

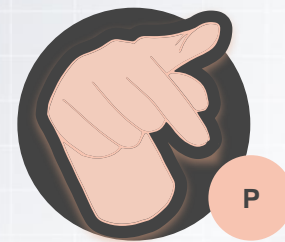
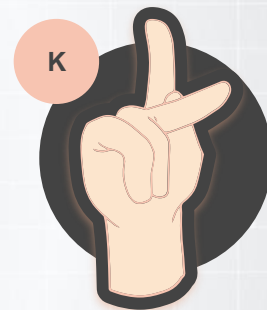
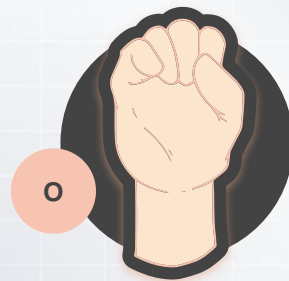
A 3D illustration of a grey rectangular block with the letters 'ai' in white, resting on a circular orange and yellow pedestal. The background features a grid pattern and a glowing orange cylinder.

RECONNAISSANCE DU LANGAGE DES SIGNES PAR ML EN TEMPS RÉEL

Halah Magri & Pr Moutacalli Mohamed Tarik
Département de mathématiques, informatique et génie, Université du
Québec à Rimouski

Problématique

La communication est essentielle pour tous, mais les personnes sourdes ou malentendantes sont souvent confrontées à des difficultés de communication. La reconnaissance de la langue des signes peut aider à surmonter ces barrières, mais il est souvent difficile de trouver des interprètes qualifiés. Dans ce contexte, la reconnaissance automatique de la langue des signes en temps réel peut offrir une solution prometteuse.

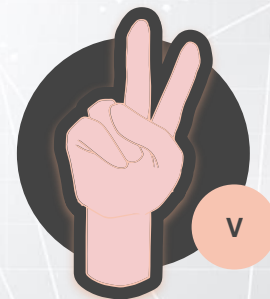


OBJECTIFS

L'objectif principal de ce projet est de développer un système de reconnaissance automatique de la langue des signes en temps réel.

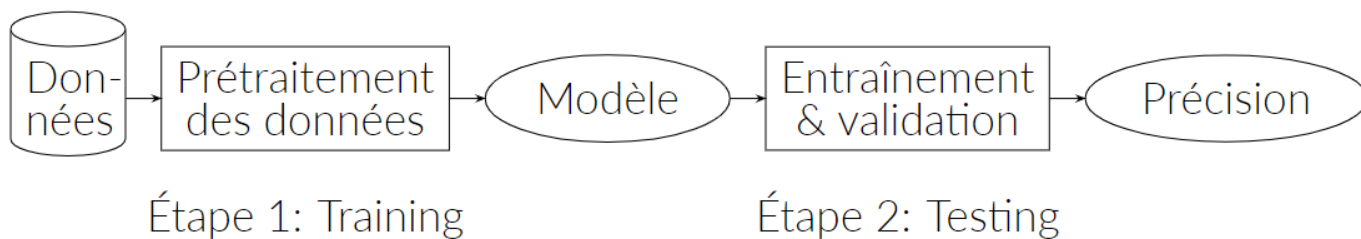
Pour atteindre cet objectif, nous avons fixé les objectifs spécifiques suivants :

- Extraire les points clés holistiques du MediaPipe
- Construire un modèle de langue des signes à l'aide d'une détection d'action alimentée par des couches LSTM
- Prédire la langue des signes en temps réel à l'aide de séquences vidéo



DESCRIPTIF DU SYSTÈME

Ce graphique présente les différentes étapes de notre projet. Le projet commence par la collecte et le prétraitement des données, où les données sont nettoyées et préparées pour le modèle. Le modèle est un réseau neuronal séquentiel qui est entraîné et validé sur les données prétraitées. Une fois le modèle formé, il est évalué pour mesurer ses performances.



MÉTHODOLOGIE SUIVIE

- Tout d'abord, nous avons extrait les points clés holistiques du MediaPipe pour obtenir une séquence de points clés qui représente les mouvements des mains, des bras et de la posture dans la vidéo. Nous avons fait appel à la bibliothèque Mediapipe Holistics.



(1) Image d'entrée



(2) Image de sortie

MÉTHODOLOGIE SUIVIE

- Ensuite, pour entraîner notre modèle, nous avons collecté et annoté un ensemble de données de langage de signes. Nous avons travaillé avec des étudiants de l'UQAR pour collecter des vidéos de personnes utilisant la langue des signes dans différents contextes. Nous avons ensuite annoté les vidéos en identifiant les mouvements de la main et du bras et en les associant à un mot ou à une phrase en langue des signes. Cette annotation a été utilisée pour entraîner notre modèle de détection d'action.
- Finalement, nous avons construit un modèle de détection d'action alimenté par des couches LSTM à l'aide de Tensorflow, Keras et skit-learn. Ce modèle est utilisé pour décoder la langue des signes à partir de la séquence de points clés.



ÉVALUATION

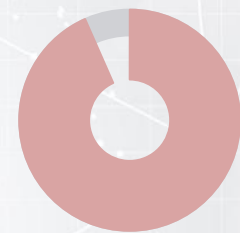
- Pour évaluer la performance de notre modèle, nous avons utilisé différentes mesures, notamment l'exactitude, la précision, le rappel et la mesure F1.
- Nous avons également utilisé des techniques de validation croisée pour tester la généralisation du modèle à de nouvelles données.
- Enfin, pour mesurer la performance en temps réel, nous avons utilisé des mesures telles que le taux de rafraîchissement des images et le temps de traitement de la séquence de points clés.

Matrice de confusion	Négative réel	Positive réel
Négative prédit	50	8
Positive prédit	2	25

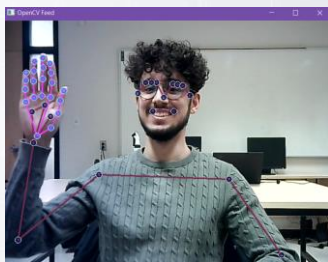
RÉSULTATS

Nous avons réalisé des tests préliminaires avec un petit ensemble de données et avons obtenu des résultats encourageants. Le système a réussi à reconnaître la langue des signes avec une précision de 88 %.

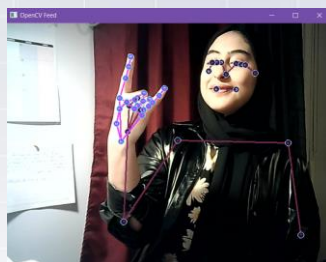
Les images suivantes illustrent les résultats de notre modèle sur différentes personnes.



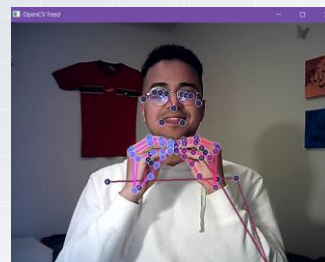
88%



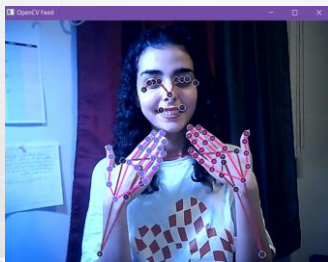
(a) Hello



(b) I Love You



(c) More



(d) Thank You



(e) Repeat

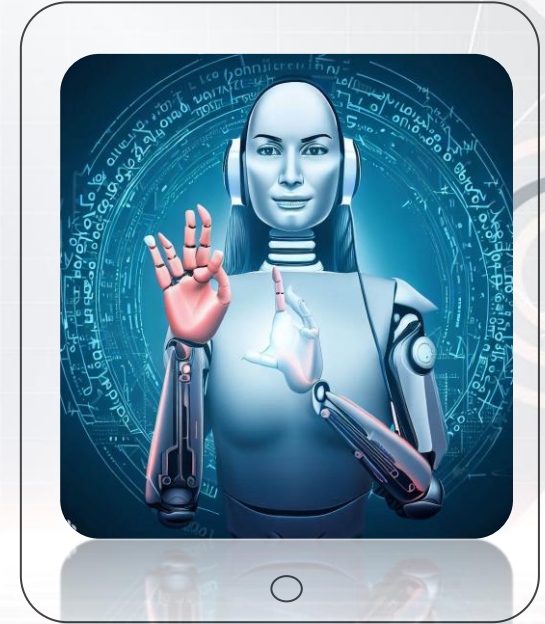


(f) Yes

CONCLUSION

Dans ce projet, nous avons développé un système de reconnaissance automatique de la langue des signes en temps réel à l'aide d'apprentissage automatique et de réseaux de neurones LSTM. Nos résultats préliminaires montrent que notre système peut être efficace pour aider à surmonter les barrières de communication pour les personnes sourdes ou malentendantes.

Cependant, nous reconnaissons que notre système doit être testé sur un ensemble de données plus grand et plus diversifié pour en évaluer pleinement les performances.



RÉFÉRENCES

- Zhang, J., Wang, J., & Lu, H. (2020). Real-time American Sign Language Recognition with MediaPipe. arXiv preprint arXiv:2012.05222.
- Graves, A., Mohamed, A. R., & Hinton, G. (2013). Speech recognition with deep recurrent neural networks. In Acoustics, speech and signal processing (icassp), 2013 ieee international conference on (pp. 6645-6649). IEEE.
- Geron, A. (2017). Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems. O'Reilly Media, Inc.