In [162…

```python
from __future__ import print_function, division

import torch
import torch.nn as nn
import torch.optim as optim
from torch.optim import lr_scheduler
import torch.backends.cudnn as cudnn
import numpy as np
import torchvision
import torchvision.transforms as transforms
from torchvision import datasets, models, transforms
import matplotlib.pyplot as plt
import time
import os
import copy

import torch.nn.functional as F
from torchvision.datasets import CIFAR100
import torchvision.transforms as tt
from torchvision.utils import make_grid
from torch.utils.data.dataloader import DataLoader
from torch.utils.data import random_split,ConcatDataset
cudnn.benchmark = True
import pickle
import pandas as pd
plt.ion()   # interactive mode
import warnings
warnings.filterwarnings("ignore")
```

Question 1

(1)

(a): Finetuning

In [2]:

```python
stats = ((0.5074,0.4867,0.4411),(0.2011,0.1987,0.2025))
train_transform = tt.Compose([
    tt.RandomHorizontalFlip(),
    tt.RandomCrop(32,padding=4,padding_mode="reflect"),
    tt.ToTensor(),
    tt.Normalize(*stats)
])

test_transform = tt.Compose([
    tt.ToTensor(),
    tt.Normalize(*stats)
])

# downloda data
#train_data = CIFAR100(download=True,root="./data",transform=train_transform)
train_data = CIFAR100(download=True,root="./data",transform=train_transform)
test_data = CIFAR100(root="./data",train=False,transform=test_transform)
```

```
Files already downloaded and verified
```

In [3]:

```python
train_data,test_data
```

Out[3]:

```
(Dataset CIFAR100
     Number of datapoints: 50000
```

```
    Root location: ./data
    Split: Train
    StandardTransform
Transform: Compose(
               RandomHorizontalFlip(p=0.5)
               RandomCrop(size=(32, 32), padding=4)
               ToTensor()
               Normalize(mean=(0.5074, 0.4867, 0.4411), std=(0.2011, 0.1987, 0.2025))
           ),
Dataset CIFAR100
    Number of datapoints: 10000
    Root location: ./data
    Split: Test
    StandardTransform
Transform: Compose(
               ToTensor()
               Normalize(mean=(0.5074, 0.4867, 0.4411), std=(0.2011, 0.1987, 0.2025))
           ))
```

In [4]:

```
test_classes_items = dict()
for test_item in test_data:
    label = test_data.classes[test_item[1]]
    if label not in test_classes_items:
        test_classes_items[label] = 1
    else:
        test_classes_items[label] += 1

test_classes_items
```

Out[4]:

```
{'mountain': 100,
 'forest': 100,
 'seal': 100,
 'mushroom': 100,
 'sea': 100,
 'tulip': 100,
 'camel': 100,
 'butterfly': 100,
 'cloud': 100,
 'apple': 100,
 'skunk': 100,
 'streetcar': 100,
 'rocket': 100,
 'lamp': 100,
 'lion': 100,
 'wolf': 100,
 'rose': 100,
 'orange': 100,
 'dinosaur': 100,
 'chimpanzee': 100,
 'can': 100,
 'keyboard': 100,
 'bicycle': 100,
 'chair': 100,
 'plate': 100,
 'lawn_mower': 100,
 'turtle': 100,
 'palm_tree': 100,
 'shark': 100,
 'pickup_truck': 100,
 'boy': 100,
 'couch': 100,
 'house': 100,
 'porcupine': 100,
 'cockroach': 100,
 'clock': 100,
 'castle': 100,
 'beaver': 100,
 'bee': 100,
 'bottle': 100,
 'pear': 100,
 'baby': 100,
 'flatfish': 100,
 'oak_tree': 100,
 'leopard': 100,
 'snail': 100,
 'crocodile': 100,
 'rabbit': 100,
 'beetle': 100,
 'girl': 100,
 'sunflower': 100,
 'raccoon': 100,
 'train': 100,
 'ray': 100,
 'trout': 100,
 'bowl': 100,
 'snake': 100,
 'orchid': 100,
 'tractor': 100,
 'caterpillar': 100,
 'bus': 100,
 'mouse': 100,
 'crab': 100,
 'sweet_pepper': 100,
 'maple_tree': 100,
 'whale': 100,
 'skyscraper': 100,
 'pine_tree': 100,
 'tank': 100,
 'cattle': 100,
 'man': 100,
 'aquarium_fish': 100,
 'shrew': 100,
 'plain': 100,
 'possum': 100,
 'lobster': 100,
 'hamster': 100,
 'television': 100,
 'dolphin': 100,
 'worm': 100,
 'squirrel': 100,
 'cup': 100,
 'woman': 100,
 'bridge': 100,
 'wardrobe': 100,
 'road': 100,
 'lizard': 100,
 'elephant': 100,
 'spider': 100,
 'fox': 100,
 'otter': 100,
 'poppy': 100,
 'willow_tree': 100,
 'table': 100,
 'kangaroo': 100,
 'telephone': 100,
 'bed': 100,
 'motorcycle': 100,
 'bear': 100,
 'tiger': 100}
```

```
In [5]: train_classes_items = dict()
        for train_item in train_data:
            label = train_data.classes[train_item[1]]
            if label not in train_classes_items:
                train_classes_items[label] = 1
            else:
                train_classes_items[label] += 1
        train_classes_items
```

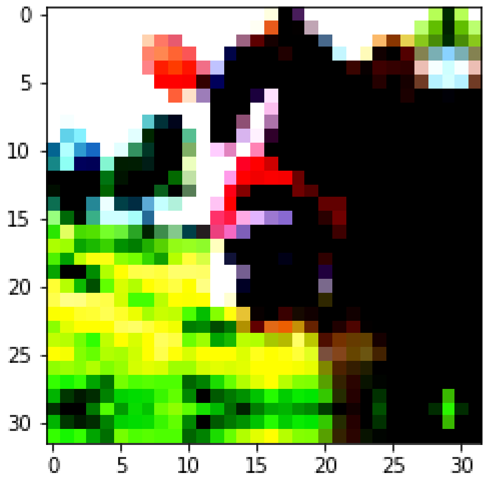
```
Out[5]: {'cattle': 500,
         'dinosaur': 500,
         'apple': 500,
         'boy': 500,
         'aquarium_fish': 500,
         'telephone': 500,
         'train': 500,
         'cup': 500,
         'cloud': 500,
         'elephant': 500,
         'keyboard': 500,
         'willow_tree': 500,
         'sunflower': 500,
         'castle': 500,
         'sea': 500,
         'bicycle': 500,
         'wolf': 500,
         'squirrel': 500,
         'shrew': 500,
         'pine_tree': 500,
         'rose': 500,
         'television': 500,
         'table': 500,
         'possum': 500,
         'oak_tree': 500,
         'leopard': 500,
         'maple_tree': 500,
         'rabbit': 500,
         'chimpanzee': 500,
         'clock': 500,
         'streetcar': 500,
         'cockroach': 500,
         'snake': 500,
         'lobster': 500,
         'mountain': 500,
         'palm_tree': 500,
         'skyscraper': 500,
         'tractor': 500,
         'shark': 500,
         'butterfly': 500,
         'bottle': 500,
         'bee': 500,
         'chair': 500,
         'woman': 500,
         'hamster': 500,
         'otter': 500,
         'seal': 500,
         'lion': 500,
         'mushroom': 500,
         'girl': 500,
         'sweet_pepper': 500,
         'forest': 500,
         'crocodile': 500,
         'orange': 500,
         'tulip': 500,
         'mouse': 500,
         'camel': 500,
         'caterpillar': 500,
         'man': 500,
         'skunk': 500,
         'kangaroo': 500,
         'raccoon': 500,
         'snail': 500,
         'rocket': 500,
         'whale': 500,
         'worm': 500,
         'turtle': 500,
         'beaver': 500,
         'plate': 500,
         'wardrobe': 500,
         'road': 500,
         'fox': 500,
         'flatfish': 500,
         'tiger': 500,
         'ray': 500,
         'dolphin': 500,
         'poppy': 500,
         'porcupine': 500,
         'lamp': 500,
         'crab': 500,
         'motorcycle': 500,
         'spider': 500,
         'tank': 500,
         'orchid': 500,
         'lizard': 500,
         'beetle': 500,
         'bridge': 500,
         'baby': 500,
         'lawn_mower': 500,
         'house': 500,
         'bus': 500,
         'couch': 500,
         'bowl': 500,
         'pear': 500,
         'bed': 500,
         'plain': 500,
         'trout': 500,
         'bear': 500,
         'pickup_truck': 500,
         'can': 500}
```

- Describe dataset
  - As shown above, there is a total of 60000 images in CIFAR-100 dataset.
  - There are 100 classes in the data set. And the data distributed evenly in the training set and testing set, each class contain 600 images.

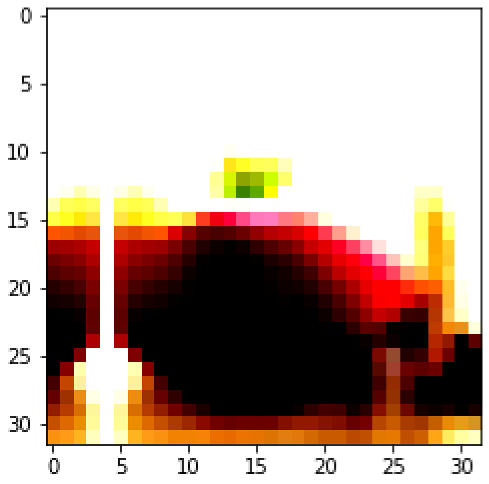
In [6]:

```
# show two classes of samples
selected_class = ['cattle','apple']
cnt_1 = 0
cnt_2 = 0
for sample in train_data:
    if train_data.classes[sample[1]] == selected_class[0] and cnt_1 < 2:
        print(f'class_name = {selected_class[0]}')
        plt.imshow(sample[0].permute(1,2,0))
        plt.show()
        cnt_1+=1
    if train_data.classes[sample[1]] == selected_class[1] and cnt_2 < 2:
        print(f'class_name = {selected_class[1]}')
        plt.imshow(sample[0].permute(1,2,0))
        plt.show()
        cnt_2+=1
    if cnt_2 == 2 and cnt_1 == 2:
        break
```

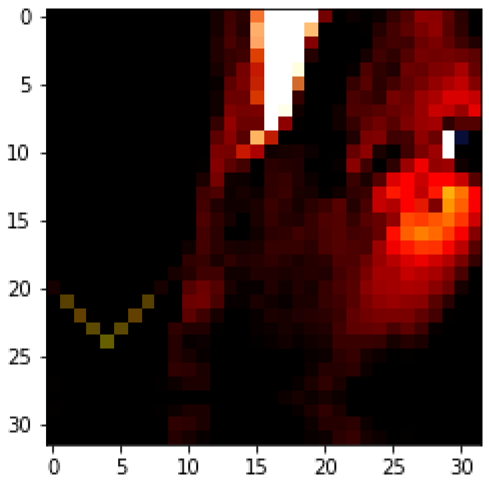
```
Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..
255] for integers).
class_name = cattle
```



```
Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..
255] for integers).
class_name = apple
```

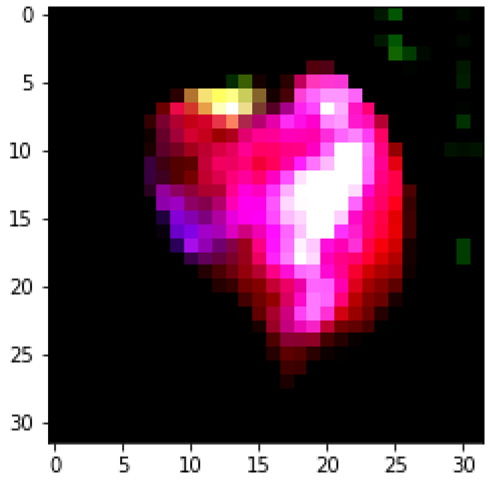


```
Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..
255] for integers).
class_name = cattle
```



```
Clipping input data to the valid range for imshow with RGB data ([0..1] for floats or [0..
255] for integers).
```

```
class_name = apple
```



(b)

In [7]:
```python
device = 'cuda'
dataloaders = {
    'train':torch.utils.data.DataLoader(train_data, batch_size=64,
                                             shuffle=True, num_workers=4),
    'val':torch.utils.data.DataLoader(test_data, batch_size=64,
                                             shuffle=True, num_workers=4)
}
dataset_sizes = {'train':len(train_data),
                 'val':len(test_data)}
```

In [8]:
```python
def train_model(model, criterion, optimizer, scheduler, num_epochs=25):
    since = time.time()

    best_model_wts = copy.deepcopy(model.state_dict())
    best_acc = 0.0

    for epoch in range(num_epochs):
        print(f'Epoch {epoch}/{num_epochs - 1}')
        print('-' * 10)

        # Each epoch has a training and validation phase
        for phase in ['train', 'val']:
            if phase == 'train':
                model.train()  # Set model to training mode
            else:
                model.eval()   # Set model to evaluate mode

            running_loss = 0.0
            running_corrects = 0

            # Iterate over data.
            for inputs, labels in dataloaders[phase]:
                total_batch = len(dataloaders[phase])
                inputs = inputs.to(device)
                labels = labels.to(device)

                # zero the parameter gradients
                optimizer.zero_grad()

                # forward
                # track history if only in train
                with torch.set_grad_enabled(phase == 'train'):
                    outputs = model(inputs)
                    _, preds = torch.max(outputs, 1)
                    loss = criterion(outputs, labels)

                    # backward + optimize only if in training phase
                    if phase == 'train':
                        loss.backward()
                        optimizer.step()

                # statistics
                running_loss += loss.item() * inputs.size(0)
                running_corrects += torch.sum(preds == labels.data)
            if phase == 'train':
                scheduler.step()

            epoch_loss = running_loss / dataset_sizes[phase]
            epoch_acc = running_corrects.double() / dataset_sizes[phase]

            print(f'{phase} Loss: {epoch_loss:.4f} Acc: {epoch_acc:.4f}')

            # deep copy the model
            if phase == 'val' and epoch_acc > best_acc:
                best_acc = epoch_acc
                best_model_wts = copy.deepcopy(model.state_dict())

        print()

    time_elapsed = time.time() - since
    print(f'Training complete in {time_elapsed // 60:.0f}m {time_elapsed % 60:.0f}s')
    print(f'Best val Acc: {best_acc:4f}')

    # load best model weights
    model.load_state_dict(best_model_wts)
    return model
```

In [89]:

```
# load model and change fc layer
learning_rate = 0.001
num_classes = len(train_data.classes)
model_ft = models.resnet50(pretrained=True)
num_ftrs = model_ft.fc.in_features
model_ft.fc = nn.Linear(num_ftrs, num_classes)
model_ft = model_ft.to(device)
criterion = nn.CrossEntropyLoss()
# Observe that all parameters are being optimized
optimizer_ft = optim.SGD(model_ft.parameters(), lr=learning_rate, momentum=0.9)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_ft, step_size=20, gamma=0.1)

model_ft = train_model(model_ft, criterion, optimizer_ft, exp_lr_scheduler,
                       num_epochs=80)
# Since the network start to converge at around 60 epochs, I trained the network for 80 ep
```

```
Epoch 0/79
----------
train Loss: 2.9764 Acc: 0.2821
val Loss: 2.0045 Acc: 0.4538

Epoch 1/79
----------
train Loss: 1.9253 Acc: 0.4755
val Loss: 1.6826 Acc: 0.5313

Epoch 2/79
----------
train Loss: 1.6377 Acc: 0.5449
```

```
val Loss: 1.5286 Acc: 0.5690

Epoch 3/79
----------
train Loss: 1.4658 Acc: 0.5858
val Loss: 1.4675 Acc: 0.5898

Epoch 4/79
----------
train Loss: 1.3365 Acc: 0.6177
val Loss: 1.4131 Acc: 0.6065

Epoch 5/79
----------
train Loss: 1.2300 Acc: 0.6456
val Loss: 1.4281 Acc: 0.6048

Epoch 6/79
----------
train Loss: 1.1362 Acc: 0.6684
val Loss: 1.3995 Acc: 0.6125

Epoch 7/79
----------
train Loss: 1.0614 Acc: 0.6861
val Loss: 1.3610 Acc: 0.6242

Epoch 8/79
----------
train Loss: 0.9826 Acc: 0.7101
val Loss: 1.3614 Acc: 0.6259

Epoch 9/79
----------
train Loss: 0.9218 Acc: 0.7251
val Loss: 1.3403 Acc: 0.6348

Epoch 10/79
----------
train Loss: 0.8680 Acc: 0.7402
val Loss: 1.3496 Acc: 0.6370

Epoch 11/79
----------
train Loss: 0.8173 Acc: 0.7543
val Loss: 1.3588 Acc: 0.6406

Epoch 12/79
----------
train Loss: 0.7583 Acc: 0.7693
val Loss: 1.3709 Acc: 0.6389

Epoch 13/79
----------
train Loss: 0.7140 Acc: 0.7817
val Loss: 1.3870 Acc: 0.6413

Epoch 14/79
----------
train Loss: 0.6791 Acc: 0.7914
val Loss: 1.3837 Acc: 0.6435

Epoch 15/79
----------
train Loss: 0.6425 Acc: 0.8017
val Loss: 1.3934 Acc: 0.6448
```

```
Epoch 16/79
----------
train Loss: 0.5978 Acc: 0.8163
val Loss: 1.4252 Acc: 0.6408

Epoch 17/79
----------
train Loss: 0.5697 Acc: 0.8259
val Loss: 1.4379 Acc: 0.6444

Epoch 18/79
----------
train Loss: 0.5401 Acc: 0.8316
val Loss: 1.4317 Acc: 0.6497

Epoch 19/79
----------
train Loss: 0.5034 Acc: 0.8428
val Loss: 1.4734 Acc: 0.6445

Epoch 20/79
----------
train Loss: 0.3743 Acc: 0.8854
val Loss: 1.3797 Acc: 0.6609

Epoch 21/79
----------
train Loss: 0.3372 Acc: 0.8987
val Loss: 1.3785 Acc: 0.6636

Epoch 22/79
----------
train Loss: 0.3093 Acc: 0.9065
val Loss: 1.3632 Acc: 0.6666

Epoch 23/79
----------
train Loss: 0.2982 Acc: 0.9114
val Loss: 1.3395 Acc: 0.6688

Epoch 24/79
----------
train Loss: 0.2833 Acc: 0.9160
val Loss: 1.3515 Acc: 0.6699

Epoch 25/79
----------
train Loss: 0.2794 Acc: 0.9150
val Loss: 1.3590 Acc: 0.6700

Epoch 26/79
----------
train Loss: 0.2667 Acc: 0.9207
val Loss: 1.3583 Acc: 0.6721

Epoch 27/79
----------
train Loss: 0.2558 Acc: 0.9241
val Loss: 1.3631 Acc: 0.6695

Epoch 28/79
----------
train Loss: 0.2497 Acc: 0.9258
val Loss: 1.3780 Acc: 0.6695
```

```
Epoch 29/79
----------
train Loss: 0.2423 Acc: 0.9289
val Loss: 1.3714 Acc: 0.6712

Epoch 30/79
----------
train Loss: 0.2372 Acc: 0.9288
val Loss: 1.3728 Acc: 0.6707

Epoch 31/79
----------
train Loss: 0.2263 Acc: 0.9327
val Loss: 1.3779 Acc: 0.6712

Epoch 32/79
----------
train Loss: 0.2287 Acc: 0.9323
val Loss: 1.3777 Acc: 0.6714

Epoch 33/79
----------
train Loss: 0.2242 Acc: 0.9334
val Loss: 1.3821 Acc: 0.6712

Epoch 34/79
----------
train Loss: 0.2191 Acc: 0.9344
val Loss: 1.3917 Acc: 0.6706

Epoch 35/79
----------
train Loss: 0.2101 Acc: 0.9387
val Loss: 1.3942 Acc: 0.6729

Epoch 36/79
----------
train Loss: 0.2066 Acc: 0.9388
val Loss: 1.4050 Acc: 0.6699

Epoch 37/79
----------
train Loss: 0.1997 Acc: 0.9422
val Loss: 1.3958 Acc: 0.6745

Epoch 38/79
----------
train Loss: 0.1982 Acc: 0.9427
val Loss: 1.4166 Acc: 0.6693

Epoch 39/79
----------
train Loss: 0.1907 Acc: 0.9436
val Loss: 1.4080 Acc: 0.6730

Epoch 40/79
----------
train Loss: 0.1911 Acc: 0.9432
val Loss: 1.3944 Acc: 0.6725

Epoch 41/79
----------
train Loss: 0.1871 Acc: 0.9457
val Loss: 1.4140 Acc: 0.6690

Epoch 42/79
```

```
----------
train Loss: 0.1859 Acc: 0.9464
val Loss: 1.3982 Acc: 0.6712

Epoch 43/79
----------
train Loss: 0.1863 Acc: 0.9450
val Loss: 1.4086 Acc: 0.6714

Epoch 44/79
----------
train Loss: 0.1857 Acc: 0.9469
val Loss: 1.3926 Acc: 0.6754

Epoch 45/79
----------
train Loss: 0.1852 Acc: 0.9450
val Loss: 1.4009 Acc: 0.6738

Epoch 46/79
----------
train Loss: 0.1821 Acc: 0.9485
val Loss: 1.4127 Acc: 0.6714

Epoch 47/79
----------
train Loss: 0.1795 Acc: 0.9480
val Loss: 1.3954 Acc: 0.6724

Epoch 48/79
----------
train Loss: 0.1801 Acc: 0.9479
val Loss: 1.4120 Acc: 0.6717

Epoch 49/79
----------
train Loss: 0.1837 Acc: 0.9476
val Loss: 1.3962 Acc: 0.6734

Epoch 50/79
----------
train Loss: 0.1796 Acc: 0.9481
val Loss: 1.4022 Acc: 0.6744

Epoch 51/79
----------
train Loss: 0.1805 Acc: 0.9484
val Loss: 1.4122 Acc: 0.6704

Epoch 52/79
----------
train Loss: 0.1781 Acc: 0.9477
val Loss: 1.4052 Acc: 0.6719

Epoch 53/79
----------
train Loss: 0.1808 Acc: 0.9480
val Loss: 1.3955 Acc: 0.6749

Epoch 54/79
----------
train Loss: 0.1810 Acc: 0.9479
val Loss: 1.3988 Acc: 0.6737

Epoch 55/79
----------
```

```
train Loss: 0.1804 Acc: 0.9476
val Loss: 1.4152 Acc: 0.6717

Epoch 56/79
----------
train Loss: 0.1792 Acc: 0.9495
val Loss: 1.4063 Acc: 0.6736

Epoch 57/79
----------
train Loss: 0.1777 Acc: 0.9483
val Loss: 1.4071 Acc: 0.6746

Epoch 58/79
----------
train Loss: 0.1753 Acc: 0.9501
val Loss: 1.4081 Acc: 0.6731

Epoch 59/79
----------
train Loss: 0.1774 Acc: 0.9488
val Loss: 1.4130 Acc: 0.6699

Epoch 60/79
----------
train Loss: 0.1784 Acc: 0.9494
val Loss: 1.4019 Acc: 0.6752

Epoch 61/79
----------
train Loss: 0.1769 Acc: 0.9488
val Loss: 1.4021 Acc: 0.6699

Epoch 62/79
----------
train Loss: 0.1776 Acc: 0.9483
val Loss: 1.4066 Acc: 0.6739

Epoch 63/79
----------
train Loss: 0.1797 Acc: 0.9479
val Loss: 1.4174 Acc: 0.6709

Epoch 64/79
----------
train Loss: 0.1734 Acc: 0.9500
val Loss: 1.4192 Acc: 0.6697

Epoch 65/79
----------
train Loss: 0.1798 Acc: 0.9490
val Loss: 1.4169 Acc: 0.6726

Epoch 66/79
----------
train Loss: 0.1736 Acc: 0.9510
val Loss: 1.4073 Acc: 0.6720

Epoch 67/79
----------
train Loss: 0.1756 Acc: 0.9496
val Loss: 1.4009 Acc: 0.6706

Epoch 68/79
----------
train Loss: 0.1772 Acc: 0.9488
```

```
val Loss: 1.4089 Acc: 0.6720

Epoch 69/79
----------
train Loss: 0.1746 Acc: 0.9491
val Loss: 1.4108 Acc: 0.6722

Epoch 70/79
----------
train Loss: 0.1743 Acc: 0.9496
val Loss: 1.4105 Acc: 0.6716

Epoch 71/79
----------
train Loss: 0.1775 Acc: 0.9491
val Loss: 1.4089 Acc: 0.6730

Epoch 72/79
----------
train Loss: 0.1782 Acc: 0.9484
val Loss: 1.4016 Acc: 0.6713

Epoch 73/79
----------
train Loss: 0.1747 Acc: 0.9499
val Loss: 1.4115 Acc: 0.6720

Epoch 74/79
----------
train Loss: 0.1758 Acc: 0.9495
val Loss: 1.3966 Acc: 0.6736

Epoch 75/79
----------
train Loss: 0.1770 Acc: 0.9493
val Loss: 1.4115 Acc: 0.6711

Epoch 76/79
----------
train Loss: 0.1770 Acc: 0.9491
val Loss: 1.4109 Acc: 0.6732

Epoch 77/79
----------
train Loss: 0.1771 Acc: 0.9496
val Loss: 1.4046 Acc: 0.6718

Epoch 78/79
----------
train Loss: 0.1779 Acc: 0.9486
val Loss: 1.4009 Acc: 0.6738

Epoch 79/79
----------
train Loss: 0.1792 Acc: 0.9482
val Loss: 1.4134 Acc: 0.6726

Training complete in 50m 40s
Best val Acc: 0.675400
```

- According to the result, the best val Acc is 0.675400.

(c)

In [90]:

```
# load model and change fc layer
learning_rate = 0.01
num_classes = len(train_data.classes)
model_ft = models.resnet50(pretrained=True)
num_ftrs = model_ft.fc.in_features
model_ft.fc = nn.Linear(num_ftrs, num_classes)
model_ft = model_ft.to(device)
criterion = nn.CrossEntropyLoss()
# Observe that all parameters are being optimized
optimizer_ft = optim.SGD(model_ft.parameters(), lr=learning_rate, momentum=0.9)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_ft, step_size=20, gamma=0.1)

model_ft = train_model(model_ft, criterion, optimizer_ft, exp_lr_scheduler,
                       num_epochs=80)
# Since the network start to converge at around 60 epochs, I trained the network for 80 ep
```

```
Epoch 0/79
----------
train Loss: 2.8204 Acc: 0.3006
val Loss: 2.3131 Acc: 0.4144

Epoch 1/79
----------
train Loss: 2.1297 Acc: 0.4357
val Loss: 2.0548 Acc: 0.4796

Epoch 2/79
----------
train Loss: 1.9906 Acc: 0.4644
val Loss: 1.8964 Acc: 0.5232

Epoch 3/79
----------
train Loss: 1.8614 Acc: 0.4946
val Loss: 3.1633 Acc: 0.4490

Epoch 4/79
----------
train Loss: 1.8372 Acc: 0.4983
val Loss: 1.6819 Acc: 0.5462

Epoch 5/79
----------
train Loss: 1.6200 Acc: 0.5498
val Loss: 1.6984 Acc: 0.5504

Epoch 6/79
----------
train Loss: 1.6798 Acc: 0.5337
val Loss: 1.7081 Acc: 0.5548

Epoch 7/79
----------
train Loss: 1.5362 Acc: 0.5705
val Loss: 1.8331 Acc: 0.5661

Epoch 8/79
----------
train Loss: 1.3747 Acc: 0.6062
val Loss: 1.7035 Acc: 0.5597

Epoch 9/79
----------
```

```
train Loss: 1.3119 Acc: 0.6239
val Loss: 1.8238 Acc: 0.5775

Epoch 10/79
----------
train Loss: 1.3332 Acc: 0.6192
val Loss: 1.6394 Acc: 0.5737

Epoch 11/79
----------
train Loss: 1.1633 Acc: 0.6563
val Loss: 1.5348 Acc: 0.5961

Epoch 12/79
----------
train Loss: 1.0736 Acc: 0.6823
val Loss: 1.4744 Acc: 0.6076

Epoch 13/79
----------
train Loss: 1.0341 Acc: 0.6911
val Loss: 1.6546 Acc: 0.5717

Epoch 14/79
----------
train Loss: 1.0212 Acc: 0.6953
val Loss: 1.6624 Acc: 0.5985

Epoch 15/79
----------
train Loss: 0.9081 Acc: 0.7240
val Loss: 1.4910 Acc: 0.6047

Epoch 16/79
----------
train Loss: 0.8494 Acc: 0.7403
val Loss: 1.5718 Acc: 0.6149

Epoch 17/79
----------
train Loss: 0.8149 Acc: 0.7502
val Loss: 1.6071 Acc: 0.6072

Epoch 18/79
----------
train Loss: 0.7304 Acc: 0.7733
val Loss: 1.5329 Acc: 0.6185

Epoch 19/79
----------
train Loss: 0.7169 Acc: 0.7766
val Loss: 1.8433 Acc: 0.6112

Epoch 20/79
----------
train Loss: 0.4640 Acc: 0.8567
val Loss: 1.3434 Acc: 0.6558

Epoch 21/79
----------
train Loss: 0.3871 Acc: 0.8800
val Loss: 1.3866 Acc: 0.6568

Epoch 22/79
----------
train Loss: 0.3538 Acc: 0.8924
```

```
val Loss: 1.4534 Acc: 0.6559

Epoch 23/79
----------
train Loss: 0.3264 Acc: 0.9005
val Loss: 1.5180 Acc: 0.6516

Epoch 24/79
----------
train Loss: 0.3079 Acc: 0.9062
val Loss: 1.5479 Acc: 0.6523

Epoch 25/79
----------
train Loss: 0.2868 Acc: 0.9106
val Loss: 1.4685 Acc: 0.6561

Epoch 26/79
----------
train Loss: 0.2682 Acc: 0.9168
val Loss: 1.5936 Acc: 0.6505

Epoch 27/79
----------
train Loss: 0.2549 Acc: 0.9197
val Loss: 1.4914 Acc: 0.6539

Epoch 28/79
----------
train Loss: 0.2405 Acc: 0.9253
val Loss: 1.5638 Acc: 0.6566

Epoch 29/79
----------
train Loss: 0.2328 Acc: 0.9269
val Loss: 1.7390 Acc: 0.6467

Epoch 30/79
----------
train Loss: 0.2209 Acc: 0.9315
val Loss: 1.6130 Acc: 0.6557

Epoch 31/79
----------
train Loss: 0.2129 Acc: 0.9331
val Loss: 1.6663 Acc: 0.6555

Epoch 32/79
----------
train Loss: 0.2042 Acc: 0.9366
val Loss: 1.7398 Acc: 0.6502

Epoch 33/79
----------
train Loss: 0.1979 Acc: 0.9378
val Loss: 1.5916 Acc: 0.6542

Epoch 34/79
----------
train Loss: 0.1892 Acc: 0.9418
val Loss: 1.7143 Acc: 0.6527

Epoch 35/79
----------
train Loss: 0.1800 Acc: 0.9448
val Loss: 1.6561 Acc: 0.6504
```

```
Epoch 36/79
----------
train Loss: 0.1748 Acc: 0.9450
val Loss: 1.7801 Acc: 0.6509

Epoch 37/79
----------
train Loss: 0.1639 Acc: 0.9490
val Loss: 1.8112 Acc: 0.6474

Epoch 38/79
----------
train Loss: 0.1627 Acc: 0.9490
val Loss: 1.8387 Acc: 0.6494

Epoch 39/79
----------
train Loss: 0.1554 Acc: 0.9515
val Loss: 1.7754 Acc: 0.6464

Epoch 40/79
----------
train Loss: 0.1400 Acc: 0.9568
val Loss: 1.7138 Acc: 0.6518

Epoch 41/79
----------
train Loss: 0.1394 Acc: 0.9576
val Loss: 1.8431 Acc: 0.6485

Epoch 42/79
----------
train Loss: 0.1379 Acc: 0.9571
val Loss: 1.8597 Acc: 0.6483

Epoch 43/79
----------
train Loss: 0.1395 Acc: 0.9572
val Loss: 1.8126 Acc: 0.6500

Epoch 44/79
----------
train Loss: 0.1308 Acc: 0.9598
val Loss: 1.8224 Acc: 0.6514

Epoch 45/79
----------
train Loss: 0.1312 Acc: 0.9598
val Loss: 1.9064 Acc: 0.6471

Epoch 46/79
----------
train Loss: 0.1287 Acc: 0.9603
val Loss: 1.9002 Acc: 0.6498

Epoch 47/79
----------
train Loss: 0.1255 Acc: 0.9612
val Loss: 1.8371 Acc: 0.6484

Epoch 48/79
----------
train Loss: 0.1263 Acc: 0.9603
val Loss: 1.9687 Acc: 0.6485
```

```
Epoch 49/79
----------
train Loss: 0.1263 Acc: 0.9602
val Loss: 1.8706 Acc: 0.6491

Epoch 50/79
----------
train Loss: 0.1224 Acc: 0.9625
val Loss: 1.9051 Acc: 0.6488

Epoch 51/79
----------
train Loss: 0.1263 Acc: 0.9611
val Loss: 1.8787 Acc: 0.6494

Epoch 52/79
----------
train Loss: 0.1232 Acc: 0.9617
val Loss: 1.8207 Acc: 0.6498

Epoch 53/79
----------
train Loss: 0.1197 Acc: 0.9627
val Loss: 1.9313 Acc: 0.6474

Epoch 54/79
----------
train Loss: 0.1249 Acc: 0.9612
val Loss: 1.8180 Acc: 0.6529

Epoch 55/79
----------
train Loss: 0.1192 Acc: 0.9638
val Loss: 1.8670 Acc: 0.6486

Epoch 56/79
----------
train Loss: 0.1169 Acc: 0.9638
val Loss: 1.8271 Acc: 0.6511

Epoch 57/79
----------
train Loss: 0.1176 Acc: 0.9640
val Loss: 1.9459 Acc: 0.6469

Epoch 58/79
----------
train Loss: 0.1178 Acc: 0.9637
val Loss: 2.0174 Acc: 0.6489

Epoch 59/79
----------
train Loss: 0.1187 Acc: 0.9636
val Loss: 2.0279 Acc: 0.6450

Epoch 60/79
----------
train Loss: 0.1129 Acc: 0.9655
val Loss: 1.8129 Acc: 0.6532

Epoch 61/79
----------
train Loss: 0.1144 Acc: 0.9651
val Loss: 1.8663 Acc: 0.6512

Epoch 62/79
```

```
----------
train Loss: 0.1118 Acc: 0.9666
val Loss: 1.8627 Acc: 0.6515

Epoch 63/79
----------
train Loss: 0.1156 Acc: 0.9650
val Loss: 1.7575 Acc: 0.6517

Epoch 64/79
----------
train Loss: 0.1135 Acc: 0.9658
val Loss: 1.9043 Acc: 0.6479

Epoch 65/79
----------
train Loss: 0.1150 Acc: 0.9649
val Loss: 1.8552 Acc: 0.6507

Epoch 66/79
----------
train Loss: 0.1180 Acc: 0.9635
val Loss: 1.7759 Acc: 0.6551

Epoch 67/79
----------
train Loss: 0.1150 Acc: 0.9646
val Loss: 1.8577 Acc: 0.6499

Epoch 68/79
----------
train Loss: 0.1187 Acc: 0.9637
val Loss: 1.8550 Acc: 0.6507

Epoch 69/79
----------
train Loss: 0.1160 Acc: 0.9649
val Loss: 2.1988 Acc: 0.6440

Epoch 70/79
----------
train Loss: 0.1141 Acc: 0.9649
val Loss: 1.8498 Acc: 0.6507

Epoch 71/79
----------
train Loss: 0.1156 Acc: 0.9645
val Loss: 2.0239 Acc: 0.6496

Epoch 72/79
----------
train Loss: 0.1153 Acc: 0.9651
val Loss: 2.0743 Acc: 0.6453

Epoch 73/79
----------
train Loss: 0.1146 Acc: 0.9652
val Loss: 1.7883 Acc: 0.6497

Epoch 74/79
----------
train Loss: 0.1094 Acc: 0.9657
val Loss: 1.7315 Acc: 0.6558

Epoch 75/79
----------
```

```
train Loss: 0.1142 Acc: 0.9649
val Loss: 2.0629 Acc: 0.6499

Epoch 76/79
----------
train Loss: 0.1132 Acc: 0.9646
val Loss: 1.8588 Acc: 0.6511

Epoch 77/79
----------
train Loss: 0.1149 Acc: 0.9647
val Loss: 1.8974 Acc: 0.6516

Epoch 78/79
----------
train Loss: 0.1124 Acc: 0.9665
val Loss: 1.8013 Acc: 0.6509

Epoch 79/79
----------
train Loss: 0.1136 Acc: 0.9647
val Loss: 1.9303 Acc: 0.6463

Training complete in 49m 38s
Best val Acc: 0.656800
```

- According to the output, the best val Acc decreased to 0.65, however, the training loss is now increased to around 0.96.

In [91]:

```python
# load model and change fc layer
learning_rate = 0.1
num_classes = len(train_data.classes)
model_ft = models.resnet50(pretrained=True)
num_ftrs = model_ft.fc.in_features
model_ft.fc = nn.Linear(num_ftrs, num_classes)
model_ft = model_ft.to(device)
criterion = nn.CrossEntropyLoss()
# Observe that all parameters are being optimized
optimizer_ft = optim.SGD(model_ft.parameters(), lr=learning_rate, momentum=0.9)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_ft, step_size=20, gamma=0.1)

model_ft = train_model(model_ft, criterion, optimizer_ft, exp_lr_scheduler,
                       num_epochs=80)
# Since the network start to converge at around 60 epochs, I trained the network for 80 ep
```

```
Epoch 0/79
----------
train Loss: 4.3980 Acc: 0.0412
val Loss: 3.9427 Acc: 0.0850

Epoch 1/79
----------
train Loss: 3.8636 Acc: 0.0950
val Loss: 3.8090 Acc: 0.1150

Epoch 2/79
----------
train Loss: 3.6908 Acc: 0.1247
val Loss: 3.5556 Acc: 0.1581

Epoch 3/79
----------
```

```
train Loss: 3.4870 Acc: 0.1600
val Loss: 3.4351 Acc: 0.1810

Epoch 4/79
----------
train Loss: 3.3370 Acc: 0.1864
val Loss: 3.1960 Acc: 0.2345

Epoch 5/79
----------
train Loss: 3.2218 Acc: 0.2094
val Loss: 3.3385 Acc: 0.2454

Epoch 6/79
----------
train Loss: 3.1062 Acc: 0.2305
val Loss: 3.2729 Acc: 0.2637

Epoch 7/79
----------
train Loss: 3.0380 Acc: 0.2438
val Loss: 2.9301 Acc: 0.2772

Epoch 8/79
----------
train Loss: 2.8995 Acc: 0.2719
val Loss: 2.8942 Acc: 0.2935

Epoch 9/79
----------
train Loss: 2.8098 Acc: 0.2839
val Loss: 3.0123 Acc: 0.3068

Epoch 10/79
----------
train Loss: 2.7267 Acc: 0.3020
val Loss: 5.0200 Acc: 0.3253

Epoch 11/79
----------
train Loss: 2.6736 Acc: 0.3131
val Loss: 3.7550 Acc: 0.1708

Epoch 12/79
----------
train Loss: 3.2325 Acc: 0.2119
val Loss: 2.8117 Acc: 0.2983

Epoch 13/79
----------
train Loss: 2.7403 Acc: 0.3004
val Loss: 2.6986 Acc: 0.3328

Epoch 14/79
----------
train Loss: 2.6065 Acc: 0.3272
val Loss: 2.6685 Acc: 0.3305

Epoch 15/79
----------
train Loss: 2.5087 Acc: 0.3458
val Loss: 2.5321 Acc: 0.3616

Epoch 16/79
----------
train Loss: 2.4389 Acc: 0.3631
```

```
val Loss: 2.5867 Acc: 0.3526

Epoch 17/79
----------
train Loss: 2.3999 Acc: 0.3699
val Loss: 2.4737 Acc: 0.3700

Epoch 18/79
----------
train Loss: 2.3240 Acc: 0.3844
val Loss: 2.4150 Acc: 0.3910

Epoch 19/79
----------
train Loss: 2.2697 Acc: 0.3947
val Loss: 2.4024 Acc: 0.3932

Epoch 20/79
----------
train Loss: 1.9637 Acc: 0.4692
val Loss: 2.2197 Acc: 0.4385

Epoch 21/79
----------
train Loss: 1.8688 Acc: 0.4915
val Loss: 2.2141 Acc: 0.4417

Epoch 22/79
----------
train Loss: 1.8322 Acc: 0.5014
val Loss: 2.1964 Acc: 0.4463

Epoch 23/79
----------
train Loss: 1.7952 Acc: 0.5077
val Loss: 2.2057 Acc: 0.4466

Epoch 24/79
----------
train Loss: 1.7593 Acc: 0.5146
val Loss: 2.2004 Acc: 0.4496

Epoch 25/79
----------
train Loss: 1.7366 Acc: 0.5205
val Loss: 2.1891 Acc: 0.4525

Epoch 26/79
----------
train Loss: 1.7248 Acc: 0.5238
val Loss: 2.1843 Acc: 0.4518

Epoch 27/79
----------
train Loss: 1.6974 Acc: 0.5320
val Loss: 2.1685 Acc: 0.4542

Epoch 28/79
----------
train Loss: 1.6789 Acc: 0.5340
val Loss: 2.1792 Acc: 0.4542

Epoch 29/79
----------
train Loss: 1.6621 Acc: 0.5408
val Loss: 2.1729 Acc: 0.4565
```

```
Epoch 30/79
----------
train Loss: 1.6337 Acc: 0.5451
val Loss: 2.1879 Acc: 0.4554

Epoch 31/79
----------
train Loss: 1.6091 Acc: 0.5531
val Loss: 2.1926 Acc: 0.4549

Epoch 32/79
----------
train Loss: 1.5819 Acc: 0.5591
val Loss: 2.1965 Acc: 0.4537

Epoch 33/79
----------
train Loss: 1.5773 Acc: 0.5589
val Loss: 2.1891 Acc: 0.4568

Epoch 34/79
----------
train Loss: 1.5462 Acc: 0.5681
val Loss: 2.1890 Acc: 0.4567

Epoch 35/79
----------
train Loss: 1.5338 Acc: 0.5689
val Loss: 2.1888 Acc: 0.4599

Epoch 36/79
----------
train Loss: 1.5121 Acc: 0.5748
val Loss: 2.2097 Acc: 0.4622

Epoch 37/79
----------
train Loss: 1.5106 Acc: 0.5750
val Loss: 2.2008 Acc: 0.4602

Epoch 38/79
----------
train Loss: 1.4693 Acc: 0.5839
val Loss: 2.1936 Acc: 0.4638

Epoch 39/79
----------
train Loss: 1.4419 Acc: 0.5921
val Loss: 2.1950 Acc: 0.4611

Epoch 40/79
----------
train Loss: 1.3825 Acc: 0.6084
val Loss: 2.1971 Acc: 0.4632

Epoch 41/79
----------
train Loss: 1.3620 Acc: 0.6165
val Loss: 2.1817 Acc: 0.4674

Epoch 42/79
----------
train Loss: 1.3548 Acc: 0.6159
val Loss: 2.1841 Acc: 0.4677
```

```
Epoch 43/79
----------
train Loss: 1.3489 Acc: 0.6199
val Loss: 2.1888 Acc: 0.4675

Epoch 44/79
----------
train Loss: 1.3491 Acc: 0.6188
val Loss: 2.1786 Acc: 0.4682

Epoch 45/79
----------
train Loss: 1.3375 Acc: 0.6233
val Loss: 2.1902 Acc: 0.4678

Epoch 46/79
----------
train Loss: 1.3358 Acc: 0.6236
val Loss: 2.1914 Acc: 0.4674

Epoch 47/79
----------
train Loss: 1.3315 Acc: 0.6226
val Loss: 2.1973 Acc: 0.4621

Epoch 48/79
----------
train Loss: 1.3242 Acc: 0.6241
val Loss: 2.1843 Acc: 0.4680

Epoch 49/79
----------
train Loss: 1.3233 Acc: 0.6230
val Loss: 2.1781 Acc: 0.4654

Epoch 50/79
----------
train Loss: 1.3234 Acc: 0.6227
val Loss: 2.1886 Acc: 0.4643

Epoch 51/79
----------
train Loss: 1.3129 Acc: 0.6258
val Loss: 2.1867 Acc: 0.4686

Epoch 52/79
----------
train Loss: 1.3127 Acc: 0.6260
val Loss: 2.1892 Acc: 0.4695

Epoch 53/79
----------
train Loss: 1.3018 Acc: 0.6293
val Loss: 2.1893 Acc: 0.4653

Epoch 54/79
----------
train Loss: 1.2995 Acc: 0.6299
val Loss: 2.1792 Acc: 0.4670

Epoch 55/79
----------
train Loss: 1.3051 Acc: 0.6294
val Loss: 2.1895 Acc: 0.4673

Epoch 56/79
```

```
----------
train Loss: 1.2905 Acc: 0.6340
val Loss: 2.1819 Acc: 0.4687

Epoch 57/79
----------
train Loss: 1.2944 Acc: 0.6334
val Loss: 2.1864 Acc: 0.4667

Epoch 58/79
----------
train Loss: 1.2871 Acc: 0.6352
val Loss: 2.1837 Acc: 0.4673

Epoch 59/79
----------
train Loss: 1.2870 Acc: 0.6347
val Loss: 2.2037 Acc: 0.4681

Epoch 60/79
----------
train Loss: 1.2697 Acc: 0.6398
val Loss: 2.1823 Acc: 0.4705

Epoch 61/79
----------
train Loss: 1.2641 Acc: 0.6411
val Loss: 2.1869 Acc: 0.4686

Epoch 62/79
----------
train Loss: 1.2649 Acc: 0.6413
val Loss: 2.2008 Acc: 0.4685

Epoch 63/79
----------
train Loss: 1.2685 Acc: 0.6410
val Loss: 2.2003 Acc: 0.4648

Epoch 64/79
----------
train Loss: 1.2671 Acc: 0.6402
val Loss: 2.1992 Acc: 0.4658

Epoch 65/79
----------
train Loss: 1.2661 Acc: 0.6400
val Loss: 2.1990 Acc: 0.4662

Epoch 66/79
----------
train Loss: 1.2617 Acc: 0.6410
val Loss: 2.1949 Acc: 0.4697

Epoch 67/79
----------
train Loss: 1.2617 Acc: 0.6416
val Loss: 2.1908 Acc: 0.4697

Epoch 68/79
----------
train Loss: 1.2605 Acc: 0.6447
val Loss: 2.2129 Acc: 0.4707

Epoch 69/79
----------
```

```
train Loss: 1.2672 Acc: 0.6399
val Loss: 2.1877 Acc: 0.4683

Epoch 70/79
----------
train Loss: 1.2640 Acc: 0.6391
val Loss: 2.1934 Acc: 0.4700

Epoch 71/79
----------
train Loss: 1.2628 Acc: 0.6404
val Loss: 2.1803 Acc: 0.4698

Epoch 72/79
----------
train Loss: 1.2593 Acc: 0.6406
val Loss: 2.1909 Acc: 0.4695

Epoch 73/79
----------
train Loss: 1.2644 Acc: 0.6386
val Loss: 2.1986 Acc: 0.4660

Epoch 74/79
----------
train Loss: 1.2649 Acc: 0.6391
val Loss: 2.1890 Acc: 0.4699

Epoch 75/79
----------
train Loss: 1.2626 Acc: 0.6403
val Loss: 2.1822 Acc: 0.4681

Epoch 76/79
----------
train Loss: 1.2658 Acc: 0.6408
val Loss: 2.2003 Acc: 0.4677

Epoch 77/79
----------
train Loss: 1.2588 Acc: 0.6408
val Loss: 2.1982 Acc: 0.4666

Epoch 78/79
----------
train Loss: 1.2619 Acc: 0.6415
val Loss: 2.2084 Acc: 0.4672

Epoch 79/79
----------
train Loss: 1.2607 Acc: 0.6386
val Loss: 2.2109 Acc: 0.4656

Training complete in 50m 1s
Best val Acc: 0.470700
```

- After increase initial lr to 0.1, the val acc decreased to 0.47, and training acc decreased yo 0.6. The performance is the worst among three settings.

(2)

In [10]:
```
# 0.001
device = 'cuda'
```

```
learning_rate = 0.001
model_conv = torchvision.models.resnet50(pretrained=True)
for param in model_conv.parameters():
    param.requires_grad = False

# Parameters of newly constructed modules have requires_grad=True by default
num_ftrs = model_conv.fc.in_features
num_classes = len(train_data.classes)
model_conv.fc = nn.Linear(num_ftrs, num_classes)

model_conv = model_conv.to(device)

criterion = nn.CrossEntropyLoss()

# Observe that only parameters of final layer are being optimized as
# opposed to before.
optimizer_conv = optim.SGD(model_conv.fc.parameters(), lr=learning_rate, momentum=0.9)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_conv, step_size=20, gamma=0.1)
model_conv = train_model(model_conv, criterion, optimizer_conv,
                         exp_lr_scheduler, num_epochs=80)

# final best accuracy is around 33%
```

```
Epoch 0/79
----------
train Loss: 4.0260 Acc: 0.1375
val Loss: 3.5425 Acc: 0.2220

Epoch 1/79
----------
train Loss: 3.4693 Acc: 0.2254
val Loss: 3.3003 Acc: 0.2540

Epoch 2/79
----------
train Loss: 3.2919 Acc: 0.2474
val Loss: 3.1521 Acc: 0.2762

Epoch 3/79
----------
train Loss: 3.1837 Acc: 0.2618
val Loss: 3.1051 Acc: 0.2852

Epoch 4/79
----------
train Loss: 3.1249 Acc: 0.2711
val Loss: 3.0424 Acc: 0.2914

Epoch 5/79
----------
train Loss: 3.0657 Acc: 0.2797
val Loss: 3.0407 Acc: 0.2940

Epoch 6/79
----------
train Loss: 3.0349 Acc: 0.2837
val Loss: 3.0350 Acc: 0.2972

Epoch 7/79
----------
train Loss: 3.0116 Acc: 0.2893
val Loss: 2.9836 Acc: 0.3003
```

```
Epoch 8/79
----------
train Loss: 2.9723 Acc: 0.2940
val Loss: 2.9369 Acc: 0.3091

Epoch 9/79
----------
train Loss: 2.9525 Acc: 0.2978
val Loss: 2.9310 Acc: 0.3113

Epoch 10/79
----------
train Loss: 2.9341 Acc: 0.2999
val Loss: 2.9557 Acc: 0.3085

Epoch 11/79
----------
train Loss: 2.9174 Acc: 0.3024
val Loss: 2.8637 Acc: 0.3176

Epoch 12/79
----------
train Loss: 2.9057 Acc: 0.3039
val Loss: 2.9345 Acc: 0.3115

Epoch 13/79
----------
train Loss: 2.8942 Acc: 0.3067
val Loss: 2.8662 Acc: 0.3167

Epoch 14/79
----------
train Loss: 2.8799 Acc: 0.3108
val Loss: 2.8651 Acc: 0.3182

Epoch 15/79
----------
train Loss: 2.8641 Acc: 0.3110
val Loss: 2.9094 Acc: 0.3159

Epoch 16/79
----------
train Loss: 2.8539 Acc: 0.3126
val Loss: 2.8836 Acc: 0.3195

Epoch 17/79
----------
train Loss: 2.8353 Acc: 0.3164
val Loss: 2.8791 Acc: 0.3209

Epoch 18/79
----------
train Loss: 2.8327 Acc: 0.3156
val Loss: 2.8251 Acc: 0.3239

Epoch 19/79
----------
train Loss: 2.8222 Acc: 0.3195
val Loss: 2.8224 Acc: 0.3229

Epoch 20/79
----------
train Loss: 2.7996 Acc: 0.3238
val Loss: 2.7968 Acc: 0.3288

Epoch 21/79
```

```
----------
train Loss: 2.7900 Acc: 0.3230
val Loss: 2.8363 Acc: 0.3247

Epoch 22/79
----------
train Loss: 2.7948 Acc: 0.3232
val Loss: 2.8343 Acc: 0.3284

Epoch 23/79
----------
train Loss: 2.7851 Acc: 0.3245
val Loss: 2.8219 Acc: 0.3262

Epoch 24/79
----------
train Loss: 2.7961 Acc: 0.3246
val Loss: 2.7785 Acc: 0.3364

Epoch 25/79
----------
train Loss: 2.7829 Acc: 0.3270
val Loss: 2.8266 Acc: 0.3290

Epoch 26/79
----------
train Loss: 2.7783 Acc: 0.3280
val Loss: 2.8008 Acc: 0.3320

Epoch 27/79
----------
train Loss: 2.7784 Acc: 0.3290
val Loss: 2.8514 Acc: 0.3243

Epoch 28/79
----------
train Loss: 2.7795 Acc: 0.3285
val Loss: 2.7823 Acc: 0.3311

Epoch 29/79
----------
train Loss: 2.7855 Acc: 0.3254
val Loss: 2.8295 Acc: 0.3226

Epoch 30/79
----------
train Loss: 2.7822 Acc: 0.3265
val Loss: 2.8025 Acc: 0.3262

Epoch 31/79
----------
train Loss: 2.7772 Acc: 0.3291
val Loss: 2.7981 Acc: 0.3309

Epoch 32/79
----------
train Loss: 2.7788 Acc: 0.3277
val Loss: 2.8074 Acc: 0.3323

Epoch 33/79
----------
train Loss: 2.7781 Acc: 0.3280
val Loss: 2.7840 Acc: 0.3293

Epoch 34/79
----------
```

```
train Loss: 2.7722 Acc: 0.3296
val Loss: 2.8161 Acc: 0.3264

Epoch 35/79
----------
train Loss: 2.7664 Acc: 0.3294
val Loss: 2.8226 Acc: 0.3266

Epoch 36/79
----------
train Loss: 2.7715 Acc: 0.3289
val Loss: 2.7861 Acc: 0.3339

Epoch 37/79
----------
train Loss: 2.7723 Acc: 0.3284
val Loss: 2.7835 Acc: 0.3276

Epoch 38/79
----------
train Loss: 2.7718 Acc: 0.3266
val Loss: 2.7704 Acc: 0.3337

Epoch 39/79
----------
train Loss: 2.7688 Acc: 0.3289
val Loss: 2.8057 Acc: 0.3308

Epoch 40/79
----------
train Loss: 2.7635 Acc: 0.3311
val Loss: 2.8092 Acc: 0.3313

Epoch 41/79
----------
train Loss: 2.7665 Acc: 0.3296
val Loss: 2.8101 Acc: 0.3318

Epoch 42/79
----------
train Loss: 2.7612 Acc: 0.3337
val Loss: 2.7690 Acc: 0.3319

Epoch 43/79
----------
train Loss: 2.7584 Acc: 0.3311
val Loss: 2.7902 Acc: 0.3279

Epoch 44/79
----------
train Loss: 2.7682 Acc: 0.3294
val Loss: 2.8170 Acc: 0.3304

Epoch 45/79
----------
train Loss: 2.7643 Acc: 0.3311
val Loss: 2.8027 Acc: 0.3331

Epoch 46/79
----------
train Loss: 2.7682 Acc: 0.3303
val Loss: 2.8111 Acc: 0.3280

Epoch 47/79
----------
train Loss: 2.7625 Acc: 0.3292
```

```
val Loss: 2.8069 Acc: 0.3304

Epoch 48/79
----------
train Loss: 2.7584 Acc: 0.3309
val Loss: 2.7823 Acc: 0.3330

Epoch 49/79
----------
train Loss: 2.7653 Acc: 0.3289
val Loss: 2.8078 Acc: 0.3309

Epoch 50/79
----------
train Loss: 2.7586 Acc: 0.3306
val Loss: 2.7895 Acc: 0.3321

Epoch 51/79
----------
train Loss: 2.7610 Acc: 0.3289
val Loss: 2.7757 Acc: 0.3338

Epoch 52/79
----------
train Loss: 2.7637 Acc: 0.3291
val Loss: 2.8215 Acc: 0.3251

Epoch 53/79
----------
train Loss: 2.7690 Acc: 0.3306
val Loss: 2.8033 Acc: 0.3278

Epoch 54/79
----------
train Loss: 2.7536 Acc: 0.3311
val Loss: 2.7722 Acc: 0.3361

Epoch 55/79
----------
train Loss: 2.7576 Acc: 0.3303
val Loss: 2.8309 Acc: 0.3306

Epoch 56/79
----------
train Loss: 2.7771 Acc: 0.3282
val Loss: 2.7899 Acc: 0.3277

Epoch 57/79
----------
train Loss: 2.7672 Acc: 0.3286
val Loss: 2.7616 Acc: 0.3334

Epoch 58/79
----------
train Loss: 2.7574 Acc: 0.3317
val Loss: 2.7850 Acc: 0.3294

Epoch 59/79
----------
train Loss: 2.7615 Acc: 0.3298
val Loss: 2.8193 Acc: 0.3272

Epoch 60/79
----------
train Loss: 2.7602 Acc: 0.3300
val Loss: 2.7950 Acc: 0.3280
```

```
Epoch 61/79
----------
train Loss: 2.7564 Acc: 0.3298
val Loss: 2.8136 Acc: 0.3276

Epoch 62/79
----------
train Loss: 2.7628 Acc: 0.3319
val Loss: 2.8145 Acc: 0.3264

Epoch 63/79
----------
train Loss: 2.7593 Acc: 0.3296
val Loss: 2.8085 Acc: 0.3305

Epoch 64/79
----------
train Loss: 2.7671 Acc: 0.3297
val Loss: 2.8214 Acc: 0.3274

Epoch 65/79
----------
train Loss: 2.7681 Acc: 0.3286
val Loss: 2.8198 Acc: 0.3290

Epoch 66/79
----------
train Loss: 2.7615 Acc: 0.3291
val Loss: 2.8006 Acc: 0.3287

Epoch 67/79
----------
train Loss: 2.7654 Acc: 0.3298
val Loss: 2.9451 Acc: 0.3221

Epoch 68/79
----------
train Loss: 2.7577 Acc: 0.3343
val Loss: 2.8253 Acc: 0.3258

Epoch 69/79
----------
train Loss: 2.7588 Acc: 0.3308
val Loss: 2.8482 Acc: 0.3258

Epoch 70/79
----------
train Loss: 2.7547 Acc: 0.3305
val Loss: 2.7868 Acc: 0.3301

Epoch 71/79
----------
train Loss: 2.7599 Acc: 0.3290
val Loss: 2.8080 Acc: 0.3306

Epoch 72/79
----------
train Loss: 2.7577 Acc: 0.3327
val Loss: 2.8069 Acc: 0.3304

Epoch 73/79
----------
train Loss: 2.7646 Acc: 0.3311
val Loss: 2.8152 Acc: 0.3310
```

```
Epoch 74/79
----------
train Loss: 2.7680 Acc: 0.3291
val Loss: 2.7854 Acc: 0.3339

Epoch 75/79
----------
train Loss: 2.7644 Acc: 0.3317
val Loss: 2.8201 Acc: 0.3300

Epoch 76/79
----------
train Loss: 2.7609 Acc: 0.3319
val Loss: 2.8664 Acc: 0.3281

Epoch 77/79
----------
train Loss: 2.7596 Acc: 0.3318
val Loss: 2.8039 Acc: 0.3293

Epoch 78/79
----------
train Loss: 2.7641 Acc: 0.3280
val Loss: 2.8662 Acc: 0.3245

Epoch 79/79
----------
train Loss: 2.7606 Acc: 0.3293
val Loss: 2.7897 Acc: 0.3330

Training complete in 25m 13s
Best val Acc: 0.336400
```

In [9]:

```
device = 'cuda'
learning_rate = 0.01
model_conv = torchvision.models.resnet50(pretrained=True)
for param in model_conv.parameters():
    param.requires_grad = False

# Parameters of newly constructed modules have requires_grad=True by default
num_ftrs = model_conv.fc.in_features
num_classes = len(train_data.classes)
model_conv.fc = nn.Linear(num_ftrs, num_classes)

model_conv = model_conv.to(device)

criterion = nn.CrossEntropyLoss()

# Observe that only parameters of final layer are being optimized as
# opposed to before.
optimizer_conv = optim.SGD(model_conv.fc.parameters(), lr=learning_rate, momentum=0.9)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_conv, step_size=20, gamma=0.1)
model_conv = train_model(model_conv, criterion, optimizer_conv,
                         exp_lr_scheduler, num_epochs=80)
```

```
C:\Learning\SoftWare\anaconda\lib\site-packages\torchvision\models\_utils.py:208: UserWarn
ing: The parameter 'pretrained' is deprecated since 0.13 and will be removed in 0.15, plea
se use 'weights' instead.
  warnings.warn(
C:\Learning\SoftWare\anaconda\lib\site-packages\torchvision\models\_utils.py:223: UserWarn
ing: Arguments other than a weight enum or `None` for 'weights' are deprecated since 0.13
and will be removed in 0.15. The current behavior is equivalent to passing `weights=ResNet
50_Weights.IMAGENET1K_V1`. You can also use `weights=ResNet50_Weights.DEFAULT` to get the
```

```
most up-to-date weights.
  warnings.warn(msg)
Epoch 0/79
----------
train Loss: 3.6140 Acc: 0.1942
val Loss: 3.3752 Acc: 0.2605

Epoch 1/79
----------
train Loss: 3.3117 Acc: 0.2505
val Loss: 3.2624 Acc: 0.2714

Epoch 2/79
----------
train Loss: 3.2358 Acc: 0.2627
val Loss: 3.2394 Acc: 0.2758

Epoch 3/79
----------
train Loss: 3.1739 Acc: 0.2721
val Loss: 3.1900 Acc: 0.2885

Epoch 4/79
----------
train Loss: 3.1435 Acc: 0.2774
val Loss: 3.1966 Acc: 0.2886

Epoch 5/79
----------
train Loss: 3.1102 Acc: 0.2846
val Loss: 3.1951 Acc: 0.2923

Epoch 6/79
----------
train Loss: 3.0675 Acc: 0.2913
val Loss: 3.1304 Acc: 0.3043

Epoch 7/79
----------
train Loss: 3.0605 Acc: 0.2955
val Loss: 3.1638 Acc: 0.2977

Epoch 8/79
----------
train Loss: 3.0446 Acc: 0.2955
val Loss: 3.2089 Acc: 0.3013

Epoch 9/79
----------
train Loss: 3.0123 Acc: 0.3021
val Loss: 3.1293 Acc: 0.3026

Epoch 10/79
----------
train Loss: 3.0117 Acc: 0.3035
val Loss: 3.0754 Acc: 0.3100

Epoch 11/79
----------
train Loss: 3.0013 Acc: 0.3052
val Loss: 3.2645 Acc: 0.3003

Epoch 12/79
----------
train Loss: 2.9842 Acc: 0.3069
val Loss: 3.0408 Acc: 0.3131
```

```
Epoch 13/79
----------
train Loss: 2.9855 Acc: 0.3087
val Loss: 3.1600 Acc: 0.3038

Epoch 14/79
----------
train Loss: 2.9885 Acc: 0.3093
val Loss: 3.1356 Acc: 0.3028

Epoch 15/79
----------
train Loss: 2.9711 Acc: 0.3096
val Loss: 3.1614 Acc: 0.3044

Epoch 16/79
----------
train Loss: 2.9714 Acc: 0.3121
val Loss: 3.0738 Acc: 0.3139

Epoch 17/79
----------
train Loss: 2.9575 Acc: 0.3154
val Loss: 3.0455 Acc: 0.3217

Epoch 18/79
----------
train Loss: 2.9458 Acc: 0.3146
val Loss: 3.1123 Acc: 0.3149

Epoch 19/79
----------
train Loss: 2.9497 Acc: 0.3154
val Loss: 3.1307 Acc: 0.3127

Epoch 20/79
----------
train Loss: 2.6829 Acc: 0.3500
val Loss: 2.8040 Acc: 0.3414

Epoch 21/79
----------
train Loss: 2.6179 Acc: 0.3571
val Loss: 2.8041 Acc: 0.3457

Epoch 22/79
----------
train Loss: 2.5959 Acc: 0.3617
val Loss: 2.8307 Acc: 0.3455

Epoch 23/79
----------
train Loss: 2.5809 Acc: 0.3650
val Loss: 2.7038 Acc: 0.3524

Epoch 24/79
----------
train Loss: 2.5690 Acc: 0.3649
val Loss: 2.7295 Acc: 0.3508

Epoch 25/79
----------
train Loss: 2.5490 Acc: 0.3661
val Loss: 2.7175 Acc: 0.3516
```

```
Epoch 26/79
----------
train Loss: 2.5564 Acc: 0.3656
val Loss: 2.7359 Acc: 0.3544

Epoch 27/79
----------
train Loss: 2.5403 Acc: 0.3681
val Loss: 2.7276 Acc: 0.3526

Epoch 28/79
----------
train Loss: 2.5315 Acc: 0.3727
val Loss: 2.7668 Acc: 0.3506

Epoch 29/79
----------
train Loss: 2.5245 Acc: 0.3737
val Loss: 2.6858 Acc: 0.3526

Epoch 30/79
----------
train Loss: 2.5309 Acc: 0.3711
val Loss: 2.6538 Acc: 0.3599

Epoch 31/79
----------
train Loss: 2.5164 Acc: 0.3736
val Loss: 2.6699 Acc: 0.3540

Epoch 32/79
----------
train Loss: 2.5184 Acc: 0.3712
val Loss: 2.6730 Acc: 0.3523

Epoch 33/79
----------
train Loss: 2.5152 Acc: 0.3742
val Loss: 2.6574 Acc: 0.3592

Epoch 34/79
----------
train Loss: 2.5104 Acc: 0.3749
val Loss: 2.6805 Acc: 0.3607

Epoch 35/79
----------
train Loss: 2.5064 Acc: 0.3746
val Loss: 2.6679 Acc: 0.3584

Epoch 36/79
----------
train Loss: 2.5092 Acc: 0.3752
val Loss: 2.7224 Acc: 0.3490

Epoch 37/79
----------
train Loss: 2.5107 Acc: 0.3733
val Loss: 2.6511 Acc: 0.3607

Epoch 38/79
----------
train Loss: 2.5097 Acc: 0.3734
val Loss: 2.7355 Acc: 0.3527

Epoch 39/79
```

```
----------
train Loss: 2.5063 Acc: 0.3724
val Loss: 2.6485 Acc: 0.3598

Epoch 40/79
----------
train Loss: 2.4889 Acc: 0.3802
val Loss: 2.6799 Acc: 0.3585

Epoch 41/79
----------
train Loss: 2.4775 Acc: 0.3787
val Loss: 2.6574 Acc: 0.3614

Epoch 42/79
----------
train Loss: 2.4847 Acc: 0.3780
val Loss: 2.6409 Acc: 0.3605

Epoch 43/79
----------
train Loss: 2.4724 Acc: 0.3814
val Loss: 2.6555 Acc: 0.3597

Epoch 44/79
----------
train Loss: 2.4763 Acc: 0.3806
val Loss: 2.6792 Acc: 0.3593

Epoch 45/79
----------
train Loss: 2.4822 Acc: 0.3791
val Loss: 2.6583 Acc: 0.3597

Epoch 46/79
----------
train Loss: 2.4822 Acc: 0.3806
val Loss: 2.6621 Acc: 0.3604

Epoch 47/79
----------
train Loss: 2.4866 Acc: 0.3776
val Loss: 2.6511 Acc: 0.3610

Epoch 48/79
----------
train Loss: 2.4828 Acc: 0.3779
val Loss: 2.6845 Acc: 0.3575

Epoch 49/79
----------
train Loss: 2.4671 Acc: 0.3800
val Loss: 2.6592 Acc: 0.3628

Epoch 50/79
----------
train Loss: 2.4835 Acc: 0.3783
val Loss: 2.6806 Acc: 0.3589

Epoch 51/79
----------
train Loss: 2.4791 Acc: 0.3793
val Loss: 2.6233 Acc: 0.3607

Epoch 52/79
----------
```

```
train Loss: 2.4736 Acc: 0.3807
val Loss: 2.6723 Acc: 0.3593

Epoch 53/79
----------
train Loss: 2.4723 Acc: 0.3811
val Loss: 2.6627 Acc: 0.3594

Epoch 54/79
----------
train Loss: 2.4705 Acc: 0.3800
val Loss: 2.6644 Acc: 0.3605

Epoch 55/79
----------
train Loss: 2.4759 Acc: 0.3821
val Loss: 2.6852 Acc: 0.3614

Epoch 56/79
----------
train Loss: 2.4772 Acc: 0.3813
val Loss: 2.7123 Acc: 0.3542

Epoch 57/79
----------
train Loss: 2.4774 Acc: 0.3785
val Loss: 2.6567 Acc: 0.3584

Epoch 58/79
----------
train Loss: 2.4704 Acc: 0.3845
val Loss: 2.6448 Acc: 0.3580

Epoch 59/79
----------
train Loss: 2.4651 Acc: 0.3832
val Loss: 2.6875 Acc: 0.3557

Epoch 60/79
----------
train Loss: 2.4720 Acc: 0.3802
val Loss: 2.6245 Acc: 0.3628

Epoch 61/79
----------
train Loss: 2.4678 Acc: 0.3820
val Loss: 2.6304 Acc: 0.3614

Epoch 62/79
----------
train Loss: 2.4646 Acc: 0.3826
val Loss: 2.6548 Acc: 0.3595

Epoch 63/79
----------
train Loss: 2.4844 Acc: 0.3792
val Loss: 2.6537 Acc: 0.3566

Epoch 64/79
----------
train Loss: 2.4654 Acc: 0.3817
val Loss: 2.6485 Acc: 0.3564

Epoch 65/79
----------
train Loss: 2.4661 Acc: 0.3808
```

```
val Loss: 2.6198 Acc: 0.3656

Epoch 66/79
----------
train Loss: 2.4654 Acc: 0.3817
val Loss: 2.6237 Acc: 0.3640

Epoch 67/79
----------
train Loss: 2.4639 Acc: 0.3839
val Loss: 2.6589 Acc: 0.3595

Epoch 68/79
----------
train Loss: 2.4511 Acc: 0.3843
val Loss: 2.6011 Acc: 0.3651

Epoch 69/79
----------
train Loss: 2.4588 Acc: 0.3852
val Loss: 2.6765 Acc: 0.3555

Epoch 70/79
----------
train Loss: 2.4770 Acc: 0.3791
val Loss: 2.6275 Acc: 0.3590

Epoch 71/79
----------
train Loss: 2.4654 Acc: 0.3810
val Loss: 2.6347 Acc: 0.3622

Epoch 72/79
----------
train Loss: 2.4523 Acc: 0.3843
val Loss: 2.6816 Acc: 0.3580

Epoch 73/79
----------
train Loss: 2.4687 Acc: 0.3825
val Loss: 2.6588 Acc: 0.3541

Epoch 74/79
----------
train Loss: 2.4741 Acc: 0.3833
val Loss: 2.6298 Acc: 0.3620

Epoch 75/79
----------
train Loss: 2.4636 Acc: 0.3811
val Loss: 2.6774 Acc: 0.3644

Epoch 76/79
----------
train Loss: 2.4641 Acc: 0.3851
val Loss: 2.6472 Acc: 0.3580

Epoch 77/79
----------
train Loss: 2.4655 Acc: 0.3835
val Loss: 2.6526 Acc: 0.3605

Epoch 78/79
----------
train Loss: 2.4682 Acc: 0.3817
val Loss: 2.6572 Acc: 0.3633
```

```
Epoch 79/79
----------
train Loss: 2.4794 Acc: 0.3820
val Loss: 2.6469 Acc: 0.3561

Training complete in 25m 20s
Best val Acc: 0.365600
```

In [10]:

```
device = 'cuda'
learning_rate = 0.1
model_conv = torchvision.models.resnet50(pretrained=True)
for param in model_conv.parameters():
    param.requires_grad = False

# Parameters of newly constructed modules have requires_grad=True by default
num_ftrs = model_conv.fc.in_features
num_classes = len(train_data.classes)
model_conv.fc = nn.Linear(num_ftrs, num_classes)

model_conv = model_conv.to(device)

criterion = nn.CrossEntropyLoss()

# Observe that only parameters of final layer are being optimized as
# opposed to before.
optimizer_conv = optim.SGD(model_conv.fc.parameters(), lr=learning_rate, momentum=0.9)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_conv, step_size=20, gamma=0.1)
model_conv = train_model(model_conv, criterion, optimizer_conv,
                         exp_lr_scheduler, num_epochs=80)
```

```
Epoch 0/79
----------
train Loss: 12.6432 Acc: 0.1422
val Loss: 14.1633 Acc: 0.1846

Epoch 1/79
----------
train Loss: 13.7550 Acc: 0.1733
val Loss: 14.7592 Acc: 0.2010

Epoch 2/79
----------
train Loss: 13.8763 Acc: 0.1908
val Loss: 13.8143 Acc: 0.2169

Epoch 3/79
----------
train Loss: 14.1087 Acc: 0.1966
val Loss: 14.6169 Acc: 0.2122

Epoch 4/79
----------
train Loss: 14.0261 Acc: 0.2070
val Loss: 14.9029 Acc: 0.2143

Epoch 5/79
----------
train Loss: 14.2065 Acc: 0.2082
val Loss: 13.9815 Acc: 0.2240

Epoch 6/79
----------
```

```
train Loss: 14.2643 Acc: 0.2133
val Loss: 15.1695 Acc: 0.2199

Epoch 7/79
----------
train Loss: 14.2622 Acc: 0.2161
val Loss: 14.9281 Acc: 0.2184

Epoch 8/79
----------
train Loss: 14.2773 Acc: 0.2184
val Loss: 15.1200 Acc: 0.2283

Epoch 9/79
----------
train Loss: 14.1678 Acc: 0.2236
val Loss: 15.9021 Acc: 0.2331

Epoch 10/79
----------
train Loss: 14.3557 Acc: 0.2262
val Loss: 16.9870 Acc: 0.2295

Epoch 11/79
----------
train Loss: 14.3424 Acc: 0.2262
val Loss: 15.5009 Acc: 0.2369

Epoch 12/79
----------
train Loss: 14.4716 Acc: 0.2260
val Loss: 15.8424 Acc: 0.2265

Epoch 13/79
----------
train Loss: 14.6237 Acc: 0.2257
val Loss: 16.1297 Acc: 0.2355

Epoch 14/79
----------
train Loss: 14.4664 Acc: 0.2300
val Loss: 14.5840 Acc: 0.2473

Epoch 15/79
----------
train Loss: 14.4088 Acc: 0.2291
val Loss: 14.4897 Acc: 0.2439

Epoch 16/79
----------
train Loss: 14.6246 Acc: 0.2317
val Loss: 15.3104 Acc: 0.2341

Epoch 17/79
----------
train Loss: 14.5524 Acc: 0.2288
val Loss: 15.8910 Acc: 0.2392

Epoch 18/79
----------
train Loss: 14.5678 Acc: 0.2339
val Loss: 15.9342 Acc: 0.2465

Epoch 19/79
----------
train Loss: 14.5794 Acc: 0.2353
```

```
val Loss: 16.9225 Acc: 0.2233

Epoch 20/79
----------
train Loss: 9.8296 Acc: 0.2922
val Loss: 10.0602 Acc: 0.2924

Epoch 21/79
----------
train Loss: 8.1507 Acc: 0.3091
val Loss: 9.0631 Acc: 0.2999

Epoch 22/79
----------
train Loss: 7.4109 Acc: 0.3123
val Loss: 8.3354 Acc: 0.3025

Epoch 23/79
----------
train Loss: 6.9785 Acc: 0.3192
val Loss: 7.8997 Acc: 0.3093

Epoch 24/79
----------
train Loss: 6.6353 Acc: 0.3192
val Loss: 7.3756 Acc: 0.3040

Epoch 25/79
----------
train Loss: 6.4145 Acc: 0.3255
val Loss: 7.2698 Acc: 0.3069

Epoch 26/79
----------
train Loss: 6.2254 Acc: 0.3203
val Loss: 7.3521 Acc: 0.3047

Epoch 27/79
----------
train Loss: 6.0775 Acc: 0.3200
val Loss: 6.9644 Acc: 0.3013

Epoch 28/79
----------
train Loss: 5.9739 Acc: 0.3197
val Loss: 7.0524 Acc: 0.3050

Epoch 29/79
----------
train Loss: 5.8055 Acc: 0.3210
val Loss: 6.8607 Acc: 0.3033

Epoch 30/79
----------
train Loss: 5.6759 Acc: 0.3240
val Loss: 6.5332 Acc: 0.3002

Epoch 31/79
----------
train Loss: 5.6525 Acc: 0.3212
val Loss: 6.7297 Acc: 0.3007

Epoch 32/79
----------
train Loss: 5.5499 Acc: 0.3204
val Loss: 6.3438 Acc: 0.3033
```

```
Epoch 33/79
----------
train Loss: 5.4376 Acc: 0.3252
val Loss: 6.5092 Acc: 0.3013

Epoch 34/79
----------
train Loss: 5.3801 Acc: 0.3213
val Loss: 6.0117 Acc: 0.3067

Epoch 35/79
----------
train Loss: 5.2941 Acc: 0.3227
val Loss: 5.9675 Acc: 0.3043

Epoch 36/79
----------
train Loss: 5.2060 Acc: 0.3250
val Loss: 6.0594 Acc: 0.3068

Epoch 37/79
----------
train Loss: 5.1152 Acc: 0.3243
val Loss: 6.0229 Acc: 0.2969

Epoch 38/79
----------
train Loss: 5.0579 Acc: 0.3255
val Loss: 5.8499 Acc: 0.3045

Epoch 39/79
----------
train Loss: 4.9990 Acc: 0.3237
val Loss: 5.6803 Acc: 0.3029

Epoch 40/79
----------
train Loss: 4.6953 Acc: 0.3366
val Loss: 5.6633 Acc: 0.3120

Epoch 41/79
----------
train Loss: 4.6003 Acc: 0.3428
val Loss: 5.5826 Acc: 0.3142

Epoch 42/79
----------
train Loss: 4.6005 Acc: 0.3372
val Loss: 5.4411 Acc: 0.3131

Epoch 43/79
----------
train Loss: 4.5642 Acc: 0.3361
val Loss: 5.3956 Acc: 0.3152

Epoch 44/79
----------
train Loss: 4.5538 Acc: 0.3412
val Loss: 5.4001 Acc: 0.3134

Epoch 45/79
----------
train Loss: 4.4985 Acc: 0.3448
val Loss: 5.3632 Acc: 0.3150
```

```
Epoch 46/79
----------
train Loss: 4.5345 Acc: 0.3388
val Loss: 5.4199 Acc: 0.3104

Epoch 47/79
----------
train Loss: 4.4694 Acc: 0.3438
val Loss: 5.4040 Acc: 0.3164

Epoch 48/79
----------
train Loss: 4.4425 Acc: 0.3431
val Loss: 5.4939 Acc: 0.3116

Epoch 49/79
----------
train Loss: 4.4571 Acc: 0.3401
val Loss: 5.2897 Acc: 0.3095

Epoch 50/79
----------
train Loss: 4.4602 Acc: 0.3420
val Loss: 5.2736 Acc: 0.3092

Epoch 51/79
----------
train Loss: 4.4283 Acc: 0.3419
val Loss: 5.2725 Acc: 0.3129

Epoch 52/79
----------
train Loss: 4.3996 Acc: 0.3420
val Loss: 5.2381 Acc: 0.3132

Epoch 53/79
----------
train Loss: 4.3962 Acc: 0.3425
val Loss: 5.4842 Acc: 0.3111

Epoch 54/79
----------
train Loss: 4.3960 Acc: 0.3432
val Loss: 5.3509 Acc: 0.3187

Epoch 55/79
----------
train Loss: 4.3918 Acc: 0.3402
val Loss: 5.3170 Acc: 0.3122

Epoch 56/79
----------
train Loss: 4.3807 Acc: 0.3415
val Loss: 5.2735 Acc: 0.3174

Epoch 57/79
----------
train Loss: 4.3736 Acc: 0.3393
val Loss: 5.1466 Acc: 0.3179

Epoch 58/79
----------
train Loss: 4.3713 Acc: 0.3428
val Loss: 5.1042 Acc: 0.3152

Epoch 59/79
```

```
----------
train Loss: 4.3632 Acc: 0.3422
val Loss: 5.3903 Acc: 0.3159

Epoch 60/79
----------
train Loss: 4.3189 Acc: 0.3451
val Loss: 5.2740 Acc: 0.3138

Epoch 61/79
----------
train Loss: 4.3270 Acc: 0.3405
val Loss: 5.1189 Acc: 0.3154

Epoch 62/79
----------
train Loss: 4.3101 Acc: 0.3447
val Loss: 5.4317 Acc: 0.3152

Epoch 63/79
----------
train Loss: 4.3343 Acc: 0.3413
val Loss: 5.1700 Acc: 0.3194

Epoch 64/79
----------
train Loss: 4.3128 Acc: 0.3457
val Loss: 5.1636 Acc: 0.3136

Epoch 65/79
----------
train Loss: 4.3621 Acc: 0.3421
val Loss: 5.1936 Acc: 0.3169

Epoch 66/79
----------
train Loss: 4.2919 Acc: 0.3442
val Loss: 5.3494 Acc: 0.3137

Epoch 67/79
----------
train Loss: 4.3178 Acc: 0.3430
val Loss: 5.3047 Acc: 0.3128

Epoch 68/79
----------
train Loss: 4.2965 Acc: 0.3457
val Loss: 5.0850 Acc: 0.3174

Epoch 69/79
----------
train Loss: 4.3247 Acc: 0.3425
val Loss: 5.2177 Acc: 0.3137

Epoch 70/79
----------
train Loss: 4.3115 Acc: 0.3443
val Loss: 5.1631 Acc: 0.3192

Epoch 71/79
----------
train Loss: 4.2784 Acc: 0.3456
val Loss: 5.0566 Acc: 0.3228

Epoch 72/79
----------
```

```
train Loss: 4.3327 Acc: 0.3432
val Loss: 5.1289 Acc: 0.3173

Epoch 73/79
----------
train Loss: 4.3123 Acc: 0.3430
val Loss: 5.2158 Acc: 0.3128

Epoch 74/79
----------
train Loss: 4.3011 Acc: 0.3407
val Loss: 5.2100 Acc: 0.3162

Epoch 75/79
----------
train Loss: 4.2622 Acc: 0.3489
val Loss: 5.1763 Acc: 0.3135

Epoch 76/79
----------
train Loss: 4.3196 Acc: 0.3456
val Loss: 5.5305 Acc: 0.3067

Epoch 77/79
----------
train Loss: 4.2841 Acc: 0.3435
val Loss: 5.4023 Acc: 0.3134

Epoch 78/79
----------
train Loss: 4.2884 Acc: 0.3444
val Loss: 5.3504 Acc: 0.3140

Epoch 79/79
----------
train Loss: 4.3099 Acc: 0.3451
val Loss: 5.2231 Acc: 0.3141

Training complete in 25m 19s
Best val Acc: 0.322800
```

In [11]:

```
device = 'cuda'
learning_rate = 1
model_conv = torchvision.models.resnet50(pretrained=True)
for param in model_conv.parameters():shangcihainengyong asd
    param.requires_grad = False

# Parameters of newly constructed modules have requires_grad=True by default
num_ftrs = model_conv.fc.in_features
num_classes = len(train_data.classes)
model_conv.fc = nn.Linear(num_ftrs, num_classes)

model_conv = model_conv.to(device)

criterion = nn.CrossEntropyLoss()

# Observe that only parameters of final layer are being optimized as
# opposed to before.
optimizer_conv = optim.SGD(model_conv.fc.parameters(), lr=learning_rate, momentum=0.9)

# Decay LR by a factor of 0.1 every 7 epochs
exp_lr_scheduler = lr_scheduler.StepLR(optimizer_conv, step_size=20, gamma=0.1)
model_conv = train_model(model_conv, criterion, optimizer_conv,
                         exp_lr_scheduler, num_epochs=80)
```

```
Epoch 0/79
----------
train Loss: 147.2992 Acc: 0.1316
val Loss: 155.1323 Acc: 0.1716

Epoch 1/79
----------
train Loss: 149.7803 Acc: 0.1721
val Loss: 160.0972 Acc: 0.1996

Epoch 2/79
----------
train Loss: 149.8357 Acc: 0.1860
val Loss: 156.4338 Acc: 0.1978

Epoch 3/79
----------
train Loss: 149.3519 Acc: 0.1936
val Loss: 159.0560 Acc: 0.2070

Epoch 4/79
----------
train Loss: 152.3486 Acc: 0.1995
val Loss: 167.1318 Acc: 0.2143

Epoch 5/79
----------
train Loss: 152.3136 Acc: 0.2065
val Loss: 157.8685 Acc: 0.2245

Epoch 6/79
----------
train Loss: 153.2660 Acc: 0.2109
val Loss: 160.6240 Acc: 0.2263

Epoch 7/79
----------
train Loss: 152.1547 Acc: 0.2125
val Loss: 144.9858 Acc: 0.2272

Epoch 8/79
----------
train Loss: 153.7249 Acc: 0.2147
val Loss: 154.4628 Acc: 0.2245

Epoch 9/79
----------
train Loss: 152.8693 Acc: 0.2174
val Loss: 182.9579 Acc: 0.2195

Epoch 10/79
----------
train Loss: 155.2207 Acc: 0.2195
val Loss: 181.2355 Acc: 0.2213

Epoch 11/79
----------
train Loss: 153.6516 Acc: 0.2252
val Loss: 180.9163 Acc: 0.2150

Epoch 12/79
----------
train Loss: 154.2816 Acc: 0.2258
val Loss: 173.4472 Acc: 0.2324

Epoch 13/79
```

```
----------
train Loss: 155.7639 Acc: 0.2244
val Loss: 172.4547 Acc: 0.2320

Epoch 14/79
----------
train Loss: 154.6159 Acc: 0.2291
val Loss: 162.6903 Acc: 0.2301

Epoch 15/79
----------
train Loss: 155.2597 Acc: 0.2283
val Loss: 168.1758 Acc: 0.2386

Epoch 16/79
----------
train Loss: 154.2294 Acc: 0.2293
val Loss: 173.9113 Acc: 0.2286

Epoch 17/79
----------
train Loss: 155.5157 Acc: 0.2289
val Loss: 170.1232 Acc: 0.2380

Epoch 18/79
----------
train Loss: 156.4090 Acc: 0.2321
val Loss: 165.8125 Acc: 0.2256

Epoch 19/79
----------
train Loss: 157.4974 Acc: 0.2302
val Loss: 161.3655 Acc: 0.2304

Epoch 20/79
----------
train Loss: 104.8162 Acc: 0.2903
val Loss: 110.6011 Acc: 0.2861

Epoch 21/79
----------
train Loss: 85.8178 Acc: 0.3047
val Loss: 94.4777 Acc: 0.2983

Epoch 22/79
----------
train Loss: 77.1242 Acc: 0.3143
val Loss: 87.3669 Acc: 0.3001

Epoch 23/79
----------
train Loss: 72.9185 Acc: 0.3141
val Loss: 83.5284 Acc: 0.3015

Epoch 24/79
----------
train Loss: 69.6343 Acc: 0.3125
val Loss: 81.6403 Acc: 0.2960

Epoch 25/79
----------
train Loss: 66.5407 Acc: 0.3179
val Loss: 74.7218 Acc: 0.3040

Epoch 26/79
----------
```

```
train Loss: 65.1050 Acc: 0.3113
val Loss: 76.3743 Acc: 0.2939

Epoch 27/79
----------
train Loss: 62.5500 Acc: 0.3139
val Loss: 74.6775 Acc: 0.2924

Epoch 28/79
----------
train Loss: 60.8655 Acc: 0.3169
val Loss: 70.4087 Acc: 0.2969

Epoch 29/79
----------
train Loss: 59.8373 Acc: 0.3127
val Loss: 67.3548 Acc: 0.3001

Epoch 30/79
----------
train Loss: 58.7354 Acc: 0.3134
val Loss: 65.9403 Acc: 0.3011

Epoch 31/79
----------
train Loss: 57.4295 Acc: 0.3128
val Loss: 68.6005 Acc: 0.2984

Epoch 32/79
----------
train Loss: 56.3085 Acc: 0.3154
val Loss: 70.1109 Acc: 0.2879

Epoch 33/79
----------
train Loss: 55.5275 Acc: 0.3124
val Loss: 63.7695 Acc: 0.2959

Epoch 34/79
----------
train Loss: 54.3400 Acc: 0.3133
val Loss: 64.9716 Acc: 0.2954

Epoch 35/79
----------
train Loss: 53.2619 Acc: 0.3118
val Loss: 61.0139 Acc: 0.2988

Epoch 36/79
----------
train Loss: 52.6067 Acc: 0.3148
val Loss: 62.1975 Acc: 0.2960

Epoch 37/79
----------
train Loss: 51.6412 Acc: 0.3154
val Loss: 60.2937 Acc: 0.2979

Epoch 38/79
----------
train Loss: 50.5247 Acc: 0.3128
val Loss: 56.9280 Acc: 0.2970

Epoch 39/79
----------
train Loss: 49.9255 Acc: 0.3157
```

```
val Loss: 59.1389 Acc: 0.2963

Epoch 40/79
----------
train Loss: 46.0064 Acc: 0.3257
val Loss: 56.1543 Acc: 0.3038

Epoch 41/79
----------
train Loss: 45.3331 Acc: 0.3299
val Loss: 55.8480 Acc: 0.3059

Epoch 42/79
----------
train Loss: 44.7831 Acc: 0.3280
val Loss: 54.9329 Acc: 0.3049

Epoch 43/79
----------
train Loss: 44.2185 Acc: 0.3329
val Loss: 54.2186 Acc: 0.3058

Epoch 44/79
----------
train Loss: 44.2757 Acc: 0.3320
val Loss: 52.2560 Acc: 0.3004

Epoch 45/79
----------
train Loss: 44.0900 Acc: 0.3321
val Loss: 53.2037 Acc: 0.3054

Epoch 46/79
----------
train Loss: 43.8885 Acc: 0.3297
val Loss: 54.7980 Acc: 0.3019

Epoch 47/79
----------
train Loss: 43.4443 Acc: 0.3332
val Loss: 51.7588 Acc: 0.3086

Epoch 48/79
----------
train Loss: 43.7752 Acc: 0.3296
val Loss: 53.3111 Acc: 0.3042

Epoch 49/79
----------
train Loss: 43.1308 Acc: 0.3298
val Loss: 50.7364 Acc: 0.3024

Epoch 50/79
----------
train Loss: 42.4810 Acc: 0.3336
val Loss: 51.5421 Acc: 0.3055

Epoch 51/79
----------
train Loss: 42.2361 Acc: 0.3348
val Loss: 49.9206 Acc: 0.3095

Epoch 52/79
----------
train Loss: 42.5470 Acc: 0.3302
val Loss: 51.7173 Acc: 0.3071
```

```
Epoch 53/79
----------
train Loss: 42.5026 Acc: 0.3331
val Loss: 51.5513 Acc: 0.3047

Epoch 54/79
----------
train Loss: 42.1903 Acc: 0.3315
val Loss: 51.2049 Acc: 0.3059

Epoch 55/79
----------
train Loss: 42.4931 Acc: 0.3287
val Loss: 50.0174 Acc: 0.3058

Epoch 56/79
----------
train Loss: 42.1288 Acc: 0.3320
val Loss: 51.0666 Acc: 0.3073

Epoch 57/79
----------
train Loss: 42.2221 Acc: 0.3304
val Loss: 51.2969 Acc: 0.3066

Epoch 58/79
----------
train Loss: 42.0567 Acc: 0.3305
val Loss: 52.1907 Acc: 0.3041

Epoch 59/79
----------
train Loss: 42.1196 Acc: 0.3317
val Loss: 52.9345 Acc: 0.3013

Epoch 60/79
----------
train Loss: 41.6383 Acc: 0.3321
val Loss: 50.2769 Acc: 0.3089

Epoch 61/79
----------
train Loss: 41.2622 Acc: 0.3330
val Loss: 53.8953 Acc: 0.3022

Epoch 62/79
----------
train Loss: 41.1455 Acc: 0.3353
val Loss: 50.6952 Acc: 0.3069

Epoch 63/79
----------
train Loss: 41.1285 Acc: 0.3319
val Loss: 51.6775 Acc: 0.3035

Epoch 64/79
----------
train Loss: 41.3798 Acc: 0.3319
val Loss: 48.6957 Acc: 0.3078

Epoch 65/79
----------
train Loss: 40.8150 Acc: 0.3374
val Loss: 50.6367 Acc: 0.2998
```

```
Epoch 66/79
----------
train Loss: 41.2728 Acc: 0.3346
val Loss: 50.7267 Acc: 0.3033

Epoch 67/79
----------
train Loss: 41.1271 Acc: 0.3351
val Loss: 50.3186 Acc: 0.3075

Epoch 68/79
----------
train Loss: 41.3596 Acc: 0.3331
val Loss: 49.6629 Acc: 0.3065

Epoch 69/79
----------
train Loss: 41.0766 Acc: 0.3337
val Loss: 50.7021 Acc: 0.3063

Epoch 70/79
----------
train Loss: 41.0759 Acc: 0.3314
val Loss: 48.8922 Acc: 0.3128

Epoch 71/79
----------
train Loss: 41.1567 Acc: 0.3339
val Loss: 51.3471 Acc: 0.3047

Epoch 72/79
----------
train Loss: 41.3470 Acc: 0.3333
val Loss: 49.7098 Acc: 0.3070

Epoch 73/79
----------
train Loss: 41.0605 Acc: 0.3324
val Loss: 49.9151 Acc: 0.3097

Epoch 74/79
----------
train Loss: 40.9780 Acc: 0.3342
val Loss: 50.3521 Acc: 0.3037

Epoch 75/79
----------
train Loss: 41.0363 Acc: 0.3354
val Loss: 49.3573 Acc: 0.3080

Epoch 76/79
----------
train Loss: 40.9961 Acc: 0.3378
val Loss: 50.8670 Acc: 0.3062

Epoch 77/79
----------
train Loss: 40.8289 Acc: 0.3370
val Loss: 49.9994 Acc: 0.3088

Epoch 78/79
----------
train Loss: 40.7757 Acc: 0.3344
val Loss: 49.0239 Acc: 0.3097

Epoch 79/79
```

```
----------
train Loss: 41.2263 Acc: 0.3326
val Loss: 52.4486 Acc: 0.3001

Training complete in 25m 37s
Best val Acc: 0.312800
```

In [16]:
```
table = pd.DataFrame({'base_learning_rate':[0.001,0.01,0.1,1],
              'FineTuneing_val':[0.675400,0.656800,0.470700,None],
              'FineTuning_train':[0.9482,0.9647,0.6386, None],
              'FixConv_val':[0.336400,0.365600,0.322800,0.312800],
              'FixConv_train':[0.3293,0.3820,0.3451,0.3326]}).transpose()
table
```

Out[16]:

	0	1	2	3
base_learning_rate	0.0010	0.0100	0.1000	1.0000
FineTuneing_val	0.6754	0.6568	0.4707	NaN
FineTuning_train	0.9482	0.9647	0.6386	NaN
FixConv_val	0.3364	0.3656	0.3228	0.3128
FixConv_train	0.3293	0.3820	0.3451	0.3326

- According to the table, Finetuning with proper learning rate can significantly improving the performance, in my experiment, the best learning rate is 0.01 for finetuing. Although over fitting, train accuracy is still higher than fixed feature extractor.
- For fixed feature extractor, the best learning rate is still 0.01, training acc and val acc is at the same level, around 37% .

Problem 2

(1)

- weakly supervised： The author treats preprocess hashtages as supervised training label and pre-train it using supervised way, and finally can use the fine-tune the pretrained model on sepecifice tasks (for example finne tune the model on Imagenet dataset to do classfication).
- semi-supervised: Unlike the weakly supervised learning, the authon didn't use the hashtag information in semi-supervised learning, instead, he trained a model using Imagenet as a teacher model and create pseudo label for the 1B image dataset. And then use pseudo labeled data to pretrain the student label and finally fine tune the model using the Imagenet dataset to do classification.

(2)

(a)

- Model robustness.
    - training on large-scale hashtag data is unexpectedly robust to label noise, and that the features learned allow a simple linear classifier to achieve state-of-the-art ImageNet-1k top-1 accuracy of 83.6% without any finetuning (compared to 84.2% with finetuning)

(b)

- Why is resampling of hashtag distribution important?
    - Hashtag frequencies follow a Zipfian distribution. For example, in the 17k hashtag vocabulary, the most frequent hashtag (#fineart) appears more than 1 million times as often as the least frequent hashtag (#shirtfront). Training on the natural distribution may not be optimal and therefore we consider two alternative ways to sample

(3)

(a) why are there two models, why it is a distillation

- Teacher model: teacher model is used to create pseudo labels on the large unlabeled data set.
- Student model: student model is first learned from the teacher model and then fine-tune the model for our task. Training the student model by the pseudo-labeled data can let the student model learn patterns from the large dataset and leverage the pattern on tasks.
- Because Distillation can be seen as a particular case of self-training, in that the teacher model makes prediction on unlabelled data, and the inferred labels are used to train the student in a supervised fashion. And this kinda the same as the work did in the paper.

(b) K&P

- P: P highest scores are retained after pseudo training.
- K: Top-K images from from each class if taken to create a new dataset.
- P>1: If P=1 we will lose the information about an image's second possible class information. Say we have an image it's most possible class is dog and second possible class is cat, if we select p=2 the cat pattern in the image will be captured by student model so we will have a more robust model.

(c)

- After we trained a teacher model, when we input an unlabeled image, it will out put a vector of possibilities.(100 dim if the teacher model is trained on Imagenet) And will will select top P possible labels as the pseudo label.
- Since P>1 an image will have several pseudo labels. after truncate by K, it's possible to see an image belong to more than one classes.

(d)

- Increase: In the early stage, when we increase K, included images have high possiblity belong to one class, there images have better quality.
- Decrease: When K is large enough, some noisy images are included, in this stage increase introduces more noise than useful information so the accuracy decrease.

Problem 3

(1)

- **Why hard to achieve peak FLOPS:** However, this assumption is unreasonable in practice. For instance, achieving peak FLOPS is a difficult proposition, usually requiring customized libraries developed by organizations with intimate knowledge of the underlying hardware, e.g., Intel's MKL (int, 2009), ATLAS (Whaley & Petitet, 2005), and cuDNN. Even these specially tuned libraries may fall short of peak execution by as much as 40% (atl). Further, any computation done outside the scope of PALEO (e.g. job scheduling, data copying) will exacerbate the observed inefficiency in practice. Sometimes such inefficiencies are warranted from the perspective of ease of programmability or maintenance of the learning platforms.

- **How PPP capture this:** PPP captures average relative inefficiency of the platfor(hardware+framework) compared to peak theoretical FLOPS.

(2)

- VGG19 has three more convolutional layers compared with VGG-16.
- The additional Conv operations = $56*56*256*3*3*256+28*28*512*3*3*512+14*14*512*3*3*512$ $=4161.798144M$
- And the previouse Conv operations in VGG-16 is 15360M, So the total conv operations of VGG-19 is 15360M+4162M = 19522M
- So the final propotion of VGG-19 is 19522M/(15503M+4162M) = 99.3%

(3)

- Why didn't match on GPU: Because TK1 and TX1 have different communcation and measurement overhead.
- What approach was adopted in Sec. 5 of the paper to mitigate the measurement overhead in GPUs: Timing measurements on GPUs can only been recorded only all cores finish their tasks. In a full forward pass, timing is only recorded at the last layer. terefore, a core may be assigned with the computation of following layers and thus it can continuously perform the computation without synchronization.

(4)

- K80 config: 1.87*10e12
- VGG-16
  - FLOPS: 1.5503*10e10
  - Time_per_forward: 1.5503*10e10 /1.87*10e12 = 0.0083 s
  - Throughput: 1/0.0083 = 120
- Google-net
  - FLOPS: 1.606*10e9
  - Time_per_forward: 1.606*10e9 /1.87*10e12 = 0.00086 s
  - Throughput: 1/0.00086 = 1164
- Resnet-50:
  - FLOPS: 3.922*10e9
  - Time_per_forward: 3.922*10e9 /1.87*10e12 = 0.0021 s
  - Throughput: 1/0.0021 = 476

Problem 4

In [68]:
```
from rest_net import *
def save_result(model, loss, model_name):
    with open(model_name+'_3070.pickle', 'wb') as handle:
        pickle.dump(loss, handle, protocol=pickle.HIGHEST_PROTOCOL)
    torch.save(model.state_dict(), model_name+'_3070.model')
```

In [61]:
```
transform_train = transforms.Compose([
    transforms.RandomCrop(32, padding=4),
    transforms.RandomHorizontalFlip(),
    transforms.ToTensor(),
    transforms.Normalize(mean=[0.485, 0.456, 0.406],
                                         std=[0.229, 0.224, 0.225]),
])
```

In [62]:
```
train_data = torchvision.datasets.CIFAR10(download=True,root="./data",transform=transform_
test_data = torchvision.datasets.CIFAR10(root="./data",train=False,transform=transform_tra
```

```
Files already downloaded and verified
```

In [53]:
```
device = 'cuda'
model = models.resnet18()
model.to(device)
trainloader = torch.utils.data.DataLoader(train_data, batch_size=128,
                                             shuffle=True, num_workers=4)
optimizer = optim.SGD(model.parameters(), lr=0.01, momentum=0.9, weight_decay=5e-4)
lr_scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer, T_max=200)
criterion = nn.CrossEntropyLoss()
final_loss = []
for epoch in range(350):  # loop over the dataset multiple times
    print(f'current epoch {epoch}', end = '\n')

    running_loss = 0.0
    total_batch = len(trainloader)
    for i, data in enumerate(trainloader, 0):
        if i < len(trainloader)-1:
            print(f'current batch: {i}, total batch: {total_batch}', end='\r')
        else:
            print(f'current batch: {i}, total batch: {total_batch}')
        # get the inputs; data is a list of [inputs, labels]
        inputs, labels = data

        # zero the parameter gradients

        inputs = inputs.to(device)
        labels = labels.to(device)
        # zero the parameter gradients
        optimizer.zero_grad()

        outputs = model(inputs)
        _, preds = torch.max(outputs, 1)
        loss = criterion(outputs, labels)

        loss.backward()
        optimizer.step()
        final_loss.append(loss.item())

        # statistics
        lr_scheduler.step()

print('Finished Training')
save_result(model, final_loss, 'resnet18')
```

```
current epoch 0
current batch: 390, total batch: 391
current epoch 1
current batch: 390, total batch: 391
current epoch 2
current batch: 390, total batch: 391
current epoch 3
current batch: 390, total batch: 391
current epoch 4
current batch: 390, total batch: 391
current epoch 5
current batch: 390, total batch: 391
current epoch 6
current batch: 390, total batch: 391
current epoch 7
current batch: 390, total batch: 391
current epoch 8
current batch: 390, total batch: 391
current epoch 9
current batch: 390, total batch: 391
current epoch 10
current batch: 390, total batch: 391
current epoch 11
current batch: 390, total batch: 391
current epoch 12
current batch: 390, total batch: 391
current epoch 13
current batch: 390, total batch: 391
current epoch 14
current batch: 390, total batch: 391
current epoch 15
current batch: 390, total batch: 391
current epoch 16
current batch: 390, total batch: 391
current epoch 17
current batch: 390, total batch: 391
current epoch 18
current batch: 390, total batch: 391
current epoch 19
current batch: 390, total batch: 391
current epoch 20
current batch: 390, total batch: 391
current epoch 21
current batch: 390, total batch: 391
current epoch 22
current batch: 390, total batch: 391
current epoch 23
current batch: 390, total batch: 391
current epoch 24
current batch: 390, total batch: 391
current epoch 25
current batch: 390, total batch: 391
current epoch 26
current batch: 390, total batch: 391
current epoch 27
current batch: 390, total batch: 391
current epoch 28
current batch: 390, total batch: 391
current epoch 29
current batch: 390, total batch: 391
current epoch 30
current batch: 390, total batch: 391
current epoch 31
current batch: 390, total batch: 391
current epoch 32
current batch: 390, total batch: 391
current epoch 33
current batch: 390, total batch: 391
current epoch 34
current batch: 390, total batch: 391
current epoch 35
current batch: 390, total batch: 391
current epoch 36
current batch: 390, total batch: 391
current epoch 37
current batch: 390, total batch: 391
current epoch 38
current batch: 390, total batch: 391
current epoch 39
current batch: 390, total batch: 391
current epoch 40
current batch: 390, total batch: 391
current epoch 41
current batch: 390, total batch: 391
current epoch 42
current batch: 390, total batch: 391
current epoch 43
current batch: 390, total batch: 391
current epoch 44
current batch: 390, total batch: 391
current epoch 45
current batch: 390, total batch: 391
current epoch 46
current batch: 390, total batch: 391
current epoch 47
current batch: 390, total batch: 391
current epoch 48
current batch: 390, total batch: 391
current epoch 49
current batch: 390, total batch: 391
current epoch 50
current batch: 390, total batch: 391
current epoch 51
current batch: 390, total batch: 391
current epoch 52
current batch: 390, total batch: 391
current epoch 53
current batch: 390, total batch: 391
current epoch 54
current batch: 390, total batch: 391
current epoch 55
current batch: 390, total batch: 391
current epoch 56
current batch: 390, total batch: 391
current epoch 57
current batch: 390, total batch: 391
current epoch 58
current batch: 390, total batch: 391
current epoch 59
current batch: 390, total batch: 391
current epoch 60
current batch: 390, total batch: 391
current epoch 61
current batch: 390, total batch: 391
current epoch 62
current batch: 390, total batch: 391
current epoch 63
current batch: 390, total batch: 391
current epoch 64
current batch: 390, total batch: 391
current epoch 65
current batch: 390, total batch: 391
current epoch 66
current batch: 390, total batch: 391
current epoch 67
current batch: 390, total batch: 391
current epoch 68
current batch: 390, total batch: 391
current epoch 69
current batch: 390, total batch: 391
current epoch 70
current batch: 390, total batch: 391
current epoch 71
current batch: 390, total batch: 391
current epoch 72
current batch: 390, total batch: 391
current epoch 73
current batch: 390, total batch: 391
current epoch 74
current batch: 390, total batch: 391
current epoch 75
current batch: 390, total batch: 391
current epoch 76
current batch: 390, total batch: 391
current epoch 77
current batch: 390, total batch: 391
current epoch 78
current batch: 390, total batch: 391
current epoch 79
current batch: 390, total batch: 391
current epoch 80
current batch: 390, total batch: 391
current epoch 81
current batch: 390, total batch: 391
current epoch 82
current batch: 390, total batch: 391
current epoch 83
current batch: 390, total batch: 391
current epoch 84
current batch: 390, total batch: 391
current epoch 85
current batch: 390, total batch: 391
current epoch 86
current batch: 390, total batch: 391
current epoch 87
current batch: 390, total batch: 391
current epoch 88
current batch: 390, total batch: 391
current epoch 89
current batch: 390, total batch: 391
current epoch 90
current batch: 390, total batch: 391
current epoch 91
current batch: 390, total batch: 391
current epoch 92
current batch: 390, total batch: 391
current epoch 93
current batch: 390, total batch: 391
current epoch 94
current batch: 390, total batch: 391
current epoch 95
current batch: 390, total batch: 391
current epoch 96
current batch: 390, total batch: 391
current epoch 97
current batch: 390, total batch: 391
current epoch 98
current batch: 390, total batch: 391
current epoch 99
current batch: 390, total batch: 391
current epoch 100
current batch: 390, total batch: 391
current epoch 101
current batch: 390, total batch: 391
current epoch 102
current batch: 390, total batch: 391
current epoch 103
current batch: 390, total batch: 391
current epoch 104
current batch: 390, total batch: 391
current epoch 105
current batch: 390, total batch: 391
current epoch 106
current batch: 390, total batch: 391
current epoch 107
current batch: 390, total batch: 391
current epoch 108
current batch: 390, total batch: 391
current epoch 109
current batch: 390, total batch: 391
current epoch 110
current batch: 390, total batch: 391
current epoch 111
current batch: 390, total batch: 391
current epoch 112
current batch: 390, total batch: 391
current epoch 113
current batch: 390, total batch: 391
current epoch 114
current batch: 390, total batch: 391
current epoch 115
current batch: 390, total batch: 391
current epoch 116
current batch: 390, total batch: 391
current epoch 117
current batch: 390, total batch: 391
current epoch 118
current batch: 390, total batch: 391
current epoch 119
current batch: 390, total batch: 391
current epoch 120
current batch: 390, total batch: 391
current epoch 121
current batch: 390, total batch: 391
current epoch 122
current batch: 390, total batch: 391
current epoch 123
current batch: 390, total batch: 391
current epoch 124
current batch: 390, total batch: 391
current epoch 125
current batch: 390, total batch: 391
current epoch 126
current batch: 390, total batch: 391
current epoch 127
current batch: 390, total batch: 391
current epoch 128
current batch: 390, total batch: 391
current epoch 129
current batch: 390, total batch: 391
current epoch 130
current batch: 390, total batch: 391
current epoch 131
current batch: 390, total batch: 391
current epoch 132
current batch: 390, total batch: 391
current epoch 133
current batch: 390, total batch: 391
current epoch 134
current batch: 390, total batch: 391
current epoch 135
current batch: 390, total batch: 391
current epoch 136
current batch: 390, total batch: 391
current epoch 137
current batch: 390, total batch: 391
current epoch 138
current batch: 390, total batch: 391
current epoch 139
current batch: 390, total batch: 391
current epoch 140
current batch: 390, total batch: 391
current epoch 141
current batch: 390, total batch: 391
current epoch 142
current batch: 390, total batch: 391
current epoch 143
current batch: 390, total batch: 391
current epoch 144
current batch: 390, total batch: 391
current epoch 145
current batch: 390, total batch: 391
current epoch 146
current batch: 390, total batch: 391
current epoch 147
current batch: 390, total batch: 391
current epoch 148
current batch: 390, total batch: 391
current epoch 149
current batch: 390, total batch: 391
current epoch 150
current batch: 390, total batch: 391
current epoch 151
current batch: 390, total batch: 391
current epoch 152
current batch: 390, total batch: 391
current epoch 153
current batch: 390, total batch: 391
current epoch 154
current batch: 390, total batch: 391
current epoch 155
current batch: 390, total batch: 391
current epoch 156
current batch: 390, total batch: 391
current epoch 157
current batch: 390, total batch: 391
current epoch 158
current batch: 390, total batch: 391
current epoch 159
current batch: 390, total batch: 391
current epoch 160
current batch: 390, total batch: 391
current epoch 161
current batch: 390, total batch: 391
current epoch 162
current batch: 390, total batch: 391
current epoch 163
current batch: 390, total batch: 391
current epoch 164
current batch: 390, total batch: 391
current epoch 165
current batch: 390, total batch: 391
current epoch 166
current batch: 390, total batch: 391
current epoch 167
current batch: 390, total batch: 391
current epoch 168
current batch: 390, total batch: 391
current epoch 169
current batch: 390, total batch: 391
current epoch 170
current batch: 390, total batch: 391
current epoch 171
current batch: 390, total batch: 391
current epoch 172
current batch: 390, total batch: 391
current epoch 173
current batch: 390, total batch: 391
current epoch 174
current batch: 390, total batch: 391
current epoch 175
current batch: 390, total batch: 391
current epoch 176
current batch: 390, total batch: 391
current epoch 177
current batch: 390, total batch: 391
current epoch 178
current batch: 390, total batch: 391
current epoch 179
current batch: 390, total batch: 391
current epoch 180
current batch: 390, total batch: 391
current epoch 181
current batch: 390, total batch: 391
current epoch 182
current batch: 390, total batch: 391
current epoch 183
current batch: 390, total batch: 391
current epoch 184
current batch: 390, total batch: 391
current epoch 185
current batch: 390, total batch: 391
current epoch 186
current batch: 390, total batch: 391
current epoch 187
current batch: 390, total batch: 391
current epoch 188
current batch: 390, total batch: 391
current epoch 189
current batch: 390, total batch: 391
current epoch 190
current batch: 390, total batch: 391
current epoch 191
current batch: 390, total batch: 391
current epoch 192
current batch: 390, total batch: 391
current epoch 193
current batch: 390, total batch: 391
current epoch 194
current batch: 390, total batch: 391
current epoch 195
current batch: 390, total batch: 391
current epoch 196
current batch: 390, total batch: 391
current epoch 197
current batch: 390, total batch: 391
current epoch 198
current batch: 390, total batch: 391
current epoch 199
current batch: 390, total batch: 391
current epoch 200
current batch: 390, total batch: 391
current epoch 201
current batch: 390, total batch: 391
current epoch 202
current batch: 390, total batch: 391
current epoch 203
current batch: 390, total batch: 391
current epoch 204
current batch: 390, total batch: 391
current epoch 205
current batch: 390, total batch: 391
current epoch 206
current batch: 390, total batch: 391
current epoch 207
current batch: 390, total batch: 391
current epoch 208
current batch: 390, total batch: 391
current epoch 209
current batch: 390, total batch: 391
current epoch 210
current batch: 390, total batch: 391
current epoch 211
current batch: 390, total batch: 391
current epoch 212
current batch: 390, total batch: 391
current epoch 213
current batch: 390, total batch: 391
current epoch 214
current batch: 390, total batch: 391
current epoch 215
current batch: 390, total batch: 391
current epoch 216
current batch: 390, total batch: 391
current epoch 217
current batch: 390, total batch: 391
current epoch 218
current batch: 390, total batch: 391
current epoch 219
current batch: 390, total batch: 391
current epoch 220
current batch: 390, total batch: 391
current epoch 221
current batch: 390, total batch: 391
current epoch 222
current batch: 390, total batch: 391
current epoch 223
current batch: 390, total batch: 391
current epoch 224
current batch: 390, total batch: 391
current epoch 225
current batch: 390, total batch: 391
current epoch 226
current batch: 390, total batch: 391
current epoch 227
current batch: 390, total batch: 391
current epoch 228
current batch: 390, total batch: 391
current epoch 229
current batch: 390, total batch: 391
current epoch 230
current batch: 390, total batch: 391
current epoch 231
current batch: 390, total batch: 391
current epoch 232
current batch: 390, total batch: 391
current epoch 233
current batch: 390, total batch: 391
current epoch 234
current batch: 390, total batch: 391
current epoch 235
current batch: 390, total batch: 391
current epoch 236
current batch: 390, total batch: 391
current epoch 237
current batch: 390, total batch: 391
current epoch 238
current batch: 390, total batch: 391
current epoch 239
current batch: 390, total batch: 391
current epoch 240
current batch: 390, total batch: 391
current epoch 241
current batch: 390, total batch: 391
current epoch 242
current batch: 390, total batch: 391
current epoch 243
current batch: 390, total batch: 391
current epoch 244
current batch: 390, total batch: 391
current epoch 245
current batch: 390, total batch: 391
current epoch 246
current batch: 390, total batch: 391
current epoch 247
current batch: 390, total batch: 391
current epoch 248
current batch: 390, total batch: 391
current epoch 249
current batch: 390, total batch: 391
current epoch 250
current batch: 390, total batch: 391
current epoch 251
current batch: 390, total batch: 391
current epoch 252
current batch: 390, total batch: 391
current epoch 253
current batch: 390, total batch: 391
current epoch 254
current batch: 390, total batch: 391
current epoch 255
current batch: 390, total batch: 391
current epoch 256
current batch: 390, total batch: 391
current epoch 257
current batch: 390, total batch: 391
current epoch 258
current batch: 390, total batch: 391
current epoch 259
current batch: 390, total batch: 391
current epoch 260
current batch: 390, total batch: 391
current epoch 261
current batch: 390, total batch: 391
current epoch 262
current batch: 390, total batch: 391
current epoch 263
current batch: 390, total batch: 391
current epoch 264
current batch: 390, total batch: 391
current epoch 265
current batch: 390, total batch: 391
current epoch 266
current batch: 390, total batch: 391
current epoch 267
current batch: 390, total batch: 391
current epoch 268
current batch: 390, total batch: 391
current epoch 269
current batch: 390, total batch: 391
current epoch 270
current batch: 390, total batch: 391
current epoch 271
current batch: 390, total batch: 391
current epoch 272
current batch: 390, total batch: 391
current epoch 273
current batch: 390, total batch: 391
current epoch 274
current batch: 390, total batch: 391
current epoch 275
current batch: 390, total batch: 391
current epoch 276
current batch: 390, total batch: 391
current epoch 277
current batch: 390, total batch: 391
current epoch 278
current batch: 390, total batch: 391
current epoch 279
current batch: 390, total batch: 391
current epoch 280
current batch: 390, total batch: 391
current epoch 281
current batch: 390, total batch: 391
current epoch 282
current batch: 390, total batch: 391
current epoch 283
current batch: 390, total batch: 391
current epoch 284
current batch: 390, total batch: 391
current epoch 285
current batch: 390, total batch: 391
current epoch 286
current batch: 390, total batch: 391
current epoch 287
current batch: 390, total batch: 391
current epoch 288
current batch: 390, total batch: 391
current epoch 289
current batch: 390, total batch: 391
current epoch 290
current batch: 390, total batch: 391
current epoch 291
current batch: 390, total batch: 391
current epoch 292
current batch: 390, total batch: 391
current epoch 293
current batch: 390, total batch: 391
current epoch 294
current batch: 390, total batch: 391
current epoch 295
current batch: 390, total batch: 391
current epoch 296
current batch: 390, total batch: 391
current epoch 297
current batch: 390, total batch: 391
current epoch 298
current batch: 390, total batch: 391
current epoch 299
current batch: 390, total batch: 391
current epoch 300
current batch: 390, total batch: 391
current epoch 301
current batch: 390, total batch: 391
current epoch 302
current batch: 390, total batch: 391
current epoch 303
current batch: 390, total batch: 391
current epoch 304
current batch: 390, total batch: 391
current epoch 305
current batch: 390, total batch: 391
current epoch 306
current batch: 390, total batch: 391
current epoch 307
current batch: 390, total batch: 391
current epoch 308
current batch: 390, total batch: 391
current epoch 309
current batch: 390, total batch: 391
current epoch 310
current batch: 390, total batch: 391
current epoch 311
current batch: 390, total batch: 391
current epoch 312
current batch: 390, total batch: 391
current epoch 313
current batch: 390, total batch: 391
current epoch 314
current batch: 390, total batch: 391
current epoch 315
current batch: 390, total batch: 391
current epoch 316
current batch: 390, total batch: 391
current epoch 317
current batch: 390, total batch: 391
current epoch 318
current batch: 390, total batch: 391
current epoch 319
current batch: 390, total batch: 391
current epoch 320
current batch: 390, total batch: 391
current epoch 330
current batch: 390, total batch: 391
current epoch 331
current batch: 390, total batch: 391
current epoch 332
current batch: 390, total batch: 391
current epoch 333
current batch: 390, total batch: 391
current epoch 334
current batch: 390, total batch: 391
current epoch 335
current batch: 390, total batch: 391
current epoch 336
current batch: 390, total batch: 391
current epoch 337
current batch: 390, total batch: 391
current epoch 338
current batch: 390, total batch: 391
current epoch 339
current batch: 390, total batch: 391
current epoch 340
current batch: 390, total batch: 391
current epoch 341
current batch: 390, total batch: 391
current epoch 342
current batch: 390, total batch: 391
current epoch 343
current batch: 390, total batch: 391
current epoch 344
current batch: 390, total batch: 391
current epoch 345
current batch: 390, total batch: 391
current epoch 346
current batch: 390, total batch: 391
current epoch 347
current batch: 390, total batch: 391
current epoch 348
current batch: 390, total batch: 391
current epoch 349
current batch: 390, total batch: 391
Finished Training
```

In [67]:

```
device = 'cuda'
model = resnet20()
model.to(device)
trainloader = torch.utils.data.DataLoader(train_data, batch_size=128,
                                          shuffle=True, num_workers=4)
optimizer = optim.SGD(model.parameters(), lr=0.01, momentum=0.9, weight_decay=5e-4)
lr_scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer, T_max=200)
criterion = nn.CrossEntropyLoss()
final_loss = []
for epoch in range(350):  # loop over the dataset multiple times
    print(f'current epoch {epoch}', end = '\n')

    running_loss = 0.0
    total_batch = len(trainloader)
    for i, data in enumerate(trainloader, 0):
        if i < len(trainloader)-1:
            print(f'current batch: {i}, total batch: {total_batch}', end='\r')
        else:
            print(f'current batch: {i}, total batch: {total_batch}')
        # get the inputs; data is a list of [inputs, labels]
        inputs, labels = data

        # zero the parameter gradients

        inputs = inputs.to(device)
        labels = labels.to(device)
        # zero the parameter gradients
        optimizer.zero_grad()

        outputs = model(inputs)
        _, preds = torch.max(outputs, 1)
        loss = criterion(outputs, labels)

        loss.backward()
        optimizer.step()
        final_loss.append(loss.item())

        # statistics
        lr_scheduler.step()

print('Finished Training')
save_result(model, final_loss, 'resnet20')
```

```
current epoch 0
current batch: 390, total batch: 391
current epoch 1
current batch: 390, total batch: 391
current epoch 2
current batch: 390, total batch: 391
current epoch 3
current batch: 390, total batch: 391
current epoch 4
current batch: 390, total batch: 391
current epoch 5
current batch: 390, total batch: 391
current epoch 6
current batch: 390, total batch: 391
current epoch 7
current batch: 390, total batch: 391
current epoch 8
current batch: 390, total batch: 391
current epoch 9
current batch: 390, total batch: 391
current epoch 10
current batch: 390, total batch: 391
current epoch 11
current batch: 390, total batch: 391
current epoch 12
current batch: 390, total batch: 391
current epoch 13
current batch: 390, total batch: 391
current epoch 14
current batch: 390, total batch: 391
current epoch 15
current batch: 390, total batch: 391
current epoch 16
current batch: 390, total batch: 391
current epoch 17
current batch: 390, total batch: 391
current epoch 18
current batch: 390, total batch: 391
current epoch 19
current batch: 390, total batch: 391
current epoch 20
current batch: 390, total batch: 391
current epoch 21
current batch: 390, total batch: 391
current epoch 22
current batch: 390, total batch: 391
current epoch 23
current batch: 390, total batch: 391
current epoch 24
current batch: 390, total batch: 391
current epoch 25
current batch: 390, total batch: 391
current epoch 26
current batch: 390, total batch: 391
current epoch 27
current batch: 390, total batch: 391
current epoch 28
current batch: 390, total batch: 391
current epoch 29
current batch: 390, total batch: 391
current epoch 30
current batch: 390, total batch: 391
current epoch 31
current batch: 390, total batch: 391
current epoch 32
current batch: 390, total batch: 391
current epoch 33
current batch: 390, total batch: 391
current epoch 34
current batch: 390, total batch: 391
current epoch 35
current batch: 390, total batch: 391
current epoch 36
current batch: 390, total batch: 391
current epoch 37
current batch: 390, total batch: 391
current epoch 38
current batch: 390, total batch: 391
current epoch 39
current batch: 390, total batch: 391
current epoch 40
current batch: 390, total batch: 391
current epoch 41
current batch: 390, total batch: 391
current epoch 42
current batch: 390, total batch: 391
current epoch 43
current batch: 390, total batch: 391
current epoch 44
current batch: 390, total batch: 391
current epoch 45
current batch: 390, total batch: 391
current epoch 46
current batch: 390, total batch: 391
current epoch 47
current batch: 390, total batch: 391
current epoch 48
current batch: 390, total batch: 391
current epoch 49
current batch: 390, total batch: 391
current epoch 50
current batch: 390, total batch: 391
current epoch 51
current batch: 390, total batch: 391
current epoch 52
current batch: 390, total batch: 391
current epoch 53
current batch: 390, total batch: 391
current epoch 54
current batch: 390, total batch: 391
current epoch 55
current batch: 390, total batch: 391
current epoch 56
current batch: 390, total batch: 391
current epoch 57
current batch: 390, total batch: 391
current epoch 58
current batch: 390, total batch: 391
current epoch 59
current batch: 390, total batch: 391
current epoch 60
current batch: 390, total batch: 391
current epoch 61
current batch: 390, total batch: 391
current epoch 62
current batch: 390, total batch: 391
current epoch 63
current batch: 390, total batch: 391
current epoch 64
current batch: 390, total batch: 391
current epoch 65
current batch: 390, total batch: 391
current epoch 66
current batch: 390, total batch: 391
current epoch 67
current batch: 390, total batch: 391
current epoch 68
current batch: 390, total batch: 391
current epoch 69
current batch: 390, total batch: 391
current epoch 70
current batch: 390, total batch: 391
current epoch 71
current batch: 390, total batch: 391
current epoch 72
current batch: 390, total batch: 391
current epoch 73
current batch: 390, total batch: 391
current epoch 74
current batch: 390, total batch: 391
current epoch 75
current batch: 390, total batch: 391
current epoch 76
current batch: 390, total batch: 391
current epoch 77
current batch: 390, total batch: 391
current epoch 78
current batch: 390, total batch: 391
current epoch 79
current batch: 390, total batch: 391
current epoch 80
current batch: 390, total batch: 391
current epoch 81
current batch: 390, total batch: 391
current epoch 82
current batch: 390, total batch: 391
current epoch 83
current batch: 390, total batch: 391
current epoch 84
current batch: 390, total batch: 391
current epoch 85
current batch: 390, total batch: 391
current epoch 86
current batch: 390, total batch: 391
current epoch 87
current batch: 390, total batch: 391
current epoch 88
current batch: 390, total batch: 391
current epoch 89
current batch: 390, total batch: 391
current epoch 90
current batch: 390, total batch: 391
current epoch 91
current batch: 390, total batch: 391
current epoch 92
current batch: 390, total batch: 391
current epoch 93
current batch: 390, total batch: 391
current epoch 94
current batch: 390, total batch: 391
current epoch 95
current batch: 390, total batch: 391
current epoch 96
current batch: 390, total batch: 391
current epoch 97
current batch: 390, total batch: 391
current epoch 98
current batch: 390, total batch: 391
current epoch 99
current batch: 390, total batch: 391
current epoch 100
current batch: 390, total batch: 391
current epoch 101
current batch: 390, total batch: 391
current epoch 102
current batch: 390, total batch: 391
current epoch 103
current batch: 390, total batch: 391
current epoch 104
current batch: 390, total batch: 391
current epoch 105
current batch: 390, total batch: 391
current epoch 106
current batch: 390, total batch: 391
current epoch 107
current batch: 390, total batch: 391
current epoch 108
current batch: 390, total batch: 391
current epoch 109
current batch: 390, total batch: 391
current epoch 110
current batch: 390, total batch: 391
current epoch 111
current batch: 390, total batch: 391
current epoch 112
current batch: 390, total batch: 391
current epoch 113
current batch: 390, total batch: 391
current epoch 114
current batch: 390, total batch: 391
current epoch 115
current batch: 390, total batch: 391
current epoch 116
current batch: 390, total batch: 391
current epoch 117
current batch: 390, total batch: 391
current epoch 118
current batch: 390, total batch: 391
current epoch 119
current batch: 390, total batch: 391
current epoch 120
current batch: 390, total batch: 391
current epoch 121
current batch: 390, total batch: 391
current epoch 122
current batch: 390, total batch: 391
current epoch 123
current batch: 390, total batch: 391
current epoch 124
current batch: 390, total batch: 391
current epoch 125
current batch: 390, total batch: 391
current epoch 126
current batch: 390, total batch: 391
current epoch 127
current batch: 390, total batch: 391
current epoch 128
current batch: 390, total batch: 391
current epoch 129
current batch: 390, total batch: 391
current epoch 130
current batch: 390, total batch: 391
current epoch 131
current batch: 390, total batch: 391
current epoch 132
current batch: 390, total batch: 391
current epoch 133
current batch: 390, total batch: 391
current epoch 134
current batch: 390, total batch: 391
current epoch 135
current batch: 390, total batch: 391
current epoch 136
current batch: 390, total batch: 391
current epoch 137
current batch: 390, total batch: 391
current epoch 138
current batch: 390, total batch: 391
current epoch 139
current batch: 390, total batch: 391
current epoch 140
current batch: 390, total batch: 391
current epoch 141
current batch: 390, total batch: 391
current epoch 142
current batch: 390, total batch: 391
current epoch 143
current batch: 390, total batch: 391
current epoch 144
current batch: 390, total batch: 391
current epoch 145
current batch: 390, total batch: 391
current epoch 146
current batch: 390, total batch: 391
current epoch 147
current batch: 390, total batch: 391
current epoch 148
current batch: 390, total batch: 391
current epoch 149
current batch: 390, total batch: 391
current epoch 150
current batch: 390, total batch: 391
current epoch 151
current batch: 390, total batch: 391
current epoch 152
current batch: 390, total batch: 391
current epoch 153
current batch: 390, total batch: 391
current epoch 154
current batch: 390, total batch: 391
current epoch 155
current batch: 390, total batch: 391
current epoch 156
current batch: 390, total batch: 391
current epoch 157
current batch: 390, total batch: 391
current epoch 158
current batch: 390, total batch: 391
current epoch 159
current batch: 390, total batch: 391
current epoch 160
current batch: 390, total batch: 391
current epoch 161
current batch: 390, total batch: 391
current epoch 162
current batch: 390, total batch: 391
current epoch 163
current batch: 390, total batch: 391
current epoch 164
current batch: 390, total batch: 391
current epoch 165
current batch: 390, total batch: 391
current epoch 166
current batch: 390, total batch: 391
current epoch 167
current batch: 390, total batch: 391
current epoch 168
current batch: 390, total batch: 391
current epoch 169
current batch: 390, total batch: 391
current epoch 170
current batch: 390, total batch: 391
current epoch 171
current batch: 390, total batch: 391
current epoch 172
current batch: 390, total batch: 391
current epoch 173
current batch: 390, total batch: 391
current epoch 174
current batch: 390, total batch: 391
current epoch 175
current batch: 390, total batch: 391
current epoch 176
current batch: 390, total batch: 391
current epoch 177
current batch: 390, total batch: 391
current epoch 178
current batch: 390, total batch: 391
current epoch 179
current batch: 390, total batch: 391
current epoch 180
current batch: 390, total batch: 391
current epoch 181
current batch: 390, total batch: 391
current epoch 182
current batch: 390, total batch: 391
current epoch 183
current batch: 390, total batch: 391
current epoch 184
current batch: 390, total batch: 391
current epoch 185
current batch: 390, total batch: 391
current epoch 186
current batch: 390, total batch: 391
current epoch 187
current batch: 390, total batch: 391
current epoch 188
current batch: 390, total batch: 391
current epoch 189
current batch: 390, total batch: 391
current epoch 190
current batch: 390, total batch: 391
current epoch 191
current batch: 390, total batch: 391
current epoch 192
current batch: 390, total batch: 391
current epoch 193
current batch: 390, total batch: 391
current epoch 194
current batch: 390, total batch: 391
current epoch 195
current batch: 390, total batch: 391
current epoch 196
current batch: 390, total batch: 391
current epoch 197
current batch: 390, total batch: 391
current epoch 198
current batch: 390, total batch: 391
current epoch 199
current batch: 390, total batch: 391
current epoch 200
current batch: 390, total batch: 391
current epoch 201
current batch: 390, total batch: 391
current epoch 202
current batch: 390, total batch: 391
current epoch 203
current batch: 390, total batch: 391
current epoch 204
current batch: 390, total batch: 391
current epoch 205
current batch: 390, total batch: 391
current epoch 206
current batch: 390, total batch: 391
current epoch 207
current batch: 390, total batch: 391
current epoch 208
current batch: 390, total batch: 391
current epoch 209
current batch: 390, total batch: 391
current epoch 210
current batch: 390, total batch: 391
current epoch 211
current batch: 390, total batch: 391
current epoch 212
current batch: 390, total batch: 391
current epoch 213
current batch: 390, total batch: 391
current epoch 214
current batch: 390, total batch: 391
current epoch 215
current batch: 390, total batch: 391
current epoch 216
current batch: 390, total batch: 391
current epoch 217
current batch: 390, total batch: 391
current epoch 218
current batch: 390, total batch: 391
current epoch 219
current batch: 390, total batch: 391
current epoch 220
current batch: 390, total batch: 391
current epoch 221
current batch: 390, total batch: 391
current epoch 222
current batch: 390, total batch: 391
current epoch 223
current batch: 390, total batch: 391
current epoch 224
current batch: 390, total batch: 391
current epoch 225
current batch: 390, total batch: 391
current epoch 226
current batch: 390, total batch: 391
current epoch 227
current batch: 390, total batch: 391
current epoch 228
current batch: 390, total batch: 391
current epoch 229
current batch: 390, total batch: 391
current epoch 230
current batch: 390, total batch: 391
current epoch 231
current batch: 390, total batch: 391
current epoch 232
current batch: 390, total batch: 391
current epoch 233
current batch: 390, total batch: 391
current epoch 234
current batch: 390, total batch: 391
current epoch 235
current batch: 390, total batch: 391
current epoch 236
current batch: 390, total batch: 391
current epoch 237
current batch: 390, total batch: 391
current epoch 238
current batch: 390, total batch: 391
current epoch 239
current batch: 390, total batch: 391
current epoch 240
current batch: 390, total batch: 391
current epoch 241
current batch: 390, total batch: 391
current epoch 242
current batch: 390, total batch: 391
current epoch 243
current batch: 390, total batch: 391
current epoch 244
current batch: 390, total batch: 391
current epoch 245
current batch: 390, total batch: 391
current epoch 246
current batch: 390, total batch: 391
current epoch 247
current batch: 390, total batch: 391
current epoch 248
current batch: 390, total batch: 391
current epoch 249
current batch: 390, total batch: 391
current epoch 250
current batch: 390, total batch: 391
current epoch 251
current batch: 390, total batch: 391
current epoch 252
current batch: 390, total batch: 391
current epoch 253
current batch: 390, total batch: 391
current epoch 254
current batch: 390, total batch: 391
current epoch 255
current batch: 390, total batch: 391
current epoch 256
current batch: 390, total batch: 391
current epoch 257
current batch: 390, total batch: 391
current epoch 258
current batch: 390, total batch: 391
current epoch 259
current batch: 390, total batch: 391
current epoch 260
current batch: 390, total batch: 391
current epoch 261
current batch: 390, total batch: 391
current epoch 262
current batch: 390, total batch: 391
current epoch 263
current batch: 390, total batch: 391
current epoch 264
current batch: 390, total batch: 391
current epoch 265
current batch: 390, total batch: 391
current epoch 266
current batch: 390, total batch: 391
current epoch 267
current batch: 390, total batch: 391
current epoch 268
current batch: 390, total batch: 391
current epoch 269
current batch: 390, total batch: 391
current epoch 270
current batch: 390, total batch: 391
current epoch 271
current batch: 390, total batch: 391
current epoch 272
current batch: 390, total batch: 391
current epoch 273
current batch: 390, total batch: 391
current epoch 274
current batch: 390, total batch: 391
current epoch 275
current batch: 390, total batch: 391
current epoch 276
current batch: 390, total batch: 391
current epoch 277
current batch: 390, total batch: 391
current epoch 278
current batch: 390, total batch: 391
current epoch 279
current batch: 390, total batch: 391
current epoch 280
current batch: 390, total batch: 391
current epoch 281
current batch: 390, total batch: 391
current epoch 282
current batch: 390, total batch: 391
current epoch 283
current batch: 390, total batch: 391
current epoch 284
current batch: 390, total batch: 391
current epoch 285
current batch: 390, total batch: 391
current epoch 286
current batch: 390, total batch: 391
current epoch 287
current batch: 390, total batch: 391
current epoch 288
current batch: 390, total batch: 391
current epoch 289
current batch: 390, total batch: 391
current epoch 290
current batch: 390, total batch: 391
current epoch 291
current batch: 390, total batch: 391
current epoch 292
current batch: 390, total batch: 391
current epoch 293
current batch: 390, total batch: 391
current epoch 294
current batch: 390, total batch: 391
current epoch 295
current batch: 390, total batch: 391
current epoch 296
current batch: 390, total batch: 391
current epoch 297
current batch: 390, total batch: 391
current epoch 298
current batch: 390, total batch: 391
current epoch 299
current batch: 390, total batch: 391
current epoch 300
current batch: 390, total batch: 391
current epoch 301
current batch: 390, total batch: 391
current epoch 302
current batch: 390, total batch: 391
current epoch 303
current batch: 390, total batch: 391
current epoch 304
current batch: 390, total batch: 391
current epoch 305
current batch: 390, total batch: 391
current epoch 306
current batch: 390, total batch: 391
current epoch 307
current batch: 390, total batch: 391
current epoch 308
current batch: 390, total batch: 391
current epoch 309
current batch: 390, total batch: 391
current epoch 310
current batch: 390, total batch: 391
current epoch 311
current batch: 390, total batch: 391
current epoch 312
current batch: 390, total batch: 391
current epoch 313
current batch: 390, total batch: 391
current epoch 314
current batch: 390, total batch: 391
current epoch 315
current batch: 390, total batch: 391
current epoch 316
current batch: 390, total batch: 391
current epoch 317
current batch: 390, total batch: 391
current epoch 318
current batch: 390, total batch: 391
current epoch 319
current batch: 390, total batch: 391
current epoch 320
current batch: 390, total batch: 391
current epoch 321
current batch: 390, total batch: 391
current epoch 322
current batch: 390, total batch: 391
current epoch 323
current batch: 390, total batch: 391
current epoch 324
current batch: 390, total batch: 391
current epoch 325
current batch: 390, total batch: 391
current epoch 326
current batch: 390, total batch: 391
current epoch 327
current batch: 390, total batch: 391
current epoch 328
current batch: 390, total batch: 391
current epoch 329
current batch: 390, total batch: 391
current epoch 330
current batch: 390, total batch: 391
current epoch 331
current batch: 390, total batch: 391
current epoch 332
current batch: 390, total batch: 391
current epoch 333
current batch: 390, total batch: 391
current epoch 334
current batch: 390, total batch: 391
current epoch 335
current batch: 390, total batch: 391
current epoch 336
current batch: 390, total batch: 391
current epoch 337
current batch: 390, total batch: 391
current epoch 338
current batch: 390, total batch: 391
current epoch 339
current batch: 390, total batch: 391
current epoch 340
current batch: 390, total batch: 391
current epoch 341
current batch: 390, total batch: 391
current epoch 342
current batch: 390, total batch: 391
current epoch 343
current batch: 390, total batch: 391
current epoch 344
current batch: 390, total batch: 391
current epoch 345
current batch: 390, total batch: 391
current epoch 346
current batch: 390, total batch: 391
current epoch 347
current batch: 390, total batch: 391
current epoch 348
current batch: 390, total batch: 391
current epoch 349
current batch: 390, total batch: 391
Finished Training
```

In [68]:

```
device = 'cuda'
model = resnet32()
model.to(device)
trainloader = torch.utils.data.DataLoader(train_data, batch_size=128,
                                          shuffle=True, num_workers=4)
optimizer = optim.SGD(model.parameters(), lr=0.01, momentum=0.9, weight_decay=5e-4)
lr_scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer, T_max=200)
criterion = nn.CrossEntropyLoss()
final_loss = []
for epoch in range(350):  # loop over the dataset multiple times
    print(f'current epoch {epoch}', end = '\n')

    running_loss = 0.0
    total_batch = len(trainloader)
    for i, data in enumerate(trainloader, 0):
        if i < len(trainloader)-1:
            print(f'current batch: {i}, total batch: {total_batch}', end='\r')
        else:
            print(f'current batch: {i}, total batch: {total_batch}')
        # get the inputs; data is a list of [inputs, labels]
        inputs, labels = data

        # zero the parameter gradients
        inputs = inputs.to(device)
        labels = labels.to(device)
        # zero the parameter gradients
        optimizer.zero_grad()

        outputs = model(inputs)
        _, preds = torch.max(outputs, 1)
        loss = criterion(outputs, labels)

        loss.backward()
        optimizer.step()
        final_loss.append(loss.item())

        # statistics
        lr_scheduler.step()

print('Finished Training')
save_result(model, final_loss, 'resnet32')
```

```
current epoch 0
current batch: 390, total batch: 391
current epoch 1
current batch: 390, total batch: 391
current epoch 2
current batch: 390, total batch: 391
current epoch 3
current batch: 390, total batch: 391
current epoch 4
current batch: 390, total batch: 391
current epoch 5
current batch: 390, total batch: 391
current epoch 6
current batch: 390, total batch: 391
current epoch 7
current batch: 390, total batch: 391
current epoch 8
current batch: 390, total batch: 391
current epoch 9
current batch: 390, total batch: 391
current epoch 10
current batch: 390, total batch: 391
current epoch 11
current batch: 390, total batch: 391
current epoch 12
current batch: 390, total batch: 391
current epoch 13
current batch: 390, total batch: 391
current epoch 14
current batch: 390, total batch: 391
current epoch 15
current batch: 390, total batch: 391
current epoch 16
current batch: 390, total batch: 391
current epoch 17
current batch: 390, total batch: 391
current epoch 18
current batch: 390, total batch: 391
current epoch 19
current batch: 390, total batch: 391
current epoch 20
current batch: 390, total batch: 391
current epoch 21
current batch: 390, total batch: 391
current epoch 22
current batch: 390, total batch: 391
current epoch 23
current batch: 390, total batch: 391
current epoch 24
current batch: 390, total batch: 391
current epoch 25
current batch: 390, total batch: 391
current epoch 26
current batch: 390, total batch: 391
current epoch 27
current batch: 390, total batch: 391
current epoch 28
current batch: 390, total batch: 391
current epoch 29
current batch: 390, total batch: 391
current epoch 30
current batch: 390, total batch: 391
current epoch 31
current batch: 390, total batch: 391
current epoch 32
current batch: 390, total batch: 391
current epoch 33
current batch: 390, total batch: 391
current epoch 34
current batch: 390, total batch: 391
current epoch 35
current batch: 390, total batch: 391
current epoch 36
current batch: 390, total batch: 391
current epoch 37
current batch: 390, total batch: 391
current epoch 38
current batch: 390, total batch: 391
current epoch 39
current batch: 390, total batch: 391
current epoch 40
current batch: 390, total batch: 391
current epoch 41
current batch: 390, total batch: 391
current epoch 42
current batch: 390, total batch: 391
current epoch 43
current batch: 390, total batch: 391
current epoch 44
current batch: 390, total batch: 391
current epoch 45
current batch: 390, total batch: 391
current epoch 46
current batch: 390, total batch: 391
current epoch 47
current batch: 390, total batch: 391
current epoch 48
current batch: 390, total batch: 391
current epoch 49
current batch: 390, total batch: 391
current epoch 50
current batch: 390, total batch: 391
current epoch 51
current batch: 390, total batch: 391
current epoch 52
current batch: 390, total batch: 391
current epoch 53
current batch: 390, total batch: 391
current epoch 54
current batch: 390, total batch: 391
current epoch 55
current batch: 390, total batch: 391
current epoch 56
current batch: 390, total batch: 391
current epoch 57
current batch: 390, total batch: 391
current epoch 58
current batch: 390, total batch: 391
current epoch 59
current batch: 390, total batch: 391
current epoch 60
current batch: 390, total batch: 391
current epoch 61
current batch: 390, total batch: 391
current epoch 62
current batch: 390, total batch: 391
current epoch 63
current batch: 390, total batch: 391
current epoch 64
current batch: 390, total batch: 391
current epoch 65
current batch: 390, total batch: 391
current epoch 66
current batch: 390, total batch: 391
current epoch 67
current batch: 390, total batch: 391
current epoch 68
current batch: 390, total batch: 391
current epoch 69
current batch: 390, total batch: 391
current epoch 70
current batch: 390, total batch: 391
current epoch 71
current batch: 390, total batch: 391
current epoch 72
current batch: 390, total batch: 391
current epoch 73
current batch: 390, total batch: 391
current epoch 74
current batch: 390, total batch: 391
current epoch 75
current batch: 390, total batch: 391
current epoch 76
current batch: 390, total batch: 391
current epoch 77
current batch: 390, total batch: 391
current epoch 78
current batch: 390, total batch: 391
current epoch 79
current batch: 390, total batch: 391
current epoch 80
current batch: 390, total batch: 391
current epoch 81
current batch: 390, total batch: 391
current epoch 82
current batch: 390, total batch: 391
current epoch 83
current batch: 390, total batch: 391
current epoch 84
current batch: 390, total batch: 391
current epoch 85
current batch: 390, total batch: 391
current epoch 86
current batch: 390, total batch: 391
current epoch 87
current batch: 390, total batch: 391
current epoch 88
current batch: 390, total batch: 391
current epoch 89
current batch: 390, total batch: 391
current epoch 90
current batch: 390, total batch: 391
current epoch 91
current batch: 390, total batch: 391
current epoch 92
current batch: 390, total batch: 391
current epoch 93
current batch: 390, total batch: 391
current epoch 94
current batch: 390, total batch: 391
current epoch 95
current batch: 390, total batch: 391
current epoch 96
current batch: 390, total batch: 391
current epoch 97
current batch: 390, total batch: 391
current epoch 98
current batch: 390, total batch: 391
current epoch 99
current batch: 390, total batch: 391
current epoch 100
current batch: 390, total batch: 391
current epoch 101
current batch: 390, total batch: 391
current epoch 102
current batch: 390, total batch: 391
current epoch 103
current batch: 390, total batch: 391
current epoch 104
current batch: 390, total batch: 391
current epoch 105
current batch: 390, total batch: 391
current epoch 106
current batch: 390, total batch: 391
current epoch 107
current batch: 390, total batch: 391
current epoch 108
current batch: 390, total batch: 391
current epoch 109
current batch: 390, total batch: 391
current epoch 110
current batch: 390, total batch: 391
current epoch 111
current batch: 390, total batch: 391
current epoch 112
current batch: 390, total batch: 391
current epoch 113
current batch: 390, total batch: 391
current epoch 114
current batch: 390, total batch: 391
current epoch 115
current batch: 390, total batch: 391
current epoch 116
current batch: 390, total batch: 391
current epoch 117
current batch: 390, total batch: 391
current epoch 118
current batch: 390, total batch: 391
current epoch 119
current batch: 390, total batch: 391
current epoch 120
current batch: 390, total batch: 391
current epoch 121
current batch: 390, total batch: 391
current epoch 122
current batch: 390, total batch: 391
current epoch 123
current batch: 390, total batch: 391
current epoch 124
current batch: 390, total batch: 391
current epoch 125
current batch: 390, total batch: 391
current epoch 126
current batch: 390, total batch: 391
current epoch 127
current batch: 390, total batch: 391
current epoch 128
current batch: 390, total batch: 391
current epoch 129
current batch: 390, total batch: 391
current epoch 130
current batch: 390, total batch: 391
current epoch 131
current batch: 390, total batch: 391
current epoch 132
current batch: 390, total batch: 391
current epoch 133
current batch: 390, total batch: 391
current epoch 134
current batch: 390, total batch: 391
current epoch 135
current batch: 390, total batch: 391
current epoch 136
current batch: 390, total batch: 391
current epoch 137
current batch: 390, total batch: 391
current epoch 138
current batch: 390, total batch: 391
current epoch 139
current batch: 390, total batch: 391
current epoch 140
current batch: 390, total batch: 391
current epoch 141
current batch: 390, total batch: 391
current epoch 142
current batch: 390, total batch: 391
current epoch 143
current batch: 390, total batch: 391
current epoch 144
current batch: 390, total batch: 391
current epoch 145
current batch: 390, total batch: 391
current epoch 146
current batch: 390, total batch: 391
current epoch 147
current batch: 390, total batch: 391
current epoch 148
current batch: 390, total batch: 391
current epoch 149
current batch: 390, total batch: 391
current epoch 150
current batch: 390, total batch: 391
current epoch 151
current batch: 390, total batch: 391
current epoch 152
current batch: 390, total batch: 391
current epoch 153
current batch: 390, total batch: 391
current epoch 154
current batch: 390, total batch: 391
current epoch 155
current batch: 390, total batch: 391
current epoch 156
current batch: 390, total batch: 391
current epoch 157
current batch: 390, total batch: 391
current epoch 158
current batch: 390, total batch: 391
current epoch 159
current batch: 390, total batch: 391
current epoch 160
current batch: 390, total batch: 391
current epoch 161
current batch: 390, total batch: 391
current epoch 162
current batch: 390, total batch: 391
current epoch 163
current batch: 390, total batch: 391
current epoch 164
current batch: 390, total batch: 391
current epoch 165
current batch: 390, total batch: 391
current epoch 166
current batch: 390, total batch: 391
current epoch 167
current batch: 390, total batch: 391
current epoch 168
current batch: 390, total batch: 391
current epoch 169
current batch: 390, total batch: 391
current epoch 170
current batch: 390, total batch: 391
current epoch 171
current batch: 390, total batch: 391
current epoch 172
current batch: 390, total batch: 391
current epoch 173
current batch: 390, total batch: 391
current epoch 174
current batch: 390, total batch: 391
current epoch 175
current batch: 390, total batch: 391
current epoch 176
current batch: 390, total batch: 391
current epoch 177
current batch: 390, total batch: 391
current epoch 178
current batch: 390, total batch: 391
current epoch 179
current batch: 390, total batch: 391
current epoch 180
current batch: 390, total batch: 391
current epoch 181
current batch: 390, total batch: 391
current epoch 182
current batch: 390, total batch: 391
current epoch 183
current batch: 390, total batch: 391
current epoch 184
current batch: 390, total batch: 391
current epoch 185
current batch: 390, total batch: 391
current epoch 186
current batch: 390, total batch: 391
current epoch 187
current batch: 390, total batch: 391
current epoch 188
current batch: 390, total batch: 391
current epoch 189
current batch: 390, total batch: 391
current epoch 190
current batch: 390, total batch: 391
current epoch 191
current batch: 390, total batch: 391
current epoch 192
current batch: 390, total batch: 391
current epoch 193
current batch: 390, total batch: 391
current epoch 194
current batch: 390, total batch: 391
current epoch 195
current batch: 390, total batch: 391
current epoch 196
current batch: 390, total batch: 391
current epoch 197
current batch: 390, total batch: 391
current epoch 198
current batch: 390, total batch: 391
current epoch 199
current batch: 390, total batch: 391
current epoch 200
current batch: 390, total batch: 391
current epoch 201
current batch: 390, total batch: 391
current epoch 202
current batch: 390, total batch: 391
current epoch 203
current batch: 390, total batch: 391
current epoch 204
current batch: 390, total batch: 391
current epoch 205
current batch: 390, total batch: 391
current epoch 206
current batch: 390, total batch: 391
current epoch 207
current batch: 390, total batch: 391
current epoch 208
current batch: 390, total batch: 391
current epoch 209
current batch: 390, total batch: 391
current epoch 210
current batch: 390, total batch: 391
current epoch 211
current batch: 390, total batch: 391
current epoch 212
current batch: 390, total batch: 391
current epoch 213
current batch: 390, total batch: 391
current epoch 214
current batch: 390, total batch: 391
current epoch 215
current batch: 390, total batch: 391
current epoch 216
current batch: 390, total batch: 391
current epoch 217
current batch: 390, total batch: 391
current epoch 218
current batch: 390, total batch: 391
current epoch 219
current batch: 390, total batch: 391
current epoch 220
current batch: 390, total batch: 391
current epoch 221
current batch: 390, total batch: 391
current epoch 222
current batch: 390, total batch: 391
current epoch 223
current batch: 390, total batch: 391
current epoch 224
current batch: 390, total batch: 391
current epoch 225
current batch: 390, total batch: 391
current epoch 226
current batch: 390, total batch: 391
current epoch 227
current batch: 390, total batch: 391
current epoch 228
current batch: 390, total batch: 391
current epoch 229
current batch: 390, total batch: 391
current epoch 230
current batch: 390, total batch: 391
current epoch 231
current batch: 390, total batch: 391
current epoch 232
current batch: 390, total batch: 391
current epoch 233
current batch: 390, total batch: 391
current epoch 234
current batch: 390, total batch: 391
current epoch 235
current batch: 390, total batch: 391
current epoch 236
current batch: 390, total batch: 391
current epoch 237
current batch: 390, total batch: 391
current epoch 238
current batch: 390, total batch: 391
current epoch 239
current batch: 390, total batch: 391
current epoch 240
current batch: 390, total batch: 391
current epoch 241
current batch: 390, total batch: 391
current epoch 242
current batch: 390, total batch: 391
current epoch 243
current batch: 390, total batch: 391
current epoch 244
current batch: 390, total batch: 391
current epoch 245
current batch: 390, total batch: 391
current epoch 246
current batch: 390, total batch: 391
current epoch 247
current batch: 390, total batch: 391
current epoch 248
current batch: 390, total batch: 391
current epoch 249
current batch: 390, total batch: 391
current epoch 250
current batch: 390, total batch: 391
current epoch 251
current batch: 390, total batch: 391
current epoch 252
current batch: 390, total batch: 391
current epoch 253
current batch: 390, total batch: 391
current epoch 254
current batch: 390, total batch: 391
current epoch 255
current batch: 390, total batch: 391
current epoch 256
current batch: 390, total batch: 391
current epoch 257
current batch: 390, total batch: 391
current epoch 258
current batch: 390, total batch: 391
current epoch 259
current batch: 390, total batch: 391
current epoch 260
current batch: 390, total batch: 391
current epoch 261
current batch: 390, total batch: 391
current epoch 262
current batch: 390, total batch: 391
current epoch 263
current batch: 390, total batch: 391
current epoch 264
current batch: 390, total batch: 391
current epoch 265
current batch: 390, total batch: 391
current epoch 266
current batch: 390, total batch: 391
current epoch 267
current batch: 390, total batch: 391
current epoch 268
current batch: 390, total batch: 391
current epoch 269
current batch: 390, total batch: 391
current epoch 270
current batch: 390, total batch: 391
current epoch 271
current batch: 390, total batch: 391
current epoch 272
current batch: 390, total batch: 391
current epoch 273
current batch: 390, total batch: 391
current epoch 274
current batch: 390, total batch: 391
current epoch 275
current batch: 390, total batch: 391
current epoch 276
current batch: 390, total batch: 391
current epoch 277
current batch: 390, total batch: 391
current epoch 278
current batch: 390, total batch: 391
current epoch 279
current batch: 390, total batch: 391
current epoch 280
current batch: 390, total batch: 391
current epoch 281
current batch: 390, total batch: 391
current epoch 282
current batch: 390, total batch: 391
current epoch 283
current batch: 390, total batch: 391
current epoch 284
current batch: 390, total batch: 391
current epoch 285
current batch: 390, total batch: 391
current epoch 286
current batch: 390, total batch: 391
current epoch 287
current batch: 390, total batch: 391
current epoch 288
current batch: 390, total batch: 391
current epoch 289
current batch: 390, total batch: 391
current epoch 290
current batch: 390, total batch: 391
current epoch 291
current batch: 390, total batch: 391
current epoch 292
current batch: 390, total batch: 391
current epoch 293
current batch: 390, total batch: 391
current epoch 294
current batch: 390, total batch: 391
current epoch 295
current batch: 390, total batch: 391
current epoch 296
current batch: 390, total batch: 391
current epoch 297
current batch: 390, total batch: 391
current epoch 298
current batch: 390, total batch: 391
current epoch 299
current batch: 390, total batch: 391
current epoch 300
current batch: 390, total batch: 391
current epoch 301
current batch: 390, total batch: 391
current epoch 302
current batch: 390, total batch: 391
current epoch 303
current batch: 390, total batch: 391
current epoch 304
current batch: 390, total batch: 391
current epoch 305
current batch: 390, total batch: 391
current epoch 306
current batch: 390, total batch: 391
current epoch 307
current batch: 390, total batch: 391
current epoch 308
current batch: 390, total batch: 391
current epoch 309
current batch: 390, total batch: 391
current epoch 310
current batch: 390, total batch: 391
current epoch 311
current batch: 390, total batch: 391
current epoch 312
current batch: 390, total batch: 391
current epoch 313
current batch: 390, total batch: 391
current epoch 314
current batch: 390, total batch: 391
current epoch 315
current batch: 390, total batch: 391
current epoch 316
current batch: 390, total batch: 391
current epoch 317
current batch: 390, total batch: 391
current epoch 318
current batch: 390, total batch: 391
current epoch 319
current batch: 390, total batch: 391
current epoch 320
current batch: 390, total batch: 391
current epoch 321
current batch: 390, total batch: 391
current epoch 322
current batch: 390, total batch: 391
current epoch 323
current batch: 390, total batch: 391
current epoch 324
current batch: 390, total batch: 391
current epoch 325
current batch: 390, total batch: 391
current epoch 326
current batch: 390, total batch: 391
current epoch 327
current batch: 390, total batch: 391
current epoch 328
current batch: 390, total batch: 391
current epoch 329
current batch: 390, total batch: 391
current epoch 330
current batch: 390, total batch: 391
current epoch 331
current batch: 390, total batch: 391
current epoch 332
current batch: 390, total batch: 391
current epoch 333
current batch: 390, total batch: 391
current epoch 334
current batch: 390, total batch: 391
current epoch 335
current batch: 390, total batch: 391
current epoch 336
current batch: 390, total batch: 391
current epoch 337
current batch: 390, total batch: 391
current epoch 338
current batch: 390, total batch: 391
current epoch 339
current batch: 390, total batch: 391
current epoch 340
current batch: 390, total batch: 391
current epoch 341
current batch: 390, total batch: 391
current epoch 342
current batch: 390, total batch: 391
current epoch 343
current batch: 390, total batch: 391
current epoch 344
current batch: 390, total batch: 391
current epoch 345
current batch: 390, total batch: 391
current epoch 346
current batch: 390, total batch: 391
current epoch 347
current batch: 390, total batch: 391
current epoch 348
current batch: 390, total batch: 391
current epoch 349
current batch: 390, total batch: 391
Finished Training
```

In [69]:

```
device = 'cuda'
model = resnet44()
model.to(device)
```

```python
trainloader = torch.utils.data.DataLoader(train_data, batch_size=128,
                                          shuffle=True, num_workers=4)
optimizer = optim.SGD(model.parameters(), lr=0.01, momentum=0.9, weight_decay=5e-4)
lr_scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer, T_max=200)
criterion = nn.CrossEntropyLoss()
final_loss = []
for epoch in range(350):  # loop over the dataset multiple times
    print(f'current epoch {epoch}', end = '\n')

    running_loss = 0.0
    total_batch = len(trainloader)
    for i, data in enumerate(trainloader, 0):
        if i < len(trainloader)-1:
            print(f'current batch: {i}, total batch: {total_batch}', end='\r')
        else:
            print(f'current batch: {i}, total batch: {total_batch}')
        # get the inputs; data is a list of [inputs, labels]
        inputs, labels = data

        # zero the parameter gradients

        inputs = inputs.to(device)
        labels = labels.to(device)
        # zero the parameter gradients
        optimizer.zero_grad()

        outputs = model(inputs)
        _, preds = torch.max(outputs, 1)
        loss = criterion(outputs, labels)

        loss.backward()
        optimizer.step()
        final_loss.append(loss.item())

        # statistics
        lr_scheduler.step()

print('Finished Training')
save_result(model, final_loss, 'resnet44')
```

```
current epoch 0
current batch: 390, total batch: 391
current epoch 1
current batch: 390, total batch: 391
current epoch 2
current batch: 390, total batch: 391
current epoch 3
current batch: 390, total batch: 391
current epoch 4
current batch: 390, total batch: 391
current epoch 5
current batch: 390, total batch: 391
current epoch 6
current batch: 390, total batch: 391
current epoch 7
current batch: 390, total batch: 391
current epoch 8
current batch: 390, total batch: 391
current epoch 9
current batch: 390, total batch: 391
current epoch 10
current batch: 390, total batch: 391
current epoch 11
current batch: 390, total batch: 391
current epoch 12
current batch: 390, total batch: 391
current epoch 13
current batch: 390, total batch: 391
current epoch 14
current batch: 390, total batch: 391
current epoch 15
current batch: 390, total batch: 391
current epoch 16
current batch: 390, total batch: 391
current epoch 17
current batch: 390, total batch: 391
current epoch 18
current batch: 390, total batch: 391
current epoch 19
current batch: 390, total batch: 391
current epoch 20
current batch: 390, total batch: 391
current epoch 21
current batch: 390, total batch: 391
current epoch 22
current batch: 390, total batch: 391
current epoch 23
current batch: 390, total batch: 391
current epoch 24
current batch: 390, total batch: 391
current epoch 25
current batch: 390, total batch: 391
current epoch 26
current batch: 390, total batch: 391
current epoch 27
current batch: 390, total batch: 391
current epoch 28
current batch: 390, total batch: 391
current epoch 29
current batch: 390, total batch: 391
current epoch 30
current batch: 390, total batch: 391
current epoch 31
current batch: 390, total batch: 391
current epoch 32
current batch: 390, total batch: 391
current epoch 33
current batch: 390, total batch: 391
current epoch 34
current batch: 390, total batch: 391
current epoch 35
current batch: 390, total batch: 391
current epoch 36
current batch: 390, total batch: 391
current epoch 37
current batch: 390, total batch: 391
current epoch 38
current batch: 390, total batch: 391
current epoch 39
current batch: 390, total batch: 391
current epoch 40
current batch: 390, total batch: 391
current epoch 41
current batch: 390, total batch: 391
current epoch 42
current batch: 390, total batch: 391
current epoch 43
current batch: 390, total batch: 391
current epoch 44
current batch: 390, total batch: 391
current epoch 45
current batch: 390, total batch: 391
current epoch 46
current batch: 390, total batch: 391
current epoch 47
current batch: 390, total batch: 391
current epoch 48
current batch: 390, total batch: 391
current epoch 49
current batch: 390, total batch: 391
current epoch 50
current batch: 390, total batch: 391
current epoch 51
current batch: 390, total batch: 391
current epoch 52
current batch: 390, total batch: 391
current epoch 53
current batch: 390, total batch: 391
current epoch 54
current batch: 390, total batch: 391
current epoch 55
current batch: 390, total batch: 391
current epoch 56
current batch: 390, total batch: 391
current epoch 57
current batch: 390, total batch: 391
current epoch 58
current batch: 390, total batch: 391
current epoch 59
current batch: 390, total batch: 391
current epoch 60
current batch: 390, total batch: 391
current epoch 61
current batch: 390, total batch: 391
current epoch 62
current batch: 390, total batch: 391
current epoch 63
current batch: 390, total batch: 391
current epoch 64
current batch: 390, total batch: 391
current epoch 65
current batch: 390, total batch: 391
current epoch 66
current batch: 390, total batch: 391
current epoch 67
current batch: 390, total batch: 391
current epoch 68
current batch: 390, total batch: 391
current epoch 69
current batch: 390, total batch: 391
current epoch 70
current batch: 390, total batch: 391
current epoch 71
current batch: 390, total batch: 391
current epoch 72
current batch: 390, total batch: 391
current epoch 73
current batch: 390, total batch: 391
current epoch 74
current batch: 390, total batch: 391
current epoch 75
current batch: 390, total batch: 391
current epoch 76
current batch: 390, total batch: 391
current epoch 77
current batch: 390, total batch: 391
current epoch 78
current batch: 390, total batch: 391
current epoch 79
current batch: 390, total batch: 391
current epoch 80
current batch: 390, total batch: 391
current epoch 81
current batch: 390, total batch: 391
current epoch 82
current batch: 390, total batch: 391
current epoch 83
current batch: 390, total batch: 391
current epoch 84
current batch: 390, total batch: 391
current epoch 85
current batch: 390, total batch: 391
current epoch 86
current batch: 390, total batch: 391
current epoch 87
current batch: 390, total batch: 391
current epoch 88
current batch: 390, total batch: 391
current epoch 89
current batch: 390, total batch: 391
current epoch 90
current batch: 390, total batch: 391
current epoch 91
current batch: 390, total batch: 391
current epoch 92
current batch: 390, total batch: 391
current epoch 93
current batch: 390, total batch: 391
current epoch 94
current batch: 390, total batch: 391
current epoch 95
current batch: 390, total batch: 391
current epoch 96
current batch: 390, total batch: 391
current epoch 97
current batch: 390, total batch: 391
current epoch 98
current batch: 390, total batch: 391
current epoch 99
current batch: 390, total batch: 391
current epoch 100
current batch: 390, total batch: 391
current epoch 101
current batch: 390, total batch: 391
current epoch 102
current batch: 390, total batch: 391
current epoch 103
current batch: 390, total batch: 391
current epoch 104
current batch: 390, total batch: 391
current epoch 105
current batch: 390, total batch: 391
current epoch 106
current batch: 390, total batch: 391
current epoch 107
current batch: 390, total batch: 391
current epoch 108
current batch: 390, total batch: 391
current epoch 109
current batch: 390, total batch: 391
current epoch 110
current batch: 390, total batch: 391
current epoch 111
current batch: 390, total batch: 391
current epoch 112
current batch: 390, total batch: 391
current epoch 113
current batch: 390, total batch: 391
current epoch 114
current batch: 390, total batch: 391
current epoch 115
current batch: 390, total batch: 391
current epoch 116
current batch: 390, total batch: 391
current epoch 117
current batch: 390, total batch: 391
current epoch 118
current batch: 390, total batch: 391
current epoch 119
current batch: 390, total batch: 391
current epoch 120
current batch: 390, total batch: 391
current epoch 121
current batch: 390, total batch: 391
current epoch 122
current batch: 390, total batch: 391
current epoch 123
current batch: 390, total batch: 391
current epoch 124
current batch: 390, total batch: 391
current epoch 125
current batch: 390, total batch: 391
current epoch 126
current batch: 390, total batch: 391
current epoch 127
current batch: 390, total batch: 391
current epoch 128
current batch: 390, total batch: 391
current epoch 129
current batch: 390, total batch: 391
current epoch 130
current batch: 390, total batch: 391
current epoch 131
current batch: 390, total batch: 391
current epoch 132
current batch: 390, total batch: 391
current epoch 133
current batch: 390, total batch: 391
current epoch 134
current batch: 390, total batch: 391
current epoch 135
current batch: 390, total batch: 391
current epoch 136
current batch: 390, total batch: 391
current epoch 137
current batch: 390, total batch: 391
current epoch 138
current batch: 390, total batch: 391
current epoch 139
current batch: 390, total batch: 391
current epoch 140
current batch: 390, total batch: 391
current epoch 141
current batch: 390, total batch: 391
current epoch 142
current batch: 390, total batch: 391
current epoch 143
current batch: 390, total batch: 391
current epoch 144
current batch: 390, total batch: 391
current epoch 145
current batch: 390, total batch: 391
current epoch 146
current batch: 390, total batch: 391
current epoch 147
current batch: 390, total batch: 391
current epoch 148
current batch: 390, total batch: 391
current epoch 149
current batch: 390, total batch: 391
current epoch 150
current batch: 390, total batch: 391
current epoch 151
current batch: 390, total batch: 391
current epoch 152
current batch: 390, total batch: 391
current epoch 153
current batch: 390, total batch: 391
current epoch 154
current batch: 390, total batch: 391
current epoch 155
current batch: 390, total batch: 391
current epoch 156
current batch: 390, total batch: 391
current epoch 157
current batch: 390, total batch: 391
current epoch 158
current batch: 390, total batch: 391
current epoch 159
current batch: 390, total batch: 391
current epoch 160
current batch: 390, total batch: 391
current epoch 161
current batch: 390, total batch: 391
current epoch 162
current batch: 390, total batch: 391
current epoch 163
current batch: 390, total batch: 391
current epoch 164
current batch: 390, total batch: 391
current epoch 165
current batch: 390, total batch: 391
current epoch 166
current batch: 390, total batch: 391
current epoch 167
current batch: 390, total batch: 391
current epoch 168
current batch: 390, total batch: 391
current epoch 169
current batch: 390, total batch: 391
current epoch 170
current batch: 390, total batch: 391
current epoch 171
current batch: 390, total batch: 391
current epoch 172
current batch: 390, total batch: 391
current epoch 173
current batch: 390, total batch: 391
current epoch 174
current batch: 390, total batch: 391
current epoch 175
current batch: 390, total batch: 391
current epoch 176
current batch: 390, total batch: 391
current epoch 177
current batch: 390, total batch: 391
current epoch 178
current batch: 390, total batch: 391
current epoch 179
current batch: 390, total batch: 391
current epoch 180
current batch: 390, total batch: 391
current epoch 181
current batch: 390, total batch: 391
current epoch 182
current batch: 390, total batch: 391
current epoch 183
current batch: 390, total batch: 391
current epoch 184
current batch: 390, total batch: 391
current epoch 185
current batch: 390, total batch: 391
current epoch 186
current batch: 390, total batch: 391
current epoch 187
current batch: 390, total batch: 391
current epoch 188
current batch: 390, total batch: 391
current epoch 189
current batch: 390, total batch: 391
current epoch 190
current batch: 390, total batch: 391
current epoch 191
current batch: 390, total batch: 391
current epoch 192
current batch: 390, total batch: 391
current epoch 193
current batch: 390, total batch: 391
current epoch 194
current batch: 390, total batch: 391
current epoch 195
current batch: 390, total batch: 391
current epoch 196
current batch: 390, total batch: 391
current epoch 197
current batch: 390, total batch: 391
current epoch 198
current batch: 390, total batch: 391
current epoch 199
current batch: 390, total batch: 391
current epoch 200
current batch: 390, total batch: 391
current epoch 201
current batch: 390, total batch: 391
current epoch 202
current batch: 390, total batch: 391
current epoch 203
current batch: 390, total batch: 391
current epoch 204
current batch: 390, total batch: 391
current epoch 205
current batch: 390, total batch: 391
current epoch 206
current batch: 390, total batch: 391
current epoch 207
current batch: 390, total batch: 391
current epoch 208
current batch: 390, total batch: 391
current epoch 209
current batch: 390, total batch: 391
current epoch 210
current batch: 390, total batch: 391
current epoch 211
current batch: 390, total batch: 391
current epoch 212
current batch: 390, total batch: 391
current epoch 213
current batch: 390, total batch: 391
current epoch 214
current batch: 390, total batch: 391
current epoch 215
current batch: 390, total batch: 391
current epoch 216
current batch: 390, total batch: 391
current epoch 217
current batch: 390, total batch: 391
current epoch 218
current batch: 390, total batch: 391
current epoch 219
current batch: 390, total batch: 391
current epoch 220
current batch: 390, total batch: 391
current epoch 221
current batch: 390, total batch: 391
current epoch 222
current batch: 390, total batch: 391
current epoch 223
current batch: 390, total batch: 391
current epoch 224
current batch: 390, total batch: 391
current epoch 225
current batch: 390, total batch: 391
current epoch 226
current batch: 390, total batch: 391
current epoch 227
current batch: 390, total batch: 391
current epoch 228
current batch: 390, total batch: 391
current epoch 229
current batch: 390, total batch: 391
current epoch 230
current batch: 390, total batch: 391
current epoch 231
current batch: 390, total batch: 391
current epoch 232
current batch: 390, total batch: 391
current epoch 233
current batch: 390, total batch: 391
current epoch 234
current batch: 390, total batch: 391
current epoch 235
current batch: 390, total batch: 391
current epoch 236
current batch: 390, total batch: 391
current epoch 237
current batch: 390, total batch: 391
current epoch 238
current batch: 390, total batch: 391
current epoch 239
current batch: 390, total batch: 391
current epoch 240
current batch: 390, total batch: 391
current epoch 241
current batch: 390, total batch: 391
current epoch 242
current batch: 390, total batch: 391
current epoch 243
current batch: 390, total batch: 391
current epoch 244
current batch: 390, total batch: 391
current epoch 245
current batch: 390, total batch: 391
current epoch 246
current batch: 390, total batch: 391
current epoch 247
current batch: 390, total batch: 391
current epoch 248
current batch: 390, total batch: 391
current epoch 249
current batch: 390, total batch: 391
current epoch 250
current batch: 390, total batch: 391
current epoch 251
current batch: 390, total batch: 391
current epoch 252
current batch: 390, total batch: 391
current epoch 253
current batch: 390, total batch: 391
current epoch 254
current batch: 390, total batch: 391
current epoch 255
current batch: 390, total batch: 391
current epoch 256
current batch: 390, total batch: 391
current epoch 257
current batch: 390, total batch: 391
current epoch 258
current batch: 390, total batch: 391
current epoch 259
current batch: 390, total batch: 391
current epoch 260
current batch: 390, total batch: 391
current epoch 261
current batch: 390, total batch: 391
current epoch 262
current batch: 390, total batch: 391
current epoch 263
current batch: 390, total batch: 391
current epoch 264
current batch: 390, total batch: 391
current epoch 265
current batch: 390, total batch: 391
current epoch 266
current batch: 390, total batch: 391
current epoch 267
current batch: 390, total batch: 391
current epoch 268
current batch: 390, total batch: 391
current epoch 269
current batch: 390, total batch: 391
current epoch 270
current batch: 390, total batch: 391
current epoch 271
current batch: 390, total batch: 391
current epoch 272
current batch: 390, total batch: 391
current epoch 273
current batch: 390, total batch: 391
current epoch 274
current batch: 390, total batch: 391
current epoch 275
current batch: 390, total batch: 391
current epoch 276
current batch: 390, total batch: 391
current epoch 277
current batch: 390, total batch: 391
current epoch 278
current batch: 390, total batch: 391
current epoch 279
current batch: 390, total batch: 391
current epoch 280
current batch: 390, total batch: 391
current epoch 281
current batch: 390, total batch: 391
current epoch 282
current batch: 390, total batch: 391
current epoch 283
current batch: 390, total batch: 391
current epoch 284
current batch: 390, total batch: 391
current epoch 285
current batch: 390, total batch: 391
current epoch 286
current batch: 390, total batch: 391
current epoch 287
current batch: 390, total batch: 391
current epoch 288
current batch: 390, total batch: 391
current epoch 289
current batch: 390, total batch: 391
current epoch 290
current batch: 390, total batch: 391
current epoch 291
current batch: 390, total batch: 391
current epoch 292
current batch: 390, total batch: 391
current epoch 293
current batch: 390, total batch: 391
current epoch 294
current batch: 390, total batch: 391
current epoch 295
current batch: 390, total batch: 391
current epoch 296
current batch: 390, total batch: 391
current epoch 297
current batch: 390, total batch: 391
current epoch 298
current batch: 390, total batch: 391
current epoch 299
current batch: 390, total batch: 391
current epoch 300
current batch: 390, total batch: 391
current epoch 301
current batch: 390, total batch: 391
current epoch 302
current batch: 390, total batch: 391
current epoch 303
current batch: 390, total batch: 391
current epoch 304
current batch: 390, total batch: 391
current epoch 305
current batch: 390, total batch: 391
current epoch 306
current batch: 390, total batch: 391
current epoch 307
current batch: 390, total batch: 391
current epoch 308
current batch: 390, total batch: 391
current epoch 309
current batch: 390, total batch: 391
current epoch 310
current batch: 390, total batch: 391
current epoch 311
current batch: 390, total batch: 391
current epoch 312
current batch: 390, total batch: 391
current epoch 313
current batch: 390, total batch: 391
current epoch 314
current batch: 390, total batch: 391
current epoch 315
current batch: 390, total batch: 391
current epoch 316
current batch: 390, total batch: 391
current epoch 317
current batch: 390, total batch: 391
current epoch 318
current batch: 390, total batch: 391
current epoch 319
current batch: 390, total batch: 391
current epoch 320
current batch: 390, total batch: 391
current epoch 321
current batch: 390, total batch: 391
current epoch 322
current batch: 390, total batch: 391
current epoch 323
current batch: 390, total batch: 391
current epoch 324
current batch: 390, total batch: 391
current epoch 325
current batch: 390, total batch: 391
current epoch 326
current batch: 390, total batch: 391
current epoch 327
current batch: 390, total batch: 391
current epoch 328
current batch: 390, total batch: 391
current epoch 329
current batch: 390, total batch: 391
current epoch 330
current batch: 390, total batch: 391
current epoch 331
current batch: 390, total batch: 391
current epoch 332
current batch: 390, total batch: 391
current epoch 333
current batch: 390, total batch: 391
current epoch 334
current batch: 390, total batch: 391
current epoch 335
current batch: 390, total batch: 391
current epoch 336
current batch: 390, total batch: 391
current epoch 337
current batch: 390, total batch: 391
current epoch 338
current batch: 390, total batch: 391
current epoch 339
current batch: 390, total batch: 391
current epoch 340
current batch: 390, total batch: 391
current epoch 341
current batch: 390, total batch: 391
current epoch 342
current batch: 390, total batch: 391
current epoch 343
current batch: 390, total batch: 391
current epoch 344
current batch: 390, total batch: 391
current epoch 345
current batch: 390, total batch: 391
current epoch 346
current batch: 390, total batch: 391
current epoch 347
current batch: 390, total batch: 391
current epoch 348
current batch: 390, total batch: 391
current epoch 349
current batch: 390, total batch: 391
Finished Training
```

In [70]:

```
device = 'cuda'
model = resnet56()
model.to(device)
trainloader = torch.utils.data.DataLoader(train_data, batch_size=128,
                                          shuffle=True, num_workers=4)
optimizer = optim.SGD(model.parameters(), lr=0.01, momentum=0.9, weight_decay=5e-4)
lr_scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer, T_max=200)
criterion = nn.CrossEntropyLoss()
final_loss = []
for epoch in range(350):  # loop over the dataset multiple times
    print(f'current epoch {epoch}', end = '\n')

    running_loss = 0.0
    total_batch = len(trainloader)
    for i, data in enumerate(trainloader, 0):
        if i < len(trainloader)-1:
            print(f'current batch: {i}, total batch: {total_batch}', end='\r')
        else:
            print(f'current batch: {i}, total batch: {total_batch}')
        # get the inputs; data is a list of [inputs, labels]
        inputs, labels = data

        # zero the parameter gradients

        inputs = inputs.to(device)
        labels = labels.to(device)
        # zero the parameter gradients
        optimizer.zero_grad()

        outputs = model(inputs)
        _, preds = torch.max(outputs, 1)
        loss = criterion(outputs, labels)

        loss.backward()
        optimizer.step()
        final_loss.append(loss.item())

        # statistics
        lr_scheduler.step()

print('Finished Training')
save_result(model, final_loss, 'resnet56')
```

```
current epoch 0
current batch: 390, total batch: 391
current epoch 1
current batch: 390, total batch: 391
current epoch 2
current batch: 390, total batch: 391
current epoch 3
current batch: 390, total batch: 391
current epoch 4
current batch: 390, total batch: 391
current epoch 5
current batch: 390, total batch: 391
current epoch 6
current batch: 390, total batch: 391
current epoch 7
current batch: 390, total batch: 391
current epoch 8
current batch: 390, total batch: 391
current epoch 9
current batch: 390, total batch: 391
current epoch 10
current batch: 390, total batch: 391
current epoch 11
current batch: 390, total batch: 391
current epoch 12
current batch: 390, total batch: 391
current epoch 13
current batch: 390, total batch: 391
current epoch 14
current batch: 390, total batch: 391
current epoch 15
current batch: 390, total batch: 391
current epoch 16
current batch: 390, total batch: 391
current epoch 17
current batch: 390, total batch: 391
current epoch 18
current batch: 390, total batch: 391
current epoch 19
current batch: 390, total batch: 391
current epoch 20
current batch: 390, total batch: 391
current epoch 21
current batch: 390, total batch: 391
current epoch 22
current batch: 390, total batch: 391
current epoch 23
current batch: 390, total batch: 391
current epoch 24
current batch: 390, total batch: 391
current epoch 25
current batch: 390, total batch: 391
current epoch 26
current batch: 390, total batch: 391
current epoch 27
current batch: 390, total batch: 391
current epoch 28
current batch: 390, total batch: 391
current epoch 29
current batch: 390, total batch: 391
current epoch 30
current batch: 390, total batch: 391
current epoch 31
current batch: 390, total batch: 391
current epoch 32
current batch: 390, total batch: 391
current epoch 33
current batch: 390, total batch: 391
current epoch 34
current batch: 390, total batch: 391
current epoch 35
current batch: 390, total batch: 391
current epoch 36
current batch: 390, total batch: 391
current epoch 37
current batch: 390, total batch: 391
current epoch 38
current batch: 390, total batch: 391
current epoch 39
current batch: 390, total batch: 391
current epoch 40
current batch: 390, total batch: 391
current epoch 41
current batch: 390, total batch: 391
current epoch 42
current batch: 390, total batch: 391
current epoch 43
current batch: 390, total batch: 391
current epoch 44
current batch: 390, total batch: 391
current epoch 45
current batch: 390, total batch: 391
current epoch 46
current batch: 390, total batch: 391
current epoch 47
current batch: 390, total batch: 391
current epoch 48
current batch: 390, total batch: 391
current epoch 49
current batch: 390, total batch: 391
current epoch 50
current batch: 390, total batch: 391
current epoch 51
current batch: 390, total batch: 391
current epoch 52
current batch: 390, total batch: 391
current epoch 53
current batch: 390, total batch: 391
current epoch 54
current batch: 390, total batch: 391
current epoch 55
current batch: 390, total batch: 391
current epoch 56
current batch: 390, total batch: 391
current epoch 57
current batch: 390, total batch: 391
current epoch 58
current batch: 390, total batch: 391
current epoch 59
current batch: 390, total batch: 391
current epoch 60
current batch: 390, total batch: 391
current epoch 61
current batch: 390, total batch: 391
current epoch 62
current batch: 390, total batch: 391
current epoch 63
current batch: 390, total batch: 391
current epoch 64
current batch: 390, total batch: 391
current epoch 65
current batch: 390, total batch: 391
current epoch 66
current batch: 390, total batch: 391
current epoch 67
current batch: 390, total batch: 391
current epoch 68
current batch: 390, total batch: 391
current epoch 69
current batch: 390, total batch: 391
current epoch 70
current batch: 390, total batch: 391
current epoch 71
current batch: 390, total batch: 391
current epoch 72
current batch: 390, total batch: 391
current epoch 73
current batch: 390, total batch: 391
current epoch 74
current batch: 390, total batch: 391
current epoch 75
current batch: 390, total batch: 391
current epoch 76
current batch: 390, total batch: 391
current epoch 77
current batch: 390, total batch: 391
current epoch 78
current batch: 390, total batch: 391
current epoch 79
current batch: 390, total batch: 391
current epoch 80
current batch: 390, total batch: 391
current epoch 81
current batch: 390, total batch: 391
current epoch 82
current batch: 390, total batch: 391
current epoch 83
current batch: 390, total batch: 391
current epoch 84
current batch: 390, total batch: 391
current epoch 85
current batch: 390, total batch: 391
current epoch 86
current batch: 390, total batch: 391
current epoch 87
current batch: 390, total batch: 391
current epoch 88
current batch: 390, total batch: 391
current epoch 89
current batch: 390, total batch: 391
current epoch 90
current batch: 390, total batch: 391
current epoch 91
current batch: 390, total batch: 391
current epoch 92
current batch: 390, total batch: 391
current epoch 93
current batch: 390, total batch: 391
current epoch 94
current batch: 390, total batch: 391
current epoch 95
current batch: 390, total batch: 391
current epoch 96
current batch: 390, total batch: 391
current epoch 97
current batch: 390, total batch: 391
current epoch 98
current batch: 390, total batch: 391
current epoch 99
current batch: 390, total batch: 391
current epoch 100
current batch: 390, total batch: 391
current epoch 101
current batch: 390, total batch: 391
current epoch 102
current batch: 390, total batch: 391
current epoch 103
current batch: 390, total batch: 391
current epoch 104
current batch: 390, total batch: 391
current epoch 105
current batch: 390, total batch: 391
current epoch 106
current batch: 390, total batch: 391
current epoch 107
current batch: 390, total batch: 391
current epoch 108
current batch: 390, total batch: 391
current epoch 109
current batch: 390, total batch: 391
current epoch 110
current batch: 390, total batch: 391
current epoch 111
current batch: 390, total batch: 391
current epoch 112
current batch: 390, total batch: 391
current epoch 113
current batch: 390, total batch: 391
current epoch 114
current batch: 390, total batch: 391
current epoch 115
current batch: 390, total batch: 391
current epoch 116
current batch: 390, total batch: 391
current epoch 117
current batch: 390, total batch: 391
current epoch 118
current batch: 390, total batch: 391
current epoch 119
current batch: 390, total batch: 391
current epoch 120
current batch: 390, total batch: 391
current epoch 121
current batch: 390, total batch: 391
current epoch 122
current batch: 390, total batch: 391
current epoch 123
current batch: 390, total batch: 391
current epoch 124
current batch: 390, total batch: 391
current epoch 125
current batch: 390, total batch: 391
current epoch 126
current batch: 390, total batch: 391
current epoch 127
current batch: 390, total batch: 391
current epoch 128
current batch: 390, total batch: 391
current epoch 129
current batch: 390, total batch: 391
current epoch 130
current batch: 390, total batch: 391
current epoch 131
current batch: 390, total batch: 391
current epoch 132
current batch: 390, total batch: 391
current epoch 133
current batch: 390, total batch: 391
current epoch 134
current batch: 390, total batch: 391
current epoch 135
current batch: 390, total batch: 391
current epoch 136
current batch: 390, total batch: 391
current epoch 137
current batch: 390, total batch: 391
current epoch 138
current batch: 390, total batch: 391
current epoch 139
current batch: 390, total batch: 391
current epoch 140
current batch: 390, total batch: 391
current epoch 141
current batch: 390, total batch: 391
current epoch 142
current batch: 390, total batch: 391
current epoch 143
current batch: 390, total batch: 391
current epoch 144
current batch: 390, total batch: 391
current epoch 145
current batch: 390, total batch: 391
current epoch 146
current batch: 390, total batch: 391
current epoch 147
current batch: 390, total batch: 391
current epoch 148
current batch: 390, total batch: 391
current epoch 149
current batch: 390, total batch: 391
current epoch 150
current batch: 390, total batch: 391
current epoch 151
current batch: 390, total batch: 391
current epoch 152
current batch: 390, total batch: 391
current epoch 153
current batch: 390, total batch: 391
current epoch 154
current batch: 390, total batch: 391
current epoch 155
current batch: 390, total batch: 391
current epoch 156
current batch: 390, total batch: 391
current epoch 157
current batch: 390, total batch: 391
current epoch 158
current batch: 390, total batch: 391
current epoch 159
current batch: 390, total batch: 391
current epoch 160
current batch: 390, total batch: 391
current epoch 161
current batch: 390, total batch: 391
current epoch 162
current batch: 390, total batch: 391
current epoch 163
current batch: 390, total batch: 391
current epoch 164
current batch: 390, total batch: 391
current epoch 165
current batch: 390, total batch: 391
current epoch 166
current batch: 390, total batch: 391
current epoch 167
current batch: 390, total batch: 391
current epoch 168
current batch: 390, total batch: 391
current epoch 169
current batch: 390, total batch: 391
current epoch 170
current batch: 390, total batch: 391
current epoch 171
current batch: 390, total batch: 391
current epoch 172
current batch: 390, total batch: 391
current epoch 173
current batch: 390, total batch: 391
current epoch 174
current batch: 390, total batch: 391
current epoch 175
current batch: 390, total batch: 391
current epoch 176
current batch: 390, total batch: 391
current epoch 177
current batch: 390, total batch: 391
current epoch 178
current batch: 390, total batch: 391
current epoch 179
current batch: 390, total batch: 391
current epoch 180
current batch: 390, total batch: 391
current epoch 181
current batch: 390, total batch: 391
current epoch 182
current batch: 390, total batch: 391
current epoch 183
current batch: 390, total batch: 391
current epoch 184
current batch: 390, total batch: 391
current epoch 185
current batch: 390, total batch: 391
current epoch 186
current batch: 390, total batch: 391
current epoch 187
current batch: 390, total batch: 391
current epoch 188
current batch: 390, total batch: 391
current epoch 189
current batch: 390, total batch: 391
current epoch 190
current batch: 390, total batch: 391
current epoch 191
current batch: 390, total batch: 391
current epoch 192
current batch: 390, total batch: 391
current epoch 193
current batch: 390, total batch: 391
current epoch 194
current batch: 390, total batch: 391
current epoch 195
current batch: 390, total batch: 391
current epoch 196
current batch: 390, total batch: 391
current epoch 197
current batch: 390, total batch: 391
current epoch 198
current batch: 390, total batch: 391
current epoch 199
current batch: 390, total batch: 391
current epoch 200
current batch: 390, total batch: 391
current epoch 201
current batch: 390, total batch: 391
current epoch 202
current batch: 390, total batch: 391
current epoch 203
current batch: 390, total batch: 391
current epoch 204
current batch: 390, total batch: 391
current epoch 205
current batch: 390, total batch: 391
current epoch 206
current batch: 390, total batch: 391
current epoch 207
current batch: 390, total batch: 391
current epoch 208
current batch: 390, total batch: 391
current epoch 209
current batch: 390, total batch: 391
current epoch 210
current batch: 390, total batch: 391
current epoch 211
current batch: 390, total batch: 391
current epoch 212
current batch: 390, total batch: 391
current epoch 213
current batch: 390, total batch: 391
current epoch 214
current batch: 390, total batch: 391
current epoch 215
current batch: 390, total batch: 391
current epoch 216
current batch: 390, total batch: 391
current epoch 217
current batch: 390, total batch: 391
current epoch 218
current batch: 390, total batch: 391
current epoch 219
current batch: 390, total batch: 391
current epoch 220
current batch: 390, total batch: 391
current epoch 221
current batch: 390, total batch: 391
current epoch 222
current batch: 390, total batch: 391
current epoch 223
current batch: 390, total batch: 391
current epoch 224
current batch: 390, total batch: 391
current epoch 225
current batch: 390, total batch: 391
current epoch 226
current batch: 390, total batch: 391
current epoch 227
current batch: 390, total batch: 391
current epoch 228
current batch: 390, total batch: 391
current epoch 229
current batch: 390, total batch: 391
current epoch 230
current batch: 390, total batch: 391
current epoch 231
current batch: 390, total batch: 391
current epoch 232
current batch: 390, total batch: 391
current epoch 233
current batch: 390, total batch: 391
current epoch 234
current batch: 390, total batch: 391
current epoch 235
current batch: 390, total batch: 391
current epoch 236
current batch: 390, total batch: 391
current epoch 237
current batch: 390, total batch: 391
current epoch 238
current batch: 390, total batch: 391
current epoch 239
current batch: 390, total batch: 391
current epoch 240
current batch: 390, total batch: 391
current epoch 241
current batch: 390, total batch: 391
current epoch 242
current batch: 390, total batch: 391
current epoch 243
current batch: 390, total batch: 391
current epoch 244
current batch: 390, total batch: 391
current epoch 245
current batch: 390, total batch: 391
current epoch 246
current batch: 390, total batch: 391
current epoch 247
current batch: 390, total batch: 391
current epoch 248
current batch: 390, total batch: 391
current epoch 249
current batch: 390, total batch: 391
current epoch 250
current batch: 390, total batch: 391
current epoch 251
current batch: 390, total batch: 391
current epoch 252
current batch: 390, total batch: 391
current epoch 253
current batch: 390, total batch: 391
current epoch 254
current batch: 390, total batch: 391
current epoch 255
current batch: 390, total batch: 391
current epoch 256
current batch: 390, total batch: 391
current epoch 257
current batch: 390, total batch: 391
current epoch 258
current batch: 390, total batch: 391
current epoch 259
current batch: 390, total batch: 391
current epoch 260
current batch: 390, total batch: 391
current epoch 261
current batch: 390, total batch: 391
current epoch 262
current batch: 390, total batch: 391
current epoch 263
current batch: 390, total batch: 391
current epoch 264
current batch: 390, total batch: 391
current epoch 265
current batch: 390, total batch: 391
current epoch 266
current batch: 390, total batch: 391
current epoch 267
current batch: 390, total batch: 391
current epoch 268
current batch: 390, total batch: 391
current epoch 269
current batch: 390, total batch: 391
current epoch 270
current batch: 390, total batch: 391
current epoch 271
current batch: 390, total batch: 391
current epoch 272
current batch: 390, total batch: 391
current epoch 273
current batch: 390, total batch: 391
current epoch 274
current batch: 390, total batch: 391
current epoch 275
current batch: 390, total batch: 391
current epoch 276
current batch: 390, total batch: 391
current epoch 277
current batch: 390, total batch: 391
current epoch 278
current batch: 390, total batch: 391
current epoch 279
current batch: 390, total batch: 391
current epoch 280
current batch: 390, total batch: 391
current epoch 281
current batch: 390, total batch: 391
current epoch 282
current batch: 390, total batch: 391
current epoch 283
current batch: 390, total batch: 391
current epoch 284
current batch: 390, total batch: 391
current epoch 285
current batch: 390, total batch: 391
current epoch 286
current batch: 390, total batch: 391
current epoch 287
current batch: 390, total batch: 391
current epoch 288
current batch: 390, total batch: 391
current epoch 289
current batch: 390, total batch: 391
current epoch 290
current batch: 390, total batch: 391
current epoch 291
current batch: 390, total batch: 391
current epoch 292
current batch: 390, total batch: 391
current epoch 293
current batch: 390, total batch: 391
current epoch 294
current batch: 390, total batch: 391
current epoch 295
current batch: 390, total batch: 391
current epoch 296
current batch: 390, total batch: 391
current epoch 297
current batch: 390, total batch: 391
current epoch 298
current batch: 390, total batch: 391
current epoch 299
current batch: 390, total batch: 391
current epoch 300
current batch: 390, total batch: 391
current epoch 301
current batch: 390, total batch: 391
current epoch 302
current batch: 390, total batch: 391
current epoch 303
current batch: 390, total batch: 391
current epoch 304
current batch: 390, total batch: 391
current epoch 305
current batch: 390, total batch: 391
current epoch 306
current batch: 390, total batch: 391
current epoch 307
current batch: 390, total batch: 391
current epoch 308
current batch: 390, total batch: 391
current epoch 309
current batch: 390, total batch: 391
current epoch 310
current batch: 390, total batch: 391
current epoch 311
current batch: 390, total batch: 391
current epoch 312
current batch: 390, total batch: 391
current epoch 313
current batch: 390, total batch: 391
current epoch 314
current batch: 390, total batch: 391
current epoch 315
current batch: 390, total batch: 391
current epoch 316
current batch: 390, total batch: 391
current epoch 317
current batch: 390, total batch: 391
current epoch 318
current batch: 390, total batch: 391
current epoch 319
current batch: 390, total batch: 391
current epoch 320
current batch: 390, total batch: 391
current epoch 321
current batch: 390, total batch: 391
current epoch 322
current batch: 390, total batch: 391
current epoch 323
current batch: 390, total batch: 391
current epoch 324
current batch: 390, total batch: 391
current epoch 325
current batch: 390, total batch: 391
current epoch 326
current batch: 390, total batch: 391
current epoch 327
current batch: 390, total batch: 391
current epoch 328
current batch: 390, total batch: 391
current epoch 329
current batch: 390, total batch: 391
current epoch 330
current batch: 390, total batch: 391
current epoch 331
current batch: 390, total batch: 391
current epoch 332
current batch: 390, total batch: 391
current epoch 333
current batch: 390, total batch: 391
current epoch 334
current batch: 390, total batch: 391
current epoch 335
current batch: 390, total batch: 391
current epoch 336
current batch: 390, total batch: 391
current epoch 337
current batch: 390, total batch: 391
current epoch 338
current batch: 390, total batch: 391
current epoch 339
current batch: 390, total batch: 391
current epoch 340
current batch: 390, total batch: 391
current epoch 341
current batch: 390, total batch: 391
current epoch 342
current batch: 390, total batch: 391
current epoch 343
current batch: 390, total batch: 391
current epoch 344
current batch: 390, total batch: 391
current epoch 345
current batch: 390, total batch: 391
current epoch 346
current batch: 390, total batch: 391
current epoch 347
current batch: 390, total batch: 391
current epoch 348
current batch: 390, total batch: 391
current epoch 349
current batch: 390, total batch: 391
Finished Training
```

Start estimating $\beta_0, \beta_1, \beta_2$

In [8]:

```
# read_data
# reference: K_80,P_100,V_100 data from Nana Antwi(nka2122), RTXGeForce data from KaiYu He
k_80_path = ['k80/resnet18_K80.pickle','k80/resnet20_K80.pickle',
             'k80/resnet32_K80.pickle','k80/resnet44_K80.pickle','k80/resnet56_K80.pickle'
p_100_path = ['p100/resnet18_p100.pickle','p100/resnet20_p100.pickle',
              'p100/resnet32_p100.pickle','p100/resnet44_p100.pickle','p100/resnet56_p100.
v_100_path = ['v100/resnet18_v100.pickle','v100/resnet20_v100.pickle',
              'v100/resnet32_v100.pickle','v100/resnet44_v100.pickle','v100/resnet56_v100.
RTX3070_path = ['RTX3070/resnet18_3070.pickle','RTX3070/resnet20_3070.pickle',
                'RTX3070/resnet32_3070.pickle','RTX3070/resnet44_3070.pickle','RTX3070/res

k80_data = []
for path in k_80_path:
    with open(path,'rb') as handle:
        k80_data.append(pickle.load(handle))

p100_data = []
for path in p_100_path:
    with open(path,'rb') as handle:
        p100_data.append(pickle.load(handle))

v100_data = []
for path in v_100_path:
    with open(path,'rb') as handle:
        v100_data.append(pickle.load(handle))
```

```
RTX3070_data = []
for path in RTX3070_path:
    with open(path,'rb') as handle:
        RTX3070_data.append(pickle.load(handle))
```

fit the beta parameters

In [57]:
```
from scipy.optimize import curve_fit
def f(k, beta_0, beta_1, beta_2):
    l = (1 / (beta_0 * k + beta_1)) + beta_2
    return l
four_gpu_data = [k80_data,p100_data,v100_data,RTX3070_data]
A = []
b = []
models_names = ['resnet_18','resnet_20','resnet_32','resnet_44','resnet_56']
gpu_names = ['K80','P100','V100','RTX3070']
models_Q4 = {}
for gpu_name_index in range(len(gpu_names)):
    for model_name_index in range(len(models_names)):
        losses = four_gpu_data[gpu_name_index][model_name_index]
        steps = list(range(len(losses)))
        loss_model_name = models_names[model_name_index] + gpu_names[gpu_name_index]
        print(f'current fitting model {loss_model_name}')
        betas, _ = curve_fit(f, xdata=steps, ydata=losses, bounds=(0, np.inf))
        models_Q4[loss_model_name] = betas
```

```
current fitting model resnet_18K80
current fitting model resnet_20K80
current fitting model resnet_32K80
current fitting model resnet_44K80
current fitting model resnet_56K80
current fitting model resnet_18P100
current fitting model resnet_20P100
current fitting model resnet_32P100
current fitting model resnet_44P100
current fitting model resnet_56P100
current fitting model resnet_18V100
current fitting model resnet_20V100
current fitting model resnet_32V100
current fitting model resnet_44V100
current fitting model resnet_56V100
current fitting model resnet_18RTX3070
current fitting model resnet_20RTX3070
current fitting model resnet_32RTX3070
current fitting model resnet_44RTX3070
current fitting model resnet_56RTX3070
```

In [58]:
```
models_Q4
```

Out[58]:
```
{'resnet_18K80': array([2.09013264e-04, 5.26713015e-01, 1.95618694e-02]),
 'resnet_20K80': array([2.84084959e-04, 5.33718844e-01, 4.02777553e-02]),
 'resnet_32K80': array([3.34481810e-04, 4.77946530e-01, 2.03401668e-02]),
 'resnet_44K80': array([3.79693614e-04, 4.58461843e-01, 2.13158387e-02]),
 'resnet_56K80': array([3.58791907e-04, 4.53133788e-01, 9.07341602e-03]),
 'resnet_18P100': array([5.47045401e-05, 6.92688614e-01, 1.29526611e-17]),
 'resnet_20P100': array([8.37824010e-05, 1.06370663e+00, 2.76914023e-16]),
 'resnet_32P100': array([1.90115352e-04, 6.01919744e-01, 4.06733964e-02]),
 'resnet_44P100': array([2.27243699e-04, 4.19269448e-01, 6.11455436e-02]),
 'resnet_56P100': array([2.84584656e-04, 3.80685602e-01, 7.38270487e-02]),
 'resnet_18V100': array([2.69111978e-04, 4.62298634e-01, 2.79970532e-01]),
 'resnet_20V100': array([4.23603485e-04, 5.62509142e-01, 1.82018854e-01]),
 'resnet_32V100': array([0.00045007, 0.3650336 , 0.16109684]),
 'resnet_44V100': array([0.00049638, 0.27942287, 0.17538129]),
 'resnet_56V100': array([0.00041467, 0.30863082, 0.15408717]),
 'resnet_18RTX3070': array([1.32464378e-04, 6.57874823e-01, 6.11972275e-02]),
 'resnet_20RTX3070': array([1.96867453e-04, 6.29378549e-01, 8.54626315e-02]),
 'resnet_32RTX3070': array([2.19561691e-04, 5.83922509e-01, 6.29022088e-02]),
 'resnet_44RTX3070': array([2.23667929e-04, 5.83097664e-01, 4.66002021e-02]),
 'resnet_56RTX3070': array([2.19173802e-04, 5.65655215e-01, 3.57220075e-02])}
```

(2)

In [66]:

```python
# train resnet 50 data
device = 'cuda'
data_size = len(train_data)
model = torchvision.models.resnet50()
model.to(device)
trainloader = torch.utils.data.DataLoader(train_data, batch_size=128,
                                             shuffle=True, num_workers=4)
optimizer = optim.SGD(model.parameters(), lr=0.01, momentum=0.9, weight_decay=5e-4)
lr_scheduler = torch.optim.lr_scheduler.CosineAnnealingLR(optimizer, T_max=200)
criterion = nn.CrossEntropyLoss()
final_loss = []
for epoch in range(350):  # loop over the dataset multiple times
    print(f'current epoch {epoch}', end = '\n')

    running_loss = 0.0
    running_corrects = 0
    total_batch = len(trainloader)
    for i, data in enumerate(trainloader, 0):
        if i < len(trainloader)-1:
            print(f'current batch: {i}, total batch: {total_batch}', end='\r')
        else:
            print(f'current batch: {i}, total batch: {total_batch}')
        # get the inputs; data is a list of [inputs, labels]
        inputs, labels = data

        # zero the parameter gradients

        inputs = inputs.to(device)
        labels = labels.to(device)
        # zero the parameter gradients
        optimizer.zero_grad()

        outputs = model(inputs)
        _, preds = torch.max(outputs, 1)
        loss = criterion(outputs, labels)

        loss.backward()
        optimizer.step()
        final_loss.append(loss.item())

        running_corrects += torch.sum(preds == labels.data)

        # statistics
        lr_scheduler.step()

    epoch_acc = running_corrects.double() / data_size
    print(f'epoch acc = {epoch_acc}')
    if epoch_acc>=0.92:
        print(f'reaches 92% acc at epoch {epoch}')
        break

print('Finished Training')
```

```
current epoch 0
current batch: 390, total batch: 391
epoch acc = 0.25464000000000003
current epoch 1
current batch: 390, total batch: 391
epoch acc = 0.33584
current epoch 2
current batch: 390, total batch: 391
epoch acc = 0.34884000000000004
current epoch 3
current batch: 390, total batch: 391
epoch acc = 0.31622
current epoch 4
current batch: 390, total batch: 391
epoch acc = 0.34296000000000004
current epoch 5
current batch: 390, total batch: 391
epoch acc = 0.33692000000000005
current epoch 6
current batch: 390, total batch: 391
epoch acc = 0.3422
current epoch 7
current batch: 390, total batch: 391
epoch acc = 0.38752000000000003
current epoch 8
current batch: 390, total batch: 391
epoch acc = 0.39438
current epoch 9
current batch: 390, total batch: 391
epoch acc = 0.34702000000000005
current epoch 10
current batch: 390, total batch: 391
epoch acc = 0.38488000000000006
current epoch 11
current batch: 390, total batch: 391
epoch acc = 0.39524000000000004
current epoch 12
current batch: 390, total batch: 391
epoch acc = 0.40662000000000004
current epoch 13
current batch: 390, total batch: 391
epoch acc = 0.42756000000000005
current epoch 14
current batch: 390, total batch: 391
epoch acc = 0.41514
current epoch 15
current batch: 390, total batch: 391
epoch acc = 0.43402
current epoch 16
current batch: 390, total batch: 391
epoch acc = 0.44966000000000006
current epoch 17
current batch: 390, total batch: 391
epoch acc = 0.45970000000000005
current epoch 18
current batch: 390, total batch: 391
epoch acc = 0.47038
current epoch 19
current batch: 390, total batch: 391
epoch acc = 0.48524000000000006
current epoch 20
current batch: 390, total batch: 391
epoch acc = 0.47052000000000005
current epoch 21
current batch: 390, total batch: 391
epoch acc = 0.44974000000000003
current epoch 22
current batch: 390, total batch: 391
epoch acc = 0.45904000000000006
current epoch 23
current batch: 390, total batch: 391
epoch acc = 0.46892000000000006
current epoch 24
current batch: 390, total batch: 391
epoch acc = 0.4772
current epoch 25
current batch: 390, total batch: 391
epoch acc = 0.49002000000000007
current epoch 26
current batch: 390, total batch: 391
epoch acc = 0.5076
current epoch 27
current batch: 390, total batch: 391
epoch acc = 0.505
current epoch 28
current batch: 390, total batch: 391
epoch acc = 0.51742
current epoch 29
current batch: 390, total batch: 391
epoch acc = 0.52986
current epoch 30
current batch: 390, total batch: 391
epoch acc = 0.5222
current epoch 31
current batch: 390, total batch: 391
epoch acc = 0.516
current epoch 32
current batch: 390, total batch: 391
epoch acc = 0.5333800000000001
current epoch 33
current batch: 390, total batch: 391
epoch acc = 0.53732
current epoch 34
current batch: 390, total batch: 391
epoch acc = 0.56164
current epoch 35
current batch: 390, total batch: 391
epoch acc = 0.5745
current epoch 36
current batch: 390, total batch: 391
epoch acc = 0.5732400000000001
current epoch 37
current batch: 390, total batch: 391
epoch acc = 0.58142
current epoch 38
current batch: 390, total batch: 391
epoch acc = 0.5961200000000001
current epoch 39
current batch: 390, total batch: 391
epoch acc = 0.6041200000000001
current epoch 40
current batch: 390, total batch: 391
epoch acc = 0.6087400000000001
current epoch 41
current batch: 390, total batch: 391
epoch acc = 0.6087600000000001
current epoch 42
current batch: 390, total batch: 391
epoch acc = 0.6242800000000001
current epoch 43
current batch: 390, total batch: 391
epoch acc = 0.6267
current epoch 44
current batch: 390, total batch: 391
epoch acc = 0.63226
current epoch 45
current batch: 390, total batch: 391
epoch acc = 0.6391800000000001
current epoch 46
current batch: 390, total batch: 391
epoch acc = 0.6401
current epoch 47
current batch: 390, total batch: 391
epoch acc = 0.6594000000000001
current epoch 48
current batch: 390, total batch: 391
epoch acc = 0.66874
current epoch 49
current batch: 390, total batch: 391
epoch acc = 0.6528200000000001
current epoch 50
current batch: 390, total batch: 391
epoch acc = 0.6729
current epoch 51
current batch: 390, total batch: 391
epoch acc = 0.6841200000000001
current epoch 52
current batch: 390, total batch: 391
epoch acc = 0.68586
current epoch 53
current batch: 390, total batch: 391
epoch acc = 0.69198
current epoch 54
current batch: 390, total batch: 391
epoch acc = 0.65878
current epoch 55
current batch: 390, total batch: 391
epoch acc = 0.6877200000000001
current epoch 56
current batch: 390, total batch: 391
epoch acc = 0.7052400000000001
current epoch 57
current batch: 390, total batch: 391
epoch acc = 0.71382
current epoch 58
current batch: 390, total batch: 391
epoch acc = 0.7186600000000001
current epoch 59
current batch: 390, total batch: 391
epoch acc = 0.72704
current epoch 60
current batch: 390, total batch: 391
epoch acc = 0.7366
current epoch 61
current batch: 390, total batch: 391
epoch acc = 0.7407800000000001
current epoch 62
current batch: 390, total batch: 391
epoch acc = 0.7471800000000001
current epoch 63
current batch: 390, total batch: 391
epoch acc = 0.7542000000000001
current epoch 64
current batch: 390, total batch: 391
epoch acc = 0.7570000000000001
current epoch 65
current batch: 390, total batch: 391
epoch acc = 0.7629600000000001
current epoch 66
current batch: 390, total batch: 391
epoch acc = 0.7678
current epoch 67
current batch: 390, total batch: 391
epoch acc = 0.7717600000000001
current epoch 68
current batch: 390, total batch: 391
epoch acc = 0.7788
current epoch 69
current batch: 390, total batch: 391
epoch acc = 0.78056
current epoch 70
current batch: 390, total batch: 391
epoch acc = 0.7878000000000001
current epoch 71
current batch: 390, total batch: 391
epoch acc = 0.7939
current epoch 72
current batch: 390, total batch: 391
epoch acc = 0.7661600000000001
current epoch 73
current batch: 390, total batch: 391
epoch acc = 0.7651600000000001
current epoch 74
current batch: 390, total batch: 391
epoch acc = 0.7883600000000001
current epoch 75
current batch: 390, total batch: 391
epoch acc = 0.7995800000000001
current epoch 76
current batch: 390, total batch: 391
epoch acc = 0.8083800000000001
current epoch 77
current batch: 390, total batch: 391
epoch acc = 0.8112400000000001
current epoch 78
current batch: 390, total batch: 391
epoch acc = 0.8141200000000001
current epoch 79
current batch: 390, total batch: 391
epoch acc = 0.8192400000000001
current epoch 80
current batch: 390, total batch: 391
epoch acc = 0.8242400000000001
current epoch 81
current batch: 390, total batch: 391
epoch acc = 0.81654
current epoch 82
current batch: 390, total batch: 391
epoch acc = 0.82452
current epoch 83
current batch: 390, total batch: 391
epoch acc = 0.8268800000000001
current epoch 84
current batch: 390, total batch: 391
epoch acc = 0.8329200000000001
current epoch 85
current batch: 390, total batch: 391
epoch acc = 0.8368800000000001
current epoch 86
current batch: 390, total batch: 391
epoch acc = 0.83996
current epoch 87
current batch: 390, total batch: 391
epoch acc = 0.8409000000000001
current epoch 88
current batch: 390, total batch: 391
epoch acc = 0.8434600000000001
current epoch 89
current batch: 390, total batch: 391
epoch acc = 0.84508
current epoch 90
current batch: 390, total batch: 391
epoch acc = 0.8501000000000001
current epoch 91
current batch: 390, total batch: 391
epoch acc = 0.8521000000000001
current epoch 92
current batch: 390, total batch: 391
epoch acc = 0.8501400000000001
current epoch 93
current batch: 390, total batch: 391
epoch acc = 0.8550000000000001
current epoch 94
current batch: 390, total batch: 391
epoch acc = 0.8557600000000001
current epoch 95
current batch: 390, total batch: 391
epoch acc = 0.8552000000000001
current epoch 96
current batch: 390, total batch: 391
epoch acc = 0.8575800000000001
current epoch 97
current batch: 390, total batch: 391
epoch acc = 0.8610200000000001
current epoch 98
current batch: 390, total batch: 391
epoch acc = 0.8613600000000001
current epoch 99
current batch: 390, total batch: 391
epoch acc = 0.8651800000000001
current epoch 100
current batch: 390, total batch: 391
epoch acc = 0.8674400000000001
current epoch 101
current batch: 390, total batch: 391
epoch acc = 0.8667
current epoch 102
current batch: 390, total batch: 391
epoch acc = 0.8722400000000001
current epoch 103
current batch: 390, total batch: 391
epoch acc = 0.8725200000000001
current epoch 104
current batch: 390, total batch: 391
epoch acc = 0.8754000000000001
current epoch 105
current batch: 390, total batch: 391
epoch acc = 0.8777600000000001
current epoch 106
current batch: 390, total batch: 391
epoch acc = 0.87826
current epoch 107
current batch: 390, total batch: 391
epoch acc = 0.8807200000000001
current epoch 108
current batch: 390, total batch: 391
epoch acc = 0.8843000000000001
current epoch 109
current batch: 390, total batch: 391
epoch acc = 0.88338
current epoch 110
current batch: 390, total batch: 391
epoch acc = 0.8859800000000001
current epoch 111
current batch: 390, total batch: 391
epoch acc = 0.8894400000000001
current epoch 112
current batch: 390, total batch: 391
epoch acc = 0.8890800000000001
current epoch 113
current batch: 390, total batch: 391
epoch acc = 0.8923000000000001
current epoch 114
current batch: 390, total batch: 391
epoch acc = 0.8960800000000001
current epoch 115
current batch: 390, total batch: 391
epoch acc = 0.8970400000000001
current epoch 116
current batch: 390, total batch: 391
epoch acc = 0.8978600000000001
current epoch 117
current batch: 390, total batch: 391
epoch acc = 0.8978600000000001
current epoch 118
current batch: 390, total batch: 391
epoch acc = 0.9003800000000001
current epoch 119
current batch: 390, total batch: 391
epoch acc = 0.90244
current epoch 120
current batch: 390, total batch: 391
epoch acc = 0.902
current epoch 121
current batch: 390, total batch: 391
epoch acc = 0.90578
current epoch 122
current batch: 390, total batch: 391
epoch acc = 0.9064800000000001
current epoch 123
current batch: 390, total batch: 391
epoch acc = 0.9060600000000001
current epoch 124
current batch: 390, total batch: 391
epoch acc = 0.9091400000000001
current epoch 125
current batch: 390, total batch: 391
epoch acc = 0.9109200000000001
current epoch 126
current batch: 390, total batch: 391
epoch acc = 0.9115000000000001
current epoch 127
current batch: 390, total batch: 391
epoch acc = 0.9130400000000001
current epoch 128
current batch: 390, total batch: 391
epoch acc = 0.9141400000000001
current epoch 129
current batch: 390, total batch: 391
epoch acc = 0.914
current epoch 130
current batch: 390, total batch: 391
epoch acc = 0.9143000000000001
current epoch 131
current batch: 390, total batch: 391
epoch acc = 0.9179
current epoch 132
current batch: 390, total batch: 391
epoch acc = 0.9169400000000001
current epoch 133
current batch: 390, total batch: 391
epoch acc = 0.9199600000000001
current epoch 134
current batch: 390, total batch: 391
epoch acc = 0.9165000000000001
current epoch 135
current batch: 390, total batch: 391
epoch acc = 0.9186400000000001
current epoch 136
current batch: 390, total batch: 391
epoch acc = 0.9177200000000001
current epoch 137
current batch: 390, total batch: 391
epoch acc = 0.9197400000000001
current epoch 138
current batch: 390, total batch: 391
epoch acc = 0.92078
reaches 92% acc at epoch 138
Finished Training
```

In [69]:
```
save_result(model, final_loss, 'resnet50')
```

read resnet50data

load data from parater and mine

My partner didn't use early stop so he run's the model with full 350 epochs

In [72]:
```
resnet_50path = ['resnet50_k80','resnet50_P100','resnet50_V100','resnet50_3070']
resnet_50_data = []
for path in resnet_50path:
    with open(path+'.pickle','rb') as handle:
        resnet_50_data.append(pickle.load(handle))
```

In [ ]:
```
four_gpu_data = [k80_data,p100_data,v100_data,RTX3070_data]
```

In [113…
```
layer_depth = [18,20,32,44,56]
regression_df = None
for gpu_name_index in range(len(gpu_names)):
    for model_name_index in range(len(models_names)):
        layer = layer_depth[model_name_index]
        current_GPU = gpu_names[gpu_name_index]
        betas = models_Q4[models_names[model_name_index]+gpu_names[gpu_name_index]]
        current_dataframe = pd.DataFrame({
            'layer':[layer],
            'GPU':[current_GPU],
            'beta0':[betas[0]],
            'beta1':[betas[1]],
            'beta2':[betas[2]]
        })
        if regression_df is None:
            regression_df = current_dataframe
        else:
            regression_df = pd.concat([regression_df,current_dataframe], axis = 0)


def categorical_K80(input_gpu):
    if input_gpu == 'K80':
        return 1
    else:
        return 0
def categorical_P100(input_gpu):
    if input_gpu == 'P100':
        return 1
    else:
        return 0
def categorical_V100(input_gpu):
    if input_gpu == 'V100':
        return 1
    else:
        return 0
def categorical_RTX3070(input_gpu):
    if input_gpu == 'RTX3070':
        return 1
    else:
        return 0
regression_df['K80'] = regression_df['GPU'].apply(categorical_K80)
regression_df['P100'] = regression_df['GPU'].apply(categorical_P100)
regression_df['V100'] = regression_df['GPU'].apply(categorical_V100)
regression_df['RTX3070'] = regression_df['GPU'].apply(categorical_RTX3070)
regression_df = regression_df.iloc[:,[0,5,6,7,8,2,3,4]]
```

In [114…

```
regression_df
```

Out[114…

	layer	K80	P100	V100	RTX3070	beta0	beta1	beta2
0	18	1	0	0	0	0.000209	0.526713	1.956187e-02
0	20	1	0	0	0	0.000284	0.533719	4.027776e-02
0	32	1	0	0	0	0.000334	0.477947	2.034017e-02
0	44	1	0	0	0	0.000380	0.458462	2.131584e-02
0	56	1	0	0	0	0.000359	0.453134	9.073416e-03
0	18	0	1	0	0	0.000055	0.692689	1.295266e-17
0	20	0	1	0	0	0.000084	1.063707	2.769140e-16
0	32	0	1	0	0	0.000190	0.601920	4.067340e-02
0	44	0	1	0	0	0.000227	0.419269	6.114554e-02
0	56	0	1	0	0	0.000285	0.380686	7.382705e-02
0	18	0	0	1	0	0.000269	0.462299	2.799705e-01
0	20	0	0	1	0	0.000424	0.562509	1.820189e-01
0	32	0	0	1	0	0.000450	0.365034	1.610968e-01
0	44	0	0	1	0	0.000496	0.279423	1.753813e-01
0	56	0	0	1	0	0.000415	0.308631	1.540872e-01
0	18	0	0	0	1	0.000132	0.657875	6.119723e-02
0	20	0	0	0	1	0.000197	0.629379	8.546263e-02
0	32	0	0	0	1	0.000220	0.583923	6.290221e-02
0	44	0	0	0	1	0.000224	0.583098	4.660020e-02
0	56	0	0	0	1	0.000219	0.565655	3.572201e-02

In [117…
```
# fit regression model
from sklearn import linear_model
beta_models = []
for i in range(3):
    regression_model = linear_model.LinearRegression()
    X = regression_df[['layer', 'K80', 'P100', 'V100', 'RTX3070']].values
    y = regression_df[[f'beta{i}']].values
    regression_model.fit(X, y)
    beta_models.append(regression_model)
beta_models
```

Out[117…
```
[LinearRegression(), LinearRegression(), LinearRegression()]
```

In [131…
```
# predict for resnet50
predict_df = pd.DataFrame({
    'layer' : [50,50,50,50],
    'K80' : [1,0,0,0],
    'P100' : [0,1,0,0],
    'V100' : [0,0,1,0],
    'RTX3070' : [0,0,0,1]
})
predict_df['beta_0'] = beta_models[0].predict(predict_df.iloc[:,[0,1,2,3,4]])
predict_df['beta_1'] = beta_models[1].predict(predict_df.iloc[:,[0,1,2,3,4]])
predict_df['beta_2'] = beta_models[2].predict(predict_df.iloc[:,[0,1,2,3,4]])
```

In [132…
```
predict_df
```

Out[132…

	layer	K80	P100	V100	RTX3070	beta_0	beta_1	beta_2
0	50	1	0	0	0	0.000369	0.391686	0.015936
1	50	0	1	0	0	0.000224	0.533345	0.028951
2	50	0	0	1	0	0.000467	0.297270	0.184333
3	50	0	0	0	1	0.000254	0.505677	0.052199

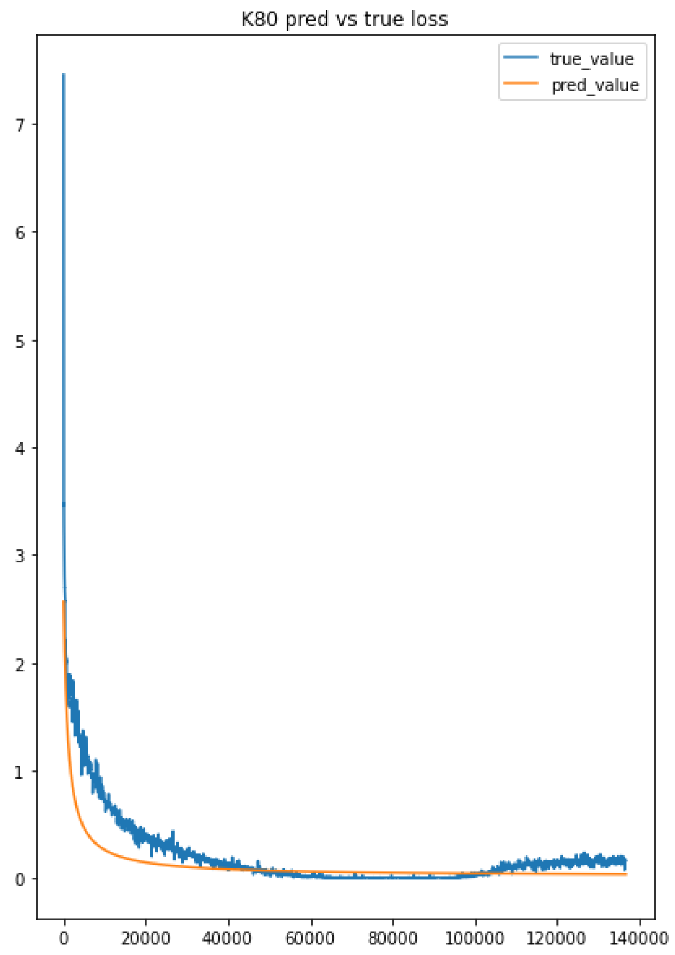
In [136…
```
predict_df.iloc[0,5]
```

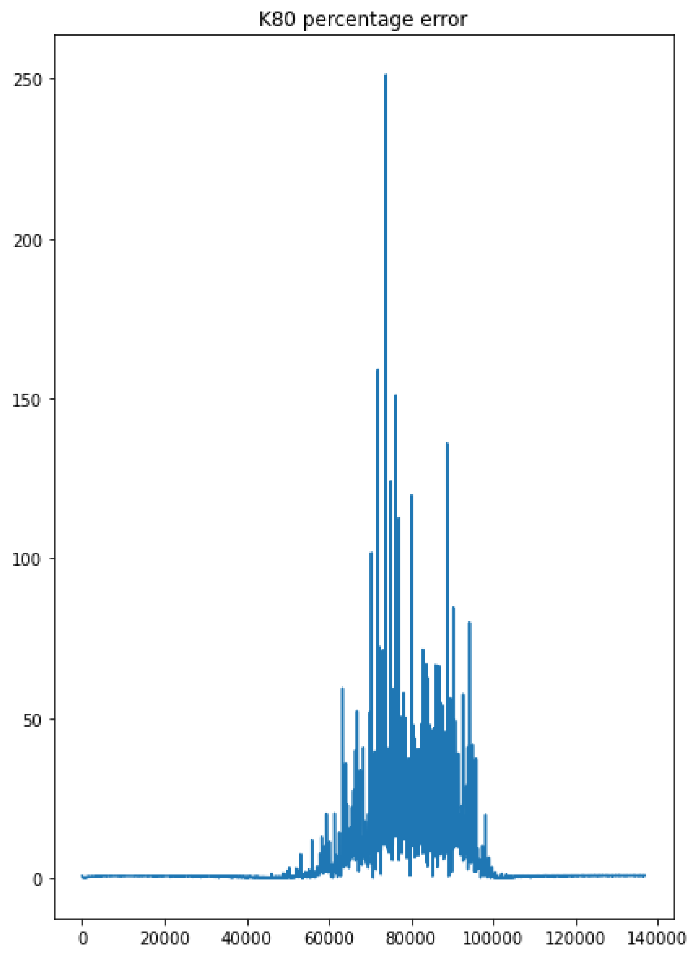
Out[136…
```
0.00036897812191694 63
```

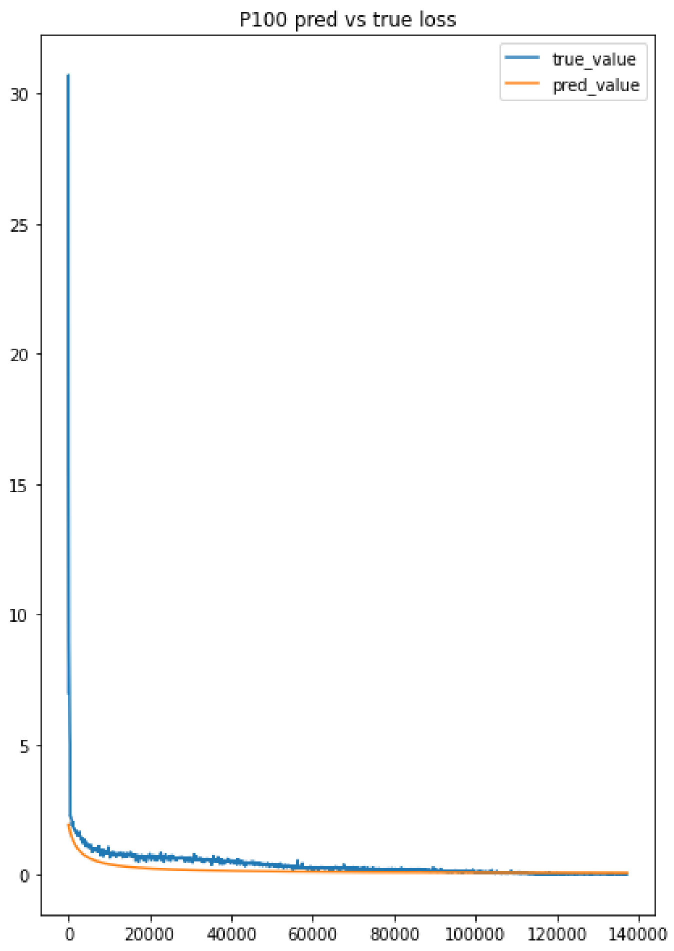
In [141…
```
pred_losses = []
for gpu_index in range(len(resnet_50_data)):
    gpu_name = gpu_names[gpu_index]
    loss_list = resnet_50_data[gpu_index]
    step = list(range(len(loss_list)))
    beta_0 = predict_df.iloc[gpu_index,5]
    beta_1 = predict_df.iloc[gpu_index,6]
    beta_2 = predict_df.iloc[gpu_index,7]
    current_pred_loss = []
    for single_step in step:
        pred_loss = f(single_step+1,beta_0,beta_1,beta_2)
        current_pred_loss.append(pred_loss)
    pred_losses.append(current_pred_loss)
percentage_error_all = []
for gpu_index in range(len(resnet_50_data)):
    percentage_error = []
    for i in range(len(resnet_50_data[gpu_index])):
        true_value = resnet_50_data[gpu_index][i]
        pred_value = pred_losses[gpu_index][i]
        error = round((abs(true_value-pred_value)/true_value),4)
        percentage_error.append(error)
    percentage_error_all.append(percentage_error)
```

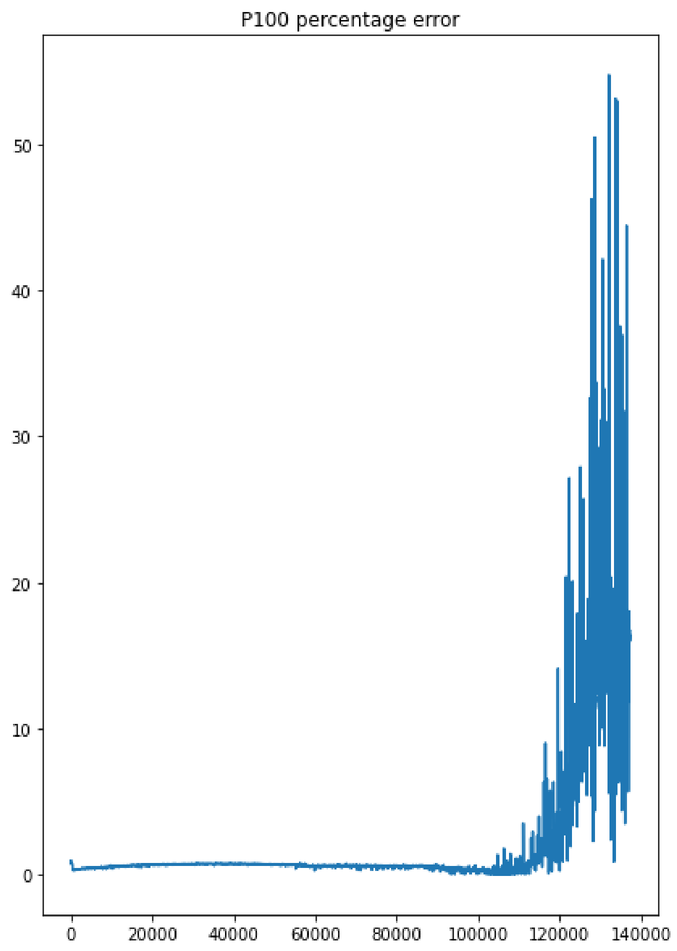
In [156…

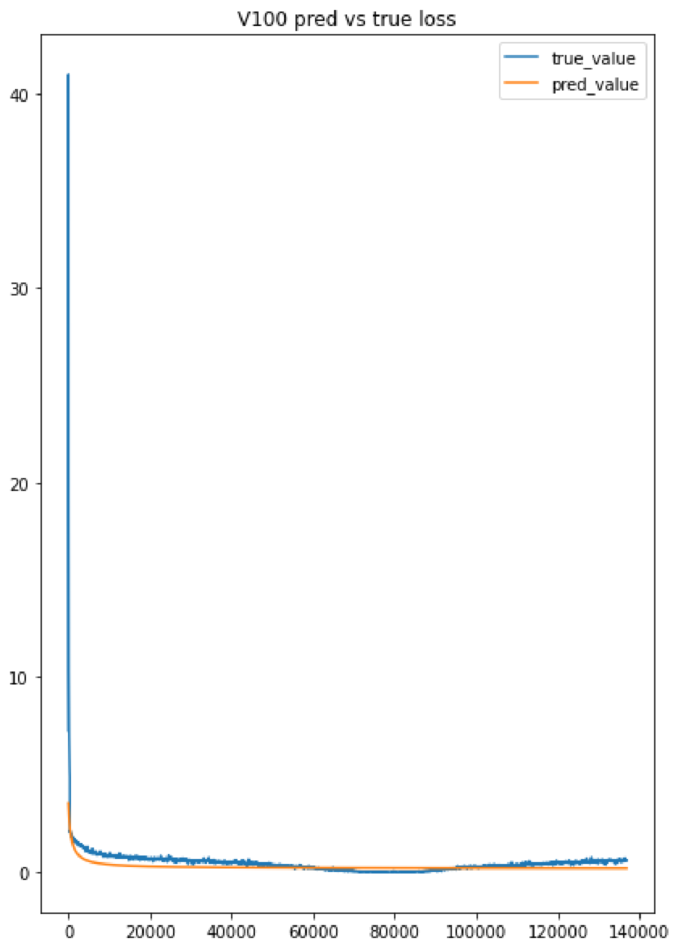
```
import matplotlib.pyplot as plt
for i in range(len(gpu_names)):
    fig,ax = plt.subplots(1,2, figsize = (15,10))
    ax[0].plot(resnet_50_data[i], label = 'true_value')
    ax[0].plot(pred_losses[i], label = 'pred_value')
    ax[0].set_title(f'{gpu_names[i]} pred vs true loss')
    ax[0].legend()
    ax[1].plot(percentage_error_all[i],)
    ax[1].set_title(f'{gpu_names[i]} percentage error')
    fig.show()
```

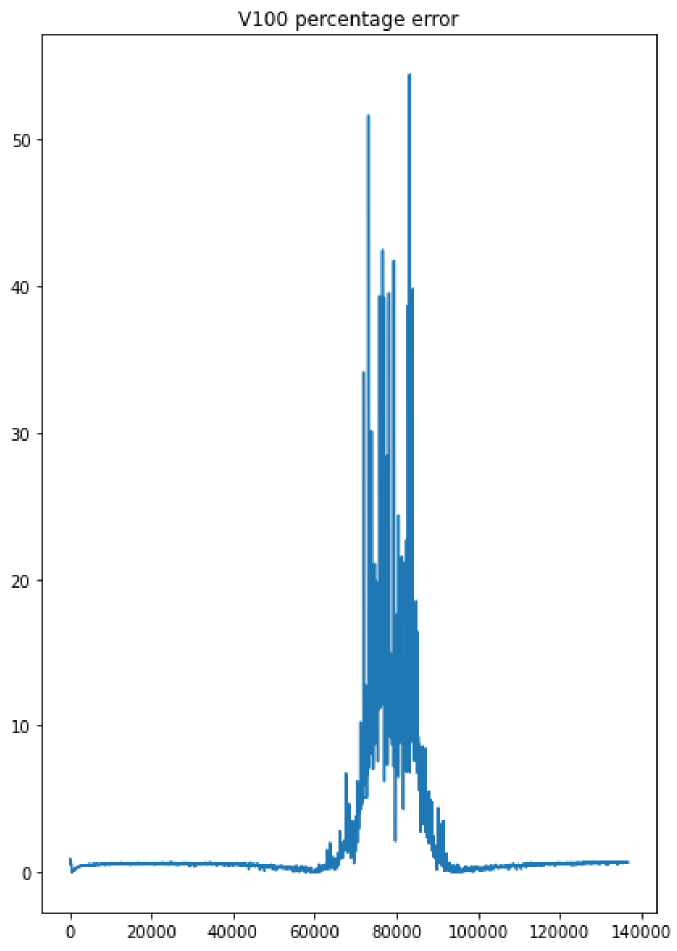


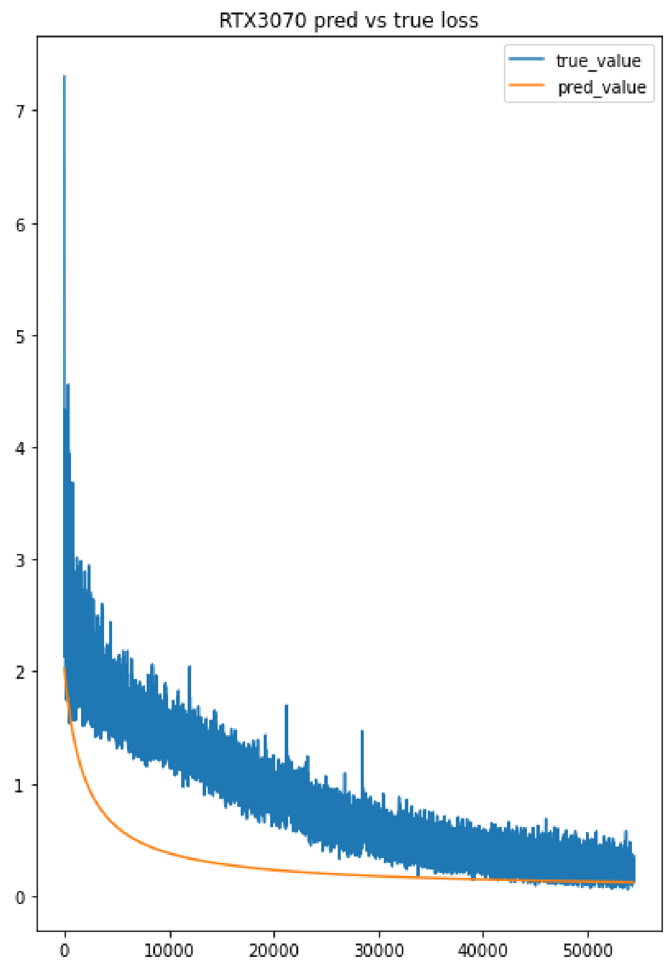


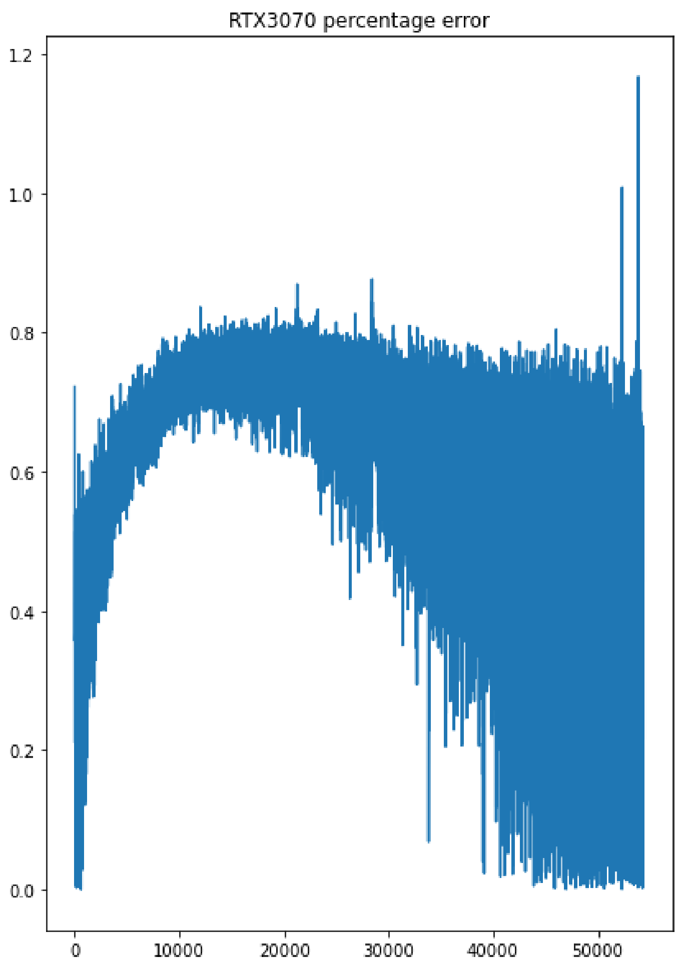

P100 pred vs true loss
true_value
pred_value


P100 percentage error


V100 pred vs true loss
true_value
pred_value


V100 percentage error





- Since the data from my team mates are trained on full 350 epochs and mine 3070 data are trained using early stop (92% acc) so the plot may seems different from each other, but the result helds same.
- Both four GPUs reachs the same level of loss and remain stable at around 50000 steps.

(3)

- Using the data from peng'spaper

In [175…

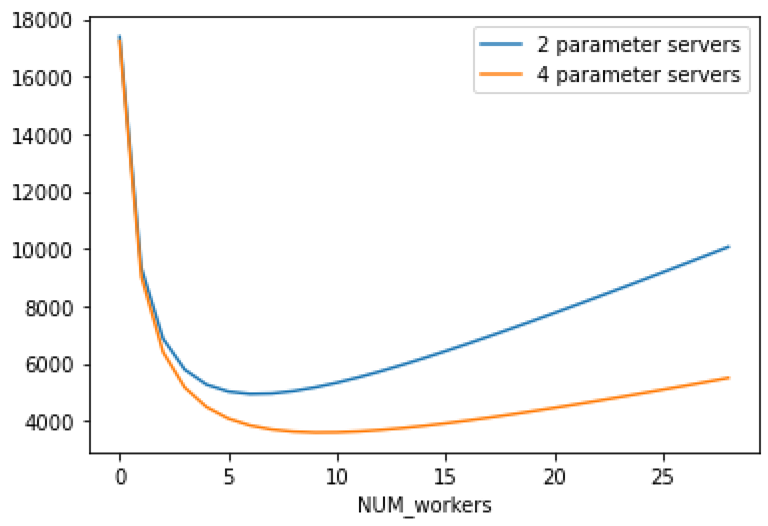
```
theta_0 = 1.02
theta_1 = 2.78
theta_2 = 4.92
theta_3 = 0.00
theta_4 = 0.02
def paper_function(w, p, m):
    return (theta_0 * m/w + theta_1 + theta_2 * w/p + theta_3 * w + theta_4 * p)**-1
parameter_servers = [2, 4]
vis_x = np.arange(1, 30)
data = []
m = 128
predicted_epochs = 128
training_data_size = 50000

f_p_2 = paper_function(vis_x, 2, M)
time_to_acc_2 = predicted_epochs / f_p_2

f_p_4 = paper_function(vis_x, 4, M)
time_to_acc_4 = predicted_epochs / f_p_4
```

In [179…

```
plt.plot(time_to_acc_2, label = '2 parameter servers')
plt.plot(time_to_acc_4, label = '4 parameter servers')
plt.legend()
plt.xlabel('NUM_workers')
```

Out[179…

```
Text(0.5, 0, 'NUM_workers')
```


NUM_workers
2 parameter servers
4 parameter servers