

C1.

```
python C1.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
```

```
[gh22620b-10-392 ~] ls
C1.py  intel  intel-oneapi-base-toolkit-2025.2.1.44_offline.sh  intel-oneapi-base-toolkit-2025.2.1.46_offline.exe  ondemand  qh  test
[gh22620b-10-392 ~] python C1.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
100%
/usr/lib64/python3.9/tarfile.py:2239: RuntimeWarning: the default behavior of tarfile extraction has been changed to disallow common exploits (including CVE-2007-4599). By default, absolute/parent paths are disallowed and some mode bits are cleared. See https://access.redhat.com/articles/7004769 for more details.
  warnings.warn(
Epoch: 1/5
Epoch [1/5] - Loss: 1.9249, Accuracy: 28.48%
Epoch: 2/5
Epoch [2/5] - Loss: 1.4737, Accuracy: 45.25%
Epoch: 3/5
Epoch [3/5] - Loss: 1.2130, Accuracy: 55.90%
Epoch: 4/5
Epoch [4/5] - Loss: 1.0007, Accuracy: 64.16%
Epoch: 5/5
Epoch [5/5] - Loss: 0.8384, Accuracy: 70.48%
[gh22620b-10-392 ~] █
```

```
python C1.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
--profiler
```

```
Host: b-10-28
The authenticity of host 'b-10-28 (<no hostip for proxy command>)' can't be established.
EDDSA key fingerprint is SHA256:eF/bk6mc/5wyamL8Wt0d1e8MkAmM01R5EM9XRoG9VCM.
Are you sure you want to continue connecting (yes/no/[fingerprint])? yes
Warning: Permanently added 'b-10-28' (EDDSA) to the list of known hosts.
Register this system with Red Hat Insights: insights-client --register
Create an account or view all your systems at https://red.ht/insights-dashboard
Last login: Fri Oct  3 07:22:25 2025
=====
- Hostname.....: b-10-28
- IP Address....: 10.144.0.117
- Disk Space.....: remaining
=====
- CPU usage.....: 0.02, 0.30, 0.20 (1, 5, 15 min)
- Memory used....: 1036 MB / 23816 MB
- Swap in use.....: 0 MB
=====

[gh22620b-10-28 ~]$ ls
ECE9143  intel  intel-oneapi-base-toolkit-2025.2.1.44_offline.sh  intel-oneapi-base-toolkit-2025.2.1.46_offline.exe  ondemand  qh  test
[gh22620b-10-28 ~]$ cd ECE9143
[gh22620b-10-28 ECE9143]$ ls
C1.py  C2.py  C3.py  C4.py  C5.py  C6.py  C7.py  data  Draw.py  Parameters.py  requirements.txt
[gh22620b-10-28 ECE9143]$ python C1.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler

Epoch: 1/5
ERROR:2025-10-03 07:28:28 6163:6163 DeviceProperties.cpp:47] gpuGetDeviceCount failed with code 100
^Z
[1]+  Stopped                  python C1.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler
[gh22620b-10-28 ECE9143]$ vim C1.py
[gh22620b-10-28 ECE9143]$ python C1.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler

Epoch: 1/5
ERROR:2025-10-03 07:33:01 6413:6413 DeviceProperties.cpp:47] gpuGetDeviceCount failed with code 100

Epoch: 2/5

Epoch: 3/5

Epoch: 4/5

Epoch: 5/5
Saved the Log
[gh22620b-10-28 ECE9143]$ python C2.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler
```



```
C2.python C2.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
```

```
Host: b-10-28
Cl.py C2.py C3.py C4.py C5.py C6.py C7.py data Draw.py Parameters.py requirements.txt
[gh2262@b-10-28 EC29143]$ python Cl.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler
Epoch: 1/5
ERROR:2025-10-03 07:28:28 6163:6163 DeviceProperties.cpp:[47] gpuGetDeviceCount failed with code 100
[gh2262@b-10-28 EC29143]$ python Cl.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler
[gh2262@b-10-28 EC29143]$ python Cl.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler
Epoch: 1/5
ERROR:2025-10-03 07:33:01 6413:6413 DeviceProperties.cpp:[47] gpuGetDeviceCount failed with code 100

Epoch: 2/5
Epoch: 3/5
Epoch: 4/5
Epoch: 5/5
Saved the Log
[gh2262@b-10-28 EC29143]$ python C2.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
Epoch: 1/5
Epoch [1/5] - Loss: 1.8074, Accuracy: 35.86%
Epoch_time: 141.8869690
Load_time: 0.00384916
Train_time: 129.39829304

Epoch: 2/5
Epoch [2/5] - Loss: 1.3231, Accuracy: 51.47%
Epoch_time: 171.54980421
Load_time: 0.00384915
Train_time: 159.0537688

Epoch: 3/5
Epoch [3/5] - Loss: 1.0533, Accuracy: 62.28%
Epoch_time: 138.692884952
Load_time: 0.00175743
Train_time: 120.14757430

Epoch: 4/5
Epoch [4/5] - Loss: 0.8717, Accuracy: 69.06%
Epoch_time: 138.692882277
Load_time: 0.00166657
Train_time: 126.33509763

Epoch: 5/5
Epoch [5/5] - Loss: 0.7210, Accuracy: 74.92%
Epoch_time: 142.26107375
Load_time: 0.00305651
Train_time: 129.93332326
[gh2262@b-10-28 EC29143]$
```

```
python C2.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5  
--profiler
```



C3.python C3.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 0 worker

```

Authentication successful
My Interactive Sessions - NYU HPC
qh2262@b-10-55-~:ECE9143
Host: b-10-55
File: '/home/qh2262/.local/lib/python3.9/site-packages/torch/autograd/graph.py', line 829, in _engine_run_backward
    return _Variable._execution_engine.run_backward( # Calls into the C++ engine to run the backward pass
KeyboardInterrupt
[qh2262@b-10-55 ECE9143]$ via C3.py
[qh2262@b-10-55 ECE9143]$ python C3.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
workers: 0

Epoch: 1/5
Epoch [1/5] - Loss: 1.7589, Accuracy: 34.50%
Epoch_time: 148.11167844
Load_time: 0.00285682
Train_time: 135.65826920

Epoch: 2/5
Epoch [2/5] - Loss: 1.2638, Accuracy: 53.82%
Epoch_time: 148.74649138
Load_time: 0.00303918
Train_time: 136.53102900

Epoch: 3/5
Epoch [3/5] - Loss: 0.8846, Accuracy: 64.95%
Epoch_time: 148.80944532
Load_time: 0.00272700
Train_time: 136.56901150

Epoch: 4/5
Epoch [4/5] - Loss: 0.8071, Accuracy: 71.52%
Epoch_time: 147.35220350
Load_time: 0.00200235
Train_time: 135.17270600

Epoch: 5/5
Epoch [5/5] - Loss: 0.6865, Accuracy: 76.07%
Epoch_time: 150.32642370
Load_time: 0.00211100
Train_time: 136.14417582
workers: 0

Epoch: 1/6
Epoch [1/6] - Loss: 0.6973, Accuracy: 79.27%
Epoch_time: 185.67859205
Load_time: 0.11415334
Train_time: 184.736553623

Epoch: 2/6
Epoch [2/6] - Loss: 0.5622, Accuracy: 80.77%
Epoch_time: 228.42607827
Load_time: 0.11733313
Train_time: 227.46283976

Epoch: 3/6

```

4 workers

```
Host: b-10-66
Authentication successful
My Interactive Sessions - NYU HPC
qh2262@b-10-55:~/ECE9143
https://cod-burst-001.hpc.nyu.edu/pun/sys/shell/ssh/b-10-55
Themes: Default

Epoch: 3/6
Epoch [3/6] - Loss: 0.9846, Accuracy: 64.95%
Epoch_time: 148.80946332
Load_time: 0.00272700
Train_time: 136.56901150

Epoch: 4/6
Epoch [4/6] - Loss: 0.8071, Accuracy: 71.52%
Epoch_time: 147.35220350
Load_time: 0.00206238
Train_time: 135.17270600
workers: 4

Epoch: 5/6
Epoch [5/6] - Loss: 0.8860, Accuracy: 76.07%
Epoch_time: 150.32642370
Load_time: 0.00211100
Train_time: 138.14417352
workers: 4

Epoch: 1/5
Epoch [1/5] - Loss: 0.6073, Accuracy: 79.27%
Epoch_time: 185.67895205
Load_time: 0.11419334
Train_time: 184.73655323

Epoch: 2/5
Epoch [2/5] - Loss: 0.5622, Accuracy: 80.77%
Epoch_time: 228.42607827
Load_time: 0.11733318
Train_time: 227.46283976

Epoch: 3/5
Epoch [3/5] - Loss: 0.5277, Accuracy: 81.84%
Epoch_time: 192.03852278
Load_time: 0.11380725
Train_time: 191.09630444

Epoch: 4/5
Epoch [4/5] - Loss: 0.5117, Accuracy: 82.47%
Epoch_time: 176.85882157
Load_time: 0.11473717
Train_time: 169.88446335
workers: 8

Epoch: 5/5
Epoch [5/5] - Loss: 0.4836, Accuracy: 83.52%
Epoch_time: 172.24995544
Load_time: 0.11499205
Train_time: 171.25881509
workers: 8

Epoch: 1/5
```

8 workers

```
Host: b-10-66
Authentication successful
My Interactive Sessions - NYU HPC
qh2262@b-10-55:~/ECE9143
https://cod-burst-001.hpc.nyu.edu/pun/sys/shell/ssh/b-10-55
Themes: Default

Epoch: 4/6
Epoch [4/6] - Loss: 0.5117, Accuracy: 82.47%
Epoch_time: 176.85882157
Load_time: 0.11473717
Train_time: 169.88446335

Epoch: 5/6
Epoch [5/6] - Loss: 0.4836, Accuracy: 83.52%
Epoch_time: 172.24995544
Load_time: 0.11499205
Train_time: 171.25881509
workers: 8

Epoch: 1/5
Epoch [1/5] - Loss: 0.4763, Accuracy: 83.72%
Epoch_time: 180.60700104
Load_time: 0.11223399
Train_time: 179.48853803

Epoch: 2/5
Epoch [2/5] - Loss: 0.4584, Accuracy: 84.30%
Epoch_time: 175.155454950
Load_time: 0.11010358
Train_time: 172.85786715

Epoch: 3/5
Epoch [3/5] - Loss: 0.4437, Accuracy: 84.90%
Epoch_time: 174.71895299
Load_time: 0.12845280
Train_time: 173.62815824

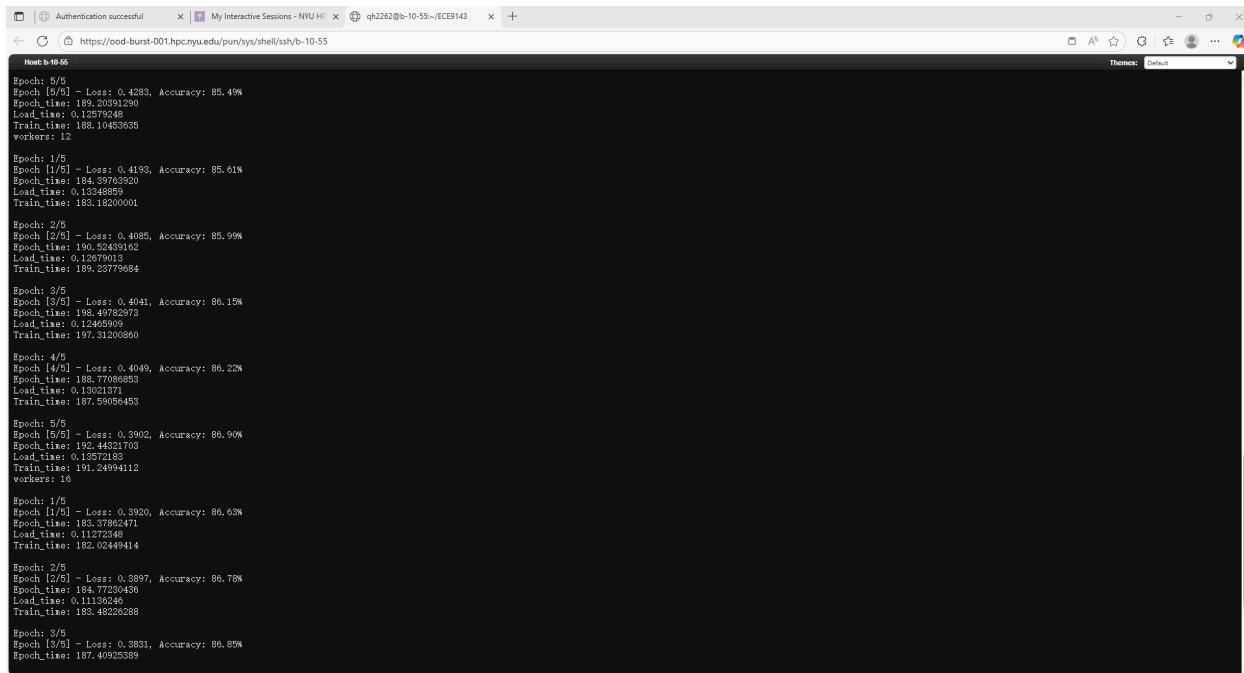
Epoch: 4/5
Epoch [4/5] - Loss: 0.4335, Accuracy: 85.27%
Epoch_time: 187.90633100
Load_time: 0.12462381
Train_time: 186.81826884

Epoch: 5/5
Epoch [5/5] - Loss: 0.4283, Accuracy: 85.49%
Epoch_time: 184.12572490
Load_time: 0.12570248
Train_time: 188.10453635
workers: 12

Epoch: 1/5
Epoch [1/5] - Loss: 0.4193, Accuracy: 85.61%
Epoch_time: 184.39763920
Load_time: 0.13346859
Train_time: 183.18200001

Epoch: 2/5
```

12 workers

```
Authentication successful | My Interactive Sessions - NYU HPC | qh2262@b-10-55-~|ECE9143 | + | https://cod-burst-001.hpc.nyu.edu/pun/sys/shell/ssh/b-10-55 | Themes: Default | 
```

Epoch: 1/5
Epoch [1/5] - Loss: 0.4283, Accuracy: 85.49%
Epoch_time: 189.20391290
Load_time: 0.12579248
Train_time: 188.10453655
workers: 12

Epoch: 1/5
Epoch [1/5] - Loss: 0.4193, Accuracy: 85.61%
Epoch_time: 184.39763920
Load_time: 0.13846859
Train_time: 183.38209001

Epoch: 2/5
Epoch [2/5] - Loss: 0.4069, Accuracy: 85.99%
Epoch_time: 190.50430142
Load_time: 0.12679013
Train_time: 189.23779084

Epoch: 3/5
Epoch [3/5] - Loss: 0.4041, Accuracy: 86.15%
Epoch_time: 198.49782973
Load_time: 0.12465909
Train_time: 197.31200860

Epoch: 4/5
Epoch [4/5] - Loss: 0.4049, Accuracy: 86.22%
Epoch_time: 188.77080853
Load_time: 0.13021371
Train_time: 187.59054533

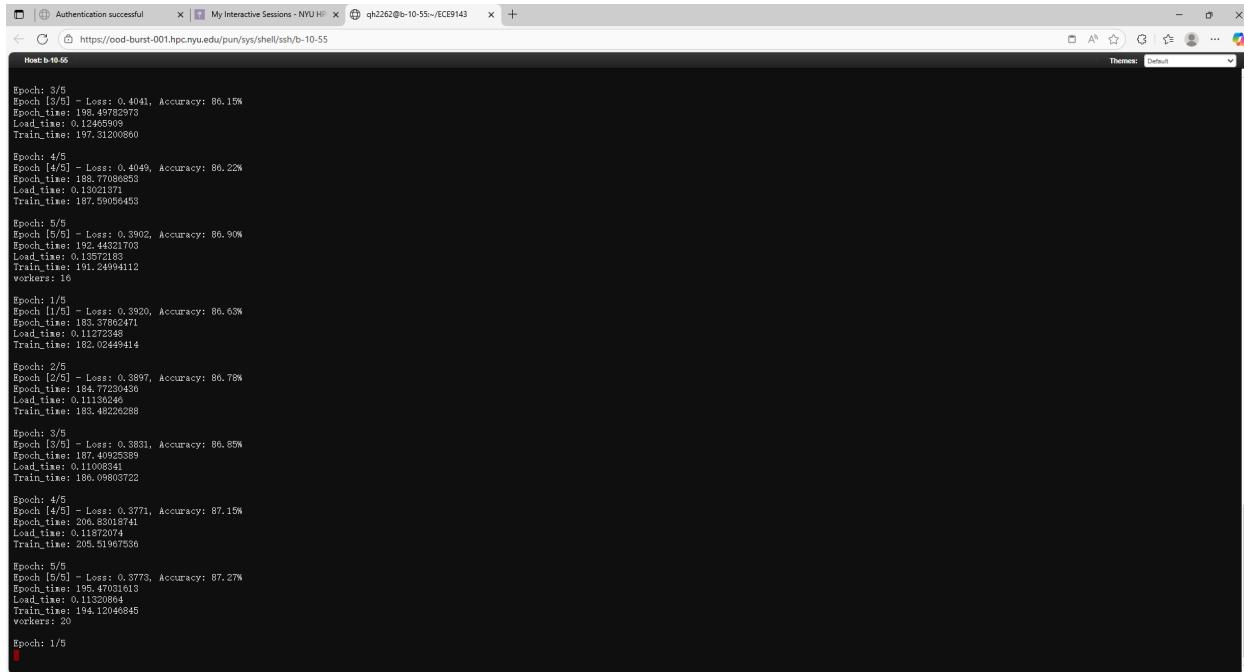
Epoch: 5/5
Epoch [5/5] - Loss: 0.3902, Accuracy: 86.90%
Epoch_time: 192.44321703
Load_time: 0.13572183
Train_time: 191.24994112
workers: 16

Epoch: 1/5
Epoch [1/5] - Loss: 0.3920, Accuracy: 86.63%
Epoch_time: 185.37862471
Load_time: 0.11272348
Train_time: 182.02449414

Epoch: 2/5
Epoch [2/5] - Loss: 0.3897, Accuracy: 86.78%
Epoch_time: 184.77230456
Load_time: 0.11362426
Train_time: 183.49220288

Epoch: 3/5
Epoch [3/5] - Loss: 0.3831, Accuracy: 86.85%
Epoch_time: 187.40925389

16 workers

```
Authentication successful | My Interactive Sessions - NYU HPC | qh2262@b-10-55-~|ECE9143 | + | https://cod-burst-001.hpc.nyu.edu/pun/sys/shell/ssh/b-10-55 | Themes: Default | 
```

Epoch: 3/5
Epoch [3/5] - Loss: 0.4041, Accuracy: 86.15%
Epoch_time: 184.39782973
Load_time: 0.13846859
Train_time: 197.31200860

Epoch: 4/5
Epoch [4/5] - Loss: 0.4049, Accuracy: 86.22%
Epoch_time: 188.77080853
Load_time: 0.13021371
Train_time: 187.59054533

Epoch: 5/5
Epoch [5/5] - Loss: 0.3902, Accuracy: 86.90%
Epoch_time: 192.44321703
Load_time: 0.13572183
Train_time: 191.24994112
workers: 16

Epoch: 1/5
Epoch [1/5] - Loss: 0.3920, Accuracy: 86.63%
Epoch_time: 185.37862471
Load_time: 0.11272348
Train_time: 182.02449414

Epoch: 2/5
Epoch [2/5] - Loss: 0.3897, Accuracy: 86.78%
Epoch_time: 184.77230456
Load_time: 0.11362426
Train_time: 183.49220288

Epoch: 3/5
Epoch [3/5] - Loss: 0.3831, Accuracy: 86.85%
Epoch_time: 187.40925389
Load_time: 0.11000341
Train_time: 186.09803722

Epoch: 4/5
Epoch [4/5] - Loss: 0.3771, Accuracy: 87.15%
Epoch_time: 182.02449414
Load_time: 0.11370274
Train_time: 205.51967556

Epoch: 5/5
Epoch [5/5] - Loss: 0.3773, Accuracy: 87.27%
Epoch_time: 198.47036163
Load_time: 0.11320844
Train_time: 194.12046845
workers: 20

Epoch: 1/5

20 workers

```

Authentications successful
My Interactive Sessions - NYU HPC
Dashboard - NYU HPC
qh2262@b-10-55-~/ECE9143
登录 - Google 账号
Host: b-10-55
Load_time: 0.11136246
Train_time: 183.48228288

Epoch: 3/5
Epoch [3/5] - Loss: 0.3831, Accuracy: 86.85%
Epoch_time: 187.3503389
Load_time: 0.11008234
Train_time: 186.09803722

Epoch: 4/5
Epoch [4/5] - Loss: 0.3771, Accuracy: 87.15%
Epoch_time: 206.83015741
Load_time: 0.11872074
Train_time: 205.51967536

Epoch: 5/5
Epoch [5/5] - Loss: 0.3773, Accuracy: 87.27%
Epoch_time: 195.47031613
Load_time: 0.11320964
Train_time: 194.12046845
workers: 20

Epoch: 1/5
Epoch [1/5] - Loss: 0.3726, Accuracy: 87.22%
Epoch_time: 195.35745753
Load_time: 0.11032448
Train_time: 195.89873720

Epoch: 2/5
Epoch [2/5] - Loss: 0.3691, Accuracy: 87.55%
Epoch_time: 183.89041617
Load_time: 0.12074880
Train_time: 182.44526344

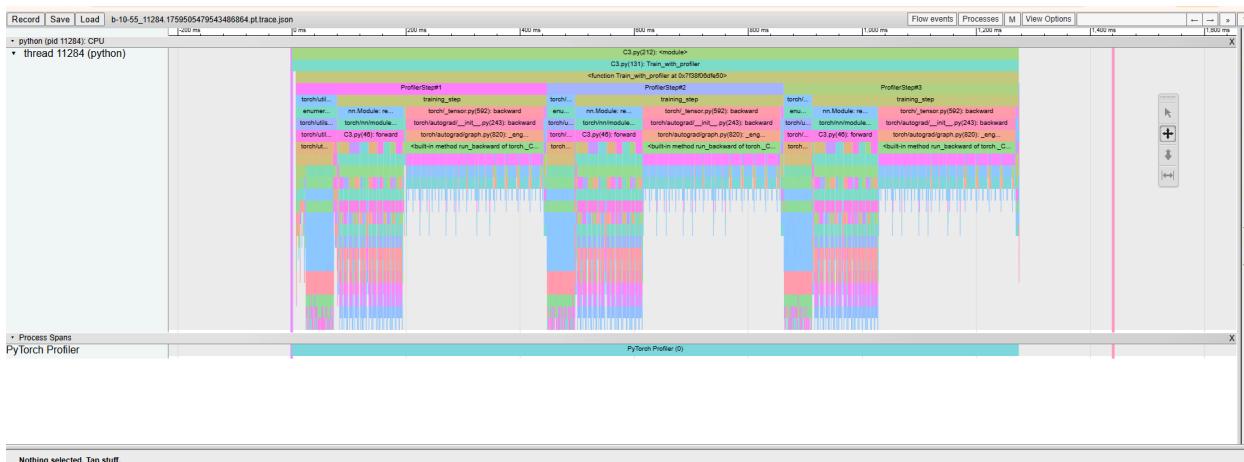
Epoch: 3/5
Epoch [3/5] - Loss: 0.3691, Accuracy: 87.57%
Epoch_time: 190.86458357
Load_time: 0.12468037
Train_time: 189.42699955

Epoch: 4/5
Epoch [4/5] - Loss: 0.3601, Accuracy: 87.82%
Epoch_time: 172.87423055
Load_time: 0.12538534
Train_time: 171.45303269

Epoch: 5/5
Epoch [5/5] - Loss: 0.3696, Accuracy: 87.71%
Epoch_time: 185.41062370
Load_time: 0.12135122
Train_time: 184.01590755
The lowest I/O time is 0.00211100 and the number of worker is 0
[qh2262@b-10-55 ECE9143]$ 

```

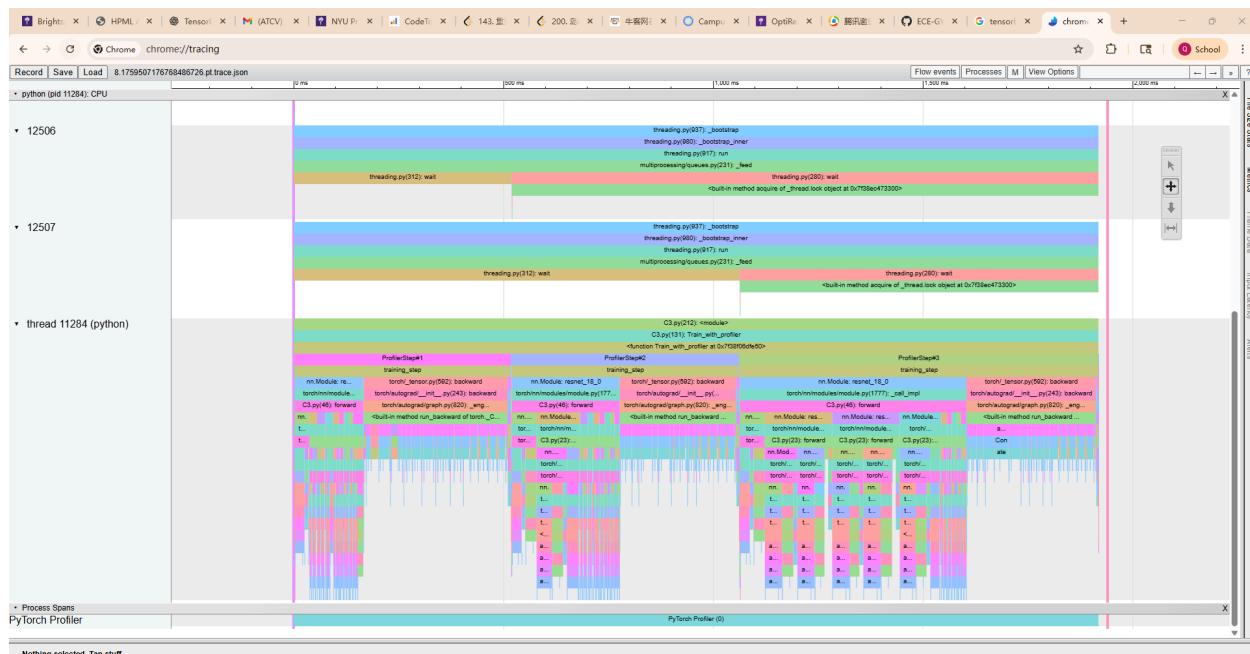
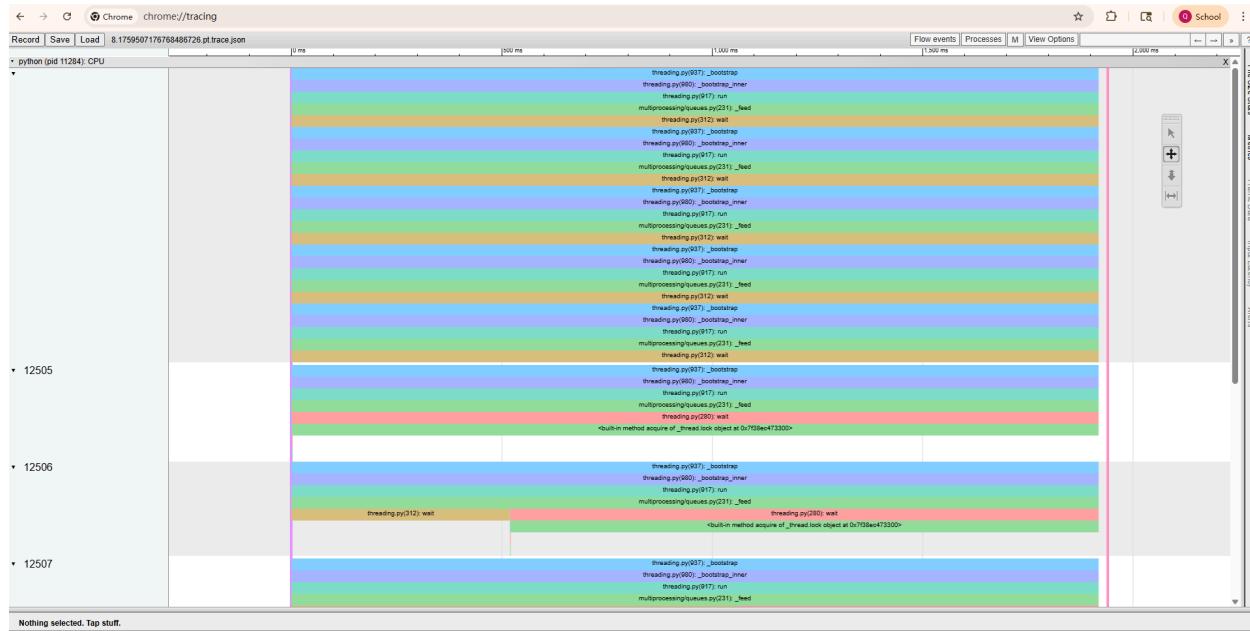
C3:python C3.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler 0 worker



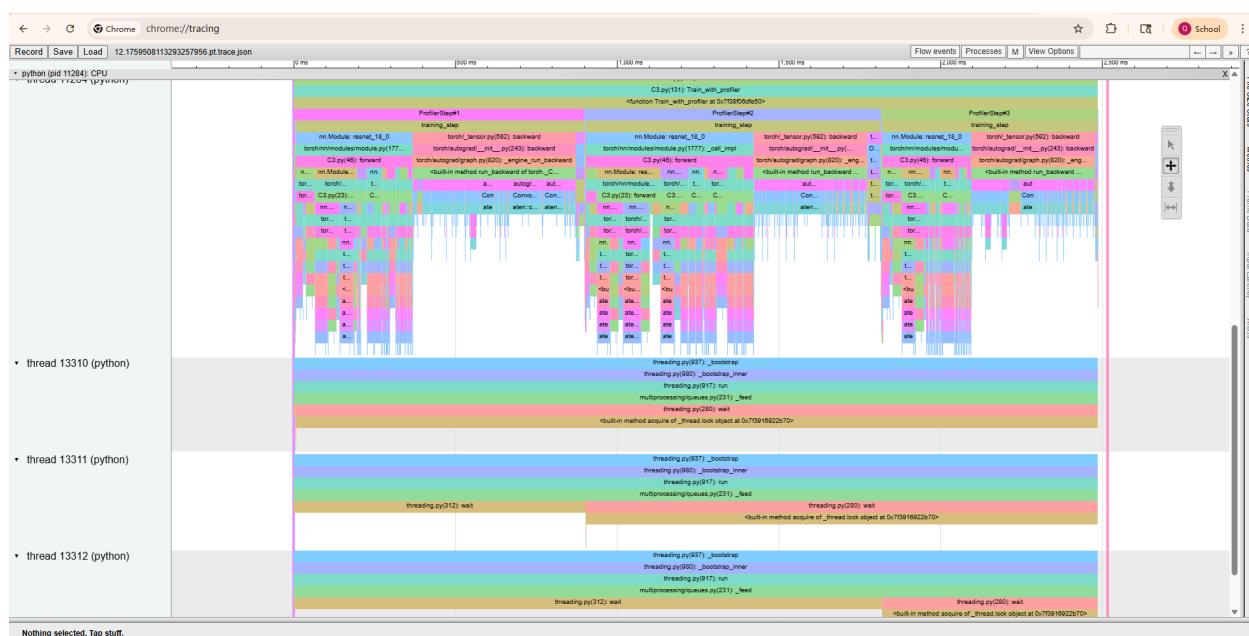
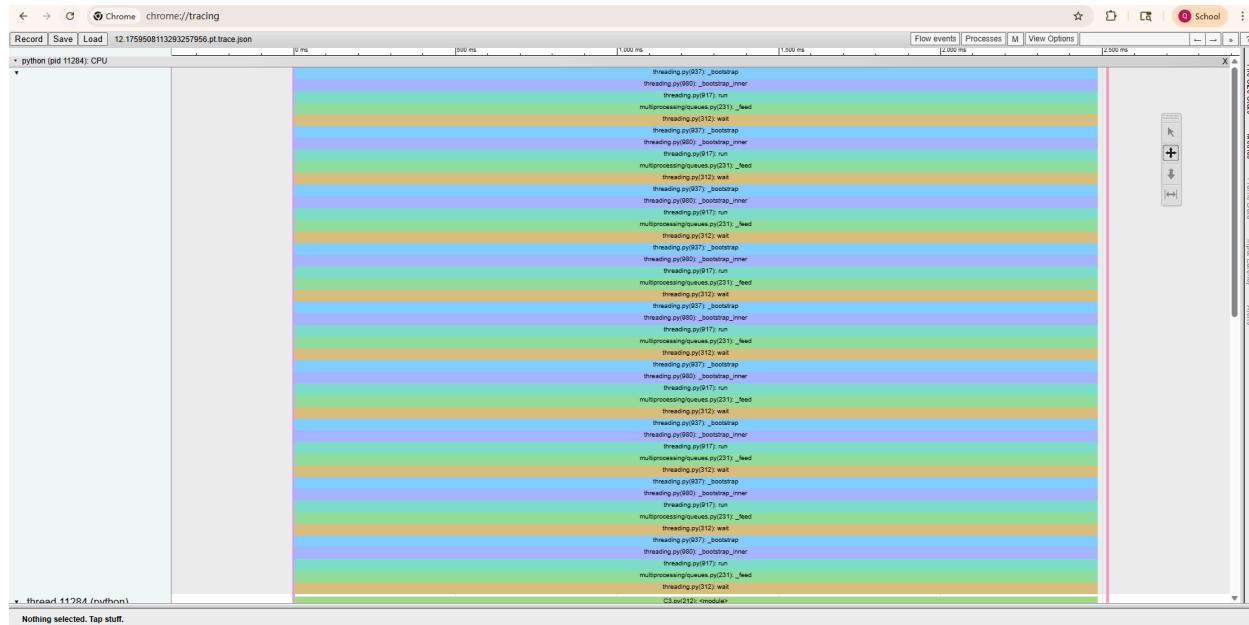
4 workers



8 workers

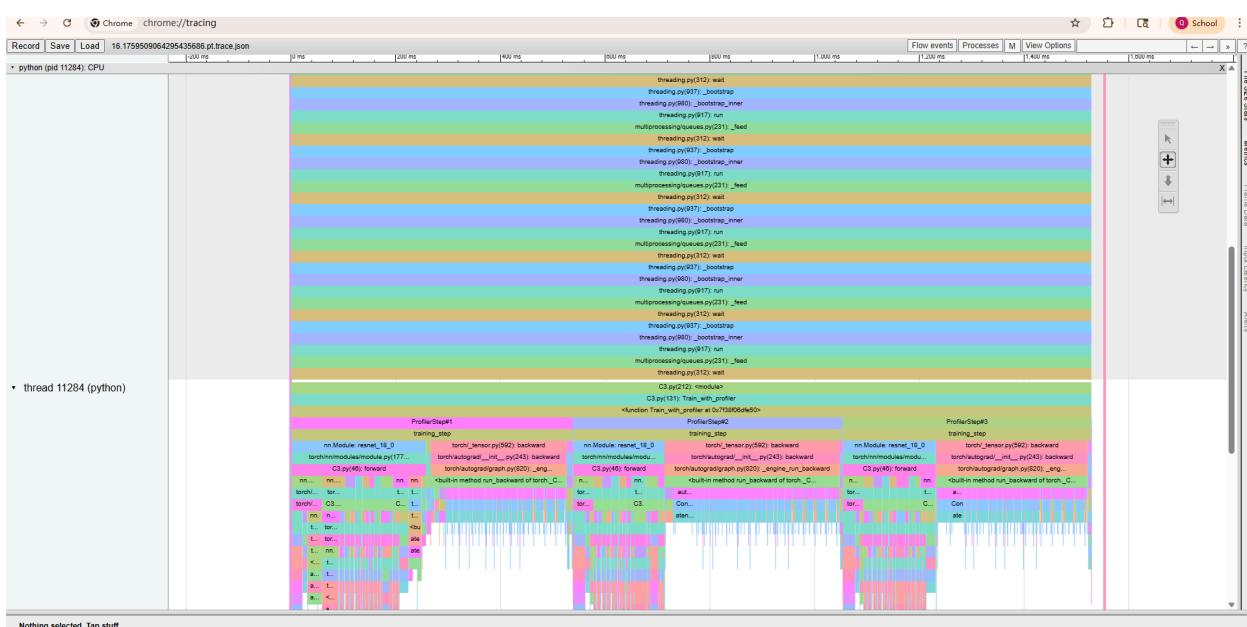
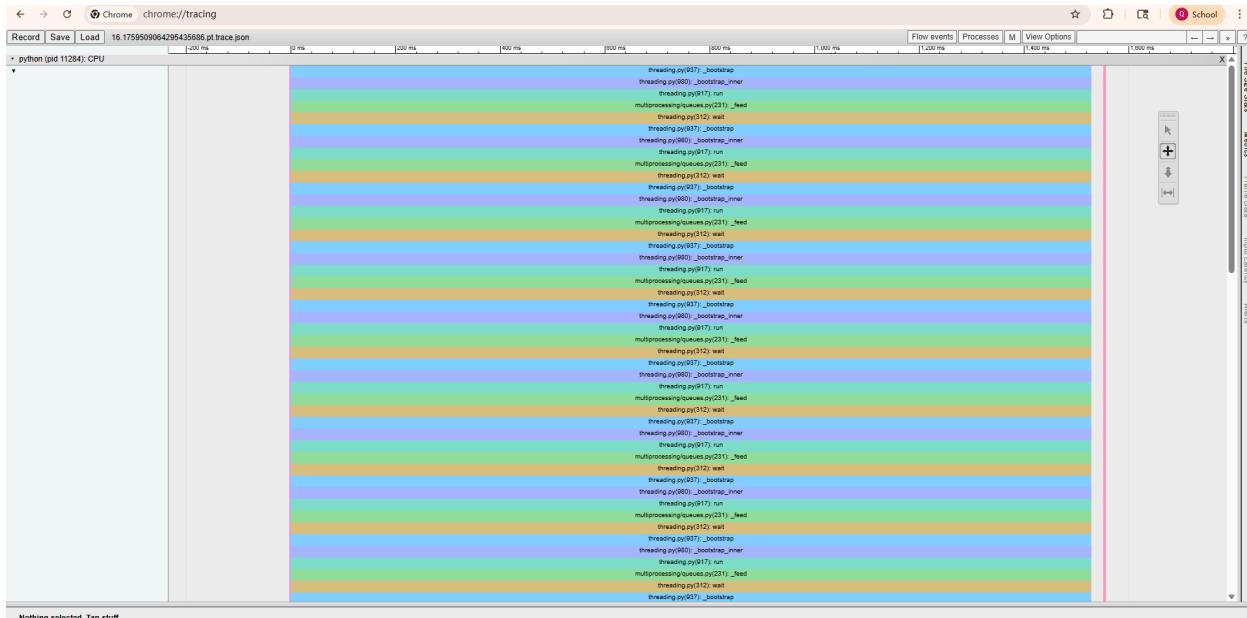


12 workers



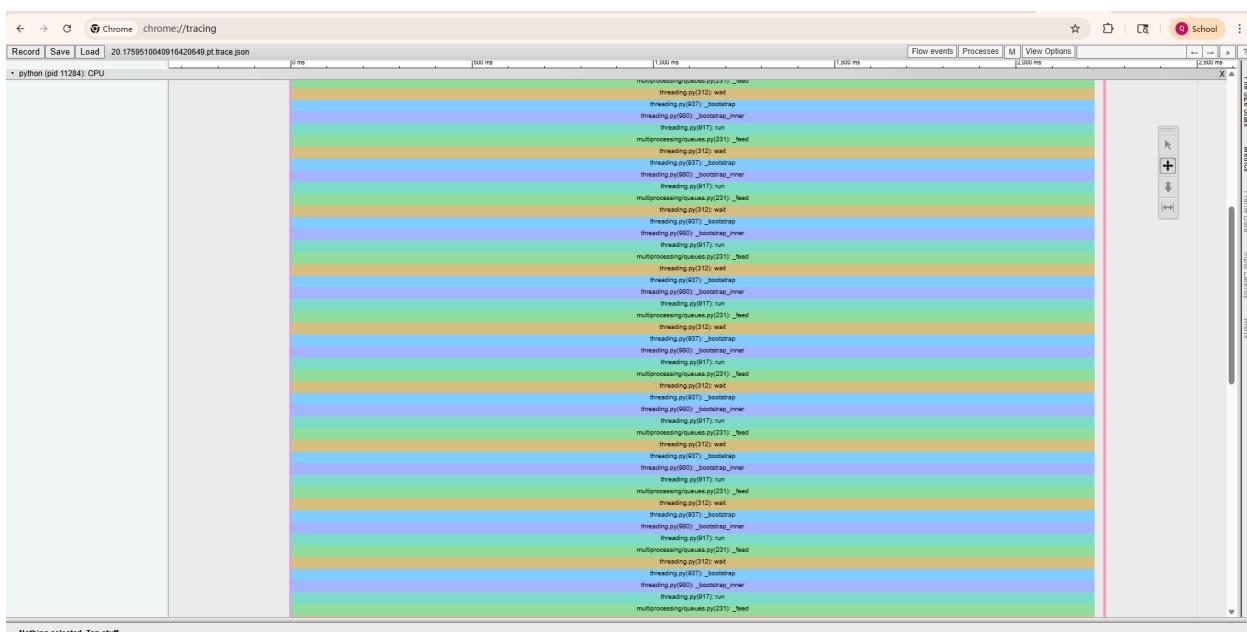
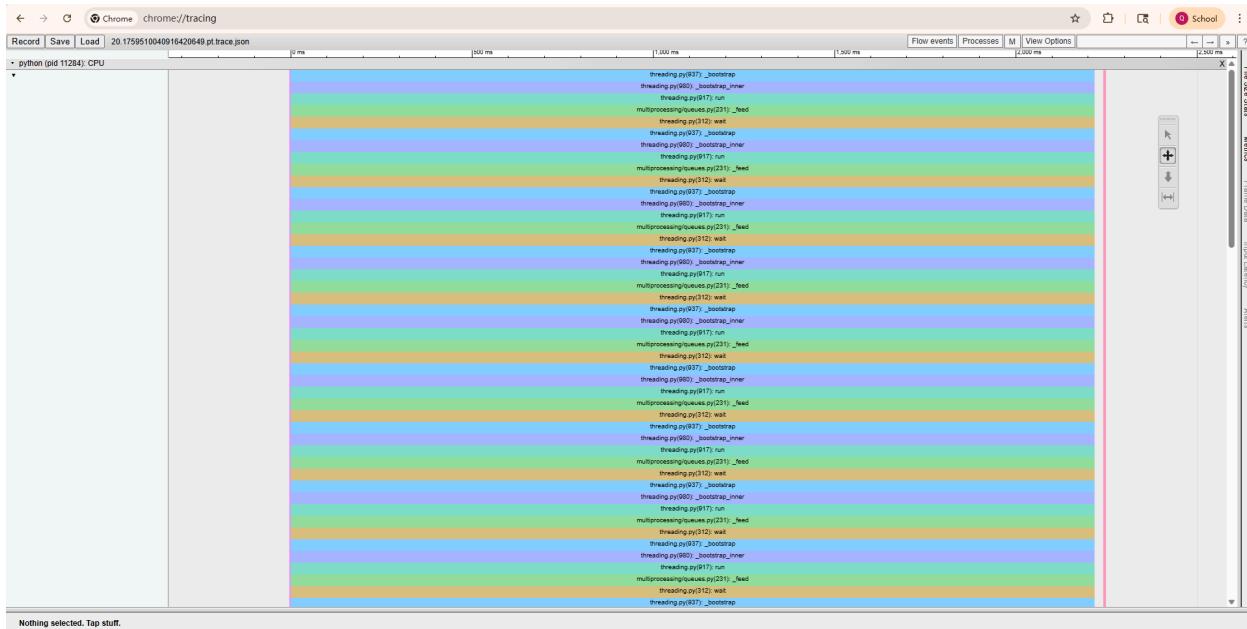


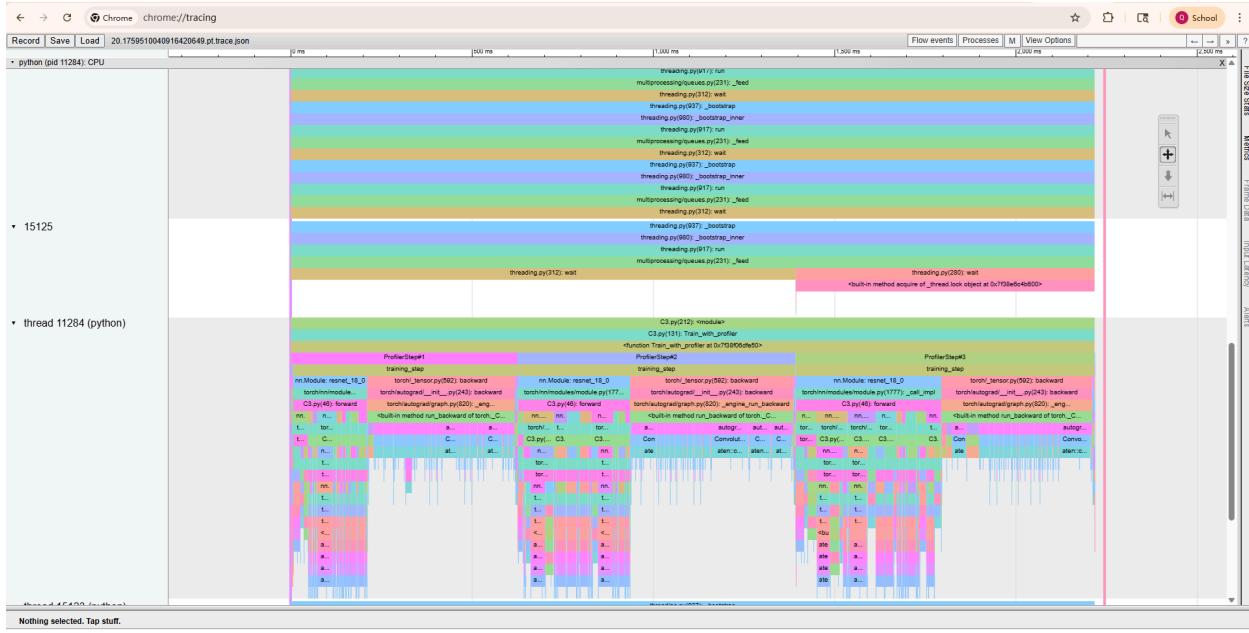
16 workers





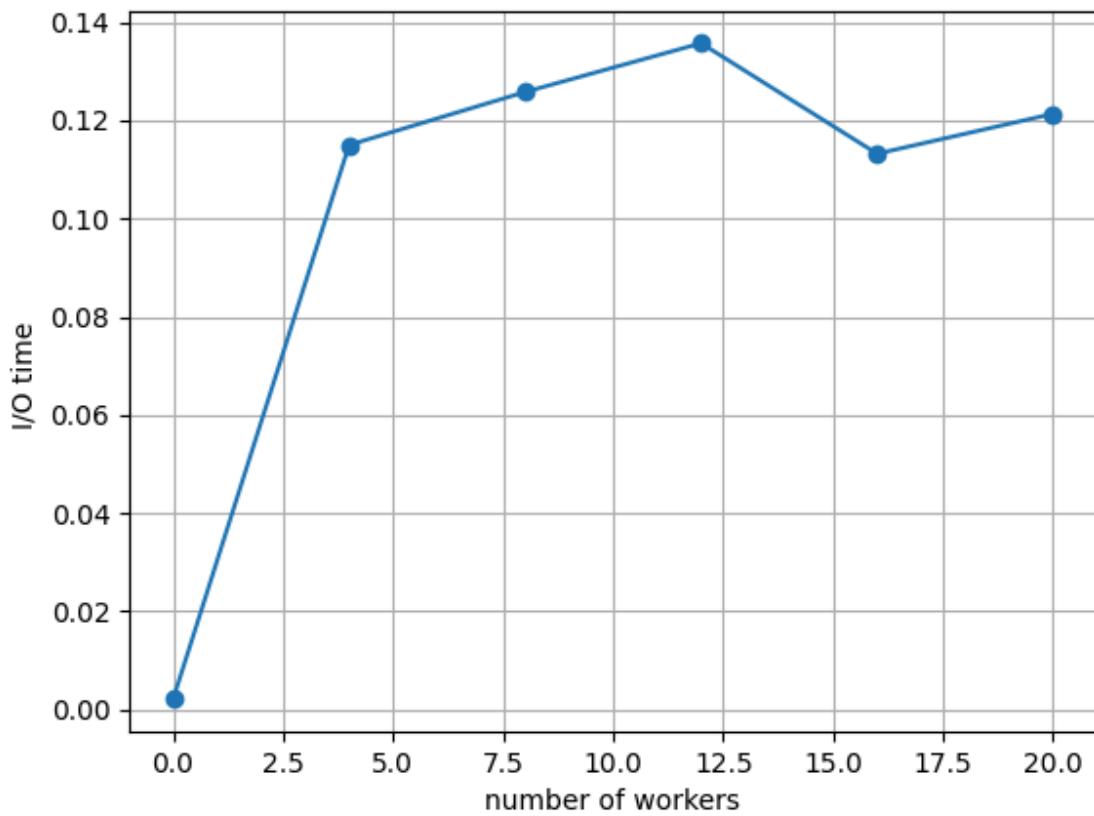
20 workers





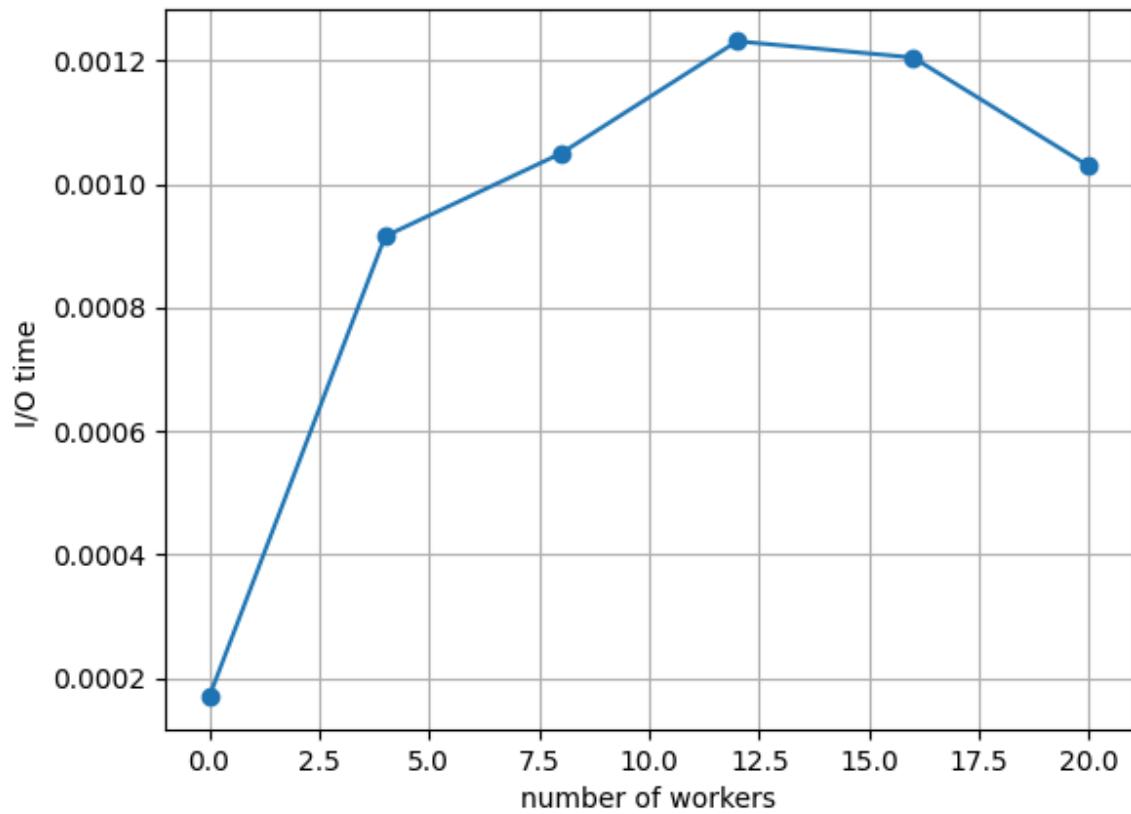
I/O time without profiler:

The I/O time vs Number of workers



I/O time with profiler

The I/O time vs Number of workers



Report how many workers are needed for the best runtime performance:

Without profiler: Worker 0 with 0.002111

```
Host: b-10-66
Load_time: 0.11136246
Train_time: 183.48228288

Epoch: 3/5
Epoch [3/5] - Loss: 0.3831, Accuracy: 86.85%
Epoch_time: 187.3203389
Load_time: 0.11008234
Train_time: 186.09803722

Epoch: 4/5
Epoch [4/5] - Loss: 0.3771, Accuracy: 87.15%
Epoch_time: 206.83015741
Load_time: 0.11872074
Train_time: 205.51967536
workers: 20

Epoch: 1/5
Epoch [1/5] - Loss: 0.3726, Accuracy: 87.22%
Epoch_time: 195.357457573
Load_time: 0.11032448
Train_time: 194.12046845

Epoch: 2/5
Epoch [2/5] - Loss: 0.3691, Accuracy: 87.55%
Epoch_time: 183.89041617
Load_time: 0.12074880
Train_time: 182.44526344

Epoch: 3/5
Epoch [3/5] - Loss: 0.3691, Accuracy: 87.57%
Epoch_time: 190.86458357
Load_time: 0.12464031
Train_time: 189.45399955

Epoch: 4/5
Epoch [4/5] - Loss: 0.3601, Accuracy: 87.82%
Epoch_time: 172.87423505
Load_time: 0.12538534
Train_time: 171.45362629

Epoch: 5/5
Epoch [5/5] - Loss: 0.3606, Accuracy: 87.71%
Epoch_time: 185.41062370
Load_time: 0.12135122
Train_time: 184.01596755
The lowest I/O time is 0.00211100 and the number of worker is 0
[qh2262@b-10-55 ECE9143]$
```

With profiler: Worker 0 with 0.00016916

```
Host: b-10-66
Authentication successful
My Interactive Sessions - NYU HPC
Dashboard - NYU HPC
qh2262@b-10-55~/.ECE9143
登录 - Google 账号
UmeTrack/lib/models/backbone
GitHub - xinghaochen/awesome-ml
Themes: Default

Epoch: 5/5
Saved the Log
workers: 8

Epoch: 1/5
Epoch: 2/5
Epoch: 3/5
Epoch: 4/5

Epoch: 5/5
Saved the Log
workers: 12

Epoch: 1/5
Epoch: 2/5
Epoch: 3/5
Epoch: 4/5

Epoch: 5/5
Saved the Log
workers: 16

Epoch: 1/5
Epoch: 2/5
Epoch: 3/5
Epoch: 4/5

Epoch: 5/5
Saved the Log
workers: 20

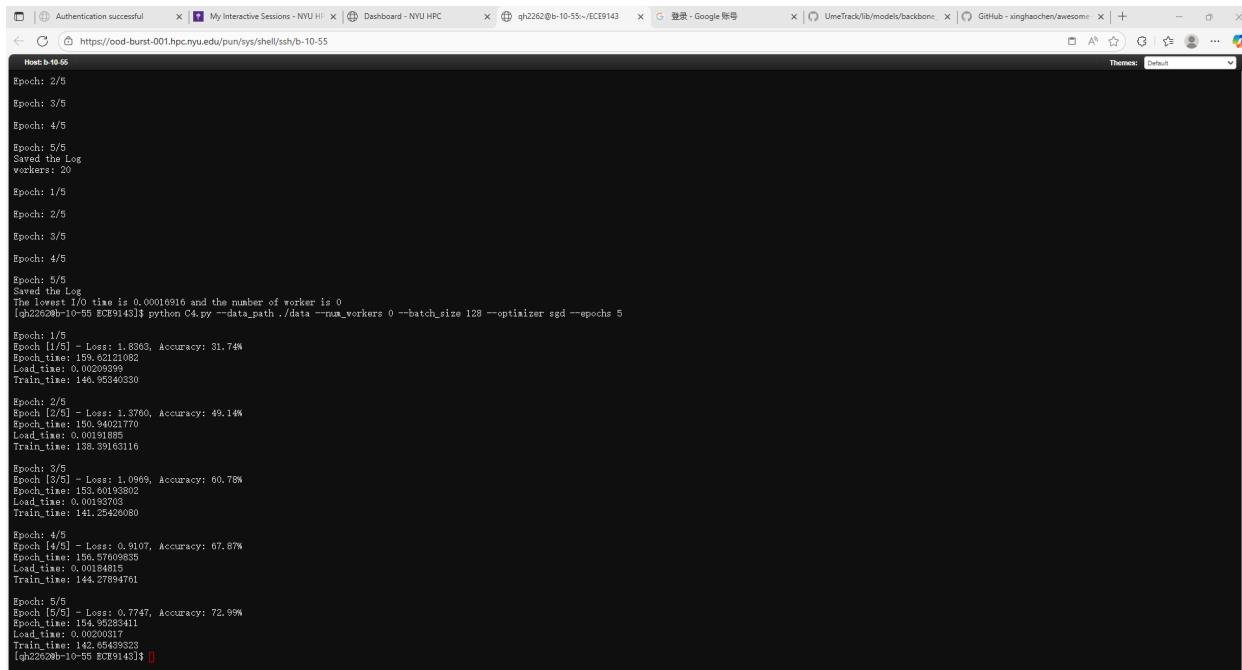
Epoch: 1/5
Epoch: 2/5
Epoch: 3/5
Epoch: 4/5

Epoch: 5/5
Saved the Log
The lowest I/O time is 0.00016916 and the number of worker is 0
[qh2262@b-10-55 ECE9143]$
```

C4:

Differences: The main difference between 0 workers and 1 workers is the asynchronous loading increases inter-process communication cost. For small batch size, the communication cost will be synchronous loading will be faster. We can check the graph in C3, from 0 worker to 4 worker also have a huge jump in loading time.

```
python C4.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
```



```
Host: b-10-55
Epoch: 2/5
Epoch: 3/5
Epoch: 4/5
Epoch: 5/5
Saved the Log
workers: 20
Epoch: 1/5
Epoch: 2/5
Epoch: 3/5
Epoch: 4/5
Epoch: 5/5
Saved the Log
The total I/O time is 0.00016916 and the number of worker is 0
[gh2262@b-10-55 EC9143]$ python C4.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5

Epoch: 1/5
Epoch [1/5] - Loss: 1.8365, Accuracy: 31.74%
Epoch_time: 159.62121082
Load_time: 0.00209399
Train_time: 146.95349330

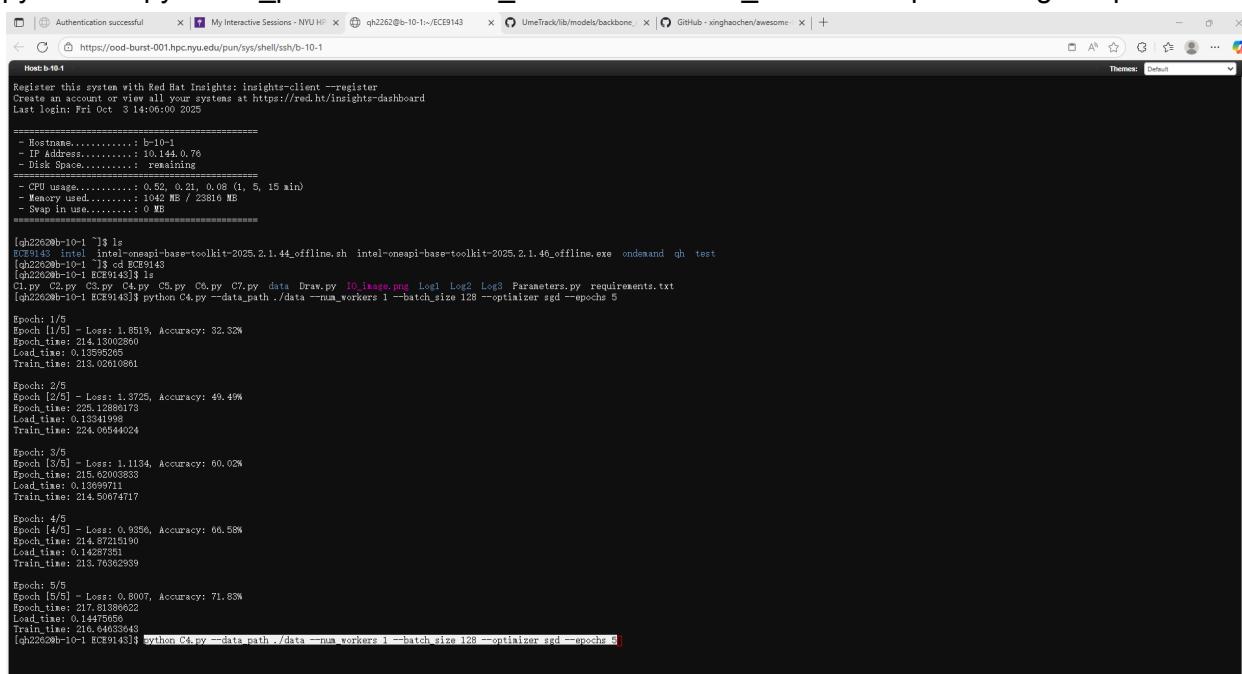
Epoch: 2/5
Epoch [2/5] - Loss: 1.3760, Accuracy: 49.14%
Epoch_time: 159.94021770
Load_time: 0.00191885
Train_time: 138.39165116

Epoch: 3/5
Epoch [3/5] - Loss: 1.0969, Accuracy: 60.78%
Epoch_time: 153.60193802
Load_time: 0.00191885
Train_time: 141.25420880

Epoch: 4/5
Epoch [4/5] - Loss: 0.9107, Accuracy: 67.87%
Epoch_time: 156.57690285
Load_time: 0.00184815
Train_time: 144.27894761

Epoch: 5/5
Epoch [5/5] - Loss: 0.7747, Accuracy: 72.99%
Epoch_time: 154.95283411
Load_time: 0.00200317
Train_time: 142.65493323
[gh2262@b-10-55 EC9143]$
```

```
python C4.py --data_path ./data --num_workers 1 --batch_size 128 --optimizer sgd --epochs 5
```



```
Host: b-10-1
Register this system with Red Hat Insights: insights-client --register
Create an account or view all your systems at https://red.ht/insights-dashboard
Last login: Fri Oct  3 14:06:00 2025

=====
- Hostname.....: b-10-1
- IP Address....: 10.144.0.76
- Disk Space.....: remaining
=====
- CPU usage.....: 0.52, 0.21, 0.08 (1, 5, 15 min)
- Memory used....: 1042 MB / 23816 MB
- Swap in use....: 0 MB
=====

[gh2262@b-10-1 ~]$ ls
EC9143  intel-oneapi-base-toolkit-2025.2.1.44_offline.sh  intel-oneapi-base-toolkit-2025.2.1.46_offline.exe  ondemand  qh  test
[gh2262@b-10-1 ~]$ cd EC9143
[gh2262@b-10-1 ~]$ python C4.py --data_path ./data --num_workers 1 --batch_size 128 --optimizer sgd --epochs 5

Epoch: 1/5
Epoch [1/5] - Loss: 1.8519, Accuracy: 32.32%
Epoch_time: 214.15002860
Load_time: 0.13595265
Train_time: 213.02610861

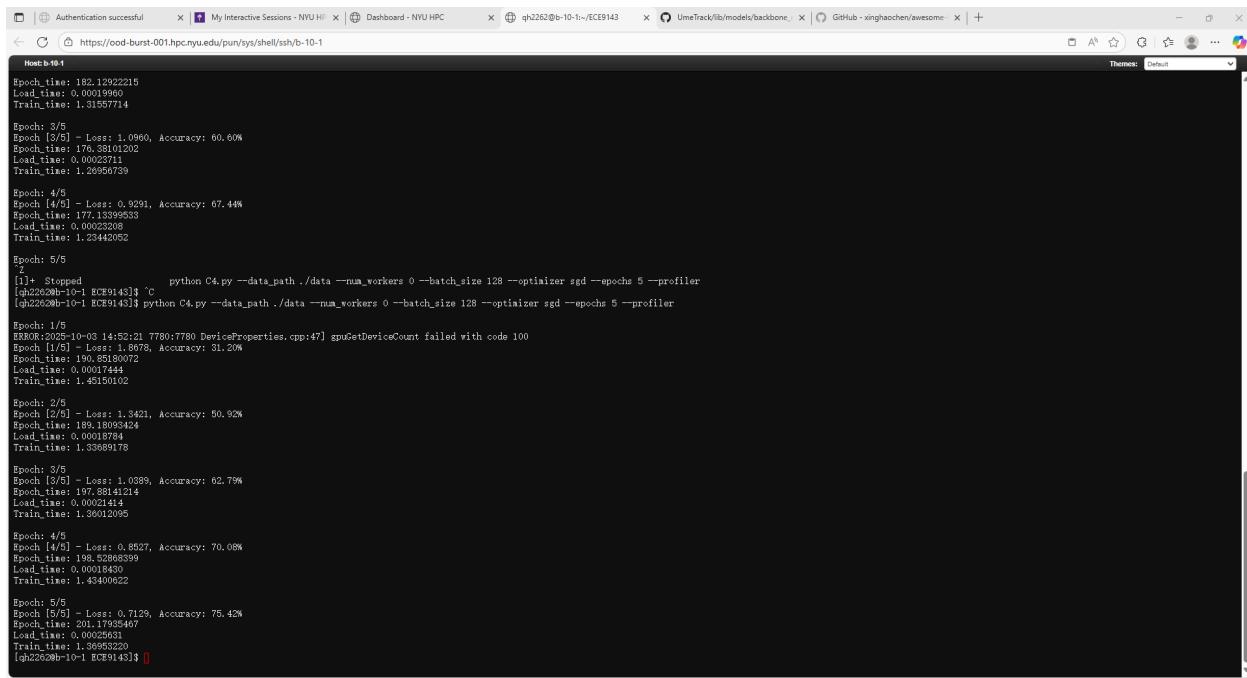
Epoch: 2/5
Epoch [2/5] - Loss: 1.3725, Accuracy: 49.49%
Epoch_time: 225.12886173
Load_time: 0.13341998
Train_time: 224.06544024

Epoch: 3/5
Epoch [3/5] - Loss: 1.1134, Accuracy: 60.02%
Epoch_time: 215.62003833
Load_time: 0.13099711
Train_time: 214.50674717

Epoch: 4/5
Epoch [4/5] - Loss: 0.9350, Accuracy: 66.58%
Epoch_time: 214.57215190
Load_time: 0.14287351
Train_time: 213.76362399

Epoch: 5/5
Epoch [5/5] - Loss: 0.8007, Accuracy: 71.83%
Epoch_time: 217.81386622
Load_time: 0.14475656
Train_time: 216.64603943
[gh2262@b-10-1 ~]$ python C4.py --data_path ./data --num_workers 1 --batch_size 128 --optimizer sgd --epochs 5
```

```
python C4.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5  
--profiler
```



Terminal window showing the execution of `python C4.py` with profiling and GPU device count errors.

```
Host: b-10-1
Epoch_time: 182.1922215
Load_time: 0.00019960
Train_time: 1.31557714

Epoch: 3/5
Epoch [3/5] - Loss: 1.0960, Accuracy: 60.60%
Epoch_time: 176.38101202
Load_time: 0.00023711
Train_time: 1.26956739

Epoch: 4/5
Epoch [4/5] - Loss: 0.9291, Accuracy: 67.44%
Epoch_time: 177.153399533
Load_time: 0.00023208
Train_time: 1.23442652

Epoch: 5/5
[1]+ Stopped python C4.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler
[qh2262@b-10-1 ECE9143]$ C
[qh2262@b-10-1 ECE9143]$ python C4.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler

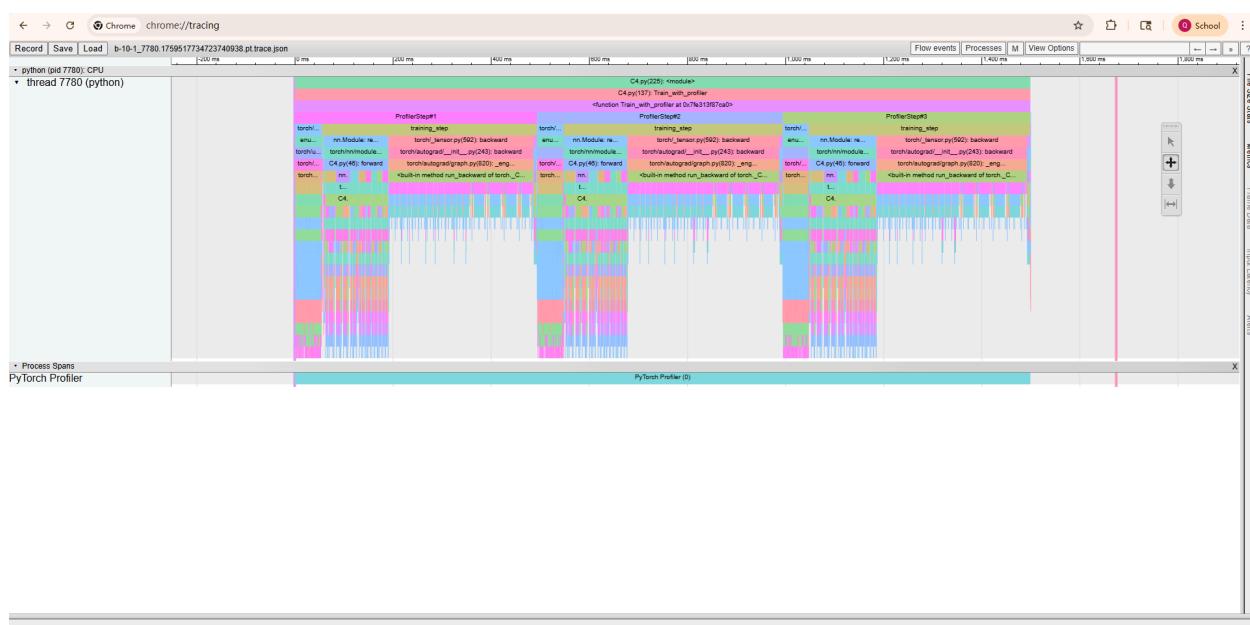
Epoch: 1/5
ERROR:2025-10-03 14:52:21 7780:7780 DeviceProperties.cpp:47] gpuGetDeviceCount failed with code 100
Epoch [1/5] - Loss: 1.8678, Accuracy: 31.20%
Epoch_time: 190.85180072
Load_time: 0.00017444
Train_time: 1.45150102

Epoch: 2/5
Epoch [2/5] - Loss: 1.3421, Accuracy: 50.92%
Epoch_time: 189.85180424
Load_time: 0.00018784
Train_time: 1.33689178

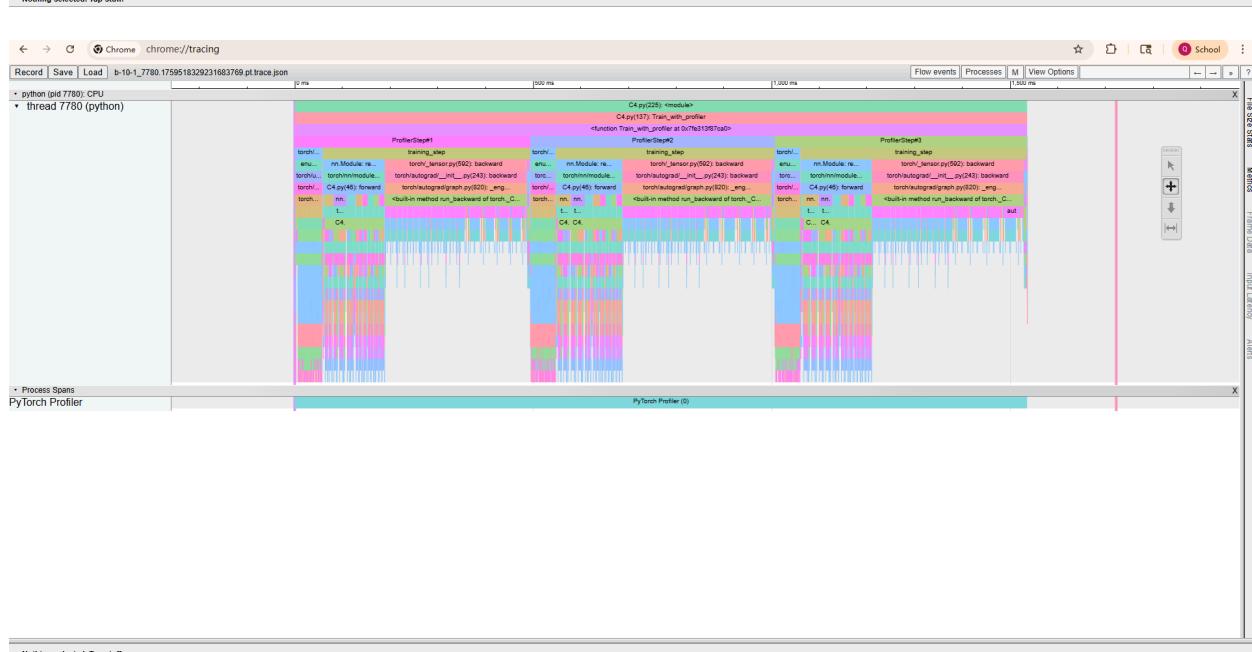
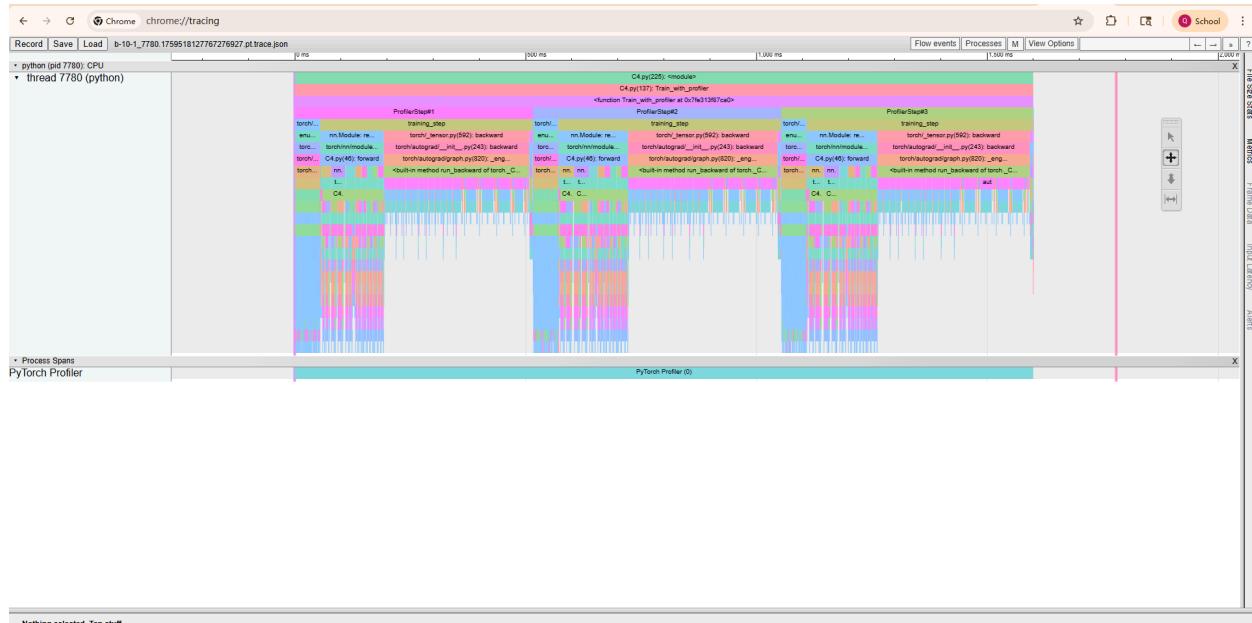
Epoch: 3/5
Epoch [3/5] - Loss: 1.0389, Accuracy: 62.79%
Epoch_time: 197.88141214
Load_time: 0.00021414
Train_time: 1.30012095

Epoch: 4/5
Epoch [4/5] - Loss: 0.8527, Accuracy: 70.08%
Epoch_time: 198.52868399
Load_time: 0.00018450
Train_time: 1.45409622

Epoch: 5/5
Epoch [5/5] - Loss: 0.7129, Accuracy: 75.42%
Epoch_time: 198.52868399
Load_time: 0.00021407
Train_time: 1.36953220
[qh2262@b-10-1 ECE9143]$
```







```
python C4.py --data_path ./data --num_workers 1 --batch_size 128 --optimizer sgd --epochs 5
--profiler
```

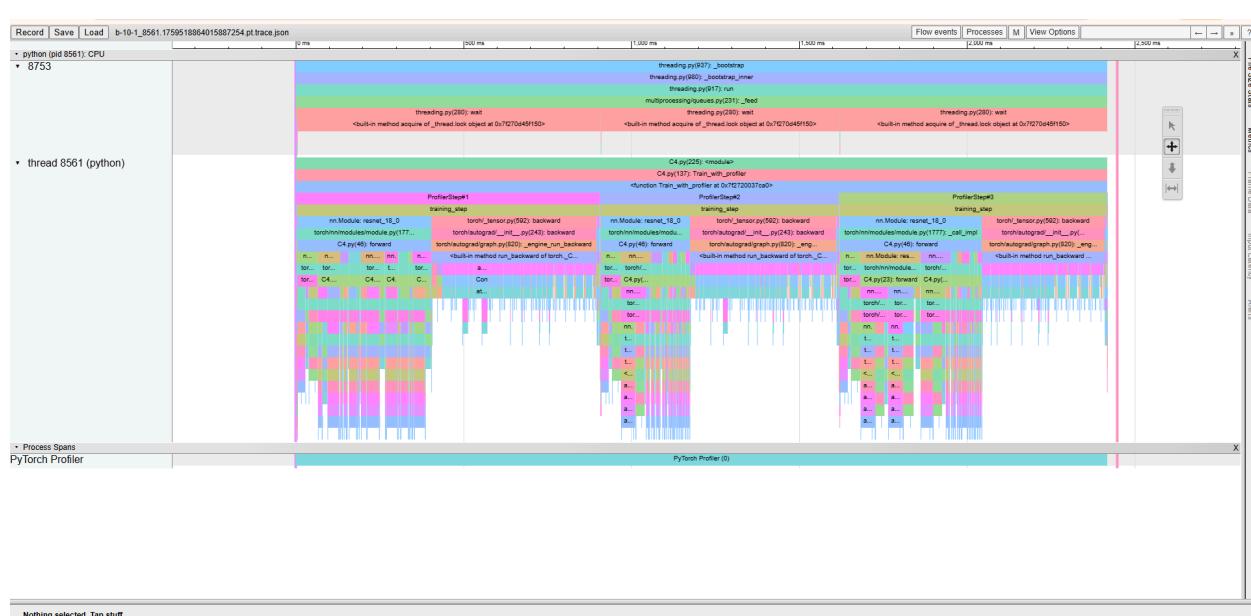
```
Epoch_time: 201.17935467
Load_time: 0.00025631
Train_time: 1.36953220
[qh2262@b-10-1 ECB9143]$ python C4.py --data_path ./data --num_workers 1 --batch_size 128 --optimizer sgd --epochs 5 --profiler

Epoch: 1/5
ERROR:2025-10-03 15:10:29 8561:8561 DeviceProperties.cpp:47] gpuGetDeviceCount failed with code 100
Epoch [1/5] - Loss: 1.8657, Accuracy: 31.26%
Epoch_time: 231.92585037
Load_time: 0.00099391
Train_time: 1.60340527

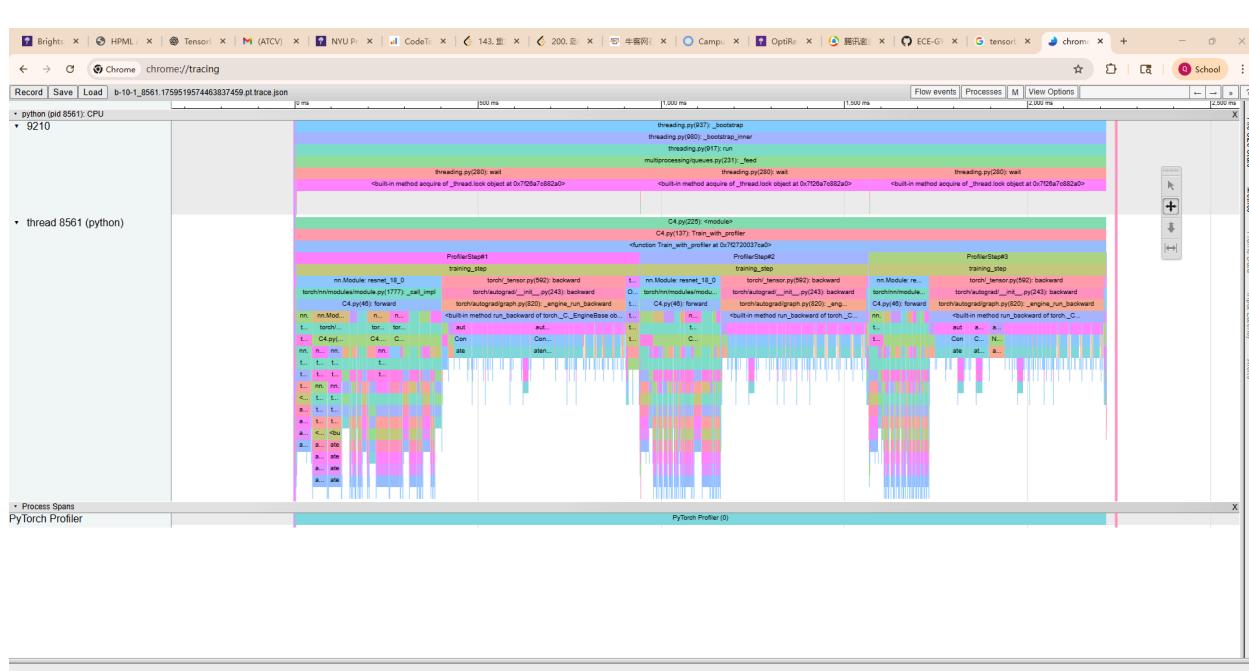
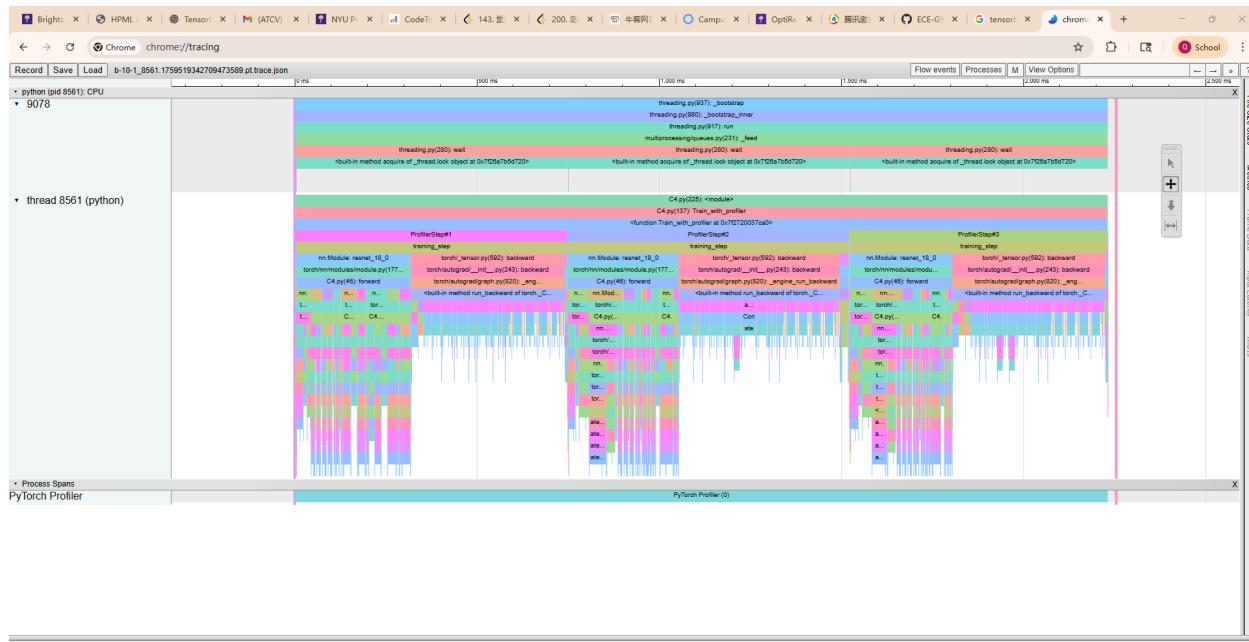
Epoch: 2/5
Epoch [2/5] - Loss: 1.3534, Accuracy: 50.28%
Epoch_time: 246.27521294
Load_time: 0.00104189
Train_time: 2.40912002

Epoch: 3/5
Epoch [3/5] - Loss: 1.0680, Accuracy: 61.65%
Epoch_time: 231.73902225
Load_time: 0.00121830
Train_time: 2.26222842

Epoch: 4/5
Epoch [4/5] - Loss: 0.8715, Accuracy: 69.31%
Epoch_time: 231.14171693
Load_time: 0.00098286
Train_time: 2.22359050
[qh2262@b-10-1 ECB9143]$ █
```







C5

```
python C5.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
```

```
[qh2262@b-10-1 ECE9143]$ vim C5.py
[qh2262@b-10-1 ECE9143]$ python C5.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5

Epoch: 1/5
Epoch [1/5] - Loss: 1.8308, Accuracy: 32.11%
Epoch_time: 174.37539533

Epoch: 2/5
Epoch [2/5] - Loss: 1.3702, Accuracy: 49.93%
Epoch_time: 176.27582715

Epoch: 3/5
Epoch [3/5] - Loss: 1.0733, Accuracy: 61.49%
Epoch_time: 176.66840814

Epoch: 4/5
Epoch [4/5] - Loss: 0.8551, Accuracy: 70.10%
Epoch_time: 175.07907607

Epoch: 5/5
Epoch [5/5] - Loss: 0.7072, Accuracy: 75.41%
Epoch_time: 177.96127058
The device is cpu: average running time over 5 epochs is : 176.07199545500006
[qh2262@b-10-1 ECE9143]$ █
```

```
python C5.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd
--epochs 5
```

```
[qh2262@b-31-1 ECE9143]$ pwd
/home/qh2262/ECE9143
[qh2262@b-31-1 ECE9143]$ ls
C1.py  C2.py  C3.py  C4.py  C5.py  C6.py  C7.py  data  Log1  Log2  Log3  Log4  Log5  Parameters.py  requirements.txt
[qh2262@b-31-1 ECE9143]$ python C5.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5

Epoch: 1/5
Epoch [1/5] - Loss: 1.6969, Accuracy: 37.54%
Epoch_time: 37.55738364

Epoch: 2/5
Epoch [2/5] - Loss: 1.1887, Accuracy: 57.16%
Epoch_time: 34.68771367

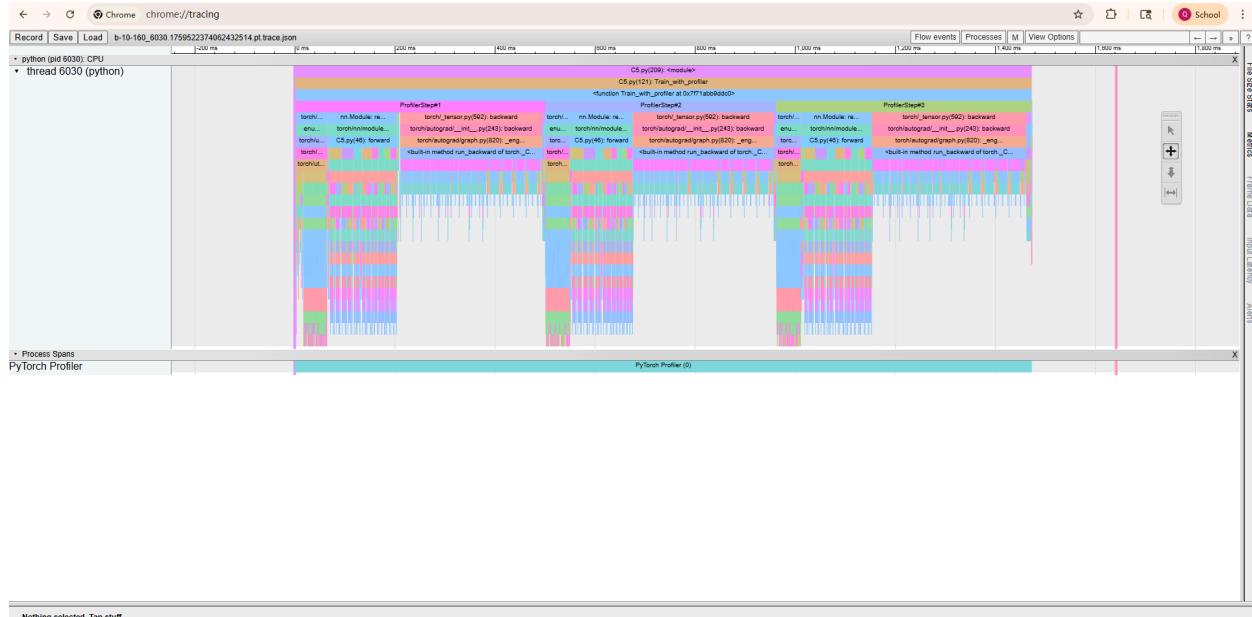
Epoch: 3/5
Epoch [3/5] - Loss: 0.9548, Accuracy: 66.17%
Epoch_time: 34.40106243

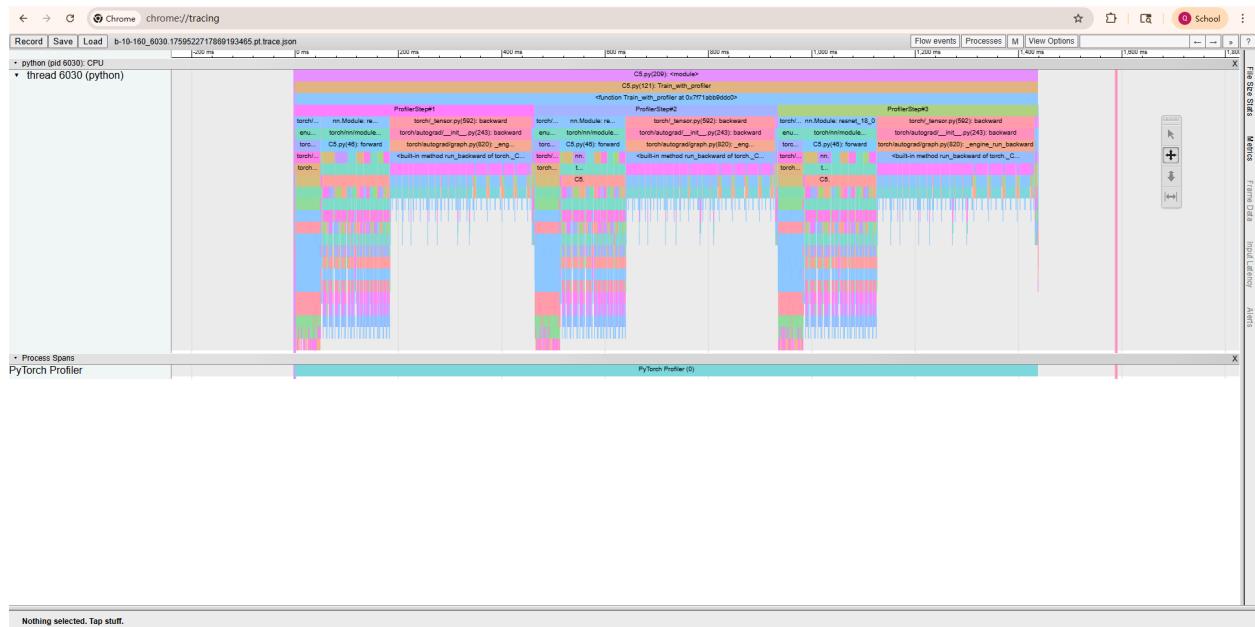
Epoch: 4/5
Epoch [4/5] - Loss: 0.7766, Accuracy: 72.70%
Epoch_time: 34.16218500

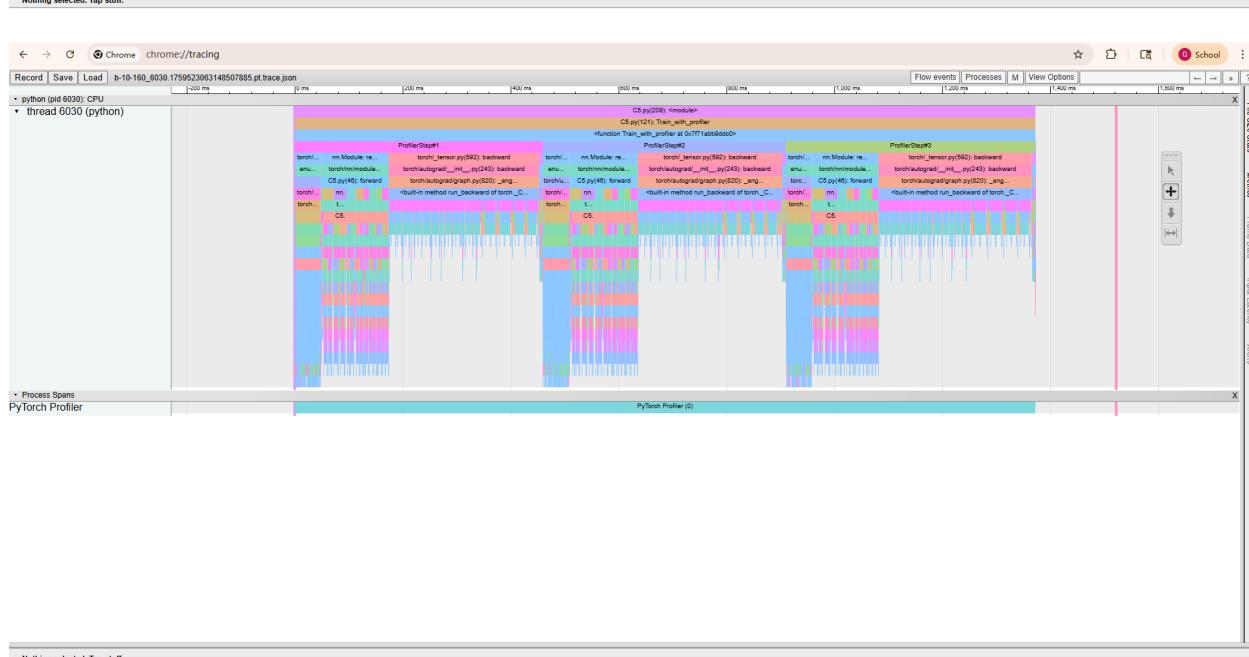
Epoch: 5/5
Epoch [5/5] - Loss: 0.6703, Accuracy: 76.83%
Epoch_time: 34.10473049
The device is cuda: average running time over 5 epochs is : 34.98261504359999
[qh2262@b-31-1 ECE9143]$ █
```

```
python C5.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
--profiler
```

```
Authentication successful
My Interactive Sessions - NYU HPC
Dashboard - NYU HPC
qh2262@b-10-160:~/ECE9143
GitHub - xinghochen/awesome
Host: b-10-160
Last login: Fri Oct 3 16:09:33 2025
=====
- Hostname.....: b-10-160
- IP Address....: 10.144.0.126
- Disk Space.....: remaining
- CPU usage.....: 0.15, 0.16, 0.07 (1. 5, 15 min)
- Memory used....: 1087 MB / 23816 MB
- Swap in use....: 0 MB
=====
[qh2262@b-10-160 ~]$ ls
-bash: 1: command not found
[qh2262@b-10-160 ~]$ cd ECE9143
[qh2262@b-10-160 ECE9143]$ ls
C1.py C2.py C3.py C4.py C5.py C6.py C7.py data Draw.py 10_image.png Log1 Log2 Log3 Log4 Parameters.py requirements.txt
[qh2262@b-10-160 ECE9143]$ python C5.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler
Epoch: 1/5
Traceback (most recent call last):
  File "/home/qh2262/ECE9143/C5.py", line 209, in <module>
    train_fn = train_fn_train_fn(model_fn, optims, loss_fn, epoch, device)
  File "/home/qh2262/ECE9143/C5.py", line 102, in train_fn_train_fn
    with torch.profiler.profile(
TypeError: __init__ got an unexpected keyword argument 'schedule'
[qh2262@b-10-160 ECE9143]$ python C5.py --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler
Epoch: 1/5
ERROR:2025-10-03 16:12:53 6030:6030 DeviceProperties.cpp:47] gpuGetDeviceCount failed with code 100
Epoch [1/5] - Loss: 1.972, Accuracy: 26.97%
Epoch_time: 172.68675499
Epoch: 2/5
Epoch [2/5] - Loss: 1.4493, Accuracy: 46.37%
Epoch_time: 171.66861219
Epoch: 3/5
Epoch [3/5] - Loss: 1.1056, Accuracy: 60.33%
Epoch_time: 174.95321393
Epoch: 4/5
Epoch [4/5] - Loss: 0.8924, Accuracy: 68.55%
Epoch_time: 170.40523973
Epoch: 5/5
Epoch [5/5] - Loss: 0.7299, Accuracy: 74.75%
Epoch_time: 181.94767053
The device is gpu: average running time over 5 epochs is : 174.3683090754
[qh2262@b-10-160 ECE9143]$
```







```
python C5.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd
--epochs 5 --profiler
```

```
Host: b-31-10
- Swap in use.....: 0 MB
=====
[gb22620b-31-10 ]$ ls
[gb9145 intel intel-oneapi-base=toolkit=2025.2.1.44_offline.sh intel-oneapi-base=toolkit=2025.2.1.46_offline.exe onedamand gh test
[gb22620b-31-10 ]$ cd ECE9143
[gb22620b-31-10 ECE9143]$ C1.py C2.py C3.py C4.py C5.py C6.py C7.py data Draw.py IO.images.png Log1 Log2 Log3 Log4 Log5 Parameters.py requirements.txt
[gb22620b-31-10 ECE9143]$ python C5.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5
Epoch: 1/5
Epoch [1/5] - Loss: 1.7472, Accuracy: 36.43%
Epoch_time: 37.67732133

Epoch: 2/5
Epoch [2/5] - Loss: 1.2566, Accuracy: 54.27%
Epoch_time: 34.36656985

Epoch: 3/5
Epoch [3/5] - Loss: 0.9887, Accuracy: 65.03%
Epoch_time: 34.14648203

Epoch: 4/5
Epoch [4/5] - Loss: 0.7897, Accuracy: 72.25%
Epoch_time: 34.04618015

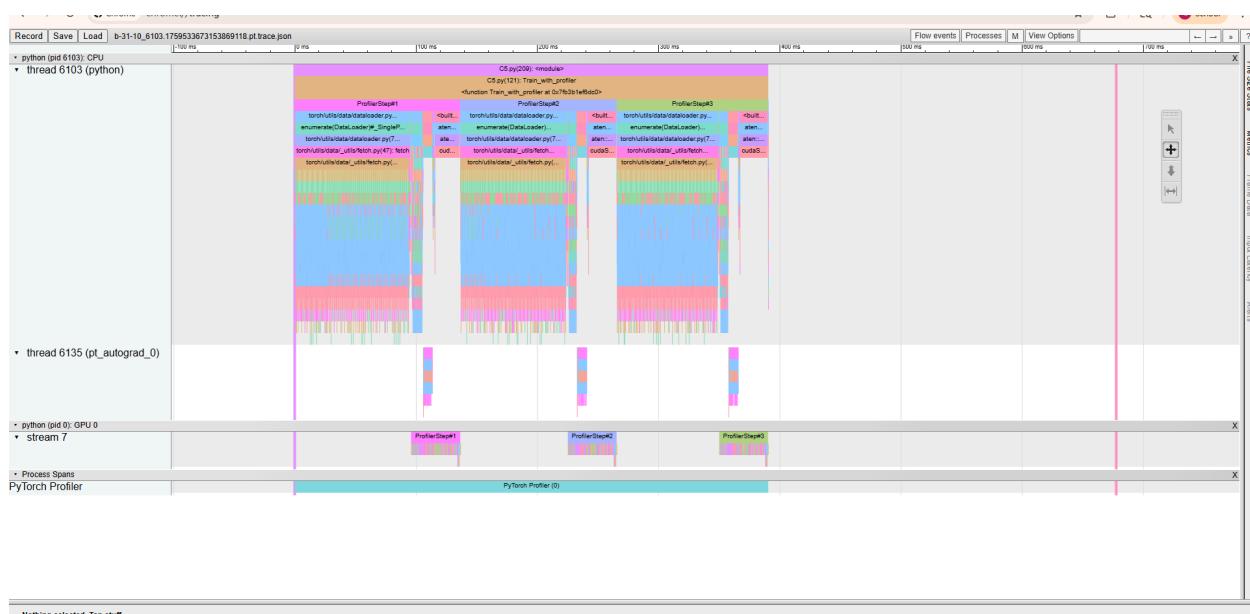
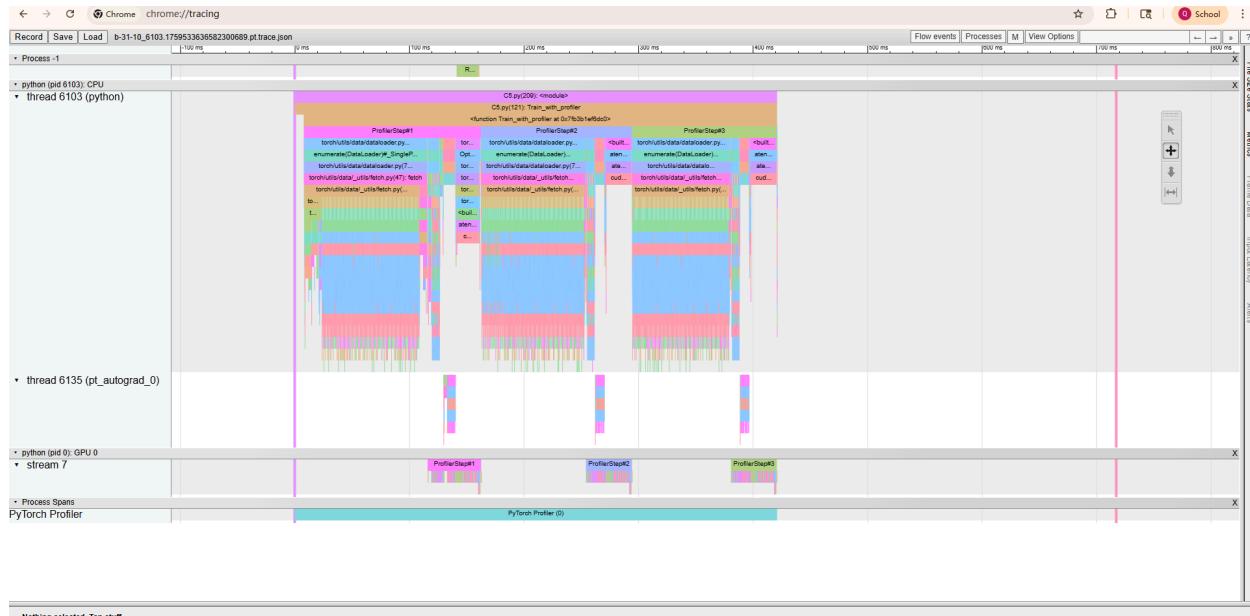
Epoch: 5/5
Epoch [5/5] - Loss: 0.6717, Accuracy: 75.87%
Epoch_time: 33.98802795
The device is cuda: average running time over 5 epochs is : 34.845316986279999
[gb22620b-31-10 ECE9143]$ python C5.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5 --profiler
Epoch: 1/5
Epoch [1/5] - Loss: 1.8091, Accuracy: 33.55%
Epoch_time: 37.10350019

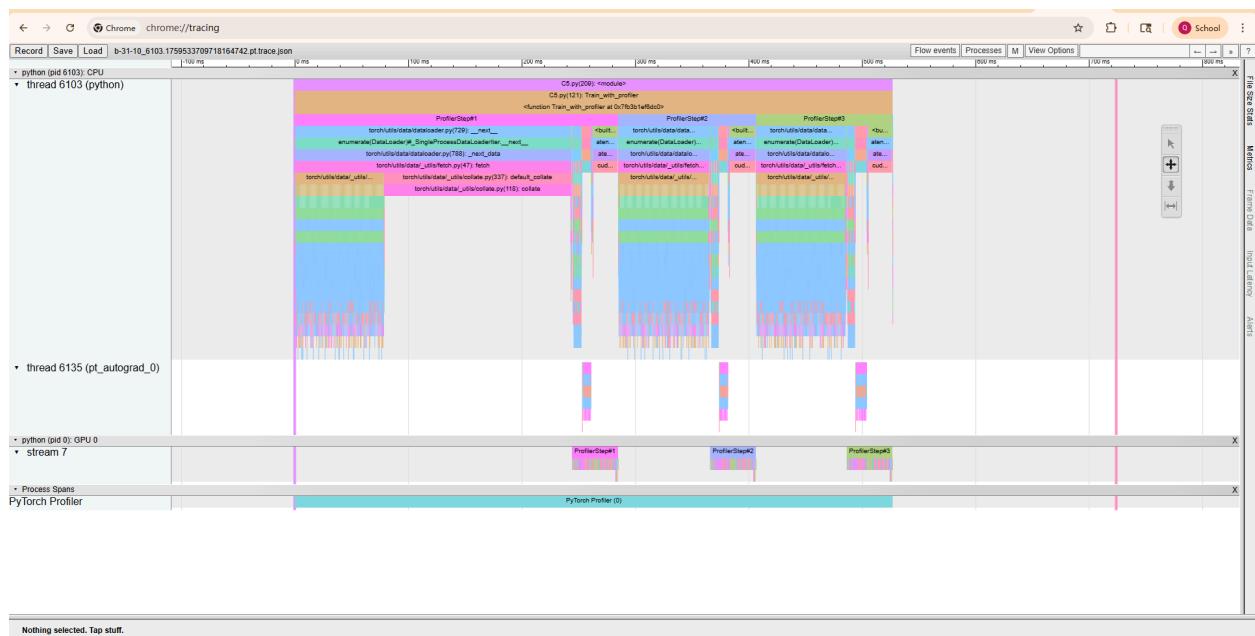
Epoch: 2/5
Epoch [2/5] - Loss: 1.3389, Accuracy: 50.80%
Epoch_time: 36.53371767

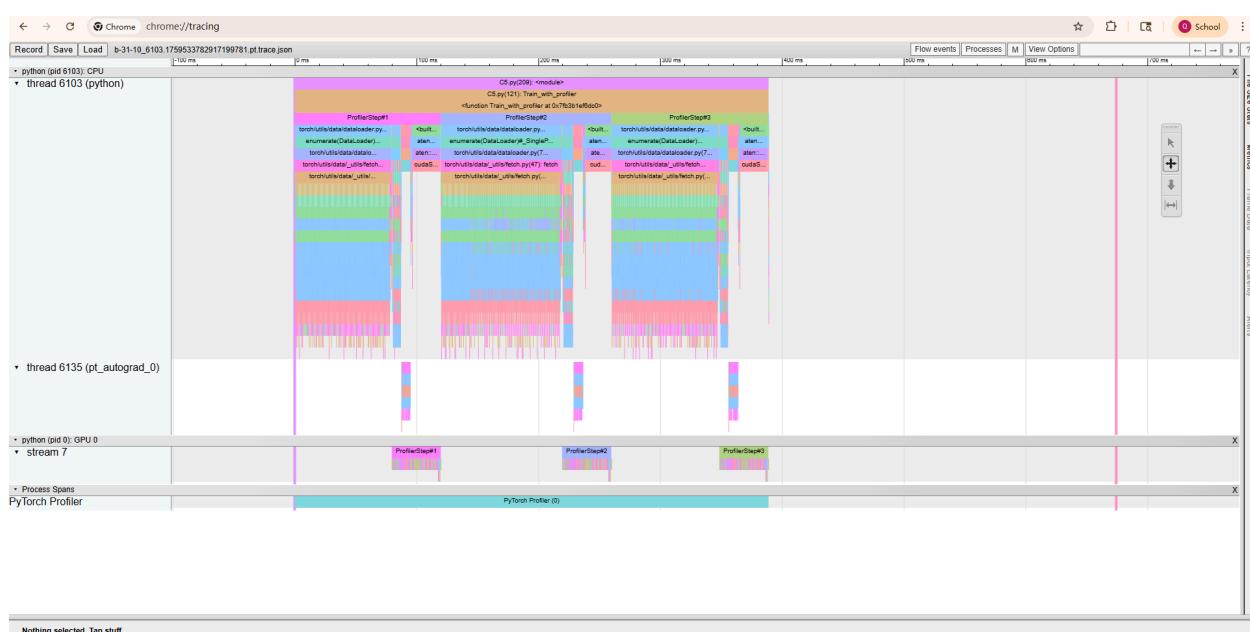
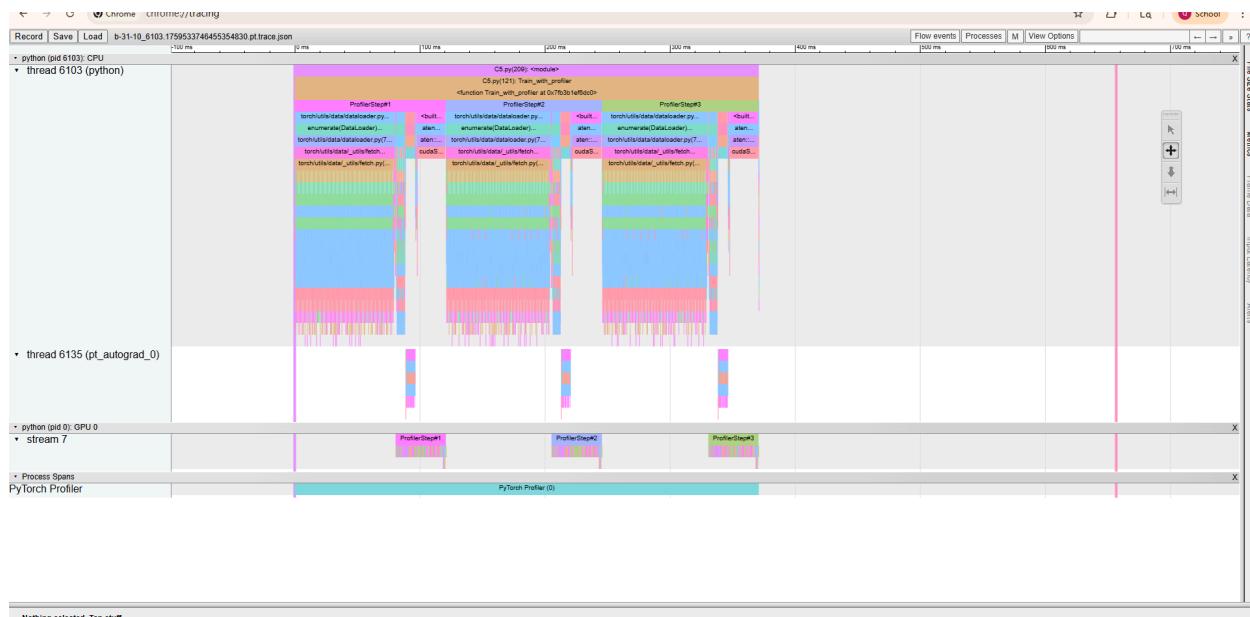
Epoch: 3/5
Epoch [3/5] - Loss: 1.0675, Accuracy: 61.87%
Epoch_time: 36.79201445

Epoch: 4/5
Epoch [4/5] - Loss: 0.8681, Accuracy: 69.37%
Epoch_time: 36.43635905

Epoch: 5/5
Epoch [5/5] - Loss: 0.7304, Accuracy: 74.68%
Epoch_time: 36.16543873
The device is cuda: average running time over 5 epochs is : 36.60580432500002
[gb22620b-31-10 ECE9143]$
```



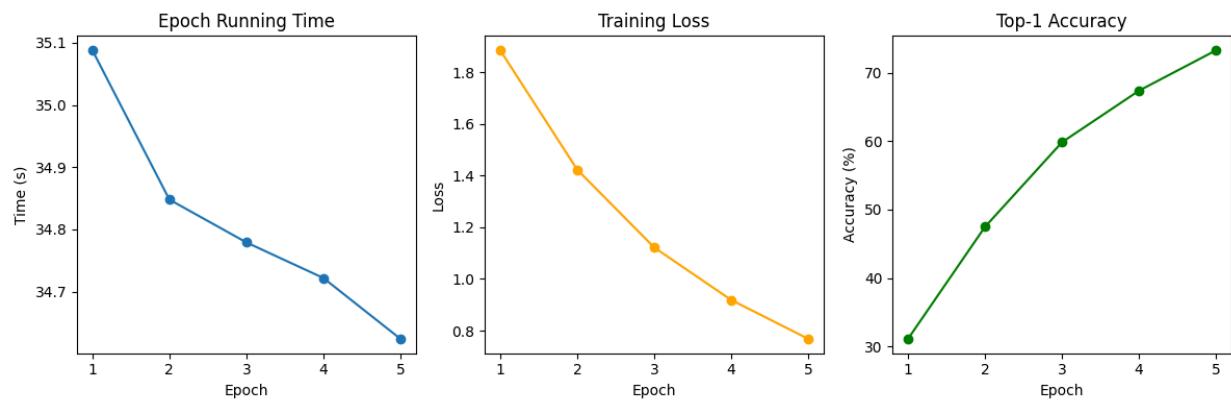




C6

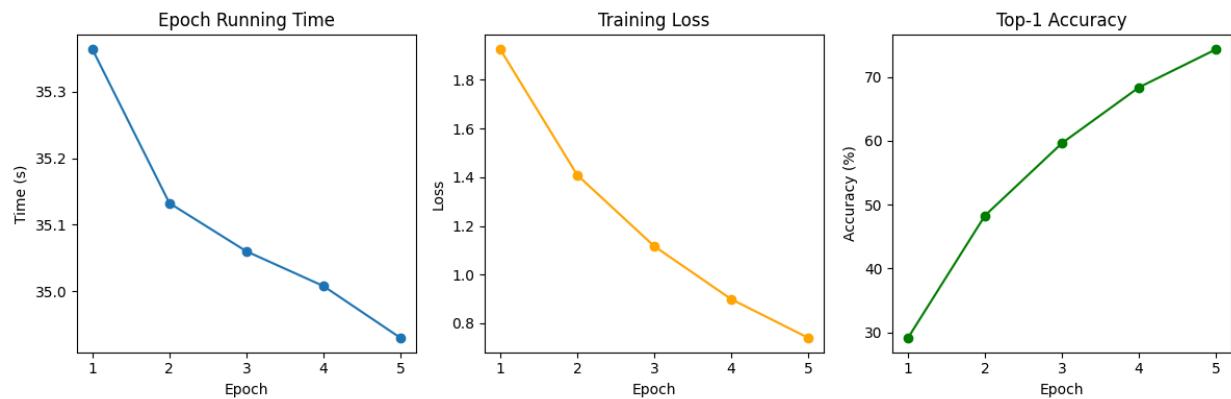
```
python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd  
--epochs 5
```

```
Epoch: 5/5  
Epoch [5/5] - Loss: 0.6703, Accuracy: 76.83%  
Epoch_time: 34.10473049  
The device is cuda: average running time over 5 epochs is : 34.98261504359999  
[qh2262@b-31-1 ECE9143]$ python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5  
  
Epoch: 1/5  
:Average training time 35.08765433  
Epoch [1/5] - Loss: 1.8844, Accuracy: 31.13%  
  
Epoch: 2/5  
:Average training time 34.84793495  
Epoch [2/5] - Loss: 1.4228, Accuracy: 47.49%  
  
Epoch: 3/5  
:Average training time 34.77917911  
Epoch [3/5] - Loss: 1.1218, Accuracy: 59.85%  
  
Epoch: 4/5  
:Average training time 34.72204001  
Epoch [4/5] - Loss: 0.9186, Accuracy: 67.38%  
  
Epoch: 5/5  
:Average training time 34.62437757  
Epoch [5/5] - Loss: 0.7683, Accuracy: 73.27%  
[qh2262@b-31-1 ECE9143]$
```



```
python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer SGD  
--epochs 5
```

```
Epoch: 5/5  
:Average training time 34.62437757  
Epoch [5/5] - Loss: 0.7683, Accuracy: 73.27%  
[qh2262@b-31-1 ECE9143]$ python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer SGD --epochs 5  
  
Epoch: 1/5  
:Average training time 35.36374897  
Epoch [1/5] - Loss: 1.9253, Accuracy: 29.15%  
  
Epoch: 2/5  
:Average training time 35.13252012  
Epoch [2/5] - Loss: 1.4092, Accuracy: 48.31%  
  
Epoch: 3/5  
:Average training time 35.06008070  
Epoch [3/5] - Loss: 1.1168, Accuracy: 59.65%  
  
Epoch: 4/5  
:Average training time 35.00789038  
Epoch [4/5] - Loss: 0.8991, Accuracy: 68.39%  
  
Epoch: 5/5  
:Average training time 34.93000874  
Epoch [5/5] - Loss: 0.7404, Accuracy: 74.34%  
[qh2262@b-31-1 ECE9143]$
```



```
python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer Adagrad  
--epochs 5
```

```

Epoch [4/5] - Loss: 0.8991, Accuracy: 68.39%
Epoch: 5/5
:Average training time 34.93000874
Epoch [5/5] - Loss: 0.7404, Accuracy: 74.34%
[qh2262@b-31-1 ECE9143]$ python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer Adagrad --epochs 5

Epoch: 1/5
:Average training time 35.58155253
Epoch [1/5] - Loss: 4.6007, Accuracy: 23.01%

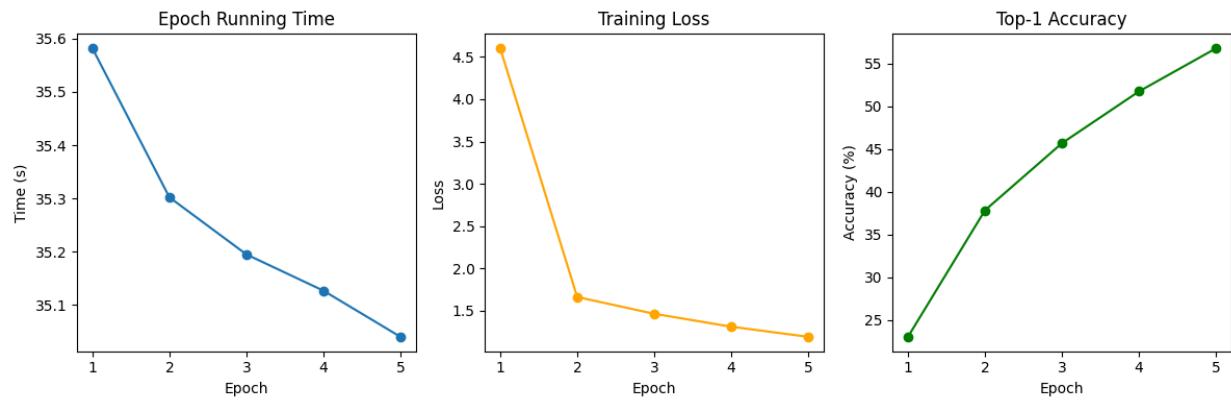
Epoch: 2/5
:Average training time 35.30209367
Epoch [2/5] - Loss: 1.6618, Accuracy: 37.82%

Epoch: 3/5
:Average training time 35.19503653
Epoch [3/5] - Loss: 1.4637, Accuracy: 45.67%

Epoch: 4/5
:Average training time 35.12701653
Epoch [4/5] - Loss: 1.3112, Accuracy: 51.73%

Epoch: 5/5
:Average training time 35.04051043
Epoch [5/5] - Loss: 1.1914, Accuracy: 56.77%
[qh2262@b-31-1 ECE9143]$ █

```



```

python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer Adadelta
--epochs 5

```

```

Epoch: 5/5
:Average training time 35.04051043
Epoch [5/5] - Loss: 1.1914, Accuracy: 56.77%
[qh2262@b-31-1 ECE9143]$ python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer Adadelta --epochs 5

Epoch: 1/5
:Average training time 36.12873974
Epoch [1/5] - Loss: 1.4096, Accuracy: 48.63%

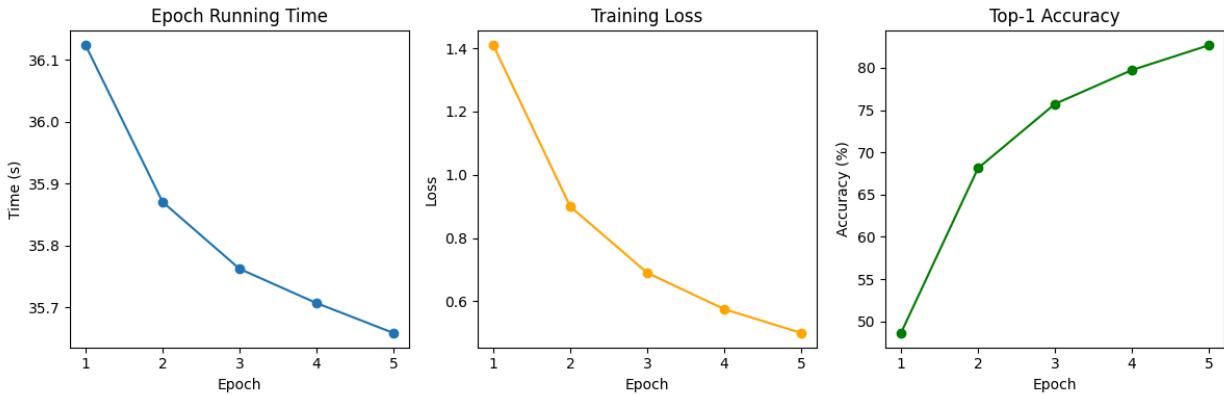
Epoch: 2/5
:Average training time 35.87036273
Epoch [2/5] - Loss: 0.8994, Accuracy: 68.12%

Epoch: 3/5
:Average training time 35.76215889
Epoch [3/5] - Loss: 0.6901, Accuracy: 75.72%

Epoch: 4/5
:Average training time 35.70678166
Epoch [4/5] - Loss: 0.5755, Accuracy: 79.73%

Epoch: 5/5
:Average training time 35.65844420
Epoch [5/5] - Loss: 0.4999, Accuracy: 82.67%
[qh2262@b-31-1 ECE9143]$ python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer Adam --epochs 5

```



```
python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer Adam --epochs 5
```

```
Epoch: 5/5
:Average training time 35.65844420
Epoch [5/5] - Loss: 0.4999, Accuracy: 82.67%
[qh2262@b-31-1 ECE9143]$ python C6.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer Adam --epochs 5

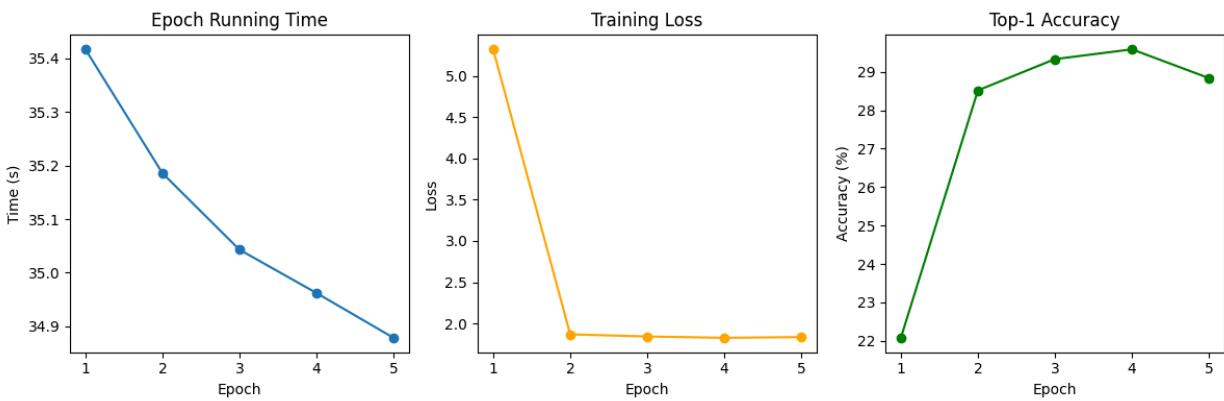
Epoch: 1/5
:Average training time 35.41665777
Epoch [1/5] - Loss: 5.3207, Accuracy: 22.08%

Epoch: 2/5
:Average training time 35.18563096
Epoch [2/5] - Loss: 1.8693, Accuracy: 28.52%

Epoch: 3/5
:Average training time 35.04297732
Epoch [3/5] - Loss: 1.8433, Accuracy: 29.33%

Epoch: 4/5
:Average training time 34.96219706
Epoch [4/5] - Loss: 1.8271, Accuracy: 29.58%

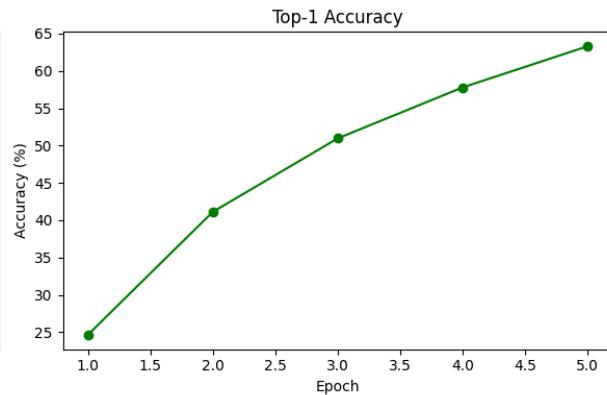
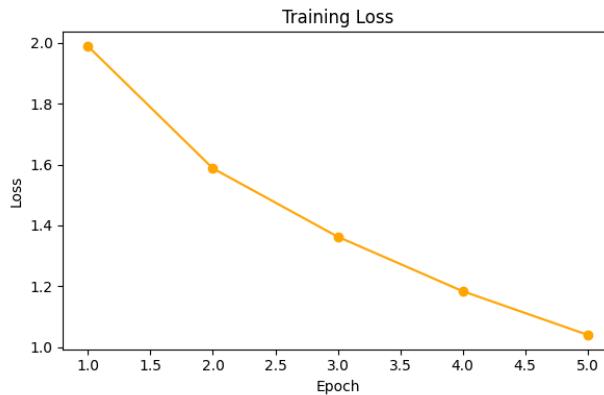
Epoch: 5/5
:Average training time 34.87865344
Epoch [5/5] - Loss: 1.8359, Accuracy: 28.84%
[qh2262@b-31-1 ECE9143]$ ]
```



C7

```
python C7.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd  
--epochs 5
```

```
Host: b-31-10  
Authentication successful  
My Interactive Sessions - NYU HPC  
Dashboard - NYU HPC  
qh2262@b-31-10:~/ECE9143  
UmeTrack/lib/models/backbone_ GitHub - xinghaochen/awesome  
https://ood-burst-001.hpc.nyu.edu/puri/sys/shell/ssh/b-31-10  
Host: b-31-10  
Epoch: 2/5 - Loss: 1.8803, Accuracy: 27.99%  
Epoch: 3/5 - Average training time 34.72241353  
Epoch: 3/5 - Loss: 1.8340, Accuracy: 29.56%  
Epoch: 4/5 - Average training time 34.57610683  
Epoch: 4/5 - Loss: 1.8145, Accuracy: 29.98%  
Epoch: 5/5 - Average training time 34.49165210  
Epoch: 5/5 - Loss: 1.8151, Accuracy: 29.81%  
[qh2262@b-31-10 ECE9143]$ python C7.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5  
Epoch: 1/5  
Epoch: 1/5 - Loss: 1.8952, Accuracy: 28.40%  
Epoch: 2/5  
Epoch: 2/5 - Loss: 1.5169, Accuracy: 44.01%  
Epoch: 3/5  
Epoch: 3/5 - Loss: 1.3952, Accuracy: 52.82%  
Epoch: 4/5  
Epoch: 4/5 - Loss: 1.1676, Accuracy: 58.50%  
Epoch: 5/5  
Epoch: 5/5 - Loss: 1.0137, Accuracy: 64.31%  
Traceback (most recent call last):  
  File "/home/qh2262/.local/lib/python3.9/site-packages/matplotlib/pyplot.py", line 612, in show  
    return _get_backend_mod().show(*args, **kwargs)  
TypeError: show() takes 1 positional argument but 2 were given  
[qh2262@b-31-10 ECE9143]$ vim C7.py  
[qh2262@b-31-10 ECE9143]$ python C7.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd --epochs 5  
Epoch: 1/5  
Epoch: 1/5 - Loss: 1.9882, Accuracy: 24.64%  
Epoch: 2/5  
Epoch: 2/5 - Loss: 1.5874, Accuracy: 41.11%  
Epoch: 3/5  
Epoch: 3/5 - Loss: 1.3630, Accuracy: 50.95%  
Epoch: 4/5  
Epoch: 4/5 - Loss: 1.1841, Accuracy: 57.78%  
Epoch: 5/5  
Epoch: 5/5 - Loss: 1.0409, Accuracy: 63.29%  
[qh2262@b-31-10 ECE9143]$
```



Q1

The total is 18 layers of convolution neural networks.

Layer 1: 1 initial convolution

Layer 2: 2 convolution

Layer3: 2 convolution

Layer4: 2 convolution

Layer5: 2 convolution
Layer6: 2 convolution
Layer7: 2 convolution
Layer8: 2 convolution
Layer9: 2 convolution

Total is $8*2+1=17$ convolutions if we count Fc too, then will have 18 convolutions which is reiset18. If we count shortcuts as well, $3+17=20$ convolutions. With last Fc, the total is 21 convolutions.

Q2

512, as we put 512,H,W into `AdaptiveAvgPool2d`. The final shape will be [B,512,1,1]. If we reshape and broadcast the last two columns it will be [B,512]. Then enter into Fc, so the final input size is 512.

Q3

```
python Q3.py --cuda --data_path ./data --num_workers 0 --batch_size 128 --optimizer sgd  
--epochs 5
```

The final parameter is 11172170. The optimizer does not change the size of trainable parameters, so the parameters in sqd still are 11172170.

Self calculation:

init conv=3*64*3*3=1728+128=1856

```

First_group:= 64*64*3*3+64*64*3*3*3+128*4=147968
Second_group:= 64*128*3*3+128*128*3*3*3+256*4+64*128*1*1=525312
Third_group:= 128*256*3*3+256*256*3*3*3+512*4+128*256*1*1=2099200
Fourth_group:= 256*512*3*3+512*512*3*3*3+1024*4+256*512*1*1=8392704
Fc:= 512*10+10=5130
sum= 147968+1856+525312+2099200+8392704+5130=11172170

```

Explain: The layer module for `resn_block` is to generate two convolution layers whose first convolution has different input dimension and out dimension if we set a different size of dimension. The second layer is fixed to output size for both `in_channel` and `out_channel`. We use padding to set the shape of the image not change. The group would has fixed pattern is $C_{in} * C_{out} * Kernel^2 + C_{out} * C_{out} * kernel^2 * 3$ for the 2 layers in one subgroup. So except the first group has the same `C_in` and `C_out`, which has no shortcut convolution cost. Other layers also add the convolution cost of $1*1$ kernel to reshape the residual($+ C_{in} * C_{out} * 1 * 1$). The init convolution only cover $3*64*kernel^2$ and batch normalized in $2*64$. The output layers is `Fc` which is 512 input size * output size as 10 classification. So, total is 11172170 which is equal to what the code given.

The trainable parameters are the same in sgd and our model, as we set the neural network to count all the trainable parameters. We use model parameters in sgd so size is the same.

```
optim=torch.optim.SGD(model.parameters(), lr=0.1, momentum=0.9, weight_decay=5e-4)
```

Q4

```

init_conv=3*64*3*3=1728+128=1856
First_group:= 64*64*3*3+64*64*3*3*3+128*4=147968
Second_group:= 64*128*3*3+128*128*3*3*3+256*4+64*128*1*1=525312
Third_group:= 128*256*3*3+256*256*3*3*3+512*4+128*256*1*1=2099200
Fourth_group:= 256*512*3*3+512*512*3*3*3+1024*4+256*512*1*1=8392704
Fc:= 512*10+10=5130
sum= 147968+1856+525312+2099200+8392704+5130=11172170

```

No matter if we use adam or sgd, and other optimizer methods, we all get the same trainable parameters.