

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: PyYAML in /opt/conda/lib/python3.8/site-packages (from wandb) (5.4.1)
Requirement already satisfied: python-dateutil>=2.6.1 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.8.2)
Requirement already satisfied: configparser>=3.8.1 in /opt/conda/lib/python3.8/site-packages (from wandb) (5.0.2)
Requirement already satisfied: protobuf>=3.12.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (3.17.3)
Requirement already satisfied: promise<3,>=2.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.3)
Requirement already satisfied: docker-pycreds>=0.4.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (0.4.0)
Requirement already satisfied: six>=1.13.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (1.16.0)
Requirement already satisfied: pathtools in /opt/conda/lib/python3.8/site-packages (from wandb) (0.1.2)
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: yaspin>=1.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.1.0)
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: GitPython>=1.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (3.1.24)
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: requests<3,>=2.0.0 in /opt/conda/lib/python3.8/site-packages (from wandb) (2.26.0)
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: idna<4,>=2.5 in /opt/conda/lib/python3.8/site-packages (from requests<3,>=2.0.0->wandb) (3.1)
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: 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: 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: 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()
```

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

In [6]:

```
wandb.init(project="w251-hw9", entity="rochelleli")
```

Syncing run **trembling-possession-5** to **Weights & Biases** ([docs](#)).

Out[6]: Display W&B run

In [7]:

```
SEED=1

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:pscwd9uw) before initializing another...

Waiting for W&B process to finish, PID 698... **(success).**

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

Synced **trembling-possession-5**: <https://wandb.ai/rochelleli/w251-hw9/runs/pscwd9uw>

Find logs at: ./wandb/run-20211031_103248-pscwd9uw/logs

Successfully finished last run (ID:pscwd9uw). Initializing new run:

Syncing run **masked-cauldron-1** to [Weights & Biases](#) (docs).

Out[12]:

```
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 = 0  
  
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 31.765 (31.765)	Data 3.818 (3.818)	Loss 7.0052e+00 (7.0052e+00)	Acc@1 0.00 (
0.00) Acc@5 1.00 (1.00)				
Epoch: [0][10/1282]	Time 0.785 (3.937)	Data 0.096 (0.725)	Loss 6.9195e+00 (6.9718e+00)	Acc@1 0.60 (
0.18) Acc@5 1.20 (0.78)				
Epoch: [0][20/1282]	Time 0.754 (2.881)	Data 0.097 (0.677)	Loss 6.9420e+00 (6.9584e+00)	Acc@1 0.20 (
0.18) Acc@5 0.80 (0.81)				
Epoch: [0][30/1282]	Time 0.763 (2.512)	Data 0.104 (0.573)	Loss 6.8143e+00 (6.9239e+00)	Acc@1 0.80 (
0.26) Acc@5 2.00 (1.10)				
Epoch: [0][40/1282]	Time 0.758 (2.296)	Data 0.096 (0.445)	Loss 6.7460e+00 (6.8903e+00)	Acc@1 0.40 (
0.37) Acc@5 3.20 (1.39)				
Epoch: [0][50/1282]	Time 0.666 (2.180)	Data 0.002 (0.380)	Loss 6.7190e+00 (6.8485e+00)	Acc@1 0.20 (
0.42) Acc@5 2.00 (1.64)				
Epoch: [0][60/1282]	Time 0.695 (2.114)	Data 0.002 (0.321)	Loss 6.5468e+00 (6.8072e+00)	Acc@1 1.20 (
0.52) Acc@5 4.00 (1.86)				
Epoch: [0][70/1282]	Time 0.755 (2.048)	Data 0.003 (0.276)	Loss 6.4231e+00 (6.7601e+00)	Acc@1 1.20 (
0.62) Acc@5 3.40 (2.21)				
Epoch: [0][80/1282]	Time 0.874 (2.009)	Data 0.002 (0.243)	Loss 6.3648e+00 (6.7166e+00)	Acc@1 0.80 (
0.69) Acc@5 4.60 (2.48)				
Epoch: [0][90/1282]	Time 1.269 (1.968)	Data 0.003 (0.216)	Loss 6.4328e+00 (6.6766e+00)	Acc@1 1.20 (
0.80) Acc@5 4.00 (2.73)				
Epoch: [0][100/1282]	Time 1.648 (1.944)	Data 0.003 (0.195)	Loss 6.1733e+00 (6.6321e+00)	Acc@1 2.40 (
0.89) Acc@5 6.80 (3.07)				
Epoch: [0][110/1282]	Time 1.124 (1.914)	Data 0.002 (0.178)	Loss 6.1747e+00 (6.5914e+00)	Acc@1 0.80 (
0.97) Acc@5 7.40 (3.39)				
Epoch: [0][120/1282]	Time 1.238 (1.900)	Data 0.003 (0.163)	Loss 6.1986e+00 (6.5519e+00)	Acc@1 3.00 (
1.07) Acc@5 6.60 (3.69)				
Epoch: [0][130/1282]	Time 1.232 (1.884)	Data 0.002 (0.151)	Loss 6.0501e+00 (6.5186e+00)	Acc@1 2.80 (
1.15) Acc@5 7.20 (3.94)				
Epoch: [0][140/1282]	Time 1.002 (1.865)	Data 0.003 (0.140)	Loss 5.8912e+00 (6.4834e+00)	Acc@1 1.40 (
1.21) Acc@5 8.60 (4.16)				
Epoch: [0][150/1282]	Time 1.012 (1.859)	Data 0.003 (0.131)	Loss 5.8954e+00 (6.4457e+00)	Acc@1 4.60 (
1.33) Acc@5 10.20 (4.52)				
Epoch: [0][160/1282]	Time 0.772 (1.845)	Data 0.003 (0.123)	Loss 5.8703e+00 (6.4114e+00)	Acc@1 2.40 (
1.42) Acc@5 9.40 (4.82)				
Epoch: [0][170/1282]	Time 0.772 (1.833)	Data 0.002 (0.116)	Loss 5.8273e+00 (6.3786e+00)	Acc@1 3.00 (
1.51) Acc@5 9.40 (5.10)				
Epoch: [0][180/1282]	Time 0.775 (1.825)	Data 0.002 (0.110)	Loss 5.7143e+00 (6.3457e+00)	Acc@1 4.20 (
1.63) Acc@5 11.20 (5.45)				
Epoch: [0][190/1282]	Time 0.688 (1.816)	Data 0.002 (0.105)	Loss 5.7807e+00 (6.3160e+00)	Acc@1 3.40 (
1.73) Acc@5 10.80 (5.75)				
Epoch: [0][200/1282]	Time 0.688 (1.815)	Data 0.002 (0.099)	Loss 5.6943e+00 (6.2852e+00)	Acc@1 4.80 (
1.84) Acc@5 12.00 (6.06)				
Epoch: [0][210/1282]	Time 0.690 (1.806)	Data 0.003 (0.095)	Loss 5.6501e+00 (6.2572e+00)	Acc@1 6.00 (

1.95)	Acc@5	12.60 (6.35)									
Epoch: [0][220/1282]	Time	0.691 (1.801)	Data	0.002 (0.091)	Loss	5.6773e+00 (6.2283e+00)	Acc@1	3.20 (2.04)			
2.04)	Acc@5	13.60 (6.64)									
Epoch: [0][230/1282]	Time	0.693 (1.795)	Data	0.003 (0.087)	Loss	5.6111e+00 (6.2014e+00)	Acc@1	3.80 (2.14)			
2.14)	Acc@5	15.00 (6.94)									
Epoch: [0][240/1282]	Time	0.697 (1.789)	Data	0.003 (0.084)	Loss	5.6436e+00 (6.1756e+00)	Acc@1	4.80 (2.23)			
2.23)	Acc@5	12.60 (7.22)									
Epoch: [0][250/1282]	Time	0.694 (1.784)	Data	0.002 (0.081)	Loss	5.4127e+00 (6.1477e+00)	Acc@1	5.20 (2.35)			
2.35)	Acc@5	15.20 (7.52)									
Epoch: [0][260/1282]	Time	0.714 (1.782)	Data	0.002 (0.078)	Loss	5.4833e+00 (6.1210e+00)	Acc@1	5.40 (2.45)			
2.45)	Acc@5	14.20 (7.82)									
Epoch: [0][270/1282]	Time	0.725 (1.777)	Data	0.002 (0.075)	Loss	5.4789e+00 (6.0954e+00)	Acc@1	4.20 (2.55)			
2.55)	Acc@5	13.80 (8.12)									
Epoch: [0][280/1282]	Time	0.682 (1.774)	Data	0.002 (0.072)	Loss	5.5131e+00 (6.0705e+00)	Acc@1	4.00 (2.65)			
2.65)	Acc@5	15.80 (8.42)									
Epoch: [0][290/1282]	Time	0.734 (1.774)	Data	0.003 (0.070)	Loss	5.2685e+00 (6.0456e+00)	Acc@1	6.00 (2.75)			
2.75)	Acc@5	17.00 (8.71)									
Epoch: [0][300/1282]	Time	0.721 (1.768)	Data	0.003 (0.068)	Loss	5.1495e+00 (6.0214e+00)	Acc@1	9.00 (2.88)			
2.88)	Acc@5	21.80 (9.03)									
Epoch: [0][310/1282]	Time	0.732 (1.768)	Data	0.003 (0.066)	Loss	5.1729e+00 (5.9975e+00)	Acc@1	5.20 (2.98)			
2.98)	Acc@5	20.20 (9.32)									
Epoch: [0][320/1282]	Time	0.718 (1.765)	Data	0.003 (0.064)	Loss	5.1785e+00 (5.9741e+00)	Acc@1	6.60 (3.10)			
3.10)	Acc@5	18.20 (9.63)									
Epoch: [0][330/1282]	Time	0.715 (1.762)	Data	0.003 (0.062)	Loss	5.0783e+00 (5.9511e+00)	Acc@1	7.80 (3.22)			
3.22)	Acc@5	20.20 (9.93)									
Epoch: [0][340/1282]	Time	0.683 (1.761)	Data	0.003 (0.060)	Loss	5.2627e+00 (5.9287e+00)	Acc@1	4.80 (3.34)			
3.34)	Acc@5	17.80 (10.23)									
Epoch: [0][350/1282]	Time	0.682 (1.758)	Data	0.002 (0.058)	Loss	5.2318e+00 (5.9082e+00)	Acc@1	7.00 (3.45)			
3.45)	Acc@5	18.40 (10.49)									
Epoch: [0][360/1282]	Time	0.728 (1.758)	Data	0.002 (0.057)	Loss	5.1501e+00 (5.8877e+00)	Acc@1	8.60 (3.55)			
3.55)	Acc@5	22.20 (10.77)									
Epoch: [0][370/1282]	Time	0.696 (1.754)	Data	0.003 (0.055)	Loss	5.0797e+00 (5.8679e+00)	Acc@1	7.20 (3.66)			
3.66)	Acc@5	21.40 (11.03)									
Epoch: [0][380/1282]	Time	0.689 (1.753)	Data	0.002 (0.054)	Loss	5.0661e+00 (5.8483e+00)	Acc@1	9.00 (3.76)			
3.76)	Acc@5	22.60 (11.31)									
Epoch: [0][390/1282]	Time	0.679 (1.752)	Data	0.002 (0.053)	Loss	5.0788e+00 (5.8295e+00)	Acc@1	9.40 (3.86)			
3.86)	Acc@5	23.00 (11.55)									
Epoch: [0][400/1282]	Time	0.711 (1.748)	Data	0.003 (0.051)	Loss	5.0956e+00 (5.8094e+00)	Acc@1	7.20 (3.95)			
3.95)	Acc@5	22.40 (11.84)									
Epoch: [0][410/1282]	Time	0.786 (1.745)	Data	0.103 (0.05							

Epoch: [0][470/1282]	Time 0.689 (1.740)	Data 0.002 (0.045)	Loss 4.8521e+00 (5.6786e+00)	Acc@1 9.20 (
4.75) Acc@5 25.00 (13.71)				
Epoch: [0][480/1282]	Time 0.716 (1.741)	Data 0.003 (0.044)	Loss 4.7670e+00 (5.6600e+00)	Acc@1 13.20 (
4.88) Acc@5 25.80 (13.97)				
Epoch: [0][490/1282]	Time 0.690 (1.739)	Data 0.002 (0.043)	Loss 4.7965e+00 (5.6420e+00)	Acc@1 13.00 (
5.00) Acc@5 25.00 (14.23)				
Epoch: [0][500/1282]	Time 0.682 (1.737)	Data 0.002 (0.043)	Loss 4.8291e+00 (5.6245e+00)	Acc@1 10.00 (
5.12) Acc@5 25.80 (14.49)				
Epoch: [0][510/1282]	Time 0.726 (1.737)	Data 0.003 (0.043)	Loss 4.8475e+00 (5.6075e+00)	Acc@1 11.20 (
5.23) Acc@5 26.40 (14.75)				
Epoch: [0][520/1282]	Time 0.687 (1.735)	Data 0.003 (0.042)	Loss 4.7205e+00 (5.5908e+00)	Acc@1 11.60 (
5.34) Acc@5 27.20 (15.00)				
Epoch: [0][530/1282]	Time 0.691 (1.736)	Data 0.002 (0.041)	Loss 4.5390e+00 (5.5743e+00)	Acc@1 11.80 (
5.47) Acc@5 30.20 (15.25)				
Epoch: [0][540/1282]	Time 0.685 (1.736)	Data 0.002 (0.040)	Loss 4.6934e+00 (5.5584e+00)	Acc@1 12.00 (
5.58) Acc@5 29.20 (15.48)				
Epoch: [0][550/1282]	Time 0.690 (1.735)	Data 0.002 (0.040)	Loss 4.7437e+00 (5.5425e+00)	Acc@1 12.20 (
5.69) Acc@5 27.20 (15.72)				
Epoch: [0][560/1282]	Time 0.684 (1.734)	Data 0.002 (0.039)	Loss 4.6569e+00 (5.5262e+00)	Acc@1 9.60 (
5.81) Acc@5 29.60 (15.97)				
Epoch: [0][570/1282]	Time 0.717 (1.733)	Data 0.003 (0.038)	Loss 4.5330e+00 (5.5102e+00)	Acc@1 11.60 (
5.92) Acc@5 30.40 (16.23)				
Epoch: [0][580/1282]	Time 0.724 (1.733)	Data 0.002 (0.038)	Loss 4.6726e+00 (5.4946e+00)	Acc@1 14.20 (
6.04) Acc@5 29.20 (16.47)				
Epoch: [0][590/1282]	Time 0.688 (1.732)	Data 0.003 (0.037)	Loss 4.6111e+00 (5.4786e+00)	Acc@1 11.80 (
6.15) Acc@5 30.80 (16.74)				
Epoch: [0][600/1282]	Time 0.680 (1.730)	Data 0.002 (0.037)	Loss 4.6452e+00 (5.4635e+00)	Acc@1 10.80 (
6.25) Acc@5 29.80 (16.98)				
Epoch: [0][610/1282]	Time 0.680 (1.730)	Data 0.003 (0.036)	Loss 4.5628e+00 (5.4475e+00)	Acc@1 11.40 (
6.37) Acc@5 33.00 (17.24)				
Epoch: [0][620/1282]	Time 0.780 (1.729)	Data 0.099 (0.036)	Loss 4.4533e+00 (5.4322e+00)	Acc@1 14.80 (
6.49) Acc@5 36.00 (17.49)				
Epoch: [0][630/1282]	Time 0.695 (1.729)	Data 0.002 (0.037)	Loss 4.4817e+00 (5.4168e+00)	Acc@1 15.40 (
6.62) Acc@5 34.60 (17.74)				
Epoch: [0][640/1282]	Time 0.686 (1.727)	Data 0.002 (0.036)	Loss 4.5289e+00 (5.4026e+00)	Acc@1 13.40 (
6.73) Acc@5 33.80 (17.98)				
Epoch: [0][650/1282]	Time 0.684 (1.727)	Data 0.002 (0.036)	Loss 4.4358e+00 (5.3879e+00)	Acc@1 14.40 (
6.84) Acc@5 34.20 (18.22)				
Epoch: [0][660/1282]	Time 0.685 (1.726)	Data 0.002 (0.035)	Loss 4.5140e+00 (5.3730e+00)	Acc@1 11.00 (
6.96) Acc@5 32.20 (18.46)				
Epoch: [0][670/1282]	Time 0.782 (1.725)	Data 0.100 (0.035)	Loss 4.4061e+00 (5.3590e+00)	Acc@1 12.60 (
7.06) Acc@5 31.60 (18.67)				
Epoch: [0][680/1282]	Time 0.684 (1.725)	Data 0.003 (0.035)	Loss 4.4132e+00 (5.3456e+00)	Acc@1 15.80 (
7.17) Acc@5 35.80 (18.89)				
Epoch: [0][690/1282]	Time 0.783 (1.724)	Data 0.099 (0.034)	Loss 4.3678e+00 (5.3311e+00)	Acc@1 13.40 (
7.28) Acc@5 35.60 (19.12)				
Epoch: [0][700/1282]	Time 0.682 (1.724)	Data 0.002 (0.034)	Loss 4.4067e+00 (5.3169e+00)	Acc@1 15.40 (
7.40) Acc@5 34.60 (19.35)				
Epoch: [0][710/1282]	Time 0.686 (1.723)	Data 0.003 (0.034)	Loss 4.2924e+00 (5.3030e+00)	Acc@1 19.20 (
7.51) Acc@5 36.00 (19.58)				
Epoch: [0][720/1282]	Time 0.778 (1.723)	Data 0.098 (0.034)	Loss 4.2623e+00 (5.2889e+00)	Acc@1 17.80 (

[illegible]

Epoch: [0][980/1282]	Time 0.778 (1.711)	Data 0.097 (0.040)	Loss 4.0263e+00 (4.9735e+00)	Acc@1 18.60 (1
0.48) Acc@5 40.80 (25.15)				
Epoch: [0][990/1282]	Time 0.794 (1.711)	Data 0.100 (0.041)	Loss 3.8458e+00 (4.9630e+00)	Acc@1 23.60 (1
0.58) Acc@5 44.40 (25.33)				
Epoch: [0][1000/1282]	Time 0.727 (1.711)	Data 0.003 (0.041)	Loss 4.0544e+00 (4.9525e+00)	Acc@1 19.80 (1
0.69) Acc@5 42.00 (25.51)				
Epoch: [0][1010/1282]	Time 0.779 (1.711)	Data 0.101 (0.040)	Loss 3.9197e+00 (4.9416e+00)	Acc@1 19.00 (1
0.79) Acc@5 45.80 (25.71)				
Epoch: [0][1020/1282]	Time 0.880 (1.710)	Data 0.202 (0.041)	Loss 3.8974e+00 (4.9310e+00)	Acc@1 20.60 (1
0.90) Acc@5 43.40 (25.89)				
Epoch: [0][1030/1282]	Time 0.679 (1.710)	Data 0.002 (0.041)	Loss 3.9300e+00 (4.9206e+00)	Acc@1 19.60 (1
1.01) Acc@5 42.20 (26.07)				
Epoch: [0][1040/1282]	Time 0.882 (1.709)	Data 0.202 (0.041)	Loss 4.0159e+00 (4.9101e+00)	Acc@1 20.20 (1
1.12) Acc@5 41.20 (26.26)				
Epoch: [0][1050/1282]	Time 0.781 (1.710)	Data 0.100 (0.042)	Loss 4.0016e+00 (4.9001e+00)	Acc@1 19.00 (1
1.22) Acc@5 42.60 (26.43)				
Epoch: [0][1060/1282]	Time 0.680 (1.709)	Data 0.003 (0.042)	Loss 3.6650e+00 (4.8898e+00)	Acc@1 23.40 (1
1.32) Acc@5 48.60 (26.62)				
Epoch: [0][1070/1282]	Time 0.699 (1.710)	Data 0.002 (0.042)	Loss 3.8824e+00 (4.8798e+00)	Acc@1 17.80 (1
1.42) Acc@5 42.20 (26.78)				
Epoch: [0][1080/1282]	Time 0.685 (1.710)	Data 0.003 (0.041)	Loss 3.7351e+00 (4.8698e+00)	Acc@1 23.00 (1
1.52) Acc@5 48.80 (26.97)				
Epoch: [0][1090/1282]	Time 0.718 (1.709)	Data 0.003 (0.041)	Loss 3.6449e+00 (4.8597e+00)	Acc@1 25.60 (1
1.62) Acc@5 50.40 (27.14)				
Epoch: [0][1100/1282]	Time 0.717 (1.708)	Data 0.003 (0.041)	Loss 3.6631e+00 (4.8496e+00)	Acc@1 22.80 (1
1.73) Acc@5 49.80 (27.32)				
Epoch: [0][1110/1282]	Time 0.710 (1.708)	Data 0.003 (0.040)	Loss 3.8230e+00 (4.8394e+00)	Acc@1 21.40 (1
1.83) Acc@5 43.00 (27.49)				
Epoch: [0][1120/1282]	Time 0.696 (1.708)	Data 0.002 (0.040)	Loss 3.7044e+00 (4.8293e+00)	Acc@1 24.00 (1
1.94) Acc@5 48.20 (27.67)				
Epoch: [0][1130/1282]	Time 0.777 (1.707)	Data 0.095 (0.040)	Loss 3.6064e+00 (4.8194e+00)	Acc@1 26.20 (1
2.04) Acc@5 49.20 (27.85)				
Epoch: [0][1140/1282]	Time 0.794 (1.707)	Data 0.097 (0.040)	Loss 3.7460e+00 (4.8095e+00)	Acc@1 22.80 (1
2.15) Acc@5 45.80 (28.03)				
Epoch: [0][1150/1282]	Time 0.785 (1.706)	Data 0.100 (0.040)	Loss 3.6300e+00 (4.7999e+00)	Acc@1 24.40 (1
2.24) Acc@5 47.60 (28.19)				
Epoch: [0][1160/1282]	Time 0.782 (1.706)	Data 0.100 (0.040)	Loss 3.4967e+00 (4.7898e+00)	Acc@1 25.40 (1
2.35) Acc@5 52.40 (28.37)				
Epoch: [0][1170/1282]	Time 0.780 (1.705)	Data 0.098 (0.041)	Loss 3.5354e+00 (4.7797e+00)	Acc@1 27.60 (1
2.46) Acc@5 50.00 (28.55)				
Epoch: [0][1180/1282]	Time 0.680 (1.705)	Data 0.003 (0.041)	Loss 3.6625e+00 (4.7700e+00)	Acc@1 24.00 (1
2.55) Acc@5 49.40 (28.72)				
Epoch: [0][1190/1282]	Time 0.692 (1.705)	Data 0.003 (0.041)	Loss 3.6948e+00 (4.7603e+00)	Acc@1 21.80 (1
2.66) Acc@5 49.80 (28.89)				
Epoch: [0][1200/1282]	Time 0.780 (1.704)	Data 0.097 (0.041)	Loss 3.4662e+00 (4.7511e+00)	Acc@1 30.20 (1
2.77) Acc@5 50.60 (29.06)				
Epoch: [0][1210/1282]	Time 0.687 (1.704)	Data 0.003 (0.042)	Loss 3.6185e+00 (4.7416e+00)	Acc@1 25.80 (1
2.87) Acc@5 48.00 (29.23)				
Epoch: [0][1220/1282]	Time 0.689 (1.704)	Data 0.002 (0.042)	Loss 3.5822e+00 (4.7323e+00)	Acc@1 25.80 (1
2.98) Acc@5 48.80 (29.41)				
Epoch: [0][1230/1282]	Time 0.716 (1.704)	Data 0.003 (0.042)	Loss 3.5782e+00 (4.7228e+00)	Acc@1 25.40 (1

```

3.08) Acc@5 50.20 ( 29.57)
Epoch: [0][1240/1282] Time 0.714 ( 1.704) Data 0.003 ( 0.041) Loss 3.7009e+00 (4.7135e+00) Acc@1 22.40 ( 1
3.18) Acc@5 48.20 ( 29.74)
Epoch: [0][1250/1282] Time 0.724 ( 1.703) Data 0.003 ( 0.041) Loss 3.6202e+00 (4.7043e+00) Acc@1 26.60 ( 1
3.28) Acc@5 49.60 ( 29.90)
Epoch: [0][1260/1282] Time 0.721 ( 1.704) Data 0.003 ( 0.041) Loss 3.4213e+00 (4.6952e+00) Acc@1 28.80 ( 1
3.37) Acc@5 54.00 ( 30.06)
Epoch: [0][1270/1282] Time 0.785 ( 1.703) Data 0.103 ( 0.041) Loss 3.3892e+00 (4.6860e+00) Acc@1 27.20 ( 1
3.47) Acc@5 55.20 ( 30.23)
Epoch: [0][1280/1282] Time 0.781 ( 1.702) Data 0.099 ( 0.041) Loss 3.6227e+00 (4.6771e+00) Acc@1 25.60 ( 1
3.57) Acc@5 48.00 ( 30.39)
Test: [ 0/100] Time 6.569 ( 6.569) Loss 3.5146e+00 (3.5146e+00) Acc@1 23.80 ( 23.80) Acc@5 54.40 ( 54.40)
Test: [ 10/100] Time 2.998 ( 2.001) Loss 4.0827e+00 (4.2301e+00) Acc@1 24.80 ( 20.73) Acc@5 44.00 ( 42.49)
Test: [ 20/100] Time 2.869 ( 1.860) Loss 4.0490e+00 (4.0471e+00) Acc@1 10.60 ( 20.67) Acc@5 39.40 ( 43.43)
Test: [ 30/100] Time 2.946 ( 1.817) Loss 4.2780e+00 (4.0942e+00) Acc@1 20.80 ( 19.25) Acc@5 42.60 ( 42.21)
Test: [ 40/100] Time 2.730 ( 1.805) Loss 3.5963e+00 (4.1573e+00) Acc@1 25.00 ( 19.18) Acc@5 52.20 ( 41.36)
Test: [ 50/100] Time 2.833 ( 1.791) Loss 5.1530e+00 (4.1964e+00) Acc@1 9.00 ( 18.33) Acc@5 22.40 ( 40.25)
Test: [ 60/100] Time 2.908 ( 1.784) Loss 4.0325e+00 (4.1941e+00) Acc@1 24.00 ( 18.45) Acc@5 42.20 ( 40.06)
Test: [ 70/100] Time 2.946 ( 1.768) Loss 4.2801e+00 (4.2172e+00) Acc@1 13.80 ( 18.17) Acc@5 37.00 ( 39.63)
Test: [ 80/100] Time 2.993 ( 1.757) Loss 3.9314e+00 (4.2321e+00) Acc@1 23.20 ( 18.11) Acc@5 43.40 ( 39.18)
Test: [ 90/100] Time 2.995 ( 1.747) Loss 4.3866e+00 (4.2455e+00) Acc@1 16.60 ( 17.87) Acc@5 34.40 ( 38.74)
* Acc@1 18.470 Acc@5 39.594
lr: 0.05
Epoch: [1][ 0/1282] Time 4.436 ( 4.436) Data 3.721 ( 3.721) Loss 3.5231e+00 (3.5231e+00) Acc@1 27.00 ( 2
7.00) Acc@5 52.00 ( 52.00)
Epoch: [1][ 10/1282] Time 2.661 ( 1.920) Data 1.944 ( 1.215) Loss 3.6332e+00 (3.5124e+00) Acc@1 26.20 ( 2
7.33) Acc@5 48.60 ( 51.07)
Epoch: [1][ 20/1282] Time 2.508 ( 1.780) Data 1.794 ( 1.078) Loss 3.3309e+00 (3.4520e+00) Acc@1 30.00 ( 2
7.88) Acc@5 55.00 ( 52.56)
Epoch: [1][ 30/1282] Time 2.763 ( 1.749) Data 2.043 ( 1.048) Loss 3.3284e+00 (3.4312e+00) Acc@1 29.60 ( 2
8.29) Acc@5 53.00 ( 53.14)
Epoch: [1][ 40/1282] Time 2.652 ( 1.732) Data 1.906 ( 1.031) Loss 3.3980e+00 (3.4232e+00) Acc@1 30.60 ( 2
8.46) Acc@5 55.00 ( 53.27)
Epoch: [1][ 50/1282] Time 2.239 ( 1.710) Data 1.562 ( 1.011) Loss 3.4100e+00 (3.4054e+00) Acc@1 30.80 ( 2
8.77) Acc@5 55.20 ( 53.44)
Epoch: [1][ 60/1282] Time 1.500 ( 1.697) Data 0.822 ( 1.000) Loss 3.3353e+00 (3.3877e+00) Acc@1 32.20 ( 2
8.99) Acc@5 56.40 ( 53.85)
Epoch: [1][ 70/1282] Time 1.500 ( 1.689) Data 0.822 ( 0.995) Loss 3.3561e+00 (3.3808e+00) Acc@1 29.80 ( 2
9.03) Acc@5 55.20 ( 54.06)
Epoch: [1][ 80/1282] Time 0.778 ( 1.678) Data 0.099 ( 0.985) Loss 3.3317e+00 (3.3750e+00) Acc@1 30.00 ( 2
9.09) Acc@5 52.60 ( 54.22)
Epoch: [1][ 90/1282] Time 0.780 ( 1.677) Data 0.101 ( 0.984) Loss 3.1071e+00 (3.3651e+00) Acc@1 32.40 ( 2
9.21) Acc@5 58.60 ( 54.40)
Epoch: [1][ 100/1282] Time 0.780 ( 1.676) Data 0.100 ( 0.982) Loss 3.3688e+00 (3.3565e+00) Acc@1 29.40 ( 2
9.35) Acc@5 53.00 ( 54.57)
Epoch: [1][ 110/1282] Time 0.685 ( 1.675) Data 0.002 ( 0.980) Loss 3.1380e+00 (3.3518e+00) Acc@1 34.80 ( 2
9.44) Acc@5 58.80 ( 54.64)
Epoch: [1][ 120/1282] Time 0.684 ( 1.670) Data 0.002 ( 0.975) Loss 3.3781e+00 (3.3457e+00) Acc@1 29.60 ( 2
9.54) Acc@5 53.60 ( 54.77)
Epoch: [1][ 130/1282] Time 0.832 ( 1.669) Data 0.155 ( 0.974) Loss 3.2498e+00 (3.3418e+00) Acc@1 35.20 ( 2
9.61) Acc@5 58.00 ( 54.82)

```

Epoch: [1][140/1282]	Time 0.777 (1.670)	Data 0.101 (0.974)	Loss 3.1167e+00 (3.3316e+00)	Acc@1 31.40 (2
9.75) Acc@5 58.00 (54.96)				
Epoch: [1][150/1282]	Time 0.682 (1.671)	Data 0.003 (0.974)	Loss 3.2931e+00 (3.3233e+00)	Acc@1 29.60 (2
9.89) Acc@5 55.40 (55.15)				
Epoch: [1][160/1282]	Time 0.684 (1.673)	Data 0.003 (0.976)	Loss 3.1277e+00 (3.3181e+00)	Acc@1 32.00 (2
9.94) Acc@5 58.40 (55.27)				
Epoch: [1][170/1282]	Time 0.678 (1.672)	Data 0.003 (0.975)	Loss 3.3448e+00 (3.3120e+00)	Acc@1 28.60 (3
0.02) Acc@5 54.00 (55.36)				
Epoch: [1][180/1282]	Time 0.683 (1.670)	Data 0.002 (0.973)	Loss 3.1771e+00 (3.3065e+00)	Acc@1 32.60 (3
0.15) Acc@5 55.60 (55.45)				
Epoch: [1][190/1282]	Time 0.907 (1.669)	Data 0.228 (0.972)	Loss 3.2173e+00 (3.3029e+00)	Acc@1 30.80 (3
0.17) Acc@5 57.20 (55.51)				
Epoch: [1][200/1282]	Time 1.179 (1.669)	Data 0.503 (0.972)	Loss 3.1715e+00 (3.2974e+00)	Acc@1 30.60 (3
0.27) Acc@5 56.60 (55.62)				
Epoch: [1][210/1282]	Time 1.285 (1.668)	Data 0.604 (0.972)	Loss 3.2809e+00 (3.2934e+00)	Acc@1 29.60 (3
0.33) Acc@5 55.80 (55.71)				
Epoch: [1][220/1282]	Time 1.653 (1.668)	Data 0.976 (0.973)	Loss 3.3392e+00 (3.2897e+00)	Acc@1 28.60 (3
0.38) Acc@5 54.40 (55.79)				
Epoch: [1][230/1282]	Time 1.625 (1.668)	Data 0.946 (0.973)	Loss 3.3343e+00 (3.2864e+00)	Acc@1 29.00 (3
0.44) Acc@5 56.80 (55.85)				
Epoch: [1][240/1282]	Time 2.014 (1.669)	Data 1.333 (0.975)	Loss 3.3462e+00 (3.2840e+00)	Acc@1 28.80 (3
0.47) Acc@5 52.60 (55.84)				
Epoch: [1][250/1282]	Time 2.283 (1.670)	Data 1.602 (0.977)	Loss 3.1500e+00 (3.2788e+00)	Acc@1 30.80 (3
0.55) Acc@5 58.00 (55.92)				
Epoch: [1][260/1282]	Time 1.959 (1.670)	Data 1.278 (0.977)	Loss 3.1291e+00 (3.2738e+00)	Acc@1 32.80 (3
0.62) Acc@5 57.40 (56.00)				
Epoch: [1][270/1282]	Time 1.807 (1.670)	Data 1.125 (0.977)	Loss 3.1749e+00 (3.2698e+00)	Acc@1 33.20 (3
0.70) Acc@5 58.20 (56.09)				
Epoch: [1][280/1282]	Time 1.857 (1.669)	Data 1.174 (0.977)	Loss 3.3133e+00 (3.2653e+00)	Acc@1 29.60 (3
0.78) Acc@5 54.20 (56.18)				
Epoch: [1][290/1282]	Time 2.455 (1.670)	Data 1.772 (0.978)	Loss 3.2941e+00 (3.2613e+00)	Acc@1 29.60 (3
0.82) Acc@5 54.40 (56.23)				
Epoch: [1][300/1282]	Time 2.264 (1.669)	Data 1.587 (0.976)	Loss 3.3094e+00 (3.2599e+00)	Acc@1 30.60 (3
0.84) Acc@5 52.60 (56.21)				
Epoch: [1][310/1282]	Time 2.420 (1.670)	Data 1.721 (0.977)	Loss 2.9646e+00 (3.2563e+00)	Acc@1 34.80 (3
0.89) Acc@5 61.20 (56.27)				
Epoch: [1][320/1282]	Time 2.797 (1.671)	Data 2.060 (0.978)	Loss 3.1252e+00 (3.2546e+00)	Acc@1 33.80 (3
0.93) Acc@5 60.60 (56.30)				
Epoch: [1][330/1282]	Time 2.498 (1.670)	Data 1.814 (0.977)	Loss 3.1207e+00 (3.2519e+00)	Acc@1 32.80 (3
0.97) Acc@5 58.20 (56.35)				
Epoch: [1][340/1282]	Time 2.610 (1.670)	Data 1.893 (0.977)	Loss 3.1021e+00 (3.2473e+00)	Acc@1 34.00 (3
1.05) Acc@5 58.80 (56.43)				
Epoch: [1][350/1282]	Time 2.737 (1.670)	Data 2.000 (0.976)	Loss 3.1783e+00 (3.2435e+00)	Acc@1 32.40 (3
1.10) Acc@5 58.80 (56.50)				
Epoch: [1][360/1282]	Time 2.537 (1.669)	Data 1.831 (0.975)	Loss 3.1603e+00 (3.2407e+00)	Acc@1 33.00 (3
1.15) Acc@5 60.20 (56.56)				
Epoch: [1][370/1282]	Time 2.488 (1.670)	Data 1.798 (0.976)	Loss 3.0371e+00 (3.2372e+00)	Acc@1 33.80 (3
1.20) Acc@5 60.20 (56.64)				
Epoch: [1][380/1282]	Time 2.634 (1.669)	Data 1.891 (0.974)	Loss 3.1002e+00 (3.2340e+00)	Acc@1 33.00 (3
1.25) Acc@5 59.20 (56.71)				
Epoch: [1][390/1282]	Time 2.856 (1.669)	Data 2.130 (0.974)	Loss 3.2313e+00 (3.2316e+00)	Acc@1 30.20 (3

[illegible]

Epoch: [1][650/1282]	Time 2.602 (1.673)	Data 1.870 (0.975)	Loss 2.8553e+00 (3.1582e+00)	Acc@1 37.20 (3
2.48) Acc@5 65.20 (58.12)				
Epoch: [1][660/1282]	Time 2.349 (1.673)	Data 1.670 (0.975)	Loss 3.1030e+00 (3.1549e+00)	Acc@1 30.60 (3
2.53) Acc@5 57.40 (58.18)				
Epoch: [1][670/1282]	Time 2.816 (1.673)	Data 2.093 (0.975)	Loss 2.9015e+00 (3.1526e+00)	Acc@1 35.20 (3
2.56) Acc@5 63.20 (58.23)				
Epoch: [1][680/1282]	Time 2.593 (1.673)	Data 1.877 (0.975)	Loss 3.0045e+00 (3.1507e+00)	Acc@1 34.80 (3
2.59) Acc@5 57.60 (58.26)				
Epoch: [1][690/1282]	Time 2.843 (1.674)	Data 2.108 (0.975)	Loss 2.8594e+00 (3.1473e+00)	Acc@1 36.80 (3
2.63) Acc@5 62.00 (58.32)				
Epoch: [1][700/1282]	Time 2.417 (1.673)	Data 1.738 (0.974)	Loss 2.9376e+00 (3.1444e+00)	Acc@1 34.60 (3
2.67) Acc@5 60.80 (58.37)				
Epoch: [1][710/1282]	Time 2.474 (1.673)	Data 1.795 (0.974)	Loss 3.0007e+00 (3.1419e+00)	Acc@1 35.20 (3
2.70) Acc@5 59.20 (58.41)				
Epoch: [1][720/1282]	Time 2.511 (1.673)	Data 1.816 (0.974)	Loss 2.9390e+00 (3.1391e+00)	Acc@1 35.00 (3
2.75) Acc@5 64.20 (58.47)				
Epoch: [1][730/1282]	Time 2.662 (1.673)	Data 1.910 (0.974)	Loss 2.9971e+00 (3.1376e+00)	Acc@1 35.40 (3
2.77) Acc@5 58.80 (58.49)				
Epoch: [1][740/1282]	Time 2.919 (1.673)	Data 2.182 (0.974)	Loss 2.9998e+00 (3.1350e+00)	Acc@1 34.00 (3
2.81) Acc@5 61.80 (58.54)				
Epoch: [1][750/1282]	Time 2.536 (1.674)	Data 1.812 (0.974)	Loss 3.1651e+00 (3.1339e+00)	Acc@1 33.80 (3
2.82) Acc@5 58.20 (58.56)				
Epoch: [1][760/1282]	Time 2.688 (1.674)	Data 1.956 (0.974)	Loss 2.8727e+00 (3.1308e+00)	Acc@1 38.40 (3
2.88) Acc@5 63.20 (58.61)				
Epoch: [1][770/1282]	Time 2.840 (1.673)	Data 2.089 (0.974)	Loss 2.8918e+00 (3.1285e+00)	Acc@1 36.80 (3
2.92) Acc@5 62.00 (58.67)				
Epoch: [1][780/1282]	Time 2.699 (1.673)	Data 1.982 (0.974)	Loss 2.9225e+00 (3.1258e+00)	Acc@1 35.00 (3
2.96) Acc@5 62.20 (58.71)				
Epoch: [1][790/1282]	Time 2.778 (1.673)	Data 2.043 (0.974)	Loss 3.0760e+00 (3.1234e+00)	Acc@1 29.80 (3
2.98) Acc@5 61.80 (58.75)				
Epoch: [1][800/1282]	Time 2.648 (1.673)	Data 1.902 (0.974)	Loss 2.8288e+00 (3.1207e+00)	Acc@1 36.80 (3
3.02) Acc@5 59.00 (58.79)				
Epoch: [1][810/1282]	Time 2.624 (1.674)	Data 1.908 (0.974)	Loss 2.9205e+00 (3.1183e+00)	Acc@1 35.20 (3
3.06) Acc@5 61.00 (58.83)				
Epoch: [1][820/1282]	Time 2.654 (1.673)	Data 1.919 (0.974)	Loss 2.7480e+00 (3.1149e+00)	Acc@1 40.00 (3
3.12) Acc@5 65.60 (58.90)				
Epoch: [1][830/1282]	Time 2.788 (1.673)	Data 2.056 (0.973)	Loss 2.8259e+00 (3.1130e+00)	Acc@1 38.80 (3
3.14) Acc@5 65.80 (58.93)				
Epoch: [1][840/1282]	Time 2.671 (1.674)	Data 1.930 (0.974)	Loss 2.9183e+00 (3.1106e+00)	Acc@1 35.20 (3
3.19) Acc@5 63.60 (58.98)				
Epoch: [1][850/1282]	Time 2.625 (1.673)	Data 1.908 (0.974)	Loss 2.7757e+00 (3.1076e+00)	Acc@1 39.20 (3
3.25) Acc@5 64.60 (59.04)				
Epoch: [1][860/1282]	Time 2.532 (1.673)	Data 1.813 (0.973)	Loss 2.9838e+00 (3.1050e+00)	Acc@1 36.20 (3
3.30) Acc@5 61.80 (59.08)				
Epoch: [1][870/1282]	Time 2.662 (1.673)	Data 1.941 (0.973)	Loss 2.9230e+00 (3.1027e+00)	Acc@1 36.60 (3
3.32) Acc@5 63.20 (59.11)				
Epoch: [1][880/1282]	Time 3.286 (1.674)	Data 2.536 (0.974)	Loss 2.8198e+00 (3.1001e+00)	Acc@1 39.20 (3
3.36) Acc@5 65.40 (59.16)				
Epoch: [1][890/1282]	Time 2.687 (1.674)	Data 1.945 (0.974)	Loss 2.9094e+00 (3.0976e+00)	Acc@1 34.80 (3
3.40) Acc@5 63.60 (59.21)				
Epoch: [1][900/1282]	Time 2.806 (1.674)	Data 2.093 (0.974)	Loss 2.9170e+00 (3.0952e+00)	Acc@1 34.20 (3

3.43)	Acc@5	60.40 (59.25)										
Epoch: [1][910/1282]	Time	2.461 (1.674)	Data	1.782 (0.974)	Loss	2.8521e+00 (3.0926e+00)	Acc@1	33.60 (3				
3.48)	Acc@5	63.00 (59.29)										
Epoch: [1][920/1282]	Time	2.635 (1.674)	Data	1.916 (0.974)	Loss	2.9959e+00 (3.0900e+00)	Acc@1	36.80 (3				
3.53)	Acc@5	59.60 (59.33)										
Epoch: [1][930/1282]	Time	2.773 (1.674)	Data	2.042 (0.974)	Loss	2.9035e+00 (3.0873e+00)	Acc@1	36.40 (3				
3.57)	Acc@5	62.60 (59.38)										
Epoch: [1][940/1282]	Time	2.615 (1.673)	Data	1.895 (0.973)	Loss	2.8417e+00 (3.0850e+00)	Acc@1	38.80 (3				
3.61)	Acc@5	66.00 (59.43)										
Epoch: [1][950/1282]	Time	2.544 (1.673)	Data	1.820 (0.973)	Loss	2.6742e+00 (3.0824e+00)	Acc@1	40.60 (3				
3.65)	Acc@5	64.80 (59.47)										
Epoch: [1][960/1282]	Time	2.648 (1.673)	Data	1.914 (0.973)	Loss	2.9388e+00 (3.0803e+00)	Acc@1	35.60 (3				
3.68)	Acc@5	61.80 (59.51)										
Epoch: [1][970/1282]	Time	2.495 (1.673)	Data	1.771 (0.973)	Loss	2.8415e+00 (3.0783e+00)	Acc@1	36.80 (3				
3.71)	Acc@5	65.20 (59.55)										
Epoch: [1][980/1282]	Time	2.587 (1.673)	Data	1.856 (0.973)	Loss	3.0179e+00 (3.0760e+00)	Acc@1	35.80 (3				
3.76)	Acc@5	60.80 (59.59)										
Epoch: [1][990/1282]	Time	3.049 (1.673)	Data	2.302 (0.973)	Loss	2.8478e+00 (3.0739e+00)	Acc@1	42.00 (3				
3.80)	Acc@5	62.80 (59.63)										
Epoch: [1][1000/1282]	Time	2.512 (1.672)	Data	1.781 (0.972)	Loss	2.9450e+00 (3.0718e+00)	Acc@1	35.40 (3				
3.83)	Acc@5	62.60 (59.66)										
Epoch: [1][1010/1282]	Time	2.743 (1.672)	Data	2.008 (0.972)	Loss	2.8890e+00 (3.0695e+00)	Acc@1	38.80 (3				
3.86)	Acc@5	62.40 (59.70)										
Epoch: [1][1020/1282]	Time	2.475 (1.672)	Data	1.794 (0.972)	Loss	2.8282e+00 (3.0672e+00)	Acc@1	35.40 (3				
3.90)	Acc@5	64.40 (59.74)										
Epoch: [1][1030/1282]	Time	2.512 (1.673)	Data	1.796 (0.973)	Loss	2.8795e+00 (3.0650e+00)	Acc@1	36.20 (3				
3.93)	Acc@5	65.40 (59.78)										
Epoch: [1][1040/1282]	Time	2.767 (1.673)	Data	2.047 (0.973)	Loss	2.9928e+00 (3.0631e+00)	Acc@1	34.40 (3				
3.96)	Acc@5	61.20 (59.82)										
Epoch: [1][1050/1282]	Time	2.525 (1.673)	Data	1.801 (0.972)	Loss	2.9927e+00 (3.0612e+00)	Acc@1	38.60 (3				
4.00)	Acc@5	62.20 (59.85)										
Epoch: [1][1060/1282]	Time	2.502 (1.673)	Data	1.799 (0.972)	Loss	2.7093e+00 (3.0588e+00)	Acc@1	37.60 (3				
4.03)	Acc@5	66.40 (59.90)										
Epoch: [1][1070/1282]	Time	2.814 (1.673)	Data	2.080 (0.972)	Loss	2.7651e+00 (3.0566e+00)	Acc@1	35.00 (3				
4.06)	Acc@5	66.20 (59.94)										
Epoch: [1][1080/1282]	Time	2.693 (1.672)	Data	1.942 (0.972)	Loss	2.7423e+00 (3.0545e+00)	Acc@1	40.40 (3				
4.10)	Acc@5	66.40 (59.97)										
Epoch: [1][1090/1282]	Time	2.513 (1.672)	Data	1.838 (0.972)	Loss	2.7287e+00 (3.0525e+00)	Acc@1	42.00 (3				
4.13)</												

```

Epoch: [1][1160/1282] Time 2.796 ( 1.672) Data 2.055 ( 0.971) Loss 2.6790e+00 (3.0360e+00) Acc@1 40.20 ( 3
4.39) Acc@5 67.60 ( 60.30)
Epoch: [1][1170/1282] Time 2.697 ( 1.672) Data 1.960 ( 0.971) Loss 2.6251e+00 (3.0333e+00) Acc@1 42.60 ( 3
4.43) Acc@5 67.20 ( 60.34)
Epoch: [1][1180/1282] Time 2.104 ( 1.671) Data 1.426 ( 0.970) Loss 2.8818e+00 (3.0311e+00) Acc@1 33.80 ( 3
4.46) Acc@5 65.80 ( 60.38)
Epoch: [1][1190/1282] Time 2.033 ( 1.671) Data 1.353 ( 0.968) Loss 2.8284e+00 (3.0287e+00) Acc@1 38.40 ( 3
4.50) Acc@5 64.40 ( 60.42)
Epoch: [1][1200/1282] Time 2.409 ( 1.671) Data 1.728 ( 0.966) Loss 2.6214e+00 (3.0264e+00) Acc@1 43.40 ( 3
4.55) Acc@5 67.80 ( 60.46)
Epoch: [1][1210/1282] Time 2.407 ( 1.671) Data 1.728 ( 0.966) Loss 2.7749e+00 (3.0243e+00) Acc@1 38.20 ( 3
4.59) Acc@5 64.00 ( 60.50)
Epoch: [1][1220/1282] Time 2.318 ( 1.671) Data 1.635 ( 0.966) Loss 2.7847e+00 (3.0221e+00) Acc@1 41.00 ( 3
4.63) Acc@5 63.40 ( 60.54)
Epoch: [1][1230/1282] Time 2.384 ( 1.671) Data 1.700 ( 0.966) Loss 2.7828e+00 (3.0200e+00) Acc@1 35.80 ( 3
4.66) Acc@5 65.60 ( 60.58)
Epoch: [1][1240/1282] Time 2.228 ( 1.671) Data 1.548 ( 0.966) Loss 2.8328e+00 (3.0178e+00) Acc@1 35.40 ( 3
4.69) Acc@5 63.80 ( 60.62)
Epoch: [1][1250/1282] Time 2.299 ( 1.671) Data 1.620 ( 0.966) Loss 2.7661e+00 (3.0156e+00) Acc@1 39.60 ( 3
4.73) Acc@5 66.40 ( 60.65)
Epoch: [1][1260/1282] Time 2.242 ( 1.671) Data 1.561 ( 0.966) Loss 2.5627e+00 (3.0135e+00) Acc@1 43.20 ( 3
4.76) Acc@5 70.00 ( 60.69)
Epoch: [1][1270/1282] Time 2.238 ( 1.671) Data 1.560 ( 0.967) Loss 2.4949e+00 (3.0111e+00) Acc@1 43.80 ( 3
4.80) Acc@5 71.20 ( 60.74)
Epoch: [1][1280/1282] Time 2.363 ( 1.671) Data 1.685 ( 0.967) Loss 2.7091e+00 (3.0089e+00) Acc@1 40.00 ( 3
4.84) Acc@5 65.00 ( 60.77)
Test: [ 0/100] Time 4.481 ( 4.481) Loss 2.2724e+00 (2.2724e+00) Acc@1 44.00 ( 44.00) Acc@5 76.00 ( 76.00)
Test: [ 10/100] Time 2.988 ( 1.991) Loss 2.4043e+00 (2.8740e+00) Acc@1 47.60 ( 37.35) Acc@5 72.40 ( 64.11)
Test: [ 20/100] Time 2.846 ( 1.854) Loss 2.6047e+00 (2.8327e+00) Acc@1 31.00 ( 35.78) Acc@5 69.20 ( 64.16)
Test: [ 30/100] Time 2.926 ( 1.803) Loss 2.5772e+00 (2.7721e+00) Acc@1 42.20 ( 35.82) Acc@5 66.20 ( 65.08)
Test: [ 40/100] Time 2.855 ( 1.792) Loss 2.7923e+00 (2.7790e+00) Acc@1 39.40 ( 36.37) Acc@5 65.80 ( 64.80)
Test: [ 50/100] Time 2.898 ( 1.781) Loss 3.9063e+00 (2.9267e+00) Acc@1 21.00 ( 34.66) Acc@5 42.60 ( 62.20)
Test: [ 60/100] Time 2.834 ( 1.776) Loss 3.4097e+00 (2.9871e+00) Acc@1 32.80 ( 34.19) Acc@5 54.20 ( 61.12)
Test: [ 70/100] Time 3.037 ( 1.764) Loss 3.5641e+00 (3.0489e+00) Acc@1 26.40 ( 33.51) Acc@5 54.00 ( 60.16)
Test: [ 80/100] Time 3.062 ( 1.756) Loss 3.1252e+00 (3.0973e+00) Acc@1 37.40 ( 33.22) Acc@5 56.60 ( 59.31)
Test: [ 90/100] Time 2.920 ( 1.745) Loss 3.3970e+00 (3.1454e+00) Acc@1 31.40 ( 32.45) Acc@5 54.20 ( 58.50)
* Acc@1 33.078 Acc@5 59.166
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	

