

Offre de Post-doctorat

Inférence du visage à partir de la voix

Contexte

Le laboratoire GREYC UMR CNRS de l'Université de Caen Normandie et ENSICAEN (Caen, France), en collaboration avec l'entreprise United Biometrics (Caen, France), lance un appel à candidatures pour un poste de recherche postdoctorale (durée entre un et 3 ans) sur l'inférence du visage à partir de la voix. Les travaux réalisés se feront dans le cadre du projet BIOPOP (BIométrie Pour les Opérations) financé par l'AID (Agence Innovation Défense).

Missions

L'objectif visé est d'inférer des informations sur le visage d'une personne à partir d'un enregistrement brut de sa voix. De récents travaux préliminaires [1, 2, 3, 4] ont montré la faisabilité de cette inférence. Il ne s'agira pas de générer le visage exact correspondant à la voix, mais de générer un visage qui permette de ressortir les caractéristiques discriminantes principales du visage (genre, âge, ethnicité, attributs crano-faciaux). Ceci peut présenter de nombreux intérêts dans le cadre d'applications liées à la sécurité. En effet, inférer un visage à partir d'une voix peut permettre ensuite à un opérateur d'effectuer différentes tâches à partir de l'image de visage inférée. Nous pouvons citer par exemple : la vérification de la cohérence entre une voix et un visage, la recherche du visage inféré dans une base de données.

Le post-doctorant réalisera un état de l'art précis des méthodes d'inférence du visage à partir de la voix. Il implémentera une solution de l'état de l'art reposant sur des modèles génératifs. Enfin il développera un nouveau modèle génératif plus efficace, qui puisse garantir la génération d'un visage réaliste pour répondre aux attentes du projet BIOPOP.

Profil attendu

- Doctorat en informatique et spécialisé en apprentissage automatique.
- De solides connaissances de l'apprentissage profond, de la vision par ordinateur.
- Des publications dans des conférences majeures du domaine.
- De solides compétences en développement logiciel/programmation, en particulier en Python/PyTorch.
- De bonnes compétences en communication écrite et verbale sont requises, le candidat doit parler couramment le français et maîtriser l'anglais écrit.
- Des compétences interpersonnelles et la capacité à travailler individuellement ou en tant que membre d'une équipe de projet sont recommandées.

Informations générales

- **Laboratoire :** Le laboratoire GREYC (UMR CNRS 6072) est une Unité Mixte de Recherche en sciences du numérique sous la tutelle de l'ENSICAEN, du CNRS et de l'Université de Caen Normandie (UNICAEN). Les travaux seront effectués au sein de l'équipe Image dont les activités de recherche sont centrées sur le développement de nouvelles méthodes de traitement et d'analyse de signaux/images/vidéos.

- **Lieu** : Caen (France), située en région Normandie, près de la mer et à environ 240 km à l'ouest de Paris, la ville possède encore de nombreux quartiers anciens, une population d'environ 120 000 habitants et une agglomération d'environ 250 000 habitants, incluant plus de 30 000 étudiants à l'Université.
- **Durée** : Deux ans, avec une extension possible d'un an.
- **Salaire** : Environ 2300 € net par mois.
- **Pour postuler** : Les candidats intéressés doivent soumettre (par courrier électronique, dans un seul fichier pdf) leur curriculum vitae, la liste de leurs publications, une lettre de motivation et les coordonnées de trois références (ne pas inclure les lettres de référence avec vos candidatures car nous ne les demanderons qu'aux candidats présélectionnés). Les candidatures seront admises jusqu'à ce que le poste soit pourvu. Le poste débutera début Octobre.
- **Contact / encadrement** :
 - Olivier Lézoray (olivier.lezoray@unicaen.fr, Professeur, UNICAEN, GREYC)
 - Sébastien Bougleux (sebastien.bougleux@unicaen.fr, Maître de Conférences, UNICAEN, GREYC)
 - Christophe Charrier (christophe.charrier@unicaen.fr, Maître de Conférences HDR, UNICAEN, GREYC)

Références

- [1] Amanda Cardoso Duarte, Francisco Roldan, Miquel Tubau, Janna Escur, Santiago Pascual, Amaia Salvador, Eva Mohedano, Kevin McGuinness, Jordi Torres, and Xavier Giró-i-Nieto, "Wav2pix : Speech-conditioned face generation using generative adversarial networks," in *IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2019, Brighton, United Kingdom, May 12-17, 2019*. 2019, pp. 8633–8637, IEEE.
- [2] Zheng Fang, Zhen Liu, Tingting Liu, Chih-Chieh Hung, Jiangjian Xiao, and Guangjin Feng, "Facial expression GAN for voice-driven face generation," *Vis. Comput.*, vol. 38, no. 3, pp. 1151–1164, 2022.
- [3] Tae-Hyun Oh, Tali Dekel, Changil Kim, Inbar Mosseri, William T. Freeman, Michael Rubinstein, and Wojciech Matusik, "Speech2face : Learning the face behind a voice," in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. 2019, pp. 7539–7548, Computer Vision Foundation / IEEE.
- [4] Yandong Wen, Bhiksha Raj, and Rita Singh, "Face reconstruction from voice using generative adversarial networks," in *Advances in Neural Information Processing Systems 32 : Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett, Eds., 2019, pp. 5266–5275.