

Analyse de l'impact des étapes de pré-processing sur les performances d'une chaîne d'apprentissage automatique sur données acoustiques

Contexte du stage

L'intégration du deep learning (DL) dans les processus d'analyse de données soulève des problématiques étroitement liées à la structure même des signaux manipulés, fréquemment organisés sous forme de séries temporelles. Ces données peuvent provenir de différents types de capteurs et représenter différents phénomènes allant de la biologie, à l'acoustique, la sismique, différents types de mouvements (stations inertielles), les systèmes industriels, etc., et présentent une forte variabilité en termes de densité spectrale, de régularité temporelle et de contenu informationnel. Ainsi, avant même l'apprentissage ou l'utilisation des réseaux, le choix de la représentation du signal conditionne et impacte notablement la capacité des modèles profonds à extraire des structures pertinentes ainsi que la forme de ces représentations.

Dans ce contexte, la phase de pré-traitement, souvent sous-estimée, devient un levier déterminant dans la performance, la généralisation et l'interprétabilité des architectures de deep learning [1]. Plusieurs stratégies de pré-traitement [2] existent, chacune mettant en évidence des propriétés spécifiques du signal : l'analyse dans le domaine temporel permet de conserver la dynamique brute de la mesure, tandis que les représentations temps-fréquence (transformée de Fourier à court terme, spectrogrammes, etc.) révèlent des régularités locales et des structures fréquentielles [3]. D'autres approches, comme les coefficients cepstraux (MFCC), privilégient une compression perceptuelle du contenu spectral particulièrement adapté à la perception humaine. Enfin, certaines méthodes reposent sur l'extraction explicite de caractéristiques spécifiques, souvent guidée par des connaissances métier, qui permettent de réduire la dimensionnalité du signal tout en ciblant des propriétés jugées discriminantes [4].

Dans de nombreux cas d'usage, notamment en environnement embarqué ou temps réel, l'intégration de modèles de DL se fait sous des contraintes strictes (calcul, mémoire, latence, énergie). Si une part importante des travaux actuels se concentre sur la compression des architectures neuronales ou la quantification des poids pour répondre à ces contraintes [5], les phases amont du traitement du signal restent souvent peu étudiées sous cet angle [6]. Or, ces opérations initiales conditionnent non seulement la charge de calcul globale, mais aussi la qualité des représentations fournies au réseau. Dans une logique d'optimisation de bout en bout, il devient donc indispensable de maîtriser et d'évaluer l'impact de chaque étape de pré-traitement, en tenant compte à la fois des performances d'apprentissage, des coûts d'exécution associés et de la résolution limitée des calculs dans ces étapes.

Le stage proposé adresse ces questions dans le cadre spécifique des données acoustiques, et en particulier dans un contexte de Passive Acoustic Monitoring (PAM) pour le suivi de la biodiversité et l'étude d'environnement à travers leur paysage sonore. En particulier les développements et tests seront effectués sur trois bases de données acoustiques distinctes, permettant ainsi d'étudier la généralité des tendances observées.

Objectifs

L'objectif principal de ce stage est d'évaluer de manière systématique l'impact de différentes stratégies de pré-traitement des séries temporelles sur les performances des modèles de deep learning, en considérant à la fois la qualité des prédictions et les coûts de calcul associés. Cette étude portera sur des signaux issus de deux bases de données acoustiques établies dans la littérature ainsi que d'une base de données issue du LECA sur les cris de Marmotte.

Une première étape consistera en la réalisation d'un état de l'art approfondi sur les méthodes de pré-traitement couramment utilisées dans ces contextes (domaine temporel, temps-fréquence, MFCC, extractions manuelles), ainsi que sur les architectures de deep learning compatibles avec ces représentations.

Dans un second temps, plusieurs chaînes de traitement complètes seront implémentées et évaluées sur les jeux de données sélectionnés. Pour chaque combinaison, on mesurera les performances du modèle en termes de précision, mais également le coût induit par le traitement en amont (temps de calcul, empreinte mémoire), ainsi que la complexité du modèle utilisé (nombre de paramètres, nombre d'opérations élémentaires – FLOPs). Cette approche permettra de construire une

cartographie précise des compromis entre expressivité du pré-traitement, efficacité du réseau, et contraintes computationnelles.

Compétences & profil recherché

- Vous êtes en dernière année d'école d'ingénieur ou en master.
- Vous avez un intérêt fort pour la recherche
- Vous avez des compétences en traitement numérique du signal et en intelligence artificielle notamment sur les opérations des différentes couches des réseaux de neurones.
- Vous avez une connaissance ou curiosité pour les domaines de l'acoustique et/ou de la bioacoustique et/ou de l'acoustique environnementale.

Informations supplémentaires

Ce stage est effectué dans le cadre de la chaire MIAI « *INSITU – Embedded Artificial Intelligence for in situ monitoring* ». Un sujet de thèse sur des thématiques connexes est d'ors-et-déjà acquis, ce stage se veut être une 1ère étude et collaboration entre le ou la candidat(e) et l'équipe encadrante et qui pourrait être poursuivie dans le cadre de la thèse.

Informations & Contacts

- Marielle Malfante : marielle.malfante@cea.fr
- Robin Gerzaguet : robin.gerzaguet@irisa.fr
- Camille Desjonquères : camille.desjonqueres@univ-grenoble-alpes.fr
- Durée : Stage de 6 mois à partir de Février 2026
- Localisation : CEA, Grenoble, France
- Gratification par mois : environ 1400 euros bruts

Pour postuler, merci de contacter l'ensemble des encadrants.

Ne pas hésitez à prendre contact pour échanger autour du sujet ou pour d'éventuelles questions

Bibliographie

- [1] H. J. D. Koresh, "Impact of the preprocessing steps in deep learning-based image classifications," *National Academy Science Letters*, pp. 1–3, 2024.
- [2] B. Bala and S. Behal, "A Brief Survey of Data Preprocessing in Machine Learning and Deep Learning Techniques," in *2024 8th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*, 2024, pp. 1755–1762. doi: 10.1109/I-SMAC61858.2024.10714767.
- [3] S. Peng, S. Sun, and Y.-D. Yao, "A Survey of Modulation Classification Using Deep Learning: Signal Representation and Data Preprocessing," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 12, pp. 7020–7038, 2022, doi: 10.1109/TNNLS.2021.3085433.
- [4] A. Parashar, A. Parashar, W. Ding, M. Shabaz, and I. Rida, "Data preprocessing and feature selection techniques in gait recognition: A comparative study of machine learning and deep learning approaches," *Pattern Recognition Letters*, vol. 172, pp. 65–73, 2023.
- [5] R. David et al., "Tensorflow lite micro: Embedded machine learning for tinyml systems," *Proceedings of machine learning and systems*, vol. 3, pp. 800–811, 2021.
- [6] P.-E. Novac, G. Boukli Hacene, A. Pegatoquet, B. Miramond, and V. Gripon, "Quantization and deployment of deep neural networks on microcontrollers," *Sensors*, vol. 21, no. 9, p. 2984, 2021.