In [1]:

```
import os
import random
import shutil
import time
import warnings

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

import torch.utils.data
import torchvision
import torchvision.transforms as transforms
import torchvision.datasets as datasets
import torchvision.models as models
```

In [2]:

```
from torch.cuda.amp import GradScaler
from torch.cuda.amp import autocast
```

In [3]:

```
from torch.utils.tensorboard import SummaryWriter
writer = SummaryWriter()
```

In [4]:

```
!pip install wandb
```

```
Collecting wandb
  Downloading wandb-0.12.6-py2.py3-none-any.whl (1.7 MB)
     |████████████████████████████████| 1.7 MB 6.6 MB/s eta 0:00:01
Requirement already satisfied: Click!=8.0.0,>=7.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (7.1.2)
Collecting shortuuid>=0.5.0
  Downloading shortuuid-1.0.1-py3-none-any.whl (7.5 kB)
Collecting yaspin>=1.0.0
  Downloading yaspin-2.1.0-py3-none-any.whl (18 kB)
Collecting sentry-sdk>=1.0.0
  Downloading sentry_sdk-1.4.3-py2.py3-none-any.whl (139 kB)
     |████████████████████████████████| 139 kB 108.3 MB/s eta 0:00:01
Requirement already satisfied: protobuf>=3.12.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (3.17.3)
Requirement already satisfied: PyYAML in /opt/conda/lib/python3.8/site-packages (from wandb) (5.4.1)
Collecting pathtools
  Downloading pathtools-0.1.2.tar.gz (11 kB)
Collecting GitPython>=1.0.0
  Downloading GitPython-3.1.24-py3-none-any.whl (180 kB)
     |████████████████████████████████| 180 kB 93.0 MB/s eta 0:00:01
Requirement already satisfied: python-dateutil>=2.6.1 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.8.2)
```

```
Requirement already satisfied: psutil>=5.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (5.8.0)
Requirement already satisfied: six>=1.13.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (1.16.0)
Collecting promise<3,>=2.0
  Downloading promise-2.3.tar.gz (19 kB)
Collecting docker-pycreds>=0.4.0
  Downloading docker_pycreds-0.4.0-py2.py3-none-any.whl (9.0 kB)
Collecting configparser>=3.8.1
  Downloading configparser-5.0.2-py3-none-any.whl (19 kB)
Requirement already satisfied: requests<3,>=2.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.26.0)
Collecting subprocess32>=3.5.3
  Downloading subprocess32-3.5.4.tar.gz (97 kB)
     |████████████████████████████████| 97 kB 13.8 MB/s eta 0:00:01
Collecting gitdb<5,>=4.0.1
  Downloading gitdb-4.0.9-py3-none-any.whl (63 kB)
     |████████████████████████████████| 63 kB 3.6 MB/s  eta 0:00:01
Requirement already satisfied: typing-extensions>=3.7.4.3 in /opt/conda/lib/python3.8/site-packages (from GitPython>=1.0.
0->wandb) (3.10.0.0)
Collecting smmap<6,>=3.0.1
  Downloading smmap-5.0.0-py3-none-any.whl (24 kB)
Requirement already satisfied: urllib3<1.27,>=1.21.1 in /opt/conda/lib/python3.8/site-packages (from requests<3,>=2.0.0->
wandb) (1.26.6)
Requirement already satisfied: charset-normalizer~=2.0.0 in /opt/conda/lib/python3.8/site-packages (from requests<3,>=2.
0.0->wandb) (2.0.0)
Requirement already satisfied: certifi>=2017.4.17 in /opt/conda/lib/python3.8/site-packages (from requests<3,>=2.0.0->wan
db) (2021.5.30)
Requirement already satisfied: idna<4,>=2.5 in /opt/conda/lib/python3.8/site-packages (from requests<3,>=2.0.0->wandb)
(3.1)
Collecting termcolor<2.0.0,>=1.1.0
  Downloading termcolor-1.1.0.tar.gz (3.9 kB)
Building wheels for collected packages: promise, subprocess32, termcolor, pathtools
  Building wheel for promise (setup.py) ... done
  Created wheel for promise: filename=promise-2.3-py3-none-any.whl size=21502 sha256=d1e96ee754b60142006ae6154cbb03c44a6c
9cadbf09017204c9db40d9f9013b
  Stored in directory: /root/.cache/pip/wheels/54/aa/01/724885182f93150035a2a91bce34a12877e8067a97baaf5dc8
  Building wheel for subprocess32 (setup.py) ... done
  Created wheel for subprocess32: filename=subprocess32-3.5.4-py3-none-any.whl size=6502 sha256=fb42cb5354f07eef2b8aa31ce
727cf7a2d8013842b8bb92b6efa16bc5de809f5
  Stored in directory: /root/.cache/pip/wheels/9f/69/d1/50b39b308a87998eaf5c1d9095e5a5bd2ad98501e2b7936d36
  Building wheel for termcolor (setup.py) ... done
  Created wheel for termcolor: filename=termcolor-1.1.0-py3-none-any.whl size=4847 sha256=bfc0d7aa5b74c50e177c51113c43646
030d2c1f18012183222795f686fb11f87
  Stored in directory: /root/.cache/pip/wheels/a0/16/9c/5473df82468f958445479c59e784896fa24f4a5fc024b0f501
  Building wheel for pathtools (setup.py) ... done
  Created wheel for pathtools: filename=pathtools-0.1.2-py3-none-any.whl size=8807 sha256=94b1cb52dd55d0928a9b7577304cf0a
fa3eb8642e6308e382f3f3f5efab6e669
  Stored in directory: /root/.cache/pip/wheels/4c/8e/7e/72fbc243e1aeecae64a96875432e70d4e92f3d2d18123be004
Successfully built promise subprocess32 termcolor pathtools
Installing collected packages: smmap, termcolor, gitdb, yaspin, subprocess32, shortuuid, sentry-sdk, promise, pathtools,
GitPython, docker-pycreds, configparser, wandb
Successfully installed GitPython-3.1.24 configparser-5.0.2 docker-pycreds-0.4.0 gitdb-4.0.9 pathtools-0.1.2 promise-2.3 s
entry-sdk-1.4.3 shortuuid-1.0.1 smmap-5.0.0 subprocess32-3.5.4 termcolor-1.1.0 wandb-0.12.6 yaspin-2.1.0
```

```
WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system packag
e manager. It is recommended to use a virtual environment instead: https://pip.pypa.io/warnings/venv
```

In [6]:
```
import wandb
wandb.login()
```

```
wandb: You can find your API key in your browser here: https://wandb.ai/authorize
wandb: Paste an API key from your profile and hit enter: ••••••••
wandb: Appending key for api.wandb.ai to your netrc file: /root/.netrc
```

Out[6]:
```
True
```

In [7]:
```
wandb.init(project="w251-hw9", entity="rochelleli")
```

```
wandb: Currently logged in as: rochelleli (use `wandb login --relogin` to force relogin)
```

Syncing run **costumed-specter-1** to Weights & Biases (docs).

Out[7]: Display W&B run

In [8]:
```
SEED=1

random.seed(SEED)
torch.manual_seed(SEED)
cudnn.deterministic = True
```

In [9]:
```
torch.cuda.device_count()
```

Out[9]:
```
1
```

In [10]:
```
START_EPOCH = 0
```

In [14]:
```
ARCH = 'resnet18'
EPOCHS = 2
LR = 0.1
MOMENTUM = 0.9
WEIGHT_DECAY = 1e-4
PRINT_FREQ = 10
TRAIN_BATCH=500
VAL_BATCH=500
WORKERS=2
TRAINDIR="/home/ubuntu/data/train"
VALDIR="/home/ubuntu/data/val"
```

In [15]:
```
if not torch.cuda.is_available():
    print('GPU not detected.. did you pass through your GPU?')
```

In [16]:
```
wandb.init(config={"epochs": EPOCHS, "batch_size": TRAIN_BATCH, "momentum": MOMENTUM, "WEIGHT_DECAY": WEIGHT_DECAY, "arc
```

Finishing last run (ID:1pgsa3v1) before initializing another...

Waiting for W&B process to finish, PID 449... **(success).**

Synced 5 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)
Synced **costumed-specter-1**: https://wandb.ai/rochelleli/w251-hw9/runs/1pgsa3v1
Find logs at: `./wandb/run-20211031_065745-1pgsa3v1/logs`
Successfully finished last run (ID:1pgsa3v1). Initializing new run:
Syncing run **eerie-poltergeist-4** to Weights & Biases (docs).

Out[16]: Display W&B run

In [24]:
```
GPU = 0
```

In [25]:
```
torch.cuda.set_device(GPU)
```

In [26]:
```
global_step = 1
```

In [17]:
```
def train(train_loader, model, criterion, optimizer, epoch):
    global global_step
    batch_time = AverageMeter('Time', ':6.3f')
    data_time = AverageMeter('Data', ':6.3f')
    losses = AverageMeter('Loss', ':.4e')
    top1 = AverageMeter('Acc@1', ':6.2f')
    top5 = AverageMeter('Acc@5', ':6.2f')
    progress = ProgressMeter(
        len(train_loader),
        [batch_time, data_time, losses, top1, top5],
        prefix="Epoch: [{}]".format(epoch))

    # Grad Scaler
    scaler = GradScaler()
    # switch to train mode
    model.train()

    end = time.time()
```

```
    for i, (images, target) in enumerate(train_loader):
        # measure data loading time
        data_time.update(time.time() - end)
        optimizer.zero_grad()

        if GPU is not None:
            images = images.cuda(GPU, non_blocking=True)
        if torch.cuda.is_available():
            target = target.cuda(GPU, non_blocking=True)

        # compute output
        with autocast():
          output = model(images)
          loss = criterion(output, target)

        # measure accuracy and record loss
        acc1, acc5 = accuracy(output, target, topk=(1, 5))
        losses.update(loss.item(), images.size(0))
        top1.update(acc1[0], images.size(0))
        top5.update(acc5[0], images.size(0))

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

        # use the scaler
        scaler.scale(loss).backward()
        scaler.step(optimizer)
        scaler.update()

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

        writer.add_scalar("Loss/train", loss, global_step = global_step)
        writer.add_scalar("acc1/train", top1.avg, global_step = global_step)
        writer.add_scalar("acc5/train", top5.avg, global_step = global_step)

        wandb.log({"Loss/train": loss, 'acc1/train': top1.avg, 'acc5/train': top5.avg})

        global_step = global_step + 1

        if i % PRINT_FREQ == 0:
            progress.display(i)
```

In [18]:

```
def validate(val_loader, model, criterion):
    global global_step
    batch_time = AverageMeter('Time', ':6.3f')
```

```
    losses = AverageMeter('Loss', ':.4e')
    top1 = AverageMeter('Acc@1', ':6.2f')
    top5 = AverageMeter('Acc@5', ':6.2f')
    progress = ProgressMeter(
        len(val_loader),
        [batch_time, losses, top1, top5],
        prefix='Test: ')

    # switch to evaluate mode
    model.eval()

    with torch.no_grad():
        end = time.time()
        for i, (images, target) in enumerate(val_loader):
            if GPU is not None:
                images = images.cuda(GPU, non_blocking=True)
            if torch.cuda.is_available():
                target = target.cuda(GPU, non_blocking=True)

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

            # measure accuracy and record loss
            acc1, acc5 = accuracy(output, target, topk=(1, 5))
            losses.update(loss.item(), images.size(0))
            top1.update(acc1[0], images.size(0))
            top5.update(acc5[0], images.size(0))

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

            if i % PRINT_FREQ == 0:
                progress.display(i)

        # TODO: this should also be done with the ProgressMeter
        print(' * Acc@1 {top1.avg:.3f} Acc@5 {top5.avg:.3f}'
              .format(top1=top1, top5=top5))
    writer.add_scalar("Loss/val", losses.avg, global_step = global_step)
    writer.add_scalar("acc1/val", top1.avg, global_step = global_step)
    writer.add_scalar("acc5/val", top5.avg, global_step = global_step)

    wandb.log({"Loss/val": losses.avg, 'acc1/val': top1.avg, 'acc5/val': top5.avg})

    global_step = global_step + 1

    return top1.avg
```

In [19]:

```
def save_checkpoint(state, is_best, filename='checkpoint.pth.tar'):
    torch.save(state, filename)
    if is_best:
        shutil.copyfile(filename, 'model_best.pth.tar')
```

In [20]:

```
class AverageMeter(object):
    """Computes and stores the average and current value"""
    def __init__(self, name, fmt=':f'):
        self.name = name
        self.fmt = fmt
        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 __str__(self):
        fmtstr = '{name} {val' + self.fmt + '} ({avg' + self.fmt + '})'
        return fmtstr.format(**self.__dict__)
```

In [21]:

```
class ProgressMeter(object):
    def __init__(self, num_batches, meters, prefix=""):
        self.batch_fmtstr = self._get_batch_fmtstr(num_batches)
        self.meters = meters
        self.prefix = prefix

    def display(self, batch):
        entries = [self.prefix + self.batch_fmtstr.format(batch)]
        entries += [str(meter) for meter in self.meters]
        print('\t'.join(entries))

    def _get_batch_fmtstr(self, num_batches):
        num_digits = len(str(num_batches // 1))
        fmt = '{:' + str(num_digits) + 'd}'
        return '[' + fmt + '/' + fmt.format(num_batches) + ']'
```

In [22]:

```
def adjust_learning_rate(optimizer, epoch):
    """Sets the learning rate to the initial LR decayed by 10 every 30 epochs"""
```

```
    lr = LR * (0.1 ** (epoch // 30))
    for param_group in optimizer.param_groups:
        param_group['lr'] = lr
```

In [23]:
```
def accuracy(output, target, topk=(1,)):
    """Computes the accuracy over the k top predictions for the specified values of k"""
    with torch.no_grad():
        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].reshape(-1).float().sum(0, keepdim=True)
            res.append(correct_k.mul_(100.0 / batch_size))
        return res
```

In [27]:
```
cudnn.benchmark = True
```

In [28]:
```
imagenet_mean_RGB = [0.47889522, 0.47227842, 0.43047404]
imagenet_std_RGB = [0.229, 0.224, 0.225]
```

In [29]:
```
normalize = transforms.Normalize(mean=imagenet_mean_RGB, std=imagenet_std_RGB)
```

In [30]:
```
IMG_SIZE = 224
NUM_CLASSES = 1000
```

In [31]:
```
model = models.__dict__[ARCH]()
```

In [32]:
```
inf = model.fc.in_features
```

In [33]:
```
model.fc = nn.Linear(inf, NUM_CLASSES)
```

In [34]:
```
model.cuda(GPU)
```

Out[34]:
```
ResNet(
  (conv1): Conv2d(3, 64, kernel_size=(7, 7), stride=(2, 2), padding=(3, 3), bias=False)
  (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
  (relu): ReLU(inplace=True)
  (maxpool): MaxPool2d(kernel_size=3, stride=2, padding=1, dilation=1, ceil_mode=False)
  (layer1): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
    (1): BasicBlock(
      (conv1): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(64, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(64, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (layer2): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(64, 128, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (downsample): Sequential(
        (0): Conv2d(64, 128, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (1): BasicBlock(
      (conv1): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(128, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(128, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (layer3): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(128, 256, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (downsample): Sequential(
        (0): Conv2d(128, 256, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
```

```
      )
    )
    (1): BasicBlock(
      (conv1): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(256, 256, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(256, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (layer4): Sequential(
    (0): BasicBlock(
      (conv1): Conv2d(256, 512, kernel_size=(3, 3), stride=(2, 2), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (downsample): Sequential(
        (0): Conv2d(256, 512, kernel_size=(1, 1), stride=(2, 2), bias=False)
        (1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      )
    )
    (1): BasicBlock(
      (conv1): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn1): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
      (relu): ReLU(inplace=True)
      (conv2): Conv2d(512, 512, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1), bias=False)
      (bn2): BatchNorm2d(512, eps=1e-05, momentum=0.1, affine=True, track_running_stats=True)
    )
  )
  (avgpool): AdaptiveAvgPool2d(output_size=(1, 1))
  (fc): Linear(in_features=512, out_features=1000, bias=True)
)
```

In [35]:

```
criterion = nn.CrossEntropyLoss().cuda(GPU)
```

In [36]:

```
optimizer = torch.optim.SGD(model.parameters(), LR,
                            momentum=MOMENTUM,
                            weight_decay=WEIGHT_DECAY)
```

In [37]:

```
scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer, T_max=EPOCHS)
```

In [38]:

```
transform_train = transforms.Compose([
    transforms.Resize((256,256)),
    transforms.RandomCrop(IMG_SIZE, padding=4),
    transforms.RandomHorizontalFlip(),
```

```
    transforms.ToTensor(),
    transforms.Normalize(imagenet_mean_RGB, imagenet_std_RGB),
])
```

In [40]:
```
train_dataset = datasets.ImageFolder(
    TRAINDIR, transform=transform_train)
```

In [41]:
```
transform_val = transforms.Compose([
    transforms.Resize((256,256)),
    transforms.RandomCrop(IMG_SIZE, padding=4),
    transforms.ToTensor(),
    transforms.Normalize(imagenet_mean_RGB, imagenet_std_RGB),
])
```

In [42]:
```
val_dataset = datasets.ImageFolder(
    VALDIR, transform=transform_val)
```

In [43]:
```
train_loader = torch.utils.data.DataLoader(
        train_dataset, batch_size=TRAIN_BATCH, shuffle=True,
        num_workers=WORKERS, pin_memory=True, sampler=None)
```

In [44]:
```
val_loader = torch.utils.data.DataLoader(
        val_dataset, batch_size=VAL_BATCH, shuffle=False,
        num_workers=WORKERS, pin_memory=True, sampler=None)
```

In [45]:
```
best_acc1 = 0
```

In [46]:
```
for epoch in range(START_EPOCH, 2):
#    adjust_learning_rate(optimizer, epoch)

    # train for one epoch
    train(train_loader, model, criterion, optimizer, epoch)

    # evaluate on validation set
    acc1 = validate(val_loader, model, criterion)

    # remember best acc@1 and save checkpoint
    is_best = acc1 > best_acc1
    best_acc1 = max(acc1, best_acc1)
```

```
    save_checkpoint({
        'epoch': epoch + 1,
        'arch': ARCH,
        'state_dict': model.state_dict(),
        'best_acc1': best_acc1,
        'optimizer' : optimizer.state_dict(),
    }, is_best)

    scheduler.step()
    print('lr: ' + str(scheduler.get_last_lr()[0]))

    writer.add_scalar("lr", scheduler.get_last_lr()[0], global_step = global_step)

    wandb.log({'lr': scheduler.get_last_lr()[0]})
```

```
Epoch: [0][   0/2563]   Time 15.712 (15.712)   Data  4.016 ( 4.016)   Loss 7.0163e+00 (7.0163e+00)   Acc@1   0.20 (
0.20)   Acc@5   0.60 (  0.60)
Epoch: [0][  10/2563]   Time  0.738 ( 2.468)   Data  0.095 ( 0.803)   Loss 7.0141e+00 (6.9886e+00)   Acc@1   0.40 (
0.09)   Acc@5   1.40 (  0.69)
Epoch: [0][  20/2563]   Time  0.816 ( 2.069)   Data  0.170 ( 0.884)   Loss 7.0331e+00 (7.0013e+00)   Acc@1   0.20 (
0.18)   Acc@5   0.60 (  0.78)
Epoch: [0][  30/2563]   Time  1.625 ( 1.966)   Data  0.979 ( 0.955)   Loss 6.8999e+00 (6.9824e+00)   Acc@1   0.00 (
0.24)   Acc@5   1.60 (  0.97)
Epoch: [0][  40/2563]   Time  1.480 ( 1.876)   Data  0.830 ( 0.953)   Loss 6.8295e+00 (6.9551e+00)   Acc@1   0.20 (
0.28)   Acc@5   1.40 (  1.13)
Epoch: [0][  50/2563]   Time  1.775 ( 1.849)   Data  1.128 ( 0.979)   Loss 6.7658e+00 (6.9228e+00)   Acc@1   0.00 (
0.31)   Acc@5   1.20 (  1.29)
Epoch: [0][  60/2563]   Time  1.806 ( 1.806)   Data  1.156 ( 0.972)   Loss 6.6769e+00 (6.8872e+00)   Acc@1   0.40 (
0.38)   Acc@5   3.40 (  1.56)
Epoch: [0][  70/2563]   Time  1.765 ( 1.783)   Data  1.114 ( 0.975)   Loss 6.6679e+00 (6.8563e+00)   Acc@1   0.20 (
0.39)   Acc@5   2.80 (  1.68)
Epoch: [0][  80/2563]   Time  1.839 ( 1.762)   Data  1.185 ( 0.973)   Loss 6.5323e+00 (6.8242e+00)   Acc@1   1.40 (
0.41)   Acc@5   3.40 (  1.83)
Epoch: [0][  90/2563]   Time  1.879 ( 1.751)   Data  1.224 ( 0.977)   Loss 6.5566e+00 (6.7969e+00)   Acc@1   1.00 (
0.44)   Acc@5   2.80 (  1.96)
Epoch: [0][ 100/2563]   Time  2.071 ( 1.739)   Data  1.412 ( 0.977)   Loss 6.5304e+00 (6.7686e+00)   Acc@1   0.20 (
0.46)   Acc@5   2.40 (  2.08)
Epoch: [0][ 110/2563]   Time  2.194 ( 1.738)   Data  1.534 ( 0.984)   Loss 6.4966e+00 (6.7412e+00)   Acc@1   1.00 (
0.52)   Acc@5   3.40 (  2.28)
Epoch: [0][ 120/2563]   Time  2.661 ( 1.729)   Data  1.963 ( 0.982)   Loss 6.4257e+00 (6.7136e+00)   Acc@1   1.00 (
0.56)   Acc@5   4.20 (  2.43)
Epoch: [0][ 130/2563]   Time  2.458 ( 1.722)   Data  1.785 ( 0.981)   Loss 6.3550e+00 (6.6890e+00)   Acc@1   1.60 (
0.62)   Acc@5   4.40 (  2.59)
Epoch: [0][ 140/2563]   Time  2.268 ( 1.711)   Data  1.604 ( 0.976)   Loss 6.2774e+00 (6.6635e+00)   Acc@1   0.80 (
0.68)   Acc@5   5.40 (  2.77)
Epoch: [0][ 150/2563]   Time  2.183 ( 1.706)   Data  1.524 ( 0.975)   Loss 6.3094e+00 (6.6417e+00)   Acc@1   0.40 (
0.72)   Acc@5   2.60 (  2.91)
Epoch: [0][ 160/2563]   Time  2.186 ( 1.704)   Data  1.524 ( 0.977)   Loss 6.3103e+00 (6.6182e+00)   Acc@1   1.60 (
0.76)   Acc@5   5.00 (  3.06)
Epoch: [0][ 170/2563]   Time  2.116 ( 1.697)   Data  1.452 ( 0.973)   Loss 6.1483e+00 (6.5953e+00)   Acc@1   2.40 (
```

```
0.81)   Acc@5   4.80 (  3.20)
Epoch: [0][ 180/2563]   Time  2.564 ( 1.692)   Data  1.871 ( 0.972)   Loss 6.0714e+00 (6.5710e+00)   Acc@1   2.60 (
0.86)   Acc@5   6.00 (  3.35)
Epoch: [0][ 190/2563]   Time  2.348 ( 1.693)   Data  1.681 ( 0.976)   Loss 6.0891e+00 (6.5489e+00)   Acc@1   2.00 (
0.93)   Acc@5   8.80 (  3.52)
Epoch: [0][ 200/2563]   Time  1.685 ( 1.688)   Data  1.020 ( 0.973)   Loss 6.0037e+00 (6.5260e+00)   Acc@1   2.00 (
1.00)   Acc@5   6.60 (  3.69)
Epoch: [0][ 210/2563]   Time  2.203 ( 1.691)   Data  1.538 ( 0.979)   Loss 6.0623e+00 (6.5038e+00)   Acc@1   1.80 (
1.06)   Acc@5   5.60 (  3.87)
Epoch: [0][ 220/2563]   Time  2.557 ( 1.689)   Data  1.844 ( 0.978)   Loss 6.1123e+00 (6.4836e+00)   Acc@1   2.80 (
1.12)   Acc@5   7.60 (  4.03)
Epoch: [0][ 230/2563]   Time  3.219 ( 1.689)   Data  2.511 ( 0.980)   Loss 6.0717e+00 (6.4632e+00)   Acc@1   2.00 (
1.19)   Acc@5   8.00 (  4.22)
Epoch: [0][ 240/2563]   Time  1.383 ( 1.685)   Data  0.716 ( 0.978)   Loss 6.0003e+00 (6.4416e+00)   Acc@1   3.40 (
1.26)   Acc@5   7.20 (  4.40)
Epoch: [0][ 250/2563]   Time  1.647 ( 1.686)   Data  0.983 ( 0.980)   Loss 5.8504e+00 (6.4208e+00)   Acc@1   1.80 (
1.32)   Acc@5   9.40 (  4.56)
Epoch: [0][ 260/2563]   Time  1.466 ( 1.684)   Data  0.803 ( 0.980)   Loss 6.0090e+00 (6.4020e+00)   Acc@1   2.40 (
1.37)   Acc@5   8.80 (  4.71)
Epoch: [0][ 270/2563]   Time  1.392 ( 1.682)   Data  0.726 ( 0.980)   Loss 5.8805e+00 (6.3829e+00)   Acc@1   3.00 (
1.41)   Acc@5   9.40 (  4.86)
Epoch: [0][ 280/2563]   Time  1.867 ( 1.679)   Data  1.199 ( 0.978)   Loss 5.7891e+00 (6.3651e+00)   Acc@1   3.20 (
1.46)   Acc@5  10.20 (  5.02)
Epoch: [0][ 290/2563]   Time  1.588 ( 1.677)   Data  0.922 ( 0.977)   Loss 5.7527e+00 (6.3461e+00)   Acc@1   2.40 (
1.50)   Acc@5   8.40 (  5.18)
Epoch: [0][ 300/2563]   Time  1.797 ( 1.679)   Data  1.131 ( 0.980)   Loss 5.7673e+00 (6.3285e+00)   Acc@1   2.40 (
1.53)   Acc@5  10.20 (  5.33)
Epoch: [0][ 310/2563]   Time  1.783 ( 1.676)   Data  1.119 ( 0.978)   Loss 5.7675e+00 (6.3109e+00)   Acc@1   3.40 (
1.58)   Acc@5   9.80 (  5.49)
Epoch: [0][ 320/2563]   Time  1.655 ( 1.676)   Data  0.987 ( 0.979)   Loss 5.7228e+00 (6.2932e+00)   Acc@1   3.00 (
1.64)   Acc@5  12.40 (  5.67)
Epoch: [0][ 330/2563]   Time  1.986 ( 1.674)   Data  1.318 ( 0.978)   Loss 5.6598e+00 (6.2752e+00)   Acc@1   2.80 (
1.69)   Acc@5  10.80 (  5.84)
Epoch: [0][ 340/2563]   Time  2.105 ( 1.675)   Data  1.440 ( 0.979)   Loss 5.6638e+00 (6.2581e+00)   Acc@1   3.00 (
1.75)   Acc@5  10.80 (  6.02)
Epoch: [0][ 350/2563]   Time  2.329 ( 1.672)   Data  1.661 ( 0.977)   Loss 5.6772e+00 (6.2418e+00)   Acc@1   4.80 (
1.81)   Acc@5  12.80 (  6.18)
Epoch: [0][ 360/2563]   Time  2.129 ( 1.672)   Data  1.465 ( 0.978)   Loss 5.6808e+00 (6.2250e+00)   Acc@1   3.20 (
1.87)   Acc@5  12.60 (  6.37)
Epoch: [0][ 370/2563]   Time  2.424 ( 1.671)   Data  1.758 ( 0.977)   Loss 5.7126e+00 (6.2088e+00)   Acc@1   2.60 (
1.92)   Acc@5  11.20 (  6.55)
Epoch: [0][ 380/2563]   Time  2.427 ( 1.669)   Data  1.758 ( 0.976)   Loss 5.6815e+00 (6.1928e+00)   Acc@1   6.00 (
1.99)   Acc@5  14.00 (  6.73)
Epoch: [0][ 390/2563]   Time  1.993 ( 1.669)   Data  1.326 ( 0.977)   Loss 5.6478e+00 (6.1777e+00)   Acc@1   3.40 (
2.03)   Acc@5  10.60 (  6.88)
Epoch: [0][ 400/2563]   Time  3.335 ( 1.671)   Data  2.601 ( 0.979)   Loss 5.4771e+00 (6.1620e+00)   Acc@1   4.80 (
2.09)   Acc@5  14.40 (  7.05)
Epoch: [0][ 410/2563]   Time  2.513 ( 1.669)   Data  1.800 ( 0.978)   Loss 5.4550e+00 (6.1467e+00)   Acc@1   5.00 (
2.13)   Acc@5  14.40 (  7.21)
Epoch: [0][ 420/2563]   Time  2.316 ( 1.666)   Data  1.651 ( 0.975)   Loss 5.5511e+00 (6.1310e+00)   Acc@1   5.00 (
2.20)   Acc@5  14.00 (  7.39)
```

```
Epoch: [0][ 430/2563]   Time  1.814 ( 1.666)   Data  1.150 ( 0.976)   Loss 5.3799e+00 (6.1151e+00)   Acc@1   5.20 (
2.27)   Acc@5  16.00 (  7.58)
Epoch: [0][ 440/2563]   Time  2.239 ( 1.665)   Data  1.574 ( 0.975)   Loss 5.4187e+00 (6.1005e+00)   Acc@1   5.80 (
2.34)   Acc@5  16.20 (  7.76)
Epoch: [0][ 450/2563]   Time  1.887 ( 1.666)   Data  1.222 ( 0.976)   Loss 5.4137e+00 (6.0851e+00)   Acc@1   5.20 (
2.40)   Acc@5  17.00 (  7.95)
Epoch: [0][ 460/2563]   Time  1.968 ( 1.665)   Data  1.301 ( 0.976)   Loss 5.4173e+00 (6.0706e+00)   Acc@1   6.20 (
2.46)   Acc@5  17.00 (  8.12)
Epoch: [0][ 470/2563]   Time  1.966 ( 1.665)   Data  1.301 ( 0.976)   Loss 5.2361e+00 (6.0565e+00)   Acc@1   5.80 (
2.52)   Acc@5  19.60 (  8.29)
Epoch: [0][ 480/2563]   Time  1.613 ( 1.665)   Data  0.948 ( 0.977)   Loss 5.2176e+00 (6.0416e+00)   Acc@1   6.00 (
2.59)   Acc@5  18.80 (  8.46)
Epoch: [0][ 490/2563]   Time  1.245 ( 1.662)   Data  0.578 ( 0.975)   Loss 5.3641e+00 (6.0272e+00)   Acc@1   5.60 (
2.67)   Acc@5  16.80 (  8.65)
Epoch: [0][ 500/2563]   Time  1.739 ( 1.664)   Data  1.073 ( 0.976)   Loss 5.3415e+00 (6.0128e+00)   Acc@1   7.20 (
2.73)   Acc@5  19.40 (  8.83)
Epoch: [0][ 510/2563]   Time  1.557 ( 1.663)   Data  0.890 ( 0.976)   Loss 5.3418e+00 (5.9981e+00)   Acc@1   5.80 (
2.80)   Acc@5  16.80 (  9.02)
Epoch: [0][ 520/2563]   Time  1.449 ( 1.663)   Data  0.784 ( 0.977)   Loss 5.1992e+00 (5.9850e+00)   Acc@1   4.80 (
2.85)   Acc@5  18.60 (  9.17)
Epoch: [0][ 530/2563]   Time  1.272 ( 1.661)   Data  0.601 ( 0.975)   Loss 5.1796e+00 (5.9712e+00)   Acc@1   7.20 (
2.92)   Acc@5  21.20 (  9.36)
Epoch: [0][ 540/2563]   Time  2.010 ( 1.662)   Data  1.343 ( 0.976)   Loss 5.2578e+00 (5.9585e+00)   Acc@1   6.40 (
2.98)   Acc@5  19.60 (  9.52)
Epoch: [0][ 550/2563]   Time  1.639 ( 1.661)   Data  0.972 ( 0.975)   Loss 5.2618e+00 (5.9460e+00)   Acc@1   5.60 (
3.03)   Acc@5  20.40 (  9.68)
Epoch: [0][ 560/2563]   Time  1.831 ( 1.660)   Data  1.164 ( 0.975)   Loss 5.2582e+00 (5.9337e+00)   Acc@1   6.80 (
3.09)   Acc@5  20.60 (  9.85)
Epoch: [0][ 570/2563]   Time  2.113 ( 1.661)   Data  1.447 ( 0.977)   Loss 5.1976e+00 (5.9206e+00)   Acc@1   7.40 (
3.16)   Acc@5  20.20 ( 10.02)
Epoch: [0][ 580/2563]   Time  2.378 ( 1.661)   Data  1.714 ( 0.976)   Loss 5.2602e+00 (5.9078e+00)   Acc@1   6.80 (
3.21)   Acc@5  19.20 ( 10.18)
Epoch: [0][ 590/2563]   Time  2.471 ( 1.661)   Data  1.806 ( 0.977)   Loss 5.1259e+00 (5.8952e+00)   Acc@1   8.20 (
3.27)   Acc@5  22.00 ( 10.35)
Epoch: [0][ 600/2563]   Time  2.455 ( 1.660)   Data  1.740 ( 0.976)   Loss 5.1782e+00 (5.8829e+00)   Acc@1   8.40 (
3.33)   Acc@5  21.40 ( 10.51)
Epoch: [0][ 610/2563]   Time  2.418 ( 1.659)   Data  1.751 ( 0.975)   Loss 5.2022e+00 (5.8709e+00)   Acc@1   6.80 (
3.39)   Acc@5  17.60 ( 10.66)
Epoch: [0][ 620/2563]   Time  2.212 ( 1.662)   Data  1.548 ( 0.978)   Loss 5.2333e+00 (5.8593e+00)   Acc@1   7.20 (
3.45)   Acc@5  18.20 ( 10.81)
Epoch: [0][ 630/2563]   Time  1.932 ( 1.661)   Data  1.269 ( 0.978)   Loss 5.2023e+00 (5.8479e+00)   Acc@1   7.80 (
3.52)   Acc@5  19.40 ( 10.96)
Epoch: [0][ 640/2563]   Time  1.783 ( 1.662)   Data  1.118 ( 0.979)   Loss 5.1473e+00 (5.8362e+00)   Acc@1   7.20 (
3.57)   Acc@5  23.40 ( 11.12)
Epoch: [0][ 650/2563]   Time  1.231 ( 1.660)   Data  0.565 ( 0.977)   Loss 5.0663e+00 (5.8247e+00)   Acc@1   8.20 (
3.63)   Acc@5  20.20 ( 11.27)
Epoch: [0][ 660/2563]   Time  1.692 ( 1.661)   Data  1.029 ( 0.978)   Loss 5.0788e+00 (5.8140e+00)   Acc@1   7.60 (
3.69)   Acc@5  22.00 ( 11.42)
Epoch: [0][ 670/2563]   Time  0.756 ( 1.660)   Data  0.093 ( 0.978)   Loss 5.0994e+00 (5.8024e+00)   Acc@1   8.80 (
3.75)   Acc@5  20.60 ( 11.57)
Epoch: [0][ 680/2563]   Time  0.822 ( 1.659)   Data  0.154 ( 0.977)   Loss 5.0015e+00 (5.7910e+00)   Acc@1   7.00 (
```

```
3.82)   Acc@5  24.20 ( 11.73)
Epoch: [0][ 690/2563]   Time  1.249 ( 1.660)    Data  0.584 ( 0.978)    Loss 5.0175e+00 (5.7797e+00)    Acc@1   7.60 (
3.88)   Acc@5  23.40 ( 11.90)
Epoch: [0][ 700/2563]   Time  1.291 ( 1.660)    Data  0.621 ( 0.977)    Loss 5.0960e+00 (5.7690e+00)    Acc@1   8.20 (
3.95)   Acc@5  21.00 ( 12.04)
Epoch: [0][ 710/2563]   Time  1.171 ( 1.661)    Data  0.504 ( 0.979)    Loss 5.0522e+00 (5.7580e+00)    Acc@1   8.00 (
4.02)   Acc@5  20.40 ( 12.19)
Epoch: [0][ 720/2563]   Time  1.562 ( 1.660)    Data  0.895 ( 0.978)    Loss 4.9474e+00 (5.7473e+00)    Acc@1   8.80 (
4.08)   Acc@5  25.00 ( 12.34)
Epoch: [0][ 730/2563]   Time  1.977 ( 1.660)    Data  1.311 ( 0.979)    Loss 5.0513e+00 (5.7367e+00)    Acc@1   8.00 (
4.14)   Acc@5  23.00 ( 12.50)
Epoch: [0][ 740/2563]   Time  1.792 ( 1.660)    Data  1.125 ( 0.979)    Loss 5.0149e+00 (5.7262e+00)    Acc@1  11.40 (
4.20)   Acc@5  24.80 ( 12.65)
Epoch: [0][ 750/2563]   Time  2.008 ( 1.660)    Data  1.339 ( 0.979)    Loss 4.8676e+00 (5.7149e+00)    Acc@1  11.20 (
4.28)   Acc@5  26.20 ( 12.81)
Epoch: [0][ 760/2563]   Time  1.775 ( 1.661)    Data  1.107 ( 0.979)    Loss 5.0174e+00 (5.7046e+00)    Acc@1  10.40 (
4.34)   Acc@5  25.00 ( 12.96)
Epoch: [0][ 770/2563]   Time  1.937 ( 1.660)    Data  1.269 ( 0.979)    Loss 4.9143e+00 (5.6939e+00)    Acc@1  11.40 (
4.41)   Acc@5  26.20 ( 13.11)
Epoch: [0][ 780/2563]   Time  2.441 ( 1.660)    Data  1.774 ( 0.980)    Loss 5.0148e+00 (5.6837e+00)    Acc@1   8.20 (
4.48)   Acc@5  23.80 ( 13.26)
Epoch: [0][ 790/2563]   Time  2.338 ( 1.659)    Data  1.676 ( 0.979)    Loss 4.8827e+00 (5.6731e+00)    Acc@1   9.60 (
4.55)   Acc@5  27.20 ( 13.42)
Epoch: [0][ 800/2563]   Time  2.297 ( 1.659)    Data  1.629 ( 0.979)    Loss 4.7108e+00 (5.6631e+00)    Acc@1  11.20 (
4.60)   Acc@5  26.80 ( 13.56)
Epoch: [0][ 810/2563]   Time  2.125 ( 1.659)    Data  1.461 ( 0.979)    Loss 4.9814e+00 (5.6532e+00)    Acc@1   8.60 (
4.67)   Acc@5  23.00 ( 13.71)
Epoch: [0][ 820/2563]   Time  1.775 ( 1.658)    Data  1.108 ( 0.978)    Loss 4.9274e+00 (5.6437e+00)    Acc@1   9.60 (
4.74)   Acc@5  25.40 ( 13.85)
Epoch: [0][ 830/2563]   Time  2.344 ( 1.659)    Data  1.691 ( 0.979)    Loss 4.7979e+00 (5.6335e+00)    Acc@1   9.20 (
4.81)   Acc@5  26.20 ( 14.00)
Epoch: [0][ 840/2563]   Time  1.750 ( 1.658)    Data  1.106 ( 0.978)    Loss 4.8107e+00 (5.6235e+00)    Acc@1   7.60 (
4.87)   Acc@5  24.20 ( 14.14)
Epoch: [0][ 850/2563]   Time  1.489 ( 1.657)    Data  0.849 ( 0.978)    Loss 4.6643e+00 (5.6134e+00)    Acc@1  10.20 (
 4.94)  Acc@5  25.60 ( 14.29)
Epoch: [0][ 860/2563]   Time  1.532 ( 1.658)    Data  0.896 ( 0.979)    Loss 4.6555e+00 (5.6038e+00)    Acc@1  12.60 (
 5.00)  Acc@5  28.60 ( 14.43)
Epoch: [0][ 870/2563]   Time  2.392 ( 1.658)    Data  1.755 ( 0.979)    Loss 4.8971e+00 (5.5944e+00)    Acc@1  12.40 (
 5.06)  Acc@5  25.80 ( 14.57)
Epoch: [0][ 880/2563]   Time  2.329 ( 1.658)    Data  1.693 ( 0.980)    Loss 4.7383e+00 (5.5846e+00)    Acc@1  11.80 (
 5.13)  Acc@5  28.20 ( 14.72)
Epoch: [0][ 890/2563]   Time  1.769 ( 1.657)    Data  1.133 ( 0.979)    Loss 4.8191e+00 (5.5745e+00)    Acc@1  11.00 (
 5.21)  Acc@5  25.80 ( 14.87)
Epoch: [0][ 900/2563]   Time  1.612 ( 1.656)    Data  0.975 ( 0.979)    Loss 4.5753e+00 (5.5646e+00)    Acc@1  13.40 (
 5.28)  Acc@5  31.40 ( 15.03)
Epoch: [0][ 910/2563]   Time  1.927 ( 1.656)    Data  1.289 ( 0.980)    Loss 4.7023e+00 (5.5551e+00)    Acc@1   9.60 (
 5.34)  Acc@5  27.80 ( 15.17)
Epoch: [0][ 920/2563]   Time  1.642 ( 1.656)    Data  1.009 ( 0.979)    Loss 4.5928e+00 (5.5459e+00)    Acc@1  14.60 (
 5.41)  Acc@5  31.60 ( 15.32)
Epoch: [0][ 930/2563]   Time  1.250 ( 1.656)    Data  0.611 ( 0.981)    Loss 4.5639e+00 (5.5367e+00)    Acc@1   9.40 (
 5.47)  Acc@5  26.20 ( 15.46)
```

```
Epoch: [0][ 940/2563]   Time  1.513 ( 1.655)   Data  0.876 ( 0.980)   Loss 4.6244e+00 (5.5273e+00)   Acc@1  12.80 (
 5.55)  Acc@5  32.60 ( 15.60)
Epoch: [0][ 950/2563]   Time  1.662 ( 1.655)   Data  1.026 ( 0.980)   Loss 4.7114e+00 (5.5182e+00)   Acc@1  10.40 (
 5.61)  Acc@5  26.40 ( 15.74)
Epoch: [0][ 960/2563]   Time  1.923 ( 1.655)   Data  1.287 ( 0.980)   Loss 4.5753e+00 (5.5092e+00)   Acc@1  13.20 (
 5.68)  Acc@5  30.20 ( 15.88)
Epoch: [0][ 970/2563]   Time  1.653 ( 1.654)   Data  1.015 ( 0.980)   Loss 4.6537e+00 (5.5001e+00)   Acc@1  12.00 (
 5.75)  Acc@5  26.60 ( 16.02)
Epoch: [0][ 980/2563]   Time  2.188 ( 1.655)   Data  1.549 ( 0.981)   Loss 4.6582e+00 (5.4912e+00)   Acc@1  11.80 (
 5.81)  Acc@5  31.20 ( 16.16)
Epoch: [0][ 990/2563]   Time  2.274 ( 1.654)   Data  1.634 ( 0.981)   Loss 4.7361e+00 (5.4825e+00)   Acc@1   9.80 (
 5.87)  Acc@5  27.40 ( 16.29)
Epoch: [0][1000/2563]   Time  2.385 ( 1.654)   Data  1.744 ( 0.980)   Loss 4.5814e+00 (5.4735e+00)   Acc@1  12.40 (
 5.94)  Acc@5  31.00 ( 16.44)
Epoch: [0][1010/2563]   Time  2.141 ( 1.654)   Data  1.496 ( 0.980)   Loss 4.5763e+00 (5.4647e+00)   Acc@1  10.60 (
 6.00)  Acc@5  33.00 ( 16.58)
Epoch: [0][1020/2563]   Time  1.865 ( 1.654)   Data  1.220 ( 0.981)   Loss 4.6481e+00 (5.4562e+00)   Acc@1  13.60 (
 6.07)  Acc@5  30.20 ( 16.72)
Epoch: [0][1030/2563]   Time  1.941 ( 1.653)   Data  1.293 ( 0.981)   Loss 4.4861e+00 (5.4478e+00)   Acc@1  12.00 (
 6.13)  Acc@5  30.00 ( 16.84)
Epoch: [0][1040/2563]   Time  1.542 ( 1.653)   Data  0.894 ( 0.981)   Loss 4.6175e+00 (5.4393e+00)   Acc@1  12.00 (
 6.20)  Acc@5  29.40 ( 16.97)
Epoch: [0][1050/2563]   Time  1.435 ( 1.653)   Data  0.787 ( 0.980)   Loss 4.5546e+00 (5.4309e+00)   Acc@1  10.20 (
 6.25)  Acc@5  31.80 ( 17.11)
Epoch: [0][1060/2563]   Time  1.149 ( 1.653)   Data  0.497 ( 0.981)   Loss 4.4541e+00 (5.4218e+00)   Acc@1  14.80 (
 6.33)  Acc@5  33.20 ( 17.25)
Epoch: [0][1070/2563]   Time  0.996 ( 1.652)   Data  0.346 ( 0.980)   Loss 4.4926e+00 (5.4132e+00)   Acc@1  13.00 (
 6.39)  Acc@5  33.20 ( 17.39)
Epoch: [0][1080/2563]   Time  0.753 ( 1.651)   Data  0.101 ( 0.980)   Loss 4.4187e+00 (5.4048e+00)   Acc@1  15.00 (
 6.45)  Acc@5  33.00 ( 17.52)
Epoch: [0][1090/2563]   Time  0.930 ( 1.652)   Data  0.274 ( 0.980)   Loss 4.6871e+00 (5.3965e+00)   Acc@1  11.80 (
 6.52)  Acc@5  30.80 ( 17.66)
Epoch: [0][1100/2563]   Time  0.755 ( 1.652)   Data  0.097 ( 0.980)   Loss 4.4457e+00 (5.3884e+00)   Acc@1  16.20 (
 6.60)  Acc@5  33.20 ( 17.80)
Epoch: [0][1110/2563]   Time  1.073 ( 1.651)   Data  0.413 ( 0.980)   Loss 4.4673e+00 (5.3795e+00)   Acc@1  14.40 (
 6.66)  Acc@5  32.20 ( 17.94)
Epoch: [0][1120/2563]   Time  0.795 ( 1.651)   Data  0.139 ( 0.979)   Loss 4.4549e+00 (5.3714e+00)   Acc@1  13.40 (
 6.73)  Acc@5  33.60 ( 18.08)
Epoch: [0][1130/2563]   Time  1.198 ( 1.651)   Data  0.541 ( 0.980)   Loss 4.4523e+00 (5.3630e+00)   Acc@1  14.60 (
 6.80)  Acc@5  31.60 ( 18.21)
Epoch: [0][1140/2563]   Time  1.132 ( 1.651)   Data  0.474 ( 0.979)   Loss 4.3032e+00 (5.3548e+00)   Acc@1  16.60 (
 6.87)  Acc@5  36.40 ( 18.34)
Epoch: [0][1150/2563]   Time  1.734 ( 1.652)   Data  1.075 ( 0.980)   Loss 4.4417e+00 (5.3470e+00)   Acc@1  13.20 (
 6.93)  Acc@5  31.00 ( 18.47)
Epoch: [0][1160/2563]   Time  1.512 ( 1.651)   Data  0.852 ( 0.980)   Loss 4.4288e+00 (5.3387e+00)   Acc@1  14.80 (
 7.00)  Acc@5  32.80 ( 18.61)
Epoch: [0][1170/2563]   Time  1.301 ( 1.651)   Data  0.646 ( 0.980)   Loss 4.4443e+00 (5.3307e+00)   Acc@1  15.40 (
 7.06)  Acc@5  35.20 ( 18.75)
Epoch: [0][1180/2563]   Time  1.753 ( 1.651)   Data  1.095 ( 0.980)   Loss 4.2035e+00 (5.3227e+00)   Acc@1  18.60 (
 7.12)  Acc@5  39.40 ( 18.88)
Epoch: [0][1190/2563]   Time  1.179 ( 1.651)   Data  0.521 ( 0.980)   Loss 4.2165e+00 (5.3146e+00)   Acc@1  15.00 (
```

```
 7.19)  Acc@5  39.00 ( 19.01)
Epoch: [0][1200/2563]   Time  1.374 ( 1.651)    Data  0.731 ( 0.981)    Loss 4.3920e+00 (5.3064e+00)    Acc@1  17.00 (
 7.26)  Acc@5  32.60 ( 19.15)
Epoch: [0][1210/2563]   Time  1.829 ( 1.651)    Data  1.190 ( 0.981)    Loss 4.2160e+00 (5.2983e+00)    Acc@1  17.60 (
 7.33)  Acc@5  37.00 ( 19.28)
Epoch: [0][1220/2563]   Time  1.606 ( 1.651)    Data  0.966 ( 0.980)    Loss 4.3465e+00 (5.2904e+00)    Acc@1  14.20 (
 7.39)  Acc@5  35.00 ( 19.41)
Epoch: [0][1230/2563]   Time  1.479 ( 1.651)    Data  0.841 ( 0.981)    Loss 4.3133e+00 (5.2824e+00)    Acc@1  13.20 (
 7.46)  Acc@5  34.20 ( 19.54)
Epoch: [0][1240/2563]   Time  2.099 ( 1.651)    Data  1.464 ( 0.981)    Loss 4.3256e+00 (5.2744e+00)    Acc@1  14.40 (
 7.52)  Acc@5  32.60 ( 19.67)
Epoch: [0][1250/2563]   Time  1.479 ( 1.650)    Data  0.845 ( 0.981)    Loss 4.3818e+00 (5.2669e+00)    Acc@1  15.60 (
 7.59)  Acc@5  35.00 ( 19.80)
Epoch: [0][1260/2563]   Time  1.198 ( 1.650)    Data  0.563 ( 0.981)    Loss 4.4859e+00 (5.2591e+00)    Acc@1  14.60 (
 7.65)  Acc@5  32.20 ( 19.93)
Epoch: [0][1270/2563]   Time  1.493 ( 1.649)    Data  0.859 ( 0.981)    Loss 4.1834e+00 (5.2517e+00)    Acc@1  15.00 (
 7.71)  Acc@5  37.60 ( 20.05)
Epoch: [0][1280/2563]   Time  1.739 ( 1.649)    Data  1.103 ( 0.981)    Loss 4.2486e+00 (5.2441e+00)    Acc@1  17.00 (
 7.78)  Acc@5  38.00 ( 20.18)
Epoch: [0][1290/2563]   Time  2.431 ( 1.649)    Data  1.795 ( 0.981)    Loss 4.3873e+00 (5.2367e+00)    Acc@1  13.40 (
 7.84)  Acc@5  34.20 ( 20.30)
Epoch: [0][1300/2563]   Time  2.372 ( 1.648)    Data  1.715 ( 0.980)    Loss 4.2243e+00 (5.2290e+00)    Acc@1  16.80 (
 7.91)  Acc@5  36.40 ( 20.43)
Epoch: [0][1310/2563]   Time  2.124 ( 1.649)    Data  1.488 ( 0.981)    Loss 4.4700e+00 (5.2214e+00)    Acc@1  11.80 (
 7.97)  Acc@5  32.60 ( 20.56)
Epoch: [0][1320/2563]   Time  1.734 ( 1.648)    Data  1.095 ( 0.980)    Loss 4.2420e+00 (5.2141e+00)    Acc@1  18.00 (
 8.04)  Acc@5  36.80 ( 20.68)
Epoch: [0][1330/2563]   Time  1.776 ( 1.648)    Data  1.135 ( 0.981)    Loss 4.2025e+00 (5.2066e+00)    Acc@1  15.00 (
 8.10)  Acc@5  40.20 ( 20.81)
Epoch: [0][1340/2563]   Time  1.796 ( 1.648)    Data  1.157 ( 0.981)    Loss 4.2131e+00 (5.1995e+00)    Acc@1  18.60 (
 8.16)  Acc@5  38.00 ( 20.93)
Epoch: [0][1350/2563]   Time  2.098 ( 1.648)    Data  1.460 ( 0.981)    Loss 4.1080e+00 (5.1922e+00)    Acc@1  17.80 (
 8.23)  Acc@5  40.40 ( 21.05)
Epoch: [0][1360/2563]   Time  1.815 ( 1.647)    Data  1.175 ( 0.981)    Loss 4.1759e+00 (5.1848e+00)    Acc@1  16.60 (
 8.29)  Acc@5  39.00 ( 21.18)
Epoch: [0][1370/2563]   Time  2.076 ( 1.647)    Data  1.434 ( 0.981)    Loss 4.1148e+00 (5.1773e+00)    Acc@1  18.20 (
 8.36)  Acc@5  36.40 ( 21.30)
Epoch: [0][1380/2563]   Time  2.576 ( 1.647)    Data  1.888 ( 0.981)    Loss 4.0611e+00 (5.1698e+00)    Acc@1  18.00 (
 8.43)  Acc@5  39.80 ( 21.43)
Epoch: [0][1390/2563]   Time  1.830 ( 1.647)    Data  1.184 ( 0.980)    Loss 4.1723e+00 (5.1625e+00)    Acc@1  17.20 (
 8.49)  Acc@5  41.60 ( 21.56)
Epoch: [0][1400/2563]   Time  1.613 ( 1.647)    Data  0.965 ( 0.981)    Loss 4.0025e+00 (5.1550e+00)    Acc@1  19.20 (
 8.56)  Acc@5  40.80 ( 21.69)
Epoch: [0][1410/2563]   Time  1.298 ( 1.646)    Data  0.652 ( 0.980)    Loss 4.0490e+00 (5.1481e+00)    Acc@1  19.60 (
 8.62)  Acc@5  44.20 ( 21.81)
Epoch: [0][1420/2563]   Time  1.238 ( 1.647)    Data  0.588 ( 0.981)    Loss 4.1697e+00 (5.1409e+00)    Acc@1  19.40 (
 8.68)  Acc@5  37.00 ( 21.93)
Epoch: [0][1430/2563]   Time  1.321 ( 1.647)    Data  0.672 ( 0.981)    Loss 4.0603e+00 (5.1334e+00)    Acc@1  17.20 (
 8.75)  Acc@5  39.60 ( 22.06)
Epoch: [0][1440/2563]   Time  1.501 ( 1.646)    Data  0.851 ( 0.980)    Loss 3.9960e+00 (5.1262e+00)    Acc@1  20.00 (
 8.82)  Acc@5  41.40 ( 22.19)
```

```
Epoch: [0][1450/2563]   Time  1.423 ( 1.647)   Data  0.769 ( 0.981)   Loss 4.1672e+00 (5.1192e+00)   Acc@1  17.80 (
 8.89)  Acc@5  39.40 ( 22.31)
Epoch: [0][1460/2563]   Time  0.751 ( 1.647)   Data  0.097 ( 0.981)   Loss 3.9982e+00 (5.1120e+00)   Acc@1  19.20 (
 8.95)  Acc@5  42.80 ( 22.44)
Epoch: [0][1470/2563]   Time  0.755 ( 1.647)   Data  0.100 ( 0.981)   Loss 3.9263e+00 (5.1050e+00)   Acc@1  20.20 (
 9.01)  Acc@5  43.40 ( 22.56)
Epoch: [0][1480/2563]   Time  0.657 ( 1.647)   Data  0.002 ( 0.981)   Loss 3.9839e+00 (5.0985e+00)   Acc@1  18.00 (
 9.07)  Acc@5  42.80 ( 22.67)
Epoch: [0][1490/2563]   Time  0.656 ( 1.647)   Data  0.002 ( 0.981)   Loss 4.1926e+00 (5.0918e+00)   Acc@1  18.60 (
 9.14)  Acc@5  37.80 ( 22.79)
Epoch: [0][1500/2563]   Time  0.662 ( 1.647)   Data  0.002 ( 0.981)   Loss 4.0177e+00 (5.0850e+00)   Acc@1  18.60 (
 9.19)  Acc@5  42.00 ( 22.91)
Epoch: [0][1510/2563]   Time  0.660 ( 1.647)   Data  0.003 ( 0.981)   Loss 4.1322e+00 (5.0784e+00)   Acc@1  17.80 (
 9.26)  Acc@5  39.60 ( 23.03)
Epoch: [0][1520/2563]   Time  0.660 ( 1.647)   Data  0.002 ( 0.980)   Loss 4.1766e+00 (5.0716e+00)   Acc@1  17.00 (
 9.32)  Acc@5  40.60 ( 23.15)
Epoch: [0][1530/2563]   Time  0.660 ( 1.647)   Data  0.002 ( 0.980)   Loss 3.8637e+00 (5.0644e+00)   Acc@1  22.60 (
 9.39)  Acc@5  44.60 ( 23.27)
Epoch: [0][1540/2563]   Time  0.667 ( 1.646)   Data  0.002 ( 0.980)   Loss 3.9573e+00 (5.0573e+00)   Acc@1  20.00 (
 9.46)  Acc@5  43.80 ( 23.39)
Epoch: [0][1550/2563]   Time  1.192 ( 1.647)   Data  0.529 ( 0.980)   Loss 4.0068e+00 (5.0502e+00)   Acc@1  19.80 (
 9.53)  Acc@5  40.60 ( 23.52)
Epoch: [0][1560/2563]   Time  1.388 ( 1.646)   Data  0.731 ( 0.980)   Loss 3.7967e+00 (5.0432e+00)   Acc@1  20.80 (
 9.60)  Acc@5  44.20 ( 23.64)
Epoch: [0][1570/2563]   Time  1.477 ( 1.646)   Data  0.833 ( 0.980)   Loss 4.0522e+00 (5.0373e+00)   Acc@1  19.00 (
 9.65)  Acc@5  41.20 ( 23.74)
Epoch: [0][1580/2563]   Time  1.643 ( 1.646)   Data  1.005 ( 0.980)   Loss 4.1465e+00 (5.0306e+00)   Acc@1  16.60 (
 9.72)  Acc@5  39.20 ( 23.86)
Epoch: [0][1590/2563]   Time  1.401 ( 1.646)   Data  0.766 ( 0.980)   Loss 4.0296e+00 (5.0239e+00)   Acc@1  19.00 (
 9.78)  Acc@5  43.60 ( 23.98)
Epoch: [0][1600/2563]   Time  1.796 ( 1.647)   Data  1.159 ( 0.981)   Loss 4.0175e+00 (5.0175e+00)   Acc@1  18.80 (
 9.84)  Acc@5  41.40 ( 24.09)
Epoch: [0][1610/2563]   Time  1.332 ( 1.646)   Data  0.697 ( 0.981)   Loss 3.9992e+00 (5.0112e+00)   Acc@1  17.80 (
 9.90)  Acc@5  43.00 ( 24.20)
Epoch: [0][1620/2563]   Time  1.613 ( 1.647)   Data  0.979 ( 0.981)   Loss 3.8597e+00 (5.0048e+00)   Acc@1  19.00 (
 9.97)  Acc@5  44.00 ( 24.31)
Epoch: [0][1630/2563]   Time  1.992 ( 1.646)   Data  1.356 ( 0.981)   Loss 3.8636e+00 (4.9981e+00)   Acc@1  18.80 ( 1
0.03)   Acc@5  46.40 ( 24.43)
Epoch: [0][1640/2563]   Time  2.504 ( 1.646)   Data  1.825 ( 0.981)   Loss 3.9545e+00 (4.9921e+00)   Acc@1  22.80 ( 1
0.09)   Acc@5  41.00 ( 24.54)
Epoch: [0][1650/2563]   Time  2.553 ( 1.646)   Data  1.886 ( 0.981)   Loss 4.0675e+00 (4.9861e+00)   Acc@1  19.60 ( 1
0.15)   Acc@5  41.20 ( 24.65)
Epoch: [0][1660/2563]   Time  2.648 ( 1.646)   Data  1.962 ( 0.981)   Loss 4.0999e+00 (4.9795e+00)   Acc@1  18.20 ( 1
0.21)   Acc@5  41.80 ( 24.76)
Epoch: [0][1670/2563]   Time  2.342 ( 1.645)   Data  1.706 ( 0.981)   Loss 3.9395e+00 (4.9733e+00)   Acc@1  18.80 ( 1
0.28)   Acc@5  43.00 ( 24.87)
Epoch: [0][1680/2563]   Time  2.652 ( 1.646)   Data  1.968 ( 0.981)   Loss 3.7528e+00 (4.9669e+00)   Acc@1  24.00 ( 1
0.35)   Acc@5  47.20 ( 24.99)
Epoch: [0][1690/2563]   Time  2.612 ( 1.646)   Data  1.926 ( 0.981)   Loss 4.0542e+00 (4.9603e+00)   Acc@1  20.40 ( 1
0.41)   Acc@5  41.00 ( 25.10)
Epoch: [0][1700/2563]   Time  3.358 ( 1.646)   Data  2.639 ( 0.981)   Loss 3.8043e+00 (4.9538e+00)   Acc@1  21.80 ( 1
```

```
0.48)   Acc@5  44.20 ( 25.22)
Epoch: [0][1710/2563]   Time  2.540 ( 1.646)    Data  1.853 ( 0.981)    Loss 3.8880e+00 (4.9476e+00)    Acc@1  22.20 ( 1
0.54)   Acc@5  44.60 ( 25.33)
Epoch: [0][1720/2563]   Time  2.518 ( 1.646)    Data  1.858 ( 0.981)    Loss 3.8443e+00 (4.9412e+00)    Acc@1  21.20 ( 1
0.61)   Acc@5  45.20 ( 25.44)
Epoch: [0][1730/2563]   Time  2.542 ( 1.646)    Data  1.866 ( 0.982)    Loss 3.7051e+00 (4.9346e+00)    Acc@1  22.60 ( 1
0.67)   Acc@5  49.20 ( 25.56)
Epoch: [0][1740/2563]   Time  2.679 ( 1.646)    Data  1.989 ( 0.981)    Loss 4.1604e+00 (4.9284e+00)    Acc@1  19.80 ( 1
0.73)   Acc@5  41.00 ( 25.67)
Epoch: [0][1750/2563]   Time  2.570 ( 1.646)    Data  1.877 ( 0.981)    Loss 3.8410e+00 (4.9222e+00)    Acc@1  20.40 ( 1
0.79)   Acc@5  43.00 ( 25.78)
Epoch: [0][1760/2563]   Time  2.486 ( 1.646)    Data  1.811 ( 0.981)    Loss 3.6163e+00 (4.9160e+00)    Acc@1  23.80 ( 1
0.85)   Acc@5  50.00 ( 25.89)
Epoch: [0][1770/2563]   Time  2.784 ( 1.647)    Data  2.093 ( 0.982)    Loss 3.6469e+00 (4.9099e+00)    Acc@1  27.00 ( 1
0.92)   Acc@5  48.20 ( 26.00)
Epoch: [0][1780/2563]   Time  2.722 ( 1.646)    Data  2.018 ( 0.982)    Loss 3.6066e+00 (4.9038e+00)    Acc@1  26.40 ( 1
0.98)   Acc@5  50.00 ( 26.11)
Epoch: [0][1790/2563]   Time  2.531 ( 1.646)    Data  1.832 ( 0.981)    Loss 3.8522e+00 (4.8976e+00)    Acc@1  24.00 ( 1
1.05)   Acc@5  45.60 ( 26.22)
Epoch: [0][1800/2563]   Time  2.422 ( 1.646)    Data  1.773 ( 0.981)    Loss 3.9128e+00 (4.8915e+00)    Acc@1  18.40 ( 1
1.11)   Acc@5  43.80 ( 26.33)
Epoch: [0][1810/2563]   Time  2.624 ( 1.646)    Data  1.921 ( 0.981)    Loss 3.8695e+00 (4.8851e+00)    Acc@1  20.80 ( 1
1.17)   Acc@5  43.20 ( 26.45)
Epoch: [0][1820/2563]   Time  2.436 ( 1.646)    Data  1.783 ( 0.981)    Loss 3.8543e+00 (4.8789e+00)    Acc@1  21.20 ( 1
1.24)   Acc@5  45.20 ( 26.56)
Epoch: [0][1830/2563]   Time  2.572 ( 1.646)    Data  1.857 ( 0.981)    Loss 3.8314e+00 (4.8729e+00)    Acc@1  20.80 ( 1
1.31)   Acc@5  45.60 ( 26.67)
Epoch: [0][1840/2563]   Time  2.697 ( 1.646)    Data  1.982 ( 0.981)    Loss 3.8130e+00 (4.8671e+00)    Acc@1  24.00 ( 1
1.36)   Acc@5  46.40 ( 26.77)
Epoch: [0][1850/2563]   Time  2.474 ( 1.645)    Data  1.822 ( 0.980)    Loss 3.8999e+00 (4.8613e+00)    Acc@1  20.20 ( 1
1.42)   Acc@5  44.60 ( 26.88)
Epoch: [0][1860/2563]   Time  2.605 ( 1.645)    Data  1.899 ( 0.980)    Loss 3.6072e+00 (4.8551e+00)    Acc@1  22.40 ( 1
1.49)   Acc@5  49.80 ( 26.99)
Epoch: [0][1870/2563]   Time  2.374 ( 1.645)    Data  1.712 ( 0.980)    Loss 3.6016e+00 (4.8490e+00)    Acc@1  27.80 ( 1
1.55)   Acc@5  49.00 ( 27.10)
Epoch: [0][1880/2563]   Time  2.554 ( 1.645)    Data  1.842 ( 0.980)    Loss 3.6917e+00 (4.8430e+00)    Acc@1  23.00 ( 1
1.61)   Acc@5  46.20 ( 27.21)
Epoch: [0][1890/2563]   Time  2.399 ( 1.645)    Data  1.740 ( 0.979)    Loss 3.6933e+00 (4.8371e+00)    Acc@1  21.20 ( 1
1.67)   Acc@5  46.20 ( 27.31)
Epoch: [0][1900/2563]   Time  2.554 ( 1.644)    Data  1.845 ( 0.979)    Loss 3.5803e+00 (4.8311e+00)    Acc@1  25.60 ( 1
1.73)   Acc@5  50.60 ( 27.42)
Epoch: [0][1910/2563]   Time  2.884 ( 1.645)    Data  2.157 ( 0.979)    Loss 3.7199e+00 (4.8250e+00)    Acc@1  25.40 ( 1
1.80)   Acc@5  48.20 ( 27.53)
Epoch: [0][1920/2563]   Time  2.642 ( 1.644)    Data  1.925 ( 0.979)    Loss 3.6995e+00 (4.8195e+00)    Acc@1  23.80 ( 1
1.85)   Acc@5  47.60 ( 27.62)
Epoch: [0][1930/2563]   Time  2.736 ( 1.644)    Data  2.026 ( 0.979)    Loss 3.5391e+00 (4.8136e+00)    Acc@1  29.00 ( 1
1.91)   Acc@5  52.20 ( 27.73)
Epoch: [0][1940/2563]   Time  2.577 ( 1.644)    Data  1.857 ( 0.978)    Loss 3.6781e+00 (4.8077e+00)    Acc@1  23.40 ( 1
1.98)   Acc@5  46.40 ( 27.83)
Epoch: [0][1950/2563]   Time  2.782 ( 1.644)    Data  2.071 ( 0.978)    Loss 3.7359e+00 (4.8019e+00)    Acc@1  22.20 ( 1
2.04)   Acc@5  47.20 ( 27.94)
```

```
Epoch: [0][1960/2563]   Time  2.478 ( 1.644)   Data  1.799 ( 0.978)   Loss 3.6846e+00 (4.7963e+00)   Acc@1  22.20 ( 1
2.09)   Acc@5  46.60 ( 28.04)
Epoch: [0][1970/2563]   Time  2.573 ( 1.644)   Data  1.868 ( 0.978)   Loss 3.8171e+00 (4.7909e+00)   Acc@1  23.40 ( 1
2.15)   Acc@5  47.40 ( 28.13)
Epoch: [0][1980/2563]   Time  2.859 ( 1.645)   Data  2.144 ( 0.979)   Loss 3.7431e+00 (4.7854e+00)   Acc@1  22.80 ( 1
2.20)   Acc@5  47.60 ( 28.23)
Epoch: [0][1990/2563]   Time  2.476 ( 1.644)   Data  1.796 ( 0.978)   Loss 3.6308e+00 (4.7796e+00)   Acc@1  21.60 ( 1
2.27)   Acc@5  49.20 ( 28.34)
Epoch: [0][2000/2563]   Time  2.721 ( 1.644)   Data  1.987 ( 0.978)   Loss 3.6888e+00 (4.7740e+00)   Acc@1  23.40 ( 1
2.33)   Acc@5  47.60 ( 28.44)
Epoch: [0][2010/2563]   Time  2.525 ( 1.645)   Data  1.824 ( 0.978)   Loss 3.7095e+00 (4.7682e+00)   Acc@1  21.80 ( 1
2.38)   Acc@5  47.80 ( 28.54)
Epoch: [0][2020/2563]   Time  2.456 ( 1.644)   Data  1.789 ( 0.978)   Loss 3.5944e+00 (4.7622e+00)   Acc@1  26.20 ( 1
2.45)   Acc@5  50.20 ( 28.64)
Epoch: [0][2030/2563]   Time  2.391 ( 1.645)   Data  1.728 ( 0.978)   Loss 3.6925e+00 (4.7568e+00)   Acc@1  22.40 ( 1
2.51)   Acc@5  46.60 ( 28.74)
Epoch: [0][2040/2563]   Time  2.726 ( 1.645)   Data  2.006 ( 0.978)   Loss 3.6648e+00 (4.7513e+00)   Acc@1  24.00 ( 1
2.57)   Acc@5  46.00 ( 28.84)
Epoch: [0][2050/2563]   Time  2.760 ( 1.644)   Data  2.055 ( 0.978)   Loss 3.5461e+00 (4.7458e+00)   Acc@1  25.60 ( 1
2.63)   Acc@5  51.00 ( 28.94)
Epoch: [0][2060/2563]   Time  2.679 ( 1.645)   Data  1.946 ( 0.978)   Loss 3.6648e+00 (4.7402e+00)   Acc@1  23.80 ( 1
2.68)   Acc@5  49.60 ( 29.03)
Epoch: [0][2070/2563]   Time  2.558 ( 1.645)   Data  1.843 ( 0.978)   Loss 3.3907e+00 (4.7346e+00)   Acc@1  26.20 ( 1
2.74)   Acc@5  55.00 ( 29.14)
Epoch: [0][2080/2563]   Time  2.465 ( 1.645)   Data  1.805 ( 0.978)   Loss 3.6459e+00 (4.7289e+00)   Acc@1  27.00 ( 1
2.81)   Acc@5  48.20 ( 29.24)
Epoch: [0][2090/2563]   Time  2.369 ( 1.645)   Data  1.704 ( 0.978)   Loss 3.4615e+00 (4.7232e+00)   Acc@1  29.80 ( 1
2.87)   Acc@5  55.60 ( 29.34)
Epoch: [0][2100/2563]   Time  2.630 ( 1.645)   Data  1.917 ( 0.978)   Loss 3.7516e+00 (4.7178e+00)   Acc@1  22.60 ( 1
2.93)   Acc@5  45.20 ( 29.44)
Epoch: [0][2110/2563]   Time  2.563 ( 1.645)   Data  1.860 ( 0.977)   Loss 3.3891e+00 (4.7123e+00)   Acc@1  29.00 ( 1
2.99)   Acc@5  52.60 ( 29.54)
Epoch: [0][2120/2563]   Time  2.640 ( 1.645)   Data  1.906 ( 0.977)   Loss 3.7213e+00 (4.7070e+00)   Acc@1  21.80 ( 1
3.05)   Acc@5  48.20 ( 29.64)
Epoch: [0][2130/2563]   Time  2.357 ( 1.644)   Data  1.695 ( 0.977)   Loss 3.6740e+00 (4.7016e+00)   Acc@1  23.80 ( 1
3.11)   Acc@5  46.40 ( 29.73)
Epoch: [0][2140/2563]   Time  2.737 ( 1.644)   Data  2.031 ( 0.977)   Loss 3.4626e+00 (4.6963e+00)   Acc@1  28.40 ( 1
3.17)   Acc@5  52.20 ( 29.83)
Epoch: [0][2150/2563]   Time  2.547 ( 1.644)   Data  1.834 ( 0.976)   Loss 3.4911e+00 (4.6913e+00)   Acc@1  27.60 ( 1
3.23)   Acc@5  50.80 ( 29.92)
Epoch: [0][2160/2563]   Time  2.630 ( 1.644)   Data  1.899 ( 0.976)   Loss 3.7201e+00 (4.6860e+00)   Acc@1  24.00 ( 1
3.29)   Acc@5  48.00 ( 30.01)
Epoch: [0][2170/2563]   Time  3.237 ( 1.644)   Data  2.526 ( 0.976)   Loss 3.4880e+00 (4.6808e+00)   Acc@1  28.60 ( 1
3.35)   Acc@5  52.60 ( 30.11)
Epoch: [0][2180/2563]   Time  2.645 ( 1.644)   Data  1.940 ( 0.976)   Loss 3.5610e+00 (4.6755e+00)   Acc@1  26.40 ( 1
3.41)   Acc@5  50.00 ( 30.21)
Epoch: [0][2190/2563]   Time  3.171 ( 1.644)   Data  2.457 ( 0.976)   Loss 3.4522e+00 (4.6702e+00)   Acc@1  29.40 ( 1
3.47)   Acc@5  53.40 ( 30.31)
Epoch: [0][2200/2563]   Time  2.461 ( 1.644)   Data  1.791 ( 0.976)   Loss 3.5716e+00 (4.6652e+00)   Acc@1  27.40 ( 1
3.53)   Acc@5  50.40 ( 30.40)
Epoch: [0][2210/2563]   Time  2.585 ( 1.644)   Data  1.866 ( 0.976)   Loss 3.5965e+00 (4.6599e+00)   Acc@1  27.20 ( 1
```

```
3.58)   Acc@5  50.20 ( 30.50)
Epoch: [0][2220/2563]   Time  2.498 ( 1.644)    Data  1.783 ( 0.976)    Loss 3.5203e+00 (4.6545e+00)    Acc@1  28.00 ( 1
3.64)   Acc@5  50.80 ( 30.60)
Epoch: [0][2230/2563]   Time  2.471 ( 1.644)    Data  1.793 ( 0.975)    Loss 3.7108e+00 (4.6495e+00)    Acc@1  22.60 ( 1
3.70)   Acc@5  47.00 ( 30.69)
Epoch: [0][2240/2563]   Time  2.677 ( 1.643)    Data  1.973 ( 0.975)    Loss 3.4724e+00 (4.6443e+00)    Acc@1  28.20 ( 1
3.76)   Acc@5  52.80 ( 30.79)
Epoch: [0][2250/2563]   Time  2.575 ( 1.644)    Data  1.882 ( 0.975)    Loss 3.4736e+00 (4.6392e+00)    Acc@1  25.00 ( 1
3.82)   Acc@5  52.80 ( 30.88)
Epoch: [0][2260/2563]   Time  2.793 ( 1.644)    Data  2.081 ( 0.975)    Loss 3.3521e+00 (4.6340e+00)    Acc@1  28.40 ( 1
3.88)   Acc@5  51.20 ( 30.97)
Epoch: [0][2270/2563]   Time  2.364 ( 1.644)    Data  1.700 ( 0.975)    Loss 3.4548e+00 (4.6289e+00)    Acc@1  24.60 ( 1
3.93)   Acc@5  51.60 ( 31.07)
Epoch: [0][2280/2563]   Time  2.587 ( 1.643)    Data  1.867 ( 0.975)    Loss 3.5622e+00 (4.6239e+00)    Acc@1  26.00 ( 1
3.98)   Acc@5  49.40 ( 31.15)
Epoch: [0][2290/2563]   Time  2.637 ( 1.644)    Data  1.929 ( 0.975)    Loss 3.4644e+00 (4.6190e+00)    Acc@1  26.60 ( 1
4.04)   Acc@5  51.00 ( 31.25)
Epoch: [0][2300/2563]   Time  2.450 ( 1.643)    Data  1.742 ( 0.975)    Loss 3.3746e+00 (4.6138e+00)    Acc@1  30.60 ( 1
4.10)   Acc@5  54.00 ( 31.34)
Epoch: [0][2310/2563]   Time  2.408 ( 1.643)    Data  1.746 ( 0.975)    Loss 3.3218e+00 (4.6084e+00)    Acc@1  30.80 ( 1
4.17)   Acc@5  54.80 ( 31.44)
Epoch: [0][2320/2563]   Time  2.468 ( 1.643)    Data  1.752 ( 0.974)    Loss 3.5190e+00 (4.6033e+00)    Acc@1  25.80 ( 1
4.22)   Acc@5  51.40 ( 31.53)
Epoch: [0][2330/2563]   Time  3.423 ( 1.643)    Data  2.708 ( 0.974)    Loss 3.3815e+00 (4.5984e+00)    Acc@1  27.00 ( 1
4.27)   Acc@5  53.40 ( 31.62)
Epoch: [0][2340/2563]   Time  2.743 ( 1.643)    Data  2.038 ( 0.974)    Loss 3.3728e+00 (4.5934e+00)    Acc@1  28.40 ( 1
4.33)   Acc@5  55.80 ( 31.71)
Epoch: [0][2350/2563]   Time  2.689 ( 1.643)    Data  1.981 ( 0.974)    Loss 3.4277e+00 (4.5887e+00)    Acc@1  25.60 ( 1
4.39)   Acc@5  51.40 ( 31.80)
Epoch: [0][2360/2563]   Time  2.548 ( 1.643)    Data  1.841 ( 0.974)    Loss 3.3551e+00 (4.5837e+00)    Acc@1  28.00 ( 1
4.44)   Acc@5  55.40 ( 31.89)
Epoch: [0][2370/2563]   Time  2.773 ( 1.643)    Data  2.068 ( 0.974)    Loss 3.5415e+00 (4.5789e+00)    Acc@1  28.80 ( 1
4.50)   Acc@5  51.40 ( 31.98)
Epoch: [0][2380/2563]   Time  2.482 ( 1.643)    Data  1.806 ( 0.974)    Loss 3.4782e+00 (4.5739e+00)    Acc@1  28.40 ( 1
4.56)   Acc@5  53.20 ( 32.07)
Epoch: [0][2390/2563]   Time  2.339 ( 1.643)    Data  1.676 ( 0.974)    Loss 3.4960e+00 (4.5690e+00)    Acc@1  26.60 ( 1
4.61)   Acc@5  52.00 ( 32.16)
Epoch: [0][2400/2563]   Time  2.688 ( 1.643)    Data  1.952 ( 0.974)    Loss 3.5244e+00 (4.5639e+00)    Acc@1  23.20 ( 1
4.67)   Acc@5  51.80 ( 32.25)
Epoch: [0][2410/2563]   Time  2.658 ( 1.643)    Data  1.927 ( 0.974)    Loss 3.2986e+00 (4.5590e+00)    Acc@1  29.00 ( 1
4.72)   Acc@5  54.60 ( 32.34)
Epoch: [0][2420/2563]   Time  2.645 ( 1.643)    Data  1.916 ( 0.974)    Loss 3.3658e+00 (4.5540e+00)    Acc@1  28.20 ( 1
4.78)   Acc@5  54.80 ( 32.43)
Epoch: [0][2430/2563]   Time  2.593 ( 1.643)    Data  1.874 ( 0.974)    Loss 3.3194e+00 (4.5493e+00)    Acc@1  30.00 ( 1
4.83)   Acc@5  54.80 ( 32.51)
Epoch: [0][2440/2563]   Time  2.836 ( 1.643)    Data  2.107 ( 0.974)    Loss 3.3547e+00 (4.5446e+00)    Acc@1  30.40 ( 1
4.89)   Acc@5  57.00 ( 32.60)
Epoch: [0][2450/2563]   Time  2.456 ( 1.644)    Data  1.785 ( 0.974)    Loss 3.2607e+00 (4.5401e+00)    Acc@1  29.80 ( 1
4.94)   Acc@5  57.00 ( 32.68)
Epoch: [0][2460/2563]   Time  2.683 ( 1.643)    Data  1.946 ( 0.974)    Loss 3.4323e+00 (4.5352e+00)    Acc@1  26.60 ( 1
5.00)   Acc@5  53.00 ( 32.77)
```

```
Epoch: [0][2470/2563]   Time  2.316 ( 1.643)    Data  1.648 ( 0.973)    Loss 3.3597e+00 (4.5307e+00)    Acc@1  25.80 ( 1
5.05)   Acc@5  55.80 ( 32.86)
Epoch: [0][2480/2563]   Time  2.860 ( 1.643)    Data  2.139 ( 0.973)    Loss 3.4392e+00 (4.5261e+00)    Acc@1  26.60 ( 1
5.10)   Acc@5  50.20 ( 32.94)
Epoch: [0][2490/2563]   Time  2.551 ( 1.643)    Data  1.856 ( 0.973)    Loss 3.6889e+00 (4.5216e+00)    Acc@1  27.20 ( 1
5.16)   Acc@5  48.40 ( 33.02)
Epoch: [0][2500/2563]   Time  2.645 ( 1.643)    Data  1.918 ( 0.973)    Loss 3.1878e+00 (4.5168e+00)    Acc@1  31.80 ( 1
5.22)   Acc@5  56.60 ( 33.11)
Epoch: [0][2510/2563]   Time  2.464 ( 1.643)    Data  1.754 ( 0.973)    Loss 3.2117e+00 (4.5121e+00)    Acc@1  31.00 ( 1
5.27)   Acc@5  54.60 ( 33.20)
Epoch: [0][2520/2563]   Time  3.013 ( 1.643)    Data  2.288 ( 0.973)    Loss 3.1778e+00 (4.5073e+00)    Acc@1  29.20 ( 1
5.33)   Acc@5  56.60 ( 33.28)
Epoch: [0][2530/2563]   Time  2.881 ( 1.643)    Data  2.165 ( 0.973)    Loss 3.4337e+00 (4.5025e+00)    Acc@1  27.60 ( 1
5.39)   Acc@5  52.80 ( 33.38)
Epoch: [0][2540/2563]   Time  2.434 ( 1.643)    Data  1.773 ( 0.973)    Loss 3.1844e+00 (4.4976e+00)    Acc@1  33.80 ( 1
5.44)   Acc@5  55.40 ( 33.46)
Epoch: [0][2550/2563]   Time  2.582 ( 1.643)    Data  1.880 ( 0.973)    Loss 3.2859e+00 (4.4931e+00)    Acc@1  30.40 ( 1
5.50)   Acc@5  54.60 ( 33.54)
Epoch: [0][2560/2563]   Time  2.478 ( 1.643)    Data  1.799 ( 0.973)    Loss 3.2529e+00 (4.4882e+00)    Acc@1  32.20 ( 1
5.56)   Acc@5  56.00 ( 33.63)
Test: [  0/100] Time  6.483 ( 6.483)    Loss 3.9590e+00 (3.9590e+00)    Acc@1  29.00 ( 29.00)   Acc@5  48.40 ( 48.40)
Test: [ 10/100] Time  2.963 ( 1.976)    Loss 3.9870e+00 (4.0170e+00)    Acc@1  21.00 ( 23.16)   Acc@5  44.40 ( 44.69)
Test: [ 20/100] Time  2.865 ( 1.842)    Loss 4.3763e+00 (3.9825e+00)    Acc@1  10.00 ( 21.34)   Acc@5  33.80 ( 44.14)
Test: [ 30/100] Time  2.885 ( 1.790)    Loss 3.3320e+00 (3.8490e+00)    Acc@1  24.80 ( 21.31)   Acc@5  54.60 ( 45.63)
Test: [ 40/100] Time  2.547 ( 1.772)    Loss 3.2501e+00 (3.8563e+00)    Acc@1  33.00 ( 21.81)   Acc@5  59.20 ( 45.50)
Test: [ 50/100] Time  2.774 ( 1.756)    Loss 5.0104e+00 (3.9624e+00)    Acc@1  10.80 ( 20.70)   Acc@5  25.20 ( 43.51)
Test: [ 60/100] Time  2.779 ( 1.748)    Loss 4.0820e+00 (3.9786e+00)    Acc@1  19.40 ( 20.48)   Acc@5  40.80 ( 43.01)
Test: [ 70/100] Time  2.950 ( 1.734)    Loss 4.8737e+00 (4.0098e+00)    Acc@1   9.40 ( 20.12)   Acc@5  27.00 ( 42.45)
Test: [ 80/100] Time  2.940 ( 1.727)    Loss 4.0539e+00 (4.0500e+00)    Acc@1  23.20 ( 20.01)   Acc@5  41.80 ( 41.83)
Test: [ 90/100] Time  2.892 ( 1.720)    Loss 4.2836e+00 (4.0748e+00)    Acc@1  20.20 ( 19.67)   Acc@5  40.40 ( 41.40)
 * Acc@1 19.706 Acc@5 41.196
lr: 0.05
Epoch: [1][   0/2563]   Time  4.411 ( 4.411)    Data  3.684 ( 3.684)    Loss 3.2238e+00 (3.2238e+00)    Acc@1  29.80 ( 2
9.80)   Acc@5  56.80 ( 56.80)
Epoch: [1][  10/2563]   Time  2.504 ( 1.825)    Data  1.804 ( 1.149)    Loss 3.1083e+00 (3.1708e+00)    Acc@1  35.60 ( 3
2.91)   Acc@5  58.40 ( 57.67)
Epoch: [1][  20/2563]   Time  2.560 ( 1.721)    Data  1.850 ( 1.046)    Loss 3.0561e+00 (3.1642e+00)    Acc@1  33.20 ( 3
2.79)   Acc@5  58.40 ( 57.75)
Epoch: [1][  30/2563]   Time  2.363 ( 1.699)    Data  1.705 ( 1.022)    Loss 3.0806e+00 (3.1456e+00)    Acc@1  37.00 ( 3
2.87)   Acc@5  58.80 ( 57.98)
Epoch: [1][  40/2563]   Time  2.563 ( 1.676)    Data  1.877 ( 0.998)    Loss 3.0295e+00 (3.1304e+00)    Acc@1  34.40 ( 3
3.28)   Acc@5  60.00 ( 58.42)
Epoch: [1][  50/2563]   Time  2.489 ( 1.657)    Data  1.776 ( 0.980)    Loss 3.0413e+00 (3.1287e+00)    Acc@1  35.20 ( 3
3.35)   Acc@5  59.80 ( 58.49)
Epoch: [1][  60/2563]   Time  2.454 ( 1.639)    Data  1.737 ( 0.963)    Loss 3.2110e+00 (3.1116e+00)    Acc@1  30.20 ( 3
3.50)   Acc@5  57.00 ( 58.84)
Epoch: [1][  70/2563]   Time  2.299 ( 1.624)    Data  1.637 ( 0.948)    Loss 2.7541e+00 (3.1015e+00)    Acc@1  37.80 ( 3
3.48)   Acc@5  64.80 ( 59.01)
Epoch: [1][  80/2563]   Time  2.354 ( 1.616)    Data  1.692 ( 0.941)    Loss 2.8702e+00 (3.0957e+00)    Acc@1  34.60 ( 3
3.47)   Acc@5  62.40 ( 59.12)
Epoch: [1][  90/2563]   Time  2.109 ( 1.608)    Data  1.443 ( 0.934)    Loss 3.0488e+00 (3.0876e+00)    Acc@1  32.80 ( 3
```

```
3.47)   Acc@5  59.40 ( 59.36)
Epoch: [1][ 100/2563]   Time  2.249 ( 1.606)    Data  1.582 ( 0.933)    Loss 2.9540e+00 (3.0817e+00)    Acc@1  37.80 ( 3
3.54)   Acc@5  63.40 ( 59.45)
Epoch: [1][ 110/2563]   Time  2.243 ( 1.604)    Data  1.581 ( 0.932)    Loss 3.0034e+00 (3.0787e+00)    Acc@1  37.00 ( 3
3.58)   Acc@5  62.80 ( 59.58)
Epoch: [1][ 120/2563]   Time  2.462 ( 1.605)    Data  1.748 ( 0.933)    Loss 3.0216e+00 (3.0733e+00)    Acc@1  36.80 ( 3
3.71)   Acc@5  61.00 ( 59.68)
Epoch: [1][ 130/2563]   Time  2.431 ( 1.604)    Data  1.766 ( 0.932)    Loss 2.9985e+00 (3.0691e+00)    Acc@1  34.60 ( 3
3.84)   Acc@5  61.60 ( 59.82)
Epoch: [1][ 140/2563]   Time  2.355 ( 1.603)    Data  1.688 ( 0.930)    Loss 2.8790e+00 (3.0637e+00)    Acc@1  33.60 ( 3
3.89)   Acc@5  65.40 ( 59.91)
Epoch: [1][ 150/2563]   Time  2.254 ( 1.599)    Data  1.588 ( 0.927)    Loss 2.9907e+00 (3.0602e+00)    Acc@1  35.00 ( 3
3.89)   Acc@5  60.20 ( 60.00)
Epoch: [1][ 160/2563]   Time  2.522 ( 1.602)    Data  1.806 ( 0.928)    Loss 3.0451e+00 (3.0572e+00)    Acc@1  32.20 ( 3
3.93)   Acc@5  61.60 ( 60.04)
Epoch: [1][ 170/2563]   Time  2.609 ( 1.601)    Data  1.872 ( 0.927)    Loss 3.0246e+00 (3.0547e+00)    Acc@1  32.00 ( 3
3.91)   Acc@5  61.60 ( 60.09)
Epoch: [1][ 180/2563]   Time  2.358 ( 1.599)    Data  1.695 ( 0.925)    Loss 3.0053e+00 (3.0514e+00)    Acc@1  35.60 ( 3
3.97)   Acc@5  59.80 ( 60.15)
Epoch: [1][ 190/2563]   Time  2.367 ( 1.599)    Data  1.707 ( 0.926)    Loss 3.0161e+00 (3.0503e+00)    Acc@1  37.00 ( 3
3.97)   Acc@5  60.60 ( 60.16)
Epoch: [1][ 200/2563]   Time  2.388 ( 1.602)    Data  1.722 ( 0.927)    Loss 3.0346e+00 (3.0477e+00)    Acc@1  35.20 ( 3
4.04)   Acc@5  59.80 ( 60.18)
Epoch: [1][ 210/2563]   Time  2.569 ( 1.603)    Data  1.855 ( 0.929)    Loss 3.0804e+00 (3.0430e+00)    Acc@1  32.60 ( 3
4.09)   Acc@5  59.60 ( 60.26)
Epoch: [1][ 220/2563]   Time  2.481 ( 1.602)    Data  1.817 ( 0.927)    Loss 2.8891e+00 (3.0401e+00)    Acc@1  37.60 ( 3
4.15)   Acc@5  59.40 ( 60.31)
Epoch: [1][ 230/2563]   Time  2.414 ( 1.601)    Data  1.748 ( 0.926)    Loss 2.8712e+00 (3.0366e+00)    Acc@1  36.20 ( 3
4.22)   Acc@5  61.40 ( 60.34)
Epoch: [1][ 240/2563]   Time  2.123 ( 1.599)    Data  1.458 ( 0.925)    Loss 2.9694e+00 (3.0339e+00)    Acc@1  35.20 ( 3
4.21)   Acc@5  61.20 ( 60.39)
Epoch: [1][ 250/2563]   Time  1.832 ( 1.598)    Data  1.165 ( 0.925)    Loss 3.0296e+00 (3.0302e+00)    Acc@1  35.40 ( 3
4.26)   Acc@5  59.80 ( 60.45)
Epoch: [1][ 260/2563]   Time  1.495 ( 1.599)    Data  0.828 ( 0.926)    Loss 2.9637e+00 (3.0289e+00)    Acc@1  35.40 ( 3
4.29)   Acc@5  61.40 ( 60.46)
Epoch: [1][ 270/2563]   Time  1.437 ( 1.599)    Data  0.771 ( 0.926)    Loss 3.0936e+00 (3.0273e+00)    Acc@1  33.40 ( 3
4.31)   Acc@5  62.00 ( 60.48)
Epoch: [1][ 280/2563]   Time  1.237 ( 1.599)    Data  0.569 ( 0.927)    Loss 2.8856e+00 (3.0249e+00)    Acc@1  35.60 ( 3
4.35)   Acc@5  63.20 ( 60.51)
Epoch: [1][ 290/2563]   Time  1.630 ( 1.601)    Data  0.964 ( 0.928)    Loss 2.8864e+00 (3.0235e+00)    Acc@1  37.60 ( 3
4.39)   Acc@5  63.20 ( 60.56)
Epoch: [1][ 300/2563]   Time  2.266 ( 1.603)    Data  1.604 ( 0.930)    Loss 2.9050e+00 (3.0226e+00)    Acc@1  36.80 ( 3
4.39)   Acc@5  62.00 ( 60.59)
Epoch: [1][ 310/2563]   Time  1.859 ( 1.601)    Data  1.196 ( 0.929)    Loss 2.8989e+00 (3.0204e+00)    Acc@1  36.20 ( 3
4.43)   Acc@5  62.60 ( 60.62)
Epoch: [1][ 320/2563]   Time  1.390 ( 1.601)    Data  0.723 ( 0.929)    Loss 2.8465e+00 (3.0186e+00)    Acc@1  38.60 ( 3
4.47)   Acc@5  63.60 ( 60.62)
Epoch: [1][ 330/2563]   Time  1.630 ( 1.601)    Data  0.964 ( 0.929)    Loss 2.9477e+00 (3.0169e+00)    Acc@1  36.40 ( 3
4.46)   Acc@5  61.80 ( 60.64)
Epoch: [1][ 340/2563]   Time  1.509 ( 1.601)    Data  0.840 ( 0.929)    Loss 3.0185e+00 (3.0171e+00)    Acc@1  33.60 ( 3
4.45)   Acc@5  60.40 ( 60.66)
```

```
Epoch: [1][ 350/2563]   Time  1.295 ( 1.600)   Data  0.629 ( 0.929)   Loss 2.9675e+00 (3.0145e+00)   Acc@1  35.00 ( 34.50)   Acc@5  64.40 ( 60.71)
Epoch: [1][ 360/2563]   Time  1.102 ( 1.599)   Data  0.436 ( 0.928)   Loss 3.0827e+00 (3.0131e+00)   Acc@1  34.20 ( 34.52)   Acc@5  60.40 ( 60.74)
Epoch: [1][ 370/2563]   Time  1.050 ( 1.599)   Data  0.385 ( 0.927)   Loss 2.9970e+00 (3.0101e+00)   Acc@1  36.20 ( 34.56)   Acc@5  61.20 ( 60.78)
Epoch: [1][ 380/2563]   Time  0.765 ( 1.598)   Data  0.098 ( 0.927)   Loss 2.9840e+00 (3.0090e+00)   Acc@1  34.00 ( 34.57)   Acc@5  61.20 ( 60.81)
Epoch: [1][ 390/2563]   Time  0.759 ( 1.598)   Data  0.098 ( 0.927)   Loss 2.6969e+00 (3.0053e+00)   Acc@1  40.60 ( 34.64)   Acc@5  64.80 ( 60.87)
Epoch: [1][ 400/2563]   Time  0.765 ( 1.599)   Data  0.097 ( 0.927)   Loss 3.0451e+00 (3.0041e+00)   Acc@1  35.00 ( 34.65)   Acc@5  59.60 ( 60.91)
Epoch: [1][ 410/2563]   Time  0.769 ( 1.599)   Data  0.104 ( 0.927)   Loss 2.9582e+00 (3.0026e+00)   Acc@1  37.00 ( 34.67)   Acc@5  60.40 ( 60.93)
Epoch: [1][ 420/2563]   Time  0.966 ( 1.598)   Data  0.301 ( 0.927)   Loss 2.9010e+00 (2.9997e+00)   Acc@1  39.60 ( 34.70)   Acc@5  64.40 ( 60.99)
Epoch: [1][ 430/2563]   Time  0.833 ( 1.599)   Data  0.166 ( 0.927)   Loss 3.0048e+00 (2.9982e+00)   Acc@1  34.60 ( 34.72)   Acc@5  60.00 ( 61.02)
Epoch: [1][ 440/2563]   Time  1.054 ( 1.599)   Data  0.388 ( 0.928)   Loss 2.8382e+00 (2.9959e+00)   Acc@1  35.00 ( 34.75)   Acc@5  63.40 ( 61.06)
Epoch: [1][ 450/2563]   Time  1.004 ( 1.599)   Data  0.339 ( 0.928)   Loss 2.7777e+00 (2.9951e+00)   Acc@1  41.20 ( 34.78)   Acc@5  65.60 ( 61.07)
Epoch: [1][ 460/2563]   Time  0.919 ( 1.600)   Data  0.254 ( 0.928)   Loss 2.9852e+00 (2.9942e+00)   Acc@1  37.20 ( 34.80)   Acc@5  61.00 ( 61.08)
Epoch: [1][ 470/2563]   Time  0.987 ( 1.600)   Data  0.320 ( 0.929)   Loss 2.9588e+00 (2.9929e+00)   Acc@1  33.60 ( 34.81)   Acc@5  62.00 ( 61.11)
Epoch: [1][ 480/2563]   Time  0.880 ( 1.600)   Data  0.213 ( 0.929)   Loss 2.9628e+00 (2.9916e+00)   Acc@1  33.20 ( 34.84)   Acc@5  61.20 ( 61.13)
Epoch: [1][ 490/2563]   Time  0.931 ( 1.600)   Data  0.262 ( 0.929)   Loss 2.8527e+00 (2.9892e+00)   Acc@1  38.20 ( 34.90)   Acc@5  63.20 ( 61.17)
Epoch: [1][ 500/2563]   Time  1.034 ( 1.600)   Data  0.372 ( 0.929)   Loss 3.0854e+00 (2.9890e+00)   Acc@1  35.40 ( 34.91)   Acc@5  59.00 ( 61.17)
Epoch: [1][ 510/2563]   Time  1.594 ( 1.601)   Data  0.930 ( 0.930)   Loss 2.8338e+00 (2.9872e+00)   Acc@1  39.00 ( 34.94)   Acc@5  61.60 ( 61.20)
Epoch: [1][ 520/2563]   Time  0.944 ( 1.601)   Data  0.278 ( 0.930)   Loss 2.9545e+00 (2.9863e+00)   Acc@1  35.20 ( 34.95)   Acc@5  60.60 ( 61.21)
Epoch: [1][ 530/2563]   Time  0.882 ( 1.601)   Data  0.218 ( 0.930)   Loss 3.0170e+00 (2.9849e+00)   Acc@1  36.60 ( 34.97)   Acc@5  60.00 ( 61.22)
Epoch: [1][ 540/2563]   Time  0.791 ( 1.600)   Data  0.127 ( 0.929)   Loss 2.8562e+00 (2.9828e+00)   Acc@1  39.60 ( 35.01)   Acc@5  66.00 ( 61.26)
Epoch: [1][ 550/2563]   Time  0.768 ( 1.600)   Data  0.101 ( 0.929)   Loss 2.7144e+00 (2.9809e+00)   Acc@1  38.00 ( 35.04)   Acc@5  66.20 ( 61.30)
Epoch: [1][ 560/2563]   Time  0.766 ( 1.600)   Data  0.099 ( 0.929)   Loss 3.0498e+00 (2.9800e+00)   Acc@1  32.40 ( 35.05)   Acc@5  59.20 ( 61.30)
Epoch: [1][ 570/2563]   Time  0.763 ( 1.599)   Data  0.096 ( 0.928)   Loss 2.9777e+00 (2.9785e+00)   Acc@1  38.00 ( 35.08)   Acc@5  60.80 ( 61.31)
Epoch: [1][ 580/2563]   Time  0.833 ( 1.598)   Data  0.168 ( 0.927)   Loss 3.0792e+00 (2.9776e+00)   Acc@1  35.20 ( 35.09)   Acc@5  60.40 ( 61.33)
Epoch: [1][ 590/2563]   Time  1.349 ( 1.599)   Data  0.680 ( 0.928)   Loss 2.9009e+00 (2.9758e+00)   Acc@1  37.60 ( 35.12)   Acc@5  60.80 ( 61.36)
Epoch: [1][ 600/2563]   Time  1.940 ( 1.599)   Data  1.276 ( 0.929)   Loss 3.0643e+00 (2.9745e+00)   Acc@1  36.00 ( 3
```

```
5.14)   Acc@5  58.40 ( 61.38)
Epoch: [1][ 610/2563]   Time  1.963 ( 1.600)    Data  1.299 ( 0.929)    Loss 2.9308e+00 (2.9729e+00)    Acc@1  36.20 ( 3
5.16)   Acc@5  59.20 ( 61.39)
Epoch: [1][ 620/2563]   Time  2.216 ( 1.600)    Data  1.551 ( 0.929)    Loss 2.8514e+00 (2.9716e+00)    Acc@1  37.60 ( 3
5.18)   Acc@5  65.40 ( 61.42)
Epoch: [1][ 630/2563]   Time  2.163 ( 1.599)    Data  1.500 ( 0.929)    Loss 2.9283e+00 (2.9703e+00)    Acc@1  34.80 ( 3
5.20)   Acc@5  63.20 ( 61.45)
Epoch: [1][ 640/2563]   Time  2.316 ( 1.599)    Data  1.650 ( 0.929)    Loss 2.8341e+00 (2.9699e+00)    Acc@1  37.40 ( 3
5.21)   Acc@5  63.80 ( 61.45)
Epoch: [1][ 650/2563]   Time  2.337 ( 1.599)    Data  1.673 ( 0.928)    Loss 2.8595e+00 (2.9682e+00)    Acc@1  39.00 ( 3
5.24)   Acc@5  63.60 ( 61.48)
Epoch: [1][ 660/2563]   Time  1.976 ( 1.599)    Data  1.312 ( 0.928)    Loss 2.8128e+00 (2.9672e+00)    Acc@1  36.80 ( 3
5.27)   Acc@5  63.40 ( 61.50)
Epoch: [1][ 670/2563]   Time  1.551 ( 1.599)    Data  0.883 ( 0.928)    Loss 2.8348e+00 (2.9653e+00)    Acc@1  39.00 ( 3
5.30)   Acc@5  62.60 ( 61.53)
Epoch: [1][ 680/2563]   Time  1.663 ( 1.599)    Data  0.994 ( 0.928)    Loss 2.6473e+00 (2.9629e+00)    Acc@1  39.80 ( 3
5.33)   Acc@5  67.60 ( 61.57)
Epoch: [1][ 690/2563]   Time  1.364 ( 1.598)    Data  0.700 ( 0.928)    Loss 2.8636e+00 (2.9615e+00)    Acc@1  36.60 ( 3
5.36)   Acc@5  64.40 ( 61.60)
Epoch: [1][ 700/2563]   Time  1.026 ( 1.598)    Data  0.362 ( 0.927)    Loss 2.9146e+00 (2.9597e+00)    Acc@1  36.80 ( 3
5.40)   Acc@5  61.60 ( 61.63)
Epoch: [1][ 710/2563]   Time  0.761 ( 1.598)    Data  0.095 ( 0.928)    Loss 2.8682e+00 (2.9584e+00)    Acc@1  38.20 ( 3
5.41)   Acc@5  64.00 ( 61.67)
Epoch: [1][ 720/2563]   Time  0.913 ( 1.598)    Data  0.249 ( 0.927)    Loss 2.8664e+00 (2.9572e+00)    Acc@1  36.40 ( 3
5.44)   Acc@5  64.60 ( 61.68)
Epoch: [1][ 730/2563]   Time  0.763 ( 1.598)    Data  0.098 ( 0.927)    Loss 2.8285e+00 (2.9553e+00)    Acc@1  38.00 ( 3
5.47)   Acc@5  63.40 ( 61.71)
Epoch: [1][ 740/2563]   Time  0.869 ( 1.598)    Data  0.204 ( 0.927)    Loss 2.8947e+00 (2.9544e+00)    Acc@1  37.20 ( 3
5.49)   Acc@5  61.40 ( 61.72)
Epoch: [1][ 750/2563]   Time  0.764 ( 1.598)    Data  0.097 ( 0.927)    Loss 2.7809e+00 (2.9529e+00)    Acc@1  40.00 ( 3
5.52)   Acc@5  68.40 ( 61.76)
Epoch: [1][ 760/2563]   Time  0.777 ( 1.598)    Data  0.112 ( 0.927)    Loss 2.7398e+00 (2.9521e+00)    Acc@1  42.80 ( 3
5.53)   Acc@5  65.60 ( 61.77)
Epoch: [1][ 770/2563]   Time  1.097 ( 1.598)    Data  0.436 ( 0.928)    Loss 2.8425e+00 (2.9511e+00)    Acc@1  36.00 ( 3
5.55)   Acc@5  64.80 ( 61.79)
Epoch: [1][ 780/2563]   Time  0.759 ( 1.598)    Data  0.096 ( 0.928)    Loss 2.8581e+00 (2.9493e+00)    Acc@1  38.60 ( 3
5.58)   Acc@5  63.60 ( 61.81)
Epoch: [1][ 790/2563]   Time  0.768 ( 1.598)    Data  0.102 ( 0.928)    Loss 2.7936e+00 (2.9482e+00)    Acc@1  38.20 ( 3
5.58)   Acc@5  63.00 ( 61.83)
Epoch: [1][ 800/2563]   Time  0.666 ( 1.599)    Data  0.002 ( 0.928)    Loss 2.8622e+00 (2.9462e+00)    Acc@1  33.80 ( 3
5.62)   Acc@5  64.80 ( 61.86)
Epoch: [1][ 810/2563]   Time  0.759 ( 1.599)    Data  0.097 ( 0.928)    Loss 2.9103e+00 (2.9454e+00)    Acc@1  36.80 ( 3
5.64)   Acc@5  61.60 ( 61.87)
Epoch: [1][ 820/2563]   Time  0.850 ( 1.599)    Data  0.185 ( 0.928)    Loss 2.9210e+00 (2.9442e+00)    Acc@1  35.80 ( 3
5.66)   Acc@5  60.20 ( 61.89)
Epoch: [1][ 830/2563]   Time  0.761 ( 1.599)    Data  0.099 ( 0.928)    Loss 2.9484e+00 (2.9439e+00)    Acc@1  37.40 ( 3
5.67)   Acc@5  62.20 ( 61.90)
Epoch: [1][ 840/2563]   Time  0.761 ( 1.599)    Data  0.097 ( 0.928)    Loss 2.7702e+00 (2.9427e+00)    Acc@1  39.20 ( 3
5.70)   Acc@5  65.60 ( 61.92)
Epoch: [1][ 850/2563]   Time  0.916 ( 1.599)    Data  0.253 ( 0.928)    Loss 2.9128e+00 (2.9414e+00)    Acc@1  35.20 ( 3
5.73)   Acc@5  63.00 ( 61.94)
```

```
Epoch: [1][ 860/2563]   Time  0.766 ( 1.599)   Data  0.099 ( 0.928)   Loss 2.9896e+00 (2.9405e+00)   Acc@1  33.40 ( 3
5.74)   Acc@5  60.20 ( 61.96)
Epoch: [1][ 870/2563]   Time  0.769 ( 1.599)   Data  0.103 ( 0.928)   Loss 2.8315e+00 (2.9396e+00)   Acc@1  38.20 ( 3
5.76)   Acc@5  63.80 ( 61.97)
Epoch: [1][ 880/2563]   Time  0.759 ( 1.599)   Data  0.097 ( 0.928)   Loss 3.0073e+00 (2.9385e+00)   Acc@1  36.20 ( 3
5.78)   Acc@5  58.40 ( 61.99)
Epoch: [1][ 890/2563]   Time  0.760 ( 1.599)   Data  0.097 ( 0.929)   Loss 2.7354e+00 (2.9364e+00)   Acc@1  41.80 ( 3
5.82)   Acc@5  65.00 ( 62.02)
Epoch: [1][ 900/2563]   Time  1.248 ( 1.600)   Data  0.581 ( 0.929)   Loss 2.5472e+00 (2.9349e+00)   Acc@1  43.20 ( 3
5.84)   Acc@5  67.80 ( 62.05)
Epoch: [1][ 910/2563]   Time  0.981 ( 1.600)   Data  0.315 ( 0.930)   Loss 2.9331e+00 (2.9337e+00)   Acc@1  37.00 ( 3
5.87)   Acc@5  63.60 ( 62.07)
Epoch: [1][ 920/2563]   Time  1.600 ( 1.601)   Data  0.935 ( 0.930)   Loss 2.8242e+00 (2.9322e+00)   Acc@1  37.00 ( 3
5.89)   Acc@5  63.80 ( 62.09)
Epoch: [1][ 930/2563]   Time  1.893 ( 1.601)   Data  1.225 ( 0.930)   Loss 2.8344e+00 (2.9310e+00)   Acc@1  36.40 ( 3
5.90)   Acc@5  63.40 ( 62.11)
Epoch: [1][ 940/2563]   Time  1.885 ( 1.601)   Data  1.217 ( 0.930)   Loss 2.8561e+00 (2.9297e+00)   Acc@1  36.20 ( 3
5.92)   Acc@5  64.40 ( 62.14)
Epoch: [1][ 950/2563]   Time  1.499 ( 1.601)   Data  0.834 ( 0.931)   Loss 2.9139e+00 (2.9285e+00)   Acc@1  36.00 ( 3
5.94)   Acc@5  61.20 ( 62.16)
Epoch: [1][ 960/2563]   Time  1.056 ( 1.601)   Data  0.390 ( 0.930)   Loss 2.8985e+00 (2.9272e+00)   Acc@1  36.60 ( 3
5.96)   Acc@5  64.00 ( 62.19)
Epoch: [1][ 970/2563]   Time  0.889 ( 1.601)   Data  0.225 ( 0.931)   Loss 2.8596e+00 (2.9258e+00)   Acc@1  36.20 ( 3
5.99)   Acc@5  62.80 ( 62.22)
Epoch: [1][ 980/2563]   Time  0.761 ( 1.601)   Data  0.097 ( 0.931)   Loss 2.7214e+00 (2.9241e+00)   Acc@1  39.60 ( 3
6.01)   Acc@5  65.60 ( 62.25)
Epoch: [1][ 990/2563]   Time  0.669 ( 1.602)   Data  0.002 ( 0.931)   Loss 2.8056e+00 (2.9232e+00)   Acc@1  39.80 ( 3
6.03)   Acc@5  65.80 ( 62.27)
Epoch: [1][1000/2563]   Time  0.670 ( 1.602)   Data  0.002 ( 0.931)   Loss 2.7865e+00 (2.9217e+00)   Acc@1  38.20 ( 3
6.05)   Acc@5  65.00 ( 62.30)
Epoch: [1][1010/2563]   Time  0.668 ( 1.602)   Data  0.002 ( 0.932)   Loss 3.0042e+00 (2.9206e+00)   Acc@1  35.80 ( 3
6.06)   Acc@5  58.40 ( 62.32)
Epoch: [1][1020/2563]   Time  0.675 ( 1.602)   Data  0.002 ( 0.931)   Loss 2.7889e+00 (2.9187e+00)   Acc@1  37.80 ( 3
6.10)   Acc@5  64.00 ( 62.35)
Epoch: [1][1030/2563]   Time  0.667 ( 1.602)   Data  0.002 ( 0.931)   Loss 2.8019e+00 (2.9176e+00)   Acc@1  42.40 ( 3
6.13)   Acc@5  65.00 ( 62.37)
Epoch: [1][1040/2563]   Time  0.666 ( 1.603)   Data  0.002 ( 0.931)   Loss 2.8362e+00 (2.9168e+00)   Acc@1  38.60 ( 3
6.14)   Acc@5  63.40 ( 62.38)
Epoch: [1][1050/2563]   Time  0.671 ( 1.603)   Data  0.002 ( 0.931)   Loss 2.7990e+00 (2.9160e+00)   Acc@1  35.40 ( 3
6.15)   Acc@5  64.20 ( 62.39)
Epoch: [1][1060/2563]   Time  0.670 ( 1.603)   Data  0.003 ( 0.932)   Loss 2.7892e+00 (2.9141e+00)   Acc@1  40.20 ( 3
6.19)   Acc@5  63.60 ( 62.43)
Epoch: [1][1070/2563]   Time  0.762 ( 1.603)   Data  0.096 ( 0.931)   Loss 2.9601e+00 (2.9131e+00)   Acc@1  36.80 ( 3
6.21)   Acc@5  63.80 ( 62.45)
Epoch: [1][1080/2563]   Time  0.669 ( 1.603)   Data  0.003 ( 0.932)   Loss 2.6844e+00 (2.9119e+00)   Acc@1  40.80 ( 3
6.25)   Acc@5  65.80 ( 62.47)
Epoch: [1][1090/2563]   Time  0.668 ( 1.603)   Data  0.003 ( 0.931)   Loss 2.7523e+00 (2.9104e+00)   Acc@1  41.40 ( 3
6.28)   Acc@5  65.00 ( 62.49)
Epoch: [1][1100/2563]   Time  0.668 ( 1.603)   Data  0.003 ( 0.931)   Loss 2.8728e+00 (2.9094e+00)   Acc@1  37.20 ( 3
6.30)   Acc@5  61.80 ( 62.51)
Epoch: [1][1110/2563]   Time  0.668 ( 1.603)   Data  0.003 ( 0.931)   Loss 2.8398e+00 (2.9085e+00)   Acc@1  38.00 ( 3
```

```
6.31)   Acc@5  64.60 ( 62.53)
Epoch: [1][1120/2563]   Time  0.666 ( 1.603)    Data  0.002 ( 0.931)    Loss 2.7927e+00 (2.9078e+00)    Acc@1  37.80 ( 3
6.34)   Acc@5  63.40 ( 62.55)
Epoch: [1][1130/2563]   Time  0.666 ( 1.603)    Data  0.003 ( 0.931)    Loss 2.6990e+00 (2.9065e+00)    Acc@1  39.80 ( 3
6.36)   Acc@5  69.00 ( 62.57)
Epoch: [1][1140/2563]   Time  0.669 ( 1.603)    Data  0.003 ( 0.931)    Loss 2.8229e+00 (2.9059e+00)    Acc@1  36.80 ( 3
6.36)   Acc@5  64.80 ( 62.58)
Epoch: [1][1150/2563]   Time  0.762 ( 1.604)    Data  0.097 ( 0.932)    Loss 2.9484e+00 (2.9050e+00)    Acc@1  33.80 ( 3
6.38)   Acc@5  61.80 ( 62.60)
Epoch: [1][1160/2563]   Time  1.020 ( 1.604)    Data  0.357 ( 0.932)    Loss 2.9742e+00 (2.9043e+00)    Acc@1  37.20 ( 3
6.39)   Acc@5  62.00 ( 62.62)
Epoch: [1][1170/2563]   Time  1.190 ( 1.604)    Data  0.525 ( 0.932)    Loss 2.7579e+00 (2.9035e+00)    Acc@1  36.40 ( 3
6.40)   Acc@5  62.40 ( 62.64)
Epoch: [1][1180/2563]   Time  1.120 ( 1.604)    Data  0.457 ( 0.932)    Loss 2.8741e+00 (2.9027e+00)    Acc@1  38.20 ( 3
6.42)   Acc@5  63.00 ( 62.65)
Epoch: [1][1190/2563]   Time  1.425 ( 1.604)    Data  0.758 ( 0.932)    Loss 2.8047e+00 (2.9015e+00)    Acc@1  38.40 ( 3
6.44)   Acc@5  64.40 ( 62.67)
Epoch: [1][1200/2563]   Time  1.648 ( 1.604)    Data  0.982 ( 0.933)    Loss 3.1625e+00 (2.9010e+00)    Acc@1  34.00 ( 3
6.46)   Acc@5  58.80 ( 62.68)
Epoch: [1][1210/2563]   Time  1.832 ( 1.605)    Data  1.168 ( 0.933)    Loss 2.6641e+00 (2.8999e+00)    Acc@1  40.00 ( 3
6.47)   Acc@5  67.00 ( 62.70)
Epoch: [1][1220/2563]   Time  1.727 ( 1.605)    Data  1.060 ( 0.933)    Loss 2.9425e+00 (2.8990e+00)    Acc@1  33.20 ( 3
6.48)   Acc@5  63.40 ( 62.72)
Epoch: [1][1230/2563]   Time  1.658 ( 1.605)    Data  0.994 ( 0.933)    Loss 2.6733e+00 (2.8982e+00)    Acc@1  41.80 ( 3
6.50)   Acc@5  66.80 ( 62.73)
Epoch: [1][1240/2563]   Time  1.371 ( 1.605)    Data  0.707 ( 0.933)    Loss 2.7537e+00 (2.8972e+00)    Acc@1  40.80 ( 3
6.51)   Acc@5  64.60 ( 62.75)
Epoch: [1][1250/2563]   Time  1.207 ( 1.604)    Data  0.542 ( 0.933)    Loss 2.9019e+00 (2.8965e+00)    Acc@1  35.00 ( 3
6.52)   Acc@5  61.20 ( 62.77)
Epoch: [1][1260/2563]   Time  1.920 ( 1.605)    Data  1.256 ( 0.933)    Loss 2.7664e+00 (2.8954e+00)    Acc@1  37.80 ( 3
6.54)   Acc@5  66.20 ( 62.79)
Epoch: [1][1270/2563]   Time  2.100 ( 1.605)    Data  1.435 ( 0.934)    Loss 2.7552e+00 (2.8945e+00)    Acc@1  38.20 ( 3
6.55)   Acc@5  65.00 ( 62.80)
Epoch: [1][1280/2563]   Time  2.089 ( 1.605)    Data  1.424 ( 0.934)    Loss 2.7762e+00 (2.8940e+00)    Acc@1  38.00 ( 3
6.57)   Acc@5  65.80 ( 62.81)
Epoch: [1][1290/2563]   Time  1.786 ( 1.605)    Data  1.122 ( 0.933)    Loss 2.7125e+00 (2.8925e+00)    Acc@1  39.60 ( 3
6.59)   Acc@5  67.60 ( 62.84)
Epoch: [1][1300/2563]   Time  2.227 ( 1.605)    Data  1.560 ( 0.934)    Loss 2.7904e+00 (2.8916e+00)    Acc@1  35.60 ( 3
6.61)   Acc@5  65.80 ( 62.85)
Epoch: [1][1310/2563]   Time  1.892 ( 1.605)    Data  1.227 ( 0.934)    Loss 2.7419e+00 (2.8906e+00)    Acc@1  41.80 ( 3
6.63)   Acc@5  65.00 ( 62.87)
Epoch: [1][1320/2563]   Time  1.388 ( 1.605)    Data  0.721 ( 0.933)    Loss 2.9030e+00 (2.8897e+00)    Acc@1  34.60 ( 3
6.64)   Acc@5  64.60 ( 62.89)
Epoch: [1][1330/2563]   Time  1.882 ( 1.605)    Data  1.218 ( 0.934)    Loss 2.8380e+00 (2.8891e+00)    Acc@1  38.20 ( 3
6.66)   Acc@5  63.80 ( 62.90)
Epoch: [1][1340/2563]   Time  2.335 ( 1.605)    Data  1.668 ( 0.934)    Loss 2.7550e+00 (2.8885e+00)    Acc@1  41.20 ( 3
6.67)   Acc@5  64.40 ( 62.91)
Epoch: [1][1350/2563]   Time  2.394 ( 1.605)    Data  1.729 ( 0.934)    Loss 2.7490e+00 (2.8876e+00)    Acc@1  41.60 ( 3
6.69)   Acc@5  63.80 ( 62.93)
Epoch: [1][1360/2563]   Time  2.364 ( 1.605)    Data  1.702 ( 0.934)    Loss 2.6449e+00 (2.8864e+00)    Acc@1  39.00 ( 3
6.71)   Acc@5  70.00 ( 62.95)
```

```
Epoch: [1][1370/2563]   Time  2.348 ( 1.605)   Data  1.680 ( 0.934)   Loss 2.6843e+00 (2.8855e+00)   Acc@1  38.40 ( 3
6.72)   Acc@5  64.40 ( 62.97)
Epoch: [1][1380/2563]   Time  2.550 ( 1.605)   Data  1.836 ( 0.934)   Loss 2.7490e+00 (2.8842e+00)   Acc@1  39.80 ( 3
6.74)   Acc@5  65.40 ( 62.99)
Epoch: [1][1390/2563]   Time  2.212 ( 1.605)   Data  1.547 ( 0.934)   Loss 2.6381e+00 (2.8830e+00)   Acc@1  40.00 ( 3
6.76)   Acc@5  68.40 ( 63.01)
Epoch: [1][1400/2563]   Time  1.459 ( 1.605)   Data  0.792 ( 0.934)   Loss 2.7574e+00 (2.8820e+00)   Acc@1  38.60 ( 3
6.78)   Acc@5  66.20 ( 63.02)
Epoch: [1][1410/2563]   Time  1.240 ( 1.605)   Data  0.575 ( 0.934)   Loss 2.6453e+00 (2.8809e+00)   Acc@1  38.80 ( 3
6.80)   Acc@5  66.40 ( 63.05)
Epoch: [1][1420/2563]   Time  1.360 ( 1.605)   Data  0.696 ( 0.934)   Loss 2.6577e+00 (2.8798e+00)   Acc@1  39.00 ( 3
6.82)   Acc@5  67.80 ( 63.07)
Epoch: [1][1430/2563]   Time  1.638 ( 1.605)   Data  0.970 ( 0.934)   Loss 2.7984e+00 (2.8787e+00)   Acc@1  40.00 ( 3
6.83)   Acc@5  64.20 ( 63.09)
Epoch: [1][1440/2563]   Time  1.811 ( 1.605)   Data  1.144 ( 0.934)   Loss 2.8922e+00 (2.8776e+00)   Acc@1  36.00 ( 3
6.85)   Acc@5  64.00 ( 63.11)
Epoch: [1][1450/2563]   Time  1.653 ( 1.605)   Data  0.988 ( 0.934)   Loss 2.7274e+00 (2.8764e+00)   Acc@1  40.00 ( 3
6.88)   Acc@5  68.80 ( 63.14)
Epoch: [1][1460/2563]   Time  1.670 ( 1.605)   Data  1.007 ( 0.934)   Loss 2.7031e+00 (2.8752e+00)   Acc@1  38.00 ( 3
6.90)   Acc@5  67.80 ( 63.16)
Epoch: [1][1470/2563]   Time  1.353 ( 1.605)   Data  0.691 ( 0.934)   Loss 2.7009e+00 (2.8738e+00)   Acc@1  40.80 ( 3
6.92)   Acc@5  66.60 ( 63.18)
Epoch: [1][1480/2563]   Time  1.478 ( 1.605)   Data  0.813 ( 0.934)   Loss 2.7567e+00 (2.8732e+00)   Acc@1  39.20 ( 3
6.92)   Acc@5  64.80 ( 63.19)
Epoch: [1][1490/2563]   Time  1.408 ( 1.605)   Data  0.742 ( 0.934)   Loss 2.7198e+00 (2.8721e+00)   Acc@1  41.00 ( 3
6.94)   Acc@5  66.00 ( 63.21)
Epoch: [1][1500/2563]   Time  0.926 ( 1.604)   Data  0.258 ( 0.934)   Loss 2.5678e+00 (2.8710e+00)   Acc@1  41.80 ( 3
6.96)   Acc@5  68.00 ( 63.23)
Epoch: [1][1510/2563]   Time  0.795 ( 1.604)   Data  0.129 ( 0.933)   Loss 2.5453e+00 (2.8698e+00)   Acc@1  38.80 ( 3
6.99)   Acc@5  71.80 ( 63.25)
Epoch: [1][1520/2563]   Time  0.676 ( 1.605)   Data  0.002 ( 0.934)   Loss 2.9442e+00 (2.8689e+00)   Acc@1  35.20 ( 3
7.01)   Acc@5  63.20 ( 63.27)
Epoch: [1][1530/2563]   Time  0.764 ( 1.604)   Data  0.096 ( 0.933)   Loss 2.6021e+00 (2.8680e+00)   Acc@1  43.60 ( 3
7.02)   Acc@5  69.60 ( 63.29)
Epoch: [1][1540/2563]   Time  0.762 ( 1.604)   Data  0.097 ( 0.933)   Loss 2.7456e+00 (2.8670e+00)   Acc@1  43.00 ( 3
7.04)   Acc@5  66.20 ( 63.31)
Epoch: [1][1550/2563]   Time  0.761 ( 1.604)   Data  0.098 ( 0.933)   Loss 2.7019e+00 (2.8658e+00)   Acc@1  38.60 ( 3
7.06)   Acc@5  64.80 ( 63.33)
Epoch: [1][1560/2563]   Time  0.761 ( 1.604)   Data  0.098 ( 0.933)   Loss 2.7233e+00 (2.8647e+00)   Acc@1  39.00 ( 3
7.08)   Acc@5  65.20 ( 63.35)
Epoch: [1][1570/2563]   Time  0.936 ( 1.604)   Data  0.271 ( 0.933)   Loss 2.7825e+00 (2.8638e+00)   Acc@1  37.20 ( 3
7.10)   Acc@5  62.80 ( 63.37)
Epoch: [1][1580/2563]   Time  0.762 ( 1.604)   Data  0.097 ( 0.933)   Loss 2.8371e+00 (2.8631e+00)   Acc@1  40.40 ( 3
7.11)   Acc@5  62.60 ( 63.38)
Epoch: [1][1590/2563]   Time  0.857 ( 1.604)   Data  0.193 ( 0.933)   Loss 2.6568e+00 (2.8622e+00)   Acc@1  40.80 ( 3
7.12)   Acc@5  67.00 ( 63.39)
Epoch: [1][1600/2563]   Time  0.759 ( 1.604)   Data  0.098 ( 0.933)   Loss 2.6923e+00 (2.8613e+00)   Acc@1  37.20 ( 3
7.14)   Acc@5  65.40 ( 63.41)
Epoch: [1][1610/2563]   Time  0.670 ( 1.604)   Data  0.002 ( 0.933)   Loss 2.7192e+00 (2.8601e+00)   Acc@1  41.00 ( 3
7.16)   Acc@5  66.40 ( 63.43)
Epoch: [1][1620/2563]   Time  0.764 ( 1.604)   Data  0.098 ( 0.933)   Loss 2.8475e+00 (2.8591e+00)   Acc@1  36.40 ( 3
```

```
7.17)   Acc@5  62.80 ( 63.44)
Epoch: [1][1630/2563]   Time  1.272 ( 1.605)    Data  0.605 ( 0.934)    Loss 2.7330e+00 (2.8581e+00)    Acc@1  38.60 ( 3
7.19)   Acc@5  65.40 ( 63.46)
Epoch: [1][1640/2563]   Time  1.254 ( 1.605)    Data  0.590 ( 0.934)    Loss 2.7299e+00 (2.8569e+00)    Acc@1  41.60 ( 3
7.21)   Acc@5  63.80 ( 63.48)
Epoch: [1][1650/2563]   Time  1.319 ( 1.605)    Data  0.654 ( 0.934)    Loss 2.5617e+00 (2.8557e+00)    Acc@1  41.20 ( 3
7.22)   Acc@5  69.00 ( 63.50)
Epoch: [1][1660/2563]   Time  1.582 ( 1.605)    Data  0.919 ( 0.934)    Loss 2.5267e+00 (2.8545e+00)    Acc@1  45.60 ( 3
7.24)   Acc@5  67.80 ( 63.52)
Epoch: [1][1670/2563]   Time  1.473 ( 1.605)    Data  0.808 ( 0.934)    Loss 2.7400e+00 (2.8535e+00)    Acc@1  39.80 ( 3
7.26)   Acc@5  66.40 ( 63.54)
Epoch: [1][1680/2563]   Time  1.469 ( 1.605)    Data  0.802 ( 0.934)    Loss 2.6285e+00 (2.8524e+00)    Acc@1  42.00 ( 3
7.27)   Acc@5  68.60 ( 63.56)
Epoch: [1][1690/2563]   Time  0.968 ( 1.605)    Data  0.304 ( 0.934)    Loss 2.6097e+00 (2.8514e+00)    Acc@1  40.20 ( 3
7.29)   Acc@5  68.60 ( 63.57)
Epoch: [1][1700/2563]   Time  0.766 ( 1.605)    Data  0.100 ( 0.934)    Loss 2.6985e+00 (2.8504e+00)    Acc@1  40.60 ( 3
7.31)   Acc@5  67.40 ( 63.59)
Epoch: [1][1710/2563]   Time  0.668 ( 1.605)    Data  0.002 ( 0.934)    Loss 2.5838e+00 (2.8497e+00)    Acc@1  42.40 ( 3
7.32)   Acc@5  67.80 ( 63.60)
Epoch: [1][1720/2563]   Time  0.669 ( 1.605)    Data  0.003 ( 0.934)    Loss 2.6009e+00 (2.8488e+00)    Acc@1  38.60 ( 3
7.34)   Acc@5  68.00 ( 63.62)
Epoch: [1][1730/2563]   Time  0.764 ( 1.605)    Data  0.100 ( 0.934)    Loss 2.6106e+00 (2.8479e+00)    Acc@1  41.40 ( 3
7.35)   Acc@5  67.00 ( 63.64)
Epoch: [1][1740/2563]   Time  0.772 ( 1.605)    Data  0.104 ( 0.934)    Loss 2.6984e+00 (2.8466e+00)    Acc@1  37.80 ( 3
7.37)   Acc@5  67.60 ( 63.66)
Epoch: [1][1750/2563]   Time  1.039 ( 1.605)    Data  0.375 ( 0.934)    Loss 2.7405e+00 (2.8457e+00)    Acc@1  37.20 ( 3
7.38)   Acc@5  65.00 ( 63.67)
Epoch: [1][1760/2563]   Time  0.772 ( 1.605)    Data  0.106 ( 0.934)    Loss 2.5125e+00 (2.8446e+00)    Acc@1  42.60 ( 3
7.41)   Acc@5  68.20 ( 63.69)
Epoch: [1][1770/2563]   Time  0.671 ( 1.605)    Data  0.002 ( 0.933)    Loss 2.6475e+00 (2.8435e+00)    Acc@1  42.00 ( 3
7.43)   Acc@5  67.00 ( 63.71)
Epoch: [1][1780/2563]   Time  0.811 ( 1.605)    Data  0.145 ( 0.933)    Loss 2.7169e+00 (2.8425e+00)    Acc@1  39.60 ( 3
7.45)   Acc@5  64.40 ( 63.72)
Epoch: [1][1790/2563]   Time  0.667 ( 1.605)    Data  0.003 ( 0.933)    Loss 2.6664e+00 (2.8414e+00)    Acc@1  38.80 ( 3
7.46)   Acc@5  65.60 ( 63.74)
Epoch: [1][1800/2563]   Time  0.669 ( 1.604)    Data  0.002 ( 0.933)    Loss 2.7516e+00 (2.8405e+00)    Acc@1  41.80 ( 3
7.48)   Acc@5  66.40 ( 63.76)
Epoch: [1][1810/2563]   Time  0.670 ( 1.604)    Data  0.002 ( 0.933)    Loss 2.5774e+00 (2.8395e+00)    Acc@1  39.80 ( 3
7.50)   Acc@5  66.00 ( 63.78)
Epoch: [1][1820/2563]   Time  0.764 ( 1.604)    Data  0.099 ( 0.933)    Loss 2.6124e+00 (2.8387e+00)    Acc@1  41.80 ( 3
7.51)   Acc@5  68.20 ( 63.79)
Epoch: [1][1830/2563]   Time  0.673 ( 1.604)    Data  0.003 ( 0.933)    Loss 2.5084e+00 (2.8376e+00)    Acc@1  42.40 ( 3
7.53)   Acc@5  67.20 ( 63.81)
Epoch: [1][1840/2563]   Time  0.667 ( 1.604)    Data  0.002 ( 0.933)    Loss 2.5446e+00 (2.8364e+00)    Acc@1  42.80 ( 3
7.55)   Acc@5  67.60 ( 63.83)
Epoch: [1][1850/2563]   Time  0.771 ( 1.605)    Data  0.104 ( 0.933)    Loss 2.6100e+00 (2.8356e+00)    Acc@1  42.40 ( 3
7.57)   Acc@5  68.40 ( 63.84)
Epoch: [1][1860/2563]   Time  0.916 ( 1.604)    Data  0.253 ( 0.933)    Loss 2.7619e+00 (2.8348e+00)    Acc@1  39.60 ( 3
7.58)   Acc@5  64.60 ( 63.85)
Epoch: [1][1870/2563]   Time  1.484 ( 1.605)    Data  0.817 ( 0.933)    Loss 2.3262e+00 (2.8336e+00)    Acc@1  47.80 ( 3
7.60)   Acc@5  71.60 ( 63.87)
```

```
Epoch: [1][1880/2563]   Time  1.963 ( 1.605)   Data  1.302 ( 0.934)   Loss 2.5158e+00 (2.8326e+00)   Acc@1  42.00 ( 3
7.62)   Acc@5  69.40 ( 63.89)
Epoch: [1][1890/2563]   Time  2.321 ( 1.605)   Data  1.658 ( 0.934)   Loss 2.5678e+00 (2.8314e+00)   Acc@1  39.00 ( 3
7.64)   Acc@5  69.60 ( 63.91)
Epoch: [1][1900/2563]   Time  2.345 ( 1.605)   Data  1.683 ( 0.934)   Loss 2.5482e+00 (2.8302e+00)   Acc@1  41.80 ( 3
7.66)   Acc@5  68.40 ( 63.93)
Epoch: [1][1910/2563]   Time  2.291 ( 1.605)   Data  1.626 ( 0.934)   Loss 2.6135e+00 (2.8291e+00)   Acc@1  39.20 ( 3
7.67)   Acc@5  67.40 ( 63.95)
Epoch: [1][1920/2563]   Time  2.467 ( 1.605)   Data  1.800 ( 0.934)   Loss 2.4869e+00 (2.8280e+00)   Acc@1  42.80 ( 3
7.70)   Acc@5  70.80 ( 63.97)
Epoch: [1][1930/2563]   Time  2.055 ( 1.605)   Data  1.392 ( 0.934)   Loss 2.6369e+00 (2.8264e+00)   Acc@1  39.20 ( 3
7.73)   Acc@5  69.20 ( 64.00)
Epoch: [1][1940/2563]   Time  2.403 ( 1.605)   Data  1.713 ( 0.934)   Loss 2.6580e+00 (2.8255e+00)   Acc@1  41.60 ( 3
7.74)   Acc@5  68.00 ( 64.01)
Epoch: [1][1950/2563]   Time  2.199 ( 1.605)   Data  1.533 ( 0.934)   Loss 2.6786e+00 (2.8246e+00)   Acc@1  40.80 ( 3
7.76)   Acc@5  66.80 ( 64.03)
Epoch: [1][1960/2563]   Time  2.812 ( 1.605)   Data  2.087 ( 0.934)   Loss 2.7480e+00 (2.8236e+00)   Acc@1  39.80 ( 3
7.78)   Acc@5  66.60 ( 64.04)
Epoch: [1][1970/2563]   Time  2.462 ( 1.605)   Data  1.793 ( 0.934)   Loss 2.7949e+00 (2.8228e+00)   Acc@1  36.40 ( 3
7.79)   Acc@5  65.40 ( 64.06)
Epoch: [1][1980/2563]   Time  2.226 ( 1.605)   Data  1.558 ( 0.934)   Loss 2.6533e+00 (2.8218e+00)   Acc@1  41.20 ( 3
7.81)   Acc@5  67.00 ( 64.08)
Epoch: [1][1990/2563]   Time  2.365 ( 1.605)   Data  1.702 ( 0.934)   Loss 2.5759e+00 (2.8208e+00)   Acc@1  42.60 ( 3
7.82)   Acc@5  68.20 ( 64.10)
Epoch: [1][2000/2563]   Time  2.129 ( 1.605)   Data  1.465 ( 0.934)   Loss 2.6182e+00 (2.8197e+00)   Acc@1  41.20 ( 3
7.85)   Acc@5  66.60 ( 64.12)
Epoch: [1][2010/2563]   Time  1.760 ( 1.605)   Data  1.093 ( 0.934)   Loss 2.6745e+00 (2.8187e+00)   Acc@1  39.60 ( 3
7.86)   Acc@5  68.20 ( 64.13)
Epoch: [1][2020/2563]   Time  1.952 ( 1.605)   Data  1.291 ( 0.934)   Loss 2.7055e+00 (2.8178e+00)   Acc@1  41.00 ( 3
7.88)   Acc@5  66.40 ( 64.15)
Epoch: [1][2030/2563]   Time  1.865 ( 1.605)   Data  1.197 ( 0.933)   Loss 2.6503e+00 (2.8165e+00)   Acc@1  41.20 ( 3
7.90)   Acc@5  67.20 ( 64.17)
Epoch: [1][2040/2563]   Time  1.887 ( 1.604)   Data  1.223 ( 0.933)   Loss 2.5236e+00 (2.8158e+00)   Acc@1  41.00 ( 3
7.91)   Acc@5  70.20 ( 64.18)
Epoch: [1][2050/2563]   Time  2.244 ( 1.605)   Data  1.577 ( 0.934)   Loss 2.7552e+00 (2.8149e+00)   Acc@1  41.60 ( 3
7.93)   Acc@5  66.80 ( 64.21)
Epoch: [1][2060/2563]   Time  2.483 ( 1.605)   Data  1.766 ( 0.934)   Loss 2.6525e+00 (2.8140e+00)   Acc@1  42.00 ( 3
7.95)   Acc@5  67.60 ( 64.22)
Epoch: [1][2070/2563]   Time  2.597 ( 1.605)   Data  1.882 ( 0.933)   Loss 2.5130e+00 (2.8129e+00)   Acc@1  43.00 ( 3
7.97)   Acc@5  67.80 ( 64.24)
Epoch: [1][2080/2563]   Time  2.329 ( 1.604)   Data  1.667 ( 0.933)   Loss 2.6315e+00 (2.8119e+00)   Acc@1  41.40 ( 3
7.98)   Acc@5  67.60 ( 64.26)
Epoch: [1][2090/2563]   Time  2.322 ( 1.604)   Data  1.656 ( 0.933)   Loss 2.6994e+00 (2.8110e+00)   Acc@1  41.00 ( 3
8.00)   Acc@5  65.40 ( 64.28)
Epoch: [1][2100/2563]   Time  2.471 ( 1.604)   Data  1.808 ( 0.933)   Loss 2.4611e+00 (2.8100e+00)   Acc@1  41.20 ( 3
8.01)   Acc@5  68.60 ( 64.29)
Epoch: [1][2110/2563]   Time  1.991 ( 1.604)   Data  1.326 ( 0.933)   Loss 2.5424e+00 (2.8092e+00)   Acc@1  44.20 ( 3
8.03)   Acc@5  69.80 ( 64.30)
Epoch: [1][2120/2563]   Time  1.819 ( 1.604)   Data  1.158 ( 0.933)   Loss 2.4770e+00 (2.8082e+00)   Acc@1  42.60 ( 3
8.04)   Acc@5  72.60 ( 64.32)
Epoch: [1][2130/2563]   Time  1.709 ( 1.604)   Data  1.049 ( 0.933)   Loss 2.7975e+00 (2.8076e+00)   Acc@1  36.60 ( 3
```

```
8.05)   Acc@5  65.00 ( 64.33)
Epoch: [1][2140/2563]   Time  1.311 ( 1.604)    Data  0.650 ( 0.933)    Loss 2.4542e+00 (2.8066e+00)    Acc@1  42.80 ( 3
8.07)   Acc@5  69.60 ( 64.35)
Epoch: [1][2150/2563]   Time  1.724 ( 1.604)    Data  1.058 ( 0.933)    Loss 2.6923e+00 (2.8058e+00)    Acc@1  40.00 ( 3
8.09)   Acc@5  64.20 ( 64.36)
Epoch: [1][2160/2563]   Time  1.820 ( 1.604)    Data  1.154 ( 0.933)    Loss 2.5058e+00 (2.8047e+00)    Acc@1  43.60 ( 3
8.11)   Acc@5  69.20 ( 64.38)
Epoch: [1][2170/2563]   Time  1.701 ( 1.604)    Data  1.035 ( 0.933)    Loss 2.5678e+00 (2.8038e+00)    Acc@1  42.60 ( 3
8.12)   Acc@5  67.60 ( 64.39)
Epoch: [1][2180/2563]   Time  1.847 ( 1.604)    Data  1.183 ( 0.933)    Loss 2.6013e+00 (2.8030e+00)    Acc@1  43.80 ( 3
8.14)   Acc@5  68.20 ( 64.41)
Epoch: [1][2190/2563]   Time  1.587 ( 1.604)    Data  0.922 ( 0.933)    Loss 2.6099e+00 (2.8018e+00)    Acc@1  40.80 ( 3
8.15)   Acc@5  68.40 ( 64.43)
Epoch: [1][2200/2563]   Time  1.814 ( 1.604)    Data  1.151 ( 0.933)    Loss 2.5237e+00 (2.8009e+00)    Acc@1  43.20 ( 3
8.17)   Acc@5  68.40 ( 64.45)
Epoch: [1][2210/2563]   Time  1.365 ( 1.603)    Data  0.699 ( 0.933)    Loss 2.6617e+00 (2.8000e+00)    Acc@1  44.40 ( 3
8.19)   Acc@5  69.20 ( 64.46)
Epoch: [1][2220/2563]   Time  1.057 ( 1.603)    Data  0.391 ( 0.933)    Loss 2.5602e+00 (2.7992e+00)    Acc@1  42.40 ( 3
8.21)   Acc@5  68.80 ( 64.48)
Epoch: [1][2230/2563]   Time  1.146 ( 1.603)    Data  0.482 ( 0.933)    Loss 2.6269e+00 (2.7983e+00)    Acc@1  40.60 ( 3
8.22)   Acc@5  69.80 ( 64.49)
Epoch: [1][2240/2563]   Time  1.017 ( 1.603)    Data  0.352 ( 0.933)    Loss 2.4331e+00 (2.7973e+00)    Acc@1  45.20 ( 3
8.24)   Acc@5  69.20 ( 64.51)
Epoch: [1][2250/2563]   Time  1.172 ( 1.603)    Data  0.507 ( 0.933)    Loss 2.5250e+00 (2.7964e+00)    Acc@1  45.40 ( 3
8.26)   Acc@5  66.00 ( 64.52)
Epoch: [1][2260/2563]   Time  1.486 ( 1.604)    Data  0.823 ( 0.933)    Loss 2.5052e+00 (2.7955e+00)    Acc@1  42.40 ( 3
8.27)   Acc@5  68.80 ( 64.54)
Epoch: [1][2270/2563]   Time  1.141 ( 1.603)    Data  0.480 ( 0.933)    Loss 2.5097e+00 (2.7947e+00)    Acc@1  45.20 ( 3
8.29)   Acc@5  66.60 ( 64.55)
Epoch: [1][2280/2563]   Time  1.204 ( 1.603)    Data  0.537 ( 0.933)    Loss 2.6549e+00 (2.7937e+00)    Acc@1  42.40 ( 3
8.30)   Acc@5  68.40 ( 64.57)
Epoch: [1][2290/2563]   Time  1.393 ( 1.603)    Data  0.726 ( 0.933)    Loss 2.4516e+00 (2.7927e+00)    Acc@1  43.40 ( 3
8.32)   Acc@5  71.20 ( 64.59)
Epoch: [1][2300/2563]   Time  1.599 ( 1.603)    Data  0.932 ( 0.933)    Loss 2.4470e+00 (2.7915e+00)    Acc@1  45.20 ( 3
8.34)   Acc@5  70.00 ( 64.61)
Epoch: [1][2310/2563]   Time  1.554 ( 1.603)    Data  0.889 ( 0.933)    Loss 2.5640e+00 (2.7906e+00)    Acc@1  42.00 ( 3
8.35)   Acc@5  68.60 ( 64.62)
Epoch: [1][2320/2563]   Time  1.697 ( 1.603)    Data  1.029 ( 0.933)    Loss 2.6074e+00 (2.7896e+00)    Acc@1  43.20 ( 3
8.37)   Acc@5  68.60 ( 64.64)
Epoch: [1][2330/2563]   Time  1.812 ( 1.603)    Data  1.144 ( 0.933)    Loss 2.5103e+00 (2.7885e+00)    Acc@1  43.40 ( 3
8.39)   Acc@5  68.20 ( 64.66)
Epoch: [1][2340/2563]   Time  1.373 ( 1.603)    Data  0.710 ( 0.933)    Loss 2.5840e+00 (2.7878e+00)    Acc@1  40.20 ( 3
8.40)   Acc@5  69.20 ( 64.67)
Epoch: [1][2350/2563]   Time  1.814 ( 1.603)    Data  1.150 ( 0.933)    Loss 2.5310e+00 (2.7869e+00)    Acc@1  43.60 ( 3
8.42)   Acc@5  70.00 ( 64.69)
Epoch: [1][2360/2563]   Time  2.118 ( 1.603)    Data  1.455 ( 0.933)    Loss 2.7038e+00 (2.7859e+00)    Acc@1  40.20 ( 3
8.44)   Acc@5  66.40 ( 64.71)
Epoch: [1][2370/2563]   Time  2.161 ( 1.603)    Data  1.497 ( 0.933)    Loss 2.6061e+00 (2.7849e+00)    Acc@1  42.00 ( 3
8.46)   Acc@5  65.80 ( 64.72)
Epoch: [1][2380/2563]   Time  2.526 ( 1.603)    Data  1.822 ( 0.933)    Loss 2.7658e+00 (2.7842e+00)    Acc@1  39.00 ( 3
8.47)   Acc@5  64.20 ( 64.74)
```

```
Epoch: [1][2390/2563]   Time  2.516 ( 1.603)   Data  1.805 ( 0.933)   Loss 2.5110e+00 (2.7831e+00)   Acc@1  43.20 ( 3
8.49)   Acc@5  69.20 ( 64.75)
Epoch: [1][2400/2563]   Time  2.381 ( 1.603)   Data  1.703 ( 0.933)   Loss 2.8270e+00 (2.7825e+00)   Acc@1  40.00 ( 3
8.50)   Acc@5  63.60 ( 64.77)
Epoch: [1][2410/2563]   Time  2.775 ( 1.603)   Data  2.075 ( 0.933)   Loss 2.4689e+00 (2.7814e+00)   Acc@1  43.60 ( 3
8.52)   Acc@5  70.00 ( 64.79)
Epoch: [1][2420/2563]   Time  2.724 ( 1.603)   Data  2.013 ( 0.933)   Loss 2.5419e+00 (2.7806e+00)   Acc@1  42.00 ( 3
8.53)   Acc@5  70.80 ( 64.80)
Epoch: [1][2430/2563]   Time  2.772 ( 1.604)   Data  2.044 ( 0.933)   Loss 2.5659e+00 (2.7796e+00)   Acc@1  43.20 ( 3
8.55)   Acc@5  67.40 ( 64.82)
Epoch: [1][2440/2563]   Time  2.427 ( 1.604)   Data  1.765 ( 0.933)   Loss 2.5278e+00 (2.7786e+00)   Acc@1  43.40 ( 3
8.56)   Acc@5  70.00 ( 64.84)
Epoch: [1][2450/2563]   Time  2.597 ( 1.604)   Data  1.872 ( 0.933)   Loss 2.4110e+00 (2.7777e+00)   Acc@1  47.20 ( 3
8.59)   Acc@5  71.20 ( 64.86)
Epoch: [1][2460/2563]   Time  2.640 ( 1.604)   Data  1.916 ( 0.933)   Loss 2.5104e+00 (2.7768e+00)   Acc@1  43.20 ( 3
8.60)   Acc@5  68.80 ( 64.87)
Epoch: [1][2470/2563]   Time  2.617 ( 1.604)   Data  1.895 ( 0.933)   Loss 2.6178e+00 (2.7759e+00)   Acc@1  42.40 ( 3
8.62)   Acc@5  68.00 ( 64.89)
Epoch: [1][2480/2563]   Time  2.677 ( 1.604)   Data  1.950 ( 0.933)   Loss 2.5048e+00 (2.7750e+00)   Acc@1  40.60 ( 3
8.64)   Acc@5  69.40 ( 64.90)
Epoch: [1][2490/2563]   Time  2.513 ( 1.604)   Data  1.795 ( 0.933)   Loss 2.5673e+00 (2.7741e+00)   Acc@1  43.20 ( 3
8.65)   Acc@5  69.20 ( 64.92)
Epoch: [1][2500/2563]   Time  2.574 ( 1.604)   Data  1.869 ( 0.933)   Loss 2.7587e+00 (2.7733e+00)   Acc@1  39.60 ( 3
8.67)   Acc@5  64.60 ( 64.93)
Epoch: [1][2510/2563]   Time  2.517 ( 1.604)   Data  1.853 ( 0.933)   Loss 2.6452e+00 (2.7726e+00)   Acc@1  44.00 ( 3
8.68)   Acc@5  69.00 ( 64.95)
Epoch: [1][2520/2563]   Time  2.428 ( 1.603)   Data  1.756 ( 0.932)   Loss 2.6210e+00 (2.7716e+00)   Acc@1  41.00 ( 3
8.70)   Acc@5  65.40 ( 64.96)
Epoch: [1][2530/2563]   Time  2.487 ( 1.603)   Data  1.778 ( 0.932)   Loss 2.6311e+00 (2.7711e+00)   Acc@1  40.00 ( 3
8.71)   Acc@5  68.40 ( 64.97)
Epoch: [1][2540/2563]   Time  2.455 ( 1.603)   Data  1.789 ( 0.932)   Loss 2.4217e+00 (2.7700e+00)   Acc@1  43.60 ( 3
8.73)   Acc@5  73.00 ( 64.99)
Epoch: [1][2550/2563]   Time  2.398 ( 1.603)   Data  1.737 ( 0.932)   Loss 2.4968e+00 (2.7690e+00)   Acc@1  42.00 ( 3
8.74)   Acc@5  70.00 ( 65.01)
Epoch: [1][2560/2563]   Time  2.516 ( 1.603)   Data  1.819 ( 0.932)   Loss 2.5480e+00 (2.7681e+00)   Acc@1  42.80 ( 3
8.76)   Acc@5  70.60 ( 65.03)
Test: [  0/100] Time  4.330 ( 4.330)   Loss 2.1737e+00 (2.1737e+00)   Acc@1  52.40 ( 52.40)   Acc@5  77.00 ( 77.00)
Test: [ 10/100] Time  2.921 ( 1.946)   Loss 2.2649e+00 (2.6023e+00)   Acc@1  50.80 ( 43.71)   Acc@5  73.00 ( 68.35)
Test: [ 20/100] Time  2.814 ( 1.824)   Loss 2.8596e+00 (2.6530e+00)   Acc@1  28.20 ( 39.93)   Acc@5  60.20 ( 67.29)
Test: [ 30/100] Time  2.884 ( 1.780)   Loss 2.2167e+00 (2.5915e+00)   Acc@1  46.20 ( 39.62)   Acc@5  73.80 ( 68.55)
Test: [ 40/100] Time  2.654 ( 1.763)   Loss 2.8088e+00 (2.6115e+00)   Acc@1  36.40 ( 39.54)   Acc@5  64.40 ( 68.16)
Test: [ 50/100] Time  2.757 ( 1.748)   Loss 3.5327e+00 (2.7415e+00)   Acc@1  25.20 ( 37.62)   Acc@5  48.80 ( 65.88)
Test: [ 60/100] Time  2.744 ( 1.742)   Loss 2.9725e+00 (2.7934e+00)   Acc@1  40.20 ( 37.20)   Acc@5  59.60 ( 64.73)
Test: [ 70/100] Time  2.839 ( 1.730)   Loss 3.5670e+00 (2.8491e+00)   Acc@1  23.60 ( 36.54)   Acc@5  53.40 ( 63.78)
Test: [ 80/100] Time  2.913 ( 1.723)   Loss 2.8902e+00 (2.9103e+00)   Acc@1  39.80 ( 35.98)   Acc@5  58.80 ( 62.63)
Test: [ 90/100] Time  2.933 ( 1.715)   Loss 3.4774e+00 (2.9609e+00)   Acc@1  28.20 ( 35.24)   Acc@5  52.60 ( 61.65)
 * Acc@1 35.722 Acc@5 62.130
lr: 0.0
```

In [47]:

```
writer.close()
```

```
%load_ext tensorboard
%tensorboard --logdir=runs
```

# Index of /

| Name | Size | Date Modified |
|---|---|---|
| .vol/ | | 1/1/20, 12:00:00 AM |
| Applications/ | | 10/28/21, 12:44:07 AM |
| bin/ | | 1/1/20, 12:00:00 AM |
| cores/ | | 6/5/20, 4:55:45 PM |
| dev/ | | 10/30/21, 11:18:08 AM |
| etc/ | | 10/15/21, 12:23:15 PM |
| home/ | | 10/30/21, 11:24:08 PM |
| Library/ | | 10/15/21, 12:22:52 PM |
| opt/ | | 2/21/21, 9:52:50 PM |
| private/ | | 1/1/20, 12:00:00 AM |
| sbin/ | | 1/1/20, 12:00:00 AM |
| System/ | | 1/1/20, 12:00:00 AM |
| tmp/ | | 10/31/21, 4:35:04 AM |
| Users/ | | 10/15/21, 11:19:49 AM |
| usr/ | | 1/1/20, 12:00:00 AM |
| var/ | | 10/15/21, 11:19:51 AM |
| Volumes/ | | 10/30/21, 11:18:19 AM |
| .file | 0 B | 1/1/20, 12:00:00 AM |
| .VolumeIcon.icns | 0 B | |

In [ ]: