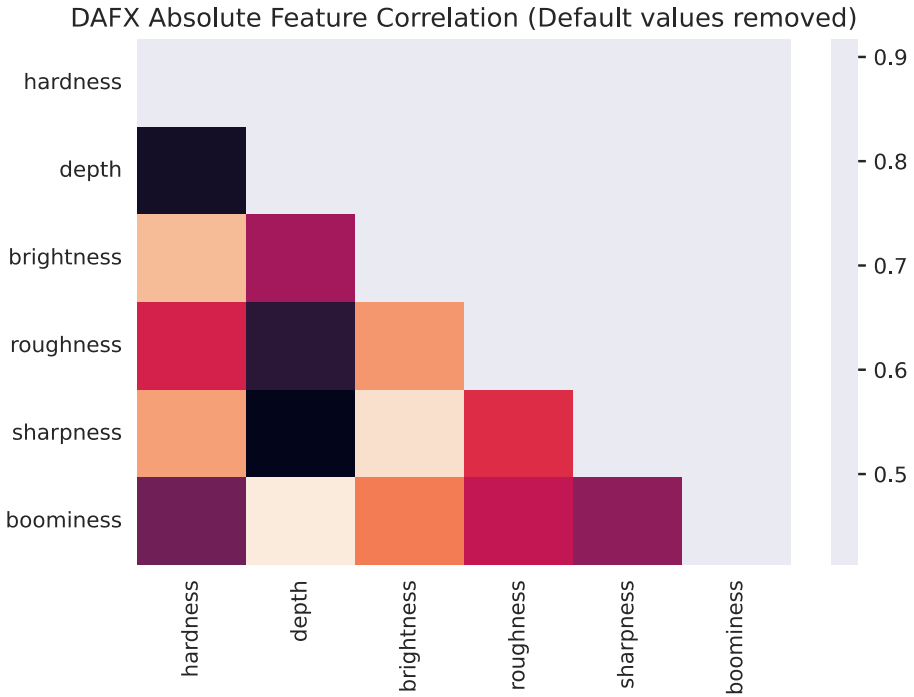
Week 1 Meeting

2357351G - MSci Half Project

Work done in final week of semester 1

- Submitted interim report.
- Implemented data generation and preprocessing pipeline.
- Added AudioCommons feature extractor to preprocessing pipelines.
    - Only using *brightness* and *depth* features for now for training efficiency.
- Implemented first version of training pipeline for both the β-VAE as well as the end-to-end model.

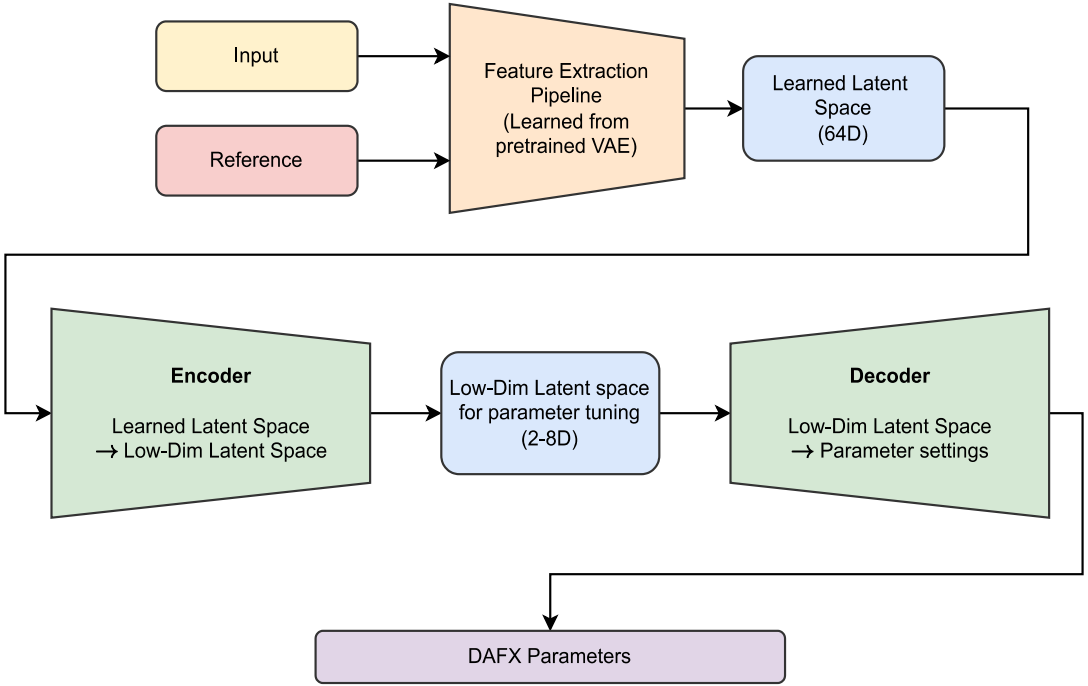
AudioCommons Feature Correlation



Questions

- For training the VAE, I need to use a DAFX for audio generation and I'm thinking about the best way to do this:
  - Either use no DAFX (just scene augmentation) or something very transparent).
  - Use different DAFX at each epoch - not sure if this would make training unstable? Or whether this would learn a richer latent space than using no DAFX?
- Need to decide on how to map from learned VAE latent space down to a low-dimensional latent space, then to parameter settings. An approach is shown in the next slide - not sure if there is an issue with not being able to use the reconstruction loss?

Training end-to-end using learned representation.



Plan for next week

- Decide on audio generation for training VAE.
- Train the VAE and perform some analysis on the learned latent space - tune the latent space dimension.
- Decide on mapping from learnt latent space to parameter settings.

Where I am in schedule

- Interim report complete.
- Made good implementation progress.
- Still a few decisions to be made around model training and mapping learned latent space to DAFX parameters for each new DAFX.