

Bias Correction with Pre-trained Audio Embeddings

We have several openings for **Master students internships** (*Master 2 or Engineer*) at Télécom Paris, Institut polytechnique de Paris, in the [Audio group \(ADASP\)](#) of the “[Signal, Statistics and Learning \(S2A\)](#)” team. *One of the topics is further described below.*

Place: Telecom Paris, 19 place Marguerite Perey, 91120 Palaiseau, France.

Start of the positions: — March/April 2024

Duration: 4 to 6 months

Possible extension to PhD (Funding on the HI-Audio project)

Context: The positions will be a part of the **ERC Advanced (2022) – HI-Audio** (*Hybrid and Interpretable Deep neural audio machines*) project, which aims at building hybrid deep approaches combining parameter-efficient and interpretable models with modern resource-efficient deep neural architectures with applications in speech/audio scene analysis, music information retrieval and sound transformation and synthesis.

Subject: Bias Correction with Pre-trained Audio Embeddings

For data and computational efficiency, it is becoming increasingly common to transfer general-purpose audio representations pre-trained from deep neural networks and data-abundant source tasks to various downstream tasks [1]. These representations, also called embeddings, such as VGGish [2] and OpenL3 [3], can achieve high accuracy on the downstream tasks with less stringent training data requirements. However, the properties of pre-trained audio embeddings are not fully understood. Specifically, and unlike traditionally engineered features, the representations may embed and propagate biases from the model’s training regime. This tendency is prevalent in the audio domain as audio is often collectively recorded under specific facilities or conditions [4]. Building upon [5], this project further investigates the phenomenon of bias propagation and aims at providing a comprehensive guide for bias correction with pre-trained audio embeddings.

In this internship it is proposed:

- to investigate the robustness of pre-trained audio embeddings against various acoustic conditions
- to quantify bias effects and the correctability of bias via domain discrepancy metrics
- to mitigate bias effect with post-processing counter-measures, i.e. without re-training the embedding model
- to provide evaluation metrics for bias mitigation performance

Candidate Profile:

- Students involved in a master or engineering program in *Data Science, Machine learning, Signal Processing, or Speech/Audio/Music processing*.

Application:

- Contact: Gaël Richard, firstname.lastname@telecom-paris.fr.

Bibliography

[1] J. Turian, J. Shier, et al. “Hear: Holistic evaluation of audio representations,” in *Competitions and Demonstrations Track of Neural Information Processing Systems (NeurIPS) conference*, 2021.

[2] A. Jansen, J. Gemmeke, et al. “Large-scale audio event discovery in one million YouTube videos,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2017.

[3] A. Cramer, H. Wu, et al. “Look, listen, and learn more: Design choices for deep audio embeddings,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019.

[4] S. Srivastava, H. Wu, et al. “A study on robustness to perturbations for representations of environmental sound,” in *Proceedings of IEEE European Signal Processing Conference (EUSIPCO)*, 2022.

[5] C. Wang, G. Richard, and B. McFee. “Transfer Learning and Bias Correction With Pre-Trained Audio Embeddings,” in *Proceedings of International Society for Music Information Retrieval (ISMIR)*, 2023.