Cosine_Annealing_LR_Scheduler

December 11, 2025

[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
print('=> Building model...')



batch_size = 128
model_name = "VGG16_quant"
model = VGG16_quant_part1()

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 warmup_lr(optimizer, base_lr, epoch, warmup_epochs):
    lr = base_lr * float(epoch + 1) / warmup_epochs
    for p in optimizer.param_groups:
        p['lr'] = lr
```

```
=> 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
```

[3]:
```
# Training Cell

# PATH = "result/VGG_quant_part2_4bit/model_best.pth.tar"
# checkpoint = torch.load(PATH)
# model.load_state_dict(checkpoint['state_dict'])
lr = 0.01
warmup_epochs = 10
weight_decay = 1e-4
epochs = 200
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)
scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(
    optimizer,
    T_max=epochs - warmup_epochs,
    eta_min=1e-5
)

#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):
    # ---- WARMUP ---------------------
    if epoch < warmup_epochs:
        warmup_lr(optimizer, lr, epoch, warmup_epochs)
    # ---- COSINE AFTER WARMUP --------
    else:
        scheduler.step()

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

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

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

```
Epoch: [0][0/391]       Time 0.775 (0.775)      Data 0.375 (0.375)      Loss
2.4753 (2.4753)    Prec 3.906% (3.906%)
Epoch: [0][100/391]     Time 0.033 (0.043)      Data 0.002 (0.005)      Loss
1.5047 (1.8725)    Prec 46.875% (29.579%)
Epoch: [0][200/391]     Time 0.034 (0.039)      Data 0.001 (0.003)      Loss
1.5098 (1.7337)    Prec 46.875% (34.939%)
Epoch: [0][300/391]     Time 0.035 (0.038)      Data 0.001 (0.003)      Loss
```

```
1.4039 (1.6497)    Prec 41.406% (38.325%)
Validation starts
Test: [0/79]    Time 0.105 (0.105)      Loss 1.3521 (1.3521)    Prec 48.438%
(48.438%)
 * Prec 50.500%
best acc: 50.500000
Epoch: [1][0/391]       Time 0.322 (0.322)      Data 0.299 (0.299)      Loss
1.3402 (1.3402)    Prec 49.219% (49.219%)
Epoch: [1][100/391]     Time 0.033 (0.038)      Data 0.001 (0.004)      Loss
1.4258 (1.3966)    Prec 46.094% (49.180%)
Epoch: [1][200/391]     Time 0.038 (0.037)      Data 0.005 (0.003)      Loss
1.2034 (1.3556)    Prec 57.812% (50.789%)
Epoch: [1][300/391]     Time 0.034 (0.036)      Data 0.004 (0.002)      Loss
1.1653 (1.3136)    Prec 59.375% (52.396%)
Validation starts
Test: [0/79]    Time 0.241 (0.241)      Loss 1.0135 (1.0135)    Prec 60.156%
(60.156%)
 * Prec 58.590%
best acc: 58.590000
Epoch: [2][0/391]       Time 0.338 (0.338)      Data 0.309 (0.309)      Loss
0.9773 (0.9773)    Prec 60.156% (60.156%)
Epoch: [2][100/391]     Time 0.039 (0.038)      Data 0.001 (0.005)      Loss
1.3377 (1.1582)    Prec 57.031% (58.632%)
Epoch: [2][200/391]     Time 0.032 (0.037)      Data 0.001 (0.003)      Loss
1.0670 (1.1271)    Prec 64.844% (59.861%)
Epoch: [2][300/391]     Time 0.032 (0.036)      Data 0.001 (0.003)      Loss
0.7997 (1.1047)    Prec 74.219% (60.629%)
Validation starts
Test: [0/79]    Time 0.083 (0.083)      Loss 0.8593 (0.8593)    Prec 70.312%
(70.312%)
 * Prec 65.100%
best acc: 65.100000
Epoch: [3][0/391]       Time 0.211 (0.211)      Data 0.189 (0.189)      Loss
0.9124 (0.9124)    Prec 64.844% (64.844%)
Epoch: [3][100/391]     Time 0.034 (0.037)      Data 0.003 (0.003)      Loss
0.8118 (0.9934)    Prec 71.094% (64.782%)
Epoch: [3][200/391]     Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
1.1648 (0.9748)    Prec 60.156% (65.442%)
Epoch: [3][300/391]     Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.8311 (0.9535)    Prec 68.750% (66.378%)
Validation starts
Test: [0/79]    Time 0.329 (0.329)      Loss 0.8829 (0.8829)    Prec 68.750%
(68.750%)
 * Prec 67.970%
best acc: 67.970000
Epoch: [4][0/391]       Time 0.345 (0.345)      Data 0.318 (0.318)      Loss
0.6883 (0.6883)    Prec 75.781% (75.781%)
Epoch: [4][100/391]     Time 0.033 (0.038)      Data 0.001 (0.005)      Loss
```

```
0.9929 (0.8982)    Prec 64.844% (68.742%)
Epoch: [4][200/391]     Time 0.034 (0.037)      Data 0.002 (0.003)      Loss
0.8074 (0.8808)    Prec 72.656% (69.298%)
Epoch: [4][300/391]     Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.8542 (0.8605)    Prec 67.969% (70.089%)
Validation starts
Test: [0/79]    Time 0.230 (0.230)      Loss 1.0539 (1.0539)    Prec 64.844%
(64.844%)
 * Prec 67.840%
best acc: 67.970000
Epoch: [5][0/391]       Time 0.309 (0.309)      Data 0.286 (0.286)      Loss
0.6760 (0.6760)    Prec 75.000% (75.000%)
Epoch: [5][100/391]     Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.8001 (0.8232)    Prec 71.094% (72.130%)
Epoch: [5][200/391]     Time 0.036 (0.037)      Data 0.002 (0.003)      Loss
0.7534 (0.8081)    Prec 71.875% (72.268%)
Epoch: [5][300/391]     Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.6687 (0.7956)    Prec 78.906% (72.711%)
Validation starts
Test: [0/79]    Time 0.129 (0.129)      Loss 0.8169 (0.8169)    Prec 71.094%
(71.094%)
 * Prec 68.460%
best acc: 68.460000
Epoch: [6][0/391]       Time 0.269 (0.269)      Data 0.245 (0.245)      Loss
0.8181 (0.8181)    Prec 72.656% (72.656%)
Epoch: [6][100/391]     Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.7346 (0.7473)    Prec 73.438% (74.497%)
Epoch: [6][200/391]     Time 0.034 (0.037)      Data 0.001 (0.003)      Loss
0.7130 (0.7404)    Prec 81.250% (74.872%)
Epoch: [6][300/391]     Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.6501 (0.7271)    Prec 78.125% (75.197%)
Validation starts
Test: [0/79]    Time 0.244 (0.244)      Loss 0.7318 (0.7318)    Prec 73.438%
(73.438%)
 * Prec 72.090%
best acc: 72.090000
Epoch: [7][0/391]       Time 0.342 (0.342)      Data 0.319 (0.319)      Loss
0.5373 (0.5373)    Prec 82.812% (82.812%)
Epoch: [7][100/391]     Time 0.035 (0.038)      Data 0.001 (0.005)      Loss
0.6230 (0.6868)    Prec 82.031% (76.764%)
Epoch: [7][200/391]     Time 0.034 (0.037)      Data 0.001 (0.003)      Loss
0.6185 (0.6919)    Prec 78.906% (76.465%)
Epoch: [7][300/391]     Time 0.037 (0.036)      Data 0.001 (0.002)      Loss
0.9112 (0.6891)    Prec 71.094% (76.762%)
Validation starts
Test: [0/79]    Time 0.240 (0.240)      Loss 0.7708 (0.7708)    Prec 75.781%
(75.781%)
 * Prec 74.150%
```

```
best acc: 74.150000
Epoch: [8][0/391]       Time 0.323 (0.323)      Data 0.299 (0.299)      Loss
0.5736 (0.5736)    Prec 81.250% (81.250%)
Epoch: [8][100/391]     Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.5834 (0.6574)    Prec 78.906% (77.916%)
Epoch: [8][200/391]     Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.6854 (0.6504)    Prec 73.438% (78.024%)
Epoch: [8][300/391]     Time 0.034 (0.036)      Data 0.001 (0.002)      Loss
0.6758 (0.6420)    Prec 76.562% (78.244%)
Validation starts
Test: [0/79]    Time 0.243 (0.243)      Loss 0.6495 (0.6495)    Prec 78.906%
(78.906%)
 * Prec 78.520%
best acc: 78.520000
Epoch: [9][0/391]       Time 0.264 (0.264)      Data 0.240 (0.240)      Loss
0.5092 (0.5092)    Prec 79.688% (79.688%)
Epoch: [9][100/391]     Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.6204 (0.6195)    Prec 75.000% (79.061%)
Epoch: [9][200/391]     Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.6260 (0.6368)    Prec 78.906% (78.553%)
Epoch: [9][300/391]     Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.3551 (0.6264)    Prec 87.500% (78.847%)
Validation starts
Test: [0/79]    Time 0.242 (0.242)      Loss 0.7443 (0.7443)    Prec 72.656%
(72.656%)
 * Prec 73.550%
best acc: 78.520000
Epoch: [10][0/391]      Time 0.405 (0.405)      Data 0.337 (0.337)      Loss
0.5586 (0.5586)    Prec 82.031% (82.031%)
Epoch: [10][100/391]    Time 0.036 (0.039)      Data 0.002 (0.005)      Loss
0.4855 (0.5745)    Prec 85.156% (80.330%)
Epoch: [10][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.5794 (0.5654)    Prec 79.688% (80.935%)
Epoch: [10][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.4763 (0.5606)    Prec 83.594% (81.022%)
Validation starts
Test: [0/79]    Time 0.223 (0.223)      Loss 0.5348 (0.5348)    Prec 78.906%
(78.906%)
 * Prec 79.420%
best acc: 79.420000
Epoch: [11][0/391]      Time 0.377 (0.377)      Data 0.352 (0.352)      Loss
0.5708 (0.5708)    Prec 81.250% (81.250%)
Epoch: [11][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.4495 (0.5210)    Prec 83.594% (82.403%)
Epoch: [11][200/391]    Time 0.033 (0.037)      Data 0.001 (0.003)      Loss
0.6876 (0.5345)    Prec 80.469% (82.008%)
Epoch: [11][300/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.5806 (0.5292)    Prec 75.781% (82.145%)
```

```
Validation starts
Test: [0/79]    Time 0.267 (0.267)      Loss 0.4903 (0.4903)    Prec 83.594%
(83.594%)
 * Prec 79.350%
best acc: 79.420000
Epoch: [12][0/391]      Time 0.395 (0.395)      Data 0.365 (0.365)      Loss
0.4641 (0.4641)    Prec 82.812% (82.812%)
Epoch: [12][100/391]    Time 0.037 (0.039)      Data 0.001 (0.005)      Loss
0.4984 (0.4841)    Prec 83.594% (83.694%)
Epoch: [12][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.4817 (0.4924)    Prec 82.031% (83.263%)
Epoch: [12][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.3234 (0.4934)    Prec 89.844% (83.251%)
Validation starts
Test: [0/79]    Time 0.215 (0.215)      Loss 0.5121 (0.5121)    Prec 81.250%
(81.250%)
 * Prec 80.640%
best acc: 80.640000
Epoch: [13][0/391]      Time 0.307 (0.307)      Data 0.282 (0.282)      Loss
0.4474 (0.4474)    Prec 86.719% (86.719%)
Epoch: [13][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.4211 (0.4676)    Prec 82.031% (84.027%)
Epoch: [13][200/391]    Time 0.035 (0.037)      Data 0.006 (0.003)      Loss
0.6269 (0.4755)    Prec 82.031% (83.671%)
Epoch: [13][300/391]    Time 0.034 (0.036)      Data 0.001 (0.002)      Loss
0.4647 (0.4695)    Prec 84.375% (83.910%)
Validation starts
Test: [0/79]    Time 0.144 (0.144)      Loss 0.5429 (0.5429)    Prec 84.375%
(84.375%)
 * Prec 80.590%
best acc: 80.640000
Epoch: [14][0/391]      Time 0.353 (0.353)      Data 0.329 (0.329)      Loss
0.3741 (0.3741)    Prec 89.844% (89.844%)
Epoch: [14][100/391]    Time 0.036 (0.038)      Data 0.001 (0.005)      Loss
0.4925 (0.4479)    Prec 80.469% (84.615%)
Epoch: [14][200/391]    Time 0.038 (0.037)      Data 0.001 (0.003)      Loss
0.3967 (0.4484)    Prec 86.719% (84.705%)
Epoch: [14][300/391]    Time 0.033 (0.036)      Data 0.001 (0.003)      Loss
0.5132 (0.4464)    Prec 82.031% (84.821%)
Validation starts
Test: [0/79]    Time 0.271 (0.271)      Loss 0.4194 (0.4194)    Prec 82.812%
(82.812%)
 * Prec 81.200%
best acc: 81.200000
Epoch: [15][0/391]      Time 0.338 (0.338)      Data 0.314 (0.314)      Loss
0.3429 (0.3429)    Prec 86.719% (86.719%)
Epoch: [15][100/391]    Time 0.036 (0.038)      Data 0.001 (0.005)      Loss
0.4079 (0.4133)    Prec 85.156% (86.239%)
```

```
Epoch: [15][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.5611 (0.4202)    Prec 80.469% (85.887%)
Epoch: [15][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.4379 (0.4238)    Prec 84.375% (85.706%)
Validation starts
Test: [0/79]    Time 0.138 (0.138)      Loss 0.4162 (0.4162)    Prec 85.938%
(85.938%)
 * Prec 82.620%
best acc: 82.620000
Epoch: [16][0/391]      Time 0.338 (0.338)      Data 0.315 (0.315)      Loss
0.3634 (0.3634)    Prec 89.062% (89.062%)
Epoch: [16][100/391]    Time 0.035 (0.038)      Data 0.001 (0.005)      Loss
0.3505 (0.4000)    Prec 87.500% (86.046%)
Epoch: [16][200/391]    Time 0.031 (0.037)      Data 0.001 (0.003)      Loss
0.4419 (0.4006)    Prec 85.938% (85.984%)
Epoch: [16][300/391]    Time 0.039 (0.036)      Data 0.001 (0.003)      Loss
0.5746 (0.4008)    Prec 82.031% (86.052%)
Validation starts
Test: [0/79]    Time 0.161 (0.161)      Loss 0.4131 (0.4131)    Prec 85.156%
(85.156%)
 * Prec 81.640%
best acc: 82.620000
Epoch: [17][0/391]      Time 0.307 (0.307)      Data 0.284 (0.284)      Loss
0.4836 (0.4836)    Prec 83.594% (83.594%)
Epoch: [17][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.3201 (0.3809)    Prec 89.844% (87.044%)
Epoch: [17][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.5456 (0.3809)    Prec 80.469% (86.991%)
Epoch: [17][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.4056 (0.3863)    Prec 86.719% (86.752%)
Validation starts
Test: [0/79]    Time 0.239 (0.239)      Loss 0.3839 (0.3839)    Prec 85.156%
(85.156%)
 * Prec 84.400%
best acc: 84.400000
Epoch: [18][200/391]    Time 0.041 (0.037)      Data 0.001 (0.003)      Loss
0.3737 (0.3697)    Prec 84.375% (87.181%)
Epoch: [18][300/391]    Time 0.032 (0.037)      Data 0.001 (0.003)      Loss
0.5148 (0.3706)    Prec 78.906% (87.269%)
Validation starts
Test: [0/79]    Time 0.222 (0.222)      Loss 0.3014 (0.3014)    Prec 89.844%
(89.844%)
 * Prec 84.370%
best acc: 84.400000
Epoch: [19][0/391]      Time 0.188 (0.188)      Data 0.166 (0.166)      Loss
0.4233 (0.4233)    Prec 85.938% (85.938%)
Epoch: [19][100/391]    Time 0.034 (0.037)      Data 0.001 (0.003)      Loss
0.3861 (0.3560)    Prec 87.500% (87.771%)
```

```
Epoch: [19][200/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.3378 (0.3630)    Prec 87.500% (87.453%)
Epoch: [19][300/391]    Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.4301 (0.3637)    Prec 83.594% (87.497%)
Validation starts
Test: [0/79]    Time 0.253 (0.253)      Loss 0.3555 (0.3555)    Prec 85.938%
(85.938%)
 * Prec 83.180%
best acc: 84.400000
Epoch: [20][0/391]      Time 0.262 (0.262)      Data 0.238 (0.238)      Loss
0.2719 (0.2719)    Prec 87.500% (87.500%)
Epoch: [20][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.3505 (0.3398)    Prec 89.062% (88.436%)
Epoch: [20][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.5034 (0.3453)    Prec 85.156% (88.219%)
Epoch: [20][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.3211 (0.3449)    Prec 89.844% (88.232%)
Validation starts
Test: [0/79]    Time 0.261 (0.261)      Loss 0.2998 (0.2998)    Prec 90.625%
(90.625%)
 * Prec 84.270%
best acc: 84.400000
Epoch: [21][0/391]      Time 0.313 (0.313)      Data 0.290 (0.290)      Loss
0.2673 (0.2673)    Prec 92.188% (92.188%)
Epoch: [21][100/391]    Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.2336 (0.3142)    Prec 91.406% (89.070%)
Epoch: [21][200/391]    Time 0.038 (0.037)      Data 0.001 (0.003)      Loss
0.3224 (0.3260)    Prec 89.844% (88.872%)
Epoch: [21][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.2814 (0.3244)    Prec 87.500% (88.925%)
Validation starts
Test: [0/79]    Time 0.149 (0.149)      Loss 0.4098 (0.4098)    Prec 85.938%
(85.938%)
 * Prec 83.240%
best acc: 84.400000
Epoch: [22][0/391]      Time 0.382 (0.382)      Data 0.358 (0.358)      Loss
0.3681 (0.3681)    Prec 86.719% (86.719%)
Epoch: [22][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.5635 (0.3240)    Prec 82.031% (88.800%)
Epoch: [22][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.3347 (0.3297)    Prec 88.281% (88.674%)
Epoch: [22][300/391]    Time 0.036 (0.036)      Data 0.001 (0.003)      Loss
0.4552 (0.3218)    Prec 84.375% (88.995%)
Validation starts
Test: [0/79]    Time 0.241 (0.241)      Loss 0.3976 (0.3976)    Prec 87.500%
(87.500%)
 * Prec 84.870%
best acc: 84.870000
```

```
Epoch: [23][0/391]      Time 0.272 (0.272)      Data 0.242 (0.242)      Loss
0.2314 (0.2314)    Prec 91.406% (91.406%)
Epoch: [23][100/391]    Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.3315 (0.3008)    Prec 86.719% (89.666%)
Epoch: [23][200/391]    Time 0.035 (0.036)      Data 0.003 (0.003)      Loss
0.2746 (0.3030)    Prec 91.406% (89.544%)
Epoch: [23][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.2636 (0.3087)    Prec 89.844% (89.312%)
Validation starts
Test: [0/79]    Time 0.139 (0.139)      Loss 0.3196 (0.3196)    Prec 89.062%
(89.062%)
 * Prec 84.890%
best acc: 84.890000
Epoch: [24][0/391]      Time 0.343 (0.343)      Data 0.319 (0.319)      Loss
0.2304 (0.2304)    Prec 93.750% (93.750%)
Epoch: [24][100/391]    Time 0.035 (0.038)      Data 0.001 (0.005)      Loss
0.2596 (0.2945)    Prec 90.625% (90.207%)
Epoch: [24][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.5271 (0.2922)    Prec 83.594% (90.054%)
Epoch: [24][300/391]    Time 0.036 (0.036)      Data 0.001 (0.003)      Loss
0.2638 (0.2987)    Prec 92.188% (89.813%)
Validation starts
Test: [0/79]    Time 0.223 (0.223)      Loss 0.2649 (0.2649)    Prec 89.844%
(89.844%)
 * Prec 84.530%
best acc: 84.890000
Epoch: [25][0/391]      Time 0.225 (0.225)      Data 0.202 (0.202)      Loss
0.4679 (0.4679)    Prec 86.719% (86.719%)
Epoch: [25][100/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.3754 (0.2921)    Prec 84.375% (89.790%)
Epoch: [25][200/391]    Time 0.036 (0.036)      Data 0.001 (0.003)      Loss
0.4582 (0.2904)    Prec 84.375% (89.968%)
Epoch: [25][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.1965 (0.2923)    Prec 93.750% (89.961%)
Validation starts
Test: [0/79]    Time 0.256 (0.256)      Loss 0.3497 (0.3497)    Prec 90.625%
(90.625%)
 * Prec 86.870%
best acc: 86.870000
Epoch: [26][0/391]      Time 0.356 (0.356)      Data 0.330 (0.330)      Loss
0.3068 (0.3068)    Prec 91.406% (91.406%)
Epoch: [26][100/391]    Time 0.035 (0.038)      Data 0.001 (0.005)      Loss
0.1633 (0.2820)    Prec 93.750% (90.385%)
Epoch: [26][200/391]    Time 0.033 (0.037)      Data 0.001 (0.003)      Loss
0.4502 (0.2765)    Prec 82.812% (90.528%)
Epoch: [26][300/391]    Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.3929 (0.2821)    Prec 84.375% (90.293%)
Validation starts
```

```
Test: [0/79]    Time 0.248 (0.248)      Loss 0.2674 (0.2674)    Prec 90.625%
(90.625%)
 * Prec 84.000%
best acc: 86.870000
Epoch: [27][0/391]      Time 0.412 (0.412)      Data 0.389 (0.389)      Loss
0.2052 (0.2052)    Prec 93.750% (93.750%)
Epoch: [27][100/391]    Time 0.035 (0.039)      Data 0.001 (0.006)      Loss
0.3730 (0.2536)    Prec 85.938% (91.120%)
Epoch: [27][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.2202 (0.2605)    Prec 89.062% (90.959%)
Epoch: [27][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.3340 (0.2680)    Prec 88.281% (90.682%)
Validation starts
Test: [0/79]    Time 0.154 (0.154)      Loss 0.2709 (0.2709)    Prec 92.188%
(92.188%)
 * Prec 85.390%
best acc: 86.870000
Epoch: [28][0/391]      Time 0.295 (0.295)      Data 0.231 (0.231)      Loss
0.1773 (0.1773)    Prec 96.094% (96.094%)
Epoch: [28][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.1579 (0.2504)    Prec 95.312% (91.491%)
Epoch: [28][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.2571 (0.2557)    Prec 89.844% (91.317%)
Epoch: [28][300/391]    Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.2545 (0.2557)    Prec 92.188% (91.370%)
Validation starts
Test: [0/79]    Time 0.253 (0.253)      Loss 0.2732 (0.2732)    Prec 90.625%
(90.625%)
 * Prec 83.690%
best acc: 86.870000
Epoch: [29][0/391]      Time 0.309 (0.309)      Data 0.285 (0.285)      Loss
0.2842 (0.2842)    Prec 90.625% (90.625%)
Epoch: [29][100/391]    Time 0.033 (0.038)      Data 0.001 (0.004)      Loss
0.2111 (0.2422)    Prec 91.406% (91.584%)
Epoch: [29][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.3285 (0.2468)    Prec 90.625% (91.426%)
Epoch: [29][300/391]    Time 0.033 (0.036)      Data 0.001 (0.002)      Loss
0.3308 (0.2499)    Prec 91.406% (91.375%)
Validation starts
Test: [0/79]    Time 0.255 (0.255)      Loss 0.3130 (0.3130)    Prec 91.406%
(91.406%)
 * Prec 84.120%
best acc: 86.870000
Epoch: [30][0/391]      Time 0.386 (0.386)      Data 0.360 (0.360)      Loss
0.2459 (0.2459)    Prec 89.844% (89.844%)
Epoch: [30][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.2706 (0.2455)    Prec 89.844% (91.306%)
Epoch: [30][200/391]    Time 0.032 (0.037)      Data 0.002 (0.003)      Loss
```

```
0.3253 (0.2489)    Prec 85.938% (91.262%)
Epoch: [30][300/391]    Time 0.038 (0.037)      Data 0.001 (0.003)      Loss
0.1771 (0.2499)    Prec 95.312% (91.240%)
Validation starts
Test: [0/79]    Time 0.243 (0.243)      Loss 0.1859 (0.1859)    Prec 92.188%
(92.188%)
 * Prec 86.750%
best acc: 86.870000
Epoch: [31][0/391]      Time 0.388 (0.388)      Data 0.364 (0.364)      Loss
0.1914 (0.1914)    Prec 94.531% (94.531%)
Epoch: [31][100/391]    Time 0.035 (0.039)      Data 0.002 (0.005)      Loss
0.1734 (0.2265)    Prec 93.750% (92.025%)
Epoch: [31][200/391]    Time 0.038 (0.037)      Data 0.001 (0.003)      Loss
0.2290 (0.2349)    Prec 91.406% (91.775%)
Epoch: [31][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.2733 (0.2374)    Prec 89.844% (91.713%)
Validation starts
Test: [0/79]    Time 0.258 (0.258)      Loss 0.3610 (0.3610)    Prec 85.938%
(85.938%)
 * Prec 83.310%
best acc: 86.870000
Epoch: [32][0/391]      Time 0.349 (0.349)      Data 0.325 (0.325)      Loss
0.2121 (0.2121)    Prec 92.969% (92.969%)
Epoch: [32][100/391]    Time 0.036 (0.038)      Data 0.001 (0.005)      Loss
0.3010 (0.2294)    Prec 89.062% (92.211%)
Epoch: [32][200/391]    Time 0.033 (0.037)      Data 0.001 (0.003)      Loss
0.1861 (0.2279)    Prec 94.531% (92.308%)
Epoch: [32][300/391]    Time 0.040 (0.036)      Data 0.002 (0.003)      Loss
0.1258 (0.2317)    Prec 94.531% (92.117%)
Validation starts
Test: [0/79]    Time 0.258 (0.258)      Loss 0.2978 (0.2978)    Prec 88.281%
(88.281%)
 * Prec 85.270%
best acc: 86.870000
Epoch: [33][0/391]      Time 0.403 (0.403)      Data 0.380 (0.380)      Loss
0.2008 (0.2008)    Prec 92.188% (92.188%)
Epoch: [33][100/391]    Time 0.041 (0.039)      Data 0.001 (0.005)      Loss
0.2063 (0.2145)    Prec 94.531% (92.791%)
Epoch: [33][200/391]    Time 0.041 (0.037)      Data 0.001 (0.003)      Loss
0.1459 (0.2153)    Prec 96.094% (92.603%)
Epoch: [33][300/391]    Time 0.034 (0.037)      Data 0.001 (0.003)      Loss
0.1413 (0.2165)    Prec 94.531% (92.551%)
Validation starts
Test: [0/79]    Time 0.249 (0.249)      Loss 0.2927 (0.2927)    Prec 92.188%
(92.188%)
 * Prec 85.410%
best acc: 86.870000
Epoch: [34][0/391]      Time 0.310 (0.310)      Data 0.287 (0.287)      Loss
```

```
0.1504 (0.1504)    Prec 94.531% (94.531%)
Epoch: [34][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.4186 (0.2189)    Prec 85.938% (92.721%)
Epoch: [34][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1995 (0.2181)    Prec 93.750% (92.646%)
Epoch: [34][300/391]    Time 0.036 (0.036)      Data 0.003 (0.002)      Loss
0.1595 (0.2166)    Prec 93.750% (92.561%)
Validation starts
Test: [0/79]    Time 0.235 (0.235)      Loss 0.3048 (0.3048)    Prec 87.500%
(87.500%)
 * Prec 86.030%
best acc: 86.870000
Epoch: [35][0/391]      Time 0.278 (0.278)      Data 0.255 (0.255)      Loss
0.2122 (0.2122)    Prec 92.969% (92.969%)
Epoch: [35][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.1937 (0.1979)    Prec 92.969% (93.224%)
Epoch: [35][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1893 (0.2048)    Prec 91.406% (92.879%)
Epoch: [35][300/391]    Time 0.035 (0.036)      Data 0.005 (0.002)      Loss
0.1230 (0.2037)    Prec 92.969% (92.891%)
Validation starts
Test: [0/79]    Time 0.252 (0.252)      Loss 0.2948 (0.2948)    Prec 90.625%
(90.625%)
 * Prec 86.780%
best acc: 86.870000
Epoch: [36][0/391]      Time 0.321 (0.321)      Data 0.295 (0.295)      Loss
0.1008 (0.1008)    Prec 97.656% (97.656%)
Epoch: [36][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.1998 (0.1956)    Prec 92.969% (93.301%)
Epoch: [36][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1759 (0.1945)    Prec 92.188% (93.284%)
Epoch: [36][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.2936 (0.2002)    Prec 89.844% (93.002%)
Validation starts
Test: [0/79]    Time 0.266 (0.266)      Loss 0.3292 (0.3292)    Prec 90.625%
(90.625%)
 * Prec 86.460%
best acc: 86.870000
Epoch: [37][0/391]      Time 0.314 (0.314)      Data 0.290 (0.290)      Loss
0.1783 (0.1783)    Prec 93.750% (93.750%)
Epoch: [37][100/391]    Time 0.030 (0.038)      Data 0.001 (0.004)      Loss
0.2403 (0.1901)    Prec 89.844% (93.410%)
Epoch: [37][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1634 (0.1893)    Prec 93.750% (93.447%)
Epoch: [37][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.2410 (0.1951)    Prec 91.406% (93.306%)
Validation starts
Test: [0/79]    Time 0.230 (0.230)      Loss 0.3523 (0.3523)    Prec 90.625%
```

```
(90.625%)
 * Prec 83.970%
best acc: 86.870000
Epoch: [38][0/391]      Time 0.317 (0.317)      Data 0.294 (0.294)      Loss
0.2798 (0.2798)    Prec 87.500% (87.500%)
Epoch: [38][100/391]    Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.2272 (0.1815)    Prec 92.188% (93.688%)
Epoch: [38][200/391]    Time 0.031 (0.037)      Data 0.001 (0.003)      Loss
0.2434 (0.1808)    Prec 89.062% (93.692%)
Epoch: [38][300/391]    Time 0.034 (0.036)      Data 0.001 (0.003)      Loss
0.2135 (0.1864)    Prec 91.406% (93.537%)
Validation starts
Test: [0/79]    Time 0.135 (0.135)      Loss 0.2832 (0.2832)    Prec 92.969%
(92.969%)
 * Prec 85.740%
best acc: 86.870000
Epoch: [39][0/391]      Time 0.410 (0.410)      Data 0.382 (0.382)      Loss
0.1398 (0.1398)    Prec 96.875% (96.875%)
Epoch: [39][100/391]    Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.1631 (0.1792)    Prec 94.531% (93.912%)
Epoch: [39][200/391]    Time 0.035 (0.037)      Data 0.007 (0.003)      Loss
0.2940 (0.1857)    Prec 90.625% (93.707%)
Epoch: [39][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1820 (0.1875)    Prec 95.312% (93.563%)
Validation starts
Test: [0/79]    Time 0.247 (0.247)      Loss 0.2373 (0.2373)    Prec 92.188%
(92.188%)
 * Prec 87.460%
best acc: 87.460000
Epoch: [40][0/391]      Time 0.260 (0.260)      Data 0.235 (0.235)      Loss
0.1624 (0.1624)    Prec 90.625% (90.625%)
Epoch: [40][100/391]    Time 0.034 (0.038)      Data 0.001 (0.004)      Loss
0.2137 (0.1684)    Prec 92.969% (93.943%)
Epoch: [40][200/391]    Time 0.040 (0.037)      Data 0.001 (0.003)      Loss
0.1510 (0.1731)    Prec 93.750% (93.902%)
Epoch: [40][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.1824 (0.1791)    Prec 92.969% (93.672%)
Validation starts
Test: [0/79]    Time 0.236 (0.236)      Loss 0.2169 (0.2169)    Prec 92.969%
(92.969%)
 * Prec 86.860%
best acc: 87.460000
Epoch: [41][0/391]      Time 0.292 (0.292)      Data 0.269 (0.269)      Loss
0.1254 (0.1254)    Prec 96.094% (96.094%)
Epoch: [41][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.1109 (0.1680)    Prec 95.312% (94.307%)
Epoch: [41][200/391]    Time 0.032 (0.037)      Data 0.001 (0.003)      Loss
0.2228 (0.1740)    Prec 95.312% (94.123%)
```

```
Epoch: [41][300/391]    Time 0.038 (0.036)      Data 0.001 (0.002)      Loss
0.1764 (0.1782)    Prec 92.969% (93.807%)
Validation starts
Test: [0/79]    Time 0.241 (0.241)      Loss 0.3143 (0.3143)    Prec 92.188%
(92.188%)
 * Prec 86.600%
best acc: 87.460000
Epoch: [42][0/391]      Time 0.400 (0.400)      Data 0.370 (0.370)      Loss
0.1960 (0.1960)    Prec 91.406% (91.406%)
Epoch: [42][100/391]    Time 0.032 (0.039)      Data 0.001 (0.005)      Loss
0.2057 (0.1633)    Prec 92.969% (94.214%)
Epoch: [42][200/391]    Time 0.038 (0.037)      Data 0.001 (0.003)      Loss
0.2824 (0.1642)    Prec 92.188% (94.224%)
Epoch: [42][300/391]    Time 0.039 (0.037)      Data 0.001 (0.003)      Loss
0.0861 (0.1684)    Prec 96.094% (94.106%)
Validation starts
Test: [0/79]    Time 0.098 (0.098)      Loss 0.2504 (0.2504)    Prec 92.188%
(92.188%)
 * Prec 87.440%
best acc: 87.460000
Epoch: [43][0/391]      Time 0.384 (0.384)      Data 0.360 (0.360)      Loss
0.1063 (0.1063)    Prec 94.531% (94.531%)
Epoch: [43][100/391]    Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.2007 (0.1609)    Prec 92.188% (94.524%)
Epoch: [43][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1689 (0.1598)    Prec 95.312% (94.520%)
Epoch: [43][300/391]    Time 0.042 (0.037)      Data 0.001 (0.003)      Loss
0.1077 (0.1617)    Prec 95.312% (94.376%)
Validation starts
Test: [0/79]    Time 0.262 (0.262)      Loss 0.2386 (0.2386)    Prec 90.625%
(90.625%)
 * Prec 88.110%
best acc: 88.110000
Epoch: [44][0/391]      Time 0.236 (0.236)      Data 0.212 (0.212)      Loss
0.1752 (0.1752)    Prec 92.969% (92.969%)
Epoch: [44][100/391]    Time 0.036 (0.038)      Data 0.001 (0.003)      Loss
0.1052 (0.1321)    Prec 95.312% (95.521%)
Epoch: [44][200/391]    Time 0.038 (0.036)      Data 0.001 (0.002)      Loss
0.0906 (0.1475)    Prec 97.656% (94.908%)
Epoch: [44][300/391]    Time 0.034 (0.036)      Data 0.001 (0.002)      Loss
0.2426 (0.1517)    Prec 91.406% (94.731%)
Validation starts
Test: [0/79]    Time 0.248 (0.248)      Loss 0.2600 (0.2600)    Prec 89.844%
(89.844%)
 * Prec 87.310%
best acc: 88.110000
Epoch: [45][0/391]      Time 0.345 (0.345)      Data 0.289 (0.289)      Loss
0.1334 (0.1334)    Prec 93.750% (93.750%)
```

```
Epoch: [45][100/391]    Time 0.035 (0.038)      Data 0.005 (0.004)      Loss
0.1122 (0.1503)    Prec 95.312% (94.678%)
Epoch: [45][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.2965 (0.1567)    Prec 92.969% (94.543%)
Epoch: [45][300/391]    Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.1582 (0.1594)    Prec 94.531% (94.401%)
Validation starts
Test: [0/79]    Time 0.249 (0.249)      Loss 0.2007 (0.2007)    Prec 92.969%
(92.969%)
 * Prec 87.410%
best acc: 88.110000
Epoch: [46][0/391]      Time 0.300 (0.300)      Data 0.277 (0.277)      Loss
0.1417 (0.1417)    Prec 96.094% (96.094%)
Epoch: [46][100/391]    Time 0.031 (0.038)      Data 0.001 (0.004)      Loss
0.1244 (0.1448)    Prec 96.875% (95.235%)
Epoch: [46][200/391]    Time 0.032 (0.037)      Data 0.005 (0.003)      Loss
0.1776 (0.1508)    Prec 93.750% (95.021%)
Epoch: [46][300/391]    Time 0.034 (0.036)      Data 0.001 (0.002)      Loss
0.1770 (0.1488)    Prec 95.312% (95.011%)
Validation starts
Test: [0/79]    Time 0.256 (0.256)      Loss 0.2424 (0.2424)    Prec 94.531%
(94.531%)
 * Prec 88.170%
best acc: 88.170000
Epoch: [47][0/391]      Time 0.289 (0.289)      Data 0.240 (0.240)      Loss
0.1368 (0.1368)    Prec 90.625% (90.625%)
Epoch: [47][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.1332 (0.1367)    Prec 93.750% (95.166%)
Epoch: [47][200/391]    Time 0.033 (0.037)      Data 0.001 (0.003)      Loss
0.1335 (0.1438)    Prec 94.531% (94.877%)
Epoch: [47][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.1256 (0.1451)    Prec 95.312% (94.806%)
Validation starts
Test: [0/79]    Time 0.244 (0.244)      Loss 0.2112 (0.2112)    Prec 92.188%
(92.188%)
 * Prec 88.480%
best acc: 88.480000
Epoch: [48][0/391]      Time 0.274 (0.274)      Data 0.250 (0.250)      Loss
0.1514 (0.1514)    Prec 95.312% (95.312%)
Epoch: [48][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.1272 (0.1376)    Prec 96.094% (95.204%)
Epoch: [48][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.1524 (0.1397)    Prec 95.312% (95.180%)
Epoch: [48][300/391]    Time 0.034 (0.036)      Data 0.001 (0.002)      Loss
0.1275 (0.1439)    Prec 96.094% (94.993%)
Validation starts
Test: [0/79]    Time 0.204 (0.204)      Loss 0.2693 (0.2693)    Prec 92.188%
(92.188%)
```

```
 * Prec 87.640%
best acc: 88.480000
Epoch: [49][0/391]      Time 0.239 (0.239)      Data 0.215 (0.215)      Loss
0.2137 (0.2137)    Prec 92.969% (92.969%)
Epoch: [49][100/391]    Time 0.035 (0.037)      Data 0.001 (0.004)      Loss
0.1132 (0.1317)    Prec 95.312% (95.235%)
Epoch: [49][200/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.1754 (0.1359)    Prec 92.188% (95.200%)
Epoch: [49][300/391]    Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.1584 (0.1354)    Prec 92.969% (95.191%)
Validation starts
Test: [0/79]    Time 0.255 (0.255)      Loss 0.1555 (0.1555)    Prec 93.750%
(93.750%)
 * Prec 88.160%
best acc: 88.480000
Epoch: [50][0/391]      Time 0.402 (0.402)      Data 0.377 (0.377)      Loss
0.1201 (0.1201)    Prec 96.094% (96.094%)
Epoch: [50][100/391]    Time 0.033 (0.039)      Data 0.005 (0.005)      Loss
0.2047 (0.1388)    Prec 91.406% (95.142%)
Epoch: [50][200/391]    Time 0.034 (0.037)      Data 0.001 (0.003)      Loss
0.1136 (0.1400)    Prec 95.312% (95.134%)
Epoch: [50][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1489 (0.1398)    Prec 95.312% (95.115%)
Validation starts
Test: [0/79]    Time 0.245 (0.245)      Loss 0.1207 (0.1207)    Prec 96.094%
(96.094%)
 * Prec 87.750%
best acc: 88.480000
Epoch: [51][0/391]      Time 0.228 (0.228)      Data 0.204 (0.204)      Loss
0.0726 (0.0726)    Prec 96.094% (96.094%)
Epoch: [51][100/391]    Time 0.032 (0.037)      Data 0.001 (0.004)      Loss
0.1220 (0.1266)    Prec 96.875% (95.475%)
Epoch: [51][200/391]    Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.1064 (0.1251)    Prec 96.094% (95.519%)
Epoch: [51][300/391]    Time 0.032 (0.036)      Data 0.001 (0.002)      Loss
0.1110 (0.1288)    Prec 95.312% (95.393%)
Validation starts
Test: [0/79]    Time 0.257 (0.257)      Loss 0.3396 (0.3396)    Prec 89.844%
(89.844%)
 * Prec 88.210%
best acc: 88.480000
Epoch: [52][0/391]      Time 0.394 (0.394)      Data 0.371 (0.371)      Loss
0.0540 (0.0540)    Prec 98.438% (98.438%)
Epoch: [52][100/391]    Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.1964 (0.1172)    Prec 92.969% (95.707%)
Epoch: [52][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1829 (0.1259)    Prec 95.312% (95.503%)
Epoch: [52][300/391]    Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
```

```
0.1037 (0.1296)    Prec 96.094% (95.385%)
Validation starts
Test: [0/79]    Time 0.244 (0.244)      Loss 0.2163 (0.2163)    Prec 92.969%
(92.969%)
 * Prec 88.410%
best acc: 88.480000
Epoch: [53][0/391]      Time 0.301 (0.301)      Data 0.277 (0.277)      Loss
0.1228 (0.1228)    Prec 96.094% (96.094%)
Epoch: [53][100/391]    Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.2287 (0.1182)    Prec 91.406% (95.838%)
Epoch: [53][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.2169 (0.1232)    Prec 93.750% (95.756%)
Epoch: [53][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0488 (0.1222)    Prec 99.219% (95.756%)
Validation starts
Test: [0/79]    Time 0.250 (0.250)      Loss 0.3279 (0.3279)    Prec 90.625%
(90.625%)
 * Prec 86.730%
best acc: 88.480000
Epoch: [54][0/391]      Time 0.297 (0.297)      Data 0.274 (0.274)      Loss
0.1048 (0.1048)    Prec 96.094% (96.094%)
Epoch: [54][100/391]    Time 0.029 (0.038)      Data 0.002 (0.004)      Loss
0.0511 (0.1115)    Prec 98.438% (96.024%)
Epoch: [54][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.1670 (0.1209)    Prec 96.094% (95.740%)
Epoch: [54][300/391]    Time 0.034 (0.036)      Data 0.002 (0.002)      Loss
0.1515 (0.1233)    Prec 96.875% (95.673%)
Validation starts
Test: [0/79]    Time 0.155 (0.155)      Loss 0.2473 (0.2473)    Prec 93.750%
(93.750%)
 * Prec 87.360%
best acc: 88.480000
Epoch: [55][0/391]      Time 0.332 (0.332)      Data 0.266 (0.266)      Loss
0.0948 (0.0948)    Prec 96.875% (96.875%)
Epoch: [55][100/391]    Time 0.034 (0.038)      Data 0.001 (0.004)      Loss
0.0650 (0.1160)    Prec 97.656% (96.225%)
Epoch: [55][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0948 (0.1186)    Prec 96.094% (95.985%)
Epoch: [55][300/391]    Time 0.039 (0.036)      Data 0.001 (0.002)      Loss
0.0564 (0.1204)    Prec 99.219% (95.925%)
Validation starts
Test: [0/79]    Time 0.239 (0.239)      Loss 0.2613 (0.2613)    Prec 92.188%
(92.188%)
 * Prec 87.860%
best acc: 88.480000
Epoch: [56][0/391]      Time 0.298 (0.298)      Data 0.272 (0.272)      Loss
0.1001 (0.1001)    Prec 96.875% (96.875%)
Epoch: [56][100/391]    Time 0.032 (0.038)      Data 0.001 (0.004)      Loss
```

```
0.1298 (0.1024)    Prec 96.094% (96.519%)
Epoch: [56][200/391]    Time 0.042 (0.037)      Data 0.001 (0.003)      Loss
0.0730 (0.1071)    Prec 97.656% (96.269%)
Epoch: [56][300/391]    Time 0.032 (0.036)      Data 0.001 (0.002)      Loss
0.1312 (0.1157)    Prec 95.312% (95.948%)
Validation starts
Test: [0/79]    Time 0.285 (0.285)      Loss 0.2997 (0.2997)    Prec 89.062%
(89.062%)
 * Prec 87.480%
best acc: 88.480000
Epoch: [57][0/391]      Time 0.384 (0.384)      Data 0.360 (0.360)      Loss
0.1337 (0.1337)    Prec 95.312% (95.312%)
Epoch: [57][100/391]    Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.0570 (0.1138)    Prec 99.219% (95.900%)
Epoch: [57][200/391]    Time 0.037 (0.037)      Data 0.001 (0.003)      Loss
0.1600 (0.1151)    Prec 94.531% (95.938%)
Epoch: [57][300/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.1037 (0.1163)    Prec 97.656% (95.943%)
Validation starts
Test: [0/79]    Time 0.253 (0.253)      Loss 0.2387 (0.2387)    Prec 88.281%
(88.281%)
 * Prec 88.100%
best acc: 88.480000
Epoch: [58][0/391]      Time 0.411 (0.411)      Data 0.382 (0.382)      Loss
0.1627 (0.1627)    Prec 96.094% (96.094%)
Epoch: [58][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0622 (0.1009)    Prec 97.656% (96.511%)
Epoch: [58][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0850 (0.1054)    Prec 95.312% (96.366%)
Epoch: [58][300/391]    Time 0.038 (0.037)      Data 0.001 (0.003)      Loss
0.1273 (0.1121)    Prec 95.312% (96.127%)
Validation starts
Test: [0/79]    Time 0.148 (0.148)      Loss 0.1936 (0.1936)    Prec 92.188%
(92.188%)
 * Prec 87.480%
best acc: 88.480000
Epoch: [59][0/391]      Time 0.357 (0.357)      Data 0.334 (0.334)      Loss
0.0812 (0.0812)    Prec 96.875% (96.875%)
Epoch: [59][100/391]    Time 0.035 (0.038)      Data 0.001 (0.005)      Loss
0.0832 (0.0993)    Prec 97.656% (96.542%)
Epoch: [59][200/391]    Time 0.032 (0.037)      Data 0.001 (0.003)      Loss
0.0513 (0.1046)    Prec 99.219% (96.366%)
Epoch: [59][300/391]    Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0870 (0.1061)    Prec 97.656% (96.291%)
Validation starts
Test: [0/79]    Time 0.158 (0.158)      Loss 0.2529 (0.2529)    Prec 92.969%
(92.969%)
 * Prec 87.690%
```

```
best acc: 88.480000
Epoch: [60][0/391]      Time 0.369 (0.369)      Data 0.346 (0.346)      Loss
0.2220 (0.2220)    Prec 93.750% (93.750%)
Epoch: [60][100/391]    Time 0.033 (0.039)      Data 0.001 (0.005)      Loss
0.0847 (0.1001)    Prec 96.875% (96.581%)
Epoch: [60][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0615 (0.0977)    Prec 98.438% (96.630%)
Epoch: [60][300/391]    Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.1337 (0.0969)    Prec 95.312% (96.665%)
Validation starts
Test: [0/79]    Time 0.254 (0.254)      Loss 0.1965 (0.1965)    Prec 92.969%
(92.969%)
 * Prec 88.260%
best acc: 88.480000
Epoch: [61][0/391]      Time 0.404 (0.404)      Data 0.379 (0.379)      Loss
0.0663 (0.0663)    Prec 98.438% (98.438%)
Epoch: [61][100/391]    Time 0.040 (0.039)      Data 0.001 (0.005)      Loss
0.0807 (0.0953)    Prec 97.656% (96.573%)
Epoch: [61][200/391]    Time 0.034 (0.037)      Data 0.001 (0.003)      Loss
0.1781 (0.0976)    Prec 94.531% (96.521%)
Epoch: [61][300/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.1262 (0.0968)    Prec 96.094% (96.551%)
Validation starts
Test: [0/79]    Time 0.256 (0.256)      Loss 0.2835 (0.2835)    Prec 92.188%
(92.188%)
 * Prec 88.080%
best acc: 88.480000
Epoch: [62][0/391]      Time 0.306 (0.306)      Data 0.279 (0.279)      Loss
0.0426 (0.0426)    Prec 97.656% (97.656%)
Epoch: [62][100/391]    Time 0.031 (0.038)      Data 0.001 (0.004)      Loss
0.1212 (0.0961)    Prec 94.531% (96.581%)
Epoch: [62][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1516 (0.0973)    Prec 95.312% (96.560%)
Epoch: [62][300/391]    Time 0.039 (0.036)      Data 0.001 (0.003)      Loss
0.0526 (0.0999)    Prec 98.438% (96.493%)
Validation starts
Test: [0/79]    Time 0.175 (0.175)      Loss 0.2750 (0.2750)    Prec 90.625%
(90.625%)
 * Prec 88.020%
best acc: 88.480000
Epoch: [63][0/391]      Time 0.222 (0.222)      Data 0.200 (0.200)      Loss
0.0862 (0.0862)    Prec 95.312% (95.312%)
Epoch: [63][100/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1120 (0.0871)    Prec 95.312% (96.875%)
Epoch: [63][200/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0556 (0.0874)    Prec 97.656% (96.968%)
Epoch: [63][300/391]    Time 0.035 (0.036)      Data 0.007 (0.002)      Loss
0.0937 (0.0930)    Prec 97.656% (96.766%)
```

```
Validation starts
Test: [0/79]    Time 0.318 (0.318)      Loss 0.2124 (0.2124)    Prec 95.312%
(95.312%)
 * Prec 88.810%
best acc: 88.810000
Epoch: [64][0/391]      Time 0.325 (0.325)      Data 0.282 (0.282)      Loss
0.1066 (0.1066)    Prec 96.875% (96.875%)
Epoch: [64][100/391]    Time 0.036 (0.038)      Data 0.005 (0.004)      Loss
0.0914 (0.0885)    Prec 96.875% (96.937%)
Epoch: [64][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0955 (0.0900)    Prec 96.875% (96.906%)
Epoch: [64][300/391]    Time 0.030 (0.036)      Data 0.001 (0.002)      Loss
0.1579 (0.0919)    Prec 93.750% (96.867%)
Validation starts
Test: [0/79]    Time 0.242 (0.242)      Loss 0.1944 (0.1944)    Prec 92.969%
(92.969%)
 * Prec 88.390%
best acc: 88.810000
Epoch: [65][0/391]      Time 0.346 (0.346)      Data 0.323 (0.323)      Loss
0.0387 (0.0387)    Prec 99.219% (99.219%)
Epoch: [65][100/391]    Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.0638 (0.0806)    Prec 97.656% (97.037%)
Epoch: [65][200/391]    Time 0.040 (0.037)      Data 0.001 (0.003)      Loss
0.1431 (0.0876)    Prec 95.312% (96.727%)
Epoch: [65][300/391]    Time 0.033 (0.036)      Data 0.001 (0.003)      Loss
0.0579 (0.0873)    Prec 98.438% (96.779%)
Validation starts
Test: [0/79]    Time 0.085 (0.085)      Loss 0.2255 (0.2255)    Prec 92.969%
(92.969%)
 * Prec 88.670%
best acc: 88.810000
Epoch: [66][0/391]      Time 0.294 (0.294)      Data 0.269 (0.269)      Loss
0.0480 (0.0480)    Prec 99.219% (99.219%)
Epoch: [66][100/391]    Time 0.038 (0.038)      Data 0.001 (0.004)      Loss
0.0583 (0.0891)    Prec 97.656% (96.906%)
Epoch: [66][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1212 (0.0875)    Prec 95.312% (96.984%)
Epoch: [66][300/391]    Time 0.032 (0.036)      Data 0.001 (0.002)      Loss
0.0252 (0.0881)    Prec 99.219% (96.940%)
Validation starts
Test: [0/79]    Time 0.238 (0.238)      Loss 0.3733 (0.3733)    Prec 89.062%
(89.062%)
 * Prec 88.650%
best acc: 88.810000
Epoch: [67][0/391]      Time 0.250 (0.250)      Data 0.189 (0.189)      Loss
0.1346 (0.1346)    Prec 96.094% (96.094%)
Epoch: [67][100/391]    Time 0.042 (0.038)      Data 0.001 (0.003)      Loss
0.1195 (0.0805)    Prec 96.094% (97.130%)
```

```
Epoch: [67][200/391]    Time 0.038 (0.036)      Data 0.001 (0.002)      Loss
0.1372 (0.0848)    Prec 94.531% (96.914%)
Epoch: [67][300/391]    Time 0.032 (0.036)      Data 0.001 (0.002)      Loss
0.0850 (0.0887)    Prec 96.875% (96.857%)
Validation starts
Test: [0/79]    Time 0.268 (0.268)      Loss 0.1488 (0.1488)    Prec 96.094%
(96.094%)
 * Prec 87.680%
best acc: 88.810000
Epoch: [68][0/391]      Time 0.282 (0.282)      Data 0.259 (0.259)      Loss
0.0797 (0.0797)    Prec 96.094% (96.094%)
Epoch: [68][100/391]    Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.0652 (0.0863)    Prec 96.875% (97.153%)
Epoch: [68][200/391]    Time 0.034 (0.037)      Data 0.001 (0.003)      Loss
0.1080 (0.0833)    Prec 96.875% (97.132%)
Epoch: [68][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.1143 (0.0838)    Prec 96.094% (97.145%)
Validation starts
Test: [0/79]    Time 0.243 (0.243)      Loss 0.2992 (0.2992)    Prec 91.406%
(91.406%)
 * Prec 88.050%
best acc: 88.810000
Epoch: [69][0/391]      Time 0.229 (0.229)      Data 0.206 (0.206)      Loss
0.1371 (0.1371)    Prec 95.312% (95.312%)
Epoch: [69][100/391]    Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.0340 (0.0762)    Prec 98.438% (97.424%)
Epoch: [69][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1615 (0.0784)    Prec 92.969% (97.275%)
Epoch: [69][300/391]    Time 0.033 (0.036)      Data 0.001 (0.002)      Loss
0.0416 (0.0815)    Prec 99.219% (97.158%)
Validation starts
Test: [0/79]    Time 0.156 (0.156)      Loss 0.2321 (0.2321)    Prec 93.750%
(93.750%)
 * Prec 88.540%
best acc: 88.810000
Epoch: [70][0/391]      Time 0.239 (0.239)      Data 0.215 (0.215)      Loss
0.0979 (0.0979)    Prec 96.094% (96.094%)
Epoch: [70][100/391]    Time 0.036 (0.037)      Data 0.001 (0.004)      Loss
0.1050 (0.0755)    Prec 95.312% (97.432%)
Epoch: [70][200/391]    Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.1019 (0.0794)    Prec 96.094% (97.303%)
Epoch: [70][300/391]    Time 0.038 (0.036)      Data 0.001 (0.002)      Loss
0.0739 (0.0807)    Prec 98.438% (97.275%)
Validation starts
Test: [0/79]    Time 0.232 (0.232)      Loss 0.2487 (0.2487)    Prec 94.531%
(94.531%)
 * Prec 88.720%
best acc: 88.810000
```

```
Epoch: [71][0/391]      Time 0.276 (0.276)      Data 0.202 (0.202)      Loss
0.0794 (0.0794)    Prec 97.656% (97.656%)
Epoch: [71][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0513 (0.0699)    Prec 97.656% (97.486%)
Epoch: [71][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1045 (0.0742)    Prec 96.875% (97.318%)
Epoch: [71][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0619 (0.0759)    Prec 96.875% (97.275%)
Validation starts
Test: [0/79]    Time 0.330 (0.330)      Loss 0.1682 (0.1682)    Prec 95.312%
(95.312%)
 * Prec 88.790%
best acc: 88.810000
Epoch: [72][0/391]      Time 0.168 (0.168)      Data 0.144 (0.144)      Loss
0.0420 (0.0420)    Prec 99.219% (99.219%)
Epoch: [72][100/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0297 (0.0697)    Prec 98.438% (97.440%)
Epoch: [72][200/391]    Time 0.032 (0.036)      Data 0.001 (0.002)      Loss
0.0617 (0.0735)    Prec 98.438% (97.435%)
Epoch: [72][300/391]    Time 0.039 (0.036)      Data 0.001 (0.002)      Loss
0.1452 (0.0741)    Prec 96.094% (97.386%)
Validation starts
Test: [0/79]    Time 0.281 (0.281)      Loss 0.2197 (0.2197)    Prec 93.750%
(93.750%)
 * Prec 88.290%
best acc: 88.810000
Epoch: [73][0/391]      Time 0.288 (0.288)      Data 0.257 (0.257)      Loss
0.1204 (0.1204)    Prec 95.312% (95.312%)
Epoch: [73][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0323 (0.0685)    Prec 98.438% (97.610%)
Epoch: [73][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1379 (0.0725)    Prec 95.312% (97.458%)
Epoch: [73][300/391]    Time 0.034 (0.036)      Data 0.001 (0.002)      Loss
0.0394 (0.0751)    Prec 97.656% (97.350%)
Validation starts
Test: [0/79]    Time 0.157 (0.157)      Loss 0.3153 (0.3153)    Prec 92.188%
(92.188%)
 * Prec 88.930%
best acc: 88.930000
Epoch: [74][0/391]      Time 0.442 (0.442)      Data 0.418 (0.418)      Loss
0.0609 (0.0609)    Prec 97.656% (97.656%)
Epoch: [74][100/391]    Time 0.031 (0.039)      Data 0.001 (0.006)      Loss
0.1216 (0.0645)    Prec 96.875% (97.780%)
Epoch: [74][200/391]    Time 0.036 (0.037)      Data 0.001 (0.004)      Loss
0.0670 (0.0657)    Prec 98.438% (97.703%)
Epoch: [74][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0850 (0.0681)    Prec 98.438% (97.571%)
Validation starts
```

```
Test: [0/79]    Time 0.227 (0.227)      Loss 0.3185 (0.3185)    Prec 92.188%
(92.188%)
 * Prec 88.260%
best acc: 88.930000
Epoch: [75][0/391]      Time 0.409 (0.409)      Data 0.385 (0.385)      Loss
0.0942 (0.0942)    Prec 96.875% (96.875%)
Epoch: [75][100/391]    Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.0593 (0.0630)    Prec 98.438% (97.803%)
Epoch: [75][200/391]    Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0593 (0.0670)    Prec 97.656% (97.629%)
Epoch: [75][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1024 (0.0676)    Prec 96.875% (97.623%)
Validation starts
Test: [0/79]    Time 0.230 (0.230)      Loss 0.3001 (0.3001)    Prec 91.406%
(91.406%)
 * Prec 89.140%
best acc: 89.140000
Epoch: [76][0/391]      Time 0.331 (0.331)      Data 0.307 (0.307)      Loss
0.1010 (0.1010)    Prec 96.875% (96.875%)
Epoch: [76][100/391]    Time 0.037 (0.038)      Data 0.001 (0.004)      Loss
0.0577 (0.0671)    Prec 99.219% (97.656%)
Epoch: [76][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0790 (0.0646)    Prec 95.312% (97.746%)
Epoch: [76][300/391]    Time 0.035 (0.036)      Data 0.008 (0.002)      Loss
0.0680 (0.0698)    Prec 96.875% (97.550%)
Validation starts
Test: [0/79]    Time 0.152 (0.152)      Loss 0.2130 (0.2130)    Prec 93.750%
(93.750%)
 * Prec 88.660%
best acc: 89.140000
Epoch: [77][0/391]      Time 0.355 (0.355)      Data 0.331 (0.331)      Loss
0.0284 (0.0284)    Prec 99.219% (99.219%)
Epoch: [77][100/391]    Time 0.035 (0.038)      Data 0.001 (0.005)      Loss
0.1113 (0.0656)    Prec 98.438% (97.695%)
Epoch: [77][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.1127 (0.0674)    Prec 96.875% (97.668%)
Epoch: [77][300/391]    Time 0.039 (0.036)      Data 0.001 (0.002)      Loss
0.0207 (0.0681)    Prec 100.000% (97.646%)
Validation starts
Test: [0/79]    Time 0.252 (0.252)      Loss 0.2739 (0.2739)    Prec 89.844%
(89.844%)
 * Prec 88.860%
best acc: 89.140000
Epoch: [78][0/391]      Time 0.389 (0.389)      Data 0.363 (0.363)      Loss
0.0607 (0.0607)    Prec 98.438% (98.438%)
Epoch: [78][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.1089 (0.0593)    Prec 96.875% (97.966%)
Epoch: [78][200/391]    Time 0.031 (0.037)      Data 0.001 (0.003)      Loss
```

```
0.0198 (0.0608)    Prec 100.000% (97.944%)
Epoch: [78][300/391]    Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0581 (0.0635)    Prec 98.438% (97.833%)
Validation starts
Test: [0/79]    Time 0.168 (0.168)      Loss 0.2055 (0.2055)    Prec 92.969%
(92.969%)
 * Prec 89.170%
best acc: 89.170000
Epoch: [79][0/391]      Time 0.286 (0.286)      Data 0.263 (0.263)      Loss
0.0528 (0.0528)    Prec 96.875% (96.875%)
Epoch: [79][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0693 (0.0567)    Prec 96.875% (97.997%)
Epoch: [79][200/391]    Time 0.035 (0.037)      Data 0.006 (0.003)      Loss
0.0663 (0.0590)    Prec 98.438% (97.909%)
Epoch: [79][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0396 (0.0607)    Prec 98.438% (97.908%)
Validation starts
Test: [0/79]    Time 0.247 (0.247)      Loss 0.2994 (0.2994)    Prec 93.750%
(93.750%)
 * Prec 88.750%
best acc: 89.170000
Epoch: [80][0/391]      Time 0.417 (0.417)      Data 0.394 (0.394)      Loss
0.0778 (0.0778)    Prec 97.656% (97.656%)
Epoch: [80][100/391]    Time 0.033 (0.039)      Data 0.001 (0.005)      Loss
0.0644 (0.0695)    Prec 97.656% (97.594%)
Epoch: [80][200/391]    Time 0.032 (0.037)      Data 0.001 (0.003)      Loss
0.0181 (0.0621)    Prec 99.219% (97.874%)
Epoch: [80][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0834 (0.0616)    Prec 96.094% (97.856%)
Validation starts
Test: [0/79]    Time 0.240 (0.240)      Loss 0.2524 (0.2524)    Prec 94.531%
(94.531%)
 * Prec 89.550%
best acc: 89.550000
Epoch: [81][0/391]      Time 0.366 (0.366)      Data 0.339 (0.339)      Loss
0.0381 (0.0381)    Prec 98.438% (98.438%)
Epoch: [81][100/391]    Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.0787 (0.0636)    Prec 96.875% (97.834%)
Epoch: [81][200/391]    Time 0.035 (0.037)      Data 0.007 (0.003)      Loss
0.0538 (0.0613)    Prec 98.438% (97.851%)
Epoch: [81][300/391]    Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0645 (0.0644)    Prec 97.656% (97.768%)
Validation starts
Test: [0/79]    Time 0.236 (0.236)      Loss 0.2263 (0.2263)    Prec 94.531%
(94.531%)
 * Prec 89.100%
best acc: 89.550000
Epoch: [82][0/391]      Time 0.312 (0.312)      Data 0.287 (0.287)      Loss
```

```
0.1145 (0.1145)    Prec 95.312% (95.312%)
Epoch: [82][100/391]    Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0686 (0.0569)    Prec 97.656% (97.935%)
Epoch: [82][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0942 (0.0525)    Prec 96.094% (98.138%)
Epoch: [82][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0276 (0.0532)    Prec 99.219% (98.118%)
Validation starts
Test: [0/79]    Time 0.262 (0.262)      Loss 0.2222 (0.2222)    Prec 92.969%
(92.969%)
 * Prec 89.000%
best acc: 89.550000
Epoch: [83][0/391]      Time 0.293 (0.293)      Data 0.270 (0.270)      Loss
0.0769 (0.0769)    Prec 97.656% (97.656%)
Epoch: [83][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0493 (0.0575)    Prec 97.656% (98.043%)
Epoch: [83][200/391]    Time 0.041 (0.037)      Data 0.001 (0.003)      Loss
0.0931 (0.0609)    Prec 96.875% (97.866%)
Epoch: [83][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0538 (0.0595)    Prec 97.656% (97.882%)
Validation starts
Test: [0/79]    Time 0.276 (0.276)      Loss 0.1622 (0.1622)    Prec 95.312%
(95.312%)
 * Prec 89.360%
best acc: 89.550000
Epoch: [84][0/391]      Time 0.403 (0.403)      Data 0.376 (0.376)      Loss
0.0321 (0.0321)    Prec 99.219% (99.219%)
Epoch: [84][100/391]    Time 0.033 (0.039)      Data 0.001 (0.005)      Loss
0.0833 (0.0514)    Prec 98.438% (98.252%)
Epoch: [84][200/391]    Time 0.037 (0.037)      Data 0.001 (0.003)      Loss
0.0223 (0.0541)    Prec 99.219% (98.134%)
Epoch: [84][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0853 (0.0559)    Prec 97.656% (98.069%)
Validation starts
Test: [0/79]    Time 0.242 (0.242)      Loss 0.2611 (0.2611)    Prec 89.844%
(89.844%)
 * Prec 89.320%
best acc: 89.550000
Epoch: [85][0/391]      Time 0.402 (0.402)      Data 0.378 (0.378)      Loss
0.0296 (0.0296)    Prec 99.219% (99.219%)
Epoch: [85][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0201 (0.0490)    Prec 99.219% (98.360%)
Epoch: [85][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0173 (0.0521)    Prec 99.219% (98.247%)
Epoch: [85][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0297 (0.0517)    Prec 99.219% (98.256%)
Validation starts
Test: [0/79]    Time 0.238 (0.238)      Loss 0.2512 (0.2512)    Prec 92.188%
```

```
(92.188%)
 * Prec 89.760%
best acc: 89.760000
Epoch: [86][0/391]      Time 0.253 (0.253)      Data 0.230 (0.230)      Loss
0.0315 (0.0315)    Prec 99.219% (99.219%)
Epoch: [86][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0656 (0.0451)    Prec 98.438% (98.430%)
Epoch: [86][200/391]    Time 0.036 (0.036)      Data 0.001 (0.003)      Loss
0.0697 (0.0465)    Prec 96.875% (98.371%)
Epoch: [86][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.1469 (0.0513)    Prec 95.312% (98.209%)
Validation starts
Test: [0/79]    Time 0.249 (0.249)      Loss 0.1967 (0.1967)    Prec 95.312%
(95.312%)
 * Prec 89.400%
best acc: 89.760000
Epoch: [87][0/391]      Time 0.393 (0.393)      Data 0.367 (0.367)      Loss
0.0267 (0.0267)    Prec 100.000% (100.000%)
Epoch: [87][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0798 (0.0489)    Prec 98.438% (98.337%)
Epoch: [87][200/391]    Time 0.032 (0.037)      Data 0.001 (0.003)      Loss
0.0659 (0.0505)    Prec 97.656% (98.220%)
Epoch: [87][300/391]    Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0292 (0.0523)    Prec 99.219% (98.126%)
Validation starts
Test: [0/79]    Time 0.241 (0.241)      Loss 0.1433 (0.1433)    Prec 93.750%
(93.750%)
 * Prec 89.480%
best acc: 89.760000
Epoch: [88][0/391]      Time 0.317 (0.317)      Data 0.291 (0.291)      Loss
0.0533 (0.0533)    Prec 97.656% (97.656%)
Epoch: [88][100/391]    Time 0.032 (0.038)      Data 0.001 (0.004)      Loss
0.0780 (0.0427)    Prec 98.438% (98.685%)
Epoch: [88][200/391]    Time 0.037 (0.037)      Data 0.001 (0.003)      Loss
0.0468 (0.0485)    Prec 97.656% (98.414%)
Epoch: [88][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.1135 (0.0502)    Prec 95.312% (98.305%)
Validation starts
Test: [0/79]    Time 0.178 (0.178)      Loss 0.3201 (0.3201)    Prec 92.188%
(92.188%)
 * Prec 89.130%
best acc: 89.760000
Epoch: [89][0/391]      Time 0.301 (0.301)      Data 0.275 (0.275)      Loss
0.0497 (0.0497)    Prec 97.656% (97.656%)
Epoch: [89][100/391]    Time 0.033 (0.038)      Data 0.001 (0.004)      Loss
0.0890 (0.0543)    Prec 96.875% (98.151%)
Epoch: [89][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0265 (0.0494)    Prec 99.219% (98.294%)
```

```
Epoch: [89][300/391]    Time 0.032 (0.036)      Data 0.001 (0.002)      Loss
0.0232 (0.0489)    Prec 99.219% (98.305%)
Validation starts
Test: [0/79]    Time 0.261 (0.261)      Loss 0.2701 (0.2701)    Prec 93.750%
(93.750%)
 * Prec 89.790%
best acc: 89.790000
Epoch: [90][0/391]      Time 0.251 (0.251)      Data 0.227 (0.227)      Loss
0.0490 (0.0490)    Prec 99.219% (99.219%)
Epoch: [90][100/391]    Time 0.033 (0.038)      Data 0.001 (0.004)      Loss
0.0224 (0.0444)    Prec 98.438% (98.538%)
Epoch: [90][200/391]    Time 0.039 (0.036)      Data 0.001 (0.003)      Loss
0.0338 (0.0449)    Prec 98.438% (98.496%)
Epoch: [90][300/391]    Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0099 (0.0444)    Prec 100.000% (98.510%)
Validation starts
Test: [0/79]    Time 0.194 (0.194)      Loss 0.2275 (0.2275)    Prec 92.188%
(92.188%)
 * Prec 89.380%
best acc: 89.790000
Epoch: [91][0/391]      Time 0.283 (0.283)      Data 0.203 (0.203)      Loss
0.0861 (0.0861)    Prec 96.094% (96.094%)
Epoch: [91][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0276 (0.0409)    Prec 99.219% (98.631%)
Epoch: [91][200/391]    Time 0.035 (0.037)      Data 0.001 (0.002)      Loss
0.0570 (0.0450)    Prec 97.656% (98.418%)
Epoch: [91][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0572 (0.0461)    Prec 97.656% (98.391%)
Validation starts
Test: [0/79]    Time 0.223 (0.223)      Loss 0.2163 (0.2163)    Prec 95.312%
(95.312%)
 * Prec 89.620%
best acc: 89.790000
Epoch: [92][0/391]      Time 0.401 (0.401)      Data 0.370 (0.370)      Loss
0.0138 (0.0138)    Prec 99.219% (99.219%)
Epoch: [92][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0586 (0.0408)    Prec 97.656% (98.623%)
Epoch: [92][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0323 (0.0416)    Prec 98.438% (98.585%)
Epoch: [92][300/391]    Time 0.036 (0.036)      Data 0.001 (0.003)      Loss
0.0263 (0.0451)    Prec 100.000% (98.456%)
Validation starts
Test: [0/79]    Time 0.245 (0.245)      Loss 0.3407 (0.3407)    Prec 92.969%
(92.969%)
 * Prec 88.930%
best acc: 89.790000
Epoch: [93][0/391]      Time 0.396 (0.396)      Data 0.363 (0.363)      Loss
0.0252 (0.0252)    Prec 100.000% (100.000%)
```

```
Epoch: [93][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0225 (0.0365)    Prec 99.219% (98.716%)
Epoch: [93][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0071 (0.0392)    Prec 100.000% (98.725%)
Epoch: [93][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0238 (0.0409)    Prec 99.219% (98.601%)
Validation starts
Test: [0/79]    Time 0.325 (0.325)      Loss 0.3852 (0.3852)    Prec 92.188%
(92.188%)
 * Prec 89.440%
best acc: 89.790000
Epoch: [94][0/391]      Time 0.298 (0.298)      Data 0.269 (0.269)      Loss
0.0679 (0.0679)    Prec 97.656% (97.656%)
Epoch: [94][100/391]    Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0686 (0.0469)    Prec 97.656% (98.399%)
Epoch: [94][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0250 (0.0441)    Prec 99.219% (98.453%)
Epoch: [94][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0524 (0.0450)    Prec 96.875% (98.450%)
Validation starts
Test: [0/79]    Time 0.207 (0.207)      Loss 0.2202 (0.2202)    Prec 95.312%
(95.312%)
 * Prec 89.460%
best acc: 89.790000
Epoch: [95][0/391]      Time 0.334 (0.334)      Data 0.305 (0.305)      Loss
0.0703 (0.0703)    Prec 96.875% (96.875%)
Epoch: [95][100/391]    Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0071 (0.0385)    Prec 100.000% (98.646%)
Epoch: [95][200/391]    Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0238 (0.0412)    Prec 99.219% (98.542%)
Epoch: [95][300/391]    Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0088 (0.0423)    Prec 100.000% (98.505%)
Validation starts
Test: [0/79]    Time 0.229 (0.229)      Loss 0.2464 (0.2464)    Prec 94.531%
(94.531%)
 * Prec 89.170%
best acc: 89.790000
Epoch: [96][0/391]      Time 0.354 (0.354)      Data 0.292 (0.292)      Loss
0.0659 (0.0659)    Prec 98.438% (98.438%)
Epoch: [96][100/391]    Time 0.035 (0.039)      Data 0.001 (0.004)      Loss
0.0320 (0.0407)    Prec 98.438% (98.569%)
Epoch: [96][200/391]    Time 0.036 (0.037)      Data 0.002 (0.003)      Loss
0.0838 (0.0425)    Prec 96.094% (98.488%)
Epoch: [96][300/391]    Time 0.036 (0.037)      Data 0.001 (0.002)      Loss
0.0522 (0.0454)    Prec 98.438% (98.373%)
Validation starts
Test: [0/79]    Time 0.131 (0.131)      Loss 0.3086 (0.3086)    Prec 92.188%
(92.188%)
```

```
 * Prec 89.500%
best acc: 89.790000
Epoch: [97][0/391]      Time 0.395 (0.395)      Data 0.365 (0.365)      Loss
0.0212 (0.0212)    Prec 99.219% (99.219%)
Epoch: [97][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0513 (0.0370)    Prec 97.656% (98.700%)
Epoch: [97][200/391]    Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0184 (0.0385)    Prec 99.219% (98.655%)
Epoch: [97][300/391]    Time 0.036 (0.036)      Data 0.001 (0.003)      Loss
0.0920 (0.0391)    Prec 95.312% (98.645%)
Validation starts
Test: [0/79]    Time 0.128 (0.128)      Loss 0.3382 (0.3382)    Prec 92.188%
(92.188%)
 * Prec 89.520%
best acc: 89.790000
Epoch: [98][0/391]      Time 0.405 (0.405)      Data 0.376 (0.376)      Loss
0.1106 (0.1106)    Prec 97.656% (97.656%)
Epoch: [98][100/391]    Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0493 (0.0399)    Prec 98.438% (98.623%)
Epoch: [98][200/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0624 (0.0400)    Prec 99.219% (98.593%)
Epoch: [98][300/391]    Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0234 (0.0399)    Prec 98.438% (98.598%)
Validation starts
Test: [0/79]    Time 0.233 (0.233)      Loss 0.2766 (0.2766)    Prec 92.188%
(92.188%)
 * Prec 89.780%
best acc: 89.790000
Epoch: [99][0/391]      Time 0.294 (0.294)      Data 0.263 (0.263)      Loss
0.0143 (0.0143)    Prec 100.000% (100.000%)
Epoch: [99][100/391]    Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.0162 (0.0374)    Prec 99.219% (98.685%)
Epoch: [99][200/391]    Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0615 (0.0352)    Prec 96.875% (98.807%)
Epoch: [99][300/391]    Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0677 (0.0373)    Prec 98.438% (98.700%)
Validation starts
Test: [0/79]    Time 0.145 (0.145)      Loss 0.2663 (0.2663)    Prec 92.969%
(92.969%)
 * Prec 89.570%
best acc: 89.790000
Epoch: [100][0/391]     Time 0.293 (0.293)      Data 0.268 (0.268)      Loss
0.0435 (0.0435)    Prec 98.438% (98.438%)
Epoch: [100][100/391]   Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0115 (0.0352)    Prec 100.000% (98.646%)
Epoch: [100][200/391]   Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0386 (0.0360)    Prec 97.656% (98.667%)
Epoch: [100][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
```

```
0.0299 (0.0364)    Prec 98.438% (98.661%)
Validation starts
Test: [0/79]    Time 0.144 (0.144)      Loss 0.2211 (0.2211)    Prec 92.969%
(92.969%)
 * Prec 89.820%
best acc: 89.820000
Epoch: [101][0/391]     Time 0.331 (0.331)      Data 0.305 (0.305)      Loss
0.0170 (0.0170)    Prec 99.219% (99.219%)
Epoch: [101][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0330 (0.0324)    Prec 98.438% (98.902%)
Epoch: [101][200/391]   Time 0.041 (0.037)      Data 0.001 (0.003)      Loss
0.0203 (0.0350)    Prec 99.219% (98.783%)
Epoch: [101][300/391]   Time 0.039 (0.036)      Data 0.001 (0.002)      Loss
0.0190 (0.0358)    Prec 99.219% (98.762%)
Validation starts
Test: [0/79]    Time 0.145 (0.145)      Loss 0.2354 (0.2354)    Prec 96.094%
(96.094%)
 * Prec 90.110%
best acc: 90.110000
Epoch: [102][0/391]     Time 0.298 (0.298)      Data 0.267 (0.267)      Loss
0.0057 (0.0057)    Prec 100.000% (100.000%)
Epoch: [102][100/391]   Time 0.038 (0.038)      Data 0.001 (0.004)      Loss
0.0042 (0.0345)    Prec 100.000% (98.755%)
Epoch: [102][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0226 (0.0343)    Prec 99.219% (98.756%)
Epoch: [102][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0062 (0.0345)    Prec 100.000% (98.736%)
Validation starts
Test: [0/79]    Time 0.094 (0.094)      Loss 0.2169 (0.2169)    Prec 93.750%
(93.750%)
 * Prec 89.740%
best acc: 90.110000
Epoch: [103][0/391]     Time 0.383 (0.383)      Data 0.350 (0.350)      Loss
0.0435 (0.0435)    Prec 97.656% (97.656%)
Epoch: [103][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0064 (0.0340)    Prec 100.000% (98.693%)
Epoch: [103][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0042 (0.0353)    Prec 100.000% (98.667%)
Epoch: [103][300/391]   Time 0.036 (0.036)      Data 0.001 (0.003)      Loss
0.0920 (0.0355)    Prec 96.875% (98.713%)
Validation starts
Test: [0/79]    Time 0.163 (0.163)      Loss 0.1887 (0.1887)    Prec 91.406%
(91.406%)
 * Prec 89.740%
best acc: 90.110000
Epoch: [104][0/391]     Time 0.315 (0.315)      Data 0.291 (0.291)      Loss
0.0179 (0.0179)    Prec 99.219% (99.219%)
Epoch: [104][100/391]   Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
```

```
0.0647 (0.0306)    Prec 97.656% (99.018%)
Epoch: [104][200/391]   Time 0.034 (0.037)      Data 0.002 (0.003)      Loss
0.0033 (0.0306)    Prec 100.000% (98.974%)
Epoch: [104][300/391]   Time 0.038 (0.036)      Data 0.001 (0.002)      Loss
0.0075 (0.0317)    Prec 100.000% (98.915%)
Validation starts
Test: [0/79]    Time 0.182 (0.182)      Loss 0.3627 (0.3627)    Prec 89.062%
(89.062%)
 * Prec 89.850%
best acc: 90.110000
Epoch: [105][0/391]     Time 0.329 (0.329)      Data 0.281 (0.281)      Loss
0.0355 (0.0355)    Prec 98.438% (98.438%)
Epoch: [105][100/391]   Time 0.037 (0.038)      Data 0.001 (0.004)      Loss
0.0170 (0.0301)    Prec 99.219% (98.948%)
Epoch: [105][200/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0790 (0.0315)    Prec 98.438% (98.908%)
Epoch: [105][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0436 (0.0307)    Prec 98.438% (98.944%)
Validation starts
Test: [0/79]    Time 0.228 (0.228)      Loss 0.2319 (0.2319)    Prec 94.531%
(94.531%)
 * Prec 90.170%
best acc: 90.170000
Epoch: [106][0/391]     Time 0.364 (0.364)      Data 0.339 (0.339)      Loss
0.0200 (0.0200)    Prec 98.438% (98.438%)
Epoch: [106][100/391]   Time 0.036 (0.038)      Data 0.002 (0.005)      Loss
0.0309 (0.0296)    Prec 99.219% (98.933%)
Epoch: [106][200/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0098 (0.0299)    Prec 100.000% (98.962%)
Epoch: [106][300/391]   Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0227 (0.0311)    Prec 99.219% (98.915%)
Validation starts
Test: [0/79]    Time 0.239 (0.239)      Loss 0.2501 (0.2501)    Prec 93.750%
(93.750%)
 * Prec 89.820%
best acc: 90.170000
Epoch: [107][0/391]     Time 0.304 (0.304)      Data 0.275 (0.275)      Loss
0.0244 (0.0244)    Prec 99.219% (99.219%)
Epoch: [107][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0297 (0.0307)    Prec 98.438% (99.025%)
Epoch: [107][200/391]   Time 0.034 (0.036)      Data 0.002 (0.003)      Loss
0.0262 (0.0322)    Prec 99.219% (98.954%)
Epoch: [107][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0347 (0.0323)    Prec 98.438% (98.918%)
Validation starts
Test: [0/79]    Time 0.149 (0.149)      Loss 0.2155 (0.2155)    Prec 92.188%
(92.188%)
 * Prec 89.980%
```

```
best acc: 90.170000
Epoch: [108][0/391]     Time 0.307 (0.307)      Data 0.276 (0.276)      Loss
0.0378 (0.0378)    Prec 97.656% (97.656%)
Epoch: [108][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0099 (0.0288)    Prec 100.000% (98.987%)
Epoch: [108][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0105 (0.0279)    Prec 100.000% (99.028%)
Epoch: [108][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0241 (0.0298)    Prec 99.219% (98.967%)
Validation starts
Test: [0/79]    Time 0.148 (0.148)      Loss 0.1964 (0.1964)    Prec 93.750%
(93.750%)
 * Prec 90.260%
best acc: 90.260000
Epoch: [109][0/391]     Time 0.198 (0.198)      Data 0.174 (0.174)      Loss
0.0208 (0.0208)    Prec 99.219% (99.219%)
Epoch: [109][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0271 (0.0279)    Prec 99.219% (98.971%)
Epoch: [109][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0603 (0.0281)    Prec 96.094% (98.966%)
Epoch: [109][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0240 (0.0291)    Prec 99.219% (98.972%)
Validation starts
Test: [0/79]    Time 0.141 (0.141)      Loss 0.1702 (0.1702)    Prec 96.094%
(96.094%)
 * Prec 90.140%
best acc: 90.260000
Epoch: [110][0/391]     Time 0.340 (0.340)      Data 0.285 (0.285)      Loss
0.0547 (0.0547)    Prec 98.438% (98.438%)
Epoch: [110][100/391]   Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0034 (0.0247)    Prec 100.000% (99.165%)
Epoch: [110][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0344 (0.0236)    Prec 99.219% (99.223%)
Epoch: [110][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0163 (0.0246)    Prec 99.219% (99.180%)
Validation starts
Test: [0/79]    Time 0.145 (0.145)      Loss 0.2635 (0.2635)    Prec 92.969%
(92.969%)
 * Prec 89.860%
best acc: 90.260000
Epoch: [111][0/391]     Time 0.393 (0.393)      Data 0.365 (0.365)      Loss
0.0064 (0.0064)    Prec 100.000% (100.000%)
Epoch: [111][100/391]   Time 0.029 (0.039)      Data 0.001 (0.005)      Loss
0.0733 (0.0269)    Prec 98.438% (99.018%)
Epoch: [111][200/391]   Time 0.029 (0.037)      Data 0.001 (0.003)      Loss
0.0302 (0.0278)    Prec 98.438% (99.001%)
Epoch: [111][300/391]   Time 0.029 (0.036)      Data 0.001 (0.003)      Loss
0.0567 (0.0283)    Prec 96.875% (99.014%)
```

```
Validation starts
Test: [0/79]    Time 0.144 (0.144)      Loss 0.3057 (0.3057)    Prec 93.750%
(93.750%)
 * Prec 90.130%
best acc: 90.260000
Epoch: [112][0/391]     Time 0.342 (0.342)      Data 0.267 (0.267)      Loss
0.0320 (0.0320)    Prec 98.438% (98.438%)
Epoch: [112][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0213 (0.0236)    Prec 99.219% (99.141%)
Epoch: [112][200/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0522 (0.0242)    Prec 96.875% (99.168%)
Epoch: [112][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0238 (0.0249)    Prec 98.438% (99.164%)
Validation starts
Test: [0/79]    Time 0.244 (0.244)      Loss 0.5111 (0.5111)    Prec 89.062%
(89.062%)
 * Prec 90.120%
best acc: 90.260000
Epoch: [113][0/391]     Time 0.218 (0.218)      Data 0.193 (0.193)      Loss
0.0866 (0.0866)    Prec 96.875% (96.875%)
Epoch: [113][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0117 (0.0258)    Prec 100.000% (99.141%)
Epoch: [113][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0434 (0.0253)    Prec 99.219% (99.125%)
Epoch: [113][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0371 (0.0255)    Prec 99.219% (99.136%)
Validation starts
Test: [0/79]    Time 0.089 (0.089)      Loss 0.2521 (0.2521)    Prec 93.750%
(93.750%)
 * Prec 90.310%
best acc: 90.310000
Epoch: [114][0/391]     Time 0.292 (0.292)      Data 0.259 (0.259)      Loss
0.0035 (0.0035)    Prec 100.000% (100.000%)
Epoch: [114][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0408 (0.0239)    Prec 99.219% (99.226%)
Epoch: [114][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0314 (0.0239)    Prec 98.438% (99.234%)
Epoch: [114][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0512 (0.0260)    Prec 99.219% (99.120%)
Validation starts
Test: [0/79]    Time 0.145 (0.145)      Loss 0.2445 (0.2445)    Prec 95.312%
(95.312%)
 * Prec 90.020%
best acc: 90.310000
Epoch: [115][0/391]     Time 0.313 (0.313)      Data 0.284 (0.284)      Loss
0.0096 (0.0096)    Prec 100.000% (100.000%)
Epoch: [115][100/391]   Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0190 (0.0256)    Prec 99.219% (99.103%)
```

```
Epoch: [115][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0094 (0.0235)    Prec 99.219% (99.168%)
Epoch: [115][300/391]   Time 0.036 (0.036)      Data 0.002 (0.002)      Loss
0.0119 (0.0238)    Prec 99.219% (99.167%)
Validation starts
Test: [0/79]    Time 0.132 (0.132)      Loss 0.3636 (0.3636)    Prec 92.188%
(92.188%)
 * Prec 90.000%
best acc: 90.310000
Epoch: [116][0/391]     Time 0.218 (0.218)      Data 0.192 (0.192)      Loss
0.0127 (0.0127)    Prec 99.219% (99.219%)
Epoch: [116][100/391]   Time 0.036 (0.037)      Data 0.002 (0.003)      Loss
0.0024 (0.0209)    Prec 100.000% (99.180%)
Epoch: [116][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0680 (0.0228)    Prec 98.438% (99.230%)
Epoch: [116][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0172 (0.0231)    Prec 99.219% (99.203%)
Validation starts
Test: [0/79]    Time 0.154 (0.154)      Loss 0.1946 (0.1946)    Prec 96.094%
(96.094%)
 * Prec 90.470%
best acc: 90.470000
Epoch: [117][0/391]     Time 0.218 (0.218)      Data 0.192 (0.192)      Loss
0.0107 (0.0107)    Prec 100.000% (100.000%)
Epoch: [117][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0065 (0.0186)    Prec 100.000% (99.412%)
Epoch: [117][200/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0315 (0.0185)    Prec 98.438% (99.390%)
Epoch: [117][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0012 (0.0189)    Prec 100.000% (99.390%)
Validation starts
Test: [0/79]    Time 0.157 (0.157)      Loss 0.1717 (0.1717)    Prec 95.312%
(95.312%)
 * Prec 90.300%
best acc: 90.470000
Epoch: [118][0/391]     Time 0.403 (0.403)      Data 0.375 (0.375)      Loss
0.0366 (0.0366)    Prec 99.219% (99.219%)
Epoch: [118][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0040 (0.0213)    Prec 100.000% (99.242%)
Epoch: [118][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0067 (0.0220)    Prec 100.000% (99.188%)
Epoch: [118][300/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0754 (0.0216)    Prec 96.875% (99.214%)
Validation starts
Test: [0/79]    Time 0.230 (0.230)      Loss 0.2753 (0.2753)    Prec 95.312%
(95.312%)
 * Prec 90.230%
best acc: 90.470000
```

```
Epoch: [119][0/391]     Time 0.301 (0.301)      Data 0.277 (0.277)      Loss
0.0013 (0.0013)    Prec 100.000% (100.000%)
Epoch: [119][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0185 (0.0197)    Prec 99.219% (99.288%)
Epoch: [119][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0588 (0.0198)    Prec 97.656% (99.293%)
Epoch: [119][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0127 (0.0193)    Prec 99.219% (99.333%)
Validation starts
Test: [0/79]    Time 0.149 (0.149)      Loss 0.3163 (0.3163)    Prec 92.188%
(92.188%)
 * Prec 90.600%
best acc: 90.600000
Epoch: [120][0/391]     Time 0.242 (0.242)      Data 0.214 (0.214)      Loss
0.0043 (0.0043)    Prec 100.000% (100.000%)
Epoch: [120][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0047 (0.0158)    Prec 100.000% (99.404%)
Epoch: [120][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0051 (0.0177)    Prec 100.000% (99.359%)
Epoch: [120][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0426 (0.0204)    Prec 97.656% (99.252%)
Validation starts
Test: [0/79]    Time 0.136 (0.136)      Loss 0.2766 (0.2766)    Prec 93.750%
(93.750%)
 * Prec 90.310%
best acc: 90.600000
Epoch: [121][0/391]     Time 0.248 (0.248)      Data 0.191 (0.191)      Loss
0.0236 (0.0236)    Prec 99.219% (99.219%)
Epoch: [121][100/391]   Time 0.035 (0.038)      Data 0.001 (0.003)      Loss
0.0594 (0.0215)    Prec 98.438% (99.350%)
Epoch: [121][200/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0133 (0.0193)    Prec 100.000% (99.355%)
Epoch: [121][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0786 (0.0201)    Prec 97.656% (99.315%)
Validation starts
Test: [0/79]    Time 0.145 (0.145)      Loss 0.2802 (0.2802)    Prec 92.969%
(92.969%)
 * Prec 90.280%
best acc: 90.600000
Epoch: [122][0/391]     Time 0.299 (0.299)      Data 0.273 (0.273)      Loss
0.0055 (0.0055)    Prec 100.000% (100.000%)
Epoch: [122][100/391]   Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0124 (0.0158)    Prec 100.000% (99.459%)
Epoch: [122][200/391]   Time 0.037 (0.036)      Data 0.001 (0.003)      Loss
0.0174 (0.0203)    Prec 100.000% (99.308%)
Epoch: [122][300/391]   Time 0.037 (0.036)      Data 0.002 (0.002)      Loss
0.0142 (0.0210)    Prec 99.219% (99.258%)
Validation starts
```

```
Test: [0/79]    Time 0.137 (0.137)      Loss 0.4020 (0.4020)    Prec 92.188%
(92.188%)
 * Prec 90.220%
best acc: 90.600000
Epoch: [123][0/391]     Time 0.293 (0.293)      Data 0.215 (0.215)      Loss
0.0012 (0.0012)    Prec 100.000% (100.000%)
Epoch: [123][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0163 (0.0157)    Prec 99.219% (99.482%)
Epoch: [123][200/391]   Time 0.036 (0.037)      Data 0.001 (0.002)      Loss
0.0178 (0.0169)    Prec 99.219% (99.444%)
Epoch: [123][300/391]   Time 0.032 (0.036)      Data 0.001 (0.002)      Loss
0.0584 (0.0180)    Prec 96.875% (99.390%)
Validation starts
Test: [0/79]    Time 0.087 (0.087)      Loss 0.2664 (0.2664)    Prec 94.531%
(94.531%)
 * Prec 90.320%
best acc: 90.600000
Epoch: [124][0/391]     Time 0.244 (0.244)      Data 0.218 (0.218)      Loss
0.0012 (0.0012)    Prec 100.000% (100.000%)
Epoch: [124][100/391]   Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.0018 (0.0152)    Prec 100.000% (99.505%)
Epoch: [124][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0247 (0.0167)    Prec 98.438% (99.417%)
Epoch: [124][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0151 (0.0176)    Prec 99.219% (99.395%)
Validation starts
Test: [0/79]    Time 0.155 (0.155)      Loss 0.2840 (0.2840)    Prec 92.188%
(92.188%)
 * Prec 90.020%
best acc: 90.600000
Epoch: [125][0/391]     Time 0.293 (0.293)      Data 0.268 (0.268)      Loss
0.0069 (0.0069)    Prec 100.000% (100.000%)
Epoch: [125][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0128 (0.0154)    Prec 99.219% (99.327%)
Epoch: [125][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0147 (0.0164)    Prec 99.219% (99.378%)
Epoch: [125][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0220 (0.0162)    Prec 98.438% (99.426%)
Validation starts
Test: [0/79]    Time 0.238 (0.238)      Loss 0.2704 (0.2704)    Prec 95.312%
(95.312%)
 * Prec 90.080%
best acc: 90.600000
Epoch: [126][0/391]     Time 0.391 (0.391)      Data 0.364 (0.364)      Loss
0.0043 (0.0043)    Prec 100.000% (100.000%)
Epoch: [126][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0101 (0.0155)    Prec 99.219% (99.404%)
Epoch: [126][200/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
```

```
0.0050 (0.0181)    Prec 100.000% (99.339%)
Epoch: [126][300/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0075 (0.0181)    Prec 100.000% (99.364%)
Validation starts
Test: [0/79]    Time 0.102 (0.102)      Loss 0.2283 (0.2283)    Prec 94.531%
(94.531%)
 * Prec 90.320%
best acc: 90.600000
Epoch: [127][0/391]     Time 0.242 (0.242)      Data 0.213 (0.213)      Loss
0.0161 (0.0161)    Prec 99.219% (99.219%)
Epoch: [127][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0388 (0.0178)    Prec 98.438% (99.366%)
Epoch: [127][200/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0276 (0.0174)    Prec 99.219% (99.394%)
Epoch: [127][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0048 (0.0174)    Prec 100.000% (99.395%)
Validation starts
Test: [0/79]    Time 0.138 (0.138)      Loss 0.2505 (0.2505)    Prec 92.969%
(92.969%)
 * Prec 90.330%
best acc: 90.600000
Epoch: [128][0/391]     Time 0.390 (0.390)      Data 0.362 (0.362)      Loss
0.0092 (0.0092)    Prec 100.000% (100.000%)
Epoch: [128][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0012 (0.0156)    Prec 100.000% (99.428%)
Epoch: [128][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0035 (0.0147)    Prec 100.000% (99.460%)
Epoch: [128][300/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0242 (0.0148)    Prec 99.219% (99.494%)
Validation starts
Test: [0/79]    Time 0.230 (0.230)      Loss 0.2586 (0.2586)    Prec 93.750%
(93.750%)
 * Prec 90.240%
best acc: 90.600000
Epoch: [129][0/391]     Time 0.308 (0.308)      Data 0.283 (0.283)      Loss
0.0276 (0.0276)    Prec 99.219% (99.219%)
Epoch: [129][100/391]   Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0037 (0.0167)    Prec 100.000% (99.482%)
Epoch: [129][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0301 (0.0169)    Prec 99.219% (99.464%)
Epoch: [129][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0095 (0.0160)    Prec 100.000% (99.483%)
Validation starts
Test: [0/79]    Time 0.139 (0.139)      Loss 0.3463 (0.3463)    Prec 93.750%
(93.750%)
 * Prec 90.540%
best acc: 90.600000
Epoch: [130][0/391]     Time 0.307 (0.307)      Data 0.280 (0.280)      Loss
```

```
0.0069 (0.0069)    Prec 100.000% (100.000%)
Epoch: [130][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0232 (0.0151)    Prec 99.219% (99.505%)
Epoch: [130][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0050 (0.0143)    Prec 100.000% (99.530%)
Epoch: [130][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0155 (0.0147)    Prec 99.219% (99.502%)
Validation starts
Test: [0/79]    Time 0.151 (0.151)      Loss 0.2482 (0.2482)    Prec 95.312%
(95.312%)
 * Prec 90.710%
best acc: 90.710000
Epoch: [131][0/391]     Time 0.229 (0.229)      Data 0.202 (0.202)      Loss
0.0029 (0.0029)    Prec 100.000% (100.000%)
Epoch: [131][100/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0106 (0.0123)    Prec 99.219% (99.567%)
Epoch: [131][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0034 (0.0131)    Prec 100.000% (99.522%)
Epoch: [131][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0119 (0.0142)    Prec 99.219% (99.496%)
Validation starts
Test: [0/79]    Time 0.122 (0.122)      Loss 0.2375 (0.2375)    Prec 96.094%
(96.094%)
 * Prec 90.610%
best acc: 90.710000
Epoch: [132][0/391]     Time 0.305 (0.305)      Data 0.279 (0.279)      Loss
0.0204 (0.0204)    Prec 98.438% (98.438%)
Epoch: [132][100/391]   Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0163 (0.0123)    Prec 99.219% (99.544%)
Epoch: [132][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0521 (0.0132)    Prec 98.438% (99.565%)
Epoch: [132][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0060 (0.0139)    Prec 100.000% (99.551%)
Validation starts
Test: [0/79]    Time 0.233 (0.233)      Loss 0.1972 (0.1972)    Prec 94.531%
(94.531%)
 * Prec 90.360%
best acc: 90.710000
Epoch: [133][0/391]     Time 0.279 (0.279)      Data 0.247 (0.247)      Loss
0.0211 (0.0211)    Prec 99.219% (99.219%)
Epoch: [133][100/391]   Time 0.039 (0.038)      Data 0.001 (0.004)      Loss
0.0067 (0.0109)    Prec 100.000% (99.660%)
Epoch: [133][200/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0008 (0.0124)    Prec 100.000% (99.580%)
Epoch: [133][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0717 (0.0123)    Prec 98.438% (99.593%)
Validation starts
Test: [0/79]    Time 0.135 (0.135)      Loss 0.2512 (0.2512)    Prec 96.094%
```

```
(96.094%)
 * Prec 90.840%
best acc: 90.840000
Epoch: [134][0/391]     Time 0.244 (0.244)      Data 0.216 (0.216)      Loss
0.0182 (0.0182)    Prec 99.219% (99.219%)
Epoch: [134][100/391]   Time 0.035 (0.037)      Data 0.001 (0.004)      Loss
0.0045 (0.0130)    Prec 100.000% (99.598%)
Epoch: [134][200/391]   Time 0.035 (0.036)      Data 0.002 (0.003)      Loss
0.0160 (0.0126)    Prec 99.219% (99.596%)
Epoch: [134][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0028 (0.0120)    Prec 100.000% (99.598%)
Validation starts
Test: [0/79]    Time 0.145 (0.145)      Loss 0.2709 (0.2709)    Prec 94.531%
(94.531%)
 * Prec 90.760%
best acc: 90.840000
Epoch: [135][0/391]     Time 0.373 (0.373)      Data 0.337 (0.337)      Loss
0.0021 (0.0021)    Prec 100.000% (100.000%)
Epoch: [135][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0240 (0.0114)    Prec 99.219% (99.613%)
Epoch: [135][200/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0055 (0.0126)    Prec 100.000% (99.584%)
Epoch: [135][300/391]   Time 0.035 (0.037)      Data 0.001 (0.002)      Loss
0.0028 (0.0120)    Prec 100.000% (99.600%)
Validation starts
Test: [0/79]    Time 0.224 (0.224)      Loss 0.3354 (0.3354)    Prec 93.750%
(93.750%)
 * Prec 90.520%
best acc: 90.840000
Epoch: [136][0/391]     Time 0.296 (0.296)      Data 0.271 (0.271)      Loss
0.0067 (0.0067)    Prec 99.219% (99.219%)
Epoch: [136][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0063 (0.0128)    Prec 99.219% (99.559%)
Epoch: [136][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0395 (0.0127)    Prec 99.219% (99.580%)
Epoch: [136][300/391]   Time 0.036 (0.036)      Data 0.002 (0.002)      Loss
0.0296 (0.0130)    Prec 99.219% (99.561%)
Validation starts
Test: [0/79]    Time 0.139 (0.139)      Loss 0.2294 (0.2294)    Prec 94.531%
(94.531%)
 * Prec 90.370%
best acc: 90.840000
Epoch: [137][0/391]     Time 0.283 (0.283)      Data 0.253 (0.253)      Loss
0.0014 (0.0014)    Prec 100.000% (100.000%)
Epoch: [137][100/391]   Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0027 (0.0113)    Prec 100.000% (99.582%)
Epoch: [137][200/391]   Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0102 (0.0111)    Prec 100.000% (99.588%)
```

```
Epoch: [137][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0246 (0.0115)    Prec 99.219% (99.582%)
Validation starts
Test: [0/79]    Time 0.236 (0.236)      Loss 0.2300 (0.2300)    Prec 94.531%
(94.531%)
 * Prec 90.550%
best acc: 90.840000
Epoch: [138][0/391]     Time 0.216 (0.216)      Data 0.190 (0.190)      Loss
0.0051 (0.0051)    Prec 100.000% (100.000%)
Epoch: [138][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0031 (0.0092)    Prec 100.000% (99.652%)
Epoch: [138][200/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0023 (0.0096)    Prec 100.000% (99.654%)
Epoch: [138][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0097 (0.0099)    Prec 100.000% (99.642%)
Validation starts
Test: [0/79]    Time 0.087 (0.087)      Loss 0.2450 (0.2450)    Prec 94.531%
(94.531%)
 * Prec 90.620%
best acc: 90.840000
Epoch: [139][0/391]     Time 0.315 (0.315)      Data 0.238 (0.238)      Loss
0.0717 (0.0717)    Prec 97.656% (97.656%)
Epoch: [139][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0081 (0.0111)    Prec 99.219% (99.567%)
Epoch: [139][200/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0074 (0.0102)    Prec 99.219% (99.627%)
Epoch: [139][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0014 (0.0105)    Prec 100.000% (99.639%)
Validation starts
Test: [0/79]    Time 0.140 (0.140)      Loss 0.2479 (0.2479)    Prec 92.969%
(92.969%)
 * Prec 90.860%
best acc: 90.860000
Epoch: [140][0/391]     Time 0.234 (0.234)      Data 0.205 (0.205)      Loss
0.0017 (0.0017)    Prec 100.000% (100.000%)
Epoch: [140][100/391]   Time 0.035 (0.037)      Data 0.001 (0.004)      Loss
0.0018 (0.0105)    Prec 100.000% (99.629%)
Epoch: [140][200/391]   Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0024 (0.0102)    Prec 100.000% (99.670%)
Epoch: [140][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0097 (0.0101)    Prec 99.219% (99.668%)
Validation starts
Test: [0/79]    Time 0.087 (0.087)      Loss 0.2428 (0.2428)    Prec 91.406%
(91.406%)
 * Prec 90.960%
best acc: 90.960000
Epoch: [141][0/391]     Time 0.288 (0.288)      Data 0.261 (0.261)      Loss
0.0166 (0.0166)    Prec 99.219% (99.219%)
```

```
Epoch: [141][100/391]   Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.0115 (0.0065)    Prec 99.219% (99.745%)
Epoch: [141][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0019 (0.0089)    Prec 100.000% (99.639%)
Epoch: [141][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0019 (0.0095)    Prec 100.000% (99.637%)
Validation starts
Test: [0/79]    Time 0.152 (0.152)      Loss 0.2689 (0.2689)    Prec 93.750%
(93.750%)
 * Prec 90.500%
best acc: 90.960000
Epoch: [142][0/391]     Time 0.294 (0.294)      Data 0.269 (0.269)      Loss
0.0005 (0.0005)    Prec 100.000% (100.000%)
Epoch: [142][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0040 (0.0097)    Prec 100.000% (99.714%)
Epoch: [142][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0018 (0.0104)    Prec 100.000% (99.670%)
Epoch: [142][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0024 (0.0100)    Prec 100.000% (99.676%)
Validation starts
Test: [0/79]    Time 0.239 (0.239)      Loss 0.3178 (0.3178)    Prec 94.531%
(94.531%)
 * Prec 90.650%
best acc: 90.960000
Epoch: [143][0/391]     Time 0.412 (0.412)      Data 0.383 (0.383)      Loss
0.0209 (0.0209)    Prec 98.438% (98.438%)
Epoch: [143][100/391]   Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.0043 (0.0107)    Prec 100.000% (99.606%)
Epoch: [143][200/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0056 (0.0108)    Prec 100.000% (99.604%)
Epoch: [143][300/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0037 (0.0102)    Prec 100.000% (99.637%)
Validation starts
Test: [0/79]    Time 0.241 (0.241)      Loss 0.3403 (0.3403)    Prec 93.750%
(93.750%)
 * Prec 90.780%
best acc: 90.960000
Epoch: [144][0/391]     Time 0.300 (0.300)      Data 0.273 (0.273)      Loss
0.0009 (0.0009)    Prec 100.000% (100.000%)
Epoch: [144][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0007 (0.0060)    Prec 100.000% (99.760%)
Epoch: [144][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0043 (0.0073)    Prec 100.000% (99.743%)
Epoch: [144][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0382 (0.0076)    Prec 99.219% (99.743%)
Validation starts
Test: [0/79]    Time 0.247 (0.247)      Loss 0.1881 (0.1881)    Prec 95.312%
(95.312%)
```

```
 * Prec 90.800%
best acc: 90.960000
Epoch: [145][0/391]     Time 0.216 (0.216)      Data 0.189 (0.189)      Loss
0.0096 (0.0096)    Prec 99.219% (99.219%)
Epoch: [145][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0010 (0.0089)    Prec 100.000% (99.722%)
Epoch: [145][200/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0022 (0.0086)    Prec 100.000% (99.716%)
Epoch: [145][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0023 (0.0086)    Prec 100.000% (99.686%)
Validation starts
Test: [0/79]    Time 0.140 (0.140)      Loss 0.3109 (0.3109)    Prec 92.969%
(92.969%)
 * Prec 90.870%
best acc: 90.960000
Epoch: [146][0/391]     Time 0.397 (0.397)      Data 0.370 (0.370)      Loss
0.0015 (0.0015)    Prec 100.000% (100.000%)
Epoch: [146][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0147 (0.0065)    Prec 99.219% (99.768%)
Epoch: [146][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0004 (0.0069)    Prec 100.000% (99.763%)
Epoch: [146][300/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0016 (0.0073)    Prec 100.000% (99.735%)
Validation starts
Test: [0/79]    Time 0.145 (0.145)      Loss 0.2753 (0.2753)    Prec 92.188%
(92.188%)
 * Prec 90.680%
best acc: 90.960000
Epoch: [147][0/391]     Time 0.408 (0.408)      Data 0.383 (0.383)      Loss
0.0728 (0.0728)    Prec 99.219% (99.219%)
Epoch: [147][100/391]   Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.0037 (0.0089)    Prec 100.000% (99.722%)
Epoch: [147][200/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0016 (0.0081)    Prec 100.000% (99.724%)
Epoch: [147][300/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0007 (0.0083)    Prec 100.000% (99.709%)
Validation starts
Test: [0/79]    Time 0.093 (0.093)      Loss 0.2579 (0.2579)    Prec 92.969%
(92.969%)
 * Prec 90.890%
best acc: 90.960000
Epoch: [148][0/391]     Time 0.244 (0.244)      Data 0.216 (0.216)      Loss
0.0077 (0.0077)    Prec 99.219% (99.219%)
Epoch: [148][100/391]   Time 0.035 (0.037)      Data 0.001 (0.004)      Loss
0.0062 (0.0083)    Prec 100.000% (99.714%)
Epoch: [148][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0020 (0.0086)    Prec 100.000% (99.712%)
Epoch: [148][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
```

```
0.0058 (0.0084)    Prec 100.000% (99.720%)
Validation starts
Test: [0/79]    Time 0.099 (0.099)      Loss 0.2485 (0.2485)    Prec 96.094%
(96.094%)
 * Prec 90.800%
best acc: 90.960000
Epoch: [149][0/391]     Time 0.234 (0.234)      Data 0.209 (0.209)      Loss
0.0097 (0.0097)    Prec 99.219% (99.219%)
Epoch: [149][100/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0082 (0.0061)    Prec 99.219% (99.760%)
Epoch: [149][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0005 (0.0062)    Prec 100.000% (99.767%)
Epoch: [149][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0012 (0.0065)    Prec 100.000% (99.769%)
Validation starts
Test: [0/79]    Time 0.145 (0.145)      Loss 0.2306 (0.2306)    Prec 94.531%
(94.531%)
 * Prec 90.880%
best acc: 90.960000
Epoch: [150][0/391]     Time 0.306 (0.306)      Data 0.276 (0.276)      Loss
0.0004 (0.0004)    Prec 100.000% (100.000%)
Epoch: [150][100/391]   Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0048 (0.0065)    Prec 100.000% (99.776%)
Epoch: [150][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0080 (0.0066)    Prec 100.000% (99.771%)
Epoch: [150][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0068 (0.0068)    Prec 100.000% (99.769%)
Validation starts
Test: [0/79]    Time 0.192 (0.192)      Loss 0.1812 (0.1812)    Prec 97.656%
(97.656%)
 * Prec 90.890%
best acc: 90.960000
Epoch: [151][0/391]     Time 0.236 (0.236)      Data 0.204 (0.204)      Loss
0.0263 (0.0263)    Prec 98.438% (98.438%)
Epoch: [151][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0087 (0.0082)    Prec 99.219% (99.698%)
Epoch: [151][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0004 (0.0077)    Prec 100.000% (99.728%)
Epoch: [151][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0127 (0.0075)    Prec 99.219% (99.735%)
Validation starts
Test: [0/79]    Time 0.140 (0.140)      Loss 0.2626 (0.2626)    Prec 94.531%
(94.531%)
 * Prec 90.890%
best acc: 90.960000
Epoch: [152][0/391]     Time 0.211 (0.211)      Data 0.188 (0.188)      Loss
0.0087 (0.0087)    Prec 99.219% (99.219%)
Epoch: [152][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
```

```
0.0081 (0.0059)    Prec 100.000% (99.799%)
Epoch: [152][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0012 (0.0055)    Prec 100.000% (99.806%)
Epoch: [152][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0120 (0.0062)    Prec 99.219% (99.774%)
Validation starts
Test: [0/79]    Time 0.135 (0.135)      Loss 0.2112 (0.2112)    Prec 96.094%
(96.094%)
 * Prec 90.750%
best acc: 90.960000
Epoch: [153][0/391]     Time 0.293 (0.293)      Data 0.264 (0.264)      Loss
0.0350 (0.0350)    Prec 99.219% (99.219%)
Epoch: [153][100/391]   Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.0041 (0.0075)    Prec 100.000% (99.722%)
Epoch: [153][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0013 (0.0069)    Prec 100.000% (99.724%)
Epoch: [153][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0069 (0.0064)    Prec 99.219% (99.759%)
Validation starts
Test: [0/79]    Time 0.144 (0.144)      Loss 0.2425 (0.2425)    Prec 94.531%
(94.531%)
 * Prec 90.660%
best acc: 90.960000
Epoch: [154][0/391]     Time 0.313 (0.313)      Data 0.281 (0.281)      Loss
0.0033 (0.0033)    Prec 100.000% (100.000%)
Epoch: [154][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0073 (0.0064)    Prec 100.000% (99.791%)
Epoch: [154][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0018 (0.0063)    Prec 100.000% (99.798%)
Epoch: [154][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0002 (0.0066)    Prec 100.000% (99.785%)
Validation starts
Test: [0/79]    Time 0.148 (0.148)      Loss 0.2420 (0.2420)    Prec 95.312%
(95.312%)
 * Prec 91.000%
best acc: 91.000000
Epoch: [155][0/391]     Time 0.277 (0.277)      Data 0.254 (0.254)      Loss
0.0018 (0.0018)    Prec 100.000% (100.000%)
Epoch: [155][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0056 (0.0060)    Prec 100.000% (99.799%)
Epoch: [155][200/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0002 (0.0067)    Prec 100.000% (99.782%)
Epoch: [155][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0011 (0.0062)    Prec 100.000% (99.800%)
Validation starts
Test: [0/79]    Time 0.225 (0.225)      Loss 0.2279 (0.2279)    Prec 94.531%
(94.531%)
 * Prec 90.840%
```

```
best acc: 91.000000
Epoch: [156][0/391]     Time 0.309 (0.309)      Data 0.282 (0.282)      Loss
0.0015 (0.0015)    Prec 100.000% (100.000%)
Epoch: [156][100/391]   Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.0011 (0.0064)    Prec 100.000% (99.822%)
Epoch: [156][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0150 (0.0059)    Prec 99.219% (99.817%)
Epoch: [156][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0014 (0.0057)    Prec 100.000% (99.829%)
Validation starts
Test: [0/79]    Time 0.129 (0.129)      Loss 0.1937 (0.1937)    Prec 95.312%
(95.312%)
 * Prec 90.710%
best acc: 91.000000
Epoch: [157][0/391]     Time 0.296 (0.296)      Data 0.272 (0.272)      Loss
0.0021 (0.0021)    Prec 100.000% (100.000%)
Epoch: [157][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0019 (0.0058)    Prec 100.000% (99.799%)
Epoch: [157][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0018 (0.0048)    Prec 100.000% (99.852%)
Epoch: [157][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0019 (0.0049)    Prec 100.000% (99.849%)
Validation starts
Test: [0/79]    Time 0.217 (0.217)      Loss 0.2199 (0.2199)    Prec 96.094%
(96.094%)
 * Prec 90.980%
best acc: 91.000000
Epoch: [158][0/391]     Time 0.290 (0.290)      Data 0.264 (0.264)      Loss
0.0012 (0.0012)    Prec 100.000% (100.000%)
Epoch: [158][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0006 (0.0058)    Prec 100.000% (99.814%)
Epoch: [158][200/391]   Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0138 (0.0070)    Prec 100.000% (99.782%)
Epoch: [158][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0025 (0.0063)    Prec 100.000% (99.803%)
Validation starts
Test: [0/79]    Time 0.240 (0.240)      Loss 0.2311 (0.2311)    Prec 94.531%
(94.531%)
 * Prec 90.640%
best acc: 91.000000
Epoch: [159][0/391]     Time 0.314 (0.314)      Data 0.285 (0.285)      Loss
0.0007 (0.0007)    Prec 100.000% (100.000%)
Epoch: [159][100/391]   Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.0009 (0.0060)    Prec 100.000% (99.783%)
Epoch: [159][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0005 (0.0057)    Prec 100.000% (99.798%)
Epoch: [159][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0407 (0.0057)    Prec 99.219% (99.798%)
```

```
Validation starts
Test: [0/79]    Time 0.152 (0.152)      Loss 0.2457 (0.2457)    Prec 92.969%
(92.969%)
 * Prec 90.760%
best acc: 91.000000
Epoch: [160][0/391]     Time 0.296 (0.296)      Data 0.272 (0.272)      Loss
0.0001 (0.0001)    Prec 100.000% (100.000%)
Epoch: [160][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0127 (0.0047)    Prec 99.219% (99.838%)
Epoch: [160][200/391]   Time 0.035 (0.036)      Data 0.002 (0.003)      Loss
0.0064 (0.0053)    Prec 100.000% (99.825%)
Epoch: [160][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0006 (0.0050)    Prec 100.000% (99.829%)
Validation starts
Test: [0/79]    Time 0.239 (0.239)      Loss 0.1847 (0.1847)    Prec 96.875%
(96.875%)
 * Prec 90.980%
best acc: 91.000000
Epoch: [161][0/391]     Time 0.331 (0.331)      Data 0.307 (0.307)      Loss
0.0005 (0.0005)    Prec 100.000% (100.000%)
Epoch: [161][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0002 (0.0063)    Prec 100.000% (99.814%)
Epoch: [161][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0052 (0.0053)    Prec 100.000% (99.833%)
Epoch: [161][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0014 (0.0050)    Prec 100.000% (99.839%)
Validation starts
Test: [0/79]    Time 0.140 (0.140)      Loss 0.2432 (0.2432)    Prec 94.531%
(94.531%)
 * Prec 90.950%
best acc: 91.000000
Epoch: [162][0/391]     Time 0.296 (0.296)      Data 0.273 (0.273)      Loss
0.0007 (0.0007)    Prec 100.000% (100.000%)
Epoch: [162][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0014 (0.0037)    Prec 100.000% (99.876%)
Epoch: [162][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0005 (0.0049)    Prec 100.000% (99.837%)
Epoch: [162][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0005 (0.0050)    Prec 100.000% (99.839%)
Validation starts
Test: [0/79]    Time 0.120 (0.120)      Loss 0.1787 (0.1787)    Prec 96.094%
(96.094%)
 * Prec 90.820%
best acc: 91.000000
Epoch: [163][0/391]     Time 0.310 (0.310)      Data 0.285 (0.285)      Loss
0.0082 (0.0082)    Prec 99.219% (99.219%)
Epoch: [163][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0090 (0.0054)    Prec 99.219% (99.807%)
```

```
Epoch: [163][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0003 (0.0045)    Prec 100.000% (99.837%)
Epoch: [163][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0274 (0.0042)    Prec 99.219% (99.855%)
Validation starts
Test: [0/79]    Time 0.135 (0.135)      Loss 0.1751 (0.1751)    Prec 96.875%
(96.875%)
 * Prec 91.210%
best acc: 91.210000
Epoch: [164][0/391]     Time 0.329 (0.329)      Data 0.305 (0.305)      Loss
0.0002 (0.0002)    Prec 100.000% (100.000%)
Epoch: [164][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0029 (0.0050)    Prec 100.000% (99.845%)
Epoch: [164][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0004 (0.0051)    Prec 100.000% (99.837%)
Epoch: [164][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0019 (0.0052)    Prec 100.000% (99.829%)
Validation starts
Test: [0/79]    Time 0.142 (0.142)      Loss 0.2089 (0.2089)    Prec 96.875%
(96.875%)
 * Prec 90.970%
best acc: 91.210000
Epoch: [165][0/391]     Time 0.377 (0.377)      Data 0.351 (0.351)      Loss
0.0005 (0.0005)    Prec 100.000% (100.000%)
Epoch: [165][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0006 (0.0052)    Prec 100.000% (99.814%)
Epoch: [165][200/391]   Time 0.035 (0.037)      Data 0.008 (0.003)      Loss
0.0466 (0.0051)    Prec 98.438% (99.813%)
Epoch: [165][300/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0001 (0.0052)    Prec 100.000% (99.811%)
Validation starts
Test: [0/79]    Time 0.133 (0.133)      Loss 0.1899 (0.1899)    Prec 95.312%
(95.312%)
 * Prec 91.130%
best acc: 91.210000
Epoch: [166][0/391]     Time 0.341 (0.341)      Data 0.274 (0.274)      Loss
0.0022 (0.0022)    Prec 100.000% (100.000%)
Epoch: [166][100/391]   Time 0.034 (0.039)      Data 0.001 (0.004)      Loss
0.0001 (0.0042)    Prec 100.000% (99.884%)
Epoch: [166][200/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0005 (0.0040)    Prec 100.000% (99.891%)
Epoch: [166][300/391]   Time 0.035 (0.037)      Data 0.001 (0.002)      Loss
0.0008 (0.0038)    Prec 100.000% (99.888%)
Validation starts
Test: [0/79]    Time 0.151 (0.151)      Loss 0.2302 (0.2302)    Prec 95.312%
(95.312%)
 * Prec 90.930%
best acc: 91.210000
```

```
Epoch: [167][0/391]     Time 0.233 (0.233)      Data 0.209 (0.209)      Loss
0.0024 (0.0024)    Prec 100.000% (100.000%)
Epoch: [167][100/391]   Time 0.038 (0.037)      Data 0.001 (0.004)      Loss
0.0076 (0.0040)    Prec 100.000% (99.876%)
Epoch: [167][200/391]   Time 0.038 (0.036)      Data 0.001 (0.002)      Loss
0.0003 (0.0041)    Prec 100.000% (99.872%)
Epoch: [167][300/391]   Time 0.041 (0.036)      Data 0.001 (0.002)      Loss
0.0021 (0.0040)    Prec 100.000% (99.873%)
Validation starts
Test: [0/79]    Time 0.136 (0.136)      Loss 0.2612 (0.2612)    Prec 93.750%
(93.750%)
 * Prec 91.020%
best acc: 91.210000
Epoch: [168][0/391]     Time 0.220 (0.220)      Data 0.193 (0.193)      Loss
0.0003 (0.0003)    Prec 100.000% (100.000%)
Epoch: [168][100/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0004 (0.0031)    Prec 100.000% (99.930%)
Epoch: [168][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0001 (0.0037)    Prec 100.000% (99.891%)
Epoch: [168][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0037 (0.0043)    Prec 100.000% (99.870%)
Validation starts
Test: [0/79]    Time 0.130 (0.130)      Loss 0.1773 (0.1773)    Prec 95.312%
(95.312%)
 * Prec 90.940%
best acc: 91.210000
Epoch: [169][0/391]     Time 0.237 (0.237)      Data 0.210 (0.210)      Loss
0.0004 (0.0004)    Prec 100.000% (100.000%)
Epoch: [169][100/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0022 (0.0045)    Prec 100.000% (99.845%)
Epoch: [169][200/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0009 (0.0036)    Prec 100.000% (99.880%)
Epoch: [169][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0194 (0.0044)    Prec 99.219% (99.849%)
Validation starts
Test: [0/79]    Time 0.153 (0.153)      Loss 0.2031 (0.2031)    Prec 96.094%
(96.094%)
 * Prec 90.960%
best acc: 91.210000
Epoch: [170][0/391]     Time 0.355 (0.355)      Data 0.278 (0.278)      Loss
0.0008 (0.0008)    Prec 100.000% (100.000%)
Epoch: [170][100/391]   Time 0.041 (0.039)      Data 0.001 (0.004)      Loss
0.0036 (0.0033)    Prec 100.000% (99.899%)
Epoch: [170][200/391]   Time 0.041 (0.037)      Data 0.001 (0.003)      Loss
0.0042 (0.0044)    Prec 100.000% (99.860%)
Epoch: [170][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0109 (0.0041)    Prec 99.219% (99.868%)
Validation starts
```

```
Test: [0/79]    Time 0.221 (0.221)      Loss 0.1738 (0.1738)    Prec 96.875%
(96.875%)
 * Prec 90.960%
best acc: 91.210000
Epoch: [171][0/391]     Time 0.215 (0.215)      Data 0.191 (0.191)      Loss
0.0005 (0.0005)    Prec 100.000% (100.000%)
Epoch: [171][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0002 (0.0025)    Prec 100.000% (99.923%)
Epoch: [171][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0010 (0.0031)    Prec 100.000% (99.903%)
Epoch: [171][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0436 (0.0037)    Prec 99.219% (99.873%)
Validation starts
Test: [0/79]    Time 0.167 (0.167)      Loss 0.1833 (0.1833)    Prec 95.312%
(95.312%)
 * Prec 90.960%
best acc: 91.210000
Epoch: [172][0/391]     Time 0.225 (0.225)      Data 0.198 (0.198)      Loss
0.0114 (0.0114)    Prec 99.219% (99.219%)
Epoch: [172][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0011 (0.0040)    Prec 100.000% (99.876%)
Epoch: [172][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0013 (0.0042)    Prec 100.000% (99.876%)
Epoch: [172][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0050 (0.0042)    Prec 100.000% (99.868%)
Validation starts
Test: [0/79]    Time 0.202 (0.202)      Loss 0.1805 (0.1805)    Prec 96.094%
(96.094%)
 * Prec 91.190%
best acc: 91.210000
Epoch: [173][0/391]     Time 0.308 (0.308)      Data 0.284 (0.284)      Loss
0.0003 (0.0003)    Prec 100.000% (100.000%)
Epoch: [173][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0010 (0.0034)    Prec 100.000% (99.892%)
Epoch: [173][200/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0020 (0.0041)    Prec 100.000% (99.864%)
Epoch: [173][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0115 (0.0040)    Prec 99.219% (99.878%)
Validation starts
Test: [0/79]    Time 0.131 (0.131)      Loss 0.1775 (0.1775)    Prec 96.094%
(96.094%)
 * Prec 91.010%
best acc: 91.210000
Epoch: [174][0/391]     Time 0.314 (0.314)      Data 0.285 (0.285)      Loss
0.0117 (0.0117)    Prec 99.219% (99.219%)
Epoch: [174][100/391]   Time 0.035 (0.039)      Data 0.002 (0.004)      Loss
0.0002 (0.0046)    Prec 100.000% (99.838%)
Epoch: [174][200/391]   Time 0.035 (0.040)      Data 0.001 (0.004)      Loss
```

```
0.0033 (0.0042)    Prec 100.000% (99.860%)
Epoch: [174][300/391]   Time 0.035 (0.038)      Data 0.001 (0.003)      Loss
0.0016 (0.0044)    Prec 100.000% (99.842%)
Validation starts
Test: [0/79]    Time 0.115 (0.115)      Loss 0.1688 (0.1688)    Prec 96.875%
(96.875%)
 * Prec 91.340%
best acc: 91.340000
Epoch: [175][0/391]     Time 0.428 (0.428)      Data 0.378 (0.378)      Loss
0.0019 (0.0019)    Prec 100.000% (100.000%)
Epoch: [175][100/391]   Time 0.035 (0.039)      Data 0.002 (0.005)      Loss
0.0040 (0.0041)    Prec 100.000% (99.907%)
Epoch: [175][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0003 (0.0033)    Prec 100.000% (99.907%)
Epoch: [175][300/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0009 (0.0038)    Prec 100.000% (99.883%)
Validation starts
Test: [0/79]    Time 0.233 (0.233)      Loss 0.1510 (0.1510)    Prec 97.656%
(97.656%)
 * Prec 91.180%
best acc: 91.340000
Epoch: [176][0/391]     Time 0.333 (0.333)      Data 0.263 (0.263)      Loss
0.0005 (0.0005)    Prec 100.000% (100.000%)
Epoch: [176][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0018 (0.0049)    Prec 100.000% (99.869%)
Epoch: [176][200/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0003 (0.0044)    Prec 100.000% (99.880%)
Epoch: [176][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0004 (0.0036)    Prec 100.000% (99.907%)
Validation starts
Test: [0/79]    Time 0.136 (0.136)      Loss 0.1795 (0.1795)    Prec 96.875%
(96.875%)
 * Prec 91.340%
best acc: 91.340000
Epoch: [177][0/391]     Time 0.289 (0.289)      Data 0.266 (0.266)      Loss
0.0002 (0.0002)    Prec 100.000% (100.000%)
Epoch: [177][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0004 (0.0031)    Prec 100.000% (99.884%)
Epoch: [177][200/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0016 (0.0033)    Prec 100.000% (99.872%)
Epoch: [177][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0005 (0.0033)    Prec 100.000% (99.873%)
Validation starts
Test: [0/79]    Time 0.222 (0.222)      Loss 0.2171 (0.2171)    Prec 94.531%
(94.531%)
 * Prec 90.910%
best acc: 91.340000
Epoch: [178][0/391]     Time 0.365 (0.365)      Data 0.339 (0.339)      Loss
```

```
0.0058 (0.0058)    Prec 100.000% (100.000%)
Epoch: [178][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0032 (0.0024)    Prec 100.000% (99.923%)
Epoch: [178][200/391]   Time 0.033 (0.037)      Data 0.003 (0.003)      Loss
0.0009 (0.0031)    Prec 100.000% (99.895%)
Epoch: [178][300/391]   Time 0.036 (0.036)      Data 0.002 (0.003)      Loss
0.0009 (0.0034)    Prec 100.000% (99.888%)
Validation starts
Test: [0/79]    Time 0.142 (0.142)      Loss 0.2216 (0.2216)    Prec 95.312%
(95.312%)
 * Prec 90.980%
best acc: 91.340000
Epoch: [179][0/391]     Time 0.228 (0.228)      Data 0.203 (0.203)      Loss
0.0119 (0.0119)    Prec 99.219% (99.219%)
Epoch: [179][100/391]   Time 0.034 (0.037)      Data 0.002 (0.004)      Loss
0.0010 (0.0043)    Prec 100.000% (99.853%)
Epoch: [179][200/391]   Time 0.035 (0.036)      Data 0.002 (0.003)      Loss
0.0002 (0.0034)    Prec 100.000% (99.895%)
Epoch: [179][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0200 (0.0034)    Prec 99.219% (99.899%)
Validation starts
Test: [0/79]    Time 0.153 (0.153)      Loss 0.2253 (0.2253)    Prec 96.094%
(96.094%)
 * Prec 91.300%
best acc: 91.340000
Epoch: [180][0/391]     Time 0.222 (0.222)      Data 0.197 (0.197)      Loss
0.0016 (0.0016)    Prec 100.000% (100.000%)
Epoch: [180][100/391]   Time 0.035 (0.037)      Data 0.001 (0.004)      Loss
0.0003 (0.0038)    Prec 100.000% (99.853%)
Epoch: [180][200/391]   Time 0.035 (0.036)      Data 0.002 (0.003)      Loss
0.0031 (0.0037)    Prec 100.000% (99.876%)
Epoch: [180][300/391]   Time 0.036 (0.036)      Data 0.002 (0.002)      Loss
0.0005 (0.0035)    Prec 100.000% (99.886%)
Validation starts
Test: [0/79]    Time 0.147 (0.147)      Loss 0.2239 (0.2239)    Prec 94.531%
(94.531%)
 * Prec 91.040%
best acc: 91.340000
Epoch: [181][0/391]     Time 0.248 (0.248)      Data 0.191 (0.191)      Loss
0.0016 (0.0016)    Prec 100.000% (100.000%)
Epoch: [181][100/391]   Time 0.036 (0.038)      Data 0.001 (0.003)      Loss
0.0010 (0.0032)    Prec 100.000% (99.853%)
Epoch: [181][200/391]   Time 0.035 (0.037)      Data 0.001 (0.002)      Loss
0.0077 (0.0037)    Prec 100.000% (99.880%)
Epoch: [181][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0049 (0.0033)    Prec 100.000% (99.888%)
Validation starts
Test: [0/79]    Time 0.152 (0.152)      Loss 0.2195 (0.2195)    Prec 95.312%
```

```
(95.312%)
 * Prec 91.130%
best acc: 91.340000
Epoch: [182][0/391]     Time 0.344 (0.344)      Data 0.285 (0.285)      Loss
0.0002 (0.0002)    Prec 100.000% (100.000%)
Epoch: [182][100/391]   Time 0.035 (0.039)      Data 0.002 (0.004)      Loss
0.0004 (0.0035)    Prec 100.000% (99.861%)
Epoch: [182][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0000 (0.0032)    Prec 100.000% (99.887%)
Epoch: [182][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0043 (0.0030)    Prec 100.000% (99.901%)
Validation starts
Test: [0/79]    Time 0.145 (0.145)      Loss 0.2502 (0.2502)    Prec 96.094%
(96.094%)
 * Prec 91.040%
best acc: 91.340000
Epoch: [183][0/391]     Time 0.417 (0.417)      Data 0.383 (0.383)      Loss
0.0012 (0.0012)    Prec 100.000% (100.000%)
Epoch: [183][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0006 (0.0030)    Prec 100.000% (99.915%)
Epoch: [183][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0006 (0.0032)    Prec 100.000% (99.895%)
Epoch: [183][300/391]   Time 0.038 (0.037)      Data 0.002 (0.003)      Loss
0.0012 (0.0035)    Prec 100.000% (99.883%)
Validation starts
Test: [0/79]    Time 0.123 (0.123)      Loss 0.2114 (0.2114)    Prec 96.094%
(96.094%)
 * Prec 91.160%
best acc: 91.340000
Epoch: [184][0/391]     Time 0.321 (0.321)      Data 0.290 (0.290)      Loss
0.0005 (0.0005)    Prec 100.000% (100.000%)
Epoch: [184][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0007 (0.0041)    Prec 100.000% (99.869%)
Epoch: [184][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0002 (0.0033)    Prec 100.000% (99.895%)
Epoch: [184][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0005 (0.0031)    Prec 100.000% (99.907%)
Validation starts
Test: [0/79]    Time 0.134 (0.134)      Loss 0.2204 (0.2204)    Prec 96.094%
(96.094%)
 * Prec 91.150%
best acc: 91.340000
Epoch: [185][0/391]     Time 0.379 (0.379)      Data 0.355 (0.355)      Loss
0.0093 (0.0093)    Prec 99.219% (99.219%)
Epoch: [185][100/391]   Time 0.036 (0.039)      Data 0.001 (0.005)      Loss
0.0213 (0.0028)    Prec 99.219% (99.938%)
Epoch: [185][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0004 (0.0029)    Prec 100.000% (99.922%)
```

```
Epoch: [185][300/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0002 (0.0031)    Prec 100.000% (99.917%)
Validation starts
Test: [0/79]    Time 0.117 (0.117)      Loss 0.2687 (0.2687)    Prec 94.531%
(94.531%)
 * Prec 91.270%
best acc: 91.340000
Epoch: [186][0/391]     Time 0.340 (0.340)      Data 0.276 (0.276)      Loss
0.0307 (0.0307)    Prec 99.219% (99.219%)
Epoch: [186][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0004 (0.0030)    Prec 100.000% (99.907%)
Epoch: [186][200/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0011 (0.0034)    Prec 100.000% (99.887%)
Epoch: [186][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0012 (0.0036)    Prec 100.000% (99.883%)
Validation starts
Test: [0/79]    Time 0.153 (0.153)      Loss 0.2984 (0.2984)    Prec 94.531%
(94.531%)
 * Prec 91.160%
best acc: 91.340000
Epoch: [187][0/391]     Time 0.247 (0.247)      Data 0.216 (0.216)      Loss
0.0271 (0.0271)    Prec 98.438% (98.438%)
Epoch: [187][100/391]   Time 0.036 (0.038)      Data 0.002 (0.004)      Loss
0.0018 (0.0036)    Prec 100.000% (99.869%)
Epoch: [187][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0192 (0.0036)    Prec 98.438% (99.872%)
Epoch: [187][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0045 (0.0039)    Prec 100.000% (99.875%)
Validation starts
Test: [0/79]    Time 0.157 (0.157)      Loss 0.2565 (0.2565)    Prec 96.094%
(96.094%)
 * Prec 91.030%
best acc: 91.340000
Epoch: [188][0/391]     Time 0.290 (0.290)      Data 0.262 (0.262)      Loss
0.0004 (0.0004)    Prec 100.000% (100.000%)
Epoch: [188][100/391]   Time 0.041 (0.038)      Data 0.001 (0.004)      Loss
0.0029 (0.0035)    Prec 100.000% (99.876%)
Epoch: [188][200/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
0.0027 (0.0038)    Prec 100.000% (99.876%)
Epoch: [188][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0002 (0.0037)    Prec 100.000% (99.891%)
Validation starts
Test: [0/79]    Time 0.140 (0.140)      Loss 0.2240 (0.2240)    Prec 95.312%
(95.312%)
 * Prec 90.990%
best acc: 91.340000
Epoch: [189][0/391]     Time 0.307 (0.307)      Data 0.280 (0.280)      Loss
0.0002 (0.0002)    Prec 100.000% (100.000%)
```

```
Epoch: [189][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0016 (0.0038)    Prec 100.000% (99.861%)
Epoch: [189][200/391]   Time 0.036 (0.037)      Data 0.002 (0.003)      Loss
0.0203 (0.0034)    Prec 99.219% (99.880%)
Epoch: [189][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0119 (0.0033)    Prec 99.219% (99.896%)
Validation starts
Test: [0/79]    Time 0.226 (0.226)      Loss 0.2595 (0.2595)    Prec 95.312%
(95.312%)
 * Prec 90.830%
best acc: 91.340000
Epoch: [190][0/391]     Time 0.135 (0.135)      Data 0.111 (0.111)      Loss
0.0084 (0.0084)    Prec 100.000% (100.000%)
Epoch: [190][100/391]   Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0131 (0.0026)    Prec 99.219% (99.954%)
Epoch: [190][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0001 (0.0028)    Prec 100.000% (99.946%)
Epoch: [190][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0004 (0.0027)    Prec 100.000% (99.943%)
Validation starts
Test: [0/79]    Time 0.284 (0.284)      Loss 0.2332 (0.2332)    Prec 96.094%
(96.094%)
 * Prec 91.080%
best acc: 91.340000
Epoch: [191][0/391]     Time 0.298 (0.298)      Data 0.273 (0.273)      Loss
0.0002 (0.0002)    Prec 100.000% (100.000%)
Epoch: [191][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0002 (0.0032)    Prec 100.000% (99.892%)
Epoch: [191][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0004 (0.0032)    Prec 100.000% (99.903%)
Epoch: [191][300/391]   Time 0.034 (0.040)      Data 0.001 (0.003)      Loss
0.0004 (0.0032)    Prec 100.000% (99.896%)
Validation starts
Test: [0/79]    Time 0.217 (0.217)      Loss 0.2390 (0.2390)    Prec 95.312%
(95.312%)
 * Prec 91.110%
best acc: 91.340000
Epoch: [192][0/391]     Time 0.294 (0.294)      Data 0.271 (0.271)      Loss
0.0024 (0.0024)    Prec 100.000% (100.000%)
Epoch: [192][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0406 (0.0032)    Prec 99.219% (99.907%)
Epoch: [192][200/391]   Time 0.035 (0.036)      Data 0.001 (0.003)      Loss
0.0013 (0.0030)    Prec 100.000% (99.918%)
Epoch: [192][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0009 (0.0031)    Prec 100.000% (99.914%)
Validation starts
Test: [0/79]    Time 0.234 (0.234)      Loss 0.2211 (0.2211)    Prec 95.312%
(95.312%)
```

```
 * Prec 91.190%
best acc: 91.340000
Epoch: [193][0/391]     Time 0.304 (0.304)      Data 0.274 (0.274)      Loss
0.0050 (0.0050)    Prec 100.000% (100.000%)
Epoch: [193][100/391]   Time 0.035 (0.038)      Data 0.002 (0.004)      Loss
0.0024 (0.0044)    Prec 100.000% (99.853%)
Epoch: [193][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0001 (0.0039)    Prec 100.000% (99.880%)
Epoch: [193][300/391]   Time 0.035 (0.036)      Data 0.002 (0.002)      Loss
0.0020 (0.0039)    Prec 100.000% (99.883%)
Validation starts
Test: [0/79]    Time 0.132 (0.132)      Loss 0.2487 (0.2487)    Prec 95.312%
(95.312%)
 * Prec 91.040%
best acc: 91.340000
Epoch: [194][0/391]     Time 0.309 (0.309)      Data 0.275 (0.275)      Loss
0.0010 (0.0010)    Prec 100.000% (100.000%)
Epoch: [194][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0004 (0.0030)    Prec 100.000% (99.892%)
Epoch: [194][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0006 (0.0032)    Prec 100.000% (99.876%)
Epoch: [194][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0192 (0.0034)    Prec 99.219% (99.878%)
Validation starts
Test: [0/79]    Time 0.138 (0.138)      Loss 0.1946 (0.1946)    Prec 96.875%
(96.875%)
 * Prec 91.160%
best acc: 91.340000
Epoch: [195][0/391]     Time 0.242 (0.242)      Data 0.218 (0.218)      Loss
0.0016 (0.0016)    Prec 100.000% (100.000%)
Epoch: [195][100/391]   Time 0.035 (0.038)      Data 0.001 (0.004)      Loss
0.0002 (0.0051)    Prec 100.000% (99.861%)
Epoch: [195][200/391]   Time 0.037 (0.037)      Data 0.001 (0.003)      Loss
0.0118 (0.0041)    Prec 99.219% (99.880%)
Epoch: [195][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0002 (0.0038)    Prec 100.000% (99.896%)
Validation starts
Test: [0/79]    Time 0.131 (0.131)      Loss 0.2599 (0.2599)    Prec 93.750%
(93.750%)
 * Prec 91.050%
best acc: 91.340000
Epoch: [196][0/391]     Time 0.377 (0.377)      Data 0.349 (0.349)      Loss
0.0004 (0.0004)    Prec 100.000% (100.000%)
Epoch: [196][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0017 (0.0041)    Prec 100.000% (99.876%)
Epoch: [196][200/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0019 (0.0044)    Prec 100.000% (99.880%)
Epoch: [196][300/391]   Time 0.036 (0.037)      Data 0.001 (0.003)      Loss
```

```
0.0097 (0.0042)    Prec 99.219% (99.883%)
Validation starts
Test: [0/79]    Time 0.223 (0.223)      Loss 0.2544 (0.2544)    Prec 96.094%
(96.094%)
 * Prec 91.130%
best acc: 91.340000
Epoch: [197][0/391]     Time 0.223 (0.223)      Data 0.198 (0.198)      Loss
0.0001 (0.0001)    Prec 100.000% (100.000%)
Epoch: [197][100/391]   Time 0.035 (0.037)      Data 0.001 (0.003)      Loss
0.0001 (0.0035)    Prec 100.000% (99.861%)
Epoch: [197][200/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0021 (0.0034)    Prec 100.000% (99.880%)
Epoch: [197][300/391]   Time 0.036 (0.036)      Data 0.001 (0.002)      Loss
0.0009 (0.0033)    Prec 100.000% (99.888%)
Validation starts
Test: [0/79]    Time 0.122 (0.122)      Loss 0.1731 (0.1731)    Prec 96.094%
(96.094%)
 * Prec 91.020%
best acc: 91.340000
Epoch: [198][0/391]     Time 0.311 (0.311)      Data 0.279 (0.279)      Loss
0.0002 (0.0002)    Prec 100.000% (100.000%)
Epoch: [198][100/391]   Time 0.036 (0.038)      Data 0.001 (0.004)      Loss
0.0005 (0.0043)    Prec 100.000% (99.830%)
Epoch: [198][200/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0003 (0.0036)    Prec 100.000% (99.856%)
Epoch: [198][300/391]   Time 0.035 (0.036)      Data 0.001 (0.002)      Loss
0.0279 (0.0037)    Prec 98.438% (99.855%)
Validation starts
Test: [0/79]    Time 0.232 (0.232)      Loss 0.2330 (0.2330)    Prec 95.312%
(95.312%)
 * Prec 91.010%
best acc: 91.340000
Epoch: [199][0/391]     Time 0.440 (0.440)      Data 0.383 (0.383)      Loss
0.0028 (0.0028)    Prec 100.000% (100.000%)
Epoch: [199][100/391]   Time 0.035 (0.039)      Data 0.001 (0.005)      Loss
0.0003 (0.0028)    Prec 100.000% (99.923%)
Epoch: [199][200/391]   Time 0.041 (0.037)      Data 0.001 (0.003)      Loss
0.0012 (0.0028)    Prec 100.000% (99.914%)
Epoch: [199][300/391]   Time 0.035 (0.037)      Data 0.002 (0.003)      Loss
0.0049 (0.0035)    Prec 100.000% (99.896%)
Validation starts
Test: [0/79]    Time 0.124 (0.124)      Loss 0.1936 (0.1936)    Prec 96.875%
(96.875%)
 * Prec 91.150%
best acc: 91.340000
```

[4]:
```
# Load trained model
# PATH = "result/VGG16_quant/model_best.pth.tar"
# checkpoint = torch.load(PATH)
# model.load_state_dict(checkpoint['state_dict'])
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")

model.cuda()
model.eval()

# Evaluate model accuracy
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()

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

```
Test set: Accuracy: 9115/10000 (91%)
```

[5]:
```
# Capture layer inputs using forward pre-hooks
class SaveOutput:
    def __init__(self):
        self.outputs = []
    def __call__(self, module, module_in):
        self.outputs.append(module_in[0])
    def clear(self):
        self.outputs = []

# Register hooks for both layers we need to capture
save_output = SaveOutput()
save_output_next = SaveOutput()
target_layer = model.features[27]  # Current conv layer
target_layer_next = model.features[29]  # Next conv layer (after ReLU)
target_layer.register_forward_pre_hook(save_output)
target_layer_next.register_forward_pre_hook(save_output_next)

# Get a sample batch
dataiter = iter(testloader)
images, labels = next(dataiter)
images = images.cuda()
```

```
# Single forward pass to capture both inputs
model.eval()
with torch.no_grad():
    _ = model(images)

x = save_output.outputs[0]  # Input to features[27]
x_next_ref = save_output_next.outputs[0]  # Input to features[29] (after ReLU)
print(f"Captured input shape (features[27]): {x.shape}")
print(f"Captured next layer input shape (features[29]): {x_next_ref.shape}")
```

```
Captured input shape (features[27]): torch.Size([128, 8, 4, 4])
Captured next layer input shape (features[29]): torch.Size([128, 8, 4, 4])
```

[6]:
```
# Convert quantized weights to integers
w_bit = 4
target_conv = model.features[27]
weight_q = target_conv.weight_q.data
w_alpha = target_conv.weight_quant.wgt_alpha.data.item()
w_delta = w_alpha / (2 ** (w_bit-1) - 1)
# divide by alpha first to get normalized values and multiply by (2^(w_bit - 1)
 ↪- 1) to get integers
weight_int = torch.round(weight_q / w_delta).int()
print(f"Weight alpha: {w_alpha:.4f}, Weight delta: {w_delta:.4f}")
print(f"Weight int shape: {weight_int.shape}")
print(f"Weight int sample (first few values): {weight_int.flatten()[:10]}")
```

```
Weight alpha: 2.1144, Weight delta: 0.3021
Weight int shape: torch.Size([8, 8, 3, 3])
Weight int sample (first few values): tensor([-4,  5,  4, -2, -2, -3, -4, -2,
-5, -3], device='cuda:0',
       dtype=torch.int32)
```

[7]:
```
# Convert quantized activations to integers
x_bit = 4
target_conv = model.features[27]
x_alpha = target_conv.act_alpha.data.item()
x_delta = x_alpha / (2 ** x_bit - 1)

act_quant_fn = act_quantization(x_bit)
x_q = act_quant_fn(x, x_alpha)
x_int = torch.round(x_q / x_delta).int()
print(f"Activation alpha: {x_alpha:.4f}, Activation delta: {x_delta:.4f}")
print(f"x_int shape: {x_int.shape}")
print(f"x_int sample (first few values): {x_int.flatten()[:10]}")
```

```
Activation alpha: 5.4116, Activation delta: 0.3608
x_int shape: torch.Size([128, 8, 4, 4])
x_int sample (first few values): tensor([0, 0, 0, 0, 0, 0, 0, 0, 0, 0],
device='cuda:0', dtype=torch.int32)
```

[8]:
```
# Perform integer convolution and recover floating-point output
conv_int = torch.nn.Conv2d(
    in_channels=8,
    out_channels=8,
    kernel_size=3,
    padding=1,
    bias=False
)
conv_int.weight = torch.nn.parameter.Parameter(weight_int.float())

output_int = conv_int(x_int.float())
output_recovered = output_int * x_delta * w_delta
print(f"Output recovered shape: {output_recovered.shape}")
print(f"Output recovered sample: {output_recovered[0, 0, :5, :5]}")
```

```
Output recovered shape: torch.Size([128, 8, 4, 4])
Output recovered sample: tensor([[ 0.8718,  0.3269,  0.9808,  0.3269],
        [ 0.2180,  0.6539, -1.3077, -0.2180],
        [ 1.5257, -3.0513, -5.0129, -0.2180],
        [ 0.7628, -5.0129, -3.4872, -0.4359]], device='cuda:0',
       grad_fn=<SliceBackward0>)
```

[10]:
```
# Apply ReLU to the recovered output and compare with reference
output_recovered_relu = F.relu(output_recovered)
difference = abs(x_next_ref - output_recovered_relu)
print(f"Difference mean: {difference.mean():.10f}")
print(f"Difference max: {difference.max():.10f}")
print(f"Difference should be < 10^-3: {difference.mean() < 1e-3}")
```

```
Difference mean: 0.0000005847
Difference max: 0.0000095367
Difference should be < 10^-3: True
```

[ ]:
```
# Prepare data for systolic array computation
# Extract single sample from batch
a_int = x_int[0,:,:,:]  # Shape: [16, 4, 4] - input channels, height, width

# Reshape weights: [out_ch, in_ch, ki, kj] -> [out_ch, in_ch, kij]
w_int = torch.reshape(weight_int, (weight_int.size(0), weight_int.size(1), -1))␣
 ↪ # [16, 16, 9]

padding = 1
stride = 1
array_size = 8  # 8x8 systolic array

# Define ranges
nig = range(a_int.size(1))  # ni (height)
njg = range(a_int.size(2))   # nj (width)
```

```
icg = range(int(w_int.size(1)))  # input channels
ocg = range(int(w_int.size(0)))  # output channels
ic_tileg = range(int(len(icg)/array_size))  # input channel tiles
oc_tileg = range(int(len(ocg)/array_size))  # output channel tiles
kijg = range(w_int.size(2))  # kernel elements (3x3 = 9)
ki_dim = int(math.sqrt(w_int.size(2)))  # kernel dimension (3)

# Pad activations
a_pad = torch.zeros(len(icg), len(nig)+padding*2, len(njg)+padding*2).cuda()
a_pad[:, padding:padding+len(nig), padding:padding+len(njg)] = a_int.cuda()
a_pad = torch.reshape(a_pad, (a_pad.size(0), -1))  # [16, 36] - flattened␣
 ↪spatial dimensions

# Create tiles for systolic array
a_tile = torch.zeros(len(ic_tileg), array_size, a_pad.size(1)).cuda()
w_tile = torch.zeros(len(oc_tileg)*len(ic_tileg), array_size, array_size,␣
 ↪len(kijg)).cuda()

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

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

# Compute partial sums using matrix multiplication (simulating systolic array)
p_nijg = range(a_pad.size(1))
psum = torch.zeros(len(ic_tileg), len(oc_tileg), array_size, len(p_nijg),␣
 ↪len(kijg)).cuda()

for kij in kijg:
    for ic_tile in ic_tileg:
        for oc_tile in oc_tileg:
            for nij in p_nijg:
                m = nn.Linear(array_size, array_size, bias=False)
                m.weight = torch.nn.
 ↪Parameter(w_tile[len(ic_tileg)*oc_tile+ic_tile,:,:,kij])
                psum[ic_tile, oc_tile, :, nij, kij] = m(a_tile[ic_tile,:,nij]).
 ↪cuda()
```

[ ]:
```
# Accumulate partial sums across tiles and kernel elements
a_pad_ni_dim = int(math.sqrt(a_pad.size(1)))  # 6 (4+2*padding)
o_ni_dim = int((a_pad_ni_dim - (ki_dim- 1) - 1)/stride + 1)  # 4
o_nijg = range(o_ni_dim**2)  # 16 output spatial positions
```

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

# Accumulate across output tiles and input tiles
for oc_tile in oc_tileg:  # Loop over output tiles (0, 1 for 16 output channels
 ↪/ 8 array_size)
    partial_out = torch.zeros(array_size, len(o_nijg)).cuda()

    # Accumulate across input tiles
    for ic_tile in ic_tileg:  # 0 = IC 0-7, 1 = IC 8-15
        for kij in kijg:
            ky = kij // ki_dim     # kernel row
            kx = kij %  ki_dim     # kernel col

            for o_nij in o_nijg:
                oy = o_nij // o_ni_dim
                ox = o_nij %  o_ni_dim
                nij_src = (oy + ky) * a_pad_ni_dim + (ox + kx)

                partial_out[:, o_nij] += psum[
                    ic_tile,    # input tile index
                    oc_tile,    # output tile index
                    :,          # 8 systolic PE rows
                    nij_src,    # spatial index
                    kij         # kernel element
                ]

    # Store this 8-channel block
    out[oc_tile * array_size:(oc_tile + 1) * array_size, :] = partial_out

# Reshape output to match original shape [16, 4, 4]
out_reshaped = out.view(len(ocg), o_ni_dim, o_ni_dim)

# Recover the output and apply ReLU
out_recovered = out_reshaped * x_delta * w_delta
out_recovered_relu = F.relu(out_recovered)

print(f"Output recovered shape: {out_recovered_relu.shape}")
print(f"Output recovered sample: {out_recovered_relu[0, :5, :5]}")

# Compare with reference
difference_tiled = abs(x_next_ref[0] - out_recovered_relu)
print(f"\nTiled computation difference mean: {difference_tiled.mean():.6f}")
print(f"Tiled computation difference max: {difference_tiled.max():.6f}")
print(f"Tiled computation difference should be < 10^-3: {difference_tiled.
 ↪mean() < 1e-3}")
```

[ ]:
```
# Generate activation data file for hardware testing
ic_tile_id = 0
nij = 0
nij_end = min(nij + 64, a_tile.size(1))
X_4b = a_tile[ic_tile_id, :, nij:nij_end]

bit_precision = 4
file = open('activation_tile0.txt', 'w')
file.write('#time0row7[msb-lsb],time0row6[msb-lst],....,time0row0[msb-lst]#\n')
file.write('#time1row7[msb-lsb],time1row6[msb-lst],....,time1row0[msb-lst]#\n')
file.write('#................#\n')

for i in range(X_4b.size(1)):        # time
    for j in range(X_4b.size(0)):    # row
        val = int(round(X_4b[7-j, i].item()))
        if val < 0:
            val = val + (2**bit_precision)  # Convert negative to unsigned
        X_bin = '{0:04b}'.format(val)
        file.write(X_bin)
    file.write('\n')
file.close()
print(f"Written activation data: shape {X_4b.shape}, nij range [{nij}:
 ↪{nij_end}]")
```

[ ]:
```
# Generate weight data files for hardware testing
# Generate files for all kij values (0-8) as expected by testbench
ic_tile_id = 0
oc_tile_id = 0
tile_id = oc_tile_id * len(ic_tileg) + ic_tile_id

bit_precision = 4
len_kij = 9  # Number of kernel iterations (0-8)

for kij in range(len_kij):
    W = w_tile[tile_id,:,:,kij]

    # Generate filename matching testbench expectation:␣
 ↪weight_itile{ic}_otile{oc}_kij{kij}.txt
    filename = f'weight_itile{ic_tile_id}_otile{oc_tile_id}_kij{kij}.txt'
    file = open(filename, 'w')

    # Write 3 comment lines (testbench skips first 3 lines)
    file.write('#col0row7[msb-lsb],col0row6[msb-lst],....,col0row0[msb-lst]#\n')
    file.write('#col1row7[msb-lsb],col1row6[msb-lst],....,col1row0[msb-lst]#\n')
    file.write('#................#\n')

    # Write 8 lines (one per column), each with 32 bits (8 rows × 4 bits)
```

```
    for i in range(W.size(0)):  # columns (8 columns)
        for j in range(W.size(1)):  # rows (8 rows)
            weight_val = round(W[7-j,i].item())
            if weight_val < 0:
                weight_val = weight_val + (2**bit_precision)  # Convert to␣
 ↪unsigned
            W_bin = '{0:04b}'.format(weight_val)
            for k in range(bit_precision):
                file.write(W_bin[k])
        file.write('\n')
    file.close()
    print(f"Written weight data: shape {W.shape}, tile_id={tile_id}, kij={kij},␣
 ↪file={filename}")

print(f"\nGenerated {len_kij} weight files for hardware testbench")
```

[ ]:
```
# Generate partial sum data file for hardware testing
ic_tile_id = 0
oc_tile_id = 0
kij = 0
nij = 0
nij_end = min(nij + 64, psum.size(3))
psum_tile = psum[ic_tile_id, oc_tile_id, :, nij:nij_end, kij]

bit_precision = 16
file = open('psum.txt', 'w')
file.write('#time0col7[msb-lsb],time0col6[msb-lst],....,time0col0[msb-lst]#\n')
file.write('#time1col7[msb-lsb],time1col6[msb-lst],....,time1col0[msb-lst]#\n')
file.write('#................#\n')

for i in range(psum_tile.size(1)):  # time
    for j in range(psum_tile.size(0)):  # column (PE)
        psum_val = round(psum_tile[7-j, i].item())
        if psum_val < 0:
            psum_val = psum_val + (2**bit_precision)  # Convert to unsigned
        psum_bin = '{0:16b}'.format(psum_val)
        for k in range(bit_precision):
            file.write(psum_bin[k])
    file.write('\n')
file.close()
print(f"Written psum data: shape {psum_tile.shape}, nij range [{nij}:
 ↪{nij_end}], kij={kij}")
```

[ ]:
```
# Generate expected output file for hardware verification (out.txt)
# This file contains the final output after accumulation and ReLU
# Uses out_recovered_relu from Cell 11 computation
```

```
ic_tile_id = 0
oc_tile_id = 0

# Extract output for first output tile (channels 0-7, i.e., first 8 channels)
# out_recovered_relu shape from Cell 11: [16, 4, 4] (16 channels, 4x4 spatial)
# For oc_tile_id=0, we need channels 0-7
out_final = out_recovered_relu[oc_tile_id * 8:(oc_tile_id + 1) * 8, :, :].cpu()
 ↪ # Shape: [8, 4, 4]

# Reshape to [8 channels, 16 spatial positions]
o_ni_dim = out_final.size(1)  # 4
out_final_flat = out_final.view(8, -1)  # [8, 16]

bit_precision = 16
file = open('out.txt', 'w')
file.write('#out0col7[msb-lsb],out0col6[msb-lst],....,out0col0[msb-lst]#\n')
file.write('#out1col7[msb-lsb],out1col6[msb-lst],....,out1col0[msb-lst]#\n')
file.write('#................#\n')

# Write 16 lines (one per output spatial position)
# Each line: 128 bits = 8 columns × 16 bits
for i in range(out_final_flat.size(1)):  # 16 output positions
    for j in range(out_final_flat.size(0)):  # 8 columns (output channels)
        out_val = round(out_final_flat[7-j, i].item())
        if out_val < 0:
            out_val = 0  # ReLU already applied, but ensure no negatives
        out_bin = '{0:016b}'.format(out_val & ((2**bit_precision)-1))  # Ensure
 ↪16 bits
        file.write(out_bin)
    file.write('\n')
file.close()
print(f"Written output data: shape {out_final_flat.shape}, file=out.txt")
```

[ ]:
```
# Generate accumulation address file (acc.txt)
# This file contains memory addresses where partial sums are stored
# Format: 11-bit binary addresses, total of len_onij × len_kij = 16 × 9 = 144
 ↪addresses

len_onij = 16  # Number of output spatial positions
len_kij = 9    # Number of kernel iterations

file = open('acc.txt', 'w')

# Generate addresses for each output position and each kij iteration
# Address calculation: kij * len_onij + output_position
for output_pos in range(len_onij):
    for kij in range(len_kij):
```

```
        address = kij * len_onij + output_pos
        # Convert to 11-bit binary
        addr_bin = '{0:011b}'.format(address & ((2**11)-1))  # Ensure 11 bits
        file.write(addr_bin + '\n')

file.close()
print(f"Written accumulation address file: {len_onij * len_kij} addresses,␣
 ↪file=acc.txt")
```

[ ]: