Kabir_Notebook_7

December 9, 2024

[ ]:

```
```

[1]:

```
import torch
from torch import nn
import torch.nn.functional as F
from collections import OrderedDict
from sklearn.metrics import accuracy_score, precision_score, recall_score,␣
↪f1_score, confusion_matrix, classification_report
import matplotlib.pyplot as plt
from torchvision import datasets
import torchvision.transforms as transforms
from torch.utils.data import DataLoader
import time
from IPython.core.magic import register_cell_magic
import gc
from torch.amp import autocast, GradScaler
from torchtnt.utils.data import CudaDataPrefetcher



@register_cell_magic
def skip(line, cell):
    return

device = 'cuda'
```

```
/home/super/.local/lib/python3.11/site-packages/tqdm/auto.py:21: TqdmWarning:
IProgress not found. Please update jupyter and ipywidgets. See
https://ipywidgets.readthedocs.io/en/stable/user_install.html
  from .autonotebook import tqdm as notebook_tqdm
```

[2]:

```
dl = False
data_path = './data'

# Load CIFAR-10 dataset with the simple transform
cifar10_train = datasets.CIFAR10(data_path, train=True, download=dl,␣
↪transform=transforms.ToTensor())
try:
```

```
    mean = torch.load('data/mean.pt')
    std = torch.load('data/std.pt')
except FileNotFoundError:
    print("Computing Mean and Std")
    train_imgs = torch.stack([img for img, _ in cifar10_train], dim=3)#.
 ↪to(device=device)
    view = train_imgs.view(3, -1)#.to(device=device)

    mean = train_imgs.view(3, -1).mean(dim=1)
    std = train_imgs.view(3, -1).std(dim=1)

    transform = transforms.Compose([
        transforms.ToTensor(),
        transforms.Normalize(mean, std)
    ])

    torch.save(mean, 'data/mean.pt')
    torch.save(std, 'data/std.pt')

# Define the transform with normalization
transform = transforms.Compose([
    transforms.ToTensor(),
    transforms.Normalize(mean, std)
])
print("Mean: ", mean)
print("Std: ", std)
cifar10_train = datasets.CIFAR10(data_path, train=True, download=dl,␣
 ↪transform=transform)
cifar10_test = datasets.CIFAR10(data_path, train=False, download=dl,␣
 ↪transform=transform)
```

```
/tmp/ipykernel_1721992/2313373535.py:7: FutureWarning: You are using
`torch.load` with `weights_only=False` (the current default value), which uses
the default pickle module implicitly. It is possible to construct malicious
pickle data which will execute arbitrary code during unpickling (See
https://github.com/pytorch/pytorch/blob/main/SECURITY.md#untrusted-models for
more details). In a future release, the default value for `weights_only` will be
flipped to `True`. This limits the functions that could be executed during
unpickling. Arbitrary objects will no longer be allowed to be loaded via this
mode unless they are explicitly allowlisted by the user via
`torch.serialization.add_safe_globals`. We recommend you start setting
`weights_only=True` for any use case where you don't have full control of the
loaded file. Please open an issue on GitHub for any issues related to this
experimental feature.
  mean = torch.load('data/mean.pt')
/tmp/ipykernel_1721992/2313373535.py:8: FutureWarning: You are using
`torch.load` with `weights_only=False` (the current default value), which uses
```

```
the default pickle module implicitly. It is possible to construct malicious
pickle data which will execute arbitrary code during unpickling (See
https://github.com/pytorch/pytorch/blob/main/SECURITY.md#untrusted-models for
more details). In a future release, the default value for `weights_only` will be
flipped to `True`. This limits the functions that could be executed during
unpickling. Arbitrary objects will no longer be allowed to be loaded via this
mode unless they are explicitly allowlisted by the user via
`torch.serialization.add_safe_globals`. We recommend you start setting
`weights_only=True` for any use case where you don't have full control of the
loaded file. Please open an issue on GitHub for any issues related to this
experimental feature.
  std = torch.load('data/std.pt')
```

```
Mean:  tensor([0.4914, 0.4822, 0.4465])
Std:  tensor([0.2470, 0.2435, 0.2616])
```

[3]:
```python
class Classifier(nn.Module):
    @classmethod
    def compare_results(cls, results1, results2):
        print('Comparing results:')
        comparisons = {
            'accuracy': 100*(results1['accuracy'] - results2['accuracy'])/
↪results1['accuracy'],
            'precision': 100*(results1['precision'] - results2['precision'])/
↪results1['precision'],
            'recall': 100*(results1['recall'] - results2['recall'])/
↪results1['recall'],
            'f1': 100*(results1['f1'] - results2['f1'])/results1['f1']
        }
        for key, value in comparisons.items():
            print(f'{key}: {value} %')

    def __init__(self):
        super().__init__()

    def forward(self, x):
        return self.stack(x)
    def predict(self, x):
        with torch.no_grad():
            self.eval()
            return self.forward(x).argmax(dim=1)

    def train_model(
        self,
        epochs,
        train_loader,
        test_loader,
        train_len,
```

```
        test_len,
        test_size,
        loss_fn=nn.CrossEntropyLoss(),
        optimizer=torch.optim.SGD,
        optimizer_args = [],
        optimizer_kwargs = {},
        print_epoch=10,
        header_epoch = 15,
        sched_factor = 0.1,
        sched_patience = 5
    ):

        scaler = GradScaler("cuda")
        optimizer = optimizer(self.parameters(), *optimizer_args, **optimizer_kwargs)
        scheduler = torch.optim.lr_scheduler.ReduceLROnPlateau(optimizer, 'max', patience=sched_patience, factor=sched_factor)
        training_time = 0
        train_hist = torch.zeros(epochs, device=device)
        test_hist = torch.zeros(epochs, device=device)
        accuracy_hist = torch.zeros(epochs, device=device)

        cell_width = 20
        header_form_spec = f'^{cell_width}'

        epoch_inspection = {
            "Epoch": 0,
            "Epoch Time (s)": 0,
            "Training Loss": 0,
            "Test Loss ": 0,
            "Overfit (%)": 0,
            "Accuracy (%)": 0,
            "Δ Accuracy (%)": 0,
            "GPU Memory (GiB)": 0
        }

        header_string = "|"
        for key in epoch_inspection.keys():
            header_string += (f"{key:{header_form_spec}}|")

        divider_string = '-'*len(header_string)
        if print_epoch:
            print(f'Training {self.__class__.__name__}\n')
            print(divider_string)
        max_accuracy = torch.zeros(1, device=device)
        for epoch in range(epochs):
            begin_epoch = time.time()
```

```
            self.train()

            start_time = time.time()
            train_loss = 0
            for X_batch, Y_batch in train_loader:
                #X_batch, Y_batch = X_batch.to(device, non_blocking=True),␣
↪Y_batch.to(device, non_blocking=True)
                optimizer.zero_grad(set_to_none=True)
                with autocast("cuda"):
                    Y_pred = self.forward(X_batch)
                    loss = loss_fn(Y_pred, Y_batch)
                scaler.scale(loss).backward()
                scaler.step(optimizer)
                scaler.update()

                train_loss += loss
            training_time += time.time() - start_time

            train_loss = train_loss/train_len
            train_hist[epoch] = train_loss


            self.eval()
            with torch.no_grad():
                test_loss = torch.zeros(1, device=device)
                correct = torch.zeros(1, device=device)

                for X_test_batch, Y_test_batch in test_loader:
                    #X_test_batch, Y_test_batch = X_test_batch.to(device,␣
↪non_blocking=True), Y_test_batch.to(device, non_blocking=True)

                    out = self.forward(X_test_batch)
                    test_loss += loss_fn(out, Y_test_batch)
                    correct += (out.argmax(dim=1) == Y_test_batch).sum()

            test_loss = test_loss/test_len
            test_hist[epoch] = test_loss
            accuracy = correct/test_size
            accuracy_hist[epoch] = accuracy

            scheduler.step(accuracy)

            end_epoch = time.time()
            if print_epoch and (epoch % print_epoch == 0 or epoch == epochs -␣
↪1) :
                mem = (torch.cuda.memory_allocated() + torch.cuda.
↪memory_reserved())/1024**3
```

```
                if header_epoch and epoch % header_epoch == 0:
                    print(header_string)
                    print(divider_string)
                epoch_duration = end_epoch - begin_epoch
                overfit = 100 * (test_loss - train_loss) / train_loss
                d_accuracy = torch.zeros(1) if max_accuracy == 0 else 100 *␣
↪(accuracy - max_accuracy) / max_accuracy
                if accuracy > max_accuracy:
                    max_accuracy = accuracy

                epoch_inspection['Epoch'] = f'{epoch}'
                epoch_inspection['Epoch Time (s)'] = f'{epoch_duration:4f}'
                epoch_inspection['Training Loss'] = f'{train_loss.item():8f}'
                epoch_inspection['Test Loss '] = f'{test_loss.item():8f}'
                epoch_inspection['Overfit (%)'] = f'{overfit.item():4f}'
                epoch_inspection['Accuracy (%)'] = f'{accuracy.item()*100:4f}'
                epoch_inspection['Δ Accuracy (%)'] = f'{d_accuracy.item():4f}'
                epoch_inspection["GPU Memory (GiB)"] = f'{mem:2f}'
                for value in epoch_inspection.values():
                    print(f"|{value:^{cell_width}}", end='')
                print('|')
                print(divider_string)


        print(f'\nTraining Time: {training_time} seconds\n')

        self.train_hist = train_hist
        self.test_hist = test_hist
        self.accuracy_hist = accuracy_hist

    def plot_training(self, title='Training Results'):
        plt.plot(self.train_hist.detach().cpu(), label='Training Loss')
        plt.plot(self.test_hist.detach().cpu(), label='Test Loss')
        plt.plot(self.accuracy_hist.detach().cpu(), label='Accuracy')
        plt.title(title)
        plt.xlabel('Epoch')
        plt.ylabel('Loss')
        plt.legend()
        plt.show()

    def get_results(self, Y_test=None, Y_pred=None):
        if Y_test is None:
            Y_test = self.last_test
        if Y_pred is None:
            Y_pred = self.last_pred

        if isinstance(Y_test, torch.Tensor):
```

```python
            Y_test = Y_test.cpu().detach().numpy()
        if isinstance(Y_pred, torch.Tensor):
            Y_pred = Y_pred.cpu().detach().numpy()
        results = {
            'accuracy': accuracy_score(Y_test, Y_pred),
            'precision': precision_score(Y_test, Y_pred, average='weighted'),
            'recall': recall_score(Y_test, Y_pred, average='weighted'),
            'f1': f1_score(Y_test, Y_pred, average='weighted'),
            'confusion_matrix': confusion_matrix(Y_test, Y_pred),
            'classification_report': classification_report(Y_test, Y_pred)
        }
        self.last_results = results
        return results
    def print_results(self, results=None):
        if results is None:
            try:
                results = self.last_results
            except:
                results = self.get_results()
        for key, value in results.items():
            if key in ['confusion_matrix', 'classification_report']:
                print(f'{key.capitalize()}:\n{value}')
            else:
                print(f'{key.capitalize()}: {value}')
```

[4]:
```python
class ConvImageClassifier(Classifier):
    def __init__(self, input_dim, conv_layers, fc_layers, activation=nn.ReLU):
        super().__init__()

        self.stack = nn.Sequential(OrderedDict(
            [
                ('conv0', nn.Conv2d(in_channels=3, out_channels=conv_layers[0],␣
↪kernel_size=3, padding=1)),
                ('activation0', activation()),
                ('maxpool0', nn.MaxPool2d(2)),
            ]
        ))

        for i in range(1, len(conv_layers)):
            self.stack.add_module(f'conv{i}', nn.
↪Conv2d(in_channels=conv_layers[i-1], out_channels=conv_layers[i],␣
↪kernel_size=3, padding=1))
            self.stack.add_module(f'activation{i}', activation())
            self.stack.add_module(f'maxpool{i}', nn.MaxPool2d(2))

        conv_out = input_dim//(2**len(conv_layers))
        self.stack.add_module('flatten', nn.Flatten())
```

```
        self.stack.add_module(f'fc0', nn.Linear(conv_out**2*conv_layers[-1],␣
↪fc_layers[0]))

        for i in range(1, len(fc_layers)):
            self.stack.add_module(f'activation_fc{i}', nn.Tanh())
            self.stack.add_module(f'fc{i}', nn.Linear(fc_layers[i-1],␣
↪fc_layers[i]))
```

[5]:

```
try:
    del train_loader
    del test_loader
    del model_1a
    del model_1b
    del resnet
    del train_loader_cuda
    del test_loader_cuda
except:
    pass

# Reset CUDA context
start = time.time()
torch.cuda.empty_cache()
torch.cuda.reset_peak_memory_stats()

gc.collect()


cifar10_train = datasets.CIFAR10(data_path, train=True, download=dl,␣
↪transform=transform)
cifar10_test = datasets.CIFAR10(data_path, train=False, download=dl,␣
↪transform=transform)

batch_size = int(2**11)
workers = 12
cpu_prefetch = 39
gpu_prefetch = 28

print('begin init train_loader')
train_loader = DataLoader(
    cifar10_train,
    batch_size=batch_size,
    shuffle=True,
    num_workers=workers,
    prefetch_factor=cpu_prefetch,
    pin_memory=True
)
```

```
X_batch = next(iter(train_loader))[0]
dtype_size = X_batch.element_size()
print(f"Batch Size: {X_batch.element_size() * X_batch.nelement() / 1024**2}␣
 ↪MiB")


print('begin init fetcher')
train_loader_cuda = CudaDataPrefetcher(
    data_iterable = train_loader,
    device = torch.device('cuda'),
    num_prefetch_batches=gpu_prefetch
)
test_loader = DataLoader(cifar10_test, batch_size=len(cifar10_test),␣
 ↪shuffle=True, num_workers=workers, pin_memory=True, prefetch_factor=1)
test_loader_cuda = CudaDataPrefetcher(
    data_iterable = test_loader,
    device = torch.device('cuda'),
    num_prefetch_batches=1
)

model_1a = ConvImageClassifier(
    input_dim = 32,
    conv_layers=[32, 64],
    fc_layers=[32, 10],
    activation=nn.ReLU
).to(device=device)
params = sum(p.numel() for p in model_1a.parameters() if p.requires_grad)
print(f"Total parameters: {params}")
print(model_1a.stack)

model_1a.train_model(
    epochs=200,
    train_loader=train_loader_cuda,
    train_len=len(train_loader),
    test_loader=test_loader_cuda,
    test_len=len(test_loader),
    test_size = len(cifar10_test),
    loss_fn=nn.CrossEntropyLoss(),
    optimizer=torch.optim.Adam,
    optimizer_kwargs={'lr': 1e-3, 'weight_decay': 1e-2},
    print_epoch=1,
    header_epoch=15,
    #Disable scheduling
    sched_patience = 200
)

del train_loader
```

```python
del test_loader

model_1a.plot_training("2 Layer CNN Training Curves")
```

```
begin init train_loader
Batch Size: 24.0 MiB
begin init fetcher
Total parameters: 150826
Sequential(
  (conv0): Conv2d(3, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (activation0): ReLU()
  (maxpool0): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (conv1): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (activation1): ReLU()
  (maxpool1): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (flatten): Flatten(start_dim=1, end_dim=-1)
  (fc0): Linear(in_features=4096, out_features=32, bias=True)
  (activation_fc1): Tanh()
  (fc1): Linear(in_features=32, out_features=10, bias=True)
)
Training ConvImageClassifier

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        Epoch         |    Epoch Time (s)    |    Training Loss     |      Test Loss
|     Overfit (%)      |     Accuracy (%)     |    Δ Accuracy (%)    |   GPU Memory
(GiB)   |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|          0           |       6.879051       |       1.937800       |       1.692992
|      -12.633269      |      42.219999       |       0.000000       |       4.213852
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|          1           |       4.994837       |       1.605037       |       1.515875
|      -5.555122       |      48.749998       |      15.466600       |       4.916979
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|          2           |       5.057521       |       1.465897       |       1.414413
|      -3.512099       |      51.569998       |       5.784616       |       5.618151
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
----------
|       3          |     4.972616      |     1.375477      |     1.350838
|    -1.791298     |    53.240001      |     3.238323      |     6.319323
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
----------
|       4          |     4.927670      |     1.308575      |     1.290943
|    -1.347484     |    55.500001      |     4.244927      |     7.020495
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
----------
|       5          |     5.044797      |     1.260512      |     1.249116
|    -0.904100     |    57.049996      |     2.792784      |     7.721667
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
----------
|       6          |     4.946557      |     1.217713      |     1.213817
|    -0.319954     |    58.370000      |     2.313768      |     8.422839
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
----------
|       7          |     4.963850      |     1.191438      |     1.185038
|    -0.537155     |    59.389997      |     1.747467      |     9.124011
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
----------
|       8          |     4.937151      |     1.141895      |     1.157622
|     1.377297     |    59.560001      |     0.286251      |     9.825182
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
----------
|       9          |     4.935848      |     1.111418      |     1.121968
|     0.949261     |    61.479998      |     3.223634      |    10.526354
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
----------
|      10          |     4.930698      |     1.078762      |     1.106483
|     2.569699     |    62.089998      |     0.992192      |    11.227526
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         11         |      4.939031      |      1.069479      |      1.099626
|      2.818901      |     61.820000      |     -0.434848      |     11.928698
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         12         |      4.886281      |      1.038594      |      1.061449
|      2.200494      |     63.379997      |      2.077629      |     12.629870
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         13         |      4.995658      |      1.013100      |      1.046757
|      3.322247      |     63.709998      |      0.520671      |     13.331042
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         14         |      4.999271      |      0.998499      |      1.042943
|      4.451067      |     64.179999      |      0.737719      |     14.032214
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         15         |      4.955255      |      0.976865      |      1.017815
|      4.191998      |     65.099996      |      1.433465      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         16         |      4.942564      |      0.963070      |      1.021136
|      6.029279      |     64.489996      |     -0.937020      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         17         |      5.041818      |      0.945570      |      0.997559
|      5.498090      |     66.139996      |      1.597542      |     14.733386
|
```

18	5.018600	0.936561	0.982521	4.907250	66.490000	0.529187	14.733386
19	4.997144	0.920925	0.977600	6.154187	66.719997	0.345911	14.733386
20	5.019781	0.907129	0.972335	7.188182	66.649997	-0.104916	14.733386
21	5.009535	0.895712	0.975230	8.877600	66.740000	0.029981	14.733386
22	5.038126	0.891375	0.960982	7.808997	67.089999	0.524421	14.733386
23	5.222510	0.878327	0.942524	7.308948	67.780000	1.028471	14.733386
24	5.124631	0.869255	0.973694	12.014804	66.719997	-1.563888	14.733386
25	5.112703	0.867551	0.953081	9.858801	67.009997	-1.136032	14.733386

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         26          |      5.065426       |       0.851950       |      0.934716
|      9.714802       |     68.000001       |       0.324581       |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         27          |      5.058987       |       0.834209       |      0.945868
|     13.385069       |     67.490000       |      -0.750002       |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         28          |      5.017909       |       0.846333       |      0.925520
|      9.356506       |     68.759996       |       1.117640       |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         29          |      5.092650       |       0.823023       |      0.908128
|     10.340587       |     69.000000       |       0.349046       |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch         |   Epoch Time (s)    |   Training Loss      |     Test Loss
|    Overfit (%)      |    Accuracy (%)     |  Δ Accuracy (%)      |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         30          |      4.918151       |       0.822712       |      0.922056
|     12.075120       |     68.349999       |      -0.942030       |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         31          |      4.989425       |       0.818894       |      0.907556
|     10.827121       |     69.330001       |       0.478262       |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         32          |      4.983510       |       0.813875       |      0.901607
|     10.779591       |     69.180000       |      -0.216358       |     14.733386
|
```

33	5.009483	0.798466	0.910532	14.035186	68.919998	-0.591378	14.733386
34	4.931305	0.800569	0.899236	12.324631	69.239998	-0.129818	14.733386
35	5.006773	0.788863	0.893350	13.245255	69.479996	0.216349	14.733386
36	4.997746	0.796416	0.895868	12.487496	69.229996	-0.359815	14.733386
37	5.044682	0.796336	0.893554	12.208151	69.309998	-0.244672	14.733386
38	5.000170	0.779699	0.882630	13.201449	69.849998	0.532530	14.733386
39	4.996383	0.775605	0.881047	13.594745	70.029998	0.257695	14.733386
40	5.043076	0.776948	0.884165	13.799740	69.940001	-0.128512	14.733386

41	5.041635	0.773695	0.893078	15.430255	69.239998	-1.128088	14.733386
42	5.058732	0.770552	0.878690	14.033852	70.269996	0.342708	14.733386
43	5.043230	0.754261	0.870302	15.384832	70.139998	-0.184997	14.733386
44	4.998542	0.752734	0.874776	16.213068	70.080000	-0.270380	14.733386
Epoch	Epoch Time (s)	Training Loss	Test Loss	Overfit (%)	Accuracy (%)	Δ Accuracy (%)	GPU Memory (GiB)
45	5.019987	0.743835	0.855158	14.966059	71.209997	1.337699	14.733386
46	5.004228	0.745706	0.872013	16.937983	70.440000	-1.081304	14.733386
47	5.011259	0.743847	0.855449	15.003299	71.039999	-0.238728	14.733386

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         48          |       4.993983       |       0.735898       |       0.864275
|     17.444962       |      70.899999       |      -0.435329       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         49          |       5.076868       |       0.732867       |       0.872597
|     19.066174       |      70.220000       |      -1.390251       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         50          |       5.019500       |       0.735458       |       0.866482
|     17.815176       |      70.419997       |      -1.109395       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         51          |       4.996568       |       0.731122       |       0.863331
|     18.082968       |      70.519996       |      -0.968967       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         52          |       4.973970       |       0.723913       |       0.852979
|     17.829021       |      71.109998       |      -0.140428       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         53          |       4.986421       |       0.720687       |       0.850710
|     18.041557       |      70.980000       |      -0.322983       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         54          |       5.057110       |       0.717377       |       0.881610
|     22.893658       |      70.099998       |      -1.558769       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         55          |       5.026342       |       0.717588       |       0.856848
|     19.406694       |      70.459998       |      -1.053222       |      14.733386
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        56          |      5.025057      |      0.713417      |      0.882737
|     23.733633      |     69.510001      |     -2.387300      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        57          |      5.034989      |      0.713721      |      0.839316
|     17.597082      |     71.660000      |      0.631938      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        58          |      5.013689      |      0.704705      |      0.855010
|     21.328808      |     70.669997      |     -1.381528      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        59          |      5.026370      |      0.708809      |      0.847124
|     19.513639      |     71.270001      |     -0.544235      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        60          |      4.995232      |      0.709052      |      0.869824
|     22.674246      |     70.379996      |     -1.786218      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        61          |      5.050303      |      0.712813      |      0.855380
|     20.000652      |     70.469999      |     -1.660621      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        62          |      5.017296      |      0.701495      |      0.868435
|     23.797775      |     70.400000      |     -1.758304      |     14.733386
|
```

63	5.017220	0.701392	0.841186	19.931078	71.590000	-0.097683	14.733386
64	5.014618	0.691178	0.828895	19.924889	72.029996	0.516321	14.733386
65	5.022207	0.691030	0.834924	20.823246	71.590000	-0.610850	14.733386
66	5.087819	0.679771	0.853366	25.537222	71.079999	-1.318890	14.733386
67	5.001268	0.692515	0.828224	19.596493	72.139996	0.152715	14.733386
68	4.985179	0.684203	0.820484	19.918194	72.349995	0.291099	14.733386
69	4.998172	0.680681	0.838111	23.128300	71.329999	-1.409809	14.733386
70	4.971375	0.687074	0.845096	22.999332	71.399999	-1.313057	14.733386

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         71          |      5.033550      |      0.677085      |      0.833596
|     23.115574      |     71.980000      |     -0.511397      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         72          |      5.027353      |      0.674671      |      0.835942
|     23.903578      |     71.669996      |     -0.939875      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         73          |      5.011705      |      0.674149      |      0.846878
|     25.621727      |     71.109998      |     -1.713887      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         74          |      5.005163      |      0.669994      |      0.822554
|     22.770489      |     72.249997      |     -0.138215      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |   Δ Accuracy (%)   |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         75          |      5.051726      |      0.661650      |      0.827408
|     25.052343      |     72.450000      |      0.138223      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         76          |      5.025922      |      0.671461      |      0.837118
|     24.671137      |     71.599996      |     -1.173228      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         77          |      5.002988      |      0.672062      |      0.834897
|     24.229179      |     71.619999      |     -1.145619      |     14.733386
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         78         |      5.037694      |      0.674381      |      0.852034
|     26.343037      |     71.029997      |     -1.959977      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         79         |      5.008406      |      0.673514      |      0.835284
|     24.018772      |     71.689999      |     -1.049001      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         80         |      5.023103      |      0.664724      |      0.814567
|     22.542067      |     72.329998      |     -0.165634      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         81         |      5.032135      |      0.660016      |      0.832222
|     26.091066      |     71.520001      |     -1.283643      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         82         |      5.014982      |      0.658289      |      0.827561
|     25.713850      |     72.069997      |     -0.524504      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         83         |      4.990252      |      0.657215      |      0.830514
|     26.368765      |     72.169995      |     -0.386480      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         84         |      5.049405      |      0.646537      |      0.816893
|     26.349035      |     72.529995      |      0.110415      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         85         |      5.018591      |      0.647542      |      0.815600
|     25.953310      |     72.700000      |      0.234392      |     14.733386
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         86          |       5.033751       |       0.647618       |       0.812683
|     25.488043       |      72.389996       |      -0.426415       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         87          |       5.045481       |       0.646377       |       0.815934
|     26.231840       |      72.779995       |       0.110035       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         88          |       4.996143       |       0.648928       |       0.810473
|     24.894083       |      72.670001       |      -0.151133       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         89          |       5.029956       |       0.639077       |       0.808710
|     26.543398       |      73.259997       |       0.659525       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch         |   Epoch Time (s)     |   Training Loss      |     Test Loss
|    Overfit (%)      |    Accuracy (%)      |  Δ Accuracy (%)      |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         90          |       5.110249       |       0.641995       |       0.831552
|     29.526297       |      71.480000       |      -2.429699       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         91          |       4.982387       |       0.649811       |       0.825403
|     27.022100       |      71.929997       |      -1.815452       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         92          |       4.985412       |       0.635004       |       0.812187
|     27.902637       |      72.479999       |      -1.064699       |      14.733386
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         93         |      5.036469      |      0.644235      |      0.817088
|     26.830687      |     72.169995      |     -1.487854      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         94         |      5.012946      |      0.632602      |      0.819317
|     29.515427      |     71.819997      |     -1.965603      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         95         |      5.045994      |      0.633972      |      0.844893
|     33.269829      |     70.599997      |     -3.630904      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         96         |      4.988159      |      0.644581      |      0.841811
|     30.598265      |     71.309996      |     -2.661755      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         97         |      5.099520      |      0.642988      |      0.819964
|     27.524033      |     72.079998      |     -1.610700      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         98         |      5.076290      |      0.630325      |      0.864446
|     37.142879      |     70.159996      |     -4.231507      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         99         |      5.029150      |      0.642412      |      0.865549
|     34.734249      |     69.690001      |     -4.873050      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        100         |      5.074078      |      0.639341      |      0.831916
|     30.120871      |     71.419996      |     -2.511605      |     14.733386
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        101          |        5.076982      |        0.641404      |        0.846540
|     31.982470       |      70.739996       |      -3.439805       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        102          |        5.054729      |        0.635157      |        0.799278
|     25.839453       |      72.939998       |      -0.436800       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        103          |        5.141601      |        0.625099      |        0.819944
|     31.170256       |      72.029996       |      -1.678954       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        104          |        5.070658      |        0.622841      |        0.818336
|     31.387720       |      72.420001       |      -1.146596       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch         |   Epoch Time (s)     |   Training Loss      |     Test Loss
|    Overfit (%)      |    Accuracy (%)      |  Δ Accuracy (%)      |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        105          |        5.019599      |        0.628627      |        0.813345
|     29.384373       |      72.560000       |      -0.955497       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        106          |        5.024177      |        0.626152      |        0.827057
|     32.085606       |      71.630001       |      -2.224947       |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        107          |        5.013819      |        0.629211      |        0.829194
|     31.783247       |      71.639997       |      -2.211303       |      14.733386
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      108       |     4.982967     |     0.627714     |     0.834247
|   32.902443    |    71.789998     |    -2.006552     |    14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      109       |     5.065071     |     0.623302     |     0.814757
|   30.716171    |    72.819996     |    -0.600603     |    14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      110       |     5.052822     |     0.620189     |     0.816181
|   31.602104    |    72.520000     |    -1.010098     |    14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      111       |     5.039487     |     0.620462     |     0.830820
|   33.903545    |    71.509999     |    -2.388750     |    14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      112       |     5.014404     |     0.628477     |     0.807500
|   28.485163    |    72.859997     |    -0.546002     |    14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      113       |     5.031634     |     0.607373     |     0.806208
|   32.736797    |    72.799999     |    -0.627899     |    14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      114       |     5.058440     |     0.613070     |     0.832420
|   35.778915    |    71.499997     |    -2.402402     |    14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      115       |     5.137480     |     0.615059     |     0.818509
|   33.077972    |    72.349995     |    -1.242154     |    14.733386
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        116         |      5.018563      |      0.609816      |      0.812268
|     33.198799      |     72.399998      |     -1.173901      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        117         |      5.109816      |      0.607323      |      0.806634
|     32.818096      |     72.839999      |     -0.573298      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        118         |      5.038103      |      0.607644      |      0.810839
|     33.439774      |     72.499996      |     -1.037402      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        119         |      5.046576      |      0.606106      |      0.822247
|     35.660679      |     71.880001      |     -1.883697      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        120         |      5.127125      |      0.604725      |      0.795232
|     31.503040      |     73.170000      |     -0.122846      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        121         |      5.092554      |      0.603490      |      0.803289
|     33.107285      |     72.639996      |     -0.846303      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        122         |      5.157858      |      0.607773      |      0.808804
|     33.076607      |     72.589999      |     -0.914548      |     14.733386
|
```

123	5.063401	0.606742	0.810017	33.502678	72.380000	-1.201197	14.733386
124	5.077441	0.600606	0.790615	31.636356	73.249996	-0.013652	14.733386
125	5.053849	0.614245	0.812236	32.233276	72.359997	-1.228502	14.733386
126	5.084872	0.599685	0.803532	33.992432	72.589999	-0.914548	14.733386
127	5.167486	0.599114	0.815815	36.170300	72.490001	-1.051046	14.733386
128	5.089398	0.602681	0.808109	34.085697	72.829998	-0.586950	14.733386
129	5.076684	0.600694	0.800499	33.262379	73.109996	-0.204752	14.733386
130	5.027541	0.597062	0.817026	36.840973	71.969998	-1.760851	14.733386

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        131         |      5.041021       |       0.594622      |       0.802022
|     34.879238      |     72.869998       |      -0.532349      |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        132         |      5.119192       |       0.598600      |       0.819156
|     36.845387      |     71.880001       |      -1.883697      |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        133         |      5.096864       |       0.595433      |       0.801509
|     34.609364      |     73.139995       |      -0.163803      |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        134         |      5.065614       |       0.587363      |       0.809745
|     37.861038      |     72.679996       |      -0.791702      |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)    |   Training Loss     |     Test Loss
|    Overfit (%)     |    Accuracy (%)     |  Δ Accuracy (%)     |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        135         |      5.141582       |       0.584925      |       0.789208
|     34.924698      |     73.619998       |       0.491401      |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        136         |      5.089416       |       0.588752      |       0.831994
|     41.314877      |     71.779996       |      -2.499323      |      14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        137         |      5.073261       |       0.602229      |       0.816510
|     35.581394      |     72.209996       |      -1.915243      |      14.733386
|
```

138	5.051284	0.588933	0.806700	36.976452	72.749996	-1.181746	14.733386
139	5.172863	0.596244	0.793793	33.132214	73.219997	-0.543332	14.733386
140	5.030095	0.589227	0.799713	35.722385	72.749996	-1.181746	14.733386
141	5.064724	0.587496	0.805443	37.097630	72.929996	-0.937247	14.733386
142	5.317110	0.589445	0.804893	36.550980	72.659999	-1.303991	14.733386
143	5.075824	0.595370	0.822761	38.193150	72.189999	-1.942405	14.733386
144	5.127365	0.595650	0.820195	37.697517	72.229999	-1.888072	14.733386
145	5.070822	0.589687	0.787928	33.618114	73.519999	-0.135831	14.733386

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       146         |      5.024504     |      0.584502     |      0.831057
|    42.182186      |     71.639997     |     -2.689488     |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       147         |      5.040956     |      0.586660     |      0.796234
|    35.723175      |     72.999996     |     -0.842165     |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       148         |      5.033256     |      0.580001     |      0.790484
|    36.290077      |     73.579997     |     -0.054334     |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       149         |      5.084164     |      0.583147     |      0.831220
|    42.540440      |     71.969998     |     -2.241238     |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      Epoch        |   Epoch Time (s)  |   Training Loss   |     Test Loss
|   Overfit (%)     |    Accuracy (%)   |  Δ Accuracy (%)   |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       150         |      5.084607     |      0.597433     |      0.805152
|    34.768665      |     72.340000     |     -1.738655     |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       151         |      5.077116     |      0.578795     |      0.797499
|    37.786106      |     72.889996     |     -0.991581     |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       152         |      5.062029     |      0.580120     |      0.799035
|    37.736263      |     73.049998     |     -0.774245     |     14.733386
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       153        |      5.031142      |      0.578993      |      0.802004
|    38.517136     |     72.829998      |     -1.073078      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       154        |      5.084552      |      0.580315      |      0.816360
|    40.675404     |     72.499996      |     -1.521327      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       155        |      5.067575      |      0.578752      |      0.804299
|    38.971237     |     72.380000      |     -1.684321      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       156        |      5.113379      |      0.579363      |      0.794566
|    37.144699     |     73.429996      |     -0.258084      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       157        |      5.017329      |      0.578061      |      0.798019
|    38.051037     |     73.069996      |     -0.747082      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       158        |      5.101545      |      0.575089      |      0.804189
|    39.837406     |     72.529995      |     -1.480579      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       159        |      5.028152      |      0.571383      |      0.781985
|    36.858311     |     73.569995      |     -0.067920      |     14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       160        |      5.082304      |      0.580265      |      0.817013
|    40.800018     |     72.259998      |     -1.847323      |     14.733386
|
```

161	5.036721	0.578746	0.811182	40.162014	72.319996	-1.765826	14.733386
162	5.045552	0.588903	0.808974	37.369617	72.439998	-1.602824	14.733386
163	5.086725	0.584779	0.819144	40.077641	72.369999	-1.697907	14.733386
164	5.100532	0.570557	0.798689	39.984184	72.950000	-0.910076	14.733386

Epoch	Epoch Time (s)	Training Loss	Test Loss	Overfit (%)	Accuracy (%)	Δ Accuracy (%)	GPU Memory (GiB)
165	5.097234	0.566995	0.822130	44.997696	72.149998	-1.996740	14.733386
166	5.046937	0.571805	0.786729	37.586933	73.299998	-0.434664	14.733386
167	5.059964	0.572341	0.806426	40.899544	72.649997	-1.317577	14.733386

168	5.083268	0.578109	0.803447	38.978401	72.819996	-1.086663	14.733386
169	5.034514	0.575779	0.808265	40.377739	72.499996	-1.521327	14.733386
170	5.096079	0.571691	0.840251	46.976452	71.059996	-3.477318	14.733386
171	5.090700	0.577162	0.804536	39.395245	72.749996	-1.181746	14.733386
172	5.041945	0.570497	0.800041	40.235794	72.889996	-0.991581	14.733386
173	5.052004	0.574199	0.788621	37.342800	73.259997	-0.488998	14.733386
174	5.052775	0.572699	0.799113	39.534668	73.420000	-0.271662	14.733386
175	5.074750	0.567056	0.807956	42.482597	72.109997	-2.051074	14.733386

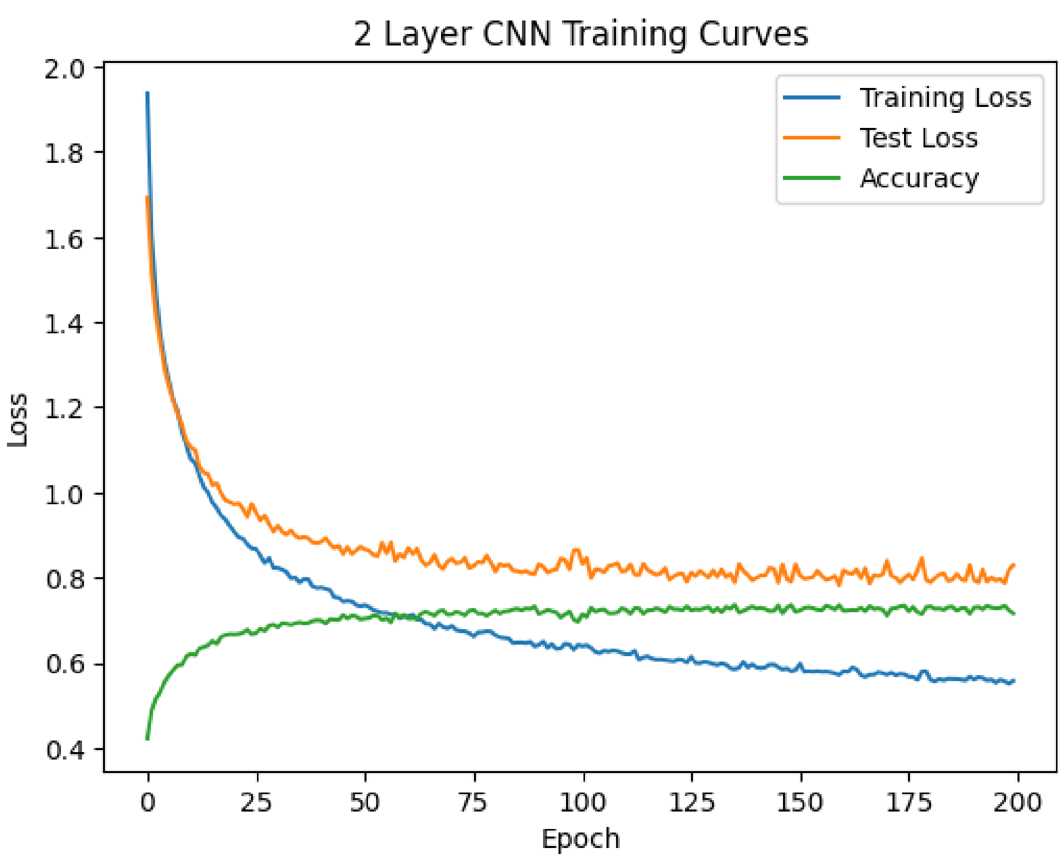
```
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|        176         |      5.063622      |      0.566776      |      0.795557
|     40.365253      |     73.119998      |     -0.679163      |     14.733386
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|        177         |      5.112973      |      0.560236      |      0.817368
|     45.897247      |     71.889997      |     -2.349906      |     14.733386
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|        178         |      5.141950      |      0.579844      |      0.846047
|     45.909477      |     71.459997      |     -2.933986      |     14.733386
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|        179         |      5.097329      |      0.579943      |      0.794550
|     37.004822      |     73.069996      |     -0.747082      |     14.733386
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|        180         |      5.048506      |      0.560042      |      0.789484
|     40.968742      |     73.219997      |     -0.543332      |     14.733386
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|        181         |      5.058635      |      0.556483      |      0.799255
|     43.626152      |     72.679996      |     -1.276828      |     14.733386
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|        182         |      5.078006      |      0.561945      |      0.807073
|     43.621426      |     72.779995      |     -1.140997      |     14.733386
|
```

183	5.123325	0.559224	0.810336	44.903713	72.469997	-1.562076	14.733386
184	5.105889	0.562512	0.792792	40.937675	73.049998	-0.774245	14.733386
185	5.135925	0.562172	0.794817	41.383297	73.170000	-0.611243	14.733386
186	5.117503	0.562027	0.808710	43.891663	72.289997	-1.806575	14.733386
187	5.105713	0.560768	0.796730	42.078255	72.819996	-1.086663	14.733386
188	5.096840	0.558971	0.793019	41.871277	72.889996	-0.991581	14.733386
189	5.079002	0.567355	0.829393	46.185753	71.509999	-2.866067	14.733386
190	5.047444	0.560887	0.794382	41.629650	72.979999	-0.869328	14.733386

```
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
---------
|         191          |       5.137882       |       0.566190       |       0.788644
|      39.289711       |      73.119998       |      -0.679163       |      14.733386
|
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
---------
|         192          |       5.176906       |       0.564188       |       0.819275
|      45.213093       |      72.579998       |      -1.412659       |      14.733386
|
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
---------
|         193          |       5.095566       |       0.558991       |       0.790708
|      41.452759       |      73.390001       |      -0.312410       |      14.733386
|
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
---------
|         194          |       5.114154       |       0.560496       |       0.800392
|      42.800682       |      72.979999       |      -0.869328       |      14.733386
|
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
---------
|        Epoch         |   Epoch Time (s)     |    Training Loss     |      Test Loss
|     Overfit (%)      |     Accuracy (%)     |   Δ Accuracy (%)     |   GPU Memory
(GiB)  |
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
---------
|         195          |       5.104700       |       0.553329       |       0.793020
|      43.317894       |      72.819996       |      -1.086663       |      14.733386
|
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
---------
|         196          |       5.060708       |       0.560255       |       0.796791
|      42.219360       |      72.920001       |      -0.950824       |      14.733386
|
-----------------------------------------------------------------------------------
-----------------------------------------------------------------------------------
---------
|         197          |       5.068001       |       0.556775       |       0.787571
|      41.452240       |      73.390001       |      -0.312410       |      14.733386
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         198          |        5.114279       |        0.551170       |        0.818402
|      48.484344       |       72.380000       |       -1.684321       |       14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         199          |        5.161066       |        0.557756       |        0.828847
|      48.603931       |       71.619999       |       -2.716650       |       14.733386
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------

Training Time: 465.7891447544098 seconds
```



[7]:
```python
try:
    del train_loader
    del test_loader
    del model_1a
    del model_1b
    del resnet
    del train_loader_cuda
    del test_loader_cuda
except:
    pass

# Reset CUDA context
start = time.time()
torch.cuda.empty_cache()
torch.cuda.reset_peak_memory_stats()

gc.collect()


cifar10_train = datasets.CIFAR10(data_path, train=True, download=dl,␣
 ↪transform=transform)
cifar10_test = datasets.CIFAR10(data_path, train=False, download=dl,␣
 ↪transform=transform)

batch_size = int(2**11)
workers = 12
cpu_prefetch = 39
gpu_prefetch = 28

print('begin init train_loader')
train_loader = DataLoader(
    cifar10_train,
    batch_size=batch_size,
    shuffle=True,
    num_workers=workers,
    prefetch_factor=cpu_prefetch,
    pin_memory=True
)

X_batch = next(iter(train_loader))[0]
dtype_size = X_batch.element_size()
print(f"Batch Size: {X_batch.element_size() * X_batch.nelement() / 1024**2}␣
 ↪MiB")


print('begin init fetcher')
train_loader_cuda = CudaDataPrefetcher(
```

```
    data_iterable = train_loader,
    device = torch.device('cuda'),
    num_prefetch_batches=gpu_prefetch
)
test_loader = DataLoader(cifar10_test, batch_size=len(cifar10_test),␣
↪shuffle=True, num_workers=workers, pin_memory=True, prefetch_factor=1)
test_loader_cuda = CudaDataPrefetcher(
    data_iterable = test_loader,
    device = torch.device('cuda'),
    num_prefetch_batches=1
)

model_1b = ConvImageClassifier(
    input_dim = 32,
    conv_layers=[32, 64, 128],
    fc_layers=[32, 10],
    activation=nn.ReLU
).to(device=device)
params = sum(p.numel() for p in model_1b.parameters() if p.requires_grad)
print(f"Total parameters: {params}")

print(model_1b.stack)

model_1b.train_model(
    epochs=200,
    train_loader=train_loader_cuda,
    train_len=len(train_loader),
    test_loader=test_loader_cuda,
    test_len=len(test_loader),
    test_size = len(cifar10_test),
    optimizer = torch.optim.Adam,
    optimizer_kwargs={'lr': 4e-4, 'weight_decay': 1e-2}, #Increase alpha to 2␣
↪next time
    loss_fn=nn.CrossEntropyLoss(),
    print_epoch=1,
    #Disable scheduling
    sched_patience = 200
)
del train_loader
del test_loader
model_1b.plot_training("3 Layer CNN Training Curves")
```

```
begin init train_loader
Batch Size: 24.0 MiB
begin init fetcher
Sequential(
  (conv0): Conv2d(3, 32, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
```

```
  (activation0): ReLU()
  (maxpool0): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (conv1): Conv2d(32, 64, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (activation1): ReLU()
  (maxpool1): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (conv2): Conv2d(64, 128, kernel_size=(3, 3), stride=(1, 1), padding=(1, 1))
  (activation2): ReLU()
  (maxpool2): MaxPool2d(kernel_size=2, stride=2, padding=0, dilation=1,
ceil_mode=False)
  (flatten): Flatten(start_dim=1, end_dim=-1)
  (fc0): Linear(in_features=2048, out_features=32, bias=True)
  (activation_fc1): Tanh()
  (fc1): Linear(in_features=32, out_features=10, bias=True)
)
Training ConvImageClassifier

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      Epoch       |   Epoch Time (s)   |   Training Loss    |     Test Loss
|   Overfit (%)    |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        0         |      6.599083      |      2.123456      |      1.948487
|    -8.239858     |     30.339998      |      0.000000      |      4.397570
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        1         |      5.095938      |      1.884181      |      1.808225
|    -4.031254     |     36.960000      |     21.819387      |      5.098743
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        2         |      5.104386      |      1.763717      |      1.702580
|    -3.466377     |     41.149998      |     11.336574      |      5.799916
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        3         |      5.259627      |      1.671804      |      1.621082
|    -3.033968     |     43.860000      |      6.585669      |      6.501088
|
```

4	5.091378	1.603726	1.563755	-2.492332	45.389998	3.488367	7.202260
5	5.196554	1.546524	1.518240	-1.828888	47.219998	4.031724	7.903431
6	5.055001	1.505353	1.491611	-0.912913	47.829998	1.291825	8.604603
7	5.300217	1.466276	1.470768	0.306398	47.700000	-0.271791	9.305775
8	5.158196	1.436028	1.435479	-0.038194	50.019997	4.578716	10.006947
9	5.092443	1.406422	1.398012	-0.598020	51.370001	2.698928	10.708119
10	5.108890	1.377094	1.370624	-0.469776	52.129996	1.479453	11.409291
11	5.151833	1.352459	1.352640	0.013336	52.759999	1.208523	12.110463

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         12          |       5.107094       |        1.329279      |       1.321973
|     -0.549647       |      53.729999       |        1.838514      |      12.811635
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         13          |       5.113696       |        1.310188      |       1.304062
|     -0.467534       |      54.240000       |        0.949193      |      13.512806
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         14          |       5.095271       |        1.286842      |       1.288354
|      0.117547       |      55.339998       |        2.028019      |      14.213978
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch         |   Epoch Time (s)     |   Training Loss      |     Test Loss
|    Overfit (%)      |    Accuracy (%)      |   Δ Accuracy (%)     |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         15          |       5.078908       |        1.272123      |       1.276889
|      0.374658       |      55.229998       |       -0.198772      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         16          |       5.201680       |        1.258695      |       1.267410
|      0.692424       |      56.049997       |        1.282975      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         17          |       5.199535       |        1.244539      |       1.255717
|      0.898146       |      56.430000       |        0.677972      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         18          |       5.049537       |        1.226315      |       1.230540
|      0.344559       |      57.080001       |        1.151871      |      14.915150
|
```

19	5.060830	1.211456	1.219420	0.657372	57.639998	0.981075	14.915150
20	5.101258	1.194798	1.198727	0.328824	58.309996	1.162384	14.915150
21	5.119519	1.180924	1.191759	0.917526	58.340001	0.051458	14.915150
22	5.043655	1.173767	1.185677	1.014699	58.649999	0.531364	14.915150
23	5.075843	1.161882	1.175911	1.207450	58.849996	0.341002	14.915150
24	5.146863	1.150066	1.162871	1.113444	59.700000	1.444357	14.915150
25	5.158943	1.138020	1.153674	1.375481	59.990001	0.485763	14.915150
26	5.087031	1.123744	1.139605	1.411456	60.240000	0.416736	14.915150

```
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|         27          |       5.090375        |        1.116595       |       1.151081
|      3.088480       |      59.689999        |       -0.913018       |      14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|         28          |       5.113190        |        1.110461       |       1.126441
|      1.439053       |      60.899997        |        1.095611       |      14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|         29          |       5.114990        |        1.096704       |       1.135921
|      3.575929       |      60.310000        |       -0.968796       |      14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       Epoch         |   Epoch Time (s)      |    Training Loss      |     Test Loss
|    Overfit (%)      |    Accuracy (%)       |   Δ Accuracy (%)      |  GPU Memory
(GiB)  |
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|         30          |       5.186446        |        1.088094       |       1.108954
|      1.917087       |      61.430001        |        0.870287       |      14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|         31          |       5.093885        |        1.077307       |       1.121839
|      4.133659       |      60.749996        |       -1.106960       |      14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|         32          |       5.126390        |        1.076997       |       1.099699
|      2.107863       |      62.000000        |        0.927884       |      14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|         33          |       5.127231        |        1.064820       |       1.102011
|      3.492626       |      61.299998        |       -1.129037       |      14.915150
|
```

34	5.073361	1.051656	1.089044	3.555117	62.140000	0.225805	14.915150
35	5.091969	1.046859	1.067175	1.940709	63.190001	1.689735	14.915150
36	5.078733	1.033740	1.061945	2.728457	63.459998	0.427278	14.915150
37	5.084619	1.026881	1.061484	3.369730	63.519996	0.094545	14.915150
38	5.107906	1.016998	1.053293	3.568783	63.290000	-0.362085	14.915150
39	5.128337	1.009513	1.045956	3.609933	63.589996	0.110201	14.915150
40	5.097540	0.998343	1.029674	3.138276	64.639997	1.651205	14.915150
41	5.108642	0.990487	1.027441	3.730927	64.609998	-0.046409	14.915150

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        42         |      5.131843      |      0.980696       |      1.030523
|     5.080703      |     64.289999      |     -0.541458       |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        43         |      5.013742      |      0.982274       |      1.052352
|     7.134252      |     63.379997      |     -1.949258       |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        44         |      5.081797      |      0.978605       |      1.017990
|     4.024603      |     65.050000      |      0.634286       |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch       |   Epoch Time (s)   |   Training Loss     |     Test Loss
|    Overfit (%)    |    Accuracy (%)    |  Δ Accuracy (%)     |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        45         |      5.021367      |      0.966912       |      1.030294
|     6.555111      |     64.709997      |     -0.522679       |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        46         |      5.044518      |      0.958193       |      1.009344
|     5.338353      |     65.270001      |      0.338203       |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        47         |      5.089929      |      0.951124       |      1.005510
|     5.718028      |     65.130001      |     -0.214493       |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        48         |      5.070228      |      0.944371       |      0.995486
|     5.412549      |     65.899998      |      0.965217       |     14.915150
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         49          |       5.104623      |       0.945438      |       1.011090
|      6.944112       |      64.840001      |      -1.608493      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         50          |       5.083852      |       0.937149      |       0.984229
|      5.023837       |      66.219997      |       0.485583      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         51          |       5.122586      |       0.925459      |       0.984319
|      6.360085       |      66.249996      |       0.045302      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         52          |       5.096201      |       0.922812      |       0.978621
|      6.047788       |      65.990001      |      -0.392446      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         53          |       5.131413      |       0.921411      |       0.997682
|      8.277654       |      65.520000      |      -1.101881      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         54          |       5.056648      |       0.908973      |       0.960075
|      5.621919       |      67.030001      |       1.177365      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         55          |       5.029018      |       0.903034      |       0.963079
|      6.649313       |      67.140001      |       0.164106      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         56          |       5.040206      |       0.902585      |       0.951875
|      5.460921       |      67.420000      |       0.417037      |      14.915150
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         57          |      5.028303       |      0.893709       |      0.948499
|      6.130551       |      67.170000      |      -0.370809      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         58          |      5.011108       |      0.891331       |      0.965768
|      8.351196       |      66.469997      |      -1.409081      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         59          |      5.085455       |      0.889393       |      0.949489
|      6.756923       |      67.379999      |      -0.059331      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch         |   Epoch Time (s)    |   Training Loss     |     Test Loss
|    Overfit (%)      |    Accuracy (%)     |  Δ Accuracy (%)     |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         60          |      5.060808       |      0.885069       |      0.935363
|      5.682459       |      67.750001      |      0.489471       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         61          |      5.040805       |      0.877112       |      0.930217
|      6.054521       |      68.099999      |      0.516603       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         62          |      5.088319       |      0.869066       |      0.938141
|      7.948235       |      67.890000      |      -0.308369      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         63          |      5.085271       |      0.873677       |      0.937699
|      7.327950       |      67.809999      |      -0.425845      |      14.915150
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         64          |       5.123614      |       0.868866      |       0.923938
|      6.338306       |      68.500000      |       0.587373      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         65          |       5.134030      |       0.859185      |       0.926281
|      7.809257       |      68.110001      |      -0.569342      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         66          |       5.078809      |       0.857151      |       0.925934
|      8.024608       |      68.129998      |      -0.540149      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         67          |       5.098463      |       0.854254      |       0.924849
|      8.264010       |      68.229997      |      -0.394165      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         68          |       5.111977      |       0.859365      |       0.935850
|      8.900144       |      68.629998      |       0.189778      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         69          |       5.039732      |       0.840639      |       0.929612
|     10.584030       |      68.009996      |      -0.903397      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         70          |       5.085504      |       0.838153      |       0.907637
|      8.290179       |      69.000000      |       0.539125      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         71          |       5.067537      |       0.843908      |       0.929506
|     10.143097       |      67.820001      |      -1.710144      |      14.915150
|
```

```
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        72          |      5.057573       |       0.835999       |       0.915937
|     9.561942       |     68.529999       |      -0.681160       |     14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        73          |      5.074442       |       0.837075       |       0.904336
|     8.035266       |     69.199997       |       0.289851       |     14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        74          |      5.095749       |       0.832735       |       0.917949
|    10.233088       |     68.949997       |      -0.361271       |     14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)    |    Training Loss     |      Test Loss
|    Overfit (%)     |    Accuracy (%)     |   Δ Accuracy (%)     |  GPU Memory
(GiB)  |
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        75          |      5.059535       |       0.821980       |       0.893973
|     8.758437       |     69.690001       |       0.708098       |     14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        76          |      5.055362       |       0.817010       |       0.903975
|    10.644279       |     69.389999       |      -0.430481       |     14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        77          |      5.104808       |       0.819375       |       0.889156
|     8.516319       |     69.499999       |      -0.272638       |     14.915150
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        78          |      5.068392       |       0.813133       |       0.900651
|    10.763136       |     69.229996       |      -0.660073       |     14.915150
|
```

79	5.061541	0.808862	0.908639	12.335466	68.930000	-1.090545	14.915150
80	5.071167	0.806705	0.894210	10.847201	69.510001	-0.258287	14.915150
81	5.105503	0.804085	0.888455	10.492674	69.369996	-0.459184	14.915150
82	5.102875	0.801373	0.897731	12.024125	69.349998	-0.487879	14.915150
83	5.091874	0.801492	0.882098	10.056890	70.269996	0.832250	14.915150
84	5.102546	0.796571	0.907992	13.987577	68.909997	-1.935391	14.915150
85	5.091497	0.793067	0.872336	9.995210	70.429999	0.227697	14.915150
86	5.108479	0.786681	0.880855	11.971060	69.870001	-0.795112	14.915150

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         87          |       5.188274      |       0.782331      |       0.870107
|      11.219831      |      70.489997      |       0.085188      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         88          |       5.092411      |       0.778280      |       0.871861
|      12.024100      |      70.300001      |      -0.269536      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         89          |       5.157857      |       0.779800      |       0.886968
|      13.743012      |      70.120001      |      -0.524891      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        Epoch        |   Epoch Time (s)    |   Training Loss     |     Test Loss
|    Overfit (%)      |    Accuracy (%)     |  Δ Accuracy (%)     |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         90          |       5.120567      |       0.777302      |       0.886984
|      14.110626      |      69.830000      |      -0.936298      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         91          |       5.096100      |       0.774991      |       0.875539
|      12.974070      |      70.400000      |      -0.127674      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         92          |       5.158041      |       0.780365      |       0.866458
|      11.032430      |      70.469999      |      -0.028369      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         93          |       5.121035      |       0.781344      |       0.915482
|      17.167654      |      68.439996      |      -2.908215      |      14.915150
|
```

94	5.098719	0.782252	0.879924	12.485970	70.099998	-0.553269	14.915150
95	5.121605	0.769990	0.859838	11.668668	70.739996	0.354660	14.915150
96	5.093006	0.763771	0.871130	14.056423	70.559996	-0.254453	14.915150
97	5.114752	0.765917	0.874804	14.216483	69.999999	-1.046081	14.915150
98	5.081268	0.765277	0.870959	13.809715	70.349997	-0.551314	14.915150
99	5.104732	0.758171	0.858317	13.208964	70.539999	-0.282722	14.915150
100	5.126863	0.755749	0.852160	12.756919	71.359998	0.876451	14.915150
101	5.160976	0.756665	0.870397	15.030764	70.319998	-1.457399	14.915150

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        102         |      5.071728      |      0.754289      |      0.870921
|     15.462502      |     70.480001      |     -1.233180      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        103         |      5.068205      |      0.752413      |      0.856732
|     13.864575      |     70.940000      |     -0.588562      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        104         |      5.067548      |      0.750012      |      0.849197
|     13.224454      |     71.039999      |     -0.448430      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        105         |      5.069998      |      0.740821      |      0.860605
|     16.169134      |     70.489997      |     -1.219172      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        106         |      5.074821      |      0.751605      |      0.856403
|     13.943147      |     71.129996      |     -0.322313      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        107         |      5.041211      |      0.748542      |      0.849530
|     13.491354      |     71.069998      |     -0.406391      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        108         |      5.067728      |      0.735770      |      0.846414
|     15.037781      |     71.219999      |     -0.196187      |     14.915150
|
```

109	5.120624	0.735236	0.834490	13.499506	71.359998	0.000000	14.915150
110	5.120245	0.736660	0.850646	15.473345	71.099997	-0.364352	14.915150
111	5.119845	0.743295	0.881617	18.609299	69.569999	-2.508407	14.915150
112	5.103891	0.732732	0.839506	14.571980	71.499997	0.196187	14.915150
113	5.176232	0.722651	0.838164	15.984736	71.480000	-0.027968	14.915150
114	5.130633	0.720923	0.845920	17.338436	71.389997	-0.153847	14.915150
115	5.117654	0.726849	0.847043	16.536243	70.980000	-0.727268	14.915150
116	5.059834	0.722787	0.836406	15.719629	71.700001	0.279725	14.915150

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       117          |       5.077767       |       0.716625       |       0.839947
|    17.208736       |      71.789998       |       0.125519       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       118          |       5.042213       |       0.711941       |       0.829023
|    16.445593       |      72.219998       |       0.598969       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       119          |       5.109068       |       0.710683       |       0.835860
|    17.613676       |      71.770000       |      -0.623092       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      Epoch         |   Epoch Time (s)     |   Training Loss      |     Test Loss
|   Overfit (%)      |    Accuracy (%)      |  Δ Accuracy (%)      |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       120          |       5.136141       |       0.708005       |       0.823399
|    16.298500       |      72.139996       |      -0.110775       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       121          |       5.081880       |       0.707609       |       0.822545
|    16.242987       |      72.119999       |      -0.138464       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       122          |       5.100162       |       0.703986       |       0.853264
|    21.204630       |      70.879996       |      -1.855444       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       123          |       5.043020       |       0.721448       |       0.820083
|    13.671699       |      72.099996       |      -0.166162       |      14.915150
|
```

124	5.151543	0.705544	0.819741	16.185585	72.329998	0.152313	14.915150
125	5.112077	0.697202	0.821662	17.851221	72.279996	-0.069131	14.915150
126	5.074937	0.706385	0.828553	17.294765	71.759999	-0.788054	14.915150
127	5.112021	0.712020	0.842574	18.335752	71.829998	-0.691275	14.915150
128	5.068683	0.701907	0.842251	19.994738	71.429998	-1.244298	14.915150
129	5.106611	0.703883	0.825850	17.327799	71.730000	-0.829529	14.915150
130	5.161842	0.693148	0.821358	18.496904	72.149998	-0.248859	14.915150
131	5.114609	0.692579	0.835119	20.580921	71.469998	-1.188994	14.915150

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        132        |      5.234236      |      0.694519      |      0.826709
|     19.033396     |     72.380000      |      0.069131      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        133        |      5.167754      |      0.683442      |      0.818009
|     19.689714     |     72.240001      |     -0.193423      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        134        |      5.169732      |      0.686514      |      0.827337
|     20.512833     |     72.119999      |     -0.359217      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch       |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)    |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        135        |      5.170857      |      0.685092      |      0.819798
|     19.662561     |     72.240001      |     -0.193423      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        136        |      5.165364      |      0.681307      |      0.813139
|     19.349766     |     72.450000      |      0.096711      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        137        |      5.184544      |      0.676408      |      0.827789
|     22.380207     |     71.980000      |     -0.648724      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        138        |      5.041432      |      0.682127      |      0.816983
|     19.769909     |     72.209996      |     -0.331269      |     14.915150
|
```

139	5.188199	0.672820	0.811393	20.595865	72.950000	0.690130	14.915150	
140	5.134880	0.672488	0.810476	20.519125	72.719997	-0.315288	14.915150	
141	5.105504	0.675568	0.805462	19.227417	72.899997	-0.068543	14.915150	
142	5.120533	0.671230	0.808244	20.412378	72.969997	0.027412	14.915150	
143	5.068957	0.665354	0.813597	22.280344	72.499996	-0.644101	14.915150	
144	5.135123	0.667591	0.809880	21.313772	72.920001	-0.068516	14.915150	
145	5.082710	0.663192	0.802908	21.067223	72.700000	-0.370011	14.915150	
146	5.110504	0.666857	0.804756	20.678925	72.799999	-0.232970	14.915150	

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       147        |      5.176138      |      0.662117      |      0.819327
|    23.743467     |     72.689998      |     -0.383718      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       148        |      5.138168      |      0.666574      |      0.826696
|    24.021763     |     71.569997      |     -1.918597      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       149        |      5.160818      |      0.664240      |      0.801299
|    20.634020     |     73.249996      |      0.383718      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      Epoch       |   Epoch Time (s)   |   Training Loss    |     Test Loss
|   Overfit (%)    |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       150        |      5.115169      |      0.659333      |      0.801870
|    21.618420     |     73.039997      |     -0.286688      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       151        |      5.125012      |      0.653944      |      0.823548
|    25.935585     |     72.270000      |     -1.337878      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       152        |      5.116081      |      0.655704      |      0.795027
|    21.247782     |     73.100001      |     -0.204771      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       153        |      5.120463      |      0.658052      |      0.818986
|    24.456205     |     72.240001      |     -1.378832      |     14.915150
|
```

154	5.186667	0.649822	0.794072	22.198418	73.359996	0.150171	14.915150
155	5.152768	0.646464	0.794968	22.971737	73.039997	-0.436204	14.915150
156	5.118167	0.651932	0.799742	22.672680	73.259997	-0.136312	14.915150
157	5.073302	0.654624	0.801499	22.436533	72.939998	-0.572517	14.915150
158	5.136588	0.644021	0.797534	23.836645	73.219997	-0.190839	14.915150
159	5.128449	0.647507	0.798428	23.307957	73.100001	-0.354410	14.915150
160	5.112539	0.641143	0.792088	23.543064	73.240000	-0.163572	14.915150
161	5.120852	0.643312	0.809783	25.877104	72.719997	-0.872409	14.915150

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       162        |      5.094538      |      0.641172      |      0.794955
|     23.984808    |     73.299998      |     -0.081786      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       163        |      5.057532      |      0.635318      |      0.784141
|     23.424995    |     73.679996      |      0.436204      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       164        |      5.093210      |      0.631926      |      0.785027
|     24.227589    |     73.759997      |      0.108579      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      Epoch       |   Epoch Time (s)   |   Training Loss    |     Test Loss
|   Overfit (%)    |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       165        |      5.111727      |      0.631431      |      0.796851
|     26.197598    |     73.299998      |     -0.623643      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       166        |      5.123473      |      0.634222      |      0.789436
|     24.473181    |     73.460001      |     -0.406719      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       167        |      5.107878      |      0.634861      |      0.785529
|     23.732395    |     73.579997      |     -0.244035      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       168        |      5.150738      |      0.629582      |      0.806329
|     28.073561    |     72.909999      |     -1.152383      |     14.915150
|
```

169	5.091628	0.628832	0.790047	25.637213	73.670000	-0.122013	14.915150
170	5.073217	0.635245	0.781749	23.062716	73.909998	0.203364	14.915150
171	5.121365	0.632413	0.813234	28.592190	72.889996	-1.380060	14.915150
172	5.029716	0.631198	0.793187	25.663792	73.299998	-0.825328	14.915150
173	5.140740	0.633403	0.781718	23.415546	73.589998	-0.432958	14.915150
174	5.088692	0.627700	0.786654	25.323292	73.569995	-0.460023	14.915150
175	5.156265	0.625505	0.783891	25.321344	73.490000	-0.568256	14.915150
176	5.202651	0.622379	0.784202	26.000751	73.490000	-0.568256	14.915150

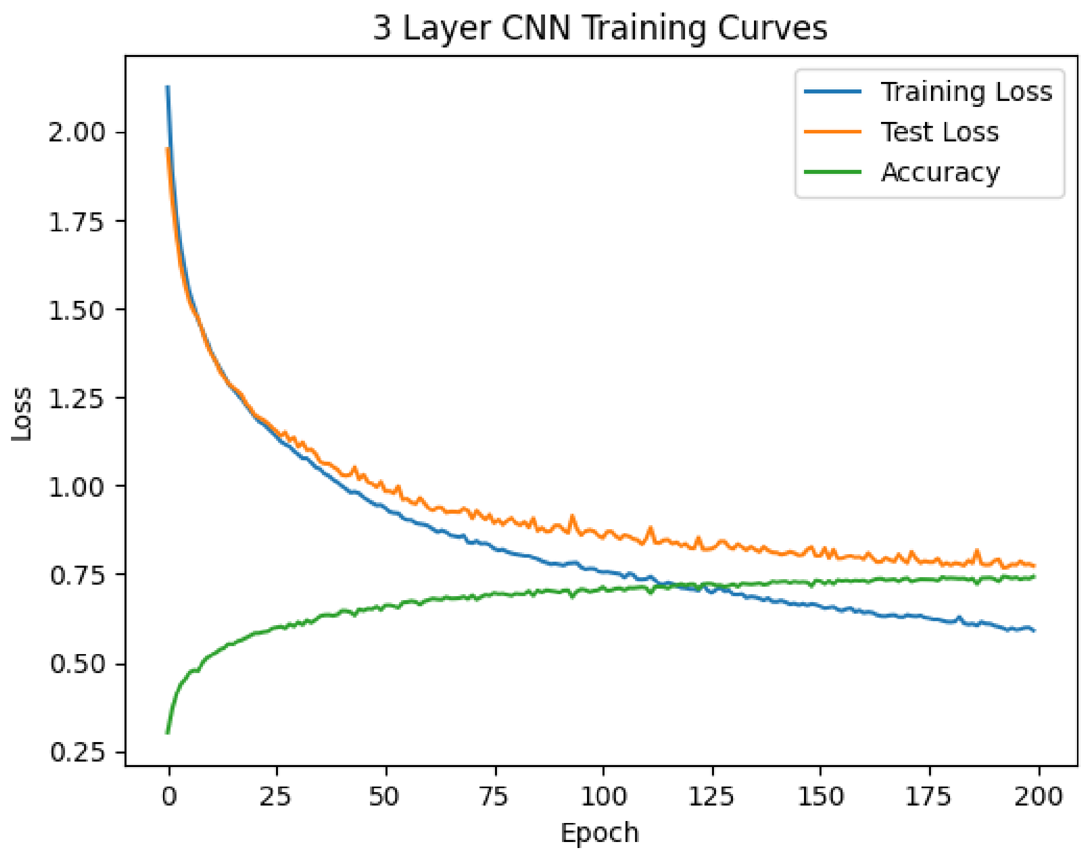
```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        177         |      5.100004      |       0.622683      |      0.793957
|     27.505703      |     73.149997      |      -1.028279      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        178         |      5.156939      |       0.618830      |      0.775399
|     25.300802      |     74.180001      |       0.365313      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        179         |      5.132620      |       0.615692      |      0.781508
|     26.931622      |     73.839998      |      -0.458348      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss     |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |  Δ Accuracy (%)     |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        180         |      5.125704      |       0.614744      |      0.775486
|     26.147738      |     73.939997      |      -0.323543      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        181         |      5.205902      |       0.616690      |      0.781392
|     26.707546      |     73.859996      |      -0.431390      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        182         |      5.122013      |       0.629987      |      0.779209
|     23.686480      |     73.759997      |      -0.566196      |     14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        183         |      5.205942      |       0.611550      |      0.774581
|     26.658596      |     73.829997      |      -0.471831      |     14.915150
|
```

184	5.180202	0.608123	0.788122	29.599031	73.390001	-1.064977	14.915150
185	5.212535	0.609963	0.781101	28.057093	73.789996	-0.525755	14.915150
186	5.164953	0.605180	0.816878	34.980904	72.209996	-2.655709	14.915150
187	5.107094	0.615284	0.779685	26.719515	73.890001	-0.390942	14.915150
188	5.131805	0.610408	0.776326	27.181515	73.939997	-0.323543	14.915150
189	5.131907	0.609715	0.777072	27.448435	74.009997	-0.229178	14.915150
190	5.079273	0.605033	0.787273	30.120705	73.699999	-0.647078	14.915150
191	5.156400	0.601853	0.791279	31.473764	72.920001	-1.698572	14.915150

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        192           |       5.183256       |       0.597795       |       0.768771
|     28.601042        |      74.379998       |       0.269611       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        193           |       5.132535       |       0.591397       |       0.770208
|     30.235384        |      74.210000       |      -0.228554       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        194           |       5.130533       |       0.598124       |       0.777528
|     29.994465        |      73.850000       |      -0.712555       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch          |   Epoch Time (s)     |   Training Loss      |     Test Loss
|    Overfit (%)       |    Accuracy (%)      |  Δ Accuracy (%)      |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        195           |       5.087523       |       0.592661       |       0.776363
|     30.996077        |      74.100000       |      -0.376444       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        196           |       5.129474       |       0.594608       |       0.786006
|     32.188976        |      73.569995       |      -1.089007       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        197           |       5.143143       |       0.599006       |       0.777057
|     29.724390        |      73.920000       |      -0.618444       |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        198           |       5.106315       |       0.599403       |       0.778661
|     29.906042        |      73.749995       |      -0.847006       |      14.915150
|
```

```
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         199          |       5.109549       |       0.591137       |       0.773772
|      30.895559       |       74.309999      |       -0.094111      |      14.915150
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------

Training Time: 472.81531620025635 seconds
```



[4]:
```
#after last bn but before last weight

class ResBlock(nn.Module):
    def __init__(self, in_chans, out_chans, nonlinearity = 'relu', stride=1,␣
 ↪dropout = 0.4):
        super().__init__()
```

```
        self.conv1 = nn.Conv2d(in_chans, out_chans, kernel_size=3, padding=1,␣
 ↪bias=False, stride=stride)
        self.batch_norm1 = nn.BatchNorm2d(num_features=out_chans)
        self.conv2 = nn.Conv2d(out_chans, out_chans, kernel_size=3, padding=1,␣
 ↪bias=False, stride=stride)
        self.batch_norm2 = nn.BatchNorm2d(num_features=out_chans)
        self.dropout = nn.Dropout(dropout)
        self.shortcut = nn.Conv2d(in_chans, out_chans, kernel_size=1, stride=1,␣
 ↪bias=False) if in_chans != out_chans else nn.Identity()


        torch.nn.init.kaiming_normal_(self.conv1.weight,␣
 ↪nonlinearity=nonlinearity)
        torch.nn.init.constant_(self.batch_norm1.weight, 0.5)
        torch.nn.init.zeros_(self.batch_norm1.bias)
        torch.nn.init.kaiming_normal_(self.conv2.weight,␣
 ↪nonlinearity=nonlinearity)
        torch.nn.init.zeros_(self.batch_norm2.bias)
    def forward(self, x):
        out = self.batch_norm1(x)
        out = F.relu(out)
        out = self.conv1(out)
        out = self.batch_norm2(out)
        out = F.relu(out)
        out = self.dropout(out)
        out = self.conv2(out)
        out += self.shortcut(x)
        return out
class ResNet(Classifier):
    def __init__(self, input_dim = 32, n_blocks = 10, conv_channels = [32,16],␣
 ↪fc_channels = [32, 10], dropout_p=0.4, dropout_h=0.4, nonlinearity='relu'):
        super().__init__()

        # Add initial convolutions
        self.h1 = nn.Sequential()
        for i in range(len(conv_channels)):
            self.h1.add_module(
                name=f'conv{i}',
                module=nn.Conv2d(
                    in_channels = 3 if i == 0 else conv_channels[i-1],
                    out_channels=conv_channels[i],
                    kernel_size=3,
                    padding=1
                )
            )
            self.h1.add_module(
```

```
                name=f'activation{i}',
                module=nn.ReLU()
            )
        self.h1.add_module(
            name=f'maxpool',
            module=nn.MaxPool2d(2)
        )
        #output of h1 before maxpool is 16 32x32 images. After maxpool, 16␣
↪16x16 images
        # h1_in: torch.Size([1024, 3, 32, 32])
        # res_block_in: torch.Size([1024, 16, 16, 16])
        # h2_in: torch.Size([1024, 64])
        #Add Resblocks
        res_block_in = conv_channels[0]//2
        self.resblocks = nn.Sequential(
            *[
                ResBlock(
                    in_chans=conv_channels[-1],
                    out_chans=conv_channels[-1],
                    nonlinearity=nonlinearity,
                    dropout=dropout_p
                ) for _ in range(n_blocks)
            ]
        )


        #output of resblocks is 16 16x16 images

        # Add final layers
        self.h2 = nn.Sequential()
        self.h2.add_module(
            name='final_batch_norm',
            module=nn.BatchNorm2d(
                num_features=conv_channels[-1]
            )
        )
        self.h2.add_module(
            name='final_relu',
            module=nn.ReLU()
        )
        self.h2.add_module(
            name = 'dropout_head',
            module=nn.Dropout(dropout_h)
        )
        self.h2.add_module(
            name = 'gap',
            module=nn.AvgPool2d(2)
```

```
        )
        self.h2.add_module(
            name = 'flatten',
            module=nn.Flatten()
        )

        #output is 16 8x8 images
        #   16 comes from conv_channels[-1]
        #   8x8 comes from input_dim // 4
        fc_in = conv_channels[-1] * (input_dim//4)**2

        for i in range(len(fc_channels)):

            self.h2.add_module(
                name=f'fc{i}',
                module=nn.Linear(
                    in_features=fc_in if i == 0 else fc_channels[i-1],
                    out_features=fc_channels[i]
                )
            )
            if i < len(fc_channels) - 1:
                self.h2.add_module(
                    name = f'fc_activation{i}',
                    module=nn.ReLU()
                )
        self.h2.add_module('softmax', nn.Softmax(dim=1))
    def forward(self, x):
        #print(f"h1_in: {x.shape}")
        out = self.h1(x)
        #print(f"res_block_in: {out.shape}")
        out = self.resblocks(out)
        #print(f"h2_in: {out.shape}")
        out = self.h2(out)
        return out
```

[5]:
```
from torch.profiler import profile, record_function, ProfilerActivity
try:
    del train_loader
    del test_loader
    del model_1a
    del model_1b
    del resnet
    del train_loader_cuda
    del test_loader_cuda
except:
    pass
```

```python
# Reset CUDA context
start = time.time()
torch.cuda.empty_cache()
torch.cuda.reset_peak_memory_stats()

gc.collect()


cifar10_train = datasets.CIFAR10(data_path, train=True, download=dl, transform=transform)
cifar10_test = datasets.CIFAR10(data_path, train=False, download=dl, transform=transform)

batch_size = int(2**11)
workers = 12
cpu_prefetch = 39
gpu_prefetch = 28

print('begin init train_loader')
train_loader = DataLoader(
    cifar10_train,
    batch_size=batch_size,
    shuffle=True,
    num_workers=workers,
    prefetch_factor=cpu_prefetch,
    pin_memory=True
)

X_batch = next(iter(train_loader))[0]
dtype_size = X_batch.element_size()
print(f"Batch Size: {X_batch.element_size() * X_batch.nelement() / 1024**2} MiB")


print('begin init fetcher')
train_loader_cuda = CudaDataPrefetcher(
    data_iterable = train_loader,
    device = torch.device('cuda'),
    num_prefetch_batches=gpu_prefetch
)
test_loader = DataLoader(cifar10_test, batch_size=len(cifar10_test), shuffle=True, num_workers=workers, pin_memory=True, prefetch_factor=1)
test_loader_cuda = CudaDataPrefetcher(
    data_iterable = test_loader,
    device = torch.device('cuda'),
    num_prefetch_batches=1
)
```

```python
resnet = ResNet(
    input_dim = 32,
    conv_channels=[16,16],
    n_blocks = 10,
    fc_channels=[16,10],
    dropout_h = 0.6,
    dropout_p = 0.4
).to(device=device)
print(f"Init time: {(time.time() - start):.2f} seconds")
params = sum(p.numel() for p in resnet.parameters() if p.requires_grad)
print(f"Total parameters: {params}")
resnet.train_model(
    epochs=200,
    train_loader=train_loader_cuda,
    train_len=len(train_loader),
    test_loader=test_loader_cuda,
    test_len=len(test_loader),
    test_size=len(cifar10_test),
    loss_fn=nn.CrossEntropyLoss(),
    optimizer = torch.optim.Adam,
    optimizer_kwargs={'lr': 2e-3, 'weight_decay': 7e-3},
    print_epoch = 1,
    sched_factor = 0.6,
    sched_patience = 15
)


resnet.plot_training("ResNet Training Curves")
```

```
begin init train_loader
Batch Size: 24.0 MiB
begin init fetcher
Init time: 4.00 seconds
Total parameters: 66090
Training ResNet

--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        Epoch         |    Epoch Time (s)    |    Training Loss     |      Test Loss
|     Overfit (%)      |     Accuracy (%)     |    Δ Accuracy (%)    |   GPU Memory
(GiB)   |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|          0           |       7.758229       |       2.262175       |       2.303453
|       1.824688       |       9.999999       |       0.000000       |       4.617005
```

```
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        1         |       5.941853     |       2.131489     |       2.312773
|     8.505051     |      10.349999     |       3.500000     |       5.320131
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        2         |       5.965214     |       2.039086     |       2.239328
|     9.820173     |      18.820000     |      81.835754     |       6.021304
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        3         |       5.957030     |       1.999411     |       2.137333
|     6.898159     |      32.960001     |      75.132843     |       6.722476
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        4         |       5.937284     |       1.963828     |       2.014531
|     2.581829     |      47.529998     |      44.205090     |       7.423647
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        5         |       5.978225     |       1.935574     |       2.026971
|     4.721979     |      46.619999     |      -1.914578     |       8.124819
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        6         |       5.998431     |       1.916032     |       2.005815
|     4.685873     |      46.309999     |      -2.566800     |       8.825991
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        7         |       5.885165     |       1.904790     |       1.935971
|     1.637028     |      55.320001     |      16.389654     |       9.527163
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        8         |       5.937993     |       1.893218     |       1.929178
|     1.899396     |      55.729997     |       0.741136     |      10.228335
```

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        9        |     5.912222     |     1.880680     |     1.889912
|     0.490889    |    59.649998     |     7.033916     |    10.929507
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        10       |     5.891649     |     1.870847     |     1.926282
|     2.963132    |    54.809999     |    -8.113996     |    11.630679
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        11       |     5.917224     |     1.862037     |     1.985746
|     6.643737    |    47.759998     |    -19.932943    |    12.331851
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        12       |     5.954925     |     1.852967     |     1.968036
|     6.210001    |    50.389999     |    -15.523887    |    13.033022
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        13       |     5.949610     |     1.850348     |     1.894329
|     2.376869    |    58.849996     |    -1.341159     |    13.734194
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        14       |     5.927939     |     1.838039     |     1.884981
|     2.553910    |    58.950001     |    -1.173507     |    14.435366
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      Epoch      |  Epoch Time (s)  |  Training Loss   |    Test Loss
|   Overfit (%)   |   Accuracy (%)   | Δ Accuracy (%)   |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        15       |     5.978298     |     1.835852     |     1.916794
|     4.408997    |    55.689996     |    -6.638729     |    15.136538
```

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        16         |      5.931443      |       1.831612      |       1.992336
|     8.774986      |     46.829998      |     -21.492037      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        17         |      5.936951      |       1.830642      |       1.878986
|     2.640804      |     60.289997      |       1.072924      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        18         |      6.016370      |       1.825897      |       1.834697
|     0.481930      |     63.940001      |       6.054079      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        19         |      5.955230      |       1.824874      |       1.879447
|     2.990498      |     58.520001      |      -8.476695      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        20         |      5.945711      |       1.819507      |       1.890054
|     3.877305      |     58.429998      |      -8.617457      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        21         |      5.960409      |       1.816705      |       1.909476
|     5.106515      |     56.019998      |     -12.386617      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        22         |      5.979871      |       1.816134      |       1.913466
|     5.359302      |     54.909998      |     -14.122619      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        23         |      5.943759      |       1.814242      |       1.888312
|     4.082688      |     58.359998      |      -8.726934      |     15.136538
```

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         24          |       5.989490        |        1.812754       |        1.932691
|      6.616277       |      53.270000        |      -16.687519       |       15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         25          |       5.976249        |        1.811979       |        1.853341
|      2.282714       |      61.919999        |       -3.159214       |       15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         26          |       5.926868        |        1.805859       |        1.890505
|      4.687298       |      57.969999        |       -9.336880       |       15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         27          |       5.957161        |        1.808381       |        1.828712
|      1.124280       |      64.550000        |        0.954019       |       15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         28          |       5.978532        |        1.808984       |        1.878929
|      3.866507       |      58.649999        |       -9.140203       |       15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         29          |       5.943629        |        1.806721       |        1.885318
|      4.350254       |      58.569998        |       -9.264141       |       15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        Epoch        |    Epoch Time (s)     |     Training Loss     |       Test Loss
|     Overfit (%)     |     Accuracy (%)      |   Δ Accuracy (%)      |   GPU Memory
(GiB)   |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         30          |       5.946413        |        1.802864       |        1.968687
|      9.197787       |      48.940000        |      -24.182804       |       15.136538
```

31	5.973595	1.804069	1.820951	0.935777	65.279996	1.130900	15.136538
32	5.941212	1.800107	1.835910	1.988929	63.870001	-2.159920	15.136538
33	5.957007	1.798617	1.939540	7.835098	52.329999	-19.837620	15.136538
34	5.956561	1.798999	1.824884	1.438882	64.569998	-1.087620	15.136538
35	6.000816	1.799080	1.926372	7.075355	53.549999	-17.968748	15.136538
36	5.929878	1.797763	1.850310	2.922882	61.909997	-5.162376	15.136538
37	5.952627	1.798736	1.868542	3.880836	60.339999	-7.567398	15.136538
38	5.983609	1.794371	1.869053	4.161981	59.729999	-8.501835	15.136538

```
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|       39        |      5.951901      |      1.791374      |      1.866068
|     4.169658    |     60.249996      |     -7.705270      |     15.136538
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|       40        |      6.068545      |      1.793264      |      1.898383
|     5.861915    |     56.849998      |     -12.913601     |     15.136538
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|       41        |      6.035099      |      1.790199      |      1.822050
|     1.779197    |     64.849997      |     -0.658701      |     15.136538
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|       42        |      5.953965      |      1.793583      |      1.866762
|     4.080087    |     60.039997      |     -8.026960      |     15.136538
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|       43        |      6.000158      |      1.793789      |      1.888961
|     5.305665    |     57.660002      |     -11.672788     |     15.136538
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|       44        |      6.012037      |      1.788270      |      1.844360
|     3.136523    |     62.459999      |     -4.319849      |     15.136538
|
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|      Epoch      |   Epoch Time (s)   |   Training Loss    |     Test Loss
|   Overfit (%)   |    Accuracy (%)    |  Δ Accuracy (%)    |  GPU Memory
(GiB)  |
-------------------------------------------------------------------------------
-------------------------------------------------------------------------------
---------
|       45        |      6.064523      |      1.786979      |      1.848547
|     3.445322    |     61.769998      |     -5.376836      |     15.136538
```

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         46         |      6.086365      |      1.789465      |      1.920969
|      7.348784      |     54.629999      |     -16.314335     |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         47         |      5.946903      |      1.786739      |      1.878974
|      5.162196      |     58.359998      |     -10.600488     |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         48         |      5.962193      |      1.780643      |      1.818045
|      2.100492      |     65.450001      |      0.260423      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         49         |      5.993754      |      1.766896      |      1.795054
|      1.593680      |     67.890000      |      3.728036      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         50         |      5.979029      |      1.765290      |      1.812099
|      2.651630      |     65.309995      |     -3.800272      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         51         |      5.931840      |      1.769414      |      1.837127
|      3.826841      |     63.080001      |     -7.084990      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         52         |      5.952838      |      1.767406      |      1.822430
|      3.113217      |     64.399999      |     -5.140670      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         53         |      5.972234      |      1.765201      |      1.797711
|      1.841746      |     66.929996      |     -1.414059      |     15.136538
```

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        54        |      5.964084       |      1.764875       |      1.791243
|      1.494045      |      67.729998      |      -0.235679      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        55        |      5.955197       |      1.766104       |      1.816597
|      2.859046      |      65.169996      |      -4.006487      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        56        |      6.029804       |      1.763746       |      1.872248
|      6.151809      |      58.849996      |      -13.315663     |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        57        |      5.953676       |      1.762739       |      1.792457
|      1.685915      |      67.839998      |      -0.073652      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        58        |      5.990416       |      1.761238       |      1.820805
|      3.382094      |      64.569998      |      -4.890267      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        59        |      5.994577       |      1.761105       |      1.813354
|      2.966815      |      65.230000      |      -3.918103      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch       |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |  Δ Accuracy (%)   |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        60        |      6.018641       |      1.758287       |      1.878296
|      6.825368      |      59.749997      |      -11.989988     |      15.136538
```

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         61         |       6.037763      |       1.745997      |       1.808743
|      3.593695      |      65.740001      |      -3.166886      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         62         |       5.997109      |       1.739772      |       1.756561
|      0.964996      |      71.319997      |       5.052286      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         63         |       5.976972      |       1.739508      |       1.791143
|      2.968363      |      68.320000      |      -4.206390      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         64         |       6.112227      |       1.732286      |       1.857396
|      7.222256      |      61.369997      |      -13.951206     |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         65         |       5.984481      |       1.726140      |       1.763518
|      2.165442      |      70.529997      |      -1.107684      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         66         |       6.100321      |       1.728138      |       1.782535
|      3.147723      |      69.150001      |      -3.042620      |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         67         |       5.998295      |       1.725812      |       1.827685
|      5.902898      |      63.919997      |      -10.375772     |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|         68         |       6.043867      |       1.727304      |       1.807588
|      4.647894      |      65.810001      |      -7.725739      |      15.136538
```

```
|
------------------------------------------------------------------------------
------------------------------------------------------------------------------
---------
|        69          |      6.017951       |      1.723430       |      1.808294
|      4.924157      |     66.399997       |     -6.898486       |     15.136538
|
------------------------------------------------------------------------------
------------------------------------------------------------------------------
---------
|        70          |      6.035763       |      1.725271       |      1.767110
|      2.425069      |     70.190001       |     -1.584404       |     15.136538
|
------------------------------------------------------------------------------
------------------------------------------------------------------------------
---------
|        71          |      5.997301       |      1.720044       |      1.785120
|      3.783398      |     68.059999       |     -4.570946       |     15.136538
|
------------------------------------------------------------------------------
------------------------------------------------------------------------------
---------
|        72          |      6.033640       |      1.718865       |      1.798987
|      4.661331      |     66.780001       |     -6.365671       |     15.136538
|
------------------------------------------------------------------------------
------------------------------------------------------------------------------
---------
|        73          |      6.025136       |      1.722847       |      1.744101
|      1.233708      |     72.560000       |      1.738647       |     15.136538
|
------------------------------------------------------------------------------
------------------------------------------------------------------------------
---------
|        74          |      6.044716       |      1.716789       |      1.828936
|      6.532368      |     63.679999       |     -12.238149      |     15.136538
|
------------------------------------------------------------------------------
------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)    |   Training Loss     |     Test Loss
|    Overfit (%)     |    Accuracy (%)     |  Δ Accuracy (%)     |  GPU Memory
(GiB)  |
------------------------------------------------------------------------------
------------------------------------------------------------------------------
---------
|        75          |      6.046995       |      1.724921       |      1.802198
|      4.480049      |     66.319996       |     -8.599786       |     15.136538
```

76	6.073751	1.722809	1.788122	3.791047	67.979997	-6.312021	15.136538
77	5.978110	1.715931	1.788931	4.254255	68.110001	-6.132854	15.136538
78	6.042846	1.714050	1.765153	2.981410	70.389998	-2.990632	15.136538
79	5.989166	1.709710	1.810159	5.875241	65.799999	-9.316429	15.136538
80	6.026500	1.713593	1.827208	6.630245	63.989997	-11.810920	15.136538
81	6.064908	1.713212	1.820815	6.280786	64.289999	-11.397466	15.136538
82	6.053454	1.713486	1.755603	2.457978	71.099997	-2.012133	15.136538
83	6.030574	1.709514	1.789637	4.686906	67.890000	-6.436053	15.136538

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        84        |      6.010733       |      1.713726       |      1.790577
|     4.484475     |     67.530000       |     -6.932194       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        85        |      5.997445       |      1.712908       |      1.741797
|     1.686579     |     72.700000       |      0.192943       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        86        |      6.078503       |      1.707408       |      1.794512
|     5.101524     |     67.109996       |     -7.689139       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        87        |      6.048886       |      1.713436       |      1.799901
|     5.046287     |     66.710001       |     -8.239338       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        88        |      6.026438       |      1.711424       |      1.845643
|     7.842492     |     61.939996       |     -14.800554      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        89        |      6.002117       |      1.707037       |      1.757974
|     2.983910     |     71.410000       |     -1.774415       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch      |   Epoch Time (s)    |   Training Loss     |     Test Loss
|    Overfit (%)   |    Accuracy (%)     |  Δ Accuracy (%)     |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        90        |      5.973038       |      1.710023       |      1.775080
|     3.804483     |     69.150001       |     -4.883080       |     15.136538
```

```
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       91        |      6.030484     |      1.706634     |      1.743197
|     2.142406    |     72.649997     |     -0.068779     |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       92        |      6.044240     |      1.706756     |      1.857970
|     8.859737    |     60.549998     |    -16.712519     |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       93        |      6.025276     |      1.709503     |      1.789289
|     4.667165    |     67.949998     |     -6.533702     |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       94        |      6.065005     |      1.709287     |      1.730637
|     1.249039    |     73.869997     |      1.609350     |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       95        |      6.063215     |      1.706813     |      1.837258
|     7.642605    |     62.809998     |    -14.972248     |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       96        |      6.046596     |      1.709644     |      1.749821
|     2.349999    |     72.299999     |     -2.125353     |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       97        |      6.041147     |      1.707956     |      1.765295
|     3.357140    |     70.229995     |     -4.927578     |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       98        |      6.140745     |      1.705159     |      1.762154
|     3.342478    |     70.159996     |     -5.022339     |     15.136538
```

```
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        99          |       5.942571       |       1.704508       |      1.816739
|      6.584328      |      64.919996       |      -12.115881      |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        100         |       5.976495       |       1.704234       |      1.753127
|      2.868941      |      71.859998       |      -2.720996       |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        101         |       6.089538       |       1.705383       |      1.850588
|      8.514546      |      61.739999       |      -16.420738      |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        102         |       6.001499       |       1.704032       |      1.744620
|      2.381902      |      72.499996       |      -1.854611       |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        103         |       6.000335       |       1.703446       |      1.796068
|      5.437298      |      67.570001       |      -8.528491       |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        104         |       5.989391       |       1.702863       |      1.846086
|      8.410723      |      61.839998       |      -16.285366      |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss    |      Test Loss
|    Overfit (%)     |    Accuracy (%)     |   Δ Accuracy (%)   |  GPU Memory
(GiB)  |
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        105         |       6.053767       |       1.704828       |      1.784842
|      4.693377      |      68.210000       |      -7.662106       |     15.136538
```

106	5.995612	1.704311	1.820918	6.841868	63.989997	-13.374848	15.136538
107	6.119347	1.703072	1.791096	5.168544	67.799997	-8.217138	15.136538
108	5.988535	1.702467	1.789317	5.101415	68.229997	-7.635035	15.136538
109	6.011374	1.704103	1.806670	6.018831	65.700001	-11.059965	15.136538
110	6.016177	1.705480	1.752431	2.752946	71.799999	-2.802217	15.136538
111	6.034240	1.686944	1.737605	3.003156	73.249996	-0.839314	15.136538
112	6.049294	1.681191	1.718573	2.223586	74.759996	1.204818	15.136538
113	6.035938	1.682705	1.719535	2.188763	75.009996	0.334403	15.136538

```
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        114          |       6.005780        |       1.682150        |       1.720372
|      2.272208       |      74.759996        |      -0.333289        |      15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        115          |       6.040393        |       1.683993        |       1.763142
|      4.700065       |      70.400000        |      -6.145842        |      15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        116          |       6.026464        |       1.679085        |       1.732503
|      3.181357       |      73.899996        |      -1.479802        |      15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        117          |       6.116028        |       1.679812        |       1.728820
|      2.917485       |      73.769999        |      -1.653109        |      15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        118          |       6.095558        |       1.680922        |       1.724223
|      2.576033       |      74.169999        |      -1.119846        |      15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        119          |       6.083001        |       1.678301        |       1.736588
|      3.472963       |      73.119998        |      -2.519661        |      15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       Epoch         |   Epoch Time (s)      |   Training Loss       |     Test Loss
|    Overfit (%)      |    Accuracy (%)       |  Δ Accuracy (%)       |  GPU Memory
(GiB)  |
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        120          |       6.071417        |       1.680773        |       1.722848
|      2.503350       |      74.419999        |      -0.786557        |      15.136538
```

121	6.101312	1.682564	1.713852	1.859566	75.769997	1.013200	15.136538
122	6.108806	1.682129	1.731231	2.919064	73.869997	-2.507588	15.136538
123	6.083366	1.681966	1.703076	1.255068	76.769996	1.319782	15.136538
124	6.085177	1.679816	1.698992	1.141518	76.580000	-0.247487	15.136538
125	6.054791	1.680250	1.748276	4.048547	71.819997	-6.447830	15.136538
126	6.054725	1.681701	1.729124	2.819957	74.149996	-3.412791	15.136538
127	6.105957	1.682942	1.785550	6.096926	67.839998	-11.632145	15.136538
128	6.027437	1.680754	1.787973	6.379204	67.739999	-11.762403	15.136538

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       129         |      6.026318       |      1.681985       |      1.742985
|     3.626660      |     72.439998       |     -5.640221       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       130         |      6.067482       |      1.681037       |      1.717516
|     2.170037      |     74.799997       |     -2.566106       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       131         |      6.053882       |      1.680817       |      1.752610
|     4.271327      |     71.630001       |     -6.695317       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       132         |      6.073949       |      1.681933       |      1.761947
|     4.757300      |     70.599997       |     -8.036992       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       133         |      6.100069       |      1.679813       |      1.729257
|     2.943421      |     73.749995       |     -3.933829       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       134         |      6.060847       |      1.681197       |      1.747969
|     3.971707      |     72.179997       |     -5.978897       |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|      Epoch        |   Epoch Time (s)    |   Training Loss     |     Test Loss
|    Overfit (%)    |    Accuracy (%)     |  Δ Accuracy (%)     |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       135         |      6.092427       |      1.676345       |      1.728331
|     3.101143      |     73.869997       |     -3.777515       |     15.136538
```

136	6.111445	1.677436	1.710195	1.952925	75.689995	-1.406800	15.136538
137	6.154238	1.680783	1.780700	5.944660	68.439996	-10.850594	15.136538
138	6.104064	1.680163	1.767590	5.203455	69.580001	-9.365631	15.136538
139	6.064308	1.679928	1.724739	2.667443	74.210000	-3.334630	15.136538
140	6.095605	1.670509	1.695053	1.469260	77.579999	1.055104	15.136538
141	6.075877	1.665254	1.685887	1.239022	78.099996	0.670272	15.136538
142	6.111793	1.662272	1.704352	2.531463	76.239997	-2.381561	15.136538
143	6.030466	1.663150	1.701710	2.318458	76.569998	-1.959024	15.136538

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        144         |       6.110031      |       1.664230      |      1.696805
|      1.957377      |      76.959997      |      -1.459665      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        145         |       6.130999      |       1.662143      |      1.703371
|      2.480386      |      76.400000      |      -2.176692      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        146         |       6.076478      |       1.659864      |      1.729352
|      4.186393      |      73.600000      |      -5.761838      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        147         |       6.191991      |       1.659926      |      1.678746
|      1.133810      |      79.029995      |       1.190780      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        148         |       6.065806      |       1.664576      |      1.687991
|      1.406648      |      78.039998      |      -1.252686      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        149         |       6.120574      |       1.660582      |      1.703497
|      2.584342      |      76.519996      |      -3.176008      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)    |    Training Loss    |      Test Loss
|    Overfit (%)     |    Accuracy (%)     |  Δ Accuracy (%)     |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        150         |       6.083641      |       1.660676      |      1.691585
|      1.861263      |      77.609998      |      -1.796783      |     15.136538
```

151	6.159505	1.661676	1.726976	3.929745	74.030000	-6.326706	15.136538
152	6.116043	1.662686	1.677465	0.888861	78.789997	-0.303680	15.136538
153	6.153685	1.661415	1.669934	0.512795	79.769999	0.936358	15.136538
154	6.083883	1.660522	1.722128	3.710041	74.159998	-7.032721	15.136538
155	6.110673	1.661754	1.710175	2.913857	75.610000	-5.214992	15.136538
156	6.137404	1.660498	1.689023	1.717822	77.899998	-2.344240	15.136538
157	6.167748	1.659847	1.681879	1.327353	78.389996	-1.729977	15.136538
158	6.080560	1.659312	1.732317	4.399757	73.509997	-7.847564	15.136538

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        159         |       6.110744       |       1.659585       |       1.717616
|      3.496686      |      74.640000       |      -6.430988       |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        160         |       6.144540       |       1.661628       |       1.708873
|      2.843304      |      75.439996       |      -5.428110       |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        161         |       6.116296       |       1.661444       |       1.721828
|      3.634423      |      74.469995       |      -6.644106       |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        162         |       6.133440       |       1.659592       |       1.694767
|      2.119488      |      76.919997       |      -3.572774       |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        163         |       6.133320       |       1.661856       |       1.685046
|      1.395393      |      78.200001       |      -1.968156       |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        164         |       6.137462       |       1.661753       |       1.698507
|      2.211732      |      76.719999       |      -3.823492       |      15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |  Δ Accuracy (%)   |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        165         |       6.100959       |       1.660505       |       1.710490
|      3.010189      |      75.489998       |      -5.365427       |      15.136538
```

166	6.162635	1.660260	1.731846	4.311771	73.479998	-7.885171	15.136538
167	6.242150	1.659401	1.695021	2.146534	77.279997	-3.121477	15.136538
168	6.126612	1.662474	1.725369	3.783221	73.850000	-7.421335	15.136538
169	6.159417	1.660429	1.695292	2.099681	77.079999	-3.372194	15.136538
170	6.097894	1.653447	1.680014	1.606785	78.639996	-1.416576	15.136538
171	6.127400	1.648021	1.669070	1.277208	79.850000	0.100290	15.136538
172	6.093162	1.646899	1.670808	1.451743	79.439998	-0.513466	15.136538
173	6.118891	1.646340	1.681919	2.161102	78.469998	-1.728244	15.136538

```
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        174          |      6.095659        |      1.646992        |      1.669149
|      1.345276       |     79.749995        |     -0.125241        |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        175          |      6.131203        |      1.648076        |      1.677630
|      1.793243       |     78.799999        |     -1.314967        |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        176          |      6.111234        |      1.649875        |      1.681801
|      1.935102       |     78.560001        |     -1.615528        |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        177          |      6.161559        |      1.647128        |      1.718898
|      4.357234       |     74.390000        |     -6.837821        |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        178          |      6.125992        |      1.647206        |      1.660054
|      0.780025       |     80.739999        |      1.114588        |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        179          |      6.156801        |      1.645703        |      1.675382
|      1.803398       |     79.189998        |     -1.919744        |     15.136538
|
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|       Epoch         |   Epoch Time (s)     |   Training Loss      |     Test Loss
|    Overfit (%)      |    Accuracy (%)      |   Δ Accuracy (%)     |  GPU Memory
(GiB)  |
---------------------------------------------------------------------------------
---------------------------------------------------------------------------------
---------
|        180          |      6.101541        |      1.646676        |      1.699421
|      3.203148       |     76.969999        |     -4.669309        |     15.136538
```

181	6.070375	1.646551	1.670855	1.476050	79.560000	-1.461480	15.136538
182	6.133426	1.645668	1.692414	2.840550	77.370000	-4.173890	15.136538
183	6.103067	1.646507	1.673047	1.611929	79.369998	-1.696806	15.136538
184	6.220975	1.645006	1.687924	2.609001	77.609998	-3.876642	15.136538
185	6.173976	1.645559	1.670266	1.501455	79.539996	-1.486255	15.136538
186	6.171250	1.644896	1.676799	1.939509	78.759998	-2.452317	15.136538
187	6.168568	1.644711	1.667325	1.374997	79.879999	-1.065147	15.136538
188	6.133537	1.644299	1.690048	2.782248	77.759999	-3.690859	15.136538

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        189         |      6.135493      |      1.645682      |      1.683321
|      2.287131      |     78.410000      |     -2.885805      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        190         |      6.136440      |      1.646676      |      1.672363
|      1.559951      |     79.189998      |     -1.919744      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        191         |      6.104340      |      1.644953      |      1.697103
|      3.170348      |     76.800001      |     -4.879859      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        192         |      6.190171      |      1.642390      |      1.673998
|      1.924519      |     79.119998      |     -2.006441      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        193         |      6.137682      |      1.644042      |      1.684601
|      2.467033      |     78.109998      |     -3.257371      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        194         |      6.132952      |      1.644127      |      1.669210
|      1.525608      |     79.549998      |     -1.473868      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|       Epoch        |   Epoch Time (s)   |   Training Loss    |     Test Loss
|    Overfit (%)     |    Accuracy (%)    |   Δ Accuracy (%)   |  GPU Memory
(GiB)  |
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        195         |      6.213376      |      1.641013      |      1.665620
|      1.499475      |     80.070001      |     -0.829822      |     15.136538
```

```
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        196         |      6.159393      |      1.635479      |      1.667431
|      1.953695      |     80.039996      |     -0.866984      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        197         |      6.220014      |      1.636976      |      1.664628
|      1.689192      |     80.129999      |     -0.755511      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        198         |      6.141409      |      1.638255      |      1.662055
|      1.452788      |     80.409998      |     -0.408721      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------
|        199         |      6.201807      |      1.636783      |      1.666127
|      1.792761      |     79.899997      |     -1.040379      |     15.136538
|
--------------------------------------------------------------------------------
--------------------------------------------------------------------------------
---------

Training Time: 667.1745541095734 seconds
```


ResNet Training Curves
Training Loss
Test Loss
Accuracy
Loss
2.0
1.5
1.0
0.5
0.0
0
25
50
75
100
125
150
175
200
Epoch