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
```

```
Requirement already satisfied: wandb in /opt/conda/lib/python3.8/site-packages (0.12.6)
Requirement already satisfied: shortuuid>=0.5.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (1.0.1)
Requirement already satisfied: sentry-sdk>=1.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (1.4.3)
Requirement already satisfied: docker-pycreds>=0.4.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (0.4.0)
Requirement already satisfied: Click!=8.0.0,>=7.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (7.1.2)
Requirement already satisfied: six>=1.13.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (1.16.0)
Requirement already satisfied: GitPython>=1.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (3.1.24)
Requirement already satisfied: promise<3,>=2.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.3)
Requirement already satisfied: requests<3,>=2.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.26.0)
Requirement already satisfied: yaspin>=1.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.1.0)
Requirement already satisfied: psutil>=5.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (5.8.0)
Requirement already satisfied: subprocess32>=3.5.3 in /opt/conda/lib/python3.8/site-packages (from wandb) (3.5.4)
Requirement already satisfied: python-dateutil>=2.6.1 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.8.2)
Requirement already satisfied: pathtools in /opt/conda/lib/python3.8/site-packages (from wandb) (0.1.2)
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)
Requirement already satisfied: configparser>=3.8.1 in /opt/conda/lib/python3.8/site-packages (from wandb) (5.0.2)
Requirement already satisfied: gitdb<5,>=4.0.1 in /opt/conda/lib/python3.8/site-packages (from GitPython>=1.0.0->wandb)
(4.0.9)
```

```
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)
Requirement already satisfied: smmap<6,>=3.0.1 in /opt/conda/lib/python3.8/site-packages (from gitdb<5,>=4.0.1->GitPython>=1.0.0->wandb) (5.0.0)
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->wandb) (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)
Requirement already satisfied: termcolor<2.0.0,>=1.1.0 in /opt/conda/lib/python3.8/site-packages (from yaspin>=1.0.0->wandb) (1.1.0)
WARNING: Running pip as the 'root' user can result in broken permissions and conflicting behaviour with the system package manager. It is recommended to use a virtual environment instead: https://pip.pypa.io/warnings/venv
```

In [5]:
```
import wandb
wandb.login()
```

Out[5]:
```
wandb: Currently logged in as: rochelleli (use `wandb login --relogin` to force relogin)
True
```

In [6]:
```
wandb.init(project="w251-hw9", entity="rochelleli")
```

Out[6]:
Syncing run **frightful-horseman-6** to Weights & Biases (docs).

Display W&B run

In [7]:
```
SEED=2

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

In [8]:
```
torch.cuda.device_count()
```

Out[8]:
```
1
```

In [9]:
```
START_EPOCH = 0
```

In [10]:
```
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 [11]:
```
if not torch.cuda.is_available():
    print('GPU not detected.. did you pass through your GPU?')
```

In [12]:
```
wandb.init(config={"epochs": EPOCHS, "batch_size": TRAIN_BATCH, "momentum": MOMENTUM, "WEIGHT_DECAY": WEIGHT_DECAY, "arc
```

Finishing last run (ID:285ki779) before initializing another...

Waiting for W&B process to finish, PID 603... **(success).**

Synced 4 W&B file(s), 0 media file(s), 0 artifact file(s) and 0 other file(s)

Synced **frightful-horseman-6**: https://wandb.ai/rochelleli/w251-hw9/runs/285ki779

Find logs at: `./wandb/run-20211031_103311-285ki779/logs`

Successfully finished last run (ID:285ki779). Initializing new run:

Syncing run **possessed-crone-1** to Weights & Biases (docs).

Out[12]: Display W&B run

In [13]:
```
GPU = 0
```

In [14]:
```
torch.cuda.set_device(GPU)
```

In [15]:
```
global_step = 0
```

In [16]:
```
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 [17]:

```
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 [18]:

```
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 [19]:

```
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 [20]:

```
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 [21]:

```
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 [22]:

```
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 [23]:

```
cudnn.benchmark = True
```

In [24]:

```
import torch.distributed as dist
import sys
```

In [25]:

```
WORLD_SIZE = 2
BACKEND = 'nccl'

URL = 'tcp://172.31.17.123:443'

RANK = 1

dist.init_process_group(backend = BACKEND, init_method=URL,rank=RANK, world_size=WORLD_SIZE)
```

In [26]:
```
imagenet_mean_RGB = [0.47889522, 0.47227842, 0.43047404]
imagenet_std_RGB = [0.229, 0.224, 0.225]
```

In [27]:
```
normalize = transforms.Normalize(mean=imagenet_mean_RGB, std=imagenet_std_RGB)
```

In [28]:
```
IMG_SIZE = 224
NUM_CLASSES = 1000
```

In [29]:
```
model = models.__dict__[ARCH]()
```

In [30]:
```
inf = model.fc.in_features
```

In [31]:
```
model.fc = nn.Linear(inf, NUM_CLASSES)
```

In [32]:
```
model.cuda(GPU)
model = torch.nn.parallel.DistributedDataParallel(model, device_ids=[GPU])
```

In [33]:
```
criterion = nn.CrossEntropyLoss().cuda(GPU)
```

In [34]:
```
optimizer = torch.optim.SGD(model.parameters(), LR,
                            momentum=MOMENTUM,
                            weight_decay=WEIGHT_DECAY)
```

In [35]:
```
scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer, T_max=EPOCHS)
```

In [36]:
```
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 [37]:
```
train_dataset = datasets.ImageFolder(
    TRAINDIR, transform=transform_train)
```

In [38]:
```
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 [39]:
```
val_dataset = datasets.ImageFolder(
    VALDIR, transform=transform_val)
```

In [40]:
```
train_loader = torch.utils.data.DataLoader(
        train_dataset, batch_size=TRAIN_BATCH, shuffle=False,
        num_workers=WORKERS, pin_memory=True, sampler=torch.utils.data.distributed.DistributedSampler(train_dataset))
```

In [41]:
```
val_loader = torch.utils.data.DataLoader(
        val_dataset, batch_size=VAL_BATCH, shuffle=False,
        num_workers=WORKERS, pin_memory=True, sampler=None)
```

In [42]:
```
best_acc1 = 0
```

In [43]:
```
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][   0/1282]   Time 15.616 (15.616)    Data  4.128 ( 4.128)    Loss 7.0156e+00 (7.0156e+00)    Acc@1   0.00 (
0.00)   Acc@5   0.60 (  0.60)
Epoch: [0][  10/1282]   Time  0.785 ( 2.468)    Data  0.127 ( 0.819)    Loss 6.9182e+00 (6.9759e+00)    Acc@1   0.00 (
0.11)   Acc@5   0.80 (  0.64)
Epoch: [0][  20/1282]   Time  0.754 ( 2.112)    Data  0.002 ( 0.916)    Loss 6.9132e+00 (6.9532e+00)    Acc@1   0.60 (
0.17)   Acc@5   1.80 (  0.84)
Epoch: [0][  30/1282]   Time  0.764 ( 1.991)    Data  0.003 ( 0.945)    Loss 6.8345e+00 (6.9262e+00)    Acc@1   0.00 (
0.23)   Acc@5   2.00 (  1.00)
Epoch: [0][  40/1282]   Time  0.758 ( 1.902)    Data  0.002 ( 0.938)    Loss 6.7156e+00 (6.8819e+00)    Acc@1   0.60 (
0.38)   Acc@5   3.20 (  1.40)
Epoch: [0][  50/1282]   Time  0.667 ( 1.863)    Data  0.003 ( 0.948)    Loss 6.6571e+00 (6.8398e+00)    Acc@1   0.60 (
0.46)   Acc@5   2.00 (  1.69)
Epoch: [0][  60/1282]   Time  0.695 ( 1.849)    Data  0.002 ( 0.968)    Loss 6.5180e+00 (6.7952e+00)    Acc@1   0.60 (
0.52)   Acc@5   3.40 (  1.95)
Epoch: [0][  70/1282]   Time  0.755 ( 1.820)    Data  0.089 ( 0.966)    Loss 6.4503e+00 (6.7562e+00)    Acc@1   1.40 (
0.57)   Acc@5   5.20 (  2.18)
Epoch: [0][  80/1282]   Time  0.875 ( 1.809)    Data  0.211 ( 0.976)    Loss 6.3870e+00 (6.7132e+00)    Acc@1   1.40 (
0.67)   Acc@5   4.40 (  2.48)
Epoch: [0][  90/1282]   Time  1.269 ( 1.791)    Data  0.602 ( 0.976)    Loss 6.3428e+00 (6.6711e+00)    Acc@1   1.40 (
0.77)   Acc@5   5.20 (  2.80)
Epoch: [0][ 100/1282]   Time  1.650 ( 1.784)    Data  0.982 ( 0.984)    Loss 6.1171e+00 (6.6282e+00)    Acc@1   2.80 (
0.84)   Acc@5   7.00 (  3.09)
Epoch: [0][ 110/1282]   Time  1.124 ( 1.769)    Data  0.454 ( 0.980)    Loss 6.0713e+00 (6.5847e+00)    Acc@1   3.60 (
0.95)   Acc@5   9.60 (  3.41)
Epoch: [0][ 120/1282]   Time  1.230 ( 1.766)    Data  0.556 ( 0.987)    Loss 6.0898e+00 (6.5466e+00)    Acc@1   4.00 (
1.04)   Acc@5   8.20 (  3.70)
Epoch: [0][ 130/1282]   Time  1.233 ( 1.760)    Data  0.557 ( 0.989)    Loss 5.9793e+00 (6.5097e+00)    Acc@1   1.60 (
1.13)   Acc@5   8.40 (  3.98)
Epoch: [0][ 140/1282]   Time  1.002 ( 1.750)    Data  0.322 ( 0.985)    Loss 6.0553e+00 (6.4755e+00)    Acc@1   2.00 (
1.21)   Acc@5   8.00 (  4.30)
Epoch: [0][ 150/1282]   Time  0.981 ( 1.752)    Data  0.300 ( 0.993)    Loss 5.9538e+00 (6.4393e+00)    Acc@1   2.80 (
1.30)   Acc@5  10.00 (  4.59)
Epoch: [0][ 160/1282]   Time  0.772 ( 1.744)    Data  0.091 ( 0.990)    Loss 5.8895e+00 (6.4059e+00)    Acc@1   2.40 (
1.38)   Acc@5   9.80 (  4.90)
Epoch: [0][ 170/1282]   Time  0.770 ( 1.738)    Data  0.090 ( 0.988)    Loss 5.7753e+00 (6.3738e+00)    Acc@1   5.20 (
1.47)   Acc@5   9.60 (  5.17)
Epoch: [0][ 180/1282]   Time  0.773 ( 1.736)    Data  0.092 ( 0.988)    Loss 5.8273e+00 (6.3423e+00)    Acc@1   2.80 (
1.58)   Acc@5   9.40 (  5.45)
Epoch: [0][ 190/1282]   Time  0.687 ( 1.732)    Data  0.002 ( 0.987)    Loss 5.7511e+00 (6.3109e+00)    Acc@1   3.60 (
1.67)   Acc@5  11.40 (  5.72)
Epoch: [0][ 200/1282]   Time  0.721 ( 1.735)    Data  0.002 ( 0.991)    Loss 5.7846e+00 (6.2832e+00)    Acc@1   4.00 (
1.77)   Acc@5  11.20 (  6.00)
Epoch: [0][ 210/1282]   Time  0.690 ( 1.729)    Data  0.003 ( 0.987)    Loss 5.6590e+00 (6.2540e+00)    Acc@1   4.20 (
```

```
1.87)   Acc@5  11.20 (  6.29)
Epoch: [0][ 220/1282]   Time  0.689 ( 1.728)   Data  0.003 ( 0.987)   Loss 5.6057e+00 (6.2269e+00)   Acc@1   4.60 (
1.99)   Acc@5  12.40 (  6.56)
Epoch: [0][ 230/1282]   Time  0.691 ( 1.725)   Data  0.002 ( 0.986)   Loss 5.5602e+00 (6.2001e+00)   Acc@1   6.60 (
2.11)   Acc@5  14.20 (  6.85)
Epoch: [0][ 240/1282]   Time  0.698 ( 1.722)   Data  0.002 ( 0.984)   Loss 5.5252e+00 (6.1724e+00)   Acc@1   5.60 (
2.22)   Acc@5  14.60 (  7.18)
Epoch: [0][ 250/1282]   Time  0.694 ( 1.720)   Data  0.002 ( 0.983)   Loss 5.5984e+00 (6.1456e+00)   Acc@1   4.40 (
2.32)   Acc@5  12.40 (  7.47)
Epoch: [0][ 260/1282]   Time  0.714 ( 1.720)   Data  0.002 ( 0.985)   Loss 5.5894e+00 (6.1197e+00)   Acc@1   4.20 (
2.43)   Acc@5  14.60 (  7.78)
Epoch: [0][ 270/1282]   Time  0.724 ( 1.717)   Data  0.002 ( 0.983)   Loss 5.3588e+00 (6.0944e+00)   Acc@1   5.40 (
2.54)   Acc@5  16.60 (  8.09)
Epoch: [0][ 280/1282]   Time  0.714 ( 1.716)   Data  0.002 ( 0.983)   Loss 5.3976e+00 (6.0702e+00)   Acc@1   7.00 (
2.66)   Acc@5  17.40 (  8.37)
Epoch: [0][ 290/1282]   Time  0.712 ( 1.718)   Data  0.003 ( 0.985)   Loss 5.3315e+00 (6.0474e+00)   Acc@1   7.40 (
2.77)   Acc@5  18.00 (  8.64)
Epoch: [0][ 300/1282]   Time  0.722 ( 1.714)   Data  0.003 ( 0.983)   Loss 5.2160e+00 (6.0225e+00)   Acc@1   8.80 (
2.91)   Acc@5  20.60 (  8.99)
Epoch: [0][ 310/1282]   Time  0.734 ( 1.716)   Data  0.002 ( 0.985)   Loss 5.2408e+00 (5.9975e+00)   Acc@1   5.60 (
3.03)   Acc@5  17.60 (  9.30)
Epoch: [0][ 320/1282]   Time  0.718 ( 1.714)   Data  0.003 ( 0.985)   Loss 5.2504e+00 (5.9747e+00)   Acc@1   7.00 (
3.13)   Acc@5  19.60 (  9.57)
Epoch: [0][ 330/1282]   Time  0.714 ( 1.713)   Data  0.003 ( 0.984)   Loss 5.2507e+00 (5.9515e+00)   Acc@1   6.80 (
3.25)   Acc@5  19.40 (  9.88)
Epoch: [0][ 340/1282]   Time  0.683 ( 1.714)   Data  0.003 ( 0.986)   Loss 5.0554e+00 (5.9281e+00)   Acc@1   6.80 (
3.37)   Acc@5  18.40 ( 10.16)
Epoch: [0][ 350/1282]   Time  0.682 ( 1.712)   Data  0.002 ( 0.984)   Loss 5.3180e+00 (5.9071e+00)   Acc@1   8.20 (
3.48)   Acc@5  19.60 ( 10.43)
Epoch: [0][ 360/1282]   Time  0.727 ( 1.713)   Data  0.002 ( 0.986)   Loss 5.1153e+00 (5.8847e+00)   Acc@1   6.60 (
3.60)   Acc@5  20.40 ( 10.73)
Epoch: [0][ 370/1282]   Time  0.694 ( 1.711)   Data  0.003 ( 0.985)   Loss 5.1310e+00 (5.8657e+00)   Acc@1   9.40 (
3.69)   Acc@5  22.00 ( 10.97)
Epoch: [0][ 380/1282]   Time  0.693 ( 1.711)   Data  0.002 ( 0.986)   Loss 4.9869e+00 (5.8452e+00)   Acc@1   8.20 (
3.81)   Acc@5  23.80 ( 11.26)
Epoch: [0][ 390/1282]   Time  0.708 ( 1.711)   Data  0.002 ( 0.987)   Loss 5.0462e+00 (5.8253e+00)   Acc@1   7.80 (
3.93)   Acc@5  21.20 ( 11.55)
Epoch: [0][ 400/1282]   Time  0.713 ( 1.708)   Data  0.002 ( 0.984)   Loss 5.0080e+00 (5.8062e+00)   Acc@1   9.20 (
4.05)   Acc@5  22.40 ( 11.79)
Epoch: [0][ 410/1282]   Time  0.785 ( 1.706)   Data  0.002 ( 0.983)   Loss 4.9063e+00 (5.7881e+00)   Acc@1   9.40 (
4.16)   Acc@5  26.60 ( 12.07)
Epoch: [0][ 420/1282]   Time  0.724 ( 1.708)   Data  0.002 ( 0.984)   Loss 4.9459e+00 (5.7690e+00)   Acc@1   9.40 (
4.28)   Acc@5  24.20 ( 12.35)
Epoch: [0][ 430/1282]   Time  0.717 ( 1.706)   Data  0.003 ( 0.983)   Loss 4.9767e+00 (5.7492e+00)   Acc@1   7.00 (
4.39)   Acc@5  23.60 ( 12.65)
Epoch: [0][ 440/1282]   Time  0.688 ( 1.706)   Data  0.003 ( 0.984)   Loss 5.0452e+00 (5.7317e+00)   Acc@1   7.60 (
4.50)   Acc@5  20.80 ( 12.91)
Epoch: [0][ 450/1282]   Time  0.686 ( 1.705)   Data  0.002 ( 0.983)   Loss 4.6303e+00 (5.7134e+00)   Acc@1  11.00 (
4.62)   Acc@5  26.40 ( 13.17)
Epoch: [0][ 460/1282]   Time  0.684 ( 1.706)   Data  0.002 ( 0.985)   Loss 4.9812e+00 (5.6956e+00)   Acc@1   9.40 (
4.73)   Acc@5  23.60 ( 13.43)
```

```
Epoch: [0][ 470/1282]   Time  0.690 ( 1.705)   Data  0.002 ( 0.985)   Loss 4.8761e+00 (5.6773e+00)   Acc@1  10.60 (
4.83)   Acc@5  23.80 ( 13.68)
Epoch: [0][ 480/1282]   Time  0.716 ( 1.708)   Data  0.002 ( 0.988)   Loss 4.7287e+00 (5.6586e+00)   Acc@1  13.40 (
4.96)   Acc@5  29.60 ( 13.95)
Epoch: [0][ 490/1282]   Time  0.689 ( 1.706)   Data  0.003 ( 0.987)   Loss 4.6827e+00 (5.6409e+00)   Acc@1   9.80 (
5.07)   Acc@5  28.60 ( 14.21)
Epoch: [0][ 500/1282]   Time  0.683 ( 1.705)   Data  0.002 ( 0.986)   Loss 4.8340e+00 (5.6235e+00)   Acc@1  12.60 (
5.19)   Acc@5  27.00 ( 14.47)
Epoch: [0][ 510/1282]   Time  0.726 ( 1.705)   Data  0.003 ( 0.986)   Loss 4.8993e+00 (5.6076e+00)   Acc@1  10.00 (
5.29)   Acc@5  23.80 ( 14.72)
Epoch: [0][ 520/1282]   Time  0.685 ( 1.704)   Data  0.002 ( 0.986)   Loss 4.7570e+00 (5.5909e+00)   Acc@1   8.60 (
5.40)   Acc@5  25.40 ( 14.97)
Epoch: [0][ 530/1282]   Time  0.691 ( 1.706)   Data  0.002 ( 0.988)   Loss 4.6879e+00 (5.5737e+00)   Acc@1  11.40 (
5.51)   Acc@5  28.80 ( 15.23)
Epoch: [0][ 540/1282]   Time  0.683 ( 1.706)   Data  0.002 ( 0.988)   Loss 4.6027e+00 (5.5574e+00)   Acc@1  13.60 (
5.64)   Acc@5  30.40 ( 15.47)
Epoch: [0][ 550/1282]   Time  0.691 ( 1.705)   Data  0.003 ( 0.988)   Loss 4.5799e+00 (5.5414e+00)   Acc@1  12.80 (
5.75)   Acc@5  31.00 ( 15.72)
Epoch: [0][ 560/1282]   Time  0.684 ( 1.705)   Data  0.002 ( 0.988)   Loss 4.5369e+00 (5.5250e+00)   Acc@1  14.80 (
5.86)   Acc@5  31.00 ( 15.98)
Epoch: [0][ 570/1282]   Time  0.716 ( 1.704)   Data  0.002 ( 0.987)   Loss 4.6132e+00 (5.5088e+00)   Acc@1  11.20 (
5.98)   Acc@5  30.80 ( 16.22)
Epoch: [0][ 580/1282]   Time  0.723 ( 1.706)   Data  0.002 ( 0.989)   Loss 4.6203e+00 (5.4932e+00)   Acc@1  14.40 (
6.09)   Acc@5  30.40 ( 16.46)
Epoch: [0][ 590/1282]   Time  0.687 ( 1.704)   Data  0.002 ( 0.988)   Loss 4.8022e+00 (5.4782e+00)   Acc@1  13.00 (
6.21)   Acc@5  26.00 ( 16.68)
Epoch: [0][ 600/1282]   Time  0.680 ( 1.704)   Data  0.002 ( 0.988)   Loss 4.7113e+00 (5.4627e+00)   Acc@1  12.40 (
6.34)   Acc@5  28.60 ( 16.92)
Epoch: [0][ 610/1282]   Time  0.679 ( 1.703)   Data  0.002 ( 0.988)   Loss 4.6303e+00 (5.4466e+00)   Acc@1  15.00 (
6.47)   Acc@5  29.60 ( 17.18)
Epoch: [0][ 620/1282]   Time  0.781 ( 1.703)   Data  0.002 ( 0.987)   Loss 4.4353e+00 (5.4315e+00)   Acc@1  11.40 (
6.57)   Acc@5  35.20 ( 17.41)
Epoch: [0][ 630/1282]   Time  0.680 ( 1.703)   Data  0.002 ( 0.987)   Loss 4.3729e+00 (5.4163e+00)   Acc@1  16.00 (
6.70)   Acc@5  36.20 ( 17.67)
Epoch: [0][ 640/1282]   Time  0.687 ( 1.702)   Data  0.002 ( 0.986)   Loss 4.5844e+00 (5.4018e+00)   Acc@1  13.20 (
6.80)   Acc@5  32.00 ( 17.89)
Epoch: [0][ 650/1282]   Time  0.682 ( 1.702)   Data  0.002 ( 0.987)   Loss 4.3705e+00 (5.3878e+00)   Acc@1  16.60 (
6.91)   Acc@5  35.20 ( 18.14)
Epoch: [0][ 660/1282]   Time  0.685 ( 1.701)   Data  0.002 ( 0.986)   Loss 4.4043e+00 (5.3734e+00)   Acc@1  14.20 (
7.03)   Acc@5  33.60 ( 18.37)
Epoch: [0][ 670/1282]   Time  0.781 ( 1.701)   Data  0.002 ( 0.986)   Loss 4.3554e+00 (5.3589e+00)   Acc@1  14.00 (
7.13)   Acc@5  33.80 ( 18.60)
Epoch: [0][ 680/1282]   Time  0.685 ( 1.701)   Data  0.002 ( 0.986)   Loss 4.3822e+00 (5.3447e+00)   Acc@1  14.60 (
7.24)   Acc@5  34.00 ( 18.84)
Epoch: [0][ 690/1282]   Time  0.784 ( 1.701)   Data  0.002 ( 0.986)   Loss 4.3757e+00 (5.3306e+00)   Acc@1  16.20 (
7.35)   Acc@5  32.00 ( 19.06)
Epoch: [0][ 700/1282]   Time  0.683 ( 1.701)   Data  0.002 ( 0.986)   Loss 4.4211e+00 (5.3172e+00)   Acc@1  13.00 (
7.46)   Acc@5  31.80 ( 19.27)
Epoch: [0][ 710/1282]   Time  0.686 ( 1.700)   Data  0.002 ( 0.985)   Loss 4.5209e+00 (5.3030e+00)   Acc@1  12.00 (
7.57)   Acc@5  28.60 ( 19.51)
Epoch: [0][ 720/1282]   Time  0.782 ( 1.701)   Data  0.002 ( 0.986)   Loss 4.2508e+00 (5.2898e+00)   Acc@1  16.40 (
```

```
7.67)   Acc@5  37.00 ( 19.72)
Epoch: [0][ 730/1282]   Time  0.685 ( 1.700)    Data  0.002 ( 0.985)    Loss 4.3283e+00 (5.2759e+00)    Acc@1  15.40 (
7.78)   Acc@5  34.80 ( 19.96)
Epoch: [0][ 740/1282]   Time  0.776 ( 1.700)    Data  0.002 ( 0.985)    Loss 4.3973e+00 (5.2624e+00)    Acc@1  14.80 (
7.89)   Acc@5  33.60 ( 20.19)
Epoch: [0][ 750/1282]   Time  0.786 ( 1.699)    Data  0.002 ( 0.984)    Loss 4.3244e+00 (5.2493e+00)    Acc@1  14.80 (
8.01)   Acc@5  34.00 ( 20.41)
Epoch: [0][ 760/1282]   Time  0.885 ( 1.699)    Data  0.002 ( 0.983)    Loss 4.3745e+00 (5.2360e+00)    Acc@1  14.80 (
8.14)   Acc@5  34.60 ( 20.63)
Epoch: [0][ 770/1282]   Time  0.883 ( 1.699)    Data  0.003 ( 0.983)    Loss 4.0093e+00 (5.2225e+00)    Acc@1  19.20 (
8.25)   Acc@5  42.00 ( 20.86)
Epoch: [0][ 780/1282]   Time  0.792 ( 1.699)    Data  0.003 ( 0.982)    Loss 4.1069e+00 (5.2092e+00)    Acc@1  18.80 (
8.36)   Acc@5  42.20 ( 21.09)
Epoch: [0][ 790/1282]   Time  0.789 ( 1.698)    Data  0.002 ( 0.981)    Loss 4.1137e+00 (5.1965e+00)    Acc@1  17.00 (
8.47)   Acc@5  39.20 ( 21.30)
Epoch: [0][ 800/1282]   Time  0.886 ( 1.699)    Data  0.003 ( 0.982)    Loss 4.1481e+00 (5.1840e+00)    Acc@1  16.40 (
8.57)   Acc@5  37.80 ( 21.50)
Epoch: [0][ 810/1282]   Time  0.883 ( 1.698)    Data  0.002 ( 0.980)    Loss 4.1009e+00 (5.1711e+00)    Acc@1  18.60 (
8.69)   Acc@5  39.60 ( 21.73)
Epoch: [0][ 820/1282]   Time  0.886 ( 1.698)    Data  0.002 ( 0.979)    Loss 4.1877e+00 (5.1582e+00)    Acc@1  16.00 (
8.81)   Acc@5  39.40 ( 21.95)
Epoch: [0][ 830/1282]   Time  0.782 ( 1.697)    Data  0.002 ( 0.978)    Loss 4.1428e+00 (5.1458e+00)    Acc@1  18.20 (
8.92)   Acc@5  40.00 ( 22.17)
Epoch: [0][ 840/1282]   Time  1.090 ( 1.696)    Data  0.002 ( 0.976)    Loss 4.1567e+00 (5.1338e+00)    Acc@1  18.40 (
9.03)   Acc@5  37.40 ( 22.37)
Epoch: [0][ 850/1282]   Time  0.886 ( 1.697)    Data  0.002 ( 0.976)    Loss 3.9054e+00 (5.1213e+00)    Acc@1  21.20 (
9.15)   Acc@5  41.20 ( 22.58)
Epoch: [0][ 860/1282]   Time  0.889 ( 1.697)    Data  0.002 ( 0.975)    Loss 4.1883e+00 (5.1095e+00)    Acc@1  17.40 (
9.27)   Acc@5  37.80 ( 22.79)
Epoch: [0][ 870/1282]   Time  0.685 ( 1.697)    Data  0.002 ( 0.975)    Loss 3.9826e+00 (5.0973e+00)    Acc@1  18.20 (
9.37)   Acc@5  42.80 ( 23.00)
Epoch: [0][ 880/1282]   Time  0.784 ( 1.697)    Data  0.002 ( 0.975)    Loss 4.0327e+00 (5.0857e+00)    Acc@1  19.40 (
9.48)   Acc@5  39.40 ( 23.18)
Epoch: [0][ 890/1282]   Time  0.791 ( 1.697)    Data  0.003 ( 0.975)    Loss 4.0477e+00 (5.0735e+00)    Acc@1  20.40 (
9.59)   Acc@5  40.40 ( 23.39)
Epoch: [0][ 900/1282]   Time  0.897 ( 1.697)    Data  0.002 ( 0.974)    Loss 4.0316e+00 (5.0619e+00)    Acc@1  20.00 (
9.70)   Acc@5  43.60 ( 23.59)
Epoch: [0][ 910/1282]   Time  0.683 ( 1.697)    Data  0.003 ( 0.975)    Loss 3.9351e+00 (5.0503e+00)    Acc@1  18.80 (
9.81)   Acc@5  45.40 ( 23.79)
Epoch: [0][ 920/1282]   Time  0.684 ( 1.696)    Data  0.002 ( 0.974)    Loss 3.8933e+00 (5.0382e+00)    Acc@1  20.80 (
9.92)   Acc@5  42.80 ( 24.00)
Epoch: [0][ 930/1282]   Time  0.682 ( 1.696)    Data  0.002 ( 0.974)    Loss 3.9266e+00 (5.0268e+00)    Acc@1  20.80 ( 1
0.03)   Acc@5  42.80 ( 24.20)
Epoch: [0][ 940/1282]   Time  0.781 ( 1.696)    Data  0.002 ( 0.974)    Loss 4.0094e+00 (5.0156e+00)    Acc@1  20.20 ( 1
0.14)   Acc@5  44.60 ( 24.39)
Epoch: [0][ 950/1282]   Time  0.783 ( 1.696)    Data  0.002 ( 0.974)    Loss 4.0414e+00 (5.0043e+00)    Acc@1  18.20 ( 1
0.25)   Acc@5  39.80 ( 24.59)
Epoch: [0][ 960/1282]   Time  0.685 ( 1.695)    Data  0.002 ( 0.973)    Loss 3.8781e+00 (4.9928e+00)    Acc@1  22.80 ( 1
0.36)   Acc@5  41.20 ( 24.79)
Epoch: [0][ 970/1282]   Time  0.778 ( 1.695)    Data  0.002 ( 0.973)    Loss 3.9890e+00 (4.9815e+00)    Acc@1  17.60 ( 1
0.47)   Acc@5  39.40 ( 24.98)
```

```
Epoch: [0][ 980/1282]   Time  0.777 ( 1.695)   Data  0.002 ( 0.973)   Loss 3.7915e+00 (4.9706e+00)   Acc@1  27.00 ( 1
0.58)   Acc@5  47.00 ( 25.17)
Epoch: [0][ 990/1282]   Time  0.793 ( 1.695)   Data  0.002 ( 0.972)   Loss 3.8234e+00 (4.9597e+00)   Acc@1  19.20 ( 1
0.69)   Acc@5  41.80 ( 25.36)
Epoch: [0][1000/1282]   Time  0.713 ( 1.695)   Data  0.002 ( 0.973)   Loss 3.9626e+00 (4.9495e+00)   Acc@1  19.20 ( 1
0.78)   Acc@5  43.60 ( 25.54)
Epoch: [0][1010/1282]   Time  0.779 ( 1.695)   Data  0.002 ( 0.973)   Loss 3.8322e+00 (4.9389e+00)   Acc@1  22.60 ( 1
0.88)   Acc@5  46.40 ( 25.73)
Epoch: [0][1020/1282]   Time  0.881 ( 1.694)   Data  0.002 ( 0.972)   Loss 3.8501e+00 (4.9287e+00)   Acc@1  22.20 ( 1
0.98)   Acc@5  46.80 ( 25.91)
Epoch: [0][1030/1282]   Time  0.680 ( 1.694)   Data  0.003 ( 0.972)   Loss 3.8828e+00 (4.9180e+00)   Acc@1  22.00 ( 1
1.09)   Acc@5  45.20 ( 26.10)
Epoch: [0][1040/1282]   Time  0.880 ( 1.694)   Data  0.002 ( 0.971)   Loss 3.7546e+00 (4.9072e+00)   Acc@1  21.00 ( 1
1.20)   Acc@5  46.60 ( 26.29)
Epoch: [0][1050/1282]   Time  0.781 ( 1.694)   Data  0.002 ( 0.971)   Loss 3.6075e+00 (4.8963e+00)   Acc@1  26.20 ( 1
1.32)   Acc@5  52.20 ( 26.48)
Epoch: [0][1060/1282]   Time  0.682 ( 1.694)   Data  0.002 ( 0.971)   Loss 3.7612e+00 (4.8858e+00)   Acc@1  21.80 ( 1
1.42)   Acc@5  44.60 ( 26.67)
Epoch: [0][1070/1282]   Time  0.698 ( 1.695)   Data  0.002 ( 0.972)   Loss 3.7519e+00 (4.8759e+00)   Acc@1  21.60 ( 1
1.51)   Acc@5  45.40 ( 26.84)
Epoch: [0][1080/1282]   Time  0.685 ( 1.695)   Data  0.002 ( 0.972)   Loss 3.6946e+00 (4.8655e+00)   Acc@1  24.80 ( 1
1.62)   Acc@5  50.80 ( 27.04)
Epoch: [0][1090/1282]   Time  0.719 ( 1.694)   Data  0.002 ( 0.971)   Loss 3.7519e+00 (4.8554e+00)   Acc@1  22.80 ( 1
1.72)   Acc@5  45.20 ( 27.21)
Epoch: [0][1100/1282]   Time  0.717 ( 1.694)   Data  0.002 ( 0.971)   Loss 3.8187e+00 (4.8448e+00)   Acc@1  23.60 ( 1
1.83)   Acc@5  46.00 ( 27.40)
Epoch: [0][1110/1282]   Time  0.714 ( 1.694)   Data  0.002 ( 0.971)   Loss 3.6997e+00 (4.8355e+00)   Acc@1  23.40 ( 1
1.93)   Acc@5  47.40 ( 27.56)
Epoch: [0][1120/1282]   Time  0.696 ( 1.694)   Data  0.002 ( 0.971)   Loss 3.7813e+00 (4.8253e+00)   Acc@1  21.40 ( 1
2.04)   Acc@5  43.40 ( 27.74)
Epoch: [0][1130/1282]   Time  0.779 ( 1.693)   Data  0.002 ( 0.971)   Loss 3.7684e+00 (4.8158e+00)   Acc@1  24.60 ( 1
2.14)   Acc@5  45.00 ( 27.92)
Epoch: [0][1140/1282]   Time  0.794 ( 1.693)   Data  0.003 ( 0.970)   Loss 3.4775e+00 (4.8056e+00)   Acc@1  28.00 ( 1
2.25)   Acc@5  54.00 ( 28.10)
Epoch: [0][1150/1282]   Time  0.785 ( 1.692)   Data  0.002 ( 0.969)   Loss 3.7545e+00 (4.7956e+00)   Acc@1  23.80 ( 1
2.35)   Acc@5  46.20 ( 28.28)
Epoch: [0][1160/1282]   Time  0.782 ( 1.692)   Data  0.003 ( 0.969)   Loss 3.8042e+00 (4.7861e+00)   Acc@1  21.80 ( 1
2.44)   Acc@5  45.40 ( 28.44)
Epoch: [0][1170/1282]   Time  0.780 ( 1.692)   Data  0.002 ( 0.969)   Loss 3.8155e+00 (4.7769e+00)   Acc@1  22.20 ( 1
2.54)   Acc@5  46.40 ( 28.60)
Epoch: [0][1180/1282]   Time  0.680 ( 1.692)   Data  0.002 ( 0.969)   Loss 3.7569e+00 (4.7673e+00)   Acc@1  24.20 ( 1
2.65)   Acc@5  46.20 ( 28.77)
Epoch: [0][1190/1282]   Time  0.692 ( 1.691)   Data  0.003 ( 0.969)   Loss 3.7393e+00 (4.7578e+00)   Acc@1  23.60 ( 1
2.75)   Acc@5  50.60 ( 28.94)
Epoch: [0][1200/1282]   Time  0.778 ( 1.690)   Data  0.002 ( 0.968)   Loss 3.5538e+00 (4.7478e+00)   Acc@1  25.60 ( 1
2.86)   Acc@5  50.20 ( 29.13)
Epoch: [0][1210/1282]   Time  0.689 ( 1.691)   Data  0.002 ( 0.968)   Loss 3.7215e+00 (4.7378e+00)   Acc@1  21.00 ( 1
2.97)   Acc@5  48.00 ( 29.31)
Epoch: [0][1220/1282]   Time  0.688 ( 1.690)   Data  0.002 ( 0.968)   Loss 3.6594e+00 (4.7287e+00)   Acc@1  24.40 ( 1
3.06)   Acc@5  47.40 ( 29.48)
Epoch: [0][1230/1282]   Time  0.715 ( 1.690)   Data  0.002 ( 0.968)   Loss 3.6392e+00 (4.7193e+00)   Acc@1  24.40 ( 1
```

```
3.16)   Acc@5  47.60 ( 29.64)
Epoch: [0][1240/1282]   Time  0.713 ( 1.691)    Data  0.002 ( 0.969)    Loss 3.5611e+00 (4.7102e+00)    Acc@1  25.60 ( 1
3.25)   Acc@5  51.40 ( 29.80)
Epoch: [0][1250/1282]   Time  0.723 ( 1.690)    Data  0.002 ( 0.969)    Loss 3.4320e+00 (4.7010e+00)    Acc@1  28.20 ( 1
3.35)   Acc@5  51.80 ( 29.96)
Epoch: [0][1260/1282]   Time  0.721 ( 1.691)    Data  0.003 ( 0.969)    Loss 3.6723e+00 (4.6922e+00)    Acc@1  24.20 ( 1
3.45)   Acc@5  47.20 ( 30.11)
Epoch: [0][1270/1282]   Time  0.784 ( 1.690)    Data  0.003 ( 0.969)    Loss 3.5192e+00 (4.6832e+00)    Acc@1  25.80 ( 1
3.55)   Acc@5  49.80 ( 30.27)
Epoch: [0][1280/1282]   Time  0.782 ( 1.690)    Data  0.002 ( 0.968)    Loss 3.3037e+00 (4.6737e+00)    Acc@1  30.20 ( 1
3.67)   Acc@5  54.40 ( 30.44)
Test: [  0/100] Time  7.625 ( 7.625)    Loss 3.4868e+00 (3.4868e+00)    Acc@1  24.00 ( 24.00)   Acc@5  55.00 ( 55.00)
Test: [ 10/100] Time  2.983 ( 2.085)    Loss 4.0612e+00 (4.2153e+00)    Acc@1  25.00 ( 20.56)   Acc@5  43.60 ( 42.16)
Test: [ 20/100] Time  2.837 ( 1.902)    Loss 4.0666e+00 (4.0327e+00)    Acc@1  10.80 ( 20.34)   Acc@5  39.40 ( 43.55)
Test: [ 30/100] Time  2.861 ( 1.831)    Loss 4.2806e+00 (4.0831e+00)    Acc@1  19.60 ( 19.03)   Acc@5  43.00 ( 42.27)
Test: [ 40/100] Time  2.579 ( 1.799)    Loss 3.5928e+00 (4.1483e+00)    Acc@1  25.20 ( 19.02)   Acc@5  51.00 ( 41.46)
Test: [ 50/100] Time  2.898 ( 1.779)    Loss 5.1258e+00 (4.1888e+00)    Acc@1   9.00 ( 18.22)   Acc@5  23.00 ( 40.33)
Test: [ 60/100] Time  2.763 ( 1.767)    Loss 3.9949e+00 (4.1866e+00)    Acc@1  24.20 ( 18.38)   Acc@5  41.00 ( 40.12)
Test: [ 70/100] Time  2.787 ( 1.750)    Loss 4.2954e+00 (4.2106e+00)    Acc@1  14.00 ( 18.12)   Acc@5  36.40 ( 39.71)
Test: [ 80/100] Time  2.922 ( 1.739)    Loss 3.9602e+00 (4.2270e+00)    Acc@1  22.60 ( 18.09)   Acc@5  42.60 ( 39.27)
Test: [ 90/100] Time  2.914 ( 1.729)    Loss 4.4120e+00 (4.2401e+00)    Acc@1  17.00 ( 17.88)   Acc@5  33.20 ( 38.81)
 * Acc@1 18.516 Acc@5 39.680
lr: 0.05
Epoch: [1][   0/1282]   Time  6.129 ( 6.129)    Data  3.830 ( 3.830)    Loss 3.5821e+00 (3.5821e+00)    Acc@1  24.00 ( 2
4.00)   Acc@5  51.20 ( 51.20)
Epoch: [1][  10/1282]   Time  2.661 ( 2.074)    Data  0.003 ( 0.445)    Loss 3.5047e+00 (3.4929e+00)    Acc@1  26.20 ( 2
5.85)   Acc@5  52.60 ( 52.00)
Epoch: [1][  20/1282]   Time  2.507 ( 1.861)    Data  0.002 ( 0.238)    Loss 3.3250e+00 (3.4390e+00)    Acc@1  27.00 ( 2
7.16)   Acc@5  55.00 ( 52.63)
Epoch: [1][  30/1282]   Time  2.762 ( 1.803)    Data  0.002 ( 0.162)    Loss 3.3054e+00 (3.4048e+00)    Acc@1  32.60 ( 2
8.21)   Acc@5  56.20 ( 53.63)
Epoch: [1][  40/1282]   Time  2.651 ( 1.774)    Data  0.002 ( 0.123)    Loss 3.3309e+00 (3.3898e+00)    Acc@1  29.00 ( 2
8.62)   Acc@5  55.40 ( 54.03)
Epoch: [1][  50/1282]   Time  2.240 ( 1.743)    Data  0.002 ( 0.100)    Loss 3.4286e+00 (3.3757e+00)    Acc@1  26.60 ( 2
8.89)   Acc@5  53.00 ( 54.12)
Epoch: [1][  60/1282]   Time  1.499 ( 1.724)    Data  0.002 ( 0.084)    Loss 3.3240e+00 (3.3696e+00)    Acc@1  30.40 ( 2
9.13)   Acc@5  55.80 ( 54.23)
Epoch: [1][  70/1282]   Time  1.500 ( 1.713)    Data  0.002 ( 0.072)    Loss 3.2612e+00 (3.3575e+00)    Acc@1  28.60 ( 2
9.33)   Acc@5  54.80 ( 54.47)
Epoch: [1][  80/1282]   Time  0.778 ( 1.699)    Data  0.003 ( 0.064)    Loss 3.3153e+00 (3.3532e+00)    Acc@1  32.20 ( 2
9.43)   Acc@5  59.00 ( 54.62)
Epoch: [1][  90/1282]   Time  0.780 ( 1.696)    Data  0.002 ( 0.057)    Loss 3.2589e+00 (3.3479e+00)    Acc@1  27.60 ( 2
9.39)   Acc@5  55.80 ( 54.62)
Epoch: [1][ 100/1282]   Time  0.783 ( 1.693)    Data  0.002 ( 0.051)    Loss 3.0238e+00 (3.3319e+00)    Acc@1  35.80 ( 2
9.67)   Acc@5  61.00 ( 54.83)
Epoch: [1][ 110/1282]   Time  0.683 ( 1.690)    Data  0.002 ( 0.047)    Loss 3.1109e+00 (3.3257e+00)    Acc@1  32.60 ( 2
9.78)   Acc@5  60.20 ( 54.99)
Epoch: [1][ 120/1282]   Time  0.683 ( 1.684)    Data  0.002 ( 0.043)    Loss 3.1702e+00 (3.3187e+00)    Acc@1  31.80 ( 2
9.88)   Acc@5  58.00 ( 55.14)
Epoch: [1][ 130/1282]   Time  0.832 ( 1.682)    Data  0.092 ( 0.042)    Loss 3.1053e+00 (3.3166e+00)    Acc@1  31.80 ( 2
9.85)   Acc@5  57.60 ( 55.18)
```

```
Epoch: [1][ 140/1282]   Time  0.779 ( 1.682)   Data  0.002 ( 0.040)   Loss 3.3112e+00 (3.3140e+00)   Acc@1  30.60 ( 2
9.92)   Acc@5  55.40 ( 55.23)
Epoch: [1][ 150/1282]   Time  0.682 ( 1.682)   Data  0.002 ( 0.038)   Loss 3.3099e+00 (3.3053e+00)   Acc@1  28.80 ( 3
0.05)   Acc@5  57.00 ( 55.38)
Epoch: [1][ 160/1282]   Time  0.684 ( 1.683)   Data  0.002 ( 0.035)   Loss 3.2475e+00 (3.3004e+00)   Acc@1  29.60 ( 3
0.16)   Acc@5  56.40 ( 55.50)
Epoch: [1][ 170/1282]   Time  0.677 ( 1.682)   Data  0.002 ( 0.034)   Loss 3.0949e+00 (3.2957e+00)   Acc@1  32.20 ( 3
0.20)   Acc@5  58.80 ( 55.56)
Epoch: [1][ 180/1282]   Time  0.681 ( 1.680)   Data  0.002 ( 0.032)   Loss 3.2992e+00 (3.2921e+00)   Acc@1  28.20 ( 3
0.25)   Acc@5  56.80 ( 55.63)
Epoch: [1][ 190/1282]   Time  0.908 ( 1.678)   Data  0.002 ( 0.033)   Loss 3.2490e+00 (3.2873e+00)   Acc@1  30.40 ( 3
0.32)   Acc@5  55.40 ( 55.73)
Epoch: [1][ 200/1282]   Time  1.180 ( 1.677)   Data  0.002 ( 0.038)   Loss 3.3293e+00 (3.2858e+00)   Acc@1  30.40 ( 3
0.32)   Acc@5  55.00 ( 55.79)
Epoch: [1][ 210/1282]   Time  1.283 ( 1.676)   Data  0.002 ( 0.044)   Loss 3.0960e+00 (3.2810e+00)   Acc@1  31.80 ( 3
0.46)   Acc@5  60.60 ( 55.89)
Epoch: [1][ 220/1282]   Time  1.654 ( 1.675)   Data  0.003 ( 0.048)   Loss 3.1168e+00 (3.2800e+00)   Acc@1  34.40 ( 3
0.50)   Acc@5  58.40 ( 55.92)
Epoch: [1][ 230/1282]   Time  1.629 ( 1.675)   Data  0.002 ( 0.049)   Loss 3.1257e+00 (3.2773e+00)   Acc@1  33.60 ( 3
0.54)   Acc@5  58.20 ( 55.97)
Epoch: [1][ 240/1282]   Time  2.015 ( 1.676)   Data  0.002 ( 0.048)   Loss 3.1253e+00 (3.2724e+00)   Acc@1  32.20 ( 3
0.58)   Acc@5  57.60 ( 56.05)
Epoch: [1][ 250/1282]   Time  2.286 ( 1.677)   Data  0.003 ( 0.046)   Loss 3.2923e+00 (3.2685e+00)   Acc@1  30.40 ( 3
0.67)   Acc@5  54.00 ( 56.12)
Epoch: [1][ 260/1282]   Time  1.960 ( 1.677)   Data  0.002 ( 0.045)   Loss 3.4016e+00 (3.2640e+00)   Acc@1  27.80 ( 3
0.74)   Acc@5  53.80 ( 56.20)
Epoch: [1][ 270/1282]   Time  1.807 ( 1.676)   Data  0.002 ( 0.046)   Loss 3.2121e+00 (3.2614e+00)   Acc@1  29.80 ( 3
0.77)   Acc@5  54.80 ( 56.24)
Epoch: [1][ 280/1282]   Time  1.858 ( 1.675)   Data  0.003 ( 0.048)   Loss 3.1782e+00 (3.2569e+00)   Acc@1  33.20 ( 3
0.84)   Acc@5  58.60 ( 56.33)
Epoch: [1][ 290/1282]   Time  2.455 ( 1.676)   Data  0.002 ( 0.052)   Loss 3.0991e+00 (3.2529e+00)   Acc@1  33.40 ( 3
0.89)   Acc@5  58.20 ( 56.38)
Epoch: [1][ 300/1282]   Time  2.264 ( 1.675)   Data  0.003 ( 0.053)   Loss 3.1766e+00 (3.2485e+00)   Acc@1  33.60 ( 3
0.99)   Acc@5  55.40 ( 56.46)
Epoch: [1][ 310/1282]   Time  2.421 ( 1.675)   Data  0.002 ( 0.054)   Loss 3.1221e+00 (3.2444e+00)   Acc@1  36.80 ( 3
1.05)   Acc@5  60.20 ( 56.56)
Epoch: [1][ 320/1282]   Time  2.794 ( 1.676)   Data  0.002 ( 0.053)   Loss 3.1300e+00 (3.2415e+00)   Acc@1  34.00 ( 3
1.08)   Acc@5  58.40 ( 56.60)
Epoch: [1][ 330/1282]   Time  2.502 ( 1.675)   Data  0.002 ( 0.051)   Loss 3.0978e+00 (3.2393e+00)   Acc@1  36.40 ( 3
1.13)   Acc@5  62.00 ( 56.64)
Epoch: [1][ 340/1282]   Time  2.611 ( 1.675)   Data  0.002 ( 0.050)   Loss 3.1287e+00 (3.2356e+00)   Acc@1  34.40 ( 3
1.18)   Acc@5  58.00 ( 56.70)
Epoch: [1][ 350/1282]   Time  2.737 ( 1.675)   Data  0.002 ( 0.048)   Loss 3.2075e+00 (3.2321e+00)   Acc@1  32.00 ( 3
1.23)   Acc@5  57.60 ( 56.77)
Epoch: [1][ 360/1282]   Time  2.536 ( 1.674)   Data  0.002 ( 0.047)   Loss 3.1240e+00 (3.2282e+00)   Acc@1  31.20 ( 3
1.27)   Acc@5  60.00 ( 56.83)
Epoch: [1][ 370/1282]   Time  2.487 ( 1.674)   Data  0.002 ( 0.046)   Loss 3.1326e+00 (3.2259e+00)   Acc@1  32.00 ( 3
1.28)   Acc@5  57.20 ( 56.86)
Epoch: [1][ 380/1282]   Time  2.636 ( 1.673)   Data  0.002 ( 0.045)   Loss 2.8880e+00 (3.2226e+00)   Acc@1  37.20 ( 3
1.32)   Acc@5  63.00 ( 56.92)
Epoch: [1][ 390/1282]   Time  2.856 ( 1.673)   Data  0.002 ( 0.044)   Loss 3.1460e+00 (3.2190e+00)   Acc@1  32.00 ( 3
```

```
1.39)   Acc@5  57.20 ( 56.97)
Epoch: [1][ 400/1282]   Time  2.514 ( 1.673)    Data  0.002 ( 0.043)    Loss 2.9460e+00 (3.2160e+00)    Acc@1  34.40 ( 3
1.44)   Acc@5  60.60 ( 57.04)
Epoch: [1][ 410/1282]   Time  2.718 ( 1.674)    Data  0.002 ( 0.042)    Loss 2.9519e+00 (3.2144e+00)    Acc@1  35.00 ( 3
1.48)   Acc@5  60.80 ( 57.06)
Epoch: [1][ 420/1282]   Time  2.471 ( 1.674)    Data  0.003 ( 0.041)    Loss 3.0618e+00 (3.2113e+00)    Acc@1  30.20 ( 3
1.53)   Acc@5  61.60 ( 57.11)
Epoch: [1][ 430/1282]   Time  2.579 ( 1.673)    Data  0.002 ( 0.040)    Loss 3.0770e+00 (3.2083e+00)    Acc@1  34.40 ( 3
1.57)   Acc@5  61.40 ( 57.16)
Epoch: [1][ 440/1282]   Time  2.591 ( 1.673)    Data  0.002 ( 0.039)    Loss 2.9669e+00 (3.2056e+00)    Acc@1  34.40 ( 3
1.61)   Acc@5  61.60 ( 57.22)
Epoch: [1][ 450/1282]   Time  2.437 ( 1.674)    Data  0.002 ( 0.038)    Loss 2.9060e+00 (3.2036e+00)    Acc@1  34.20 ( 3
1.64)   Acc@5  60.40 ( 57.24)
Epoch: [1][ 460/1282]   Time  2.657 ( 1.674)    Data  0.002 ( 0.038)    Loss 3.1396e+00 (3.2007e+00)    Acc@1  30.80 ( 3
1.69)   Acc@5  59.00 ( 57.29)
Epoch: [1][ 470/1282]   Time  2.674 ( 1.675)    Data  0.002 ( 0.037)    Loss 3.1846e+00 (3.1978e+00)    Acc@1  30.80 ( 3
1.72)   Acc@5  57.40 ( 57.34)
Epoch: [1][ 480/1282]   Time  2.767 ( 1.675)    Data  0.002 ( 0.036)    Loss 2.9806e+00 (3.1947e+00)    Acc@1  36.00 ( 3
1.78)   Acc@5  59.80 ( 57.40)
Epoch: [1][ 490/1282]   Time  2.575 ( 1.675)    Data  0.003 ( 0.035)    Loss 3.0998e+00 (3.1925e+00)    Acc@1  33.80 ( 3
1.81)   Acc@5  58.80 ( 57.44)
Epoch: [1][ 500/1282]   Time  2.951 ( 1.676)    Data  0.003 ( 0.035)    Loss 3.1043e+00 (3.1900e+00)    Acc@1  34.40 ( 3
1.87)   Acc@5  59.60 ( 57.49)
Epoch: [1][ 510/1282]   Time  2.818 ( 1.677)    Data  0.002 ( 0.034)    Loss 3.1869e+00 (3.1883e+00)    Acc@1  31.00 ( 3
1.90)   Acc@5  58.20 ( 57.53)
Epoch: [1][ 520/1282]   Time  2.662 ( 1.677)    Data  0.002 ( 0.033)    Loss 3.0819e+00 (3.1851e+00)    Acc@1  32.80 ( 3
1.94)   Acc@5  59.40 ( 57.59)
Epoch: [1][ 530/1282]   Time  2.603 ( 1.677)    Data  0.002 ( 0.033)    Loss 3.1018e+00 (3.1826e+00)    Acc@1  33.00 ( 3
1.97)   Acc@5  57.80 ( 57.62)
Epoch: [1][ 540/1282]   Time  2.696 ( 1.677)    Data  0.002 ( 0.032)    Loss 2.8581e+00 (3.1807e+00)    Acc@1  34.60 ( 3
2.01)   Acc@5  62.20 ( 57.64)
Epoch: [1][ 550/1282]   Time  2.521 ( 1.676)    Data  0.002 ( 0.032)    Loss 2.8900e+00 (3.1783e+00)    Acc@1  39.20 ( 3
2.05)   Acc@5  64.00 ( 57.69)
Epoch: [1][ 560/1282]   Time  2.723 ( 1.678)    Data  0.002 ( 0.031)    Loss 2.9404e+00 (3.1757e+00)    Acc@1  35.40 ( 3
2.09)   Acc@5  62.00 ( 57.75)
Epoch: [1][ 570/1282]   Time  2.511 ( 1.677)    Data  0.002 ( 0.031)    Loss 2.9945e+00 (3.1729e+00)    Acc@1  35.00 ( 3
2.13)   Acc@5  61.40 ( 57.79)
Epoch: [1][ 580/1282]   Time  2.598 ( 1.677)    Data  0.003 ( 0.030)    Loss 3.1908e+00 (3.1702e+00)    Acc@1  33.40 ( 3
2.17)   Acc@5  57.00 ( 57.84)
Epoch: [1][ 590/1282]   Time  2.848 ( 1.677)    Data  0.003 ( 0.030)    Loss 3.2621e+00 (3.1680e+00)    Acc@1  32.40 ( 3
2.21)   Acc@5  55.20 ( 57.89)
Epoch: [1][ 600/1282]   Time  2.481 ( 1.676)    Data  0.003 ( 0.029)    Loss 3.2347e+00 (3.1660e+00)    Acc@1  31.40 ( 3
2.24)   Acc@5  55.60 ( 57.91)
Epoch: [1][ 610/1282]   Time  2.710 ( 1.676)    Data  0.003 ( 0.029)    Loss 3.0421e+00 (3.1629e+00)    Acc@1  35.60 ( 3
2.30)   Acc@5  58.40 ( 57.97)
Epoch: [1][ 620/1282]   Time  2.750 ( 1.676)    Data  0.002 ( 0.028)    Loss 2.9053e+00 (3.1604e+00)    Acc@1  37.80 ( 3
2.33)   Acc@5  61.80 ( 58.01)
Epoch: [1][ 630/1282]   Time  2.475 ( 1.676)    Data  0.002 ( 0.028)    Loss 2.9965e+00 (3.1581e+00)    Acc@1  35.60 ( 3
2.37)   Acc@5  61.80 ( 58.05)
Epoch: [1][ 640/1282]   Time  2.678 ( 1.676)    Data  0.002 ( 0.028)    Loss 3.0966e+00 (3.1558e+00)    Acc@1  33.60 ( 3
2.41)   Acc@5  56.80 ( 58.08)
```

```
Epoch: [1][ 650/1282]   Time  2.626 ( 1.676)   Data  0.002 ( 0.027)   Loss 2.9100e+00 (3.1537e+00)   Acc@1  34.80 ( 3
2.44)   Acc@5  63.20 ( 58.12)
Epoch: [1][ 660/1282]   Time  2.349 ( 1.676)   Data  0.003 ( 0.027)   Loss 3.0147e+00 (3.1515e+00)   Acc@1  34.00 ( 3
2.48)   Acc@5  60.40 ( 58.17)
Epoch: [1][ 670/1282]   Time  2.816 ( 1.676)   Data  0.003 ( 0.027)   Loss 3.0517e+00 (3.1489e+00)   Acc@1  34.80 ( 3
2.53)   Acc@5  60.60 ( 58.21)
Epoch: [1][ 680/1282]   Time  2.592 ( 1.676)   Data  0.003 ( 0.026)   Loss 2.9541e+00 (3.1465e+00)   Acc@1  37.00 ( 3
2.57)   Acc@5  64.40 ( 58.25)
Epoch: [1][ 690/1282]   Time  2.867 ( 1.676)   Data  0.002 ( 0.026)   Loss 2.9330e+00 (3.1442e+00)   Acc@1  36.60 ( 3
2.61)   Acc@5  60.60 ( 58.31)
Epoch: [1][ 700/1282]   Time  2.420 ( 1.676)   Data  0.002 ( 0.026)   Loss 3.0042e+00 (3.1418e+00)   Acc@1  33.80 ( 3
2.65)   Acc@5  59.80 ( 58.34)
Epoch: [1][ 710/1282]   Time  2.475 ( 1.675)   Data  0.002 ( 0.025)   Loss 3.0659e+00 (3.1387e+00)   Acc@1  31.60 ( 3
2.70)   Acc@5  60.20 ( 58.40)
Epoch: [1][ 720/1282]   Time  2.511 ( 1.675)   Data  0.003 ( 0.025)   Loss 2.9184e+00 (3.1363e+00)   Acc@1  34.80 ( 3
2.73)   Acc@5  61.80 ( 58.45)
Epoch: [1][ 730/1282]   Time  2.658 ( 1.675)   Data  0.002 ( 0.025)   Loss 2.9163e+00 (3.1335e+00)   Acc@1  35.20 ( 3
2.77)   Acc@5  63.20 ( 58.50)
Epoch: [1][ 740/1282]   Time  2.918 ( 1.675)   Data  0.003 ( 0.024)   Loss 3.0503e+00 (3.1313e+00)   Acc@1  34.40 ( 3
2.81)   Acc@5  60.40 ( 58.53)
Epoch: [1][ 750/1282]   Time  2.537 ( 1.676)   Data  0.003 ( 0.024)   Loss 2.9764e+00 (3.1288e+00)   Acc@1  34.40 ( 3
2.86)   Acc@5  60.20 ( 58.58)
Epoch: [1][ 760/1282]   Time  2.689 ( 1.676)   Data  0.003 ( 0.024)   Loss 3.0750e+00 (3.1256e+00)   Acc@1  34.60 ( 3
2.91)   Acc@5  58.60 ( 58.63)
Epoch: [1][ 770/1282]   Time  2.840 ( 1.676)   Data  0.002 ( 0.023)   Loss 2.7935e+00 (3.1225e+00)   Acc@1  39.80 ( 3
2.95)   Acc@5  63.20 ( 58.69)
Epoch: [1][ 780/1282]   Time  2.697 ( 1.675)   Data  0.002 ( 0.023)   Loss 2.7364e+00 (3.1196e+00)   Acc@1  42.20 ( 3
3.01)   Acc@5  66.20 ( 58.75)
Epoch: [1][ 790/1282]   Time  2.778 ( 1.675)   Data  0.003 ( 0.023)   Loss 2.8273e+00 (3.1169e+00)   Acc@1  36.60 ( 3
3.05)   Acc@5  64.40 ( 58.79)
Epoch: [1][ 800/1282]   Time  2.648 ( 1.675)   Data  0.003 ( 0.023)   Loss 2.8606e+00 (3.1143e+00)   Acc@1  37.40 ( 3
3.09)   Acc@5  63.60 ( 58.84)
Epoch: [1][ 810/1282]   Time  2.624 ( 1.676)   Data  0.002 ( 0.022)   Loss 2.9015e+00 (3.1116e+00)   Acc@1  34.40 ( 3
3.13)   Acc@5  62.40 ( 58.88)
Epoch: [1][ 820/1282]   Time  2.650 ( 1.675)   Data  0.003 ( 0.022)   Loss 2.9289e+00 (3.1089e+00)   Acc@1  35.80 ( 3
3.17)   Acc@5  62.00 ( 58.93)
Epoch: [1][ 830/1282]   Time  2.788 ( 1.675)   Data  0.003 ( 0.022)   Loss 2.9288e+00 (3.1070e+00)   Acc@1  39.20 ( 3
3.21)   Acc@5  63.20 ( 58.96)
Epoch: [1][ 840/1282]   Time  2.671 ( 1.676)   Data  0.002 ( 0.022)   Loss 2.9609e+00 (3.1046e+00)   Acc@1  36.60 ( 3
3.25)   Acc@5  61.40 ( 59.00)
Epoch: [1][ 850/1282]   Time  2.625 ( 1.675)   Data  0.002 ( 0.021)   Loss 2.7431e+00 (3.1018e+00)   Acc@1  41.00 ( 3
3.31)   Acc@5  66.40 ( 59.06)
Epoch: [1][ 860/1282]   Time  2.532 ( 1.675)   Data  0.003 ( 0.021)   Loss 2.9754e+00 (3.0993e+00)   Acc@1  33.80 ( 3
3.36)   Acc@5  62.20 ( 59.10)
Epoch: [1][ 870/1282]   Time  2.660 ( 1.675)   Data  0.003 ( 0.021)   Loss 2.7761e+00 (3.0965e+00)   Acc@1  39.40 ( 3
3.40)   Acc@5  67.00 ( 59.15)
Epoch: [1][ 880/1282]   Time  3.284 ( 1.675)   Data  0.002 ( 0.021)   Loss 2.8696e+00 (3.0945e+00)   Acc@1  38.00 ( 3
3.42)   Acc@5  64.40 ( 59.18)
Epoch: [1][ 890/1282]   Time  2.683 ( 1.676)   Data  0.002 ( 0.021)   Loss 2.9443e+00 (3.0919e+00)   Acc@1  35.60 ( 3
3.46)   Acc@5  60.00 ( 59.23)
Epoch: [1][ 900/1282]   Time  2.806 ( 1.676)   Data  0.003 ( 0.020)   Loss 2.8574e+00 (3.0896e+00)   Acc@1  39.00 ( 3
```

```
3.50)   Acc@5  65.00 ( 59.28)
Epoch: [1][ 910/1282]   Time  2.464 ( 1.676)    Data  0.002 ( 0.020)    Loss 2.7528e+00 (3.0872e+00)    Acc@1  40.60 ( 3
3.54)   Acc@5  67.60 ( 59.32)
Epoch: [1][ 920/1282]   Time  2.634 ( 1.676)    Data  0.003 ( 0.020)    Loss 2.8887e+00 (3.0845e+00)    Acc@1  36.40 ( 3
3.58)   Acc@5  60.80 ( 59.37)
Epoch: [1][ 930/1282]   Time  2.769 ( 1.676)    Data  0.003 ( 0.020)    Loss 2.8925e+00 (3.0823e+00)    Acc@1  36.40 ( 3
3.62)   Acc@5  61.40 ( 59.41)
Epoch: [1][ 940/1282]   Time  2.613 ( 1.675)    Data  0.002 ( 0.020)    Loss 2.9197e+00 (3.0801e+00)    Acc@1  36.40 ( 3
3.65)   Acc@5  61.00 ( 59.46)
Epoch: [1][ 950/1282]   Time  2.542 ( 1.675)    Data  0.003 ( 0.019)    Loss 2.8878e+00 (3.0776e+00)    Acc@1  34.60 ( 3
3.69)   Acc@5  61.80 ( 59.51)
Epoch: [1][ 960/1282]   Time  2.648 ( 1.675)    Data  0.002 ( 0.019)    Loss 2.9306e+00 (3.0752e+00)    Acc@1  36.00 ( 3
3.74)   Acc@5  59.80 ( 59.55)
Epoch: [1][ 970/1282]   Time  2.495 ( 1.675)    Data  0.002 ( 0.019)    Loss 2.8924e+00 (3.0728e+00)    Acc@1  34.00 ( 3
3.78)   Acc@5  62.00 ( 59.60)
Epoch: [1][ 980/1282]   Time  2.587 ( 1.675)    Data  0.002 ( 0.019)    Loss 2.8246e+00 (3.0702e+00)    Acc@1  39.00 ( 3
3.83)   Acc@5  64.00 ( 59.64)
Epoch: [1][ 990/1282]   Time  3.048 ( 1.675)    Data  0.002 ( 0.019)    Loss 2.7851e+00 (3.0674e+00)    Acc@1  38.00 ( 3
3.88)   Acc@5  63.40 ( 59.70)
Epoch: [1][1000/1282]   Time  2.512 ( 1.674)    Data  0.002 ( 0.019)    Loss 2.9208e+00 (3.0656e+00)    Acc@1  37.40 ( 3
3.90)   Acc@5  61.20 ( 59.73)
Epoch: [1][1010/1282]   Time  2.742 ( 1.674)    Data  0.003 ( 0.018)    Loss 2.8769e+00 (3.0634e+00)    Acc@1  35.80 ( 3
3.93)   Acc@5  62.20 ( 59.77)
Epoch: [1][1020/1282]   Time  2.478 ( 1.674)    Data  0.003 ( 0.018)    Loss 2.8871e+00 (3.0617e+00)    Acc@1  35.80 ( 3
3.96)   Acc@5  64.00 ( 59.80)
Epoch: [1][1030/1282]   Time  2.512 ( 1.675)    Data  0.002 ( 0.018)    Loss 2.7990e+00 (3.0591e+00)    Acc@1  39.00 ( 3
4.01)   Acc@5  65.00 ( 59.84)
Epoch: [1][1040/1282]   Time  2.765 ( 1.675)    Data  0.002 ( 0.018)    Loss 2.7864e+00 (3.0570e+00)    Acc@1  36.00 ( 3
4.05)   Acc@5  67.20 ( 59.88)
Epoch: [1][1050/1282]   Time  2.521 ( 1.674)    Data  0.002 ( 0.018)    Loss 2.7155e+00 (3.0544e+00)    Acc@1  40.40 ( 3
4.10)   Acc@5  67.80 ( 59.93)
Epoch: [1][1060/1282]   Time  2.503 ( 1.674)    Data  0.003 ( 0.018)    Loss 2.7283e+00 (3.0518e+00)    Acc@1  39.00 ( 3
4.13)   Acc@5  67.60 ( 59.99)
Epoch: [1][1070/1282]   Time  2.816 ( 1.674)    Data  0.002 ( 0.018)    Loss 2.8005e+00 (3.0500e+00)    Acc@1  37.80 ( 3
4.16)   Acc@5  62.60 ( 60.01)
Epoch: [1][1080/1282]   Time  2.694 ( 1.674)    Data  0.002 ( 0.017)    Loss 2.7506e+00 (3.0478e+00)    Acc@1  42.00 ( 3
4.20)   Acc@5  64.40 ( 60.05)
Epoch: [1][1090/1282]   Time  2.512 ( 1.673)    Data  0.002 ( 0.017)    Loss 2.7566e+00 (3.0456e+00)    Acc@1  38.40 ( 3
4.24)   Acc@5  65.00 ( 60.09)
Epoch: [1][1100/1282]   Time  2.546 ( 1.673)    Data  0.002 ( 0.017)    Loss 2.8414e+00 (3.0434e+00)    Acc@1  36.20 ( 3
4.27)   Acc@5  65.20 ( 60.13)
Epoch: [1][1110/1282]   Time  2.637 ( 1.673)    Data  0.003 ( 0.017)    Loss 2.7252e+00 (3.0418e+00)    Acc@1  36.00 ( 3
4.29)   Acc@5  66.20 ( 60.16)
Epoch: [1][1120/1282]   Time  2.814 ( 1.673)    Data  0.002 ( 0.017)    Loss 2.8529e+00 (3.0396e+00)    Acc@1  34.00 ( 3
4.32)   Acc@5  61.00 ( 60.20)
Epoch: [1][1130/1282]   Time  2.680 ( 1.674)    Data  0.002 ( 0.017)    Loss 2.8477e+00 (3.0378e+00)    Acc@1  37.60 ( 3
4.36)   Acc@5  63.80 ( 60.24)
Epoch: [1][1140/1282]   Time  2.516 ( 1.673)    Data  0.002 ( 0.017)    Loss 2.5762e+00 (3.0356e+00)    Acc@1  43.60 ( 3
4.40)   Acc@5  68.80 ( 60.27)
Epoch: [1][1150/1282]   Time  2.566 ( 1.673)    Data  0.002 ( 0.016)    Loss 2.8064e+00 (3.0331e+00)    Acc@1  39.00 ( 3
4.44)   Acc@5  62.60 ( 60.31)
```

```
Epoch: [1][1160/1282]   Time  2.795 ( 1.673)    Data  0.002 ( 0.016)    Loss 2.9946e+00 (3.0313e+00)    Acc@1  36.80 ( 3
4.46)   Acc@5  61.40 ( 60.34)
Epoch: [1][1170/1282]   Time  2.698 ( 1.673)    Data  0.002 ( 0.016)    Loss 2.9095e+00 (3.0295e+00)    Acc@1  37.80 ( 3
4.50)   Acc@5  61.80 ( 60.37)
Epoch: [1][1180/1282]   Time  2.105 ( 1.673)    Data  0.002 ( 0.017)    Loss 2.8175e+00 (3.0272e+00)    Acc@1  37.60 ( 3
4.54)   Acc@5  63.40 ( 60.42)
Epoch: [1][1190/1282]   Time  2.030 ( 1.672)    Data  0.002 ( 0.019)    Loss 2.8928e+00 (3.0250e+00)    Acc@1  40.20 ( 3
4.58)   Acc@5  62.40 ( 60.45)
Epoch: [1][1200/1282]   Time  2.407 ( 1.672)    Data  0.003 ( 0.021)    Loss 2.8183e+00 (3.0227e+00)    Acc@1  37.60 ( 3
4.62)   Acc@5  64.60 ( 60.49)
Epoch: [1][1210/1282]   Time  2.407 ( 1.672)    Data  0.002 ( 0.022)    Loss 2.9153e+00 (3.0200e+00)    Acc@1  35.00 ( 3
4.66)   Acc@5  63.60 ( 60.54)
Epoch: [1][1220/1282]   Time  2.317 ( 1.672)    Data  0.002 ( 0.022)    Loss 2.9005e+00 (3.0180e+00)    Acc@1  37.80 ( 3
4.71)   Acc@5  62.20 ( 60.57)
Epoch: [1][1230/1282]   Time  2.380 ( 1.673)    Data  0.002 ( 0.022)    Loss 2.7364e+00 (3.0156e+00)    Acc@1  38.80 ( 3
4.74)   Acc@5  66.20 ( 60.62)
Epoch: [1][1240/1282]   Time  2.232 ( 1.672)    Data  0.002 ( 0.021)    Loss 2.7297e+00 (3.0136e+00)    Acc@1  40.80 ( 3
4.77)   Acc@5  66.20 ( 60.65)
Epoch: [1][1250/1282]   Time  2.298 ( 1.672)    Data  0.002 ( 0.021)    Loss 2.5840e+00 (3.0111e+00)    Acc@1  41.20 ( 3
4.81)   Acc@5  68.20 ( 60.70)
Epoch: [1][1260/1282]   Time  2.242 ( 1.672)    Data  0.002 ( 0.021)    Loss 2.8281e+00 (3.0089e+00)    Acc@1  36.40 ( 3
4.85)   Acc@5  65.60 ( 60.74)
Epoch: [1][1270/1282]   Time  2.237 ( 1.672)    Data  0.002 ( 0.021)    Loss 2.7000e+00 (3.0069e+00)    Acc@1  40.80 ( 3
4.88)   Acc@5  64.40 ( 60.78)
Epoch: [1][1280/1282]   Time  2.363 ( 1.672)    Data  0.002 ( 0.021)    Loss 2.5770e+00 (3.0044e+00)    Acc@1  41.40 ( 3
4.93)   Acc@5  69.00 ( 60.82)
Test: [  0/100] Time  4.481 ( 4.481)    Loss 2.2630e+00 (2.2630e+00)    Acc@1  44.20 ( 44.20)   Acc@5  76.40 ( 76.40)
Test: [ 10/100] Time  2.920 ( 1.950)    Loss 2.3561e+00 (2.8683e+00)    Acc@1  49.40 ( 37.80)   Acc@5  72.40 ( 63.98)
Test: [ 20/100] Time  2.768 ( 1.820)    Loss 2.6376e+00 (2.8284e+00)    Acc@1  29.00 ( 35.91)   Acc@5  68.40 ( 64.20)
Test: [ 30/100] Time  2.896 ( 1.774)    Loss 2.5615e+00 (2.7674e+00)    Acc@1  41.40 ( 35.83)   Acc@5  67.60 ( 65.27)
Test: [ 40/100] Time  2.577 ( 1.759)    Loss 2.7604e+00 (2.7731e+00)    Acc@1  39.40 ( 36.46)   Acc@5  66.40 ( 65.25)
Test: [ 50/100] Time  2.842 ( 1.749)    Loss 3.8991e+00 (2.9221e+00)    Acc@1  22.20 ( 34.74)   Acc@5  42.40 ( 62.52)
Test: [ 60/100] Time  2.803 ( 1.740)    Loss 3.4073e+00 (2.9845e+00)    Acc@1  31.60 ( 34.23)   Acc@5  54.00 ( 61.42)
Test: [ 70/100] Time  2.888 ( 1.724)    Loss 3.5829e+00 (3.0476e+00)    Acc@1  26.60 ( 33.56)   Acc@5  52.60 ( 60.39)
Test: [ 80/100] Time  2.903 ( 1.717)    Loss 3.1231e+00 (3.0971e+00)    Acc@1  38.20 ( 33.20)   Acc@5  57.20 ( 59.49)
Test: [ 90/100] Time  2.893 ( 1.709)    Loss 3.4361e+00 (3.1464e+00)    Acc@1  31.80 ( 32.51)   Acc@5  53.40 ( 58.62)
 * Acc@1 33.104 Acc@5 59.296
lr: 0.0
```

In [44]:

```
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 [ ]: