

Stéganographie et tatouage audio : Robustesse, Détection et Identification



Mots-clés : stéganographie, tatouage, watermarking, apprentissage profond, traitement de la parole

Description

Au sein du Pôle Recherche de l'AMIAD (Agence Ministérielle pour l'IA de Défense), vous rejoindrez notre équipe, où vous mettrez en place plusieurs solutions à l'état de l'art de stéganographie audio opensource, dont vous comparerez la robustesse face à diverses transformations du signal audio. Vous développerez ensuite une solution de détection de la présence de stéganographie audio dans un enregistrement (détection de présence), pour poursuivre avec une solution d'identification de la méthode de stéganographie utilisée (identification de source).

La stéganographie audio consiste à dissimuler un message secret ou un fichier à l'intérieur d'un fichier audio, comme un enregistrement de parole, de manière imperceptible pour l'oreille humaine. Ce message secret permet souvent d'identifier l'origine de l'enregistrement. Il doit rester autant que possible robuste à des manipulations de l'enregistrement, comme l'ajout de bruit ou l'emploi de coupures et d'extractions de segments. La stéganographie est proche du "tatouage", aussi appelé "watermarking".

Différentes solutions ont été développées au cours des dernières décennies, impliquant principalement des méthodes comme LSB, Echo Hiding, Phase Coding ou des approches spectrales [Bassia et al., 2001, Arnold, 2000, Djebbar et al., 2012, Gopalan, 2003]. Plus récemment, la génération ou la transformation d'enregistrements audio à l'aide d'approches issues du deep learning ont transformé le paysage et ont également amené le développement d'approches de stéganographie/tatouage à base de deep learning [Jiang et al., 2020, Liu et al., 2024, Salah et al., 2025, Singh et al., 2024, Zong et al., 2025, Li et al., 2024]. L'usage de la stéganographie/tatouage se généralise également rapidement, au vu de la récente réglementation européenne.

Travail à réaliser

- Etude comparative des solutions récemment proposées ;
- Réalisation d'un benchmark réaliste permettant de mesurer la robustesse des approches proposées à diverses manipulations des enregistrements audio ;
- Définition, mise en place et évaluation d'une solution neuronale, inspirée de la détection de deepfake audio, de détection de la présence de stéganographie/tatouage dans un enregistrement audio. Cette phase nécessitera l'apprentissage ou l'adaptation d'un "encoder" ainsi que d'un "classifieur". L'encoder transforme un enregistrement audio de longueur variable en un vecteur de caractéristiques de taille fixe. Le classifieur estime si un enregistrement, représenté par un tel vecteur, possède ou non des informations relevant d'une technique de stéganographie ;
- Définition, mise en place et évaluation d'une solution neuronale d'identification de stéganographie. La tâche consiste ici à identifier quelle est la méthode de stéganographie utilisée dans un enregistrement donné. L'application reste proche de la précédente, avec un encodeur et un classifieur mais, cette fois, le classifieur indique la méthode utilisée parmi les n connues. Les deux cas spéciaux (pas de stéganographie et stéganographie inconnue) pourront être traités dans un deuxième temps.

Profil recherché

Vous suivez une formation d'ingénieur ou de master en mathématiques ou informatique et vous disposez :

- de connaissances préalables dans les domaines de l'apprentissage automatique, de l'apprentissage profond ou du traitement de la parole ;
- d'un intérêt prononcé pour la recherche et les nouvelles technologies ;
- de solides compétences en programmation *python* et *PyTorch*.

Informations générales

- Durée : 6 mois
- Niveau : Master 2 ou dernière année d'école d'ingénieur
- Début : dès que possible
- Lieu : AMIAD - X-Innovation Center, École Polytechnique, Palaiseau
- Ressources mises à disposition : l'AMIAD dispose de ressources computationnelles dédiées (Serveurs et GPU).
- Contacts : Jean-François Bonastre (<https://cv.hal.science/jean-francois-bonastre>), jean-francois.bonastre@polytechnique.edu

Point de contact

Jean-François Bonastre : jean-francois.bonastre@polytechnique.edu

Références

- [Arnold, 2000] Arnold, M. (2000). Audio watermarking : Features, applications, and algorithms. In *IEEE international conference on multimedia and expo (II)*, pages 1013–1016.
- [Bassia et al., 2001] Bassia, P., Pitas, I., and Nikolaidis, N. (2001). Robust audio watermarking in the time domain. *IEEE Transactions on multimedia*, 3(2) :232–241.
- [Djebbar et al., 2012] Djebbar, F., Ayad, B., Meraim, K. A., and Hamam, H. (2012). Comparative study of digital audio steganography techniques. *EURASIP Journal on Audio, Speech, and Music Processing*, 2012(1) :25.
- [Gopalan, 2003] Gopalan, K. (2003). Audio steganography using bit modification. In *2003 International Conference on Multimedia and Expo. ICME'03. Proceedings (Cat. No. 03TH8698)*, volume 1, pages I–629. IEEE.
- [Jiang et al., 2020] Jiang, S., Ye, D., Huang, J., Shang, Y., and Zheng, Z. (2020). Smartsteganography : Light-weight generative audio steganography model for smart embedding application. *Journal of Network and Computer Applications*, 165 :102689.
- [Li et al., 2024] Li, P., Zhang, X., Xiao, J., and Wang, J. (2024). Ideaw : Robust neural audio watermarking with invertible dual-embedding. In *Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing*, pages 4500–4511.
- [Liu et al., 2024] Liu, H., Guo, M., Jiang, Z., Wang, L., and Gong, N. Z. (2024). Audiomarkbench : Benchmarking robustness of audio watermarking. *Advances in Neural Information Processing Systems*, 37 :52241–52265.
- [Salah et al., 2025] Salah, E., Narima, Z., Khaldi, A., and Redouane, K. M. (2025). Survey of imperceptible and robust digital audio watermarking systems. *Multimedia Tools and Applications*, 84(7) :3635–3681.
- [Singh et al., 2024] Singh, M. K., Takahashi, N., Liao, W., and Mitsufuji, Y. (2024). Silencipher : Deep audio watermarking. *arXiv preprint arXiv :2406.03822*.
- [Zong et al., 2025] Zong, W., Chow, Y.-W., Susilo, W., Baek, J., and Camtepe, S. (2025). {AudioMarkNet} : Audio watermarking for deepfake speech detection. In *34th USENIX Security Symposium (USENIX Security 25)*, pages 4663–4682.