

Traduction vidéo de langage des signes en texte

Bertrand Awenze¹, Christian Nebot², John-William Lebel³, Gabriel Anjuère⁴, Lucas Ambinintsoa⁵ ¹

^{1,3} GLO-4030 ^{2,4,5} GLO-7030

April 29, 2025

Objectif

L'objectif de ce projet est de créer et d'entraîner un modèle capable de traduire une vidéo en langage des signes en texte.

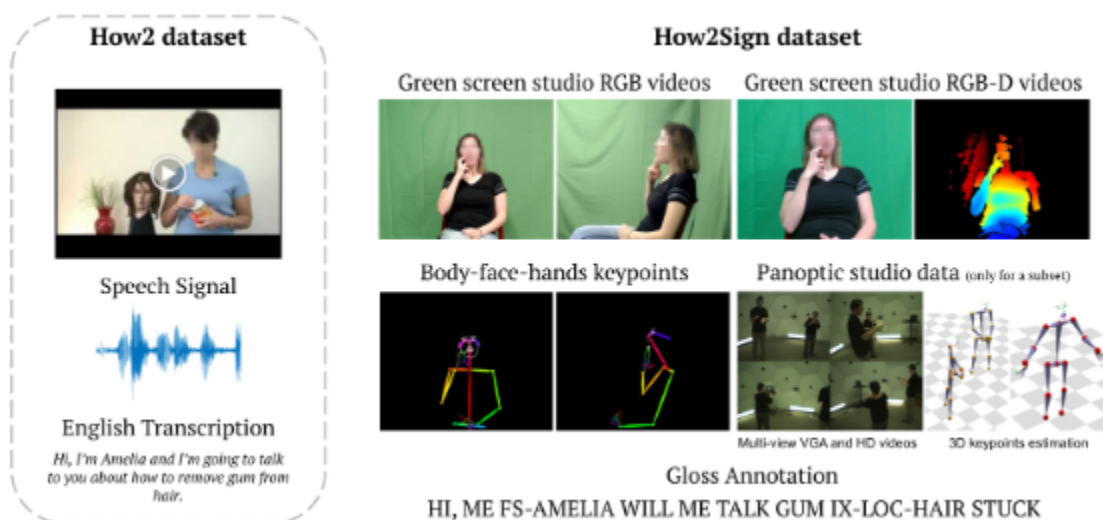


Figure 1: Présentation du projet

Pourquoi avoir choisi ce projet pour GLO-7030 ?

- Peu de modèles existants pour cette tâche.
- Disponibilité d'un nouveau jeu de données de haute qualité.
- Possibilité d'explorer plusieurs approches et architectures.

Proposition de méthodologie

Notre approche consiste à combiner différentes architectures, notamment des autoencodeurs et des transformers, afin d'effectuer la conversion des images et du son en texte.

Dans un premier temps, un prétraitement des données sera nécessaire. La grande dimensionnalité des données, due à leur haute qualité, nécessite une réduction. Nous envisageons de réduire la résolution des vidéos afin d'améliorer la généralisation des modèles. De plus, le fond vert présent dans le jeu de données pourra être changé pour augmenter les données.

Nous débuterons l'entraînement sur un sous-ensemble du jeu de données, ce qui nous permettra d'évaluer les architectures avant de passer à l'ensemble des données. Une fois la bonne architecture choisie, nous procéderons à l'entraînement sur l'intégralité du jeu de données.

Les données nécessaires à cette tâche sont accessibles via le site suivant : How2Sign