

Proposition de stage M2 – IRIT, équipe SAMoVA

Adaptations de systèmes de reconnaissance automatique de parole en contexte pathologique (cancer ORL)

1. Contexte scientifique

L'utilisation d'outils d'aide à la transcription, notamment via les outils de reconnaissance automatique de la parole (ASR, Automatic Speech Recognition), est un champ de recherche en plein développement dans le contexte hospitalier (Zhang et al., 2023). En effet, pouvoir bénéficier d'une transcription de la parole des patients permettrait de réaliser des analyses de contenu de leur discours. Nous pourrions ainsi étudier si le choix des mots utilisés, les thématiques abordées ou le sentiment se dégageant de leur discours donnent des indices sur leur qualité de vie (Balaguer, 2021). Ce point est essentiel pour des médecins ou des orthophonistes car l'objectif du soin de façon générale renvoie à l'amélioration de la qualité de vie des patients (Reich, 2009).

Or, actuellement, les systèmes de reconnaissance automatique de parole fonctionnent mal sur la parole pathologique. Appliqués à la parole après cancer ORL, les systèmes présentent des taux d'erreurs mots élevés (54,9 % à partir d'un système Conformer-Transducer Large¹, 67,3 % avec Whisper d'OpenAI² ou encore 71,1 % avec un système wav2vec 2.0 avec CTC/Attention³). Cela peut s'expliquer par les dégradations acoustiques de l'articulation de la parole liées aux modifications anatomiques consécutives à la maladie ou à son traitement (chirurgical notamment).

De plus, le contexte de l'étude de la parole spontanée (qui est la plus proche des « vraies » conditions de productions de parole quotidiennes (Prins & Bastiaanse, 2004)) représente de multiples challenges en raison de l'absence de référence (contrairement à l'étude de la parole sur un texte lu par exemple (Evain et al., 2022)) et à la variabilité qui caractérise ce type de parole (multiples locuteurs, parole moins précise et plus relâchée...).

2. Objectif(s) du stage

L'objectif de ce stage est d'améliorer la reconnaissance automatique de la parole spontanée pathologique après cancer ORL en utilisant de nouveaux systèmes ASR et/ou en adaptant des systèmes existants aux spécificités de la parole cancérologique.

Ce stage pourrait être poursuivi en doctorat dans une thématique relative à l'analyse automatique de la parole spontanée cancérologique en contexte clinique pour étudier la communication et la qualité de vie des patients.

3. Environnement scientifique

Le stage se déroulera à l'IRIT, dans l'équipe SAMoVA. Cette équipe a développé depuis quelques années des thématiques de recherche autour de l'analyse automatique de la parole pathologique. Elle a notamment participé aux projets C2SI (Institut National du Cancer 2014-2017), RUGBI (ANR, 2017-2022) et contribue au projet DAPADAF-E depuis 2019 (Ministère de la Santé). Ainsi, le stagiaire bénéficiera de l'expertise des enseignants chercheurs de l'équipe SAMoVA, ainsi que des liens étroits avec les équipes hospitalières du CHU de Toulouse (médecins et orthophonistes).

Ce stage s'appuie sur une base de données déjà constituée (au sein de la Parolothèque) comprenant un corpus de parole spontanée de 25 patients ayant été traités pour un cancer ORL (enregistrements d'entretiens entre patient et orthophoniste) avec annotations complètes des entretiens. Ce corpus est le seul de ce type à ce jour en langue française.

4. Lieu du stage

IRIT, Université Paul Sabatier Toulouse III, 118 Route de Narbonne

¹ Site : https://catalog.ngc.nvidia.com/orgs/nvidia/teams/nemo/models/stt_fr_conformer_transducer_large, consulté le 02/10/2023

² Site : <https://github.com/openai/whisper>, consulté le 02/10/2023

³ Site : <https://huggingface.co/speechbrain/asr-wav2vec2-commonvoice-fr>, consulté le 02/10/2023

5. Durée du stage

5 à 6 mois

6. Connaissances et compétences requises

- Conditions requises : Master 2 ou dernière année d'école d'ingénieur
- Savoirs : connaissances en programmation (Python notamment), en traitement du signal et en apprentissage automatique
- Savoir-faire : bonnes pratiques de développement informatique
- Savoir-être : capacités d'organisation, autonomie, initiative, intérêt pour les aspects cliniques

7. Encadrement et contact

Les candidatures doivent être envoyées à Julien Pinquier, enseignant-chercheur (julien.pinquier@irit.fr) et Mathieu Balaguer, ingénieur de recherche et orthophoniste (mathieu.balaguer@irit.fr).

8. Références

- Balaguer, M. (2021). *Mesure de l'altération de la communication par analyses automatiques de la parole spontanée après traitement d'un cancer oral ou oropharyngé* [Université Toulouse III Paul Sabatier]. <http://www.theses.fr/2021TOU30109>
- Evain, S., Rossato, S., Lecouteux, B., Portet, F., Evain, S., Rossato, S., Lecouteux, B., Portet, F., & De, E. (2022). *Exploration de systèmes end-to-end pour la reconnaissance automatique de la parole spontanée* To cite this version : HAL Id : hal-03507681.
- Prins, R., & Bastiaanse, R. (2004). Analysing the spontaneous speech of aphasic speakers. *Aphasiology*, 18(12), 1075–1091. <https://doi.org/10.1080/02687030444000534>
- Reich, M. (2009). Cancer et image du corps : identité, représentation et symbolique : Le corps retrouvé (French). *Cancer and Body Image : Identity, Representation (English)*, 85(3), 247–254. <https://doi.org/10.3917/inpsy.8503.0247>
- Zhang, J., Wu, J., Qiu, Y., Song, A., Li, W., Li, X., & Liu, Y. (2023). Intelligent speech technologies for transcription, disease diagnosis, and medical equipment interactive control in smart hospitals: A review. *Computers in Biology and Medicine*, 153(September 2022), 106517. <https://doi.org/10.1016/j.combiomed.2022.106517>