# Week 3 Meeting

2357351G - MSci Half Project

What I've done this week

- Summarised *Bridging Audio Analysis, Perception and Synthesis with Perceptually-Regulated Variational Timbre Spaces (2018)*.
- Simplified VAE training to try to debug issues discussed last week.
- Performed a longer training run without using the AudioCommons feature extractor.
- Created an initial (very high level) model evaluation plan.
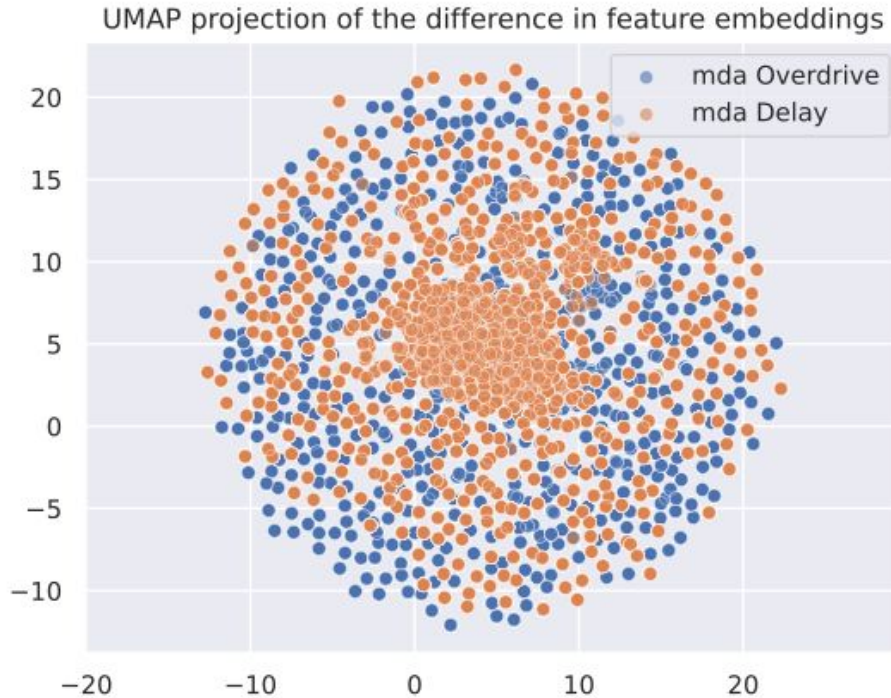- Began working on final report.
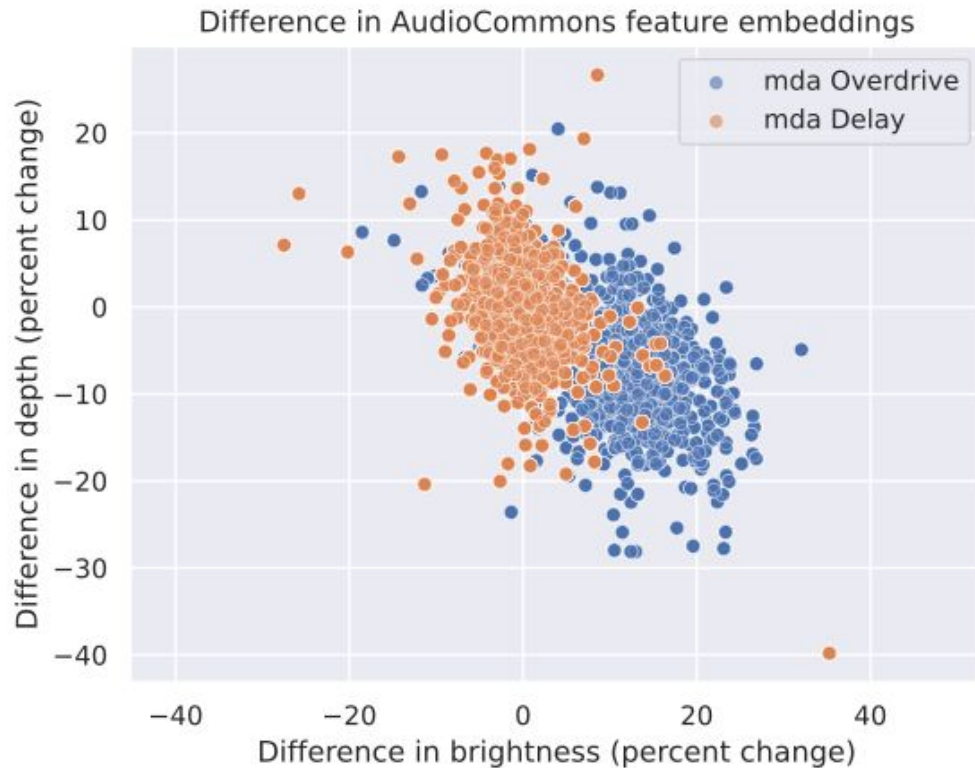
Projection of clean feature embeddings

UMAP projection of the clean feature embeddings
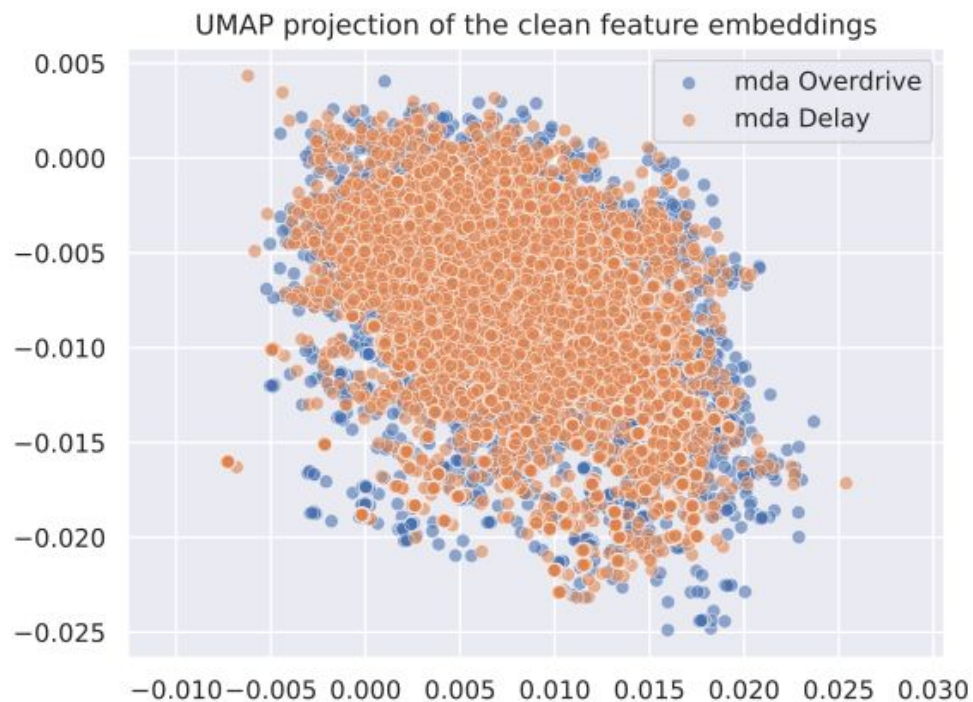
Projection of effected feature embeddings

UMAP projection of the effected feature embeddings
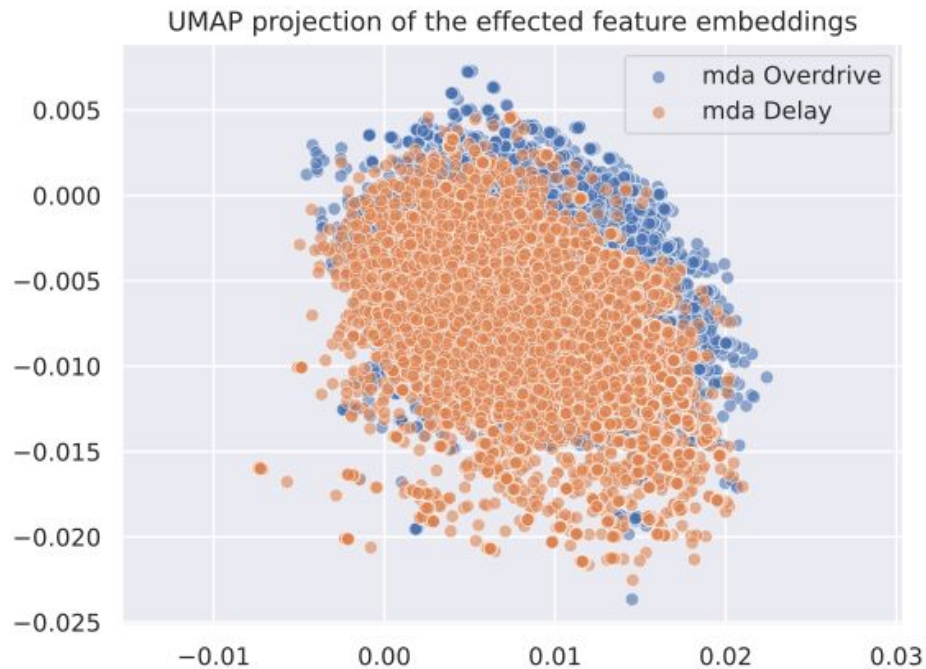
Projection of difference in feature embeddings

Difference in clean/effected AC features

Difference in AudioCommons feature embeddings

Projection of learned latent space

UMAP projection of the learned embeddings (500 datapoints per effect)

Questions

- Any feedback on the initial evaluation plan would be appreciated! Anything you think I should add? Or if I should compare my approach against more than just the DeepAFX-ST model (which my model is based on)?
- The relatively poor reconstruction loss and lack of structure in both the audio embeddings and learned latent embeddings seems to be an issue. Do you have any guidance or advice for further debugging that could be performed?

Plan for next week

- Further debugging of VAE.
- An even longer training run (without AudioCommons features).
- Continue to work on sections of the final report which I can write at the moment.

Where I am in schedule

- Slightly concerned at the lack of structure in both latent space and embeddings of audio - worried that there isn't a lot of time remaining in the project considering end-to-end training needs to be performed and experiments conducted.
- Writing out the evaluation plan has been useful for solidifying the scope and aims of the project.