Project_Q1_VGG16

November 26, 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...')
device = torch.device("cuda")


batch_size = 128
model_name = "VGG16_quant_project"
model = VGG16_quant_project()

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)
```

```
=> 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
```

[2]:
```
print_freq = 100
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,adjust_list):
    """For resnet, the lr starts from 0.1, and is divided by 10 at 80 and 120␣
 ↪epochs"""

    if epoch in adjust_list:
        for param_group in optimizer.param_groups:
            param_group['lr'] = param_group['lr'] * 0.1
```

[3]:
```
lr = 4e-2
weight_decay = 1e-4
epochs = 100
best_prec = 0
model.cuda()
criterion = nn.CrossEntropyLoss().cuda()
optimizer = torch.optim.SGD(model.parameters(), lr=lr, momentum=0.
 ↪9,weight_decay=weight_decay)

if not os.path.exists('result'):
    os.makedirs('result')

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

adjust_list = [80,90]
```

```python
for epoch in range(0, epochs):
    adjust_learning_rate(optimizer, epoch,adjust_list)

    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 0.412 (0.412)      Data 0.251 (0.251)      Loss
2.4524 (2.4524)    Prec 10.938% (10.938%)
Epoch: [0][100/391]     Time 0.054 (0.058)      Data 0.002 (0.004)      Loss
2.2622 (3.4320)    Prec 8.594% (10.582%)
Epoch: [0][200/391]     Time 0.055 (0.057)      Data 0.002 (0.003)      Loss
2.0569 (2.8186)    Prec 17.188% (13.254%)
Epoch: [0][300/391]     Time 0.060 (0.056)      Data 0.002 (0.003)      Loss
2.0206 (2.5567)    Prec 21.875% (15.695%)
Validation starts
Test: [0/79]    Time 0.205 (0.205)      Loss 1.9066 (1.9066)    Prec 29.688%
(29.688%)
 * Prec 24.470%
best acc: 24.470000
Epoch: [1][0/391]       Time 0.308 (0.308)      Data 0.257 (0.257)      Loss
1.9837 (1.9837)    Prec 21.094% (21.094%)
Epoch: [1][100/391]     Time 0.058 (0.057)      Data 0.002 (0.004)      Loss
1.9341 (1.9230)    Prec 18.750% (22.741%)
Epoch: [1][200/391]     Time 0.054 (0.056)      Data 0.003 (0.003)      Loss
1.8041 (1.9032)    Prec 31.250% (23.865%)
```

```
Epoch: [1][300/391]     Time 0.053 (0.056)      Data 0.001 (0.002)      Loss
1.9593 (1.8844)    Prec 17.969% (24.886%)
Validation starts
Test: [0/79]    Time 0.235 (0.235)      Loss 1.9639 (1.9639)    Prec 25.781%
(25.781%)
 * Prec 29.070%
best acc: 29.070000
Epoch: [2][0/391]       Time 0.297 (0.297)      Data 0.255 (0.255)      Loss
1.8296 (1.8296)    Prec 32.812% (32.812%)
Epoch: [2][100/391]     Time 0.049 (0.057)      Data 0.001 (0.004)      Loss
1.7276 (1.8146)    Prec 34.375% (29.061%)
Epoch: [2][200/391]     Time 0.056 (0.056)      Data 0.001 (0.003)      Loss
1.7497 (1.7968)    Prec 32.812% (30.092%)
Epoch: [2][300/391]     Time 0.058 (0.056)      Data 0.002 (0.002)      Loss
1.8450 (1.7774)    Prec 32.031% (31.157%)
Validation starts
Test: [0/79]    Time 0.234 (0.234)      Loss 1.6269 (1.6269)    Prec 40.625%
(40.625%)
 * Prec 35.490%
best acc: 35.490000
Epoch: [3][0/391]       Time 0.319 (0.319)      Data 0.264 (0.264)      Loss
1.6004 (1.6004)    Prec 35.938% (35.938%)
Epoch: [3][100/391]     Time 0.056 (0.057)      Data 0.001 (0.004)      Loss
1.6629 (1.6449)    Prec 39.062% (36.781%)
Epoch: [3][200/391]     Time 0.052 (0.056)      Data 0.002 (0.003)      Loss
1.4720 (1.6014)    Prec 49.219% (38.930%)
Epoch: [3][300/391]     Time 0.061 (0.056)      Data 0.001 (0.002)      Loss
1.4026 (1.5702)    Prec 50.000% (40.503%)
Validation starts
Test: [0/79]    Time 0.243 (0.243)      Loss 1.4099 (1.4099)    Prec 46.094%
(46.094%)
 * Prec 46.430%
best acc: 46.430000
Epoch: [4][0/391]       Time 0.303 (0.303)      Data 0.258 (0.258)      Loss
1.4986 (1.4986)    Prec 45.312% (45.312%)
Epoch: [4][100/391]     Time 0.053 (0.057)      Data 0.001 (0.004)      Loss
1.3995 (1.4259)    Prec 50.000% (46.782%)
Epoch: [4][200/391]     Time 0.056 (0.056)      Data 0.002 (0.003)      Loss
1.3926 (1.3982)    Prec 50.000% (47.936%)
Epoch: [4][300/391]     Time 0.060 (0.056)      Data 0.001 (0.002)      Loss
1.3221 (1.3802)    Prec 48.438% (48.816%)
Validation starts
Test: [0/79]    Time 0.243 (0.243)      Loss 1.3974 (1.3974)    Prec 45.312%
(45.312%)
 * Prec 48.850%
best acc: 48.850000
Epoch: [5][0/391]       Time 0.328 (0.328)      Data 0.282 (0.282)      Loss
1.3581 (1.3581)    Prec 52.344% (52.344%)
```

```
Epoch: [5][100/391]     Time 0.054 (0.058)      Data 0.001 (0.004)      Loss
1.0712 (1.2298)    Prec 65.625% (54.680%)
Epoch: [5][200/391]     Time 0.058 (0.056)      Data 0.001 (0.003)      Loss
1.1835 (1.2244)    Prec 55.469% (54.936%)
Epoch: [5][300/391]     Time 0.053 (0.056)      Data 0.001 (0.002)      Loss
1.1784 (1.2124)    Prec 55.469% (55.383%)
Validation starts
Test: [0/79]    Time 0.205 (0.205)      Loss 1.3428 (1.3428)    Prec 57.031%
(57.031%)
 * Prec 51.820%
best acc: 51.820000
Epoch: [6][0/391]       Time 0.290 (0.290)      Data 0.242 (0.242)      Loss
1.2816 (1.2816)    Prec 60.938% (60.938%)
Epoch: [6][100/391]     Time 0.054 (0.057)      Data 0.001 (0.004)      Loss
1.1514 (1.1086)    Prec 57.031% (59.723%)
Epoch: [6][200/391]     Time 0.050 (0.056)      Data 0.002 (0.003)      Loss
1.0208 (1.0889)    Prec 64.844% (60.421%)
Epoch: [6][300/391]     Time 0.054 (0.056)      Data 0.001 (0.002)      Loss
1.1456 (1.0790)    Prec 57.812% (60.800%)
Validation starts
Test: [0/79]    Time 0.231 (0.231)      Loss 0.9607 (0.9607)    Prec 64.062%
(64.062%)
 * Prec 62.770%
best acc: 62.770000
Epoch: [7][0/391]       Time 0.307 (0.307)      Data 0.255 (0.255)      Loss
1.0684 (1.0684)    Prec 62.500% (62.500%)
Epoch: [7][100/391]     Time 0.050 (0.058)      Data 0.001 (0.004)      Loss
0.9600 (0.9873)    Prec 71.875% (64.140%)
Epoch: [7][200/391]     Time 0.051 (0.057)      Data 0.002 (0.003)      Loss
1.1038 (0.9836)    Prec 59.375% (64.307%)
Epoch: [7][300/391]     Time 0.050 (0.057)      Data 0.001 (0.002)      Loss
0.7966 (0.9731)    Prec 69.531% (64.685%)
Validation starts
Test: [0/79]    Time 0.236 (0.236)      Loss 0.9273 (0.9273)    Prec 64.062%
(64.062%)
 * Prec 64.210%
best acc: 64.210000
Epoch: [8][0/391]       Time 0.278 (0.278)      Data 0.236 (0.236)      Loss
1.2173 (1.2173)    Prec 62.500% (62.500%)
Epoch: [8][100/391]     Time 0.055 (0.060)      Data 0.001 (0.004)      Loss
1.1647 (0.9021)    Prec 56.250% (67.567%)
Epoch: [8][200/391]     Time 0.054 (0.057)      Data 0.001 (0.003)      Loss
0.8566 (0.9000)    Prec 71.875% (67.716%)
Epoch: [8][300/391]     Time 0.050 (0.057)      Data 0.002 (0.002)      Loss
0.8749 (0.8899)    Prec 70.312% (68.236%)
Validation starts
Test: [0/79]    Time 0.257 (0.257)      Loss 0.7249 (0.7249)    Prec 75.781%
(75.781%)
```

```
 * Prec 67.000%
best acc: 67.000000
Epoch: [9][0/391]       Time 0.313 (0.313)      Data 0.266 (0.266)      Loss
0.8011 (0.8011)    Prec 71.875% (71.875%)
Epoch: [9][100/391]     Time 0.055 (0.058)      Data 0.001 (0.004)      Loss
0.9520 (0.8259)    Prec 63.281% (70.568%)
Epoch: [9][200/391]     Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.8286 (0.8244)    Prec 71.094% (70.721%)
Epoch: [9][300/391]     Time 0.054 (0.056)      Data 0.001 (0.002)      Loss
0.8589 (0.8209)    Prec 67.969% (70.878%)
Validation starts
Test: [0/79]    Time 0.221 (0.221)      Loss 1.0047 (1.0047)    Prec 64.844%
(64.844%)
 * Prec 61.930%
best acc: 67.000000
Epoch: [10][0/391]      Time 0.275 (0.275)      Data 0.235 (0.235)      Loss
0.7420 (0.7420)    Prec 74.219% (74.219%)
Epoch: [10][100/391]    Time 0.053 (0.058)      Data 0.002 (0.004)      Loss
0.7695 (0.7629)    Prec 69.531% (73.329%)
Epoch: [10][200/391]    Time 0.059 (0.057)      Data 0.002 (0.003)      Loss
0.8913 (0.7705)    Prec 66.406% (73.002%)
Epoch: [10][300/391]    Time 0.058 (0.056)      Data 0.001 (0.002)      Loss
0.7936 (0.7678)    Prec 71.094% (73.144%)
Validation starts
Test: [0/79]    Time 0.243 (0.243)      Loss 0.7571 (0.7571)    Prec 77.344%
(77.344%)
 * Prec 71.550%
best acc: 71.550000
Epoch: [11][0/391]      Time 0.276 (0.276)      Data 0.230 (0.230)      Loss
0.6766 (0.6766)    Prec 76.562% (76.562%)
Epoch: [11][100/391]    Time 0.051 (0.058)      Data 0.001 (0.004)      Loss
0.7478 (0.7299)    Prec 72.656% (74.776%)
Epoch: [11][200/391]    Time 0.055 (0.056)      Data 0.002 (0.003)      Loss
0.6359 (0.7215)    Prec 77.344% (75.051%)
Epoch: [11][300/391]    Time 0.055 (0.056)      Data 0.002 (0.002)      Loss
0.5976 (0.7135)    Prec 78.125% (75.392%)
Validation starts
Test: [0/79]    Time 0.207 (0.207)      Loss 0.7488 (0.7488)    Prec 71.094%
(71.094%)
 * Prec 74.360%
best acc: 74.360000
Epoch: [12][0/391]      Time 0.289 (0.289)      Data 0.242 (0.242)      Loss
0.6230 (0.6230)    Prec 78.125% (78.125%)
Epoch: [12][100/391]    Time 0.052 (0.058)      Data 0.001 (0.004)      Loss
0.5143 (0.6707)    Prec 84.375% (77.073%)
Epoch: [12][200/391]    Time 0.059 (0.057)      Data 0.001 (0.003)      Loss
0.7863 (0.6661)    Prec 73.438% (76.998%)
Epoch: [12][300/391]    Time 0.056 (0.056)      Data 0.001 (0.003)      Loss
```

```
0.7007 (0.6657)    Prec 76.562% (77.100%)
Validation starts
Test: [0/79]    Time 0.222 (0.222)      Loss 0.7022 (0.7022)    Prec 75.781%
(75.781%)
 * Prec 74.370%
best acc: 74.370000
Epoch: [13][0/391]      Time 0.313 (0.313)      Data 0.266 (0.266)      Loss
0.5199 (0.5199)    Prec 82.031% (82.031%)
Epoch: [13][100/391]    Time 0.052 (0.058)      Data 0.003 (0.004)      Loss
0.7801 (0.6298)    Prec 77.344% (78.434%)
Epoch: [13][200/391]    Time 0.068 (0.057)      Data 0.002 (0.003)      Loss
0.6878 (0.6310)    Prec 74.219% (78.172%)
Epoch: [13][300/391]    Time 0.056 (0.056)      Data 0.001 (0.003)      Loss
0.6724 (0.6257)    Prec 76.562% (78.307%)
Validation starts
Test: [0/79]    Time 0.236 (0.236)      Loss 0.7516 (0.7516)    Prec 71.875%
(71.875%)
 * Prec 75.800%
best acc: 75.800000
Epoch: [14][0/391]      Time 0.262 (0.262)      Data 0.220 (0.220)      Loss
0.6366 (0.6366)    Prec 78.906% (78.906%)
Epoch: [14][100/391]    Time 0.048 (0.057)      Data 0.002 (0.004)      Loss
0.5546 (0.5854)    Prec 82.031% (79.827%)
Epoch: [14][200/391]    Time 0.051 (0.056)      Data 0.001 (0.003)      Loss
0.5106 (0.5856)    Prec 84.375% (80.061%)
Epoch: [14][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.5514 (0.5886)    Prec 79.688% (79.880%)
Validation starts
Test: [0/79]    Time 0.233 (0.233)      Loss 0.6609 (0.6609)    Prec 80.469%
(80.469%)
 * Prec 76.290%
best acc: 76.290000
Epoch: [15][0/391]      Time 0.258 (0.258)      Data 0.211 (0.211)      Loss
0.4898 (0.4898)    Prec 88.281% (88.281%)
Epoch: [15][100/391]    Time 0.059 (0.057)      Data 0.002 (0.004)      Loss
0.4130 (0.5532)    Prec 82.031% (81.343%)
Epoch: [15][200/391]    Time 0.053 (0.057)      Data 0.002 (0.003)      Loss
0.6900 (0.5536)    Prec 76.562% (81.207%)
Epoch: [15][300/391]    Time 0.054 (0.057)      Data 0.001 (0.002)      Loss
0.5596 (0.5621)    Prec 78.125% (80.915%)
Validation starts
Test: [0/79]    Time 0.229 (0.229)      Loss 0.5778 (0.5778)    Prec 80.469%
(80.469%)
 * Prec 78.930%
best acc: 78.930000
Epoch: [16][0/391]      Time 0.308 (0.308)      Data 0.260 (0.260)      Loss
0.6491 (0.6491)    Prec 78.125% (78.125%)
Epoch: [16][100/391]    Time 0.053 (0.058)      Data 0.002 (0.004)      Loss
```

```
0.4911 (0.5290)    Prec 81.250% (81.915%)
Epoch: [16][200/391]    Time 0.056 (0.057)      Data 0.002 (0.003)      Loss
0.5030 (0.5356)    Prec 82.812% (81.573%)
Epoch: [16][300/391]    Time 0.064 (0.056)      Data 0.002 (0.003)      Loss
0.5472 (0.5347)    Prec 80.469% (81.689%)
Validation starts
Test: [0/79]    Time 0.220 (0.220)      Loss 0.6140 (0.6140)    Prec 79.688%
(79.688%)
 * Prec 78.650%
best acc: 78.930000
Epoch: [17][0/391]      Time 0.288 (0.288)      Data 0.239 (0.239)      Loss
0.4693 (0.4693)    Prec 84.375% (84.375%)
Epoch: [17][100/391]    Time 0.055 (0.058)      Data 0.002 (0.004)      Loss
0.4789 (0.5159)    Prec 84.375% (82.387%)
Epoch: [17][200/391]    Time 0.055 (0.057)      Data 0.002 (0.003)      Loss
0.2796 (0.5146)    Prec 89.062% (82.280%)
Epoch: [17][300/391]    Time 0.053 (0.056)      Data 0.002 (0.002)      Loss
0.6301 (0.5128)    Prec 81.250% (82.460%)
Validation starts
Test: [0/79]    Time 0.238 (0.238)      Loss 0.5777 (0.5777)    Prec 82.812%
(82.812%)
 * Prec 80.090%
best acc: 80.090000
Epoch: [18][0/391]      Time 0.278 (0.278)      Data 0.232 (0.232)      Loss
0.4675 (0.4675)    Prec 82.812% (82.812%)
Epoch: [18][100/391]    Time 0.057 (0.059)      Data 0.002 (0.004)      Loss
0.4296 (0.4729)    Prec 82.031% (84.189%)
Epoch: [18][200/391]    Time 0.055 (0.057)      Data 0.002 (0.003)      Loss
0.4280 (0.4844)    Prec 88.281% (83.773%)
Epoch: [18][300/391]    Time 0.056 (0.057)      Data 0.002 (0.002)      Loss
0.6612 (0.4837)    Prec 75.781% (83.711%)
Validation starts
Test: [0/79]    Time 0.221 (0.221)      Loss 0.5731 (0.5731)    Prec 79.688%
(79.688%)
 * Prec 79.990%
best acc: 80.090000
Epoch: [19][0/391]      Time 0.312 (0.312)      Data 0.265 (0.265)      Loss
0.4987 (0.4987)    Prec 82.812% (82.812%)
Epoch: [19][100/391]    Time 0.051 (0.058)      Data 0.003 (0.004)      Loss
0.4900 (0.4619)    Prec 85.938% (84.112%)
Epoch: [19][200/391]    Time 0.058 (0.056)      Data 0.001 (0.003)      Loss
0.5233 (0.4733)    Prec 80.469% (83.788%)
Epoch: [19][300/391]    Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.5465 (0.4721)    Prec 85.938% (84.025%)
Validation starts
Test: [0/79]    Time 0.211 (0.211)      Loss 0.8041 (0.8041)    Prec 72.656%
(72.656%)
 * Prec 77.030%
```

```
best acc: 80.090000
Epoch: [20][0/391]      Time 0.314 (0.314)      Data 0.274 (0.274)      Loss
0.4098 (0.4098)    Prec 86.719% (86.719%)
Epoch: [20][100/391]    Time 0.053 (0.058)      Data 0.001 (0.004)      Loss
0.6044 (0.4395)    Prec 80.469% (85.102%)
Epoch: [20][200/391]    Time 0.054 (0.057)      Data 0.002 (0.003)      Loss
0.3906 (0.4492)    Prec 85.938% (84.643%)
Epoch: [20][300/391]    Time 0.056 (0.056)      Data 0.001 (0.002)      Loss
0.5448 (0.4474)    Prec 80.469% (84.738%)
Validation starts
Test: [0/79]    Time 0.238 (0.238)      Loss 0.5606 (0.5606)    Prec 81.250%
(81.250%)
 * Prec 80.420%
best acc: 80.420000
Epoch: [21][0/391]      Time 0.305 (0.305)      Data 0.256 (0.256)      Loss
0.3029 (0.3029)    Prec 92.969% (92.969%)
Epoch: [21][100/391]    Time 0.058 (0.059)      Data 0.002 (0.004)      Loss
0.3994 (0.4313)    Prec 86.719% (85.574%)
Epoch: [21][200/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.4682 (0.4309)    Prec 83.594% (85.242%)
Epoch: [21][300/391]    Time 0.053 (0.057)      Data 0.001 (0.002)      Loss
0.5610 (0.4361)    Prec 82.812% (85.208%)
Validation starts
Test: [0/79]    Time 0.225 (0.225)      Loss 0.4311 (0.4311)    Prec 85.938%
(85.938%)
 * Prec 83.150%
best acc: 83.150000
Epoch: [22][0/391]      Time 0.315 (0.315)      Data 0.266 (0.266)      Loss
0.3743 (0.3743)    Prec 86.719% (86.719%)
Epoch: [22][100/391]    Time 0.054 (0.058)      Data 0.001 (0.004)      Loss
0.4179 (0.4104)    Prec 86.719% (85.999%)
Epoch: [22][200/391]    Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.4948 (0.4137)    Prec 85.156% (85.926%)
Epoch: [22][300/391]    Time 0.053 (0.056)      Data 0.001 (0.002)      Loss
0.4352 (0.4143)    Prec 82.031% (85.909%)
Validation starts
Test: [0/79]    Time 0.249 (0.249)      Loss 0.5266 (0.5266)    Prec 82.812%
(82.812%)
 * Prec 82.170%
best acc: 83.150000
Epoch: [23][0/391]      Time 0.303 (0.303)      Data 0.257 (0.257)      Loss
0.2009 (0.2009)    Prec 94.531% (94.531%)
Epoch: [23][100/391]    Time 0.056 (0.058)      Data 0.002 (0.004)      Loss
0.3400 (0.3766)    Prec 88.281% (87.299%)
Epoch: [23][200/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.2960 (0.3815)    Prec 89.062% (87.034%)
Epoch: [23][300/391]    Time 0.058 (0.056)      Data 0.001 (0.002)      Loss
0.3805 (0.3883)    Prec 84.375% (86.911%)
```

```
Validation starts
Test: [0/79]    Time 0.264 (0.264)      Loss 0.5492 (0.5492)    Prec 84.375%
(84.375%)
 * Prec 81.850%
best acc: 83.150000
Epoch: [24][0/391]      Time 0.287 (0.287)      Data 0.244 (0.244)      Loss
0.3930 (0.3930)    Prec 85.938% (85.938%)
Epoch: [24][100/391]    Time 0.060 (0.058)      Data 0.002 (0.004)      Loss
0.3227 (0.3758)    Prec 89.844% (87.260%)
Epoch: [24][200/391]    Time 0.056 (0.057)      Data 0.002 (0.003)      Loss
0.4895 (0.3834)    Prec 82.812% (86.968%)
Epoch: [24][300/391]    Time 0.059 (0.056)      Data 0.002 (0.002)      Loss
0.4819 (0.3854)    Prec 84.375% (86.869%)
Validation starts
Test: [0/79]    Time 0.227 (0.227)      Loss 0.4969 (0.4969)    Prec 84.375%
(84.375%)
 * Prec 84.400%
best acc: 84.400000
Epoch: [25][0/391]      Time 0.298 (0.298)      Data 0.254 (0.254)      Loss
0.3934 (0.3934)    Prec 87.500% (87.500%)
Epoch: [25][100/391]    Time 0.055 (0.057)      Data 0.001 (0.004)      Loss
0.5359 (0.3674)    Prec 80.469% (87.577%)
Epoch: [25][200/391]    Time 0.053 (0.056)      Data 0.001 (0.003)      Loss
0.3100 (0.3659)    Prec 89.844% (87.488%)
Epoch: [25][300/391]    Time 0.051 (0.056)      Data 0.002 (0.002)      Loss
0.4745 (0.3724)    Prec 82.812% (87.370%)
Validation starts
Test: [0/79]    Time 0.241 (0.241)      Loss 0.4263 (0.4263)    Prec 85.156%
(85.156%)
 * Prec 82.950%
best acc: 84.400000
Epoch: [26][0/391]      Time 0.294 (0.294)      Data 0.249 (0.249)      Loss
0.2791 (0.2791)    Prec 88.281% (88.281%)
Epoch: [26][100/391]    Time 0.056 (0.058)      Data 0.002 (0.004)      Loss
0.3177 (0.3473)    Prec 89.062% (87.964%)
Epoch: [26][200/391]    Time 0.055 (0.057)      Data 0.002 (0.003)      Loss
0.4085 (0.3525)    Prec 85.938% (88.064%)
Epoch: [26][300/391]    Time 0.062 (0.056)      Data 0.001 (0.003)      Loss
0.2845 (0.3566)    Prec 91.406% (87.920%)
Validation starts
Test: [0/79]    Time 0.251 (0.251)      Loss 0.4616 (0.4616)    Prec 82.031%
(82.031%)
 * Prec 83.830%
best acc: 84.400000
Epoch: [27][0/391]      Time 0.271 (0.271)      Data 0.230 (0.230)      Loss
0.2215 (0.2215)    Prec 91.406% (91.406%)
Epoch: [27][100/391]    Time 0.056 (0.058)      Data 0.001 (0.004)      Loss
0.1674 (0.3217)    Prec 94.531% (89.233%)
```

```
Epoch: [27][200/391]    Time 0.056 (0.057)      Data 0.002 (0.003)      Loss
0.1803 (0.3330)    Prec 95.312% (88.697%)
Epoch: [27][300/391]    Time 0.058 (0.056)      Data 0.002 (0.002)      Loss
0.3084 (0.3386)    Prec 89.062% (88.486%)
Validation starts
Test: [0/79]    Time 0.238 (0.238)      Loss 0.4120 (0.4120)    Prec 87.500%
(87.500%)
 * Prec 84.470%
best acc: 84.470000
Epoch: [28][0/391]      Time 0.285 (0.285)      Data 0.240 (0.240)      Loss
0.2873 (0.2873)    Prec 92.188% (92.188%)
Epoch: [28][100/391]    Time 0.055 (0.058)      Data 0.001 (0.004)      Loss
0.3800 (0.3211)    Prec 86.719% (89.310%)
Epoch: [28][200/391]    Time 0.049 (0.057)      Data 0.002 (0.003)      Loss
0.2520 (0.3219)    Prec 90.625% (89.175%)
Epoch: [28][300/391]    Time 0.062 (0.056)      Data 0.001 (0.003)      Loss
0.4851 (0.3291)    Prec 82.812% (88.876%)
Validation starts
Test: [0/79]    Time 0.215 (0.215)      Loss 0.3923 (0.3923)    Prec 89.062%
(89.062%)
 * Prec 83.970%
best acc: 84.470000
Epoch: [29][0/391]      Time 0.286 (0.286)      Data 0.239 (0.239)      Loss
0.4689 (0.4689)    Prec 82.031% (82.031%)
Epoch: [29][100/391]    Time 0.058 (0.058)      Data 0.001 (0.004)      Loss
0.1427 (0.3062)    Prec 96.875% (89.643%)
Epoch: [29][200/391]    Time 0.053 (0.057)      Data 0.002 (0.003)      Loss
0.3308 (0.3154)    Prec 88.281% (89.300%)
Epoch: [29][300/391]    Time 0.056 (0.056)      Data 0.002 (0.003)      Loss
0.2776 (0.3204)    Prec 90.625% (89.104%)
Validation starts
Test: [0/79]    Time 0.265 (0.265)      Loss 0.5007 (0.5007)    Prec 85.156%
(85.156%)
 * Prec 83.520%
best acc: 84.470000
Epoch: [30][0/391]      Time 0.272 (0.272)      Data 0.225 (0.225)      Loss
0.1515 (0.1515)    Prec 96.094% (96.094%)
Epoch: [30][100/391]    Time 0.059 (0.057)      Data 0.002 (0.004)      Loss
0.3969 (0.3042)    Prec 85.156% (89.712%)
Epoch: [30][200/391]    Time 0.058 (0.056)      Data 0.002 (0.003)      Loss
0.3082 (0.3144)    Prec 89.844% (89.354%)
Epoch: [30][300/391]    Time 0.058 (0.056)      Data 0.002 (0.002)      Loss
0.3177 (0.3149)    Prec 85.938% (89.348%)
Validation starts
Test: [0/79]    Time 0.194 (0.194)      Loss 0.3138 (0.3138)    Prec 89.844%
(89.844%)
 * Prec 84.310%
best acc: 84.470000
```

```
Epoch: [31][0/391]      Time 0.312 (0.312)      Data 0.263 (0.263)      Loss
0.1730 (0.1730)    Prec 93.750% (93.750%)
Epoch: [31][100/391]    Time 0.056 (0.058)      Data 0.001 (0.004)      Loss
0.2912 (0.2980)    Prec 90.625% (90.231%)
Epoch: [31][200/391]    Time 0.066 (0.057)      Data 0.002 (0.003)      Loss
0.3191 (0.3014)    Prec 91.406% (89.887%)
Epoch: [31][300/391]    Time 0.053 (0.057)      Data 0.002 (0.003)      Loss
0.4346 (0.3064)    Prec 86.719% (89.693%)
Validation starts
Test: [0/79]    Time 0.247 (0.247)      Loss 0.3946 (0.3946)    Prec 87.500%
(87.500%)
 * Prec 84.860%
best acc: 84.860000
Epoch: [32][0/391]      Time 0.305 (0.305)      Data 0.261 (0.261)      Loss
0.3022 (0.3022)    Prec 90.625% (90.625%)
Epoch: [32][100/391]    Time 0.056 (0.058)      Data 0.002 (0.004)      Loss
0.3824 (0.2789)    Prec 89.062% (90.416%)
Epoch: [32][200/391]    Time 0.052 (0.057)      Data 0.002 (0.003)      Loss
0.1973 (0.2926)    Prec 93.750% (90.108%)
Epoch: [32][300/391]    Time 0.052 (0.056)      Data 0.001 (0.002)      Loss
0.3149 (0.2921)    Prec 89.062% (90.090%)
Validation starts
Test: [0/79]    Time 0.248 (0.248)      Loss 0.3584 (0.3584)    Prec 85.156%
(85.156%)
 * Prec 85.450%
best acc: 85.450000
Epoch: [33][0/391]      Time 0.291 (0.291)      Data 0.242 (0.242)      Loss
0.2092 (0.2092)    Prec 93.750% (93.750%)
Epoch: [33][100/391]    Time 0.054 (0.058)      Data 0.001 (0.004)      Loss
0.1712 (0.2726)    Prec 91.406% (90.888%)
Epoch: [33][200/391]    Time 0.052 (0.057)      Data 0.002 (0.003)      Loss
0.2241 (0.2774)    Prec 92.188% (90.641%)
Epoch: [33][300/391]    Time 0.049 (0.056)      Data 0.002 (0.002)      Loss
0.1724 (0.2817)    Prec 93.750% (90.436%)
Validation starts
Test: [0/79]    Time 0.202 (0.202)      Loss 0.3363 (0.3363)    Prec 92.188%
(92.188%)
 * Prec 85.690%
best acc: 85.690000
Epoch: [34][0/391]      Time 0.337 (0.337)      Data 0.295 (0.295)      Loss
0.2481 (0.2481)    Prec 92.969% (92.969%)
Epoch: [34][100/391]    Time 0.050 (0.058)      Data 0.001 (0.005)      Loss
0.1935 (0.2614)    Prec 93.750% (91.074%)
Epoch: [34][200/391]    Time 0.052 (0.057)      Data 0.001 (0.003)      Loss
0.3491 (0.2689)    Prec 86.719% (90.765%)
Epoch: [34][300/391]    Time 0.056 (0.056)      Data 0.002 (0.003)      Loss
0.3100 (0.2692)    Prec 90.625% (90.859%)
Validation starts
```

```
Test: [0/79]    Time 0.214 (0.214)      Loss 0.3819 (0.3819)    Prec 89.062%
(89.062%)
 * Prec 86.120%
best acc: 86.120000
Epoch: [35][0/391]      Time 0.278 (0.278)      Data 0.230 (0.230)      Loss
0.2992 (0.2992)    Prec 88.281% (88.281%)
Epoch: [35][100/391]    Time 0.055 (0.058)      Data 0.001 (0.004)      Loss
0.1528 (0.2572)    Prec 93.750% (91.166%)
Epoch: [35][200/391]    Time 0.056 (0.057)      Data 0.002 (0.003)      Loss
0.2080 (0.2714)    Prec 91.406% (90.769%)
Epoch: [35][300/391]    Time 0.054 (0.056)      Data 0.001 (0.002)      Loss
0.2866 (0.2746)    Prec 90.625% (90.690%)
Validation starts
Test: [0/79]    Time 0.233 (0.233)      Loss 0.3601 (0.3601)    Prec 89.062%
(89.062%)
 * Prec 85.940%
best acc: 86.120000
Epoch: [36][0/391]      Time 0.290 (0.290)      Data 0.247 (0.247)      Loss
0.2230 (0.2230)    Prec 90.625% (90.625%)
Epoch: [36][100/391]    Time 0.053 (0.058)      Data 0.002 (0.004)      Loss
0.2726 (0.2411)    Prec 92.969% (91.770%)
Epoch: [36][200/391]    Time 0.056 (0.057)      Data 0.001 (0.003)      Loss
0.2912 (0.2562)    Prec 90.625% (91.212%)
Epoch: [36][300/391]    Time 0.059 (0.056)      Data 0.002 (0.002)      Loss
0.3717 (0.2599)    Prec 85.938% (91.040%)
Validation starts
Test: [0/79]    Time 0.211 (0.211)      Loss 0.2782 (0.2782)    Prec 88.281%
(88.281%)
 * Prec 86.490%
best acc: 86.490000
Epoch: [37][0/391]      Time 0.288 (0.288)      Data 0.237 (0.237)      Loss
0.2222 (0.2222)    Prec 92.969% (92.969%)
Epoch: [37][100/391]    Time 0.057 (0.058)      Data 0.004 (0.004)      Loss
0.2642 (0.2360)    Prec 90.625% (92.071%)
Epoch: [37][200/391]    Time 0.057 (0.057)      Data 0.001 (0.003)      Loss
0.2132 (0.2437)    Prec 91.406% (91.810%)
Epoch: [37][300/391]    Time 0.058 (0.057)      Data 0.001 (0.003)      Loss
0.1908 (0.2499)    Prec 90.625% (91.549%)
Validation starts
Test: [0/79]    Time 0.192 (0.192)      Loss 0.3635 (0.3635)    Prec 85.938%
(85.938%)
 * Prec 85.880%
best acc: 86.490000
Epoch: [38][0/391]      Time 0.280 (0.280)      Data 0.241 (0.241)      Loss
0.3278 (0.3278)    Prec 88.281% (88.281%)
Epoch: [38][100/391]    Time 0.052 (0.058)      Data 0.002 (0.004)      Loss
0.2117 (0.2395)    Prec 90.625% (91.491%)
Epoch: [38][200/391]    Time 0.061 (0.057)      Data 0.001 (0.003)      Loss
```

```
0.2350 (0.2413)    Prec 92.188% (91.531%)
Epoch: [38][300/391]    Time 0.053 (0.056)      Data 0.002 (0.002)      Loss
0.2857 (0.2474)    Prec 88.281% (91.409%)
Validation starts
Test: [0/79]    Time 0.217 (0.217)      Loss 0.4327 (0.4327)    Prec 85.938%
(85.938%)
 * Prec 85.200%
best acc: 86.490000
Epoch: [39][0/391]      Time 0.272 (0.272)      Data 0.224 (0.224)      Loss
0.2212 (0.2212)    Prec 93.750% (93.750%)
Epoch: [39][100/391]    Time 0.055 (0.058)      Data 0.001 (0.004)      Loss
0.2369 (0.2385)    Prec 87.500% (91.754%)
Epoch: [39][200/391]    Time 0.053 (0.057)      Data 0.002 (0.003)      Loss
0.1847 (0.2421)    Prec 91.406% (91.717%)
Epoch: [39][300/391]    Time 0.060 (0.056)      Data 0.003 (0.002)      Loss
0.2771 (0.2427)    Prec 87.500% (91.616%)
Validation starts
Test: [0/79]    Time 0.267 (0.267)      Loss 0.3195 (0.3195)    Prec 90.625%
(90.625%)
 * Prec 87.090%
best acc: 87.090000
Epoch: [40][0/391]      Time 0.285 (0.285)      Data 0.236 (0.236)      Loss
0.2626 (0.2626)    Prec 91.406% (91.406%)
Epoch: [40][100/391]    Time 0.053 (0.058)      Data 0.001 (0.004)      Loss
0.1840 (0.2239)    Prec 92.188% (92.381%)
Epoch: [40][200/391]    Time 0.053 (0.056)      Data 0.002 (0.003)      Loss
0.2696 (0.2254)    Prec 89.062% (92.234%)
Epoch: [40][300/391]    Time 0.058 (0.056)      Data 0.001 (0.002)      Loss
0.1695 (0.2346)    Prec 94.531% (91.972%)
Validation starts
Test: [0/79]    Time 0.232 (0.232)      Loss 0.3097 (0.3097)    Prec 89.062%
(89.062%)
 * Prec 83.610%
best acc: 87.090000
Epoch: [41][0/391]      Time 0.275 (0.275)      Data 0.228 (0.228)      Loss
0.2205 (0.2205)    Prec 93.750% (93.750%)
Epoch: [41][100/391]    Time 0.056 (0.058)      Data 0.002 (0.004)      Loss
0.1915 (0.2193)    Prec 95.312% (92.327%)
Epoch: [41][200/391]    Time 0.059 (0.057)      Data 0.002 (0.003)      Loss
0.2071 (0.2240)    Prec 92.188% (92.339%)
Epoch: [41][300/391]    Time 0.056 (0.056)      Data 0.003 (0.002)      Loss
0.3847 (0.2297)    Prec 88.281% (92.190%)
Validation starts
Test: [0/79]    Time 0.231 (0.231)      Loss 0.3599 (0.3599)    Prec 88.281%
(88.281%)
 * Prec 85.990%
best acc: 87.090000
Epoch: [42][0/391]      Time 0.296 (0.296)      Data 0.244 (0.244)      Loss
```

```
0.1269 (0.1269)    Prec 96.094% (96.094%)
Epoch: [42][100/391]    Time 0.055 (0.058)      Data 0.002 (0.004)      Loss
0.2062 (0.2405)    Prec 92.969% (91.793%)
Epoch: [42][200/391]    Time 0.056 (0.057)      Data 0.002 (0.003)      Loss
0.2028 (0.2302)    Prec 92.969% (92.137%)
Epoch: [42][300/391]    Time 0.055 (0.056)      Data 0.002 (0.002)      Loss
0.3650 (0.2297)    Prec 87.500% (92.172%)
Validation starts
Test: [0/79]    Time 0.217 (0.217)      Loss 0.3222 (0.3222)    Prec 87.500%
(87.500%)
 * Prec 86.280%
best acc: 87.090000
Epoch: [43][0/391]      Time 0.304 (0.304)      Data 0.254 (0.254)      Loss
0.2030 (0.2030)    Prec 92.188% (92.188%)
Epoch: [43][100/391]    Time 0.056 (0.058)      Data 0.001 (0.004)      Loss
0.1199 (0.2050)    Prec 93.750% (92.922%)
Epoch: [43][200/391]    Time 0.053 (0.057)      Data 0.002 (0.003)      Loss
0.1501 (0.2114)    Prec 95.312% (92.716%)
Epoch: [43][300/391]    Time 0.055 (0.057)      Data 0.002 (0.003)      Loss
0.1489 (0.2140)    Prec 96.875% (92.683%)
Validation starts
Test: [0/79]    Time 0.242 (0.242)      Loss 0.2368 (0.2368)    Prec 90.625%
(90.625%)
 * Prec 87.210%
best acc: 87.210000
Epoch: [44][0/391]      Time 0.303 (0.303)      Data 0.252 (0.252)      Loss
0.1307 (0.1307)    Prec 94.531% (94.531%)
Epoch: [44][100/391]    Time 0.052 (0.058)      Data 0.004 (0.004)      Loss
0.1698 (0.2035)    Prec 95.312% (92.969%)
Epoch: [44][200/391]    Time 0.060 (0.057)      Data 0.002 (0.003)      Loss
0.1628 (0.2137)    Prec 95.312% (92.662%)
Epoch: [44][300/391]    Time 0.055 (0.056)      Data 0.002 (0.003)      Loss
0.2380 (0.2156)    Prec 92.188% (92.694%)
Validation starts
Test: [0/79]    Time 0.236 (0.236)      Loss 0.3560 (0.3560)    Prec 89.844%
(89.844%)
 * Prec 86.300%
best acc: 87.210000
Epoch: [45][0/391]      Time 0.285 (0.285)      Data 0.239 (0.239)      Loss
0.1260 (0.1260)    Prec 95.312% (95.312%)
Epoch: [45][100/391]    Time 0.070 (0.058)      Data 0.001 (0.004)      Loss
0.1973 (0.1900)    Prec 94.531% (93.603%)
Epoch: [45][200/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.1932 (0.1964)    Prec 93.750% (93.256%)
Epoch: [45][300/391]    Time 0.060 (0.057)      Data 0.002 (0.003)      Loss
0.2032 (0.2009)    Prec 93.750% (93.080%)
Validation starts
Test: [0/79]    Time 0.223 (0.223)      Loss 0.3133 (0.3133)    Prec 90.625%
```

```
(90.625%)
 * Prec 86.920%
best acc: 87.210000
Epoch: [46][0/391]      Time 0.265 (0.265)      Data 0.218 (0.218)      Loss
0.1727 (0.1727)    Prec 94.531% (94.531%)
Epoch: [46][100/391]    Time 0.051 (0.057)      Data 0.002 (0.004)      Loss
0.1823 (0.1945)    Prec 92.969% (93.394%)
Epoch: [46][200/391]    Time 0.061 (0.057)      Data 0.001 (0.003)      Loss
0.1787 (0.1937)    Prec 92.188% (93.342%)
Epoch: [46][300/391]    Time 0.064 (0.056)      Data 0.001 (0.002)      Loss
0.2201 (0.2019)    Prec 89.844% (93.062%)
Validation starts
Test: [0/79]    Time 0.245 (0.245)      Loss 0.3573 (0.3573)    Prec 87.500%
(87.500%)
 * Prec 85.060%
best acc: 87.210000
Epoch: [47][0/391]      Time 0.289 (0.289)      Data 0.242 (0.242)      Loss
0.2177 (0.2177)    Prec 91.406% (91.406%)
Epoch: [47][100/391]    Time 0.056 (0.058)      Data 0.001 (0.004)      Loss
0.1460 (0.1930)    Prec 94.531% (93.557%)
Epoch: [47][200/391]    Time 0.056 (0.057)      Data 0.001 (0.003)      Loss
0.1566 (0.1956)    Prec 92.969% (93.369%)
Epoch: [47][300/391]    Time 0.054 (0.056)      Data 0.002 (0.002)      Loss
0.2015 (0.1987)    Prec 92.188% (93.223%)
Validation starts
Test: [0/79]    Time 0.216 (0.216)      Loss 0.3023 (0.3023)    Prec 90.625%
(90.625%)
 * Prec 86.790%
best acc: 87.210000
Epoch: [48][0/391]      Time 0.315 (0.315)      Data 0.264 (0.264)      Loss
0.1537 (0.1537)    Prec 94.531% (94.531%)
Epoch: [48][100/391]    Time 0.053 (0.058)      Data 0.001 (0.004)      Loss
0.2074 (0.1799)    Prec 92.969% (94.013%)
Epoch: [48][200/391]    Time 0.055 (0.057)      Data 0.002 (0.003)      Loss
0.1530 (0.1906)    Prec 96.094% (93.560%)
Epoch: [48][300/391]    Time 0.057 (0.056)      Data 0.001 (0.002)      Loss
0.1977 (0.1940)    Prec 92.969% (93.355%)
Validation starts
Test: [0/79]    Time 0.223 (0.223)      Loss 0.2220 (0.2220)    Prec 91.406%
(91.406%)
 * Prec 88.140%
best acc: 88.140000
Epoch: [49][0/391]      Time 0.268 (0.268)      Data 0.221 (0.221)      Loss
0.1763 (0.1763)    Prec 93.750% (93.750%)
Epoch: [49][100/391]    Time 0.055 (0.058)      Data 0.001 (0.004)      Loss
0.1851 (0.1815)    Prec 93.750% (93.742%)
Epoch: [49][200/391]    Time 0.052 (0.057)      Data 0.002 (0.003)      Loss
0.1861 (0.1835)    Prec 92.969% (93.696%)
```

```
Epoch: [49][300/391]    Time 0.054 (0.056)      Data 0.001 (0.002)      Loss
0.1506 (0.1893)    Prec 96.094% (93.550%)
Validation starts
Test: [0/79]    Time 0.228 (0.228)      Loss 0.1614 (0.1614)    Prec 94.531%
(94.531%)
 * Prec 87.540%
best acc: 88.140000
Epoch: [50][0/391]      Time 0.268 (0.268)      Data 0.220 (0.220)      Loss
0.0797 (0.0797)    Prec 97.656% (97.656%)
Epoch: [50][100/391]    Time 0.056 (0.057)      Data 0.002 (0.004)      Loss
0.1913 (0.1729)    Prec 92.969% (93.990%)
Epoch: [50][200/391]    Time 0.056 (0.056)      Data 0.003 (0.003)      Loss
0.1338 (0.1789)    Prec 94.531% (93.766%)
Epoch: [50][300/391]    Time 0.056 (0.056)      Data 0.001 (0.002)      Loss
0.1505 (0.1825)    Prec 96.094% (93.695%)
Validation starts
Test: [0/79]    Time 0.246 (0.246)      Loss 0.2810 (0.2810)    Prec 89.844%
(89.844%)
 * Prec 87.320%
best acc: 88.140000
Epoch: [51][0/391]      Time 0.273 (0.273)      Data 0.232 (0.232)      Loss
0.1065 (0.1065)    Prec 96.875% (96.875%)
Epoch: [51][100/391]    Time 0.055 (0.058)      Data 0.002 (0.004)      Loss
0.1885 (0.1833)    Prec 93.750% (93.704%)
Epoch: [51][200/391]    Time 0.054 (0.057)      Data 0.001 (0.003)      Loss
0.2415 (0.1845)    Prec 92.188% (93.731%)
Epoch: [51][300/391]    Time 0.057 (0.056)      Data 0.001 (0.002)      Loss
0.1429 (0.1815)    Prec 96.875% (93.877%)
Validation starts
Test: [0/79]    Time 0.234 (0.234)      Loss 0.3303 (0.3303)    Prec 90.625%
(90.625%)
 * Prec 86.120%
best acc: 88.140000
Epoch: [52][0/391]      Time 0.283 (0.283)      Data 0.242 (0.242)      Loss
0.1348 (0.1348)    Prec 95.312% (95.312%)
Epoch: [52][100/391]    Time 0.071 (0.058)      Data 0.001 (0.004)      Loss
0.1662 (0.1658)    Prec 92.188% (94.330%)
Epoch: [52][200/391]    Time 0.061 (0.057)      Data 0.002 (0.003)      Loss
0.1964 (0.1704)    Prec 92.969% (94.189%)
Epoch: [52][300/391]    Time 0.063 (0.056)      Data 0.002 (0.002)      Loss
0.1225 (0.1701)    Prec 96.094% (94.204%)
Validation starts
Test: [0/79]    Time 0.226 (0.226)      Loss 0.4285 (0.4285)    Prec 86.719%
(86.719%)
 * Prec 86.560%
best acc: 88.140000
Epoch: [53][0/391]      Time 0.278 (0.278)      Data 0.227 (0.227)      Loss
0.1629 (0.1629)    Prec 94.531% (94.531%)
```

```
Epoch: [53][100/391]    Time 0.052 (0.057)      Data 0.004 (0.004)      Loss
0.1540 (0.1606)    Prec 96.094% (94.485%)
Epoch: [53][200/391]    Time 0.055 (0.056)      Data 0.002 (0.003)      Loss
0.3587 (0.1718)    Prec 87.500% (94.022%)
Epoch: [53][300/391]    Time 0.061 (0.056)      Data 0.002 (0.002)      Loss
0.3081 (0.1775)    Prec 90.625% (93.888%)
Validation starts
Test: [0/79]    Time 0.254 (0.254)      Loss 0.3578 (0.3578)    Prec 86.719%
(86.719%)
 * Prec 86.960%
best acc: 88.140000
Epoch: [54][0/391]      Time 0.273 (0.273)      Data 0.226 (0.226)      Loss
0.0926 (0.0926)    Prec 97.656% (97.656%)
Epoch: [54][100/391]    Time 0.063 (0.058)      Data 0.001 (0.004)      Loss
0.2125 (0.1630)    Prec 92.969% (94.268%)
Epoch: [54][200/391]    Time 0.065 (0.057)      Data 0.002 (0.003)      Loss
0.1723 (0.1712)    Prec 96.094% (94.045%)
Epoch: [54][300/391]    Time 0.055 (0.056)      Data 0.002 (0.002)      Loss
0.1594 (0.1727)    Prec 93.750% (94.064%)
Validation starts
Test: [0/79]    Time 0.227 (0.227)      Loss 0.4193 (0.4193)    Prec 85.938%
(85.938%)
 * Prec 86.690%
best acc: 88.140000
Epoch: [55][0/391]      Time 0.326 (0.326)      Data 0.286 (0.286)      Loss
0.1431 (0.1431)    Prec 96.094% (96.094%)
Epoch: [55][100/391]    Time 0.056 (0.058)      Data 0.001 (0.004)      Loss
0.1005 (0.1603)    Prec 96.875% (94.524%)
Epoch: [55][200/391]    Time 0.054 (0.057)      Data 0.001 (0.003)      Loss
0.2057 (0.1626)    Prec 92.188% (94.488%)
Epoch: [55][300/391]    Time 0.057 (0.056)      Data 0.002 (0.003)      Loss
0.2630 (0.1668)    Prec 91.406% (94.256%)
Validation starts
Test: [0/79]    Time 0.218 (0.218)      Loss 0.1924 (0.1924)    Prec 92.969%
(92.969%)
 * Prec 87.350%
best acc: 88.140000
Epoch: [56][0/391]      Time 0.274 (0.274)      Data 0.229 (0.229)      Loss
0.0521 (0.0521)    Prec 99.219% (99.219%)
Epoch: [56][100/391]    Time 0.057 (0.058)      Data 0.003 (0.004)      Loss
0.1541 (0.1514)    Prec 92.969% (94.771%)
Epoch: [56][200/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.2161 (0.1576)    Prec 89.844% (94.562%)
Epoch: [56][300/391]    Time 0.054 (0.057)      Data 0.001 (0.002)      Loss
0.2075 (0.1612)    Prec 93.750% (94.456%)
Validation starts
Test: [0/79]    Time 0.228 (0.228)      Loss 0.2591 (0.2591)    Prec 92.188%
(92.188%)
```

```
 * Prec 87.810%
best acc: 88.140000
Epoch: [57][0/391]      Time 0.273 (0.273)      Data 0.230 (0.230)      Loss
0.0767 (0.0767)    Prec 97.656% (97.656%)
Epoch: [57][100/391]    Time 0.054 (0.058)      Data 0.002 (0.004)      Loss
0.0927 (0.1479)    Prec 96.875% (95.088%)
Epoch: [57][200/391]    Time 0.059 (0.057)      Data 0.001 (0.003)      Loss
0.1660 (0.1518)    Prec 91.406% (94.939%)
Epoch: [57][300/391]    Time 0.053 (0.056)      Data 0.002 (0.002)      Loss
0.1710 (0.1571)    Prec 94.531% (94.723%)
Validation starts
Test: [0/79]    Time 0.213 (0.213)      Loss 0.1943 (0.1943)    Prec 92.969%
(92.969%)
 * Prec 87.510%
best acc: 88.140000
Epoch: [58][0/391]      Time 0.315 (0.315)      Data 0.267 (0.267)      Loss
0.0856 (0.0856)    Prec 96.094% (96.094%)
Epoch: [58][100/391]    Time 0.053 (0.058)      Data 0.002 (0.004)      Loss
0.2058 (0.1539)    Prec 94.531% (94.640%)
Epoch: [58][200/391]    Time 0.053 (0.057)      Data 0.002 (0.003)      Loss
0.1569 (0.1497)    Prec 91.406% (94.838%)
Epoch: [58][300/391]    Time 0.055 (0.056)      Data 0.002 (0.002)      Loss
0.2331 (0.1534)    Prec 90.625% (94.661%)
Validation starts
Test: [0/79]    Time 0.216 (0.216)      Loss 0.3154 (0.3154)    Prec 89.844%
(89.844%)
 * Prec 87.170%
best acc: 88.140000
Epoch: [59][0/391]      Time 0.272 (0.272)      Data 0.224 (0.224)      Loss
0.2159 (0.2159)    Prec 92.969% (92.969%)
Epoch: [59][100/391]    Time 0.055 (0.058)      Data 0.001 (0.004)      Loss
0.1545 (0.1559)    Prec 96.875% (94.609%)
Epoch: [59][200/391]    Time 0.054 (0.057)      Data 0.002 (0.003)      Loss
0.1323 (0.1508)    Prec 95.312% (94.881%)
Epoch: [59][300/391]    Time 0.057 (0.057)      Data 0.001 (0.002)      Loss
0.1391 (0.1587)    Prec 94.531% (94.562%)
Validation starts
Test: [0/79]    Time 0.345 (0.345)      Loss 0.3938 (0.3938)    Prec 88.281%
(88.281%)
 * Prec 87.670%
best acc: 88.140000
Epoch: [60][0/391]      Time 0.271 (0.271)      Data 0.225 (0.225)      Loss
0.2015 (0.2015)    Prec 92.969% (92.969%)
Epoch: [60][100/391]    Time 0.055 (0.057)      Data 0.001 (0.004)      Loss
0.1456 (0.1397)    Prec 94.531% (95.119%)
Epoch: [60][200/391]    Time 0.056 (0.056)      Data 0.002 (0.003)      Loss
0.1555 (0.1463)    Prec 95.312% (94.819%)
Epoch: [60][300/391]    Time 0.057 (0.057)      Data 0.002 (0.002)      Loss
```

```
0.1114 (0.1489)    Prec 96.875% (94.791%)
Validation starts
Test: [0/79]    Time 0.224 (0.224)      Loss 0.4133 (0.4133)    Prec 86.719%
(86.719%)
 * Prec 86.080%
best acc: 88.140000
Epoch: [61][0/391]      Time 0.299 (0.299)      Data 0.247 (0.247)      Loss
0.0937 (0.0937)    Prec 97.656% (97.656%)
Epoch: [61][100/391]    Time 0.059 (0.058)      Data 0.001 (0.004)      Loss
0.1269 (0.1482)    Prec 95.312% (94.964%)
Epoch: [61][200/391]    Time 0.056 (0.057)      Data 0.002 (0.003)      Loss
0.1110 (0.1423)    Prec 96.875% (95.145%)
Epoch: [61][300/391]    Time 0.056 (0.056)      Data 0.002 (0.002)      Loss
0.0617 (0.1459)    Prec 97.656% (95.043%)
Validation starts
Test: [0/79]    Time 0.229 (0.229)      Loss 0.2397 (0.2397)    Prec 90.625%
(90.625%)
 * Prec 87.100%
best acc: 88.140000
Epoch: [62][0/391]      Time 0.302 (0.302)      Data 0.253 (0.253)      Loss
0.1530 (0.1530)    Prec 95.312% (95.312%)
Epoch: [62][100/391]    Time 0.058 (0.058)      Data 0.002 (0.004)      Loss
0.1306 (0.1343)    Prec 96.875% (95.359%)
Epoch: [62][200/391]    Time 0.054 (0.058)      Data 0.002 (0.003)      Loss
0.0606 (0.1377)    Prec 97.656% (95.285%)
Epoch: [62][300/391]    Time 0.058 (0.057)      Data 0.002 (0.003)      Loss
0.1315 (0.1432)    Prec 96.094% (95.063%)
Validation starts
Test: [0/79]    Time 0.222 (0.222)      Loss 0.3128 (0.3128)    Prec 88.281%
(88.281%)
 * Prec 87.650%
best acc: 88.140000
Epoch: [63][0/391]      Time 0.273 (0.273)      Data 0.228 (0.228)      Loss
0.1102 (0.1102)    Prec 94.531% (94.531%)
Epoch: [63][100/391]    Time 0.055 (0.057)      Data 0.002 (0.004)      Loss
0.0728 (0.1348)    Prec 96.094% (95.452%)
Epoch: [63][200/391]    Time 0.057 (0.057)      Data 0.001 (0.003)      Loss
0.2077 (0.1408)    Prec 92.969% (95.266%)
Epoch: [63][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0969 (0.1442)    Prec 96.875% (95.105%)
Validation starts
Test: [0/79]    Time 0.273 (0.273)      Loss 0.4675 (0.4675)    Prec 85.938%
(85.938%)
 * Prec 85.720%
best acc: 88.140000
Epoch: [64][0/391]      Time 0.294 (0.294)      Data 0.247 (0.247)      Loss
0.1204 (0.1204)    Prec 95.312% (95.312%)
Epoch: [64][100/391]    Time 0.057 (0.058)      Data 0.002 (0.004)      Loss
```

```
0.2016 (0.1250)    Prec 92.188% (95.614%)
Epoch: [64][200/391]    Time 0.059 (0.057)      Data 0.002 (0.003)      Loss
0.0938 (0.1312)    Prec 96.094% (95.487%)
Epoch: [64][300/391]    Time 0.056 (0.057)      Data 0.001 (0.002)      Loss
0.1890 (0.1352)    Prec 95.312% (95.380%)
Validation starts
Test: [0/79]    Time 0.228 (0.228)      Loss 0.3814 (0.3814)    Prec 86.719%
(86.719%)
 * Prec 87.950%
best acc: 88.140000
Epoch: [65][0/391]      Time 0.287 (0.287)      Data 0.239 (0.239)      Loss
0.0954 (0.0954)    Prec 96.875% (96.875%)
Epoch: [65][100/391]    Time 0.050 (0.058)      Data 0.001 (0.004)      Loss
0.2069 (0.1393)    Prec 92.188% (95.336%)
Epoch: [65][200/391]    Time 0.055 (0.056)      Data 0.002 (0.003)      Loss
0.1180 (0.1449)    Prec 96.094% (95.126%)
Epoch: [65][300/391]    Time 0.050 (0.056)      Data 0.002 (0.002)      Loss
0.1513 (0.1434)    Prec 96.094% (95.128%)
Validation starts
Test: [0/79]    Time 0.236 (0.236)      Loss 0.3535 (0.3535)    Prec 88.281%
(88.281%)
 * Prec 86.800%
best acc: 88.140000
Epoch: [66][0/391]      Time 0.297 (0.297)      Data 0.250 (0.250)      Loss
0.1536 (0.1536)    Prec 92.969% (92.969%)
Epoch: [66][100/391]    Time 0.055 (0.057)      Data 0.001 (0.004)      Loss
0.1921 (0.1328)    Prec 93.750% (95.521%)
Epoch: [66][200/391]    Time 0.054 (0.056)      Data 0.002 (0.003)      Loss
0.1971 (0.1329)    Prec 92.188% (95.491%)
Epoch: [66][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.1569 (0.1339)    Prec 95.312% (95.460%)
Validation starts
Test: [0/79]    Time 0.226 (0.226)      Loss 0.3051 (0.3051)    Prec 89.844%
(89.844%)
 * Prec 87.390%
best acc: 88.140000
Epoch: [67][0/391]      Time 0.298 (0.298)      Data 0.243 (0.243)      Loss
0.1145 (0.1145)    Prec 96.094% (96.094%)
Epoch: [67][100/391]    Time 0.060 (0.058)      Data 0.001 (0.004)      Loss
0.1130 (0.1298)    Prec 96.875% (95.699%)
Epoch: [67][200/391]    Time 0.050 (0.057)      Data 0.002 (0.003)      Loss
0.1081 (0.1325)    Prec 96.094% (95.600%)
Epoch: [67][300/391]    Time 0.052 (0.056)      Data 0.002 (0.002)      Loss
0.1847 (0.1333)    Prec 90.625% (95.523%)
Validation starts
Test: [0/79]    Time 0.213 (0.213)      Loss 0.3055 (0.3055)    Prec 92.188%
(92.188%)
 * Prec 88.740%
```

```
best acc: 88.740000
Epoch: [68][0/391]      Time 0.298 (0.298)      Data 0.252 (0.252)      Loss
0.1845 (0.1845)    Prec 93.750% (93.750%)
Epoch: [68][100/391]    Time 0.055 (0.058)      Data 0.002 (0.004)      Loss
0.1128 (0.1289)    Prec 96.094% (95.738%)
Epoch: [68][200/391]    Time 0.064 (0.057)      Data 0.002 (0.003)      Loss
0.1257 (0.1316)    Prec 94.531% (95.499%)
Epoch: [68][300/391]    Time 0.059 (0.057)      Data 0.001 (0.003)      Loss
0.1046 (0.1349)    Prec 96.875% (95.466%)
Validation starts
Test: [0/79]    Time 0.254 (0.254)      Loss 0.4792 (0.4792)    Prec 88.281%
(88.281%)
 * Prec 87.100%
best acc: 88.740000
Epoch: [69][0/391]      Time 0.269 (0.269)      Data 0.224 (0.224)      Loss
0.1117 (0.1117)    Prec 96.094% (96.094%)
Epoch: [69][100/391]    Time 0.053 (0.058)      Data 0.002 (0.004)      Loss
0.1306 (0.1205)    Prec 92.188% (95.815%)
Epoch: [69][200/391]    Time 0.063 (0.057)      Data 0.001 (0.003)      Loss
0.1340 (0.1274)    Prec 96.875% (95.674%)
Epoch: [69][300/391]    Time 0.059 (0.056)      Data 0.002 (0.002)      Loss
0.1431 (0.1314)    Prec 95.312% (95.562%)
Validation starts
Test: [0/79]    Time 0.240 (0.240)      Loss 0.2813 (0.2813)    Prec 92.969%
(92.969%)
 * Prec 88.300%
best acc: 88.740000
Epoch: [70][0/391]      Time 0.242 (0.242)      Data 0.193 (0.193)      Loss
0.1011 (0.1011)    Prec 94.531% (94.531%)
Epoch: [70][100/391]    Time 0.055 (0.058)      Data 0.001 (0.004)      Loss
0.1480 (0.1288)    Prec 96.094% (95.552%)
Epoch: [70][200/391]    Time 0.051 (0.057)      Data 0.003 (0.003)      Loss
0.0880 (0.1282)    Prec 96.875% (95.678%)
Epoch: [70][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.1368 (0.1288)    Prec 95.312% (95.671%)
Validation starts
Test: [0/79]    Time 0.235 (0.235)      Loss 0.3531 (0.3531)    Prec 90.625%
(90.625%)
 * Prec 86.350%
best acc: 88.740000
Epoch: [71][0/391]      Time 0.273 (0.273)      Data 0.229 (0.229)      Loss
0.0895 (0.0895)    Prec 97.656% (97.656%)
Epoch: [71][100/391]    Time 0.065 (0.058)      Data 0.002 (0.004)      Loss
0.1004 (0.1211)    Prec 95.312% (95.777%)
Epoch: [71][200/391]    Time 0.049 (0.056)      Data 0.003 (0.003)      Loss
0.1453 (0.1299)    Prec 96.875% (95.557%)
Epoch: [71][300/391]    Time 0.056 (0.056)      Data 0.001 (0.002)      Loss
0.1687 (0.1296)    Prec 96.094% (95.546%)
```

```
Validation starts
Test: [0/79]    Time 0.236 (0.236)      Loss 0.3855 (0.3855)    Prec 87.500%
(87.500%)
 * Prec 87.650%
best acc: 88.740000
Epoch: [72][0/391]      Time 0.296 (0.296)      Data 0.246 (0.246)      Loss
0.0877 (0.0877)    Prec 98.438% (98.438%)
Epoch: [72][100/391]    Time 0.059 (0.058)      Data 0.002 (0.004)      Loss
0.0990 (0.1290)    Prec 95.312% (95.637%)
Epoch: [72][200/391]    Time 0.058 (0.057)      Data 0.002 (0.003)      Loss
0.2168 (0.1301)    Prec 93.750% (95.623%)
Epoch: [72][300/391]    Time 0.052 (0.057)      Data 0.001 (0.003)      Loss
0.1653 (0.1304)    Prec 94.531% (95.580%)
Validation starts
Test: [0/79]    Time 0.221 (0.221)      Loss 0.2881 (0.2881)    Prec 90.625%
(90.625%)
 * Prec 86.800%
best acc: 88.740000
Epoch: [73][0/391]      Time 0.272 (0.272)      Data 0.222 (0.222)      Loss
0.1463 (0.1463)    Prec 93.750% (93.750%)
Epoch: [73][100/391]    Time 0.060 (0.058)      Data 0.001 (0.004)      Loss
0.1619 (0.1201)    Prec 92.969% (95.823%)
Epoch: [73][200/391]    Time 0.054 (0.057)      Data 0.002 (0.003)      Loss
0.1411 (0.1224)    Prec 96.094% (95.771%)
Epoch: [73][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.1530 (0.1224)    Prec 94.531% (95.790%)
Validation starts
Test: [0/79]    Time 0.209 (0.209)      Loss 0.3841 (0.3841)    Prec 89.062%
(89.062%)
 * Prec 87.860%
best acc: 88.740000
Epoch: [74][0/391]      Time 0.283 (0.283)      Data 0.242 (0.242)      Loss
0.0445 (0.0445)    Prec 99.219% (99.219%)
Epoch: [74][100/391]    Time 0.053 (0.058)      Data 0.001 (0.004)      Loss
0.0772 (0.1149)    Prec 97.656% (96.009%)
Epoch: [74][200/391]    Time 0.057 (0.057)      Data 0.001 (0.003)      Loss
0.1759 (0.1196)    Prec 95.312% (95.911%)
Epoch: [74][300/391]    Time 0.056 (0.056)      Data 0.001 (0.002)      Loss
0.0605 (0.1270)    Prec 98.438% (95.665%)
Validation starts
Test: [0/79]    Time 0.221 (0.221)      Loss 0.2868 (0.2868)    Prec 91.406%
(91.406%)
 * Prec 88.270%
best acc: 88.740000
Epoch: [75][0/391]      Time 0.260 (0.260)      Data 0.213 (0.213)      Loss
0.1210 (0.1210)    Prec 96.094% (96.094%)
Epoch: [75][100/391]    Time 0.060 (0.057)      Data 0.002 (0.004)      Loss
0.1328 (0.1046)    Prec 96.875% (96.573%)
```

```
Epoch: [75][200/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.1259 (0.1171)    Prec 93.750% (96.105%)
Epoch: [75][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.1329 (0.1221)    Prec 93.750% (95.847%)
Validation starts
Test: [0/79]    Time 0.226 (0.226)      Loss 0.1898 (0.1898)    Prec 90.625%
(90.625%)
 * Prec 88.470%
best acc: 88.740000
Epoch: [76][0/391]      Time 0.269 (0.269)      Data 0.221 (0.221)      Loss
0.0483 (0.0483)    Prec 97.656% (97.656%)
Epoch: [76][100/391]    Time 0.055 (0.058)      Data 0.002 (0.004)      Loss
0.0898 (0.1089)    Prec 96.875% (96.310%)
Epoch: [76][200/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.0978 (0.1118)    Prec 98.438% (96.226%)
Epoch: [76][300/391]    Time 0.056 (0.056)      Data 0.001 (0.002)      Loss
0.0760 (0.1169)    Prec 96.875% (96.140%)
Validation starts
Test: [0/79]    Time 0.233 (0.233)      Loss 0.3449 (0.3449)    Prec 92.188%
(92.188%)
 * Prec 86.700%
best acc: 88.740000
Epoch: [77][0/391]      Time 0.279 (0.279)      Data 0.232 (0.232)      Loss
0.1696 (0.1696)    Prec 94.531% (94.531%)
Epoch: [77][100/391]    Time 0.053 (0.057)      Data 0.001 (0.004)      Loss
0.1624 (0.1200)    Prec 93.750% (96.040%)
Epoch: [77][200/391]    Time 0.056 (0.056)      Data 0.001 (0.003)      Loss
0.1407 (0.1219)    Prec 95.312% (95.927%)
Epoch: [77][300/391]    Time 0.053 (0.056)      Data 0.001 (0.002)      Loss
0.1175 (0.1209)    Prec 95.312% (95.948%)
Validation starts
Test: [0/79]    Time 0.252 (0.252)      Loss 0.1998 (0.1998)    Prec 92.188%
(92.188%)
 * Prec 86.050%
best acc: 88.740000
Epoch: [78][0/391]      Time 0.281 (0.281)      Data 0.232 (0.232)      Loss
0.1439 (0.1439)    Prec 96.094% (96.094%)
Epoch: [78][100/391]    Time 0.054 (0.057)      Data 0.001 (0.004)      Loss
0.1353 (0.1088)    Prec 93.750% (96.457%)
Epoch: [78][200/391]    Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.1391 (0.1134)    Prec 94.531% (96.191%)
Epoch: [78][300/391]    Time 0.055 (0.056)      Data 0.002 (0.002)      Loss
0.1496 (0.1167)    Prec 95.312% (96.065%)
Validation starts
Test: [0/79]    Time 0.221 (0.221)      Loss 0.2815 (0.2815)    Prec 92.188%
(92.188%)
 * Prec 87.920%
best acc: 88.740000
```

```
Epoch: [79][0/391]      Time 0.292 (0.292)      Data 0.247 (0.247)      Loss
0.1242 (0.1242)    Prec 94.531% (94.531%)
Epoch: [79][100/391]    Time 0.056 (0.057)      Data 0.002 (0.004)      Loss
0.1523 (0.1077)    Prec 96.094% (96.395%)
Epoch: [79][200/391]    Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.0771 (0.1132)    Prec 97.656% (96.140%)
Epoch: [79][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0680 (0.1133)    Prec 97.656% (96.078%)
Validation starts
Test: [0/79]    Time 0.239 (0.239)      Loss 0.3236 (0.3236)    Prec 89.844%
(89.844%)
 * Prec 86.900%
best acc: 88.740000
Epoch: [80][0/391]      Time 0.281 (0.281)      Data 0.234 (0.234)      Loss
0.0890 (0.0890)    Prec 96.875% (96.875%)
Epoch: [80][100/391]    Time 0.057 (0.058)      Data 0.001 (0.004)      Loss
0.0697 (0.0730)    Prec 96.875% (97.703%)
Epoch: [80][200/391]    Time 0.061 (0.057)      Data 0.001 (0.003)      Loss
0.1069 (0.0639)    Prec 95.312% (97.913%)
Epoch: [80][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0371 (0.0603)    Prec 99.219% (98.051%)
Validation starts
Test: [0/79]    Time 0.214 (0.214)      Loss 0.2292 (0.2292)    Prec 92.969%
(92.969%)
 * Prec 90.920%
best acc: 90.920000
Epoch: [81][0/391]      Time 0.286 (0.286)      Data 0.239 (0.239)      Loss
0.0844 (0.0844)    Prec 97.656% (97.656%)
Epoch: [81][100/391]    Time 0.055 (0.057)      Data 0.002 (0.004)      Loss
0.0196 (0.0422)    Prec 99.219% (98.708%)
Epoch: [81][200/391]    Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.0081 (0.0402)    Prec 100.000% (98.748%)
Epoch: [81][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0852 (0.0408)    Prec 96.875% (98.718%)
Validation starts
Test: [0/79]    Time 0.259 (0.259)      Loss 0.1943 (0.1943)    Prec 93.750%
(93.750%)
 * Prec 91.280%
best acc: 91.280000
Epoch: [82][0/391]      Time 0.290 (0.290)      Data 0.241 (0.241)      Loss
0.0227 (0.0227)    Prec 100.000% (100.000%)
Epoch: [82][100/391]    Time 0.054 (0.058)      Data 0.001 (0.004)      Loss
0.0562 (0.0347)    Prec 98.438% (99.002%)
Epoch: [82][200/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.0038 (0.0343)    Prec 100.000% (98.982%)
Epoch: [82][300/391]    Time 0.054 (0.056)      Data 0.001 (0.002)      Loss
0.0214 (0.0349)    Prec 99.219% (98.933%)
Validation starts
```

```
Test: [0/79]    Time 0.233 (0.233)      Loss 0.1684 (0.1684)    Prec 93.750%
(93.750%)
 * Prec 90.880%
best acc: 91.280000
Epoch: [83][0/391]      Time 0.290 (0.290)      Data 0.238 (0.238)      Loss
0.0581 (0.0581)    Prec 96.875% (96.875%)
Epoch: [83][100/391]    Time 0.054 (0.058)      Data 0.001 (0.004)      Loss
0.0353 (0.0324)    Prec 99.219% (98.917%)
Epoch: [83][200/391]    Time 0.054 (0.057)      Data 0.002 (0.003)      Loss
0.0206 (0.0321)    Prec 99.219% (98.927%)
Epoch: [83][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0123 (0.0317)    Prec 99.219% (98.941%)
Validation starts
Test: [0/79]    Time 0.208 (0.208)      Loss 0.2055 (0.2055)    Prec 94.531%
(94.531%)
 * Prec 91.110%
best acc: 91.280000
Epoch: [84][0/391]      Time 0.256 (0.256)      Data 0.211 (0.211)      Loss
0.0217 (0.0217)    Prec 99.219% (99.219%)
Epoch: [84][100/391]    Time 0.055 (0.057)      Data 0.001 (0.004)      Loss
0.0101 (0.0277)    Prec 100.000% (99.165%)
Epoch: [84][200/391]    Time 0.054 (0.056)      Data 0.001 (0.002)      Loss
0.0666 (0.0273)    Prec 97.656% (99.188%)
Epoch: [84][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0230 (0.0274)    Prec 99.219% (99.164%)
Validation starts
Test: [0/79]    Time 0.231 (0.231)      Loss 0.1926 (0.1926)    Prec 94.531%
(94.531%)
 * Prec 91.220%
best acc: 91.280000
Epoch: [85][0/391]      Time 0.266 (0.266)      Data 0.226 (0.226)      Loss
0.0138 (0.0138)    Prec 100.000% (100.000%)
Epoch: [85][100/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.0402 (0.0241)    Prec 99.219% (99.211%)
Epoch: [85][200/391]    Time 0.060 (0.056)      Data 0.001 (0.002)      Loss
0.0239 (0.0238)    Prec 99.219% (99.227%)
Epoch: [85][300/391]    Time 0.050 (0.056)      Data 0.001 (0.002)      Loss
0.0140 (0.0241)    Prec 99.219% (99.190%)
Validation starts
Test: [0/79]    Time 0.219 (0.219)      Loss 0.2357 (0.2357)    Prec 92.969%
(92.969%)
 * Prec 90.850%
best acc: 91.280000
Epoch: [86][0/391]      Time 0.267 (0.267)      Data 0.225 (0.225)      Loss
0.0101 (0.0101)    Prec 100.000% (100.000%)
Epoch: [86][100/391]    Time 0.058 (0.058)      Data 0.001 (0.004)      Loss
0.0053 (0.0228)    Prec 100.000% (99.281%)
Epoch: [86][200/391]    Time 0.055 (0.057)      Data 0.001 (0.002)      Loss
```

```
0.0083 (0.0214)    Prec 100.000% (99.308%)
Epoch: [86][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0046 (0.0213)    Prec 100.000% (99.289%)
Validation starts
Test: [0/79]    Time 0.233 (0.233)      Loss 0.1956 (0.1956)    Prec 93.750%
(93.750%)
 * Prec 91.320%
best acc: 91.320000
Epoch: [87][0/391]      Time 0.313 (0.313)      Data 0.270 (0.270)      Loss
0.0038 (0.0038)    Prec 100.000% (100.000%)
Epoch: [87][100/391]    Time 0.055 (0.057)      Data 0.001 (0.004)      Loss
0.0131 (0.0249)    Prec 99.219% (99.165%)
Epoch: [87][200/391]    Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.0126 (0.0235)    Prec 100.000% (99.219%)
Epoch: [87][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0129 (0.0229)    Prec 99.219% (99.252%)
Validation starts
Test: [0/79]    Time 0.216 (0.216)      Loss 0.1839 (0.1839)    Prec 93.750%
(93.750%)
 * Prec 91.040%
best acc: 91.320000
Epoch: [88][0/391]      Time 0.272 (0.272)      Data 0.232 (0.232)      Loss
0.0105 (0.0105)    Prec 100.000% (100.000%)
Epoch: [88][100/391]    Time 0.058 (0.057)      Data 0.001 (0.004)      Loss
0.0327 (0.0201)    Prec 99.219% (99.319%)
Epoch: [88][200/391]    Time 0.056 (0.056)      Data 0.001 (0.003)      Loss
0.0167 (0.0202)    Prec 99.219% (99.335%)
Epoch: [88][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0117 (0.0202)    Prec 100.000% (99.328%)
Validation starts
Test: [0/79]    Time 0.241 (0.241)      Loss 0.2084 (0.2084)    Prec 93.750%
(93.750%)
 * Prec 91.200%
best acc: 91.320000
Epoch: [89][0/391]      Time 0.271 (0.271)      Data 0.224 (0.224)      Loss
0.0149 (0.0149)    Prec 100.000% (100.000%)
Epoch: [89][100/391]    Time 0.055 (0.057)      Data 0.001 (0.004)      Loss
0.0074 (0.0197)    Prec 100.000% (99.381%)
Epoch: [89][200/391]    Time 0.054 (0.056)      Data 0.001 (0.003)      Loss
0.0209 (0.0196)    Prec 99.219% (99.351%)
Epoch: [89][300/391]    Time 0.054 (0.056)      Data 0.001 (0.002)      Loss
0.0083 (0.0193)    Prec 100.000% (99.362%)
Validation starts
Test: [0/79]    Time 0.225 (0.225)      Loss 0.1716 (0.1716)    Prec 95.312%
(95.312%)
 * Prec 90.950%
best acc: 91.320000
Epoch: [90][0/391]      Time 0.275 (0.275)      Data 0.234 (0.234)      Loss
```

```
0.0494 (0.0494)    Prec 98.438% (98.438%)
Epoch: [90][100/391]    Time 0.056 (0.058)      Data 0.001 (0.004)      Loss
0.0477 (0.0188)    Prec 98.438% (99.350%)
Epoch: [90][200/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.0167 (0.0184)    Prec 99.219% (99.370%)
Epoch: [90][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0110 (0.0188)    Prec 99.219% (99.369%)
Validation starts
Test: [0/79]    Time 0.228 (0.228)      Loss 0.1840 (0.1840)    Prec 95.312%
(95.312%)
 * Prec 91.260%
best acc: 91.320000
Epoch: [91][0/391]      Time 0.266 (0.266)      Data 0.219 (0.219)      Loss
0.0039 (0.0039)    Prec 100.000% (100.000%)
Epoch: [91][100/391]    Time 0.055 (0.057)      Data 0.001 (0.004)      Loss
0.0295 (0.0185)    Prec 98.438% (99.381%)
Epoch: [91][200/391]    Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.0072 (0.0177)    Prec 100.000% (99.390%)
Epoch: [91][300/391]    Time 0.056 (0.056)      Data 0.001 (0.002)      Loss
0.0057 (0.0176)    Prec 100.000% (99.421%)
Validation starts
Test: [0/79]    Time 0.253 (0.253)      Loss 0.1996 (0.1996)    Prec 96.094%
(96.094%)
 * Prec 91.380%
best acc: 91.380000
Epoch: [92][0/391]      Time 0.272 (0.272)      Data 0.230 (0.230)      Loss
0.0136 (0.0136)    Prec 99.219% (99.219%)
Epoch: [92][100/391]    Time 0.054 (0.057)      Data 0.001 (0.004)      Loss
0.0074 (0.0162)    Prec 100.000% (99.497%)
Epoch: [92][200/391]    Time 0.053 (0.056)      Data 0.001 (0.003)      Loss
0.0251 (0.0158)    Prec 99.219% (99.499%)
Epoch: [92][300/391]    Time 0.061 (0.056)      Data 0.001 (0.002)      Loss
0.0315 (0.0166)    Prec 99.219% (99.496%)
Validation starts
Test: [0/79]    Time 0.233 (0.233)      Loss 0.1921 (0.1921)    Prec 96.094%
(96.094%)
 * Prec 91.380%
best acc: 91.380000
Epoch: [93][0/391]      Time 0.284 (0.284)      Data 0.237 (0.237)      Loss
0.0277 (0.0277)    Prec 98.438% (98.438%)
Epoch: [93][100/391]    Time 0.055 (0.057)      Data 0.002 (0.004)      Loss
0.0112 (0.0143)    Prec 100.000% (99.520%)
Epoch: [93][200/391]    Time 0.053 (0.056)      Data 0.001 (0.003)      Loss
0.0232 (0.0140)    Prec 99.219% (99.537%)
Epoch: [93][300/391]    Time 0.056 (0.056)      Data 0.002 (0.002)      Loss
0.0072 (0.0149)    Prec 100.000% (99.504%)
Validation starts
Test: [0/79]    Time 0.217 (0.217)      Loss 0.2127 (0.2127)    Prec 96.094%
```

```
(96.094%)
 * Prec 91.340%
best acc: 91.380000
Epoch: [94][0/391]      Time 0.280 (0.280)      Data 0.232 (0.232)      Loss
0.0344 (0.0344)    Prec 99.219% (99.219%)
Epoch: [94][100/391]    Time 0.059 (0.057)      Data 0.002 (0.004)      Loss
0.0162 (0.0167)    Prec 99.219% (99.513%)
Epoch: [94][200/391]    Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.0128 (0.0154)    Prec 99.219% (99.526%)
Epoch: [94][300/391]    Time 0.053 (0.056)      Data 0.001 (0.002)      Loss
0.0131 (0.0149)    Prec 100.000% (99.546%)
Validation starts
Test: [0/79]    Time 0.210 (0.210)      Loss 0.2090 (0.2090)    Prec 95.312%
(95.312%)
 * Prec 91.320%
best acc: 91.380000
Epoch: [95][0/391]      Time 0.294 (0.294)      Data 0.249 (0.249)      Loss
0.0032 (0.0032)    Prec 100.000% (100.000%)
Epoch: [95][100/391]    Time 0.056 (0.058)      Data 0.002 (0.004)      Loss
0.0031 (0.0152)    Prec 100.000% (99.513%)
Epoch: [95][200/391]    Time 0.055 (0.056)      Data 0.001 (0.003)      Loss
0.0818 (0.0151)    Prec 96.875% (99.526%)
Epoch: [95][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0053 (0.0161)    Prec 100.000% (99.483%)
Validation starts
Test: [0/79]    Time 0.235 (0.235)      Loss 0.2050 (0.2050)    Prec 95.312%
(95.312%)
 * Prec 91.410%
best acc: 91.410000
Epoch: [96][0/391]      Time 0.267 (0.267)      Data 0.218 (0.218)      Loss
0.0111 (0.0111)    Prec 99.219% (99.219%)
Epoch: [96][100/391]    Time 0.056 (0.057)      Data 0.001 (0.004)      Loss
0.0271 (0.0150)    Prec 99.219% (99.536%)
Epoch: [96][200/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.0369 (0.0147)    Prec 99.219% (99.537%)
Epoch: [96][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0187 (0.0157)    Prec 99.219% (99.507%)
Validation starts
Test: [0/79]    Time 0.221 (0.221)      Loss 0.1923 (0.1923)    Prec 94.531%
(94.531%)
 * Prec 91.350%
best acc: 91.410000
Epoch: [97][0/391]      Time 0.266 (0.266)      Data 0.221 (0.221)      Loss
0.0034 (0.0034)    Prec 100.000% (100.000%)
Epoch: [97][100/391]    Time 0.055 (0.057)      Data 0.001 (0.004)      Loss
0.0098 (0.0145)    Prec 99.219% (99.513%)
Epoch: [97][200/391]    Time 0.060 (0.057)      Data 0.001 (0.003)      Loss
0.0065 (0.0149)    Prec 100.000% (99.479%)
```

```
Epoch: [97][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0180 (0.0153)    Prec 100.000% (99.489%)
Validation starts
Test: [0/79]    Time 0.224 (0.224)      Loss 0.2169 (0.2169)    Prec 94.531%
(94.531%)
 * Prec 91.340%
best acc: 91.410000
Epoch: [98][0/391]      Time 0.274 (0.274)      Data 0.233 (0.233)      Loss
0.0144 (0.0144)    Prec 100.000% (100.000%)
Epoch: [98][100/391]    Time 0.055 (0.059)      Data 0.001 (0.004)      Loss
0.0020 (0.0144)    Prec 100.000% (99.575%)
Epoch: [98][200/391]    Time 0.055 (0.057)      Data 0.001 (0.003)      Loss
0.0238 (0.0154)    Prec 98.438% (99.514%)
Epoch: [98][300/391]    Time 0.055 (0.056)      Data 0.001 (0.002)      Loss
0.0104 (0.0151)    Prec 100.000% (99.512%)
Validation starts
Test: [0/79]    Time 0.233 (0.233)      Loss 0.2104 (0.2104)    Prec 95.312%
(95.312%)
 * Prec 91.410%
best acc: 91.410000
Epoch: [99][0/391]      Time 0.261 (0.261)      Data 0.218 (0.218)      Loss
0.0469 (0.0469)    Prec 98.438% (98.438%)
Epoch: [99][100/391]    Time 0.054 (0.057)      Data 0.001 (0.004)      Loss
0.0216 (0.0139)    Prec 99.219% (99.544%)
Epoch: [99][200/391]    Time 0.052 (0.056)      Data 0.002 (0.003)      Loss
0.0088 (0.0151)    Prec 100.000% (99.499%)
Epoch: [99][300/391]    Time 0.058 (0.056)      Data 0.001 (0.002)      Loss
0.0170 (0.0153)    Prec 100.000% (99.517%)
Validation starts
Test: [0/79]    Time 0.228 (0.228)      Loss 0.1903 (0.1903)    Prec 96.094%
(96.094%)
 * Prec 91.570%
best acc: 91.570000
```

[3]:
```python
PATH = "result/VGG16_quant_project/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)))
```

```
/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)


Test set: Accuracy: 9157/10000 (92%)
```

[4]:
```
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()
i = 0

for layer in model.modules():
    i = i+1
    if isinstance(layer, QuantConv2d):
        print(i,"-th layer prehooked")
        layer.register_forward_pre_hook(save_output)
####################################################

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

```
3 -th layer prehooked
```

```
7 -th layer prehooked
12 -th layer prehooked
16 -th layer prehooked
21 -th layer prehooked
25 -th layer prehooked
29 -th layer prehooked
34 -th layer prehooked
38 -th layer prehooked
41 -th layer prehooked
46 -th layer prehooked
50 -th layer prehooked
54 -th layer prehooked
```

[5]:
```
w_bit = 4
weight_q = model.features[27].weight_q
w_alpha = model.features[27].weight_quant.wgt_alpha
w_delta = w_alpha/(2**(w_bit-1)-1)
weight_int = weight_q/w_delta

x_bit = 4
x = save_output.outputs[8][0]
x_alpha = model.features[27].act_alpha
x_delta = x_alpha/(2**x_bit-1)
act_quant = act_quantization(x_bit)
x_q = act_quant(x,x_alpha)
x_int = x_q/x_delta

conv_int = nn.Conv2d(8,8,kernel_size=3, padding=1, bias=False)
conv_int.weight = torch.nn.parameter.Parameter(weight_int)
output_int = conv_int(x_int)
psum_recovered = output_int*w_delta*x_delta
relu = nn.ReLU(inplace=True)
psum_after_relu = relu(psum_recovered)

difference = (save_output.outputs[9][0] - psum_after_relu).mean()

print("The difference between psum original and psum recovered = {}".
 ↪format(difference))
```

```
The difference between psum original and psum recovered = -2.893951744908918e-08
```

[178]:
```
padding = 1
stride = 1
array_size = 8
x_size = x_int.size()
nig = range(x_size[2])
njg = range(x_size[3])
```

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

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

a_pad = torch.zeros((x_size[1],x_size[2]+2*padding,x_size[3]+2*padding))
act_int = x_int[0,:,:,:]
a_pad[ :, padding:padding+len(nig), padding:padding+len(njg)] = act_int.cuda()
a_pad = torch.reshape(a_pad, (a_pad.size(0), -1))
w_int = torch.reshape(weight_int, (weight_int.size(0), weight_int.size(1), -1))

ic_tile = range(int(int(act_int.size(0))/array_size))
oc_tile = range(int(int(w_int.size(0))/array_size))


a_tile = torch.
 ↪zeros(len(ic_tile),array_size,len(nig)+padding*2,len(njg)+padding*2).cuda()
a_tile = torch.reshape(a_tile,(a_tile.size(0),a_tile.size(1),-1))

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

w_tile = torch.zeros(len(ic_tile),len(oc_tile),array_size,array_size,len(kijg)).
 ↪cuda()



for oct in oc_tile:
    for ict in ic_tile:
        w_tile[ict,oct,:,:,:] = w_int[(oct*array_size):
 ↪((oct+1)*array_size),(ict*array_size):((ict+1)*array_size),:]



p_nijg = range(a_pad.size(1)) ## paded activation's nij group

psum = torch.zeros(len(ic_tile),len(oc_tile),array_size,len(p_nijg),len(kijg)).
 ↪cuda()

for kij in kijg:
    for ict in ic_tile:
        for oct in oc_tile:
            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_tile[ict,oct,:,:,kij])
```

```
                psum[ict,oct,:, nij, kij] = m(a_tile[ict,:,nij]).cuda()
```

[179]:
```
import math

a_pad_ni_dim = int(math.sqrt(a_pad.size(1))) # 32 + 2*pad = 34

o_ni_dim = int((a_pad_ni_dim - (ki_dim- 1) - 1)/stride + 1) #34 - 2 - 1 + 1 = 32
o_nijg = range(o_ni_dim**2)

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


### SFP accumulation ###
for o_nij in o_nijg:
    for kij in kijg:
        for ict in ic_tile:
            for oct in oc_tile:
                out[oct*array_size:(oct+1)*array_size,o_nij] =␣
↪out[oct*array_size:(oct+1)*array_size,o_nij] + \
                psum[ict,oct,:, int(o_nij/o_ni_dim)*a_pad_ni_dim +␣
↪o_nij%o_ni_dim + int(kij/ki_dim)*a_pad_ni_dim + kij%ki_dim, kij]
                ## 2nd index = (int(o_nij/30)*32 + o_nij%30) + (int(kij/3)*32 +␣
↪kij%3)
```

[180]:
```
## Helper functions to dump files
def dec_to_bin(arr,bit):
    bin_arr = []
    for a in arr:
        a = int(a)
        if a < 0:
            a+=bit
        b = '{0:04b}'.format(int(a))
        bin_arr.append(b)
    return bin_arr

def dec_to_bin_psum(arr,bit):
    bin_arr = []
    for a in arr:
        a = int(a)
        sign = 0
        if a < 0:
            sign = 1
            a+=bit
        b = '{0:16b}'.format(int(a))
        if sign == 1:
            b = b.replace(" ", "1")
```

```
        else:
            b = b.replace(" ","0")
        bin_arr.append(b)
    return bin_arr


def convert_to_list(arr):
    return arr.tolist()
```

[181]:
```
## Activation dump

fp_act = open('activation_project.txt','w')
fp_act.write("#####\n")
fp_act.write("#####\n")
fp_act.write("#####\n")
fp_act_dec = open('activation_dec.txt','w')
for i in range(36):
    act_line = a_pad[:,i]
    act_arr = [int(j+0.001) for j in convert_to_list(act_line)]
    for a in act_arr[::-1]:
        #print(int(a))
        fp_act_dec.write(str(int(a)))
        fp_act_dec.write(" ")
    fp_act_dec.write("\n")
    bin_act = dec_to_bin(act_arr,16)
    for b in bin_act[::-1]:
        #print(b)
        fp_act.write(b)
    fp_act.write('\n')


fp_act.close()
fp_act_dec.close()
```

[182]:
```
## Weight dump
fp_wgt = open('weight_project.txt','w')
fp_wgt.write("#####\n")
fp_wgt.write("#####\n")
fp_wgt.write("#####\n")
fp_wgt_dec = open('weight_dec.txt','w')

for kij in range(9):
    for w in range(8):
        w_line = w_int[w,:,kij]
        w_arr = []
        for i in convert_to_list(w_line):
            if i < 0:
```

```
                w_arr.append(int(i-0.001))
            else:
                w_arr.append(int(i+0.001))

        for ww in w_arr[::-1]:
            #print(int(ww))
            fp_wgt_dec.write(str(int(ww)))
            fp_wgt_dec.write(" ")
        fp_wgt_dec.write("\n")
        bin_wgt = dec_to_bin(w_arr,16)
        for b in bin_wgt[::-1]:
            #print("Count = {} , b = {}".format(count,b))
            fp_wgt.write(b)
        fp_wgt.write('\n')

fp_wgt.close()
fp_wgt_dec.close()
```

[183]:
```
## psum dump
fp_psum = open('psum_project.txt','w')
fp_psum.write("#####\n")
fp_psum.write("#####\n")
fp_psum.write("#####\n")
fp_psum_dec = open('psum_dec.txt','w')
fp_psum_relu = open('output_project.txt','w')
for kij in range(16):
    psum_line = out[:,kij]
    psum_arr = []
    out_arr = []
    for i in convert_to_list(psum_line):
        if i < 0:
            psum_arr.append(int(i-0.001))
            out_arr.append(0)
        else:
            psum_arr.append(int(i+0.001))
            out_arr.append(int(i+0.001))
    for p in psum_arr:
        #print(int(a))
        fp_psum_dec.write(str(int(p)))
        fp_psum_dec.write(" ")
    fp_psum_dec.write("\n")
    bin_act = dec_to_bin_psum(psum_arr,65536)
    out_bin_act = dec_to_bin_psum(out_arr,65536)
    for b in bin_act:
        #print(b)
        fp_psum.write(b)
        fp_psum.write(' ')
```

```
    fp_psum.write('\n')
    fp_psum_relu.write("#")
    fp_psum_relu.write("#")

    for b in out_bin_act:
        fp_psum_relu.write(b)
        fp_psum_relu.write(' ')
    fp_psum_relu.write("\n")
fp_psum.close()
fp_psum_dec.close()
fp_psum_relu.close()
```