

Neural audio coding for speech enhancement

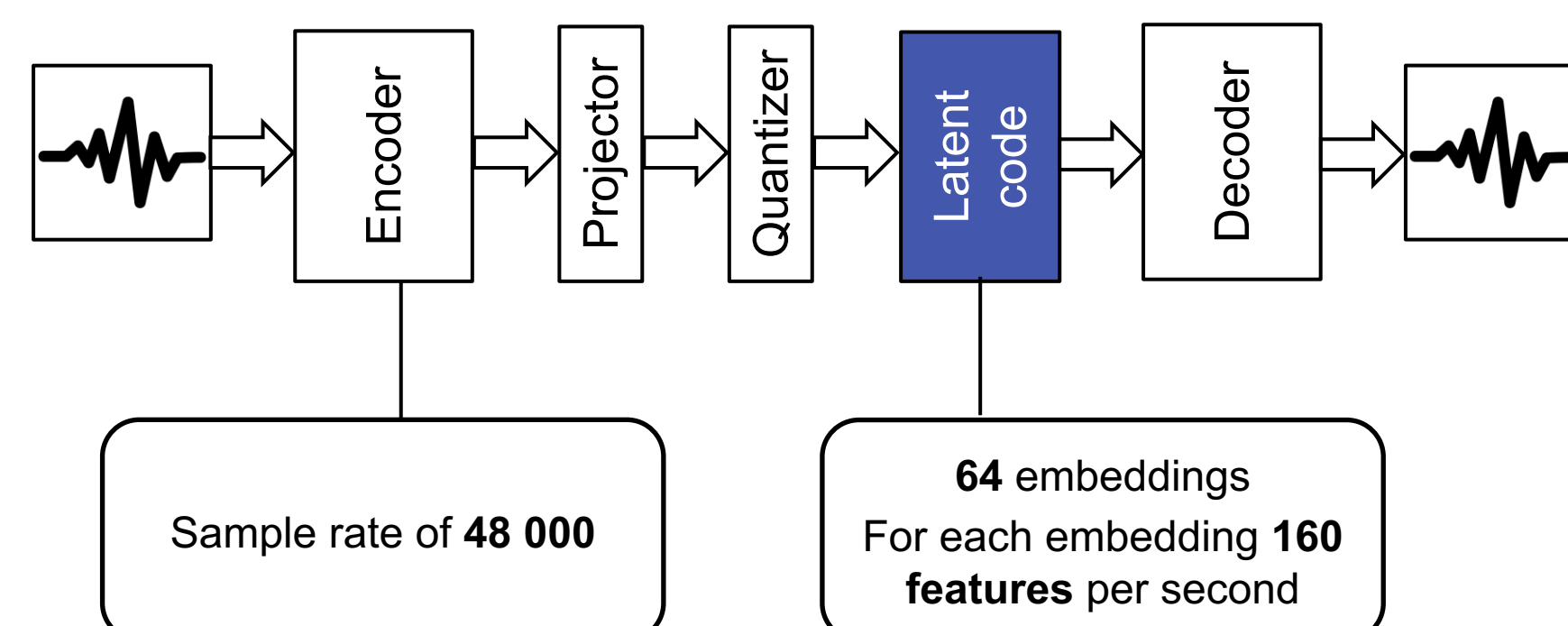
Boldizsár Elek (s233085) | István László Mádi (s232971) | András Bence Schin (s233084)

Supervisor: Kenny Falkær Olsen

PROBLEM

Enhancing speech quality is a fundamental challenge across various applications involving recorded audio, spanning from headphones to hearing aids.. In this project, we investigated whether we can specialize neural audio codecs for the task of denoising audio signals.

AUDIODEC



EXPERIMENTATION

PROS

Fine tuning encoder block of AudioDec

Introducing a network in the latent domain, calculating loss between reconstructed (decoded) audio clips

Introducing a network in the latent domain, calculating loss between latent representations

No need to introduce a new network

Sophisticated loss functions such as SNR, SDR, PESQ (+ SI-SNR, SI-SDR)

Quicker training (especially with pre-encoded data)

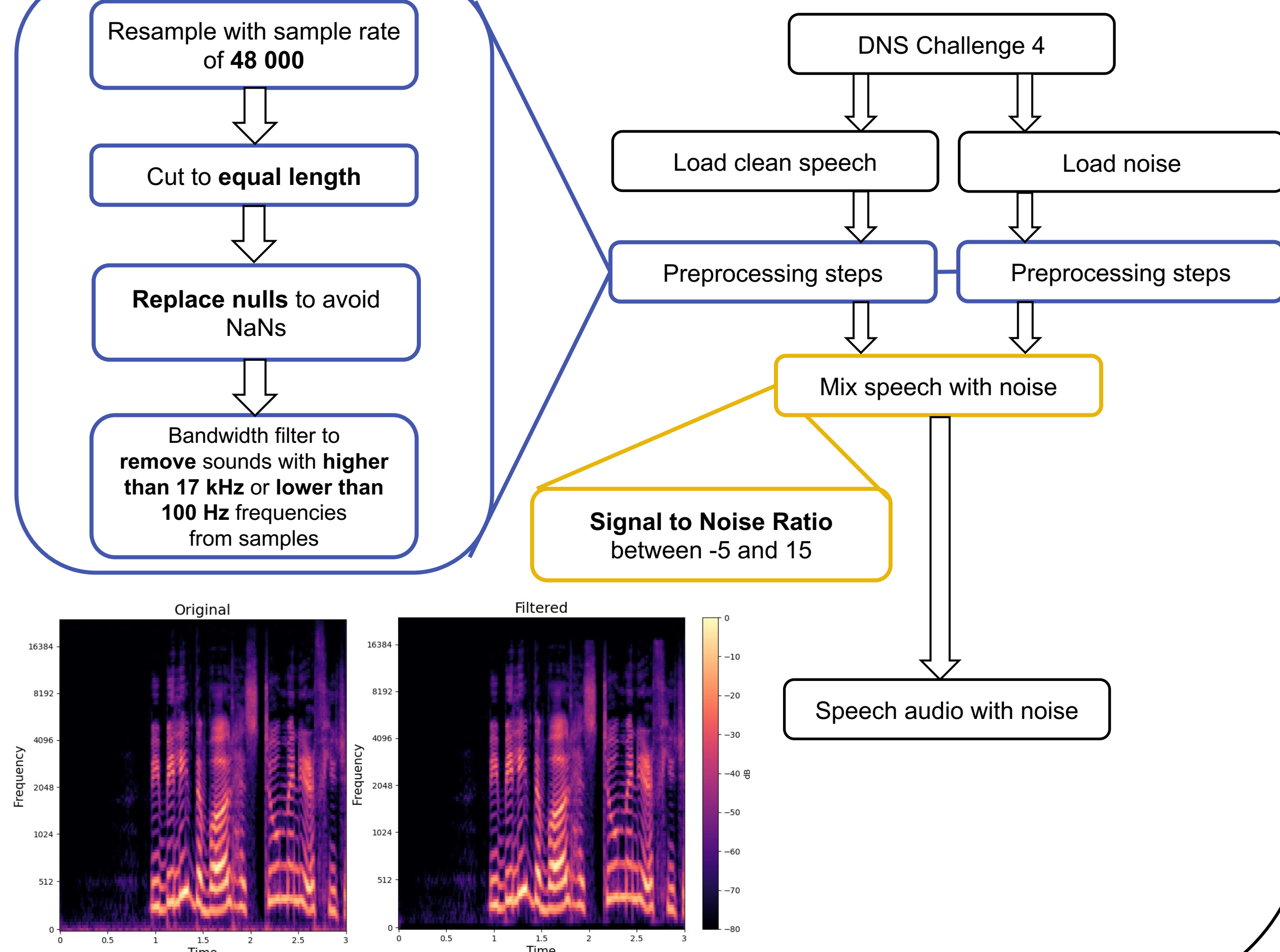
CONS

During our experiments, the audio quality significantly degraded. Computationally expensive

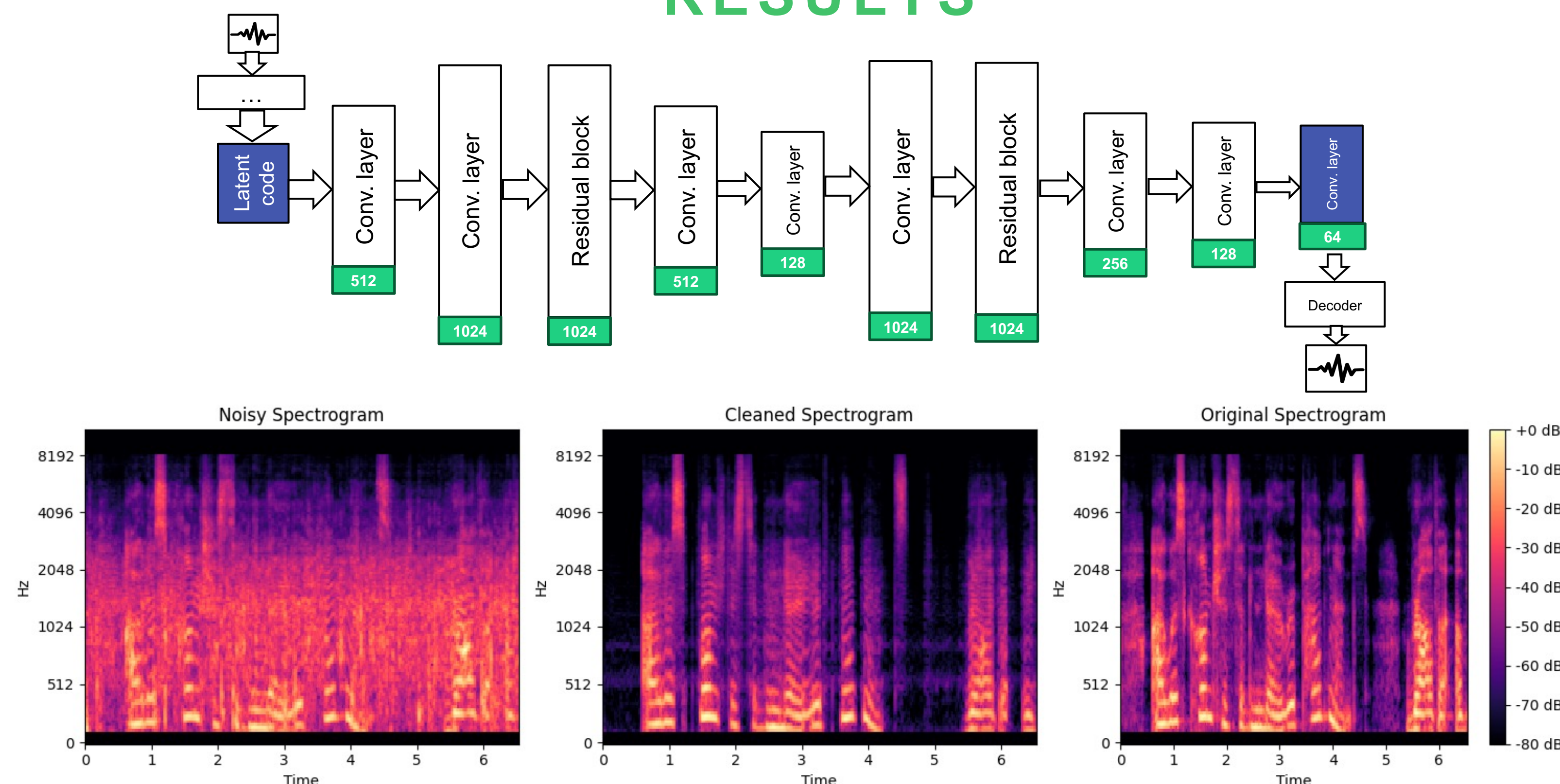
Slower training, as we have to forward call on the decoder. Requires even more computation

Limited choice in loss functions

DATA



RESULTS



References

Y. -C. Wu, I. D. Gebru, D. Marković and A. Richard, "Audiodec: An Open-Source Streaming High-Fidelity Neural Audio Codec," *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSP49357.2023.10096509.

Dubey, H., Aazami, A., Gopal, V., Naderi, B., Braun, S., Cutler, R., Gamper, H., Golestaneh, M., & Aichner, R. (2023). ICASSP 2023 Deep Noise Suppression Challenge. In *ICASSP*.