
DEEP LEARNING APPLIQUE A LA REDUCTION DE BRUIT MULTICANALE DE SIGNAUX AUDIO CAPTES PAR UNE ANTENNE DE MICROPHONES EMBARQUEE SUR MICRO-DRONE

DIRECTEUR DE THESE (CNAM PARIS) : ÉRIC BAVU (eric.bavu@lecnam.net)
LABORATOIRE DE MECANIQUE DES STRUCTURES ET DES SYSTEMES COUPLES (LMSSC – EA 3196)

CO-ENCADRANTS : THOMAS JOUBAUD / SEBASTIEN HENGY
INSTITUT DE RECHERCHES FRANCO-ALLEMAND DE RECHERCHES SAINT-LOUIS – APC TEAM

English version follows (page 7)

CONTEXTE

Le développement des drones volants a permis un accroissement des cas d'usage dans les domaines civils et militaires. Si leur utilisation appliquée à la vidéo pour la surveillance, la détection, ou simplement les loisirs, est désormais bien connue des professionnels et du grand public, l'exploitation de **signaux acoustiques** provenant d'une **antenne de microphones embarquée** représente un nouveau défi pour les chercheurs. La facilité de déploiement des drones, couplée avec des techniques connues de traitement d'antenne, permettrait l'application de méthodes de **restitution d'environnement sonore ou de détection d'évènements** sur une vaste zone d'opération. La recherche et le sauvetage de blessés, en cas de sinistre ou sur le champ de bataille, constitue un cas d'usage d'intérêt croissant.

La problématique principale résultant de l'intégration d'une antenne de microphones à un drone demeure la **captation du bruit produit par le moteur et les pales de ce dernier, qui viennent masquer les signaux sonores d'intérêt**. Afin de limiter ce phénomène, il est donc nécessaire de réaliser une opération de **débruitage multicanal des données** captées par l'ensemble des microphones composant l'antenne. Pour cela, nous proposons l'utilisation de méthodes **d'apprentissage profond**. En effet, Les méthodes standard de réduction de bruit de type soustraction spectrale, filtrage de Wiener, exploitent des estimations de la source sonore à éliminer. Pour appliquer ces méthodes, il est donc nécessaire de connaître les moments où le signal d'intérêt est présent, nécessitant la mise en place de mécanismes adaptés (détection d'activité). Dans les applications envisagées, les caractéristiques du signal d'intérêt ne sont pas connues, cette mise en place est alors compromise. De plus, ces méthodes ont tendance à dégrader la phase des signaux, empêchant alors les traitements d'antenne réalisés à partir des données débruitées. Ces limites motivent l'utilisation de méthodes de **réduction de bruit** par Deep Learning afin d'éliminer les perturbations dues au drone.

COFINANCEMENT ET COLLABORATION

Cette thèse s'inscrit dans le cadre du projet AHEAD (Artificial Intelligence for Health, Physics, Transportation and Defense) porté par le Cnam Paris et financé par l'ANR dans le cadre de contrats doctoraux en intelligence artificielle.

Les recherches menées par l'équipe acoustique du LMSSC concernent principalement le développement de la synthèse en temps réel de champs sonores, les méthodes de Deep Learning pour l'acoustique, les problèmes inverses utilisant des méthodes fréquentielles et temporelles, l'imagerie acoustique et la localisation en milieu réverbérant. Depuis 2013, l'équipe acoustique collabore avec l'ISL sur des thématiques liées à la protection du combattant, à travers le co-encadrement de 5 thèses de doctorat, et est le coordinateur d'un projet ANR sur la localisation et la reconnaissance de sources en mouvement par Deep Learning.

Cette thèse est cofinancée par l'Institut de recherche franco-allemand de Saint-Louis. Les recherches du groupe APC (Acoustique et Protection du Combattant) portent sur l'amélioration de la protection du soldat à différents niveaux. Le groupe de recherche cherche à réduire la vulnérabilité des individus en travaillant sur différents aspects de la protection auditive tout en améliorant ses performances opérationnelles (communications audio et sécurité).

CONTEXTE NATIONAL ET INTERNATIONAL

Depuis 2006, le Deep Learning rassemble les efforts scientifiques et économiques des chercheurs universitaires et des grandes entreprises de nouvelles technologies (en particulier les géants « GAFAM » : Google, Apple, Facebook, Amazon et Microsoft), qui y voient un énorme potentiel applicatif. Ces dernières années, nous avons commencé à profiter de ces méthodes dans notre vie quotidienne, et elles deviennent omniprésentes dans les applications de reconnaissance d'images, ainsi que dans la reconnaissance vocale et les assistants vocaux virtuels avec adaptabilité et reconnaissance du locuteur. La modélisation générative des signaux audio est un problème fondamental à l'intersection du traitement du signal et de l'apprentissage automatique, et l'une des avancées récentes les plus significatives dans le traitement audio basé sur l'IA a été la capacité de modéliser directement les signaux bruts dans le domaine temporel en utilisant des réseaux de neurones. Dans ce projet, nous explorons de nouveaux algorithmes de modélisation pour l'audio. En particulier, nous nous concentrons sur le problème spécifique de débruitage dans une situation très défavorable, puisque l'antenne microphonique est placée à proximité des sources de bruit, avec pour objectif de traiter le problème comme un problème « audio-to-audio » multicanal (tant du point de vue des entrées que des sorties).

OBJECTIFS

L'objectif de ce travail de thèse est de mettre en œuvre des techniques de débruitage des signaux mesurés à l'aide de microphones embarqués sur un drone. Le bruit du drone en vol vient perturber les capacités d'écoute des microphones embarqués pour des applications de détection d'événements au sol (déplacement de véhicules, appels à l'aide, conversation,...). Les techniques de débruitage développées dans le cadre de cette thèse, basées

sur des technologies impliquant l'utilisation de réseaux de neurones profonds, devront permettre la détection d'événements d'intérêt se produisant au sol.

On souhaite également confronter les méthodes d'apprentissage profond à des méthodes standards de débruitage dans diverses phases du traitement des signaux fournis par l'antenne de microphones et les capteurs d'attitude embarqués sur le drone. Des études récentes montrent que les techniques standard de débruitage basées sur les filtres de Wiener peuvent en effet s'avérer pertinentes en fonction du type de source que l'on souhaite extraire : pour la détection et la localisation de bruits large bandes (type bruit blanc) elles peuvent être adaptées, mais elles possèdent des performances limitées pour la localisation de signaux harmoniques.

Cette limitation peut être problématique dans la détection et la localisation de signaux de type parole, ou générés par des véhicules, ce qui motive cette étude. Pour cela, on pourra au début des travaux s'appuyer sur des **bases de données publiques existantes** qui ont été mises à disposition pour des applications de type « search and rescue ». La base de données DREGON comprend notamment des données intégrant les informations de vol d'un drone embarquant une antenne de 8 microphones. Cette base de données intègre des informations synchronisées des enregistrements des microphones et des paramètres de vol d'un drone de type quadcoptère, et des positionnements de chaque élément : haut-parleur émettant le signal d'intérêt, et le drone porteur de l'antenne microphonique. Cette **base de données sera complétée avec des mesures faites à l'ISL** en chambre anéchoïque et en environnement représentatif extérieur, **et au LMSSC** à l'aide d'une sphère de haut-parleurs permettant la génération d'un environnement acoustique contrôlé.

La performance des algorithmes proposés pourra être mesurée selon deux méthodologies :

- Méthode statistique mesurant les capacités des algorithmes sur des métriques connues telles que le rapport signal sur bruit, l'application sur les signaux débruités d'algorithmes de détection d'activité et de localisation de sources de type formation de voies
- Méthode subjective, avec mesure sur un panel de testeurs de la qualité perçue des signaux débruités, et sur leur intelligibilité si le signal cible est de la parole, ...

On évaluera notamment au cours du travail de thèse l'intérêt des **différents paramètres d'entrée à disposition pour la base de données d'apprentissage et l'architecture des réseaux** de neurones testés. On pourra ainsi déterminer les paramètres d'entrée les plus impactants sur la capacité à débruiter efficacement un signal sonore mesuré à l'aide de microphones embarqués sur drone.

TACHES

Concernant l'architecture des réseaux de neurones développés pendant la thèse, une attention particulière sera portée à la finalité d'implémentation sur système embarqué. Les contraintes de performance privilégient les architectures « **end-to-end** » utilisant des signaux bruts (dans le domaine temporel) comme entrées.

En premier lieu, une étude bibliographique des méthodes de Deep Learning pertinentes (type de données d'entrée : temporelles brutes, spectrogrammes, / type d'architectures proposées pour la réduction de

bruit multicanale / approches spécifiques aux drones) afin de proposer un cadre et une architecture pour le développement de ces méthodes dans la cadre de la thèse.

En considérant que les caractéristiques du signal d'intérêt doivent être extraites du bruit pour ensuite reconstruire un signal nettoyé, nous nous orientons vers une **architecture de type encodeur-décodeur**, qui pourrait constituer la partie générative d'un réseau adverse génératif dont le discriminateur différencierait les signaux nettoyés de signaux silencieux réels. Ce type d'architecture a été appliqué avec succès au LMSSC pour l'extension de bande passante de signaux vocaux (audio super-résolution). Inclure une dépendance temporelle dans le réseau de neurones serait aussi particulièrement adapté au problème de traitement de signaux acoustiques. Une étude sera alors consacrée à la faisabilité d'implémentation de Transformers causaux intégré au modèle profond.

Pour la constitution de la base de données, un prototype de drone sera à disposition permettant l'acquisition synchrone de données mesurées à l'aide de microphones embarqués et des différents paramètres de vol tels que l'altitude, les informations des capteurs d'attitude, la vitesse rotation des différents moteurs, ...

Une chambre anéchoïque est à disposition pour mesurer les bruits générés par une maquette intégrant quatre moteurs de drone dont la vitesse de rotation est variable. On pourra y mesurer les bruits propres des moteurs, puis ajouter des bruits d'intérêt à l'aide d'un haut-parleur à rapport signal sur bruit variable.

Un terrain d'expériences extérieur permettra la mesure de données à l'aide d'une antenne embarquée sur un drone en vol. On pourra y appliquer des scénarios de vol stationnaires ou mobiles jusqu'à une hauteur de 80 mètres par rapport au sol.

ÉCHEANCIER PREVISIONNEL

Première année :

- Bibliographie sur les techniques de débruitage, la détection d'événement, les méthodes de deep learning appliquées à l'acoustique,
- Constitution d'une base de données d'intérêt en chambre anéchoïque et en environnement extérieur sur sites de l'ISL,
- Choix des capteurs et définition de leur implémentation sur porteur de type drone.
- Test méthodes standard de débruitage sur données chambre anéchoïque et définition « performance de référence »
- Développement de premières architectures de réseaux de neurones adaptées au débruitage temps réel embarqué sur micro-drone.

Deuxième année :

- Test réseau de neurones proposés et confrontation aux performances de méthodes standard sur BDD (RSB, détection d'activité, psycho-acoustique/perception du signal débruité)
- Analyse des performances « temps réel » des différents algorithmes de débruitage sur données chambre anéchoïque.
- Compléter BDD avec mesures en environnement représentatif (T.E. Baldersheim), en présence de silence ou avec activité au sol.
- Data augmentation dans le laboratoire du LMSSC avec outil spherebedev
- Définition des paramètres de vol d'intérêt à exploiter dans le réseau de neurones en complément des données acoustiques brutes, et définition nouvelle architecture de réseau de neurones les intégrant.
- Participation à un congrès international pour présenter les résultats des tests réalisés en première année de thèse
- Rédaction d'un article scientifique dans une revue de rang A sur les résultats obtenus sur la classification et la localisation de plusieurs sources en temps réel.

Troisième année :

- Analyse des performances des méthodes proposées sur des signaux issus des mesures prévues au cours de la seconde année,
- Comparaison performances vol stationnaire/déplacement
- Test de l'impact du débruitage sur la qualité du signal restitué en évaluant les distorsions de phase ou les effets sur des méthodes de traitement d'antenne
- Rédaction d'un article scientifique dans une revue de rang A pour présenter les résultats de l'étude sur l'utilisation simultanée d'algorithmes de classification et de localisation.
- Rédaction du manuscrit de thèse.

CANDIDATURE

Les candidats (possédant une citoyenneté européenne) doivent être titulaires d'un master ou d'un diplôme d'ingénieurs avec une spécialisation en informatique, traitement du signal, acoustique ou mathématiques appliquées. Expérience nécessaire en Deep Learning, si possible appliqués à des signaux audio. Très bonnes compétences en programmation, avec une expérience de l'utilisation des bibliothèques d'apprentissage profond Pytorch ou Tensorflow. Bonnes capacités de synthèse écrite et orale pour la présentation des travaux de recherche. Une expérience dans la rédaction d'un article serait un plus.

Nous avons une flexibilité pour la date de début de thèse, entre Septembre 2023 et Décembre 2023.

Merci d'envoyer votre CV et lettre de motivation à eric.bavu@lecnam.net avant le **15/Mai/2023**.

BIBLIOGRAPHIE

- [1] Markovic, D., Defossez, A., & Richard, A. (2022). **Implicit Neural Spatial Filtering for Multichannel Source Separation in the Waveform Domain.** *arXiv preprint arXiv:2206.15423*.
- [2] Defossez, A., Synnaeve, G., & Adi, Y. (2020). **Real time speech enhancement in the waveform domain.** *arXiv preprint arXiv:2006.12847*.
- [3] Manamperi, W., Abhayapala, T., Samarasinghe, P., & Zhang, J. A. (2023). **Drone Audition: Audio Signal Enhancement from Drone Embedded Microphones.**
- [4] Song, Y., Kindt, S., & Madhu, N. (2022, July). **Drone Ego-Noise Cancellation for Improved Speech Capture using Deep Convolutional Autoencoder Assisted Multistage Beamforming.** In *2022 25th International Conference on Information Fusion (FUSION)* (pp. 1-8). IEEE.
- [5] Spadini, T., Aldeia, G. S. I., Barreto, G., Alves, K., Ferreira, H., Suyama, R., & Nose-Filho, K. (2019, October). **On the application of SEGAN for the attenuation of the ego-noise in the speech sound source localization problem.** In *2019 Workshop on Communication Networks and Power Systems (WCNPS)* (pp. 1-4). IEEE.
- [6] A. Deleforge, **Drone audition for search and rescue: Datasets and challenges**, *Quiet drones international symposium*, 25-27 May 2020, Paris.
- [7] E. Bavu *et al.*, **Deepomatics: A deep-learning based multimodal approach for aerial drone detection and localization**, *Quiet drones international symposium*, June 2022, Paris.
- [8] H. Pujol, E. Bavu, A. Garcia, **BeamLearning: An end-to-end deep learning approach for the angular localization of sound sources using raw multichannel acoustic pressure data**, *The journal of the acoustical society of America*, 149(6), p.4248-4263.
- [9] P. Naz, S. Hengy, A. Ramamonjy, O. Rassy, E. Bavu, **Outdoor field trials for the measurement of the acoustic signals of mini UAVs**, *e-forum Acousticum*, December 2020, Lyon.
- [10] Hauret, J., Joubaud, T., Zimpfer, V., & Bavu, E. (2023, June). **EBEN: Extreme bandwidth extension network applied to speech signals captured with noise-resilient body-conduction microphones.** In *ICASSP 2023: 2023 IEEE International Conference on Acoustics, Speech and Signal Processing*.
- [11] T. Joubaud, G. Pallone, **Electroacoustic method for the calibration of a heterogeneous distributed speaker system**, *European Signal Processing conference*, January 2021.
- [12] T. Joubaud, V. Zimpfer, A. Garcia, C. Langrenne, **Sound localization models as evaluation tools for tactical communication and protective systems**, *The journal of acoustical society of America*, 141(4)
- [13] Hengy S., Naz P., Gounon P., **Source localization using widely separated special microphone arrays: study of the applicability of higher order statistics**, *Proceedings of SPIE - The International Society for Optical Engineering* 01/2004.
- [14] Hengy S. *et al.*, **Acoustic Shooter Localization Using a Network of Asynchronous Acoustic Nodes**, *IET Radar, Sonar & Navigation*, vol. 10(issue 9), 2016, p. 1528-1535.
- [15] Hengy S., **Amélioration des procédés de détection acoustique et développement de nouveaux concepts d'antenne**, rapport ISL R107/2005

DEEP LEARNING APPLIED TO MULTICHANNEL NOISE REDUCTION OF AUDIO SIGNALS CAPTURED BY A MICRO- DRONE MOUNTED MICROPHONE ARRAY

PHD SUPERVISOR (CNAM PARIS) : ÉRIC BAVU (eric.bavu@lecnam.net)
LABORATOIRE DE MECANIQUE DES STRUCTURES ET DES SYSTEMES COUPLES (LMSSC – EA 3196)

CO-SUPERVISORS : THOMAS JOUBAUD / SÉBASTIEN HENGY
INSTITUT DE RECHERCHES FRANCO-ALLEMAND DE RECHERCHES SAINT-LOUIS – APC TEAM

CONTEXT

The development of flying drones has led to an increase in the number of use cases in the civilian and military fields. While their use applied to video for surveillance, detection, or simply leisure is now well known to professionals and the public, the use of acoustic signals from an onboard microphone array represents a new challenge for researchers. The easy deployment of UAVs, combined with microphone array processing techniques, could allow the application of methods for **sound environment restitution or event detection over a large area of operation**. Search and rescue of casualties, in case of disaster or on the battlefield, is a use case of growing interest.

The main problem resulting from the integration of a microphone array into a UAV lies in the **(ego)-noise produced by the engine and the blades** of the latter, which mask the sound signals of interest. To limit this phenomenon, it is therefore necessary to carry out a **multi-channel denoising operation** of the data captured by all the microphones within the array. For that purpose, we propose the use of **deep learning methods**. Indeed, standard noise reduction methods such as spectral subtraction and Wiener filtering use estimates of the sound source to be eliminated. To apply these standard methods, it is therefore necessary to know the time intervals when the signal of interest is present, requiring the implementation of specific mechanisms (activity detection). Given that the characteristics of the signal of interest are not known in the intended applications, this implementation is compromised. In addition, traditional methods tend to deteriorate the phase of the signals, thus preventing array signal processing from being performed using the denoised audio signals. These limitations motivate the use of Deep Learning noise reduction methods to **eliminate UAV noise interference**.

CO-FUNDING AND COLLABORATION

This thesis is part of the AHEAD project (Artificial Intelligence for Health, Physics, Transportation and Defense) led by Cnam Paris and financed by the ANR within the framework of doctoral contracts in Artificial Intelligence.

The research carried out by the LMSSC acoustics team mainly concerns the development of real-time synthesis of sound fields, Deep Learning methods for acoustics, inverse problems using frequency and time methods, acoustic imaging and localization in reverberant environments. Since 2013, the acoustics team has been collaborating with the ISL on themes related to the protection of the combatant, through the co-supervision of 5 doctoral theses, and is the leader of an ANR project on the localization and recognition of moving sources using Deep Learning.

The thesis is co-financed by the French-German Research Institute of Saint-Louis. The research of the APC Group (Acoustics and Protection of the Combatant) group is focused on improving the protection of the soldier at various levels. The research group seeks to reduce the vulnerability of the soldier by working on various aspects of hearing protection while improving his/her operational performance (audio communications and safety).

NATIONAL AND INTERNATIONAL CONTEXT

Since 2006, Deep Learning has brought together the scientific and economic efforts of academic researchers and major new technology companies (in particular the "GAFAM" giants: Google, Apple, Facebook, Amazon and Microsoft), which see enormous application potential. In recent years, we have begun to take advantage of these methods in everyday life, and they are becoming ubiquitous in image recognition applications, as well as in voice recognition and virtual voice assistants with adaptability and speaker recognition.

The generative modeling of audio signals is a fundamental problem at the intersection of signal processing and machine learning, and one of the most significant recent advances in machine learning-based audio processing has been the ability to directly model *raw signals* in the time domain using neural networks. In this project, we explore new modeling algorithms for audio.

In particular, we focus on the specific **multichannel denoising problem in a very adverse situation**, since the microphone array is placed in vicinity of the noise sources, with the objective of treating the problem as a multi-channel "audio-to-audio" problem (both from the input and output point of view).

OBJECTIVES

The objective of this PhD research is to implement denoising techniques for signals measured using microphones mounted on an unmanned aerial vehicle. The noise produced by the flying aircraft affects the ability to use embedded microphones to monitor events on the ground (vehicle movements, calls for help, conversations, etc.). The denoising techniques developed in this project, which are based on deep neural network techniques, should allow to detect events of interest occurring on the ground.

We also wish to evaluate the deep learning methods against standard denoising methods in various stages of the signal processing, using the microphone array and the attitude sensors data embedded on the UAV. Recent studies show that standard denoising techniques based on Wiener filters can indeed be relevant depending on the type of source to be extracted: for the detection and localization of broadband noise (white noise type) they can be adapted, but they have limited performances for the localization of harmonic signals.

This limitation can be problematic for the detection and localization of speech type signals, or vehicles generated signals, which motivates this study. For this, we will initially leverage publicly available databases that have been made available for "search and rescue" type applications. The DREGON database includes data including flight information from a drone with an 8 microphone antenna. This database gathers synchronized information from microphones recordings and flight parameters of a quadcopter type UAV, and the position of each element: the loudspeaker emitting the signal of interest, and the UAV carrying the microphone antenna. This database will be enriched with measurements made at the ISL in an anechoic chamber and in an outdoor representative environment, and at the LMSSC using a sphere of loudspeakers allowing the generation of a controlled acoustic environment.

The performance of the proposed algorithms can be measured using two methodologies:

- Statistical methods measuring the capabilities of the algorithms on known metrics such as the signal-to-noise ratio, the application of activity detection algorithms and source localization algorithms.
- Subjective methods, with measurement of the perceived quality of denoised signals on a panel of testers, and their intelligibility if the target signal is speech, ...

In particular, we will evaluate the benefits of the various input parameters available in the training dataset and the architecture of targeted neural networks. We shall thus be able to identify the input parameters that have the greatest impact on the ability to efficiently denoise a sound signal measured using microphones on board a drone.

TASKS

Regarding the architecture of the neural networks developed during the thesis, a specific attention will be paid to the implementation on embedded systems. Performance constraints tend to be in favor of end-to-end architectures using raw signals (in the time domain) as inputs.

First, a bibliographic survey of relevant Deep Learning methods (type of input data: raw time data, spectrograms, / type of architectures proposed for multichannel noise reduction / UAV specific studies ...) in order to propose a framework and an architecture for the development of these methods in the context of this PhD thesis.

Considering that the features of the signal of interest must be extracted from the noise to reconstruct a cleaned signal, we move towards an encoder-decoder type architecture, which could constitute the generative part of a generative adversarial network whose discriminator would differentiate cleaned signals from real "silent" signals. This type of architecture has been successfully applied at the LMSSC for bandwidth extension of speech signals (super-resolution audio). Including a time dependency in the neural network would also be particularly adapted to the problem of acoustic signal processing. A study will then be devoted to the feasibility of implementing causal transformers integrated into the deep model. For the constitution of the database, a prototype drone will be available allowing the synchronous acquisition of data measured with on-

board microphones and various flight parameters such as altitude, attitude sensor information, speed of rotation of the various engines, ...

An anechoic chamber is available to measure the noise generated by a model integrating four drone engines whose rotation speed is variable. We will be able to measure the clean noise of the motors, then add noises of interest using a loudspeaker with variable signal to noise ratio. An outdoor experimental field will allow the measurement of data using an antenna embedded on a drone in flight. It will be possible to apply stationary or mobile flight scenarios up to a height of 80 meters from the ground.

PREVISIONAL TIMESCALE / SCHEDULE

First year:

- Bibliography on denoising techniques, event detection, deep learning methods applied to acoustics,
- Constitution of a database of interest in anechoic chamber and in outdoor environment on ISL sites,
- Choice of sensors and implementation on a drone type carrier.
- Testing standard denoising methods on anechoic chamber data and defining "reference performance".
- Development of the first neural network architectures adapted to real-time denoising on micro-drone.

Second Year

- Testing of proposed neural networks and comparison with the performance of standard methods on Dataset (SNR, activity detection, psychoacoustics/signal perception)
- Analysis of the "real time" performances of the different denoising algorithms on anechoic chamber data.
- Complete Dataset with measurements in representative environment (Baldersheim), in presence of silence or with activity on the ground.
- Data augmentation in the LMSSC laboratory with spherebedev tool
- Definition of the flight parameters of interest to be used in the neural network in addition to the raw acoustic data, and definition of a new neural network architecture integrating them.
- Participation in an international congress to present the results of the tests carried out during the first year of the thesis
- Writing of a scientific article in a rank A journal on the results obtained on the classification and localization of several sources in real time.

Third Year :

- Analysis of the performance of the proposed methods on signals from the measurements planned during the second year,
- Comparison of hover/displacement performance
- Test of the impact of denoising on the quality of the restored signal by evaluating the phase distortions or the effects on antenna processing methods
- Writing of a scientific paper in a top-tier journal to present the results of the study on the simultaneous use of classification and localisation algorithms.
- Writing of the thesis manuscript.

APPLICATION

Candidates (holding European citizenship) must hold a master's degree or engineering school with a major in computer science, signal processing, acoustics, or applied mathematics. Experience in machine learning and deep learning, in particular convolutional networks.

The candidate should have very good programming skills, with experience using the Pytorch or Tensorflow deep learning libraries. Good written and oral synthesis skills for the presentation of research work. Experience in writing an article would be a plus.

We have a flexible start date for the thesis, between September 2023 and December 2023.

Please send your CV and cover letter to eric.bavu@lecnam.net before **15/May/2023**.

REFERENCES

- [1] Markovic, D., Defossez, A., & Richard, A. (2022). **Implicit Neural Spatial Filtering for Multichannel Source Separation in the Waveform Domain**. *arXiv preprint arXiv:2206.15423*.
- [2] Defossez, A., Synnaeve, G., & Adi, Y. (2020). **Real time speech enhancement in the waveform domain**. *arXiv preprint arXiv:2006.12847*.
- [3] Manamperi, W., Abhayapala, T., Samarasinghe, P., & Zhang, J. A. (2023). **Drone Audition: Audio Signal Enhancement from Drone Embedded Microphones**.
- [4] Song, Y., Kindt, S., & Madhu, N. (2022, July). **Drone Ego-Noise Cancellation for Improved Speech Capture using Deep Convolutional Autoencoder Assisted Multistage Beamforming**. In *2022 25th International Conference on Information Fusion (FUSION)* (pp. 1-8). IEEE.
- [5] Spadini, T., Aldeia, G. S. I., Barreto, G., Alves, K., Ferreira, H., Suyama, R., & Nose-Filho, K. (2019, October). **On the application of SEGAN for the attenuation of the ego-noise in the speech sound source localization problem**. In *2019 Workshop on Communication Networks and Power Systems (WCNPS)* (pp. 1-4). IEEE.
- [6] A. Deleforge, **Drone audition for search and rescue: Datasets and challenges**, *Quiet drones international symposium*, 25-27 May 2020, Paris.
- [7] E. Bavu *et al.*, **Deplomatics: A deep-learning based multimodal approach for aerial drone detection and localization**, *Quiet drones international symposium*, June 2022, Paris.
- [8] H. Pujol, E. Bavu, A. Garcia, **BeamLearning: An end-to-end deep learning approach for the angular localization of sound sources using raw multichannel acoustic pressure data**, *The journal of the acoustical society of America*, 149(6), p.4248-4263.
- [9] P. Naz, S. Hengy, A. Ramamonjy, O. Rassy, E. Bavu, **Outdoor field trials for the measurement of the acoustic signals of mini UAVs**, *e-forum Acousticum*, December 2020, Lyon.
- [10] Hauret, J., Joubaud, T., Zimpfer, V., & Bavu, E. (2023, June). **EBEN: Extreme bandwidth extension network applied to speech signals captured with noise-resilient body-conduction microphones**. In *ICASSP 2023: 2023 IEEE International Conference on Acoustics, Speech and Signal Processing*.

- [11] T. Joubaud, G. Pallone, **Electroacoustic method for the calibration of a heterogeneous distributed speaker system**, *European Signal Processing conference*, January 2021.
- [12] T. Joubaud, V. Zimpfer, A. Garcia, C. Langrenne, **Sound localization models as evaluation tools for tactical communication and protective systems**, *The journal of acoustical society of America*, 141(4)
- [13] Hengy S., Naz P., Gounon P., **Source localization using widely separated special microphone arrays: study of the applicability of higher order statistics**, *Proceedings of SPIE - The International Society for Optical Engineering* 01/2004.
- [14] Hengy S. *et al.*, **Acoustic Shooter Localization Using a Network of Asynchronous Acoustic Nodes**, *IET Radar, Sonar & Navigation*, vol. 10(issue 9), 2016, p. 1528-1535.
- [15] Hengy S., **Amélioration des procédés de détection acoustique et développement de nouveaux concepts d'antenne**, rapport ISL R107/2005