A ZOOM FILTER FOR APPLAUSE AND LAUGHTER

Meeting 09.02.22

Where we left off

- Only 30 batches in 2 hours
  - *Which equals 30*32 = 960 segments*
    - 8s audio per minute

- Create dataloader using lhotse
- Is that going to speed up the training enough?

Today

1. **FYI:** No project renaming
2. New dataloader is fast enough
3. New evaluation -> good recall, precision bad
    1. *Choose feature representation (Fbank vs Spectrogram)*
    2. *Train on mixed snippets (why didn't they do that?)*
        - change in logic
4. Realtime prediction
    - change in logic
5. Contribution to lhotse – mention in thesis?
6. Can't do all of that **and** train on Mobilenet

New Evaluation

- On 2 meetings

threshold	precision		recall		valid_pred_laughs	
	mean	median	mean	median	mean	median
0.1	0.009289	0.009289	0.926635	0.926635	2342.0	2342.0
0.2	0.018489	0.018489	0.917084	0.917084	1897.0	1897.0
0.3	0.033927	0.033927	0.897891	0.897891	1103.0	1103.0
0.4	0.057323	0.057323	0.860460	0.860460	623.0	623.0
0.5	0.089995	0.089995	0.776256	0.776256	344.0	344.0
0.6	0.136735	0.136735	0.649098	0.649098	202.5	202.5
0.7	0.199281	0.199281	0.477080	0.477080	112.5	112.5
0.8	0.275497	0.275497	0.346883	0.346883	59.0	59.0
0.9	0.373416	0.373416	0.176774	0.176774	26.0	26.0

Today

1. **FYI:** No project renaming
2. New dataloader is fast enough
3. New evaluation -> good recall, precision bad
    1. *Choose feature representation (Fbank vs Spectrogram)*
    2. *Train on mixed snippets (why didn't they do that?)*
        - change in logic
4. Realtime prediction
    - change in logic
5. Contribution to lhotse – mention in thesis?
6. Can't do all of that **and** train on Mobilenet

New method performance

- Only 30 batches in 2 hours
  - *Which equals 30*32 = 960 segments [1s each]*
    - 8s audio per minute

- 523 batches in ~2minutes
  - *Which equals 523*32 = 16 736 segments [1s each]*
    - 139s audio per minute = 2.3s audio per second

GPU Utilisation

- 50% utilisation
  - *But sufficient speed*
- When I open an interactive session, it doesn't show the usage
- Is that important for my thesis?