Offre n°2025-09161

engineer position: Automatic speech recognition for non-natives speakers in a noisy environment

Type de contrat : Fixed-term contract

Niveau de diplôme exigé : Graduate degree or equivalent

Fonction : Temporary scientific engineer

Niveau d'expérience souhaité : Recently graduated

Contexte et atouts du poste

The work will be performed at MultiSpeech Team of INRIA-LORIA, Nancy.

MULTISPEECH is a joint research team between the Université of Lorraine, Inria, and CNRS. It is part of department D4 “Natural language and knowledge processing” of LORIA.

Its research focuses on speech processing, with particular emphasis to multisource (source separation, robust speech recognition), multilingual (computer assisted language learning), and multimodal aspects.

Mission confiée

Context
When a person has their hands busy performing a task like driving a car or piloting an airplane, voice is a fast and efficient way to achieve interaction. In aeronautical communications, the English language is most often compulsory. Unfortunately, a large part of the pilots are not native English and speak with an accent dependent on their native language and are therefore influenced by the pronunciation mechanisms of this language. Inside an aircraft cockpit, non-native voice of the pilots and the surrounding noises are the most difficult challenges to overcome in order to have efficient automatic speech recognition (ASR). The problems of non-native speech are numerous: incorrect or approximate pronunciations, errors of agreement in gender and number, use of non-existent words, missing articles, grammatically incorrect sentences, etc. The acoustic environment adds a disturbing component to the speech signal. Much of the success of speech recognition relies on the ability to take into account different accents and ambient noises into the models used by ARP.
Automatic speech recognition has made great progress thanks to the spectacular development of deep learning. In recent years, end-to-end automatic speech recognition, which directly optimizes the probability of the output character

sequence based on the input acoustic characteristics, has made great progress [Chan et al., 2016; Baevski et al., 2020; Gulati, et al., 2020].

Objectives
The recruited person will have to develop methodologies and tools to obtain high-performance non-native automatic speech recognition in the aeronautical context and more specifically in a (noisy) aircraft cockpit.
This project will be based on an end-to-end automatic speech recognition system.

References
[Baevski et al., 2020] A. Baevski, H. Zhou, A. Mohamed, and M. Auli. Wav2vec 2.0: A framework for self-supervised learning of speech representations, 34th Conference on Neural Information Processing Systems (NeurIPS 2020), 2020.
[Chan et al., 2016] W. Chan, N. Jaitly, Q. Le and O. Vinyals. Listen, attend and spell: A neural network for large vocabulary conversational speech recognition. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2016, pp. 4960-4964, 2016.
[Chorowski et al., 2017] J. Chorowski, N. Jaitly. Towards better decoding and language model integration in sequence to sequence models. Interspeech, 2017.
[Houlsby et al., 2019] N. Houlsby, A. Giurgiu, S. Jastrzebski, B. Morrone, Q. De Laroussilhe, A. Gesmundo, M. Attariyan, S. Gelly. Parameter-efficient transfer learning for NLP. International Conference on Machine Learning, PMLR, pp. 2790–2799, 2019.
[Gulati et al., 2020] A. Gulati, J. Qin, C.-C. Chiu, N. Parmar, Y. Zhang, J. Yu, W. Han, S. Wang, Z. Zhang, Y. Wu, and R. Pang. Conformer: Convolution-augmented transformer for speech recognition. Interspeech, 2020.
[Shi et al., 2021] X. Shi, F. Yu, Y. Lu, Y. Liang, Q. Feng, D. Wang, Y. Qian, and L. Xie. The accented english speech recognition challenge 2020: open datasets, tracks, baselines, results and methods. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 6918–6922, 2021.

Principales activités

The main activities are those typical of a engineer. They include: literature reading, scientific development, programming and simulation, data processing, reporting and presentation, paper writing, collaboration with the team, the supervisors and other scientific partners.

Duration: 8-10 months

Applications will be examined subject to the requirements of the Directorate General of Armaments (DGA).

Compétences

- M.Sc. or engineer degree in speech/audio processing, computer vision, machine learning, or in a related field,
- ability to work independently as well as in a team,
- solid programming skills (Python, PyTorch), and deep learning knowledge,
- good level of written and spoken English.

Avantages

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

Rémunération

From 2692€ gross/month according to experience and qualifications

Informations générales

- **Thème/Domaine :** Language, Speech and Audio
  IT Technical and production engineering (BAP E)
- **Ville :** Villers lès Nancy
- **Centre Inria :** Centre Inria de l'Université de Lorraine
- **Date de prise de fonction souhaitée :** 2025-10-01
- **Durée de contrat :** 12 months
- **Date limite pour postuler :** 2025-08-15

Contacts

- **Équipe Inria :** MULTISPEECH
- **Recruteur :**
  Illina Irina / irina.illina@loria.fr

A propos d'Inria

Inria est l'institut national de recherche dédié aux sciences et technologies du numérique. Il emploie 2600 personnes. Ses 215 équipes-projets agiles, en général communes avec des partenaires académiques, impliquent plus de 3900 scientifiques pour relever les défis du numérique, souvent à l'interface d'autres disciplines. L'institut fait appel à de nombreux talents dans plus d'une quarantaine de métiers différents. 900 personnels d'appui à la recherche et à l'innovation contribuent à faire émerger et grandir des projets scientifiques ou entrepreneuriaux qui impactent le monde. Inria travaille avec de nombreuses entreprises et a accompagné la création de plus de 200 start-up. L'institut s'e?orce ainsi de répondre aux enjeux de la transformation numérique de la science, de la société et de l'économie.

Attention: Les candidatures doivent être déposées en ligne sur le site Inria. Le traitement des candidatures adressées par d'autres canaux n'est pas garanti.

Consignes pour postuler

Sécurité défense :
Ce poste est susceptible d'être affecté dans une zone à régime restrictif (ZRR), telle que définie dans le décret n°2011-1425 relatif à la protection du potentiel scientifique et technique de la nation (PPST). L'autorisation d'accès à une zone est délivrée par le chef d'établissement, après avis ministériel favorable, tel que défini dans l'arrêté du 03 juillet 2012, relatif à la PPST. Un avis ministériel défavorable pour un poste affecté dans une ZRR aurait pour conséquence l'annulation du recrutement.

Politique de recrutement :
Dans le cadre de sa politique diversité, tous les postes Inria sont accessibles aux personnes en situation de handicap.