0.044)      Data 0.001 (0.003)      Loss 0.2732 (0.0917)    Prec 91.40
6% (96.926%)
Epoch: [135][300/391]   Time 0.043 (0.042)      Data 0.002 (0.003)      Loss 0.1160 (0.0980)    Prec 95.31
2% (96.696%)
Validation starts
Test: [0/79]    Time 0.222 (0.222)      Loss 0.1211 (0.1211)    Prec 96.875% (96.875%)
```

```
 * Prec 89.020%
best acc: 89.950000
Epoch: [136][0/391]     Time 0.416 (0.416)      Data 0.387 (0.387)      Loss 0.0531 (0.0531)    Prec 98.43
8% (98.438%)
Epoch: [136][100/391]   Time 0.043 (0.047)      Data 0.002 (0.005)      Loss 0.0853 (0.0949)    Prec 96.09
4% (96.689%)
Epoch: [136][200/391]   Time 0.098 (0.046)      Data 0.002 (0.004)      Loss 0.0436 (0.0924)    Prec 96.87
5% (96.813%)
Epoch: [136][300/391]   Time 0.043 (0.043)      Data 0.001 (0.003)      Loss 0.0886 (0.0955)    Prec 97.65
6% (96.737%)
Validation starts
Test: [0/79]    Time 0.214 (0.214)      Loss 0.2024 (0.2024)    Prec 95.312% (95.312%)
 * Prec 88.810%
best acc: 89.950000
Epoch: [137][0/391]     Time 0.396 (0.396)      Data 0.371 (0.371)      Loss 0.1429 (0.1429)    Prec 94.53
1% (94.531%)
Epoch: [137][100/391]   Time 0.041 (0.047)      Data 0.002 (0.005)      Loss 0.0919 (0.0931)    Prec 96.87
5% (96.612%)
Epoch: [137][200/391]   Time 0.044 (0.045)      Data 0.001 (0.003)      Loss 0.1068 (0.0950)    Prec 98.43
8% (96.685%)
Epoch: [137][300/391]   Time 0.043 (0.043)      Data 0.002 (0.003)      Loss 0.0458 (0.0964)    Prec 96.87
5% (96.737%)
Validation starts
Test: [0/79]    Time 0.215 (0.215)      Loss 0.2970 (0.2970)    Prec 92.188% (92.188%)
 * Prec 88.390%
best acc: 89.950000
Epoch: [138][0/391]     Time 0.429 (0.429)      Data 0.402 (0.402)      Loss 0.0991 (0.0991)    Prec 96.87
5% (96.875%)
Epoch: [138][100/391]   Time 0.043 (0.047)      Data 0.001 (0.005)      Loss 0.0995 (0.0969)    Prec 98.43
8% (96.481%)
Epoch: [138][200/391]   Time 0.044 (0.045)      Data 0.001 (0.003)      Loss 0.1203 (0.0973)    Prec 97.65
6% (96.615%)
Epoch: [138][300/391]   Time 0.043 (0.043)      Data 0.002 (0.003)      Loss 0.1376 (0.0991)    Prec 95.31
2% (96.561%)
Validation starts
Test: [0/79]    Time 0.312 (0.312)      Loss 0.1748 (0.1748)    Prec 96.094% (96.094%)
 * Prec 88.700%
best acc: 89.950000
Epoch: [139][0/391]     Time 0.339 (0.339)      Data 0.312 (0.312)      Loss 0.0838 (0.0838)    Prec 96.87
5% (96.875%)
Epoch: [139][100/391]   Time 0.043 (0.046)      Data 0.002 (0.005)      Loss 0.0574 (0.0836)    Prec 98.43
8% (97.355%)
Epoch: [139][200/391]   Time 0.043 (0.044)      Data 0.002 (0.003)      Loss 0.1125 (0.0895)    Prec 94.53
1% (97.100%)
Epoch: [139][300/391]   Time 0.153 (0.043)      Data 0.001 (0.003)      Loss 0.1112 (0.0945)    Prec 97.65
6% (96.870%)
Validation starts
Test: [0/79]    Time 0.213 (0.213)      Loss 0.1879 (0.1879)    Prec 95.312% (95.312%)
 * Prec 90.190%
best acc: 90.190000
Epoch: [140][0/391]     Time 0.367 (0.367)      Data 0.306 (0.306)      Loss 0.0599 (0.0599)    Prec 98.43
8% (98.438%)
Epoch: [140][100/391]   Time 0.043 (0.046)      Data 0.002 (0.005)      Loss 0.0838 (0.0744)    Prec 98.43
8% (97.602%)
Epoch: [140][200/391]   Time 0.043 (0.045)      Data 0.002 (0.003)      Loss 0.1659 (0.0867)    Prec 96.87
5% (97.023%)
Epoch: [140][300/391]   Time 0.046 (0.043)      Data 0.001 (0.003)      Loss 0.0499 (0.0929)    Prec 97.65
6% (96.831%)
Validation starts
Test: [0/79]    Time 0.211 (0.211)      Loss 0.2185 (0.2185)    Prec 92.969% (92.969%)
 * Prec 89.510%
best acc: 90.190000
Epoch: [141][0/391]     Time 0.394 (0.394)      Data 0.365 (0.365)      Loss 0.0658 (0.0658)    Prec 99.21
9% (99.219%)
Epoch: [141][100/391]   Time 0.044 (0.047)      Data 0.001 (0.005)      Loss 0.1315 (0.0873)    Prec 95.31
2% (96.991%)
Epoch: [141][200/391]   Time 0.044 (0.045)      Data 0.001 (0.003)      Loss 0.0538 (0.0918)    Prec 98.43
8% (96.809%)
Epoch: [141][300/391]   Time 0.020 (0.043)      Data 0.001 (0.003)      Loss 0.1841 (0.0965)    Prec 95.31
2% (96.717%)
Validation starts
Test: [0/79]    Time 0.207 (0.207)      Loss 0.2640 (0.2640)    Prec 92.188% (92.188%)
 * Prec 88.960%
best acc: 90.190000
Epoch: [142][0/391]     Time 0.441 (0.441)      Data 0.414 (0.414)      Loss 0.1330 (0.1330)    Prec 96.09
4% (96.094%)
Epoch: [142][100/391]   Time 0.043 (0.047)      Data 0.002 (0.006)      Loss 0.0493 (0.0876)    Prec 98.43
8% (97.123%)
Epoch: [142][200/391]   Time 0.043 (0.045)      Data 0.002 (0.004)      Loss 0.0599 (0.0875)    Prec 98.43
8% (97.023%)
Epoch: [142][300/391]   Time 0.021 (0.044)      Data 0.002 (0.003)      Loss 0.1544 (0.0906)    Prec 96.09
4% (96.935%)
```

```
Validation starts
Test: [0/79]    Time 0.314 (0.314)      Loss 0.3384 (0.3384)    Prec 90.625% (90.625%)
 * Prec 88.440%
best acc: 90.190000
Epoch: [143][0/391]     Time 0.391 (0.391)      Data 0.361 (0.361)      Loss 0.1284 (0.1284)    Prec 95.31
2% (95.312%)
Epoch: [143][100/391]   Time 0.043 (0.047)      Data 0.002 (0.005)      Loss 0.1297 (0.0908)    Prec 94.53
1% (96.883%)
Epoch: [143][200/391]   Time 0.045 (0.045)      Data 0.002 (0.003)      Loss 0.0368 (0.0940)    Prec 98.43
8% (96.859%)
Epoch: [143][300/391]   Time 0.019 (0.044)      Data 0.002 (0.003)      Loss 0.0341 (0.0931)    Prec 98.43
8% (96.839%)
Validation starts
Test: [0/79]    Time 0.295 (0.295)      Loss 0.3183 (0.3183)    Prec 89.844% (89.844%)
 * Prec 89.360%
best acc: 90.190000
Epoch: [144][0/391]     Time 0.326 (0.326)      Data 0.299 (0.299)      Loss 0.0713 (0.0713)    Prec 96.87
5% (96.875%)
Epoch: [144][100/391]   Time 0.043 (0.046)      Data 0.001 (0.004)      Loss 0.0971 (0.0894)    Prec 96.09
4% (96.898%)
Epoch: [144][200/391]   Time 0.044 (0.045)      Data 0.001 (0.003)      Loss 0.0452 (0.0959)    Prec 97.65
6% (96.727%)
Epoch: [144][300/391]   Time 0.019 (0.044)      Data 0.001 (0.002)      Loss 0.0326 (0.0963)    Prec 99.21
9% (96.675%)
Validation starts
Test: [0/79]    Time 0.210 (0.210)      Loss 0.2809 (0.2809)    Prec 91.406% (91.406%)
 * Prec 89.050%
best acc: 90.190000
Epoch: [145][0/391]     Time 0.441 (0.441)      Data 0.413 (0.413)      Loss 0.0677 (0.0677)    Prec 99.21
9% (99.219%)
Epoch: [145][100/391]   Time 0.041 (0.047)      Data 0.001 (0.006)      Loss 0.0374 (0.0881)    Prec 99.21
9% (96.999%)
Epoch: [145][200/391]   Time 0.046 (0.045)      Data 0.001 (0.003)      Loss 0.1081 (0.0903)    Prec 95.31
2% (96.875%)
Epoch: [145][300/391]   Time 0.024 (0.044)      Data 0.001 (0.003)      Loss 0.0284 (0.0939)    Prec 99.21
9% (96.795%)
Validation starts
Test: [0/79]    Time 0.307 (0.307)      Loss 0.2079 (0.2079)    Prec 95.312% (95.312%)
 * Prec 89.630%
best acc: 90.190000
Epoch: [146][0/391]     Time 0.409 (0.409)      Data 0.384 (0.384)      Loss 0.0962 (0.0962)    Prec 96.09
4% (96.094%)
Epoch: [146][100/391]   Time 0.043 (0.047)      Data 0.001 (0.005)      Loss 0.0891 (0.0905)    Prec 96.09
4% (96.921%)
Epoch: [146][200/391]   Time 0.044 (0.045)      Data 0.001 (0.003)      Loss 0.1063 (0.0917)    Prec 96.87
5% (96.840%)
Epoch: [146][300/391]   Time 0.037 (0.045)      Data 0.001 (0.003)      Loss 0.1336 (0.0933)    Prec 96.09
4% (96.789%)
Validation starts
Test: [0/79]    Time 0.292 (0.292)      Loss 0.2147 (0.2147)    Prec 92.969% (92.969%)
 * Prec 88.670%
best acc: 90.190000
Epoch: [147][0/391]     Time 0.413 (0.413)      Data 0.384 (0.384)      Loss 0.1132 (0.1132)    Prec 96.87
5% (96.875%)
Epoch: [147][100/391]   Time 0.044 (0.047)      Data 0.001 (0.005)      Loss 0.0812 (0.0922)    Prec 96.09
4% (96.829%)
Epoch: [147][200/391]   Time 0.043 (0.045)      Data 0.001 (0.003)      Loss 0.0908 (0.0940)    Prec 96.87
5% (96.778%)
Epoch: [147][300/391]   Time 0.046 (0.045)      Data 0.001 (0.003)      Loss 0.1021 (0.0968)    Prec 97.65
6% (96.678%)
Validation starts
Test: [0/79]    Time 0.309 (0.309)      Loss 0.1869 (0.1869)    Prec 93.750% (93.750%)
 * Prec 88.360%
best acc: 90.190000
Epoch: [148][0/391]     Time 0.315 (0.315)      Data 0.285 (0.285)      Loss 0.1244 (0.1244)    Prec 96.09
4% (96.094%)
Epoch: [148][100/391]   Time 0.044 (0.046)      Data 0.001 (0.004)      Loss 0.1064 (0.0855)    Prec 96.09
4% (97.161%)
Epoch: [148][200/391]   Time 0.044 (0.044)      Data 0.001 (0.003)      Loss 0.0682 (0.0869)    Prec 98.43
8% (97.139%)
Epoch: [148][300/391]   Time 0.041 (0.044)      Data 0.001 (0.002)      Loss 0.0596 (0.0907)    Prec 98.43
8% (96.976%)
Validation starts
Test: [0/79]    Time 0.201 (0.201)      Loss 0.2947 (0.2947)    Prec 91.406% (91.406%)
 * Prec 88.930%
best acc: 90.190000
Epoch: [149][0/391]     Time 0.386 (0.386)      Data 0.305 (0.305)      Loss 0.1028 (0.1028)    Prec 96.09
4% (96.094%)
Epoch: [149][100/391]   Time 0.043 (0.047)      Data 0.002 (0.004)      Loss 0.0734 (0.0884)    Prec 97.65
6% (97.107%)
Epoch: [149][200/391]   Time 0.043 (0.045)      Data 0.001 (0.003)      Loss 0.1176 (0.0851)    Prec 96.09
4% (97.120%)
```

```
Epoch: [149][300/391]   Time 0.043 (0.044)      Data 0.002 (0.003)      Loss 0.1029 (0.0889)    Prec 95.31
2% (96.935%)
Validation starts
Test: [0/79]    Time 0.206 (0.206)      Loss 0.3260 (0.3260)    Prec 90.625% (90.625%)
 * Prec 88.730%
best acc: 90.190000
```

In [3]:
```
# HW

#  1. Train with 4 bits for both weight and activation to achieve >90% accuracy
#  2. Find x_int and w_int for the 2nd convolution layer
#  3. Check the recovered psum has similar value to the un-quantized original psum
#     (such as example 1 in W3S2)
```

In [4]:
```
PATH = "result/final_VGG16_quant/model_best.pth.tar"
checkpoint = torch.load(PATH)
model.load_state_dict(checkpoint['state_dict'])
device = torch.device("cuda")

model.cuda()
model.eval()

test_loss = 0
correct = 0

with torch.no_grad():
    for data, target in testloader:
        data, target = data.to(device), target.to(device) # loading to GPU
        output = model(data)
        pred = output.argmax(dim=1, keepdim=True)
        correct += pred.eq(target.view_as(pred)).sum().item()

test_loss /= len(testloader.dataset)

print('\nTest set: Accuracy: {}/{} ({:.0f}%)\n'.format(
        correct, len(testloader.dataset),
        100. * correct / len(testloader.dataset)))
```

```
Test set: Accuracy: 9019/10000 (90%)
```

In [5]:
```
#send an input and grap the value by using prehook like HW3
```

In [6]:
```
class SaveOutput:
    def __init__(self):
        self.outputs = []
    def __call__(self, module, module_in):
        self.outputs.append(module_in)
    def clear(self):
        self.outputs = []

######### Save inputs from selected layer ##########
save_output = SaveOutput()

for layer in model.modules():
    if isinstance(layer, torch.nn.Conv2d):
        print("prehooked")
        layer.register_forward_pre_hook(save_output)       ## Input for the module will be grapped
####################################################

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

dataiter = iter(trainloader)
images, labels = next(dataiter)
images = images.to(device)
out = model(images)

print("1st convolution's input size:", save_output.outputs[0][0].size())
print("2nd convolution's input size:", save_output.outputs[1][0].size())
```

```
prehooked
prehooked
prehooked
prehooked
prehooked
prehooked
prehooked
prehooked
prehooked
prehooked
prehooked
prehooked
prehooked
1st convolution's input size: torch.Size([128, 3, 32, 32])
2nd convolution's input size: torch.Size([128, 64, 32, 32])
```

In [7]:
```
for index in range(len(save_output.outputs)):
    print(str(index) + ": " + str(save_output.outputs[index][0].size()))
```

```
0: torch.Size([128, 3, 32, 32])
1: torch.Size([128, 64, 32, 32])
2: torch.Size([128, 64, 16, 16])
3: torch.Size([128, 128, 16, 16])
4: torch.Size([128, 128, 8, 8])
5: torch.Size([128, 256, 8, 8])
6: torch.Size([128, 256, 8, 8])
7: torch.Size([128, 256, 4, 4])
8: torch.Size([128, 8, 4, 4])
9: torch.Size([128, 8, 4, 4])
10: torch.Size([128, 512, 2, 2])
11: torch.Size([128, 512, 2, 2])
12: torch.Size([128, 512, 2, 2])
```

In [8]:
```
w_bit = 4
weight_q = model.features[27].weight_q # quantized value is stored during the training
w_alpha = model.features[27].weight_quant.wgt_alpha  # alpha is defined in your model already. bring it out
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
```

In [9]:
```
x_bit = 4
x = save_output.outputs[8][0]  # input of the 2nd conv layer
x_alpha  = model.features[27].act_alpha
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
```

In [11]:
```
conv_int = torch.nn.Conv2d(in_channels = 8, out_channels=8, kernel_size = 3, padding=1, bias = False)
conv_int.weight = torch.nn.parameter.Parameter(weight_int)

output_int =  conv_int.forward(x_int)    # output_int can be calculated with conv_int and x_int
output_recovered = output_int * x_delta * w_delta  # recover with x_delta and w_delta

relu = torch.nn.ReLU()
relu_output_recovered = relu(output_recovered)
#print(output_recovered)
```

In [12]:
```
next_input = save_output.outputs[9][0]

print(next_input.size())
print(relu_output_recovered.size())
```

```
torch.Size([128, 8, 4, 4])
torch.Size([128, 8, 4, 4])
```

In [13]:
```
difference = abs( next_input - relu_output_recovered )
print(difference.mean())  ## It should be small, e.g.,2.3 in my trainned model
```

```
tensor(3.1711e-08, device='cuda:0', grad_fn=<MeanBackward0>)
```

In [14]:
```
print(x_int[0,:,:,:].size())
```

```
torch.Size([8, 4, 4])
```

In [15]:
```
print(x_int.size())
```

```
torch.Size([128, 8, 4, 4])
```

In [16]:
```
# act_int.size = torch.Size([128, 64, 32, 32])  <- batch_size, input_ch, ni, nj
```

```
a_int = x_int[0,:,:,:]  # pick only one input out of batch
# a_int.size() = [64, 32, 32]

# conv_int.weight.size() = torch.Size([64, 64, 3, 3])  <- output_ch, input_ch, ki, kj
w_int = torch.reshape(weight_int, (weight_int.size(0), weight_int.size(1), -1))  # merge ki, kj index to ki
# w_int.weight.size() = torch.Size([64, 64, 9])

padding = 1
stride = 1
array_size = 8 # row and column number

nig = range(a_int.size(1))  ## ni group
njg = range(a_int.size(2))  ## nj group

icg = range(int(w_int.size(1)))  ## input channel
ocg = range(int(w_int.size(0)))  ## output channel

ic_tileg = range(int(len(icg)/array_size))
oc_tileg = range(int(len(ocg)/array_size))

kijg = range(w_int.size(2))
ki_dim = int(math.sqrt(w_int.size(2)))  ## Kernel's 1 dim size

######## Padding before Convolution #######
a_pad = torch.zeros(len(icg), len(nig)+padding*2, len(nig)+padding*2).cuda()
# a_pad.size() = [64, 32+2pad, 32+2pad]
a_pad[ :, padding:padding+len(nig), padding:padding+len(njg)] = a_int.cuda()
a_pad = torch.reshape(a_pad, (a_pad.size(0), -1))
# a_pad.size() = [64, (32+2pad)*(32+2pad)]


a_tile = torch.zeros(len(ic_tileg), array_size,    a_pad.size(1)).cuda()
w_tile = torch.zeros(len(oc_tileg)*len(ic_tileg), array_size, array_size, len(kijg)).cuda()

for ic_tile in ic_tileg:
    a_tile[ic_tile,:,:] = a_pad[ic_tile*array_size:(ic_tile+1)*array_size,:]

for ic_tile in ic_tileg:
    for oc_tile in oc_tileg:
        w_tile[oc_tile*len(oc_tileg) + ic_tile,:,:,:] = w_int[oc_tile*array_size:(oc_tile+1)*array_size, ic



##########################################

p_nijg = range(a_pad.size(1)) ## psum nij group

psum = torch.zeros(len(ic_tileg), len(oc_tileg), array_size, len(p_nijg), len(kijg)).cuda()

for kij in kijg:
    for ic_tile in ic_tileg:       # Tiling into array_sizeXarray_size array
        for oc_tile in oc_tileg:   # Tiling into array_sizeXarray_size array
            for nij in p_nijg:       # time domain, sequentially given input
                    m = nn.Linear(array_size, array_size, bias=False)
                    #m.weight = torch.nn.Parameter(w_int[oc_tile*array_size:(oc_tile+1)*array_size, ic_tile
                    m.weight = torch.nn.Parameter(w_tile[len(oc_tileg)*oc_tile+ic_tile,:,:,kij])
                    psum[ic_tile, oc_tile, :, nij, kij] = m(a_tile[ic_tile,:,nij]).cuda()
```

In [17]:
```
print(a_tile.size())
print(psum.size())
```

```
torch.Size([1, 8, 36])
torch.Size([1, 1, 8, 36, 9])
```

In [18]:
```
import math

a_pad_ni_dim = int(math.sqrt(a_pad.size(1))) # 32

o_ni_dim = int((a_pad_ni_dim - (ki_dim- 1) - 1)/stride + 1)
o_nijg = range(o_ni_dim**2)

print(len(o_nijg))

out = torch.zeros(len(ocg), len(o_nijg)).cuda()


### SFP accumulation ###
for o_nij in o_nijg:
    for kij in kijg:
        for ic_tile in ic_tileg:
            for oc_tile in oc_tileg:
```

```
                out[oc_tile*array_size:(oc_tile+1)*array_size, o_nij] = out[oc_tile*array_size:(oc_tile+1)*
                psum[ic_tile, oc_tile, :, int(o_nij/o_ni_dim)*a_pad_ni_dim + o_nij%o_ni_dim + int(kij/ki_di
                ## 4th index = (int(o_nij/30)*32 + o_nij%30) + (int(kij/3)*32 + kij%3)
```

16

In [19]:
```
print(o_ni_dim)
print(a_pad_ni_dim)
print(ki_dim)
```

```
4
6
3
```

In [20]:
```
temp_acc = 0
for row in range(8):
    for col in range(1):
        print(a_tile[0,row,7])
        print(w_tile[0,col,row,0])
        temp_acc = temp_acc + a_tile[0,row,7] * w_tile[0, col, row, 0]
print(temp_acc)
```

```
tensor(0., device='cuda:0', grad_fn=<SelectBackward0>)
tensor(-7.0000, device='cuda:0', grad_fn=<SelectBackward0>)
tensor(0., device='cuda:0', grad_fn=<SelectBackward0>)
tensor(-7.0000, device='cuda:0', grad_fn=<SelectBackward0>)
tensor(0., device='cuda:0', grad_fn=<SelectBackward0>)
tensor(-7.0000, device='cuda:0', grad_fn=<SelectBackward0>)
tensor(2., device='cuda:0', grad_fn=<SelectBackward0>)
tensor(-7.0000, device='cuda:0', grad_fn=<SelectBackward0>)
tensor(6., device='cuda:0', grad_fn=<SelectBackward0>)
tensor(-7.0000, device='cuda:0', grad_fn=<SelectBackward0>)
tensor(0., device='cuda:0', grad_fn=<SelectBackward0>)
tensor(7.0000, device='cuda:0', grad_fn=<SelectBackward0>)
tensor(11.0000, device='cuda:0', grad_fn=<SelectBackward0>)
tensor(7.0000, device='cuda:0', grad_fn=<SelectBackward0>)
tensor(0., device='cuda:0', grad_fn=<SelectBackward0>)
tensor(-7.0000, device='cuda:0', grad_fn=<SelectBackward0>)
tensor(21.0000, device='cuda:0', grad_fn=<AddBackward0>)
```

In [21]:
```
### show this cell partially. The following cells should be printed by students ###
tile_id = 0
nij = 200 # just a random number
X = a_tile[tile_id,:,:]  # [tile_num, array row num, time_steps]

bit_precision = 4
file = open('activation.txt', 'w') #write to file
file.write('#time0row7[msb-lsb],time0row6[msb-lst],....,time0row0[msb-lst]#\n')
file.write('#time1row7[msb-lsb],time1row6[msb-lst],....,time1row0[msb-lst]#\n')
file.write('#................#\n')

for i in range(X.size(1)):  # time step
    for j in range(X.size(0)): # row #
        X_bin = '{0:04b}'.format(round(X[X.size(0)-1-j,i].item()))
        for k in range(bit_precision):
            file.write(X_bin[k])
        #file.write(' ')  # for visibility with blank between words, you can use
    file.write('\n')
file.close() #close file
```

In [22]:
```
### Complete this cell ###
tile_id = 0
kij = 0


bit_precision = 4
for kij in range(9):
    W = w_tile[tile_id,:,:,kij]  # w_tile[tile_num, array col num, array row num, kij]
    file = open('weight_' + str(kij) + '.txt', 'w') #write to file
    file.write('#col0row7[msb-lsb],col0row6[msb-lst],....,col0row0[msb-lst]#\n')
    file.write('#col1row7[msb-lsb],col1row6[msb-lst],....,col1row0[msb-lst]#\n')
    file.write('#................#\n')

    for i in range(W.size(1)):  # col
        for j in range(W.size(0)): # row #
            W_bin = '{0:04b}'.format(round(W[i,7-j].item()) + (16 if (round(W[i,7-j].item()) < 0) else 0))
            for k in range(bit_precision):
                file.write(W_bin[k])
            #file.write(' ')  # for visibility with blank between words, you can use
        file.write('\n')
    file.close() #close file
```

In [23]:
```
### Complete this cell ###
ic_tile_id = 0
oc_tile_id = 0


kij = 0
nij = 200
# psum[len(ic_tileg), len(oc_tileg), array_size, len(p_nijg), len(kijg)]


bit_precision = 16

for kij in range(9):
    psum_tile = psum[ic_tile_id,oc_tile_id,:,:,kij]
    file = open('psum_' + str(kij) + '.txt', 'w') #write to file
    file.write('#time0col7[msb-lsb],time0col6[msb-lst],....,time0col0[msb-lst]#\n')
    file.write('#time1col7[msb-lsb],time1col6[msb-lst],....,time1col0[msb-lst]#\n')
    file.write('#................#\n')

    for i in range(psum_tile.size(1)):  # time step
        for j in range(psum_tile.size(0)): # col #
            #psum_bin = '{0:016b}'.format(round(psum_tile[psum_tile.size(0)-1-j,i].item()) + (2**bit_precis
            curr_psum = round(psum_tile[psum_tile.size(0)-1-j,i].item())
            if (i == 7 and kij == 0):
                print(curr_psum)
            if (curr_psum < 0):
                if (i == 7 and kij == 0):
                    print('{0:016b}'.format(curr_psum))
                curr_psum = curr_psum + 2**bit_precision
                if (i == 7 and kij == 0):
                    print('{0:016b}'.format(curr_psum))
            psum_bin = '{0:016b}'.format(curr_psum)

            for k in range(bit_precision):
                file.write(psum_bin[k])
            #file.write(' ')  # for visibility with blank between words, you can use
        file.write('\n')
    file.close() #close file
```

```
-21
-000000000010101
1111111111101011
21
-49
-000000000110001
1111111111001111
-105
-000000001101001
1111111110010111
-21
-000000000010101
1111111111101011
105
-105
-000000001101001
1111111110010111
21
```

In [24]:
```
### Complete this cell ###


bit_precision = 16
# out is array of size columns x len(o_nijg)

file = open('out.txt', 'w') #write to file
file.write('#time0col7[msb-lsb],time0col6[msb-lst],....,time0col0[msb-lst]#\n')
file.write('#time1col7[msb-lsb],time1col6[msb-lst],....,time1col0[msb-lst]#\n')
file.write('#................#\n')

for i in range(out.size(1)):  # time step
    for j in range(out.size(0)): # row #
        out_bin = '{0:016b}'.format(0 if (round(out[out.size(0)-1-j,i].item()) < 0) else (round(out[out.siz
        for k in range(bit_precision):
            file.write(out_bin[k])
        #file.write(' ')  # for visibility with blank between words, you can use
    file.write('\n')
file.close() #close file
```

In [25]:
```
print(out[0,:])
```

```
tensor([ 273.0000,  154.0000,  182.0000,  147.0000,  308.0000,  363.9999,
         154.0000,  252.0000,   49.0000,  133.0000,   77.0000,  -63.0000,
         -84.0000, -329.0000, -238.0000, -210.0000], device='cuda:0',
       grad_fn=<SliceBackward0>)
```

In [26]:
```
print(out[0,:])
```

```
tensor([ 273.0000,  154.0000,  182.0000,  147.0000,  308.0000,  363.9999,
         154.0000,  252.0000,   49.0000,  133.0000,   77.0000,  -63.0000,
         -84.0000, -329.0000, -238.0000, -210.0000], device='cuda:0',
       grad_fn=<SliceBackward0>)
```

In [27]:
```
print(X.size())
```

```
torch.Size([8, 36])
```

In [28]:
```
print(o_ni_dim)
print(a_pad_ni_dim)
```

```
4
6
```

In [29]:
```
print(weight_int.size())
```

```
torch.Size([8, 8, 3, 3])
```

In [ ]:

In [ ]:

In [ ]: