جامعة الأميرة
نورة بنت عبدالرحمن
كلية علوم الحاسب والمعلومات

واصف

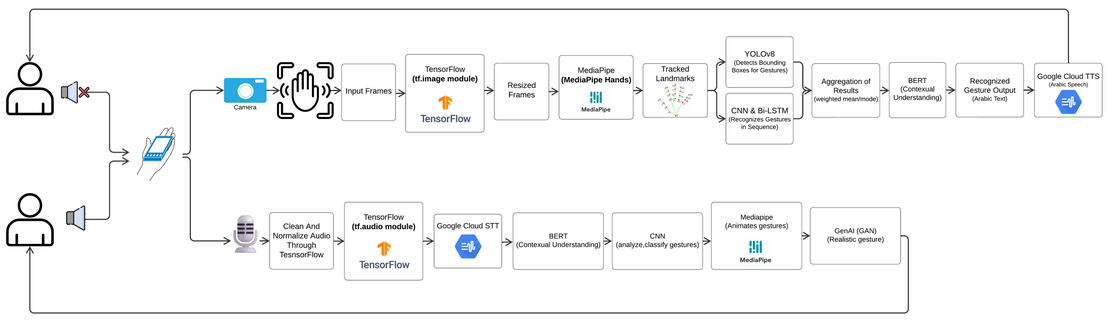
Mobile Application for Deaf and Mute individuals

Layan Almegbil, Abeer Alabdullatif, Rimas AlAqil, Haifa AlKhalaf
Supervisor: Dr. Eman AlDakheel

Abstract

Our System, Wasif, offers a communication application for deaf and mute individuals, we aim to develop cross-platform application that bridges these communication gaps by translating spoken language into sign language symbols and written text, while converting sign language gestures into audible speech. Wasif application analyzes hand and facial movements in real-time, ensuring accurate and efficient communication, built using robust frameworks which is TensorFlow, and OpenCV, seeks to empower deaf and mute individuals to fully engage in all aspects of life, it is have been validated through rigorous testing with target user groups, ensuring that it meets the real-world needs.

Methodology



Wasif captures video input and processes it through MediaPipe to extract hand landmarks, then uses YOLOv8 to detect gesture regionsm these are passed to a hybrid CNN + Bi-LSTM model for classification, and results are aggregated for higher accuracy.
For audio, the system cleans and normalizes the input using TensorFlow Audio, converts it to text via Google STT, and enhances context understanding using BERT.
The final output appears either as spoken Arabic through Google TTS, or as animated sign gestures using MediaPipe and Generative AI (GAN) for natural, expressive communication.

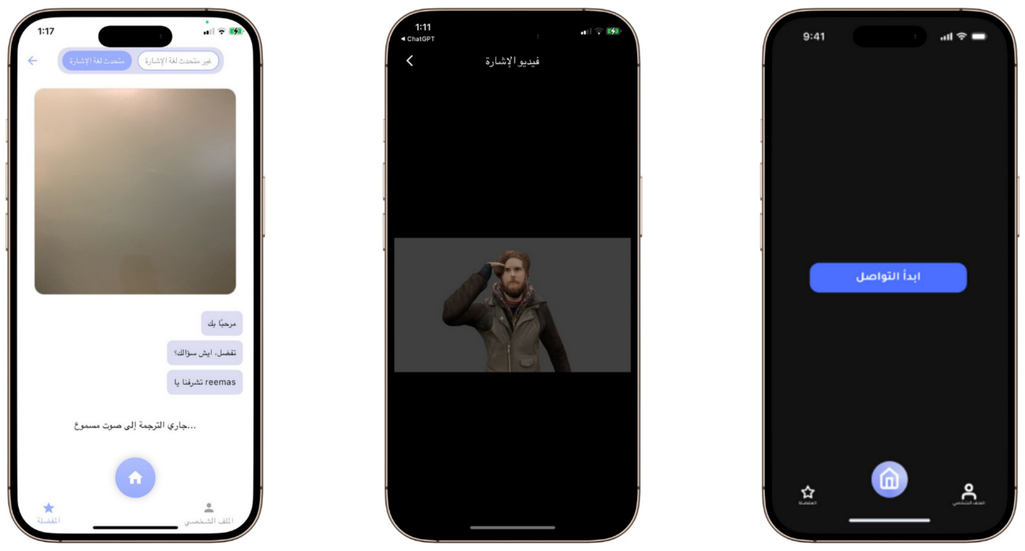
Classification Model

Accuracy	F1	Recall	Precision
97.96%	96.3%	96.2%	96.5%

The classification model in Wasif is built using a combination of CNN for spatial feature extraction and Bi-LSTM for temporal sequence recognition, enabling the system to accurately interpret both static and dynamic gestures, hand landmarks detected by MediaPipe are used as input features, and the model is trained on the full KARSL-502 dataset, containing 502 Arabic sign language classes.
After extensive training and testing, the model achieved a high results that reflect the system's high reliability in recognizing real-time gestures and ensuring smooth, responsive interaction.

System Design

The application of Wasif is built on a modular architecture that integrates real-time AI models with a mobile app developed using Flutter for both Android and iOS, the system processes user input through MediaPipe to track hand landmarks and uses YOLOv8 for gesture detection. A hybrid model of CNN and Bi-LSTM classifies gestures accurately in real-time.
For speech input, TensorFlow Audio cleans and processes the signal, which is converted to text using Google STT, then analyzed using BERT for contextual understanding. The translated output is presented as either Arabic speech using Google TTS or animated sign gesture.
To support visual output, Wasif includes a curated avatar video database, where each recognized word or gesture is linked to a pre-recorded video of a signing avatar. These videos are automatically fetched and displayed in sync with recognized speech or text, enabling deaf users to understand responses in visual sign language. This ensures accessibility and keeps the experience smooth and responsive.



Conclusion

Wasif successfully bridges the gap between deaf/mute and hearing individuals through accurate, real-time translation of gestures and speech, the app achieved high performance in gesture recognition and user interaction, with its intuitive design and reliable results, Wasif sets the foundation for a more accessible communication experience and holds strong potential for future development and expansion.