Meeting 5: Corti

**Time and data:** 05/04/2021 (16.00-16.30)

**Participants:** All group members

**Questions:**

1. Vi har problemer at vi ikke kan generere data hurtigt nok. Et eksempel er at vi ikke kan sammenligne med den pre-trained ASR model (trænet på 960 timer). Vi kan generere 1 minut tale pr. sekund (kan måske gøres lidt bedre). Vi har en upper-limit på 24 trænings timer.
2. Hvad er din holdning til vores 24 timers upper-limit på trænings timer. Lige nu skal vi træne rigtig mange modeller. Det er urealistisk at skulle træne mere end 24 timer når HPC har den limit som den har.
3. Vi skal have hjælp til at vælge hyper-parametre (antal lag, learning rate, kernel size og alt muligt andet) og generelt idéer til hvordan vi kan forbedre modellen.
4. Learning rate der overskriver BERT?
5. Sammenligning med tacotron 2 mel-specs som grundlag for vores rapport.

**Opfølgende spørgsmål:**

* Når vi skal træne ASR modellen (baseline) ville det så ikke give mere mening at træne på LJSpeech datasættet for sammenlignelighed (Tacotron 2 og WaveNet er trænet på dette datasæt og ikke et uddrag af LibriSpeech).

Idéer:

2 learning rates

Seed?

Tacotron mel-spectogram

Mixed-precision

CNU-ARCTIC