

Modèles génératifs pour la synthèse sonore

- Laboratoire LISAA (EA4120) - UGE
- Équipe CCAMAN
- à définir, des liens existent avec les centres majeurs en informatique musicale (CCRMA, Stanford University; C4DM, QMUL...)
- Kevin Dahan (kevin.dahan@univ-eiffel.fr)
- Filières visées : Datascience et intelligence artificielle, Arts et Sciences

- Présentation générale du sujet (environ 5 à 10 lignes)

Dans le cadre de nos recherches autour de l'intelligence artificielle appliquée au sonore et aux musiques électroacoustiques, ce projet explorera l'application de modèles génératifs pour la synthèse sonore, avec un accent particulier sur les techniques innovantes de génération de sons dans le cadre des musiques électroacoustiques. L'objectif est de recenser et d'évaluer les modèles existants, en réalisant une veille technologique pour suivre les dernières avancées dans le domaine.

- Objectif du projet (environ 10 à 20 lignes)

Ce stage vise à examiner en profondeur les modèles de synthèse sonore utilisant des approches d'intelligence artificielle, notamment les réseaux de neurones génératifs comme les GANs, VAEs, et autres architectures pertinentes (par exemple, LLM pour code). L'étudiant(e) devra :

- Effectuer une étude approfondie des modèles génératifs de synthèse sonore existants ;
- Réaliser une évaluation de ces modèles à travers des expérimentations concrètes, en prenant en compte des critères comme la qualité sonore et l'efficacité computationnelle ;
- Mettre en place une veille technologique pour identifier les nouvelles approches et modèles ;
- Explorer l'utilisation de ces modèles dans le domaine des musiques électroacoustiques, en lien avec les travaux d'un doctorant du laboratoire sur ce sujet ;
- Identifier les possibilités de collaboration avec des acteurs internationaux, afin d'envisager des applications et stages futurs sur la thématique de la synthèse sonore.

Il est attendu que les résultats donnent lieu à des publications dans les conférences internationales du domaine (ICMC, ISMIR...)

- Bibliographie indicative

- Engel, J., Agrawal, K., Chen, S., Gulrajani, I., Donahue, C., & Roberts, A. (2019). GANSynth: Adversarial Neural Audio Synthesis. arXiv preprint arXiv:1902.08710.
- Oord, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., Kavukcuoglu, K. WaveNet: A Generative Model for Raw Audio: arXiv preprint arXiv:1608.03499 (2016).
- Vasquez, J., & Lewis, E. (2019). MelNet: A Generative Model for Audio in the Frequency Domain. arXiv preprint arXiv:1906.01083.
- Sturm, B. L., Santos, J. F., Ben-Tal, O., & Korshunova, I. (2016). Music transcription modelling and composition using deep learning. In 1st Conference on Computer Simulation of Musical Creativity. arXiv: 1608.08723
