

VIRTUAL LISTENER PANEL

Niels Asp Fuglsang*, Guillaume Picard, Benjamin Pedersen, Søren Vase Legarth

FORCE Technology

Motivation

Enhancing and quantifying the perceptual quality of audio products typically involves human listener panels to identify key sound characteristics. SenseLab, with its specialized expertise in perceptual audio evaluation, utilizes and maintains an expert panel to gain insights into various audio products. However, conducting listening tests with an expert panel can be both costly and time-consuming. Building on previous work [2, 3], the proposed Virtual Listener Panel (VLP) leverages machine learning to enable manufacturers to compare different sound characteristics across systems within minutes. In contrast, using a human expert panel usually takes days or weeks to complete a test, and maintaining such a panel requires significant effort.

Theory

We implemented a siamese-inspired neural network that learns its own rating scale based on no-reference ratings.

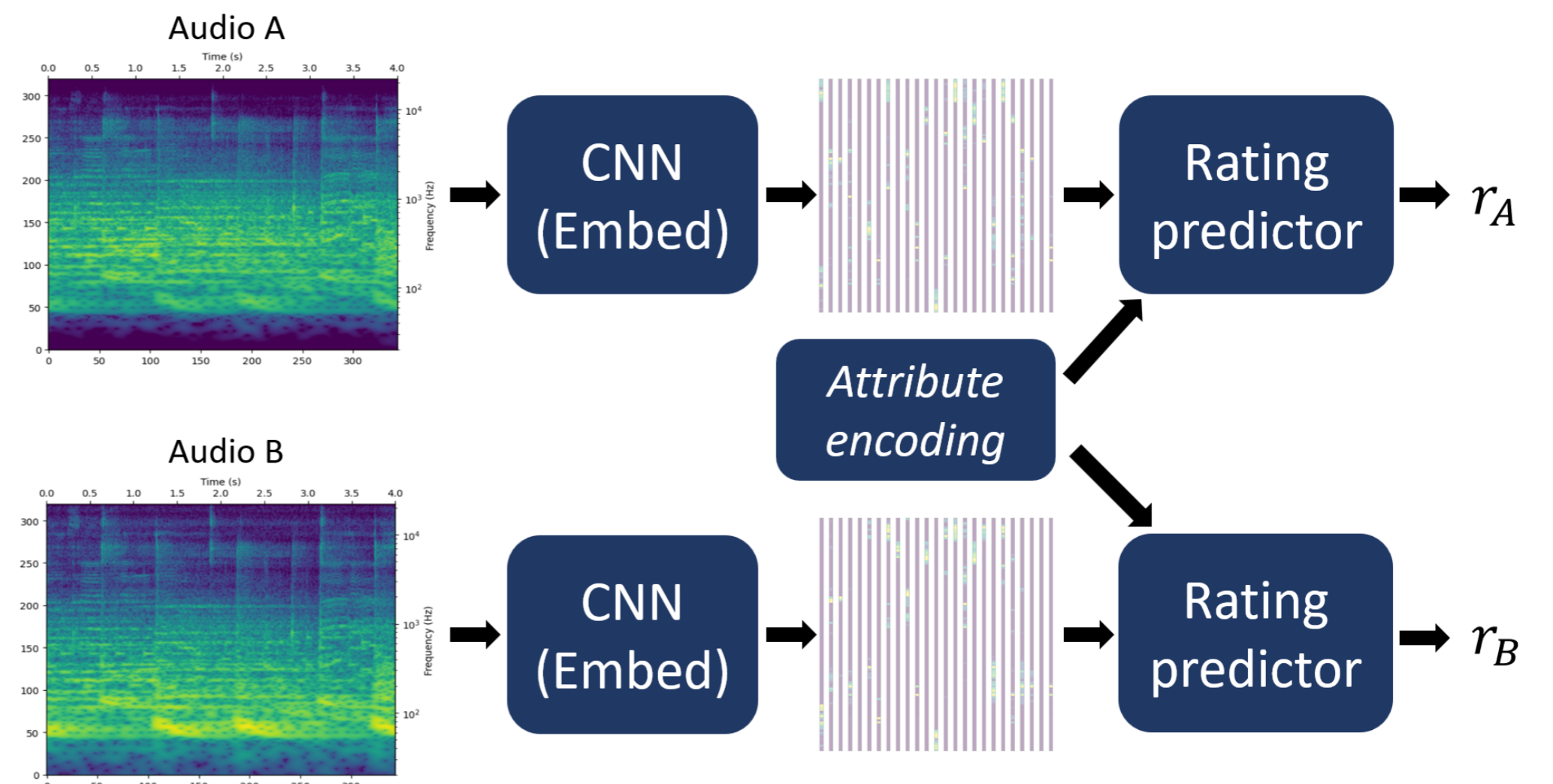


Fig 1: Model architecture inspired by SQAPP[1].

Data

Effective data collection to build a dataset with extensive variation

- Data collected from SenseLab's expert panel, evaluating different music genre samples and pink noise.
- Listening test were all conducted with identical sound cards and with Sennheiser HD650 headphones.
- A total of $\approx 25,000$ ratings comparing random hybrid/augmented systems against each other.
- Transform the data into $\approx 150,000$ pairwise ratings for model training and validation.
- Create a pretraining dataset based on 36 well-known audio features.

| | |
|------------------------|----------------------|
| Compression | None / WAV |
| Sample rate | 48 kHz |
| Loudness normalization | ITU, -20 dB |
| Spectrum type | Variable-Q Transform |

Table 1: Processed Audio Configuration

Experiments and Results

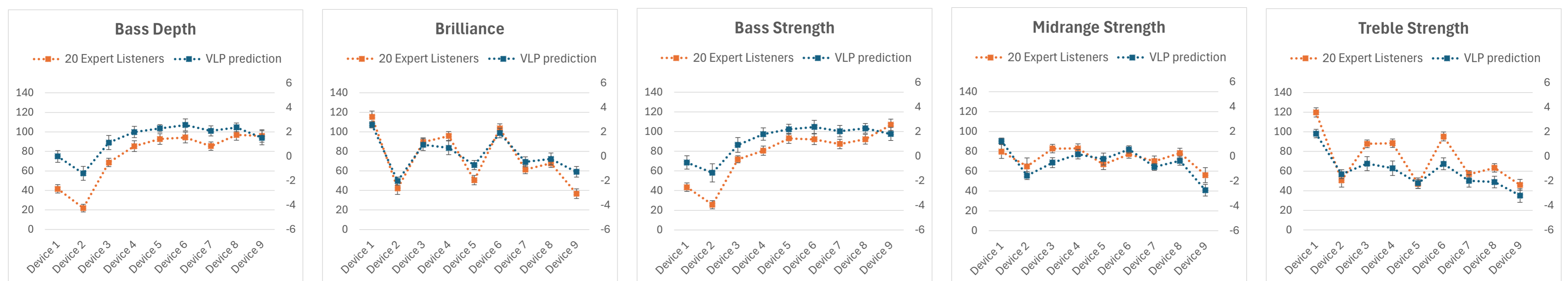


Fig 2: Validating the ratings of real systems of the model trained on hybrid systems.

Conclusion

- The VLP introduces a promising and scalable framework for estimating the perceptual characteristics of sound.
- The VLP can ensure perceptually-informed decisions from the very beginning of the audio product development process.

References

- [1] P. Manocha, Z. Jin, and A. Finkelstein. Sqapp: No-reference speech quality assessment via pairwise preference. In *ICASSP 2022 - 2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2022.
- [2] C. Volk, J. Nordby, T. Stegenborg-Andersen, and N. Zacharov. Efficient data collection pipeline for audio machine learning of audio quality. *Journal of the Audio Engineering Society*, (10488), 2021.
- [3] C. Volk, J. Nordby, T. Stegenborg-Andersen, and N. Zacharov. Predicting audio quality for different assessor types using machine learning. *Journal of the Audio Engineering Society*, (10494), 2021.

