
HW3 - ASR

Team 5

B09901116 陳守仁、B09602017 白宗民、B08901207 呂俐君

Outline

- **Basic Experiments**
 - Vox + Large Clean
 - Vox + Large Noisy
 - Wiki3 + Large Clean
- **Experiment of Different Batch Size**

Basic Experiments

Basic Experiments

- Vox + large clean

```
PER: 33.38290171431778 %  
DEL RATE: 9.368985157543278 %  
INS RATE: 12.344771471058612 %  
SUB RATE: 11.669145085715888 %  
1842
```

- Vox + large noisy

```
PER: 29.000869784798404 %  
DEL RATE: 8.073847535142113 %  
INS RATE: 9.420049942481972 %  
SUB RATE: 11.506972307174323 %  
1842
```

- wiki3 + large clean

```
PER: 48.541847871833 %  
DEL RATE: 8.235459162200836 %  
INS RATE: 16.977638113408712 %  
SUB RATE: 23.32875059622345 %  
1842
```

- **PER:** Phone Error Rate, which represents the percentage of phoneme-level errors in the ASR system's output relative to the reference transcription.
- **DEL RATE:** Deletion Rate, which represents the percentage of phonemes in the reference transcription that were not correctly recognized by the ASR system and were omitted.
- **INS RATE:** Insertion Rate, which represents the percentage of additional phonemes in the ASR output that were not present in the reference transcription.
- **SUB RATE:** Substitution Rate, which represents the percentage of phonemes in the ASR output that were different from the phonemes in the reference transcription.

Experiment of Different Batch Size

Different Batch Size

- Batch size = 100

```
PER: 31.329087287113154 %  
DEL RATE: 7.750063129541821 %  
INS RATE: 12.290339777222862 %  
SUB RATE: 11.288684380348474 %  
1842
```

- Batch size = 160

```
PER: 29.000869784798404 %  
DEL RATE: 8.073847535142113 %  
INS RATE: 9.420049942481972 %  
SUB RATE: 11.506972307174323 %  
1842
```

- Batch size = 256

```
PER: 96.01526332033332 %  
DEL RATE: 3.685642939311467 %  
INS RATE: 14.05740579669482 %  
SUB RATE: 78.27221458432703 %  
1842
```

- **large_noisy + voxpopuli_trans**
- Epoch \approx 3000
- Performance of batch size = 256 **sucks** \rightarrow still not find out the answer
- Performance of batch size = 100 is **similar** with original one(160)
- Time of batch size = 256 is almost two times longer than original
- **Conclusion:** Find out the best batch size need some expertise or trails

Different Batch Size

- The error message appear in the trail of batch size = 256

→ might be the reason of the poor performance

```
[2023-10-09 04:56:08,802][unsupervised.tasks.unpaired_audio_text][INFO] - split valid has unpaired text? False
[2023-10-09 04:56:08,810][fairseq.utils][INFO] - *****CUDA environments for all 1 workers*****
[2023-10-09 04:56:08,810][fairseq.utils][INFO] - rank 0: capabilities = 7.0 ; total memory = 31.749 GB ; name = Tesla V100-SXM2-32GB
[2023-10-09 04:56:08,810][fairseq.utils][INFO] - *****CUDA environments for all 1 workers*****
[2023-10-09 04:56:08,810][fairseq_cli.train][INFO] - training on 1 devices (GPUs/TPUs)
[2023-10-09 04:56:08,810][fairseq_cli.train][INFO] - max tokens per device = None and max sentences per device = 100
[2023-10-09 04:56:08,811][fairseq.trainer][INFO] - Preparing to load checkpoint ./checkpoint_last.pt
[2023-10-09 04:56:08,811][fairseq.trainer][INFO] - No existing checkpoint found ./checkpoint_last.pt
[2023-10-09 04:56:08,811][fairseq.trainer][INFO] - loading train data for epoch 1
[2023-10-09 04:56:08,941][unsupervised.data.extracted_features_dataset][INFO] - loaded 5920, skipped 0 samples
[2023-10-09 04:56:08,942][unsupervised.tasks.unpaired_audio_text][INFO] - split train has unpaired text? True
[2023-10-09 04:56:09,040][fairseq.data.data_utils][INFO] - loaded 155,341 examples from: /home/u9655801/hw/hw3/u-s2p/text/voxpophli_trans/phones/train
[2023-10-09 04:56:09,127][fairseq.trainer][INFO] - NOTE: your device may support faster training with --fp16 or --amp
[2023-10-09 04:56:09,225][fairseq.trainer][INFO] - begin training epoch 1
[2023-10-09 04:56:09,225][fairseq_cli.train][INFO] - Start iterating over samples
[2023-10-09 04:56:29,429][fairseq_cli.train][INFO] - end of epoch 1 (average epoch stats below)
Exception in thread Thread-2:
Traceback (most recent call last):
  File "/home/u9655801/miniconda3/envs/hw3_ASR/lib/python3.8/threading.py", line 932, in _bootstrap_inner
    self.run()
  File "/home/u9655801/miniconda3/envs/hw3_ASR/lib/python3.8/site-packages/tensorboard/summary/writer/event_file_writer.py", line 233, in run
    self._record_writer.write(data)
  File "/home/u9655801/miniconda3/envs/hw3_ASR/lib/python3.8/site-packages/tensorboard/summary/writer/record_writer.py", line 40, in write
    self._writer.write(header + header_crc + data + footer_crc)
  File "/home/u9655801/miniconda3/envs/hw3_ASR/lib/python3.8/site-packages/tensorboard/compat/tensorflow_stub/io/gfile.py", line 766, in write
    self.fs.append(self.filename, file_content, self.binary_mode)
  File "/home/u9655801/miniconda3/envs/hw3_ASR/lib/python3.8/site-packages/tensorboard/compat/tensorflow_stub/io/gfile.py", line 160, in append
    self._write(filename, file_content, "ab" if binary_mode else "a")
  File "/home/u9655801/miniconda3/envs/hw3_ASR/lib/python3.8/site-packages/tensorboard/compat/tensorflow_stub/io/gfile.py", line 164, in _write
    with io.open(filename, mode, encoding=encoding) as f:
FileNotFoundError: [Errno 2] No such file or directory: 'b'tb/train/events.out.tfevents.1696751235.cjnz7yctr1696744979977-stz7t.2177.0'
```

Thanks!