[HW4_prob2_1]_VGG16_Loss

October 30, 2022

[1]:
```
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

from models import *


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



batch_size = 128
model_name = "VGG16_1"
model = VGG16()

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 => every 5 batch, prints out
                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


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'))


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(
  (features): Sequential(
    (0): Conv2d(3, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (2): ReLU(inplace=True)
    (3): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (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): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (8): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True,
```

```
track_running_stats=True)
    (9): ReLU(inplace=True)
    (10): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (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): Conv2d(128, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (15): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (16): ReLU(inplace=True)
    (17): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (18): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (19): ReLU(inplace=True)
    (20): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (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): Conv2d(256, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (25): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (26): ReLU(inplace=True)
    (27): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (28): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (29): ReLU(inplace=True)
    (30): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (31): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (32): ReLU(inplace=True)
    (33): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
    (34): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (35): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (36): ReLU(inplace=True)
```

```
    (37): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (38): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (39): ReLU(inplace=True)
    (40): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1),
bias=False)
    (41): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True,
track_running_stats=True)
    (42): ReLU(inplace=True)
    (43): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
    (44): 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
```

[2]:
```
# This cell won't be given, but students will complete the training

lr = 4.4e-2
weight_decay = 1.0e-4
epochs = 80
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)


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)
```

```
/opt/conda/lib/python3.9/site-packages/torch/nn/functional.py:718: UserWarning:
Named tensors and all their associated APIs are an experimental feature and
subject to change. Please do not use them for anything important until they are
released as stable. (Triggered internally at
/pytorch/c10/core/TensorImpl.h:1156.)
  return torch.max_pool2d(input, kernel_size, stride, padding, dilation,
ceil_mode)
Epoch: [0][0/391]       Time 0.771 (0.771)      Data 0.677 (0.677)      Loss
2.5002 (2.5002)    Prec 9.375% (9.375%)
Epoch: [0][100/391]     Time 0.044 (0.051)      Data 0.002 (0.008)      Loss
2.2862 (2.8068)    Prec 15.625% (11.170%)
Epoch: [0][200/391]     Time 0.044 (0.048)      Data 0.002 (0.005)      Loss
2.2531 (2.5453)    Prec 12.500% (11.909%)
Epoch: [0][300/391]     Time 0.094 (0.057)      Data 0.002 (0.004)      Loss
2.2484 (2.4498)    Prec 8.594% (12.246%)
Validation starts
Test: [0/79]    Time 0.458 (0.458)      Loss 2.2971 (2.2971)    Prec 19.531%
(19.531%)
 * Prec 17.030%
best acc: 17.030000
Epoch: [1][0/391]       Time 0.852 (0.852)      Data 0.812 (0.812)      Loss
2.2252 (2.2252)    Prec 14.844% (14.844%)
Epoch: [1][100/391]     Time 0.094 (0.101)      Data 0.002 (0.010)      Loss
1.9321 (2.0466)    Prec 23.438% (20.158%)
Epoch: [1][200/391]     Time 0.061 (0.097)      Data 0.002 (0.006)      Loss
1.8629 (2.0084)    Prec 21.094% (20.064%)
Epoch: [1][300/391]     Time 0.094 (0.094)      Data 0.001 (0.005)      Loss
1.8712 (1.9743)    Prec 21.875% (20.767%)
Validation starts
Test: [0/79]    Time 0.993 (0.993)      Loss 1.9068 (1.9068)    Prec 25.000%
(25.000%)
 * Prec 24.500%
best acc: 24.500000
Epoch: [2][0/391]       Time 0.693 (0.693)      Data 0.650 (0.650)      Loss
```

```
1.8748 (1.8748)    Prec 23.438% (23.438%)
Epoch: [2][100/391]     Time 0.094 (0.100)      Data 0.002 (0.008)      Loss
1.8215 (1.8681)    Prec 21.094% (25.131%)
Epoch: [2][200/391]     Time 0.066 (0.097)      Data 0.002 (0.005)      Loss
1.8717 (1.8582)    Prec 16.406% (25.323%)
Epoch: [2][300/391]     Time 0.093 (0.096)      Data 0.002 (0.005)      Loss
1.8278 (1.8483)    Prec 25.781% (25.548%)
Validation starts
Test: [0/79]    Time 0.471 (0.471)      Loss 1.9274 (1.9274)    Prec 28.125%
(28.125%)
 * Prec 29.200%
best acc: 29.200000
Epoch: [3][0/391]       Time 0.768 (0.768)      Data 0.727 (0.727)      Loss
1.7947 (1.7947)    Prec 26.562% (26.562%)
Epoch: [3][100/391]     Time 0.094 (0.101)      Data 0.001 (0.009)      Loss
1.6133 (1.7753)    Prec 34.375% (30.221%)
Epoch: [3][200/391]     Time 0.044 (0.097)      Data 0.002 (0.005)      Loss
1.7007 (1.7480)    Prec 37.500% (31.592%)
Epoch: [3][300/391]     Time 0.094 (0.094)      Data 0.002 (0.004)      Loss
1.5147 (1.7195)    Prec 42.188% (32.654%)
Validation starts
Test: [0/79]    Time 0.556 (0.556)      Loss 1.8219 (1.8219)    Prec 31.250%
(31.250%)
 * Prec 34.940%
best acc: 34.940000
Epoch: [4][0/391]       Time 0.956 (0.956)      Data 0.911 (0.911)      Loss
1.5012 (1.5012)    Prec 42.188% (42.188%)
Epoch: [4][100/391]     Time 0.094 (0.103)      Data 0.001 (0.011)      Loss
1.6523 (1.5781)    Prec 33.594% (39.488%)
Epoch: [4][200/391]     Time 0.203 (0.100)      Data 0.005 (0.006)      Loss
1.5947 (1.5481)    Prec 38.281% (40.322%)
Epoch: [4][300/391]     Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
1.4125 (1.5176)    Prec 48.438% (41.580%)
Validation starts
Test: [0/79]    Time 0.797 (0.797)      Loss 1.3442 (1.3442)    Prec 54.688%
(54.688%)
 * Prec 49.320%
best acc: 49.320000
Epoch: [5][0/391]       Time 0.741 (0.741)      Data 0.699 (0.699)      Loss
1.2436 (1.2436)    Prec 56.250% (56.250%)
Epoch: [5][100/391]     Time 0.094 (0.101)      Data 0.001 (0.008)      Loss
1.1835 (1.3409)    Prec 50.000% (49.706%)
Epoch: [5][200/391]     Time 0.044 (0.097)      Data 0.002 (0.005)      Loss
1.3148 (1.2985)    Prec 49.219% (51.543%)
Epoch: [5][300/391]     Time 0.094 (0.094)      Data 0.001 (0.004)      Loss
1.0009 (1.2636)    Prec 66.406% (53.333%)
Validation starts
Test: [0/79]    Time 0.452 (0.452)      Loss 1.1711 (1.1711)    Prec 54.688%
```

```
(54.688%)
 * Prec 55.900%
best acc: 55.900000
Epoch: [6][0/391]       Time 1.057 (1.057)      Data 1.012 (1.012)      Loss
1.1657 (1.1657)    Prec 60.938% (60.938%)
Epoch: [6][100/391]     Time 0.095 (0.104)      Data 0.001 (0.012)      Loss
1.0198 (1.0688)    Prec 60.938% (61.564%)
Epoch: [6][200/391]     Time 0.193 (0.099)      Data 0.001 (0.007)      Loss
0.8349 (1.0546)    Prec 70.312% (62.030%)
Epoch: [6][300/391]     Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
1.0403 (1.0303)    Prec 66.406% (63.094%)
Validation starts
Test: [0/79]    Time 0.622 (0.622)      Loss 0.9697 (0.9697)    Prec 65.625%
(65.625%)
 * Prec 63.860%
best acc: 63.860000
Epoch: [7][0/391]       Time 1.022 (1.022)      Data 0.983 (0.983)      Loss
0.9353 (0.9353)    Prec 66.406% (66.406%)
Epoch: [7][100/391]     Time 0.097 (0.103)      Data 0.002 (0.011)      Loss
0.9957 (0.9054)    Prec 67.188% (68.000%)
Epoch: [7][200/391]     Time 0.088 (0.099)      Data 0.001 (0.006)      Loss
0.6970 (0.8840)    Prec 74.219% (68.703%)
Epoch: [7][300/391]     Time 0.095 (0.096)      Data 0.003 (0.005)      Loss
0.7720 (0.8747)    Prec 76.562% (69.132%)
Validation starts
Test: [0/79]    Time 0.387 (0.387)      Loss 1.0452 (1.0452)    Prec 57.812%
(57.812%)
 * Prec 65.010%
best acc: 65.010000
Epoch: [8][0/391]       Time 1.086 (1.086)      Data 1.036 (1.036)      Loss
0.9526 (0.9526)    Prec 65.625% (65.625%)
Epoch: [8][100/391]     Time 0.094 (0.104)      Data 0.002 (0.012)      Loss
0.6766 (0.7731)    Prec 75.781% (73.105%)
Epoch: [8][200/391]     Time 0.093 (0.099)      Data 0.002 (0.007)      Loss
0.6302 (0.7662)    Prec 76.562% (73.278%)
Epoch: [8][300/391]     Time 0.095 (0.096)      Data 0.001 (0.006)      Loss
0.6925 (0.7555)    Prec 71.875% (73.640%)
Validation starts
Test: [0/79]    Time 0.646 (0.646)      Loss 0.7521 (0.7521)    Prec 71.094%
(71.094%)
 * Prec 72.800%
best acc: 72.800000
Epoch: [9][0/391]       Time 0.670 (0.670)      Data 0.616 (0.616)      Loss
0.6084 (0.6084)    Prec 79.688% (79.688%)
Epoch: [9][100/391]     Time 0.094 (0.100)      Data 0.001 (0.007)      Loss
0.8063 (0.6784)    Prec 72.656% (76.957%)
Epoch: [9][200/391]     Time 0.090 (0.097)      Data 0.002 (0.005)      Loss
0.6651 (0.6797)    Prec 78.125% (76.710%)
```

```
Epoch: [9][300/391]     Time 0.096 (0.094)      Data 0.001 (0.004)      Loss
0.6531 (0.6784)    Prec 76.562% (76.799%)
Validation starts
Test: [0/79]    Time 0.642 (0.642)      Loss 0.7252 (0.7252)    Prec 77.344%
(77.344%)
 * Prec 75.600%
best acc: 75.600000
Epoch: [10][0/391]      Time 0.965 (0.965)      Data 0.923 (0.923)      Loss
0.7227 (0.7227)    Prec 71.875% (71.875%)
Epoch: [10][100/391]    Time 0.094 (0.103)      Data 0.002 (0.011)      Loss
0.4168 (0.6112)    Prec 86.719% (79.270%)
Epoch: [10][200/391]    Time 0.092 (0.099)      Data 0.002 (0.007)      Loss
0.7274 (0.6088)    Prec 75.781% (79.454%)
Epoch: [10][300/391]    Time 0.095 (0.096)      Data 0.002 (0.005)      Loss
0.6083 (0.6035)    Prec 77.344% (79.529%)
Validation starts
Test: [0/79]    Time 0.672 (0.672)      Loss 0.4928 (0.4928)    Prec 84.375%
(84.375%)
 * Prec 78.240%
best acc: 78.240000
Epoch: [11][0/391]      Time 0.651 (0.651)      Data 0.611 (0.611)      Loss
0.5884 (0.5884)    Prec 80.469% (80.469%)
Epoch: [11][100/391]    Time 0.095 (0.100)      Data 0.002 (0.008)      Loss
0.5358 (0.5518)    Prec 83.594% (81.173%)
Epoch: [11][200/391]    Time 0.091 (0.098)      Data 0.002 (0.006)      Loss
0.5326 (0.5525)    Prec 75.000% (81.040%)
Epoch: [11][300/391]    Time 0.095 (0.096)      Data 0.002 (0.004)      Loss
0.4956 (0.5470)    Prec 78.125% (81.260%)
Validation starts
Test: [0/79]    Time 0.826 (0.826)      Loss 0.4245 (0.4245)    Prec 85.938%
(85.938%)
 * Prec 81.740%
best acc: 81.740000
Epoch: [12][0/391]      Time 1.107 (1.107)      Data 1.063 (1.063)      Loss
0.5173 (0.5173)    Prec 83.594% (83.594%)
Epoch: [12][100/391]    Time 0.094 (0.104)      Data 0.001 (0.012)      Loss
0.5682 (0.5102)    Prec 78.906% (82.867%)
Epoch: [12][200/391]    Time 0.091 (0.099)      Data 0.001 (0.007)      Loss
0.5058 (0.5102)    Prec 83.594% (82.809%)
Epoch: [12][300/391]    Time 0.094 (0.095)      Data 0.001 (0.006)      Loss
0.5054 (0.5116)    Prec 81.250% (82.807%)
Validation starts
Test: [0/79]    Time 0.456 (0.456)      Loss 0.5500 (0.5500)    Prec 80.469%
(80.469%)
 * Prec 78.170%
best acc: 81.740000
Epoch: [13][0/391]      Time 0.775 (0.775)      Data 0.733 (0.733)      Loss
0.6353 (0.6353)    Prec 79.688% (79.688%)
```

```
Epoch: [13][100/391]    Time 0.095 (0.101)      Data 0.001 (0.009)      Loss
0.6341 (0.4738)    Prec 83.594% (84.104%)
Epoch: [13][200/391]    Time 0.094 (0.097)      Data 0.001 (0.006)      Loss
0.5444 (0.4616)    Prec 78.906% (84.527%)
Epoch: [13][300/391]    Time 0.094 (0.095)      Data 0.001 (0.004)      Loss
0.7323 (0.4679)    Prec 74.219% (84.147%)
Validation starts
Test: [0/79]    Time 0.493 (0.493)      Loss 0.4554 (0.4554)    Prec 85.156%
(85.156%)
 * Prec 82.480%
best acc: 82.480000
Epoch: [14][0/391]      Time 0.714 (0.714)      Data 0.669 (0.669)      Loss
0.3938 (0.3938)    Prec 88.281% (88.281%)
Epoch: [14][100/391]    Time 0.094 (0.100)      Data 0.002 (0.008)      Loss
0.4001 (0.4204)    Prec 88.281% (85.752%)
Epoch: [14][200/391]    Time 0.093 (0.098)      Data 0.001 (0.005)      Loss
0.4333 (0.4309)    Prec 84.375% (85.487%)
Epoch: [14][300/391]    Time 0.095 (0.095)      Data 0.001 (0.005)      Loss
0.5235 (0.4352)    Prec 83.594% (85.421%)
Validation starts
Test: [0/79]    Time 0.642 (0.642)      Loss 0.4689 (0.4689)    Prec 83.594%
(83.594%)
 * Prec 80.930%
best acc: 82.480000
Epoch: [15][0/391]      Time 0.772 (0.772)      Data 0.731 (0.731)      Loss
0.4740 (0.4740)    Prec 83.594% (83.594%)
Epoch: [15][100/391]    Time 0.094 (0.101)      Data 0.002 (0.009)      Loss
0.5188 (0.4012)    Prec 83.594% (86.371%)
Epoch: [15][200/391]    Time 0.091 (0.098)      Data 0.002 (0.006)      Loss
0.4334 (0.3982)    Prec 86.719% (86.493%)
Epoch: [15][300/391]    Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
0.3428 (0.4028)    Prec 89.062% (86.270%)
Validation starts
Test: [0/79]    Time 0.701 (0.701)      Loss 0.3792 (0.3792)    Prec 87.500%
(87.500%)
 * Prec 85.810%
best acc: 85.810000
Epoch: [16][0/391]      Time 1.042 (1.042)      Data 1.002 (1.002)      Loss
0.4960 (0.4960)    Prec 85.156% (85.156%)
Epoch: [16][100/391]    Time 0.093 (0.104)      Data 0.002 (0.011)      Loss
0.4147 (0.3741)    Prec 86.719% (87.338%)
Epoch: [16][200/391]    Time 0.088 (0.099)      Data 0.002 (0.007)      Loss
0.3314 (0.3841)    Prec 88.281% (86.940%)
Epoch: [16][300/391]    Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
0.4159 (0.3803)    Prec 85.156% (87.051%)
Validation starts
Test: [0/79]    Time 0.529 (0.529)      Loss 0.4409 (0.4409)    Prec 86.719%
(86.719%)
```

```
 * Prec 80.770%
best acc: 85.810000
Epoch: [17][0/391]      Time 0.746 (0.746)      Data 0.706 (0.706)      Loss
0.3660 (0.3660)    Prec 86.719% (86.719%)
Epoch: [17][100/391]    Time 0.095 (0.101)      Data 0.001 (0.009)      Loss
0.2817 (0.3556)    Prec 86.719% (87.987%)
Epoch: [17][200/391]    Time 0.092 (0.097)      Data 0.002 (0.006)      Loss
0.3741 (0.3558)    Prec 84.375% (87.970%)
Epoch: [17][300/391]    Time 0.094 (0.095)      Data 0.002 (0.004)      Loss
0.3383 (0.3594)    Prec 90.625% (87.822%)
Validation starts
Test: [0/79]    Time 0.517 (0.517)      Loss 0.6008 (0.6008)    Prec 79.688%
(79.688%)
 * Prec 82.630%
best acc: 85.810000
Epoch: [18][0/391]      Time 0.814 (0.814)      Data 0.766 (0.766)      Loss
0.3056 (0.3056)    Prec 86.719% (86.719%)
Epoch: [18][100/391]    Time 0.094 (0.101)      Data 0.001 (0.009)      Loss
0.2693 (0.3380)    Prec 89.062% (88.660%)
Epoch: [18][200/391]    Time 0.092 (0.098)      Data 0.002 (0.006)      Loss
0.2693 (0.3379)    Prec 91.406% (88.604%)
Epoch: [18][300/391]    Time 0.094 (0.094)      Data 0.001 (0.004)      Loss
0.2443 (0.3373)    Prec 92.188% (88.611%)
Validation starts
Test: [0/79]    Time 0.488 (0.488)      Loss 0.4600 (0.4600)    Prec 85.938%
(85.938%)
 * Prec 84.200%
best acc: 85.810000
Epoch: [19][0/391]      Time 0.764 (0.764)      Data 0.723 (0.723)      Loss
0.2582 (0.2582)    Prec 92.969% (92.969%)
Epoch: [19][100/391]    Time 0.095 (0.101)      Data 0.001 (0.009)      Loss
0.2132 (0.2944)    Prec 91.406% (90.029%)
Epoch: [19][200/391]    Time 0.083 (0.097)      Data 0.002 (0.005)      Loss
0.3116 (0.3087)    Prec 89.844% (89.463%)
Epoch: [19][300/391]    Time 0.095 (0.094)      Data 0.001 (0.004)      Loss
0.2322 (0.3163)    Prec 92.188% (89.226%)
Validation starts
Test: [0/79]    Time 0.594 (0.594)      Loss 0.4575 (0.4575)    Prec 83.594%
(83.594%)
 * Prec 85.120%
best acc: 85.810000
Epoch: [20][0/391]      Time 0.764 (0.764)      Data 0.714 (0.714)      Loss
0.3101 (0.3101)    Prec 90.625% (90.625%)
Epoch: [20][100/391]    Time 0.094 (0.101)      Data 0.001 (0.008)      Loss
0.2744 (0.2976)    Prec 89.844% (89.991%)
Epoch: [20][200/391]    Time 0.088 (0.097)      Data 0.001 (0.006)      Loss
0.2396 (0.2991)    Prec 93.750% (89.879%)
Epoch: [20][300/391]    Time 0.096 (0.095)      Data 0.001 (0.005)      Loss
```

```
0.3982 (0.2991)    Prec 86.719% (89.828%)
Validation starts
Test: [0/79]    Time 0.478 (0.478)      Loss 0.3966 (0.3966)    Prec 88.281%
(88.281%)
 * Prec 84.380%
best acc: 85.810000
Epoch: [21][0/391]      Time 0.656 (0.656)      Data 0.612 (0.612)      Loss
0.2384 (0.2384)    Prec 90.625% (90.625%)
Epoch: [21][100/391]    Time 0.094 (0.100)      Data 0.002 (0.008)      Loss
0.2608 (0.2783)    Prec 93.750% (90.517%)
Epoch: [21][200/391]    Time 0.088 (0.098)      Data 0.001 (0.005)      Loss
0.3552 (0.2850)    Prec 89.844% (90.431%)
Epoch: [21][300/391]    Time 0.095 (0.095)      Data 0.001 (0.004)      Loss
0.2360 (0.2841)    Prec 92.969% (90.454%)
Validation starts
Test: [0/79]    Time 0.989 (0.989)      Loss 0.2768 (0.2768)    Prec 92.969%
(92.969%)
 * Prec 86.920%
best acc: 86.920000
Epoch: [22][0/391]      Time 0.776 (0.776)      Data 0.732 (0.732)      Loss
0.2091 (0.2091)    Prec 92.188% (92.188%)
Epoch: [22][100/391]    Time 0.094 (0.101)      Data 0.002 (0.009)      Loss
0.2479 (0.2574)    Prec 91.406% (91.252%)
Epoch: [22][200/391]    Time 0.091 (0.098)      Data 0.001 (0.005)      Loss
0.2320 (0.2619)    Prec 94.531% (91.142%)
Epoch: [22][300/391]    Time 0.095 (0.096)      Data 0.001 (0.005)      Loss
0.2696 (0.2641)    Prec 92.188% (91.035%)
Validation starts
Test: [0/79]    Time 0.496 (0.496)      Loss 0.4001 (0.4001)    Prec 86.719%
(86.719%)
 * Prec 85.570%
best acc: 86.920000
Epoch: [23][0/391]      Time 0.729 (0.729)      Data 0.672 (0.672)      Loss
0.2507 (0.2507)    Prec 91.406% (91.406%)
Epoch: [23][100/391]    Time 0.094 (0.101)      Data 0.002 (0.008)      Loss
0.1179 (0.2541)    Prec 96.094% (91.151%)
Epoch: [23][200/391]    Time 0.081 (0.098)      Data 0.001 (0.005)      Loss
0.1138 (0.2555)    Prec 96.094% (91.325%)
Epoch: [23][300/391]    Time 0.097 (0.095)      Data 0.002 (0.004)      Loss
0.2421 (0.2610)    Prec 90.625% (91.084%)
Validation starts
Test: [0/79]    Time 0.442 (0.442)      Loss 0.3750 (0.3750)    Prec 87.500%
(87.500%)
 * Prec 86.550%
best acc: 86.920000
Epoch: [24][0/391]      Time 0.718 (0.718)      Data 0.676 (0.676)      Loss
0.1489 (0.1489)    Prec 96.875% (96.875%)
Epoch: [24][100/391]    Time 0.094 (0.100)      Data 0.001 (0.008)      Loss
```

```
0.2853 (0.2253)    Prec 90.625% (92.311%)
Epoch: [24][200/391]    Time 0.093 (0.097)      Data 0.002 (0.005)      Loss
0.2824 (0.2362)    Prec 92.188% (91.966%)
Epoch: [24][300/391]    Time 0.094 (0.095)      Data 0.003 (0.004)      Loss
0.3216 (0.2400)    Prec 92.188% (91.746%)
Validation starts
Test: [0/79]    Time 0.466 (0.466)      Loss 0.2620 (0.2620)    Prec 92.188%
(92.188%)
 * Prec 85.280%
best acc: 86.920000
Epoch: [25][0/391]      Time 0.787 (0.787)      Data 0.673 (0.673)      Loss
0.2451 (0.2451)    Prec 92.188% (92.188%)
Epoch: [25][100/391]    Time 0.095 (0.101)      Data 0.002 (0.009)      Loss
0.2161 (0.2336)    Prec 93.750% (91.762%)
Epoch: [25][200/391]    Time 0.085 (0.097)      Data 0.002 (0.005)      Loss
0.2060 (0.2376)    Prec 93.750% (91.698%)
Epoch: [25][300/391]    Time 0.094 (0.095)      Data 0.002 (0.005)      Loss
0.4143 (0.2400)    Prec 85.938% (91.624%)
Validation starts
Test: [0/79]    Time 0.496 (0.496)      Loss 0.3185 (0.3185)    Prec 91.406%
(91.406%)
 * Prec 87.540%
best acc: 87.540000
Epoch: [26][0/391]      Time 0.784 (0.784)      Data 0.742 (0.742)      Loss
0.1025 (0.1025)    Prec 97.656% (97.656%)
Epoch: [26][100/391]    Time 0.094 (0.101)      Data 0.002 (0.009)      Loss
0.3475 (0.2146)    Prec 87.500% (92.752%)
Epoch: [26][200/391]    Time 0.091 (0.097)      Data 0.001 (0.005)      Loss
0.2800 (0.2192)    Prec 89.062% (92.522%)
Epoch: [26][300/391]    Time 0.095 (0.094)      Data 0.002 (0.004)      Loss
0.1463 (0.2220)    Prec 94.531% (92.489%)
Validation starts
Test: [0/79]    Time 0.654 (0.654)      Loss 0.3578 (0.3578)    Prec 89.062%
(89.062%)
 * Prec 86.670%
best acc: 87.540000
Epoch: [27][0/391]      Time 0.642 (0.642)      Data 0.602 (0.602)      Loss
0.1251 (0.1251)    Prec 96.094% (96.094%)
Epoch: [27][100/391]    Time 0.094 (0.100)      Data 0.001 (0.007)      Loss
0.1763 (0.2084)    Prec 94.531% (92.868%)
Epoch: [27][200/391]    Time 0.092 (0.096)      Data 0.002 (0.004)      Loss
0.1486 (0.2111)    Prec 94.531% (92.825%)
Epoch: [27][300/391]    Time 0.094 (0.094)      Data 0.001 (0.004)      Loss
0.1742 (0.2156)    Prec 94.531% (92.650%)
Validation starts
Test: [0/79]    Time 0.531 (0.531)      Loss 0.2687 (0.2687)    Prec 93.750%
(93.750%)
 * Prec 86.990%
```

```
best acc: 87.540000
Epoch: [28][0/391]      Time 0.710 (0.710)      Data 0.669 (0.669)      Loss
0.1635 (0.1635)    Prec 94.531% (94.531%)
Epoch: [28][100/391]    Time 0.096 (0.101)      Data 0.001 (0.008)      Loss
0.1244 (0.1913)    Prec 96.875% (93.510%)
Epoch: [28][200/391]    Time 0.091 (0.097)      Data 0.002 (0.005)      Loss
0.1759 (0.2016)    Prec 92.188% (93.179%)
Epoch: [28][300/391]    Time 0.094 (0.095)      Data 0.001 (0.004)      Loss
0.1956 (0.2062)    Prec 93.750% (93.013%)
Validation starts
Test: [0/79]    Time 0.623 (0.623)      Loss 0.4007 (0.4007)    Prec 87.500%
(87.500%)
 * Prec 87.660%
best acc: 87.660000
Epoch: [29][0/391]      Time 1.106 (1.106)      Data 1.064 (1.064)      Loss
0.1767 (0.1767)    Prec 91.406% (91.406%)
Epoch: [29][100/391]    Time 0.095 (0.104)      Data 0.001 (0.012)      Loss
0.2422 (0.1875)    Prec 92.969% (93.611%)
Epoch: [29][200/391]    Time 0.092 (0.099)      Data 0.002 (0.007)      Loss
0.1752 (0.1950)    Prec 95.312% (93.377%)
Epoch: [29][300/391]    Time 0.094 (0.096)      Data 0.001 (0.006)      Loss
0.1671 (0.1920)    Prec 94.531% (93.553%)
Validation starts
Test: [0/79]    Time 0.630 (0.630)      Loss 0.2707 (0.2707)    Prec 90.625%
(90.625%)
 * Prec 88.760%
best acc: 88.760000
Epoch: [30][0/391]      Time 0.686 (0.686)      Data 0.643 (0.643)      Loss
0.1352 (0.1352)    Prec 94.531% (94.531%)
Epoch: [30][100/391]    Time 0.093 (0.100)      Data 0.002 (0.008)      Loss
0.1488 (0.1845)    Prec 92.969% (93.673%)
Epoch: [30][200/391]    Time 0.090 (0.097)      Data 0.001 (0.005)      Loss
0.1917 (0.1890)    Prec 93.750% (93.505%)
Epoch: [30][300/391]    Time 0.094 (0.095)      Data 0.002 (0.004)      Loss
0.1915 (0.1920)    Prec 92.969% (93.358%)
Validation starts
Test: [0/79]    Time 0.501 (0.501)      Loss 0.3876 (0.3876)    Prec 87.500%
(87.500%)
 * Prec 87.230%
best acc: 88.760000
Epoch: [31][0/391]      Time 0.734 (0.734)      Data 0.694 (0.694)      Loss
0.1716 (0.1716)    Prec 92.969% (92.969%)
Epoch: [31][100/391]    Time 0.094 (0.101)      Data 0.002 (0.008)      Loss
0.1751 (0.1689)    Prec 96.094% (94.121%)
Epoch: [31][200/391]    Time 0.086 (0.097)      Data 0.001 (0.005)      Loss
0.2834 (0.1730)    Prec 90.625% (94.030%)
Epoch: [31][300/391]    Time 0.095 (0.095)      Data 0.002 (0.004)      Loss
0.1157 (0.1764)    Prec 96.094% (93.869%)
```

```
Validation starts
Test: [0/79]    Time 0.559 (0.559)      Loss 0.3797 (0.3797)    Prec 88.281%
(88.281%)
 * Prec 87.100%
best acc: 88.760000
Epoch: [32][0/391]      Time 0.595 (0.595)      Data 0.552 (0.552)      Loss
0.1144 (0.1144)    Prec 95.312% (95.312%)
Epoch: [32][100/391]    Time 0.094 (0.099)      Data 0.002 (0.007)      Loss
0.2038 (0.1664)    Prec 92.969% (94.106%)
Epoch: [32][200/391]    Time 0.093 (0.097)      Data 0.002 (0.004)      Loss
0.2425 (0.1708)    Prec 92.969% (94.045%)
Epoch: [32][300/391]    Time 0.094 (0.094)      Data 0.002 (0.004)      Loss
0.2165 (0.1710)    Prec 92.188% (94.030%)
Validation starts
Test: [0/79]    Time 0.592 (0.592)      Loss 0.3371 (0.3371)    Prec 89.062%
(89.062%)
 * Prec 88.320%
best acc: 88.760000
Epoch: [33][0/391]      Time 1.148 (1.148)      Data 1.047 (1.047)      Loss
0.1018 (0.1018)    Prec 97.656% (97.656%)
Epoch: [33][100/391]    Time 0.095 (0.106)      Data 0.001 (0.013)      Loss
0.1081 (0.1555)    Prec 96.875% (94.694%)
Epoch: [33][200/391]    Time 0.090 (0.100)      Data 0.002 (0.008)      Loss
0.1243 (0.1622)    Prec 96.094% (94.450%)
Epoch: [33][300/391]    Time 0.097 (0.097)      Data 0.002 (0.006)      Loss
0.1435 (0.1699)    Prec 92.969% (94.132%)
Validation starts
Test: [0/79]    Time 0.715 (0.715)      Loss 0.3039 (0.3039)    Prec 89.062%
(89.062%)
 * Prec 88.350%
best acc: 88.760000
Epoch: [34][0/391]      Time 0.691 (0.691)      Data 0.650 (0.650)      Loss
0.1324 (0.1324)    Prec 96.875% (96.875%)
Epoch: [34][100/391]    Time 0.094 (0.100)      Data 0.001 (0.008)      Loss
0.0556 (0.1527)    Prec 97.656% (94.949%)
Epoch: [34][200/391]    Time 0.090 (0.099)      Data 0.001 (0.006)      Loss
0.1398 (0.1587)    Prec 95.312% (94.722%)
Epoch: [34][300/391]    Time 0.094 (0.096)      Data 0.001 (0.004)      Loss
0.0664 (0.1594)    Prec 97.656% (94.664%)
Validation starts
Test: [0/79]    Time 0.561 (0.561)      Loss 0.3172 (0.3172)    Prec 89.844%
(89.844%)
 * Prec 87.210%
best acc: 88.760000
Epoch: [35][0/391]      Time 0.731 (0.731)      Data 0.689 (0.689)      Loss
0.1097 (0.1097)    Prec 96.094% (96.094%)
Epoch: [35][100/391]    Time 0.096 (0.101)      Data 0.002 (0.008)      Loss
0.1533 (0.1499)    Prec 95.312% (94.856%)
```

```
Epoch: [35][200/391]    Time 0.085 (0.098)      Data 0.002 (0.005)      Loss
0.1919 (0.1545)    Prec 93.750% (94.609%)
Epoch: [35][300/391]    Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
0.1461 (0.1571)    Prec 93.750% (94.573%)
Validation starts
Test: [0/79]    Time 0.518 (0.518)      Loss 0.3127 (0.3127)    Prec 89.844%
(89.844%)
 * Prec 87.740%
best acc: 88.760000
Epoch: [36][0/391]      Time 0.636 (0.636)      Data 0.595 (0.595)      Loss
0.1383 (0.1383)    Prec 94.531% (94.531%)
Epoch: [36][100/391]    Time 0.095 (0.100)      Data 0.002 (0.007)      Loss
0.1282 (0.1417)    Prec 96.875% (95.251%)
Epoch: [36][200/391]    Time 0.050 (0.098)      Data 0.003 (0.005)      Loss
0.3567 (0.1484)    Prec 89.062% (94.912%)
Epoch: [36][300/391]    Time 0.094 (0.095)      Data 0.001 (0.004)      Loss
0.1887 (0.1501)    Prec 93.750% (94.874%)
Validation starts
Test: [0/79]    Time 0.549 (0.549)      Loss 0.3492 (0.3492)    Prec 89.844%
(89.844%)
 * Prec 87.920%
best acc: 88.760000
Epoch: [37][0/391]      Time 0.732 (0.732)      Data 0.678 (0.678)      Loss
0.0734 (0.0734)    Prec 96.875% (96.875%)
Epoch: [37][100/391]    Time 0.094 (0.101)      Data 0.002 (0.008)      Loss
0.1773 (0.1406)    Prec 92.969% (95.320%)
Epoch: [37][200/391]    Time 0.090 (0.097)      Data 0.002 (0.005)      Loss
0.2371 (0.1436)    Prec 92.188% (95.149%)
Epoch: [37][300/391]    Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
0.0807 (0.1467)    Prec 96.875% (94.975%)
Validation starts
Test: [0/79]    Time 0.548 (0.548)      Loss 0.4087 (0.4087)    Prec 86.719%
(86.719%)
 * Prec 86.220%
best acc: 88.760000
Epoch: [38][0/391]      Time 0.737 (0.737)      Data 0.695 (0.695)      Loss
0.1488 (0.1488)    Prec 95.312% (95.312%)
Epoch: [38][100/391]    Time 0.096 (0.101)      Data 0.001 (0.008)      Loss
0.1306 (0.1293)    Prec 93.750% (95.521%)
Epoch: [38][200/391]    Time 0.090 (0.097)      Data 0.001 (0.005)      Loss
0.1977 (0.1383)    Prec 92.188% (95.285%)
Epoch: [38][300/391]    Time 0.094 (0.095)      Data 0.001 (0.004)      Loss
0.1412 (0.1441)    Prec 95.312% (95.087%)
Validation starts
Test: [0/79]    Time 0.742 (0.742)      Loss 0.2541 (0.2541)    Prec 88.281%
(88.281%)
 * Prec 87.970%
best acc: 88.760000
```

```
Epoch: [39][0/391]      Time 0.622 (0.622)      Data 0.580 (0.580)      Loss
0.1637 (0.1637)    Prec 95.312% (95.312%)
Epoch: [39][100/391]    Time 0.094 (0.100)      Data 0.001 (0.007)      Loss
0.1323 (0.1340)    Prec 96.094% (95.545%)
Epoch: [39][200/391]    Time 0.029 (0.097)      Data 0.004 (0.004)      Loss
0.1459 (0.1341)    Prec 96.875% (95.398%)
Epoch: [39][300/391]    Time 0.094 (0.095)      Data 0.002 (0.004)      Loss
0.1644 (0.1365)    Prec 95.312% (95.305%)
Validation starts
Test: [0/79]    Time 0.588 (0.588)      Loss 0.3584 (0.3584)    Prec 89.844%
(89.844%)
 * Prec 88.200%
best acc: 88.760000
Epoch: [40][0/391]      Time 0.622 (0.622)      Data 0.580 (0.580)      Loss
0.1060 (0.1060)    Prec 96.875% (96.875%)
Epoch: [40][100/391]    Time 0.095 (0.100)      Data 0.002 (0.007)      Loss
0.2146 (0.1310)    Prec 95.312% (95.614%)
Epoch: [40][200/391]    Time 0.116 (0.097)      Data 0.091 (0.005)      Loss
0.1475 (0.1277)    Prec 95.312% (95.682%)
Epoch: [40][300/391]    Time 0.094 (0.095)      Data 0.002 (0.004)      Loss
0.1732 (0.1292)    Prec 92.969% (95.616%)
Validation starts
Test: [0/79]    Time 0.594 (0.594)      Loss 0.4393 (0.4393)    Prec 89.844%
(89.844%)
 * Prec 88.250%
best acc: 88.760000
Epoch: [41][0/391]      Time 0.743 (0.743)      Data 0.705 (0.705)      Loss
0.0935 (0.0935)    Prec 97.656% (97.656%)
Epoch: [41][100/391]    Time 0.095 (0.101)      Data 0.002 (0.009)      Loss
0.1267 (0.1210)    Prec 96.094% (95.862%)
Epoch: [41][200/391]    Time 0.183 (0.098)      Data 0.002 (0.006)      Loss
0.2342 (0.1287)    Prec 92.188% (95.550%)
Epoch: [41][300/391]    Time 0.094 (0.095)      Data 0.001 (0.005)      Loss
0.2287 (0.1328)    Prec 93.750% (95.484%)
Validation starts
Test: [0/79]    Time 0.793 (0.793)      Loss 0.2964 (0.2964)    Prec 89.062%
(89.062%)
 * Prec 87.930%
best acc: 88.760000
Epoch: [42][0/391]      Time 1.163 (1.163)      Data 1.119 (1.119)      Loss
0.1028 (0.1028)    Prec 96.094% (96.094%)
Epoch: [42][100/391]    Time 0.094 (0.105)      Data 0.001 (0.013)      Loss
0.0730 (0.1056)    Prec 96.094% (96.210%)
Epoch: [42][200/391]    Time 0.207 (0.100)      Data 0.004 (0.007)      Loss
0.0887 (0.1155)    Prec 99.219% (95.938%)
Epoch: [42][300/391]    Time 0.094 (0.097)      Data 0.002 (0.006)      Loss
0.2328 (0.1204)    Prec 92.188% (95.842%)
Validation starts
```

```
Test: [0/79]    Time 0.577 (0.577)      Loss 0.3217 (0.3217)    Prec 90.625%
(90.625%)
 * Prec 88.370%
best acc: 88.760000
Epoch: [43][0/391]      Time 0.644 (0.644)      Data 0.602 (0.602)      Loss
0.1408 (0.1408)    Prec 94.531% (94.531%)
Epoch: [43][100/391]    Time 0.095 (0.100)      Data 0.002 (0.008)      Loss
0.1419 (0.1194)    Prec 95.312% (96.032%)
Epoch: [43][200/391]    Time 0.124 (0.097)      Data 0.001 (0.005)      Loss
0.1684 (0.1224)    Prec 94.531% (95.775%)
Epoch: [43][300/391]    Time 0.094 (0.095)      Data 0.001 (0.004)      Loss
0.1073 (0.1251)    Prec 96.094% (95.671%)
Validation starts
Test: [0/79]    Time 0.490 (0.490)      Loss 0.4017 (0.4017)    Prec 89.062%
(89.062%)
 * Prec 87.520%
best acc: 88.760000
Epoch: [44][0/391]      Time 0.510 (0.510)      Data 0.466 (0.466)      Loss
0.2033 (0.2033)    Prec 93.750% (93.750%)
Epoch: [44][100/391]    Time 0.094 (0.098)      Data 0.002 (0.006)      Loss
0.1726 (0.1180)    Prec 96.094% (95.908%)
Epoch: [44][200/391]    Time 0.178 (0.096)      Data 0.004 (0.004)      Loss
0.1078 (0.1231)    Prec 96.875% (95.662%)
Epoch: [44][300/391]    Time 0.095 (0.095)      Data 0.001 (0.003)      Loss
0.1210 (0.1238)    Prec 96.875% (95.689%)
Validation starts
Test: [0/79]    Time 0.484 (0.484)      Loss 0.3070 (0.3070)    Prec 89.844%
(89.844%)
 * Prec 88.060%
best acc: 88.760000
Epoch: [45][0/391]      Time 0.798 (0.798)      Data 0.761 (0.761)      Loss
0.1138 (0.1138)    Prec 96.094% (96.094%)
Epoch: [45][100/391]    Time 0.094 (0.101)      Data 0.001 (0.009)      Loss
0.0923 (0.1187)    Prec 96.875% (96.078%)
Epoch: [45][200/391]    Time 0.028 (0.097)      Data 0.003 (0.005)      Loss
0.0950 (0.1150)    Prec 97.656% (96.094%)
Epoch: [45][300/391]    Time 0.091 (0.093)      Data 0.002 (0.004)      Loss
0.1527 (0.1163)    Prec 92.188% (96.055%)
Validation starts
Test: [0/79]    Time 0.478 (0.478)      Loss 0.3078 (0.3078)    Prec 92.188%
(92.188%)
 * Prec 89.690%
best acc: 89.690000
Epoch: [46][0/391]      Time 0.763 (0.763)      Data 0.722 (0.722)      Loss
0.0814 (0.0814)    Prec 97.656% (97.656%)
Epoch: [46][100/391]    Time 0.094 (0.101)      Data 0.002 (0.009)      Loss
0.0587 (0.1120)    Prec 99.219% (96.163%)
Epoch: [46][200/391]    Time 0.207 (0.098)      Data 0.003 (0.005)      Loss
```

```
0.1481 (0.1094)    Prec 95.312% (96.284%)
Epoch: [46][300/391]    Time 0.094 (0.095)      Data 0.001 (0.004)      Loss
0.0926 (0.1102)    Prec 96.094% (96.221%)
Validation starts
Test: [0/79]    Time 0.880 (0.880)      Loss 0.2920 (0.2920)    Prec 91.406%
(91.406%)
 * Prec 89.090%
best acc: 89.690000
Epoch: [47][0/391]      Time 0.803 (0.803)      Data 0.765 (0.765)      Loss
0.0883 (0.0883)    Prec 97.656% (97.656%)
Epoch: [47][100/391]    Time 0.094 (0.101)      Data 0.001 (0.009)      Loss
0.1278 (0.1001)    Prec 96.094% (96.542%)
Epoch: [47][200/391]    Time 0.028 (0.097)      Data 0.003 (0.005)      Loss
0.0871 (0.1066)    Prec 97.656% (96.385%)
Epoch: [47][300/391]    Time 0.094 (0.095)      Data 0.002 (0.004)      Loss
0.0833 (0.1115)    Prec 96.875% (96.148%)
Validation starts
Test: [0/79]    Time 0.547 (0.547)      Loss 0.3582 (0.3582)    Prec 89.062%
(89.062%)
 * Prec 88.180%
best acc: 89.690000
Epoch: [48][0/391]      Time 0.685 (0.685)      Data 0.644 (0.644)      Loss
0.0585 (0.0585)    Prec 98.438% (98.438%)
Epoch: [48][100/391]    Time 0.097 (0.100)      Data 0.001 (0.008)      Loss
0.1104 (0.1016)    Prec 96.094% (96.442%)
Epoch: [48][200/391]    Time 0.238 (0.098)      Data 0.002 (0.005)      Loss
0.1011 (0.1070)    Prec 96.094% (96.249%)
Epoch: [48][300/391]    Time 0.094 (0.094)      Data 0.001 (0.004)      Loss
0.1086 (0.1078)    Prec 96.094% (96.239%)
Validation starts
Test: [0/79]    Time 0.441 (0.441)      Loss 0.3242 (0.3242)    Prec 89.062%
(89.062%)
 * Prec 88.810%
best acc: 89.690000
Epoch: [49][0/391]      Time 0.666 (0.666)      Data 0.626 (0.626)      Loss
0.1058 (0.1058)    Prec 96.094% (96.094%)
Epoch: [49][100/391]    Time 0.094 (0.100)      Data 0.002 (0.008)      Loss
0.0963 (0.1043)    Prec 96.875% (96.643%)
Epoch: [49][200/391]    Time 0.081 (0.097)      Data 0.002 (0.005)      Loss
0.0704 (0.0996)    Prec 96.875% (96.646%)
Epoch: [49][300/391]    Time 0.095 (0.095)      Data 0.002 (0.004)      Loss
0.1393 (0.1029)    Prec 93.750% (96.517%)
Validation starts
Test: [0/79]    Time 0.436 (0.436)      Loss 0.3873 (0.3873)    Prec 88.281%
(88.281%)
 * Prec 87.680%
best acc: 89.690000
Epoch: [50][0/391]      Time 0.803 (0.803)      Data 0.767 (0.767)      Loss
```

```
0.0764 (0.0764)    Prec 96.875% (96.875%)
Epoch: [50][100/391]    Time 0.091 (0.101)      Data 0.002 (0.009)      Loss
0.0300 (0.0980)    Prec 99.219% (96.558%)
Epoch: [50][200/391]    Time 0.095 (0.098)      Data 0.002 (0.005)      Loss
0.0846 (0.0996)    Prec 96.094% (96.545%)
Epoch: [50][300/391]    Time 0.095 (0.095)      Data 0.002 (0.004)      Loss
0.0607 (0.1023)    Prec 99.219% (96.457%)
Validation starts
Test: [0/79]    Time 0.454 (0.454)      Loss 0.3894 (0.3894)    Prec 89.062%
(89.062%)
 * Prec 88.510%
best acc: 89.690000
Epoch: [51][0/391]      Time 0.812 (0.812)      Data 0.771 (0.771)      Loss
0.1258 (0.1258)    Prec 96.094% (96.094%)
Epoch: [51][100/391]    Time 0.095 (0.101)      Data 0.001 (0.009)      Loss
0.1227 (0.1013)    Prec 95.312% (96.341%)
Epoch: [51][200/391]    Time 0.094 (0.098)      Data 0.002 (0.005)      Loss
0.1984 (0.1013)    Prec 92.969% (96.447%)
Epoch: [51][300/391]    Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
0.0545 (0.1016)    Prec 98.438% (96.436%)
Validation starts
Test: [0/79]    Time 0.575 (0.575)      Loss 0.3686 (0.3686)    Prec 91.406%
(91.406%)
 * Prec 88.960%
best acc: 89.690000
Epoch: [52][0/391]      Time 0.698 (0.698)      Data 0.660 (0.660)      Loss
0.0338 (0.0338)    Prec 100.000% (100.000%)
Epoch: [52][100/391]    Time 0.095 (0.100)      Data 0.002 (0.008)      Loss
0.0383 (0.1015)    Prec 98.438% (96.457%)
Epoch: [52][200/391]    Time 0.094 (0.097)      Data 0.001 (0.005)      Loss
0.1554 (0.0982)    Prec 95.312% (96.634%)
Epoch: [52][300/391]    Time 0.095 (0.095)      Data 0.001 (0.004)      Loss
0.0639 (0.1004)    Prec 97.656% (96.538%)
Validation starts
Test: [0/79]    Time 0.490 (0.490)      Loss 0.2123 (0.2123)    Prec 90.625%
(90.625%)
 * Prec 89.860%
best acc: 89.860000
Epoch: [53][0/391]      Time 1.002 (1.002)      Data 0.957 (0.957)      Loss
0.1364 (0.1364)    Prec 93.750% (93.750%)
Epoch: [53][100/391]    Time 0.092 (0.103)      Data 0.001 (0.011)      Loss
0.1039 (0.0985)    Prec 96.094% (96.550%)
Epoch: [53][200/391]    Time 0.094 (0.099)      Data 0.002 (0.006)      Loss
0.1414 (0.0997)    Prec 96.094% (96.498%)
Epoch: [53][300/391]    Time 0.095 (0.096)      Data 0.001 (0.005)      Loss
0.0913 (0.1004)    Prec 96.094% (96.491%)
Validation starts
Test: [0/79]    Time 0.549 (0.549)      Loss 0.2652 (0.2652)    Prec 90.625%
```

```
(90.625%)
 * Prec 89.950%
best acc: 89.950000
Epoch: [54][0/391]      Time 1.172 (1.172)      Data 1.131 (1.131)      Loss
0.1057 (0.1057)    Prec 96.875% (96.875%)
Epoch: [54][100/391]    Time 0.094 (0.105)      Data 0.002 (0.013)      Loss
0.1146 (0.0927)    Prec 94.531% (96.860%)
Epoch: [54][200/391]    Time 0.190 (0.100)      Data 0.003 (0.007)      Loss
0.0899 (0.0945)    Prec 95.312% (96.801%)
Epoch: [54][300/391]    Time 0.092 (0.096)      Data 0.001 (0.006)      Loss
0.1644 (0.0984)    Prec 92.969% (96.688%)
Validation starts
Test: [0/79]    Time 0.589 (0.589)      Loss 0.3274 (0.3274)    Prec 90.625%
(90.625%)
 * Prec 88.610%
best acc: 89.950000
Epoch: [55][0/391]      Time 0.695 (0.695)      Data 0.661 (0.661)      Loss
0.1619 (0.1619)    Prec 95.312% (95.312%)
Epoch: [55][100/391]    Time 0.097 (0.100)      Data 0.001 (0.008)      Loss
0.0739 (0.0806)    Prec 98.438% (97.239%)
Epoch: [55][200/391]    Time 0.112 (0.097)      Data 0.004 (0.005)      Loss
0.0611 (0.0885)    Prec 97.656% (97.003%)
Epoch: [55][300/391]    Time 0.094 (0.094)      Data 0.001 (0.004)      Loss
0.1543 (0.0910)    Prec 93.750% (96.893%)
Validation starts
Test: [0/79]    Time 0.552 (0.552)      Loss 0.2658 (0.2658)    Prec 91.406%
(91.406%)
 * Prec 88.320%
best acc: 89.950000
Epoch: [56][0/391]      Time 0.895 (0.895)      Data 0.860 (0.860)      Loss
0.0541 (0.0541)    Prec 99.219% (99.219%)
Epoch: [56][100/391]    Time 0.094 (0.102)      Data 0.001 (0.010)      Loss
0.0973 (0.0862)    Prec 96.094% (97.076%)
Epoch: [56][200/391]    Time 0.044 (0.098)      Data 0.002 (0.006)      Loss
0.1350 (0.0875)    Prec 96.094% (97.038%)
Epoch: [56][300/391]    Time 0.098 (0.096)      Data 0.001 (0.005)      Loss
0.0624 (0.0899)    Prec 96.875% (96.942%)
Validation starts
Test: [0/79]    Time 0.490 (0.490)      Loss 0.3228 (0.3228)    Prec 92.188%
(92.188%)
 * Prec 89.180%
best acc: 89.950000
Epoch: [57][0/391]      Time 0.968 (0.968)      Data 0.921 (0.921)      Loss
0.0677 (0.0677)    Prec 96.875% (96.875%)
Epoch: [57][100/391]    Time 0.094 (0.103)      Data 0.001 (0.011)      Loss
0.0742 (0.0940)    Prec 96.094% (96.805%)
Epoch: [57][200/391]    Time 0.057 (0.098)      Data 0.002 (0.006)      Loss
0.0301 (0.0888)    Prec 98.438% (96.926%)
```

```
Epoch: [57][300/391]    Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
0.1007 (0.0895)    Prec 96.094% (96.880%)
Validation starts
Test: [0/79]    Time 0.486 (0.486)      Loss 0.4299 (0.4299)    Prec 89.062%
(89.062%)
 * Prec 88.490%
best acc: 89.950000
Epoch: [58][0/391]      Time 0.695 (0.695)      Data 0.654 (0.654)      Loss
0.1275 (0.1275)    Prec 96.875% (96.875%)
Epoch: [58][100/391]    Time 0.091 (0.100)      Data 0.002 (0.008)      Loss
0.0861 (0.0854)    Prec 97.656% (97.146%)
Epoch: [58][200/391]    Time 0.057 (0.097)      Data 0.001 (0.005)      Loss
0.0584 (0.0824)    Prec 96.875% (97.194%)
Epoch: [58][300/391]    Time 0.095 (0.095)      Data 0.002 (0.004)      Loss
0.0581 (0.0855)    Prec 97.656% (97.093%)
Validation starts
Test: [0/79]    Time 0.519 (0.519)      Loss 0.3380 (0.3380)    Prec 89.062%
(89.062%)
 * Prec 89.300%
best acc: 89.950000
Epoch: [59][0/391]      Time 0.684 (0.684)      Data 0.642 (0.642)      Loss
0.0912 (0.0912)    Prec 97.656% (97.656%)
Epoch: [59][100/391]    Time 0.094 (0.100)      Data 0.001 (0.008)      Loss
0.0773 (0.0806)    Prec 96.094% (97.293%)
Epoch: [59][200/391]    Time 0.093 (0.097)      Data 0.001 (0.005)      Loss
0.0847 (0.0889)    Prec 97.656% (96.964%)
Epoch: [59][300/391]    Time 0.092 (0.095)      Data 0.002 (0.005)      Loss
0.0908 (0.0925)    Prec 97.656% (96.857%)
Validation starts
Test: [0/79]    Time 0.695 (0.695)      Loss 0.2203 (0.2203)    Prec 92.969%
(92.969%)
 * Prec 89.000%
best acc: 89.950000
Epoch: [60][0/391]      Time 0.589 (0.589)      Data 0.548 (0.548)      Loss
0.0217 (0.0217)    Prec 100.000% (100.000%)
Epoch: [60][100/391]    Time 0.094 (0.099)      Data 0.002 (0.007)      Loss
0.0625 (0.0891)    Prec 98.438% (97.045%)
Epoch: [60][200/391]    Time 0.096 (0.097)      Data 0.002 (0.004)      Loss
0.0471 (0.0897)    Prec 96.875% (97.034%)
Epoch: [60][300/391]    Time 0.097 (0.095)      Data 0.002 (0.003)      Loss
0.0675 (0.0882)    Prec 97.656% (97.057%)
Validation starts
Test: [0/79]    Time 0.517 (0.517)      Loss 0.3858 (0.3858)    Prec 88.281%
(88.281%)
 * Prec 88.590%
best acc: 89.950000
Epoch: [61][0/391]      Time 0.702 (0.702)      Data 0.657 (0.657)      Loss
0.0769 (0.0769)    Prec 96.875% (96.875%)
```

```
Epoch: [61][100/391]    Time 0.094 (0.100)      Data 0.001 (0.008)      Loss
0.0909 (0.0891)    Prec 96.875% (96.898%)
Epoch: [61][200/391]    Time 0.094 (0.097)      Data 0.002 (0.005)      Loss
0.1368 (0.0851)    Prec 96.094% (97.112%)
Epoch: [61][300/391]    Time 0.094 (0.095)      Data 0.001 (0.004)      Loss
0.0452 (0.0859)    Prec 99.219% (97.093%)
Validation starts
Test: [0/79]    Time 0.604 (0.604)      Loss 0.2534 (0.2534)    Prec 92.188%
(92.188%)
 * Prec 89.820%
best acc: 89.950000
Epoch: [62][0/391]      Time 0.599 (0.599)      Data 0.550 (0.550)      Loss
0.0641 (0.0641)    Prec 97.656% (97.656%)
Epoch: [62][100/391]    Time 0.094 (0.099)      Data 0.001 (0.007)      Loss
0.0758 (0.0767)    Prec 96.875% (97.440%)
Epoch: [62][200/391]    Time 0.094 (0.097)      Data 0.001 (0.004)      Loss
0.1273 (0.0822)    Prec 96.094% (97.209%)
Epoch: [62][300/391]    Time 0.094 (0.095)      Data 0.002 (0.004)      Loss
0.1502 (0.0836)    Prec 93.750% (97.083%)
Validation starts
Test: [0/79]    Time 0.387 (0.387)      Loss 0.2711 (0.2711)    Prec 92.188%
(92.188%)
 * Prec 89.380%
best acc: 89.950000
Epoch: [63][0/391]      Time 0.738 (0.738)      Data 0.693 (0.693)      Loss
0.1195 (0.1195)    Prec 96.875% (96.875%)
Epoch: [63][100/391]    Time 0.094 (0.101)      Data 0.001 (0.008)      Loss
0.0796 (0.0696)    Prec 96.875% (97.633%)
Epoch: [63][200/391]    Time 0.094 (0.098)      Data 0.002 (0.005)      Loss
0.0562 (0.0786)    Prec 97.656% (97.283%)
Epoch: [63][300/391]    Time 0.098 (0.094)      Data 0.002 (0.004)      Loss
0.0828 (0.0822)    Prec 97.656% (97.166%)
Validation starts
Test: [0/79]    Time 0.592 (0.592)      Loss 0.2578 (0.2578)    Prec 93.750%
(93.750%)
 * Prec 88.480%
best acc: 89.950000
Epoch: [64][0/391]      Time 0.769 (0.769)      Data 0.727 (0.727)      Loss
0.0856 (0.0856)    Prec 97.656% (97.656%)
Epoch: [64][100/391]    Time 0.095 (0.101)      Data 0.001 (0.009)      Loss
0.1272 (0.0854)    Prec 95.312% (97.099%)
Epoch: [64][200/391]    Time 0.094 (0.098)      Data 0.002 (0.005)      Loss
0.1749 (0.0825)    Prec 92.188% (97.217%)
Epoch: [64][300/391]    Time 0.095 (0.095)      Data 0.002 (0.004)      Loss
0.1159 (0.0869)    Prec 94.531% (97.036%)
Validation starts
Test: [0/79]    Time 0.525 (0.525)      Loss 0.3188 (0.3188)    Prec 92.188%
(92.188%)
```

```
 * Prec 90.150%
best acc: 90.150000
Epoch: [65][0/391]      Time 0.689 (0.689)      Data 0.650 (0.650)      Loss
0.0276 (0.0276)    Prec 98.438% (98.438%)
Epoch: [65][100/391]    Time 0.094 (0.100)      Data 0.001 (0.008)      Loss
0.1792 (0.0754)    Prec 95.312% (97.339%)
Epoch: [65][200/391]    Time 0.043 (0.097)      Data 0.003 (0.005)      Loss
0.0653 (0.0779)    Prec 98.438% (97.330%)
Epoch: [65][300/391]    Time 0.094 (0.095)      Data 0.002 (0.004)      Loss
0.0520 (0.0831)    Prec 99.219% (97.194%)
Validation starts
Test: [0/79]    Time 0.745 (0.745)      Loss 0.3964 (0.3964)    Prec 89.062%
(89.062%)
 * Prec 88.990%
best acc: 90.150000
Epoch: [66][0/391]      Time 0.656 (0.656)      Data 0.604 (0.604)      Loss
0.0441 (0.0441)    Prec 98.438% (98.438%)
Epoch: [66][100/391]    Time 0.095 (0.100)      Data 0.002 (0.008)      Loss
0.0895 (0.0802)    Prec 96.875% (97.355%)
Epoch: [66][200/391]    Time 0.060 (0.097)      Data 0.002 (0.005)      Loss
0.1123 (0.0820)    Prec 97.656% (97.279%)
Epoch: [66][300/391]    Time 0.094 (0.097)      Data 0.001 (0.004)      Loss
0.0404 (0.0828)    Prec 98.438% (97.199%)
Validation starts
Test: [0/79]    Time 0.538 (0.538)      Loss 0.2997 (0.2997)    Prec 90.625%
(90.625%)
 * Prec 89.430%
best acc: 90.150000
Epoch: [67][0/391]      Time 0.766 (0.766)      Data 0.698 (0.698)      Loss
0.1510 (0.1510)    Prec 94.531% (94.531%)
Epoch: [67][100/391]    Time 0.092 (0.101)      Data 0.001 (0.008)      Loss
0.0730 (0.0811)    Prec 97.656% (97.161%)
Epoch: [67][200/391]    Time 0.042 (0.098)      Data 0.003 (0.005)      Loss
0.0514 (0.0833)    Prec 97.656% (97.135%)
Epoch: [67][300/391]    Time 0.094 (0.097)      Data 0.002 (0.004)      Loss
0.1186 (0.0834)    Prec 94.531% (97.101%)
Validation starts
Test: [0/79]    Time 0.465 (0.465)      Loss 0.2281 (0.2281)    Prec 90.625%
(90.625%)
 * Prec 89.580%
best acc: 90.150000
Epoch: [68][0/391]      Time 0.922 (0.922)      Data 0.878 (0.878)      Loss
0.0528 (0.0528)    Prec 98.438% (98.438%)
Epoch: [68][100/391]    Time 0.094 (0.103)      Data 0.002 (0.010)      Loss
0.0723 (0.0752)    Prec 96.094% (97.386%)
Epoch: [68][200/391]    Time 0.088 (0.099)      Data 0.001 (0.006)      Loss
0.0491 (0.0761)    Prec 98.438% (97.357%)
Epoch: [68][300/391]    Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
```

```
0.0756 (0.0793)    Prec 97.656% (97.257%)
Validation starts
Test: [0/79]    Time 0.577 (0.577)      Loss 0.3233 (0.3233)    Prec 89.844%
(89.844%)
 * Prec 89.350%
best acc: 90.150000
Epoch: [69][0/391]      Time 0.808 (0.808)      Data 0.764 (0.764)      Loss
0.0764 (0.0764)    Prec 96.094% (96.094%)
Epoch: [69][100/391]    Time 0.094 (0.101)      Data 0.001 (0.009)      Loss
0.0726 (0.0798)    Prec 96.875% (97.370%)
Epoch: [69][200/391]    Time 0.188 (0.098)      Data 0.002 (0.005)      Loss
0.0871 (0.0829)    Prec 96.875% (97.244%)
Epoch: [69][300/391]    Time 0.094 (0.095)      Data 0.001 (0.004)      Loss
0.0620 (0.0832)    Prec 98.438% (97.168%)
Validation starts
Test: [0/79]    Time 0.498 (0.498)      Loss 0.3857 (0.3857)    Prec 90.625%
(90.625%)
 * Prec 89.690%
best acc: 90.150000
Epoch: [70][0/391]      Time 0.672 (0.672)      Data 0.630 (0.630)      Loss
0.0300 (0.0300)    Prec 99.219% (99.219%)
Epoch: [70][100/391]    Time 0.094 (0.100)      Data 0.002 (0.008)      Loss
0.0493 (0.0657)    Prec 98.438% (97.649%)
Epoch: [70][200/391]    Time 0.044 (0.097)      Data 0.001 (0.005)      Loss
0.0998 (0.0706)    Prec 95.312% (97.524%)
Epoch: [70][300/391]    Time 0.094 (0.094)      Data 0.002 (0.004)      Loss
0.0310 (0.0750)    Prec 98.438% (97.332%)
Validation starts
Test: [0/79]    Time 0.569 (0.569)      Loss 0.3029 (0.3029)    Prec 91.406%
(91.406%)
 * Prec 90.010%
best acc: 90.150000
Epoch: [71][0/391]      Time 1.139 (1.139)      Data 1.097 (1.097)      Loss
0.0402 (0.0402)    Prec 98.438% (98.438%)
Epoch: [71][100/391]    Time 0.095 (0.105)      Data 0.002 (0.012)      Loss
0.1285 (0.0747)    Prec 95.312% (97.401%)
Epoch: [71][200/391]    Time 0.094 (0.099)      Data 0.002 (0.007)      Loss
0.0159 (0.0719)    Prec 100.000% (97.520%)
Epoch: [71][300/391]    Time 0.094 (0.097)      Data 0.001 (0.006)      Loss
0.0645 (0.0747)    Prec 97.656% (97.376%)
Validation starts
Test: [0/79]    Time 0.455 (0.455)      Loss 0.4269 (0.4269)    Prec 90.625%
(90.625%)
 * Prec 88.800%
best acc: 90.150000
Epoch: [72][0/391]      Time 0.731 (0.731)      Data 0.689 (0.689)      Loss
0.0464 (0.0464)    Prec 97.656% (97.656%)
Epoch: [72][100/391]    Time 0.094 (0.101)      Data 0.002 (0.008)      Loss
```

```
0.0976 (0.0742)    Prec 96.875% (97.416%)
Epoch: [72][200/391]    Time 0.094 (0.098)      Data 0.002 (0.005)      Loss
0.0483 (0.0822)    Prec 98.438% (97.178%)
Epoch: [72][300/391]    Time 0.092 (0.095)      Data 0.002 (0.004)      Loss
0.1090 (0.0803)    Prec 97.656% (97.244%)
Validation starts
Test: [0/79]    Time 0.617 (0.617)      Loss 0.2731 (0.2731)    Prec 92.188%
(92.188%)
 * Prec 89.840%
best acc: 90.150000
Epoch: [73][0/391]      Time 0.624 (0.624)      Data 0.583 (0.583)      Loss
0.0811 (0.0811)    Prec 97.656% (97.656%)
Epoch: [73][100/391]    Time 0.095 (0.100)      Data 0.001 (0.007)      Loss
0.0671 (0.0674)    Prec 96.875% (97.672%)
Epoch: [73][200/391]    Time 0.094 (0.097)      Data 0.002 (0.004)      Loss
0.1456 (0.0751)    Prec 94.531% (97.396%)
Epoch: [73][300/391]    Time 0.094 (0.095)      Data 0.001 (0.004)      Loss
0.0353 (0.0757)    Prec 99.219% (97.456%)
Validation starts
Test: [0/79]    Time 0.407 (0.407)      Loss 0.2800 (0.2800)    Prec 92.969%
(92.969%)
 * Prec 89.330%
best acc: 90.150000
Epoch: [74][0/391]      Time 0.836 (0.836)      Data 0.794 (0.794)      Loss
0.1059 (0.1059)    Prec 95.312% (95.312%)
Epoch: [74][100/391]    Time 0.094 (0.102)      Data 0.001 (0.009)      Loss
0.0180 (0.0722)    Prec 100.000% (97.641%)
Epoch: [74][200/391]    Time 0.094 (0.098)      Data 0.002 (0.005)      Loss
0.0956 (0.0731)    Prec 96.875% (97.555%)
Epoch: [74][300/391]    Time 0.095 (0.096)      Data 0.002 (0.005)      Loss
0.0725 (0.0770)    Prec 98.438% (97.410%)
Validation starts
Test: [0/79]    Time 0.506 (0.506)      Loss 0.2762 (0.2762)    Prec 92.188%
(92.188%)
 * Prec 89.240%
best acc: 90.150000
Epoch: [75][0/391]      Time 0.815 (0.815)      Data 0.772 (0.772)      Loss
0.0789 (0.0789)    Prec 97.656% (97.656%)
Epoch: [75][100/391]    Time 0.095 (0.101)      Data 0.001 (0.009)      Loss
0.0430 (0.0689)    Prec 98.438% (97.734%)
Epoch: [75][200/391]    Time 0.095 (0.098)      Data 0.001 (0.005)      Loss
0.0825 (0.0745)    Prec 97.656% (97.524%)
Epoch: [75][300/391]    Time 0.094 (0.096)      Data 0.002 (0.004)      Loss
0.0793 (0.0741)    Prec 96.875% (97.464%)
Validation starts
Test: [0/79]    Time 0.590 (0.590)      Loss 0.3640 (0.3640)    Prec 89.844%
(89.844%)
 * Prec 88.500%
```

```
best acc: 90.150000
Epoch: [76][0/391]      Time 0.814 (0.814)      Data 0.769 (0.769)      Loss
0.0492 (0.0492)    Prec 98.438% (98.438%)
Epoch: [76][100/391]    Time 0.095 (0.101)      Data 0.002 (0.009)      Loss
0.0948 (0.0716)    Prec 96.094% (97.610%)
Epoch: [76][200/391]    Time 0.097 (0.098)      Data 0.001 (0.005)      Loss
0.1388 (0.0797)    Prec 93.750% (97.252%)
Epoch: [76][300/391]    Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
0.0715 (0.0780)    Prec 96.875% (97.298%)
Validation starts
Test: [0/79]    Time 0.589 (0.589)      Loss 0.3419 (0.3419)    Prec 92.969%
(92.969%)
 * Prec 89.040%
best acc: 90.150000
Epoch: [77][0/391]      Time 0.947 (0.947)      Data 0.906 (0.906)      Loss
0.1051 (0.1051)    Prec 96.094% (96.094%)
Epoch: [77][100/391]    Time 0.091 (0.103)      Data 0.002 (0.011)      Loss
0.0131 (0.0634)    Prec 100.000% (97.672%)
Epoch: [77][200/391]    Time 0.094 (0.099)      Data 0.001 (0.006)      Loss
0.0346 (0.0723)    Prec 98.438% (97.400%)
Epoch: [77][300/391]    Time 0.094 (0.096)      Data 0.002 (0.005)      Loss
0.0481 (0.0720)    Prec 98.438% (97.423%)
Validation starts
Test: [0/79]    Time 0.550 (0.550)      Loss 0.3288 (0.3288)    Prec 90.625%
(90.625%)
 * Prec 89.970%
best acc: 90.150000
Epoch: [78][0/391]      Time 1.031 (1.031)      Data 0.990 (0.990)      Loss
0.0792 (0.0792)    Prec 97.656% (97.656%)
Epoch: [78][100/391]    Time 0.094 (0.104)      Data 0.001 (0.011)      Loss
0.1045 (0.0653)    Prec 94.531% (97.672%)
Epoch: [78][200/391]    Time 0.094 (0.099)      Data 0.001 (0.006)      Loss
0.0903 (0.0721)    Prec 96.875% (97.547%)
Epoch: [78][300/391]    Time 0.094 (0.096)      Data 0.001 (0.005)      Loss
0.0843 (0.0760)    Prec 95.312% (97.412%)
Validation starts
Test: [0/79]    Time 0.551 (0.551)      Loss 0.2955 (0.2955)    Prec 92.188%
(92.188%)
 * Prec 89.710%
best acc: 90.150000
Epoch: [79][0/391]      Time 0.595 (0.595)      Data 0.552 (0.552)      Loss
0.0718 (0.0718)    Prec 96.094% (96.094%)
Epoch: [79][100/391]    Time 0.094 (0.099)      Data 0.002 (0.007)      Loss
0.0868 (0.0739)    Prec 97.656% (97.471%)
Epoch: [79][200/391]    Time 0.096 (0.097)      Data 0.002 (0.004)      Loss
0.0596 (0.0775)    Prec 98.438% (97.322%)
Epoch: [79][300/391]    Time 0.096 (0.095)      Data 0.001 (0.004)      Loss
0.0931 (0.0776)    Prec 95.312% (97.402%)
```

```
Validation starts
Test: [0/79]    Time 0.674 (0.674)      Loss 0.2899 (0.2899)    Prec 89.844%
(89.844%)
 * Prec 89.660%
best acc: 90.150000
```

[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)
```

[3]:
```
PATH = "result/VGG16_1/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: 9015/10000 (90%)
```

[4]:
```
weight = model.features[0].weight.abs().sum()
print(weight)
```

```
tensor(316.1834, device='cuda:0', grad_fn=<SumBackward0>)
```

[ ]:
```

```