Feedback Project Work

Feedback group: Generating Speech from Transcripts (group 2)

6th of April 2022

Feedback

Almost all of the introduction is very on-point, it provides background knowledge and a scope that gives a very polished outlook and really brings us in depth with what the project is about. It is commendable to see how open you guys have been with what is superficial and the only thing we had to google to understand from the introduction was the definition of HRS and LRS. Much of what has been covered though, is very in-depth.
It is a bit unclear as to how data in a HRL turns into data in a LRL(2. page of introduction) - 'combined' is a bit of a vague word to use; Do we reuse pre-trained weights or is some other technique implemented? Another good pointer that could be recommended would be to avoid using 'we' as much as possible in the paper to keep it more professional. Other than that, the idea of creating an acronym section was very thoughtful and the use of the references was very well placed.
In the method section, it was a really good idea to implement prerequisites and pipelines before going through all the heavy reading models/material. Very fleshed out segment for the mathematical operations, from input to output of WaveNet. The procedure is exceptionally well written/creative that it is hard to miss out on a step and the approach follows an engineering perspective.

Constructive Questions

1. Have you checked up on the licenses for all of your imported figures?(DTU normally very strict about it) You could go as far to attach the license type in your references/caption.
2. Have you looked into if there are cases(e.g. competitions) where your model structure (e.g. Tacotron2) has performed well for solving similar problems? Maybe
3. What was the decision to choosing LibriSpeech to evaluate TTS compared to other datasets?
4. When you trained your own ASR-model, did you train it with randomly initalized weights or pretrained?(It is a bit unclear when reading and this important in regard to your 3. research question).
5. How is it that data in a HRL turns into data in a LRL? (This needs to clarified better in the introduction. A figure at the end would've done wonders)