ishara
Arabic Continuous Sign Language Translation System

Ahmed Nasser , Amgad Salem , Abdulrahman Badr , Abdulrahman Dessoky , Asem Saber
Esraa Salem , Amany Karam , Aya Ragab

Dr/ Ibrahim Zaghloul

Faculty of Computers & Artificial Intelligence

Benha University

Abstract

Deaf people have big problem to communicate with all people especially people who don't know sign language

Our objective is to create a mobile and a web application that translate Arabic continuous sign language to text and vice versa to help deaf people and those of disability of not hearing and not speaking have a better communication with ordinary people.

Our solution is to build application that will offer a low-cost approach to translate Arabic continuous sign language into text and vice versa in real-time introducing ISHARA.

Our application offers a live translator and a chat conversation between the deaf and ordinary people to facilitate communication in real-world and remotely

Objectives

Our aim is to :

1- facilitate communication by providing tools and technologies that ease the communication between deaf individuals and those who do not use sign language.

2- reduce manual efforts required for deaf individuals to access information and engage with others by leveraging technology to automate and streamline communication processes.
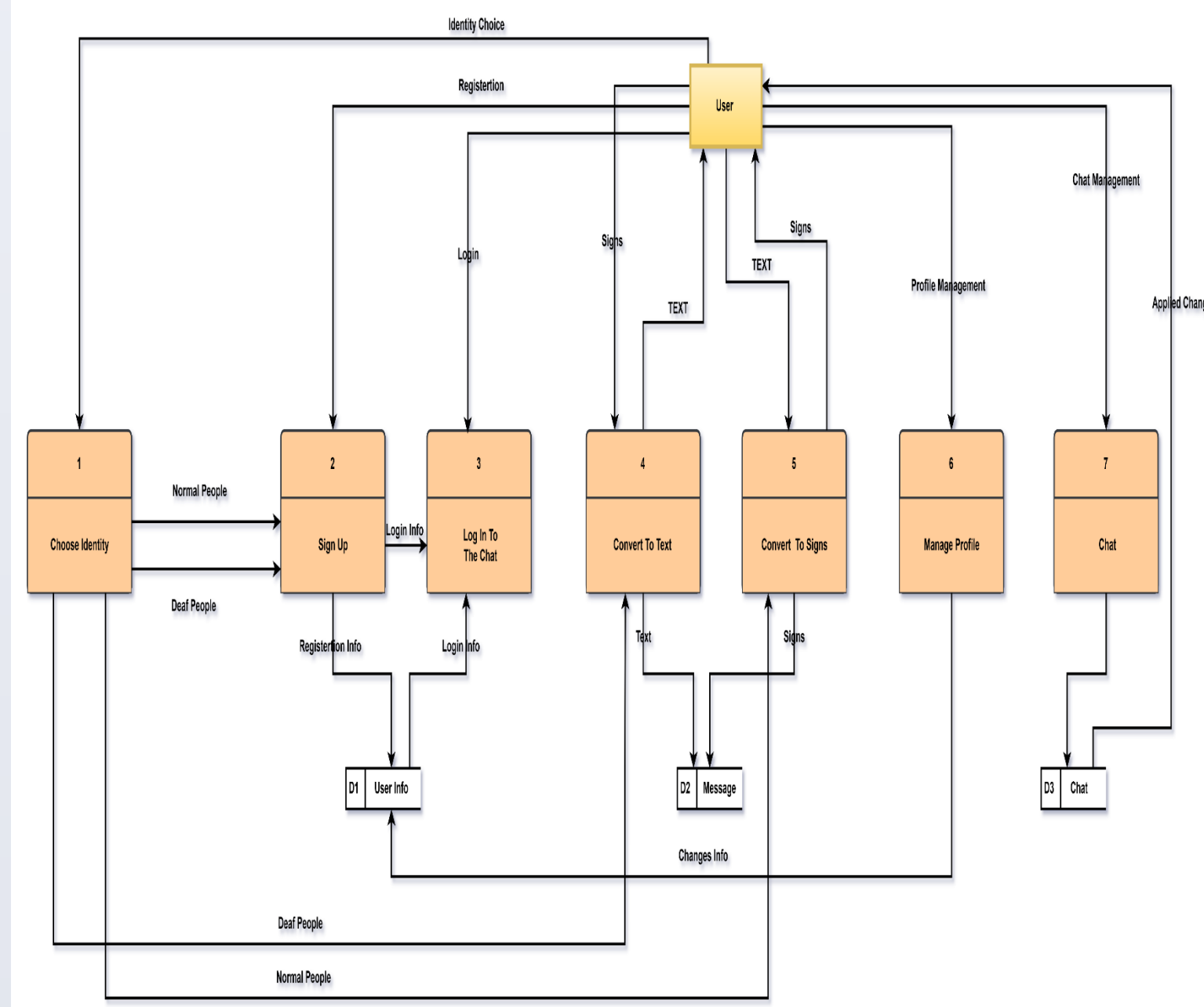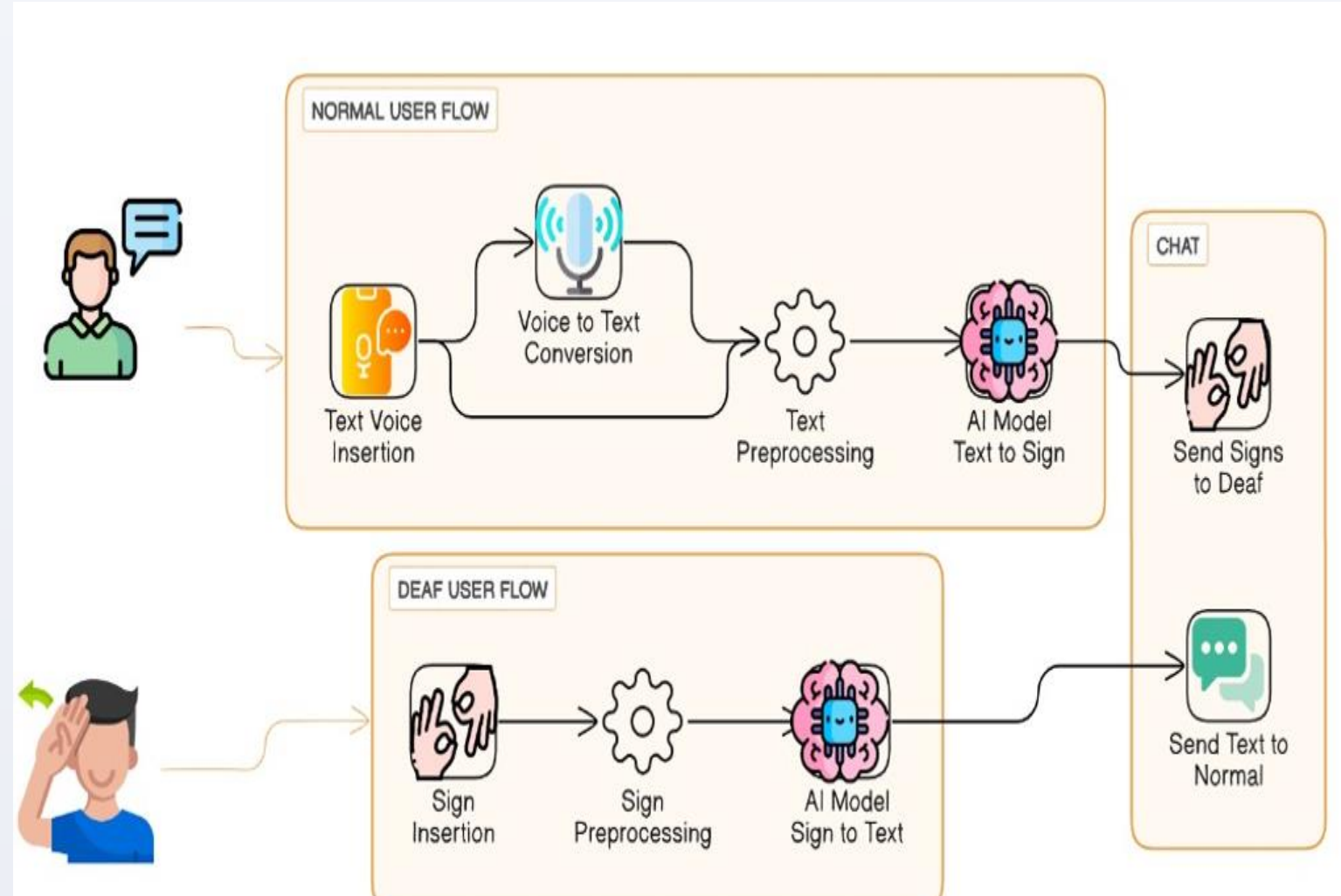
Methods

This project seeks to make it easy to communicate with deaf people and to make it easier to understand them quickly, the way which the project works is to convert Arabic sign language to text and vice versa.

The proposed system is a chat application ISHARA chat application

ISHARA Logo

Data Flow Diagram DFD (Level 1)

References

- **Mediapipe** : framework developed by Google gives solutions to pose estimation technique
  (face , pose (body) , hands).

- **Keras ASR** : Model developed by Keras used for audio speech recognition ,its architecture inspires us to get our ISHARA model.

- **JUMLA-QSL-22** : A Dataset of Qatari Sign Language Sentences.

Tools

python   MediaPipe

TensorFlow

MoviePy   Flutter

node   React

OpenCV   PyTorch

Results

Sign2Text : We implement transformer from scratch converting continuous sign language video into text using pose estimation technique to extract motion of body and facial features (expressions) Model Trained on Arabic datasets which are Qatarien dataset (JUMLA) and on Egyptian dataset (ISHARA)
these are the results scores using WER & BLEU metrics

Model	Dataset Size	Batch Size	WER Score (Word Error Rate) ↓ Better 0% – 100%			BLEU Score (Bilingual Evaluation Understudy) ↑ Better 0 – 1		
			Train (80 %)	Val (10 %)	Test (10 %)	Train (80 %)	Val (10 %)	Test (10 %)
Qatarian Model	6300 Samples	32	0.83 %	2.38 %	17.05 %	0.9933	0.9892	0.8374
Egyptian Model NO.1	600 Samples	32	0 %	13.68 %	31.90 %	1.0	0.8802	0.7308
Egyptian Model NO.2	600 Samples	64	0 %	24.69 %	24.69 %	1.0	0.7958	0.7866
Egyptian Model NO.3	600 Samples	32	0 %	12.71 %	18.18 %	1.0	0.8752	0.7772

Qatar Sign Video → مرحبا Continuous Text
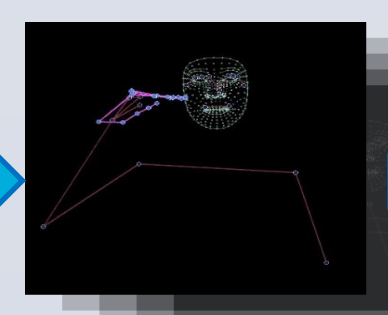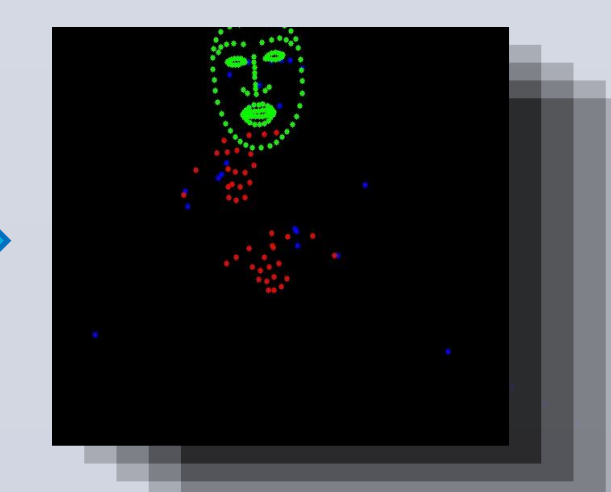
Egyptian Sign Video → Pose Estimation Features → انا اسمي حسام حسن Continuous Text

Text2Sign : Model trained for converting continuous Sentence to Pose video
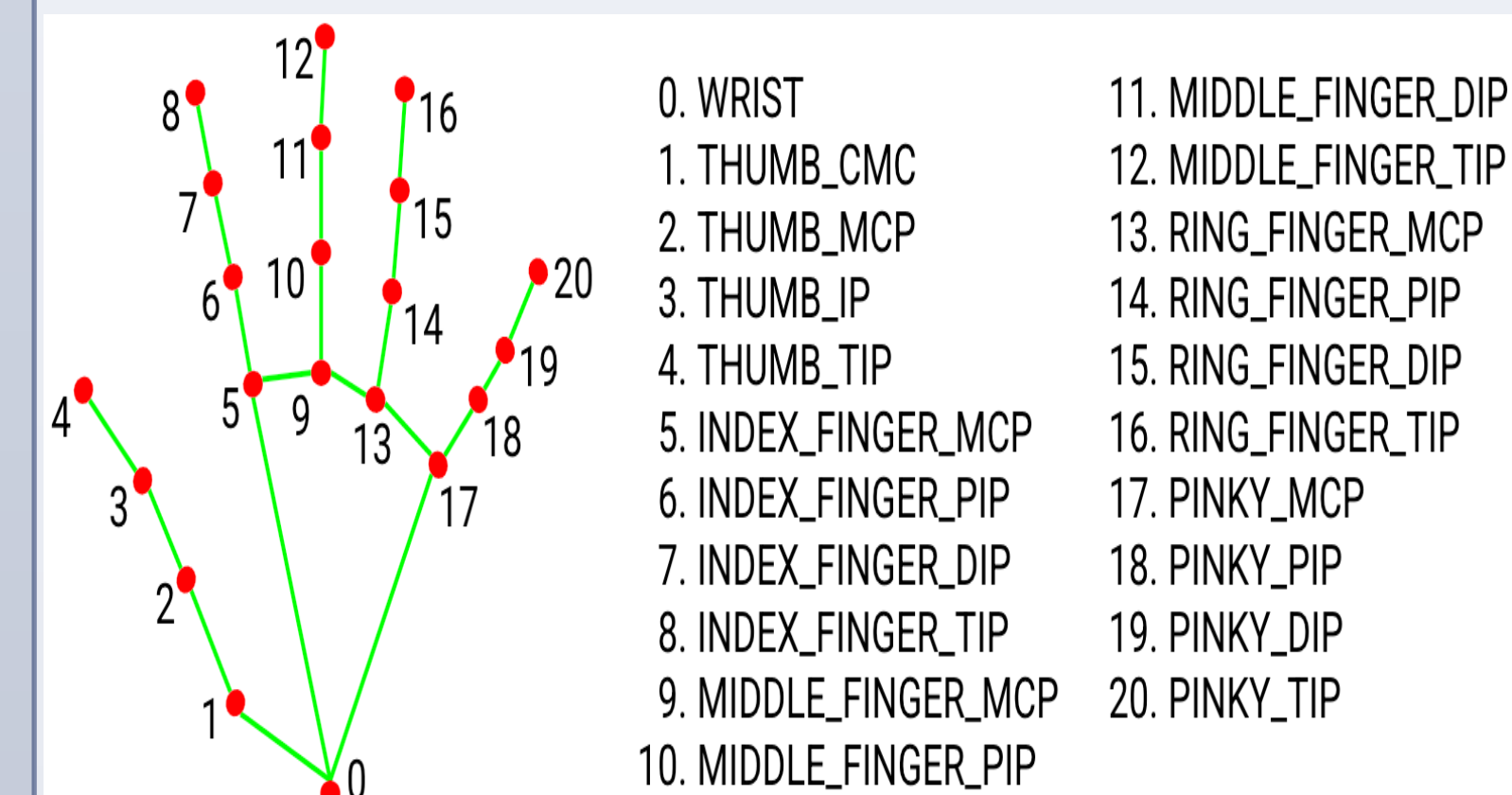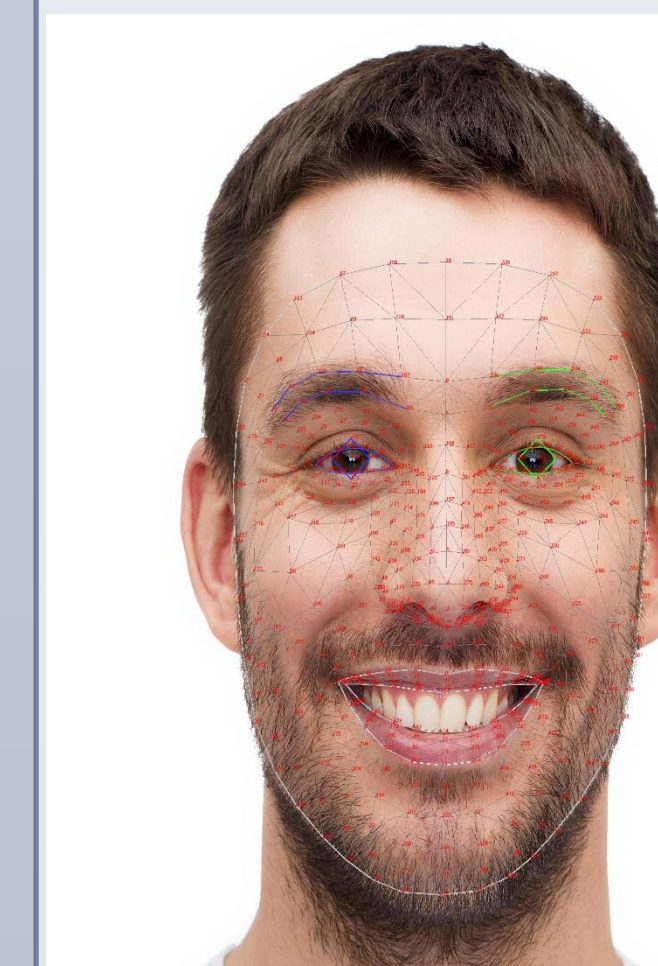
انا طالب → Pose Video (Avatar)

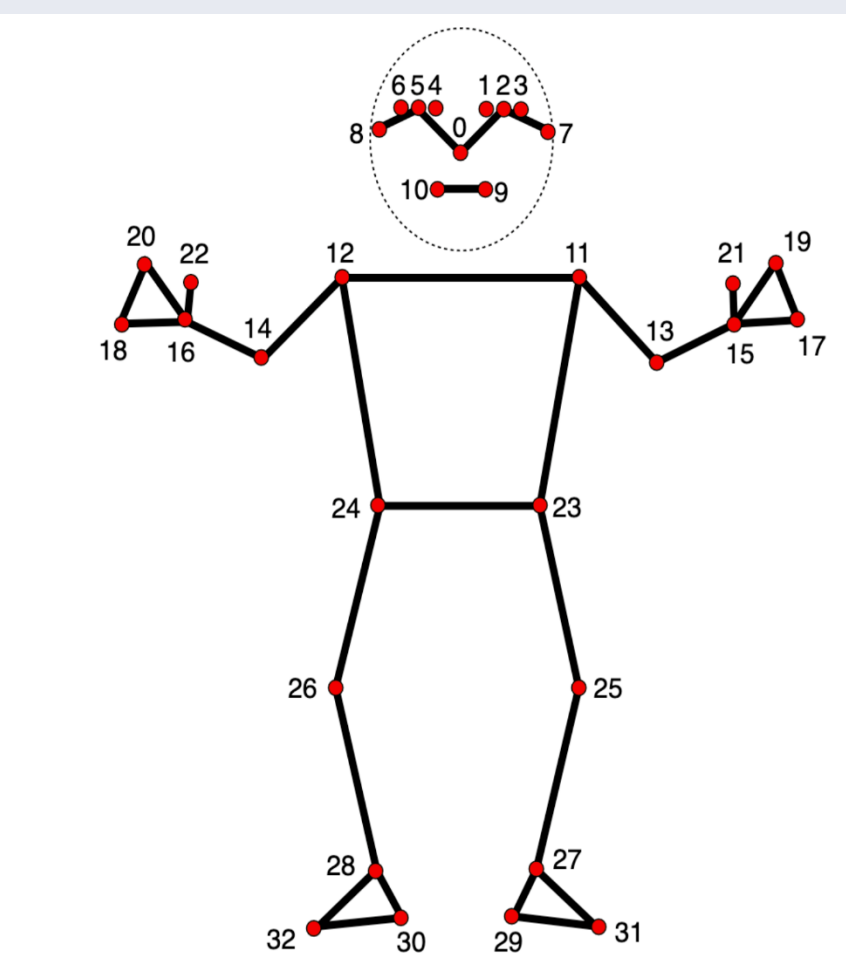Continuous Text

Conclusions

Our idea to develop application used by deaf people and ordinary people (talking) to facilitate communication between them using pose estimation through Mediapipe framework to extract landmarks (features) for pose of body , hands and face mesh (all of them are important for understanding sign language).

0. WRIST	11. MIDDLE_FINGER_DIP
1. THUMB_CMC	12. MIDDLE_FINGER_TIP
2. THUMB_MCP	13. RING_FINGER_MCP
3. THUMB_IP	14. RING_FINGER_PIP
4. THUMB_TIP	15. RING_FINGER_DIP
5. INDEX_FINGER_MCP	16. RING_FINGER_TIP
6. INDEX_FINGER_PIP	17. PINKY_MCP
7. INDEX_FINGER_DIP	18. PINKY_PIP
8. INDEX_FINGER_TIP	19. PINKY_DIP
9. MIDDLE_FINGER_MCP	20. PINKY_TIP
10. MIDDLE_FINGER_PIP	

Hands Landmarks

Face Mesh Landmarks   Pose Landmarks

We used JUMLA Dataset as standards to collect the Egyptian Dataset ISHARA.

Implementing Transformer from scratch based on Keras ASR architecture with some modifications to handle sign videos after extraction of landmarks from them we get our Sign2Text model then Text2Sign model using GAN.
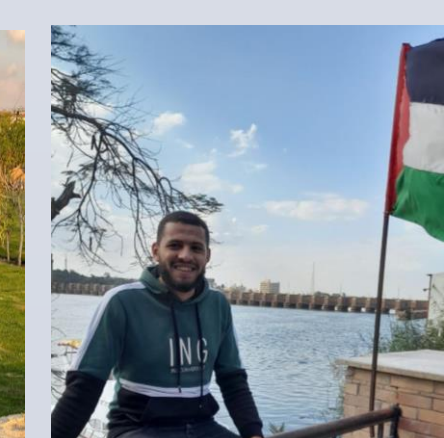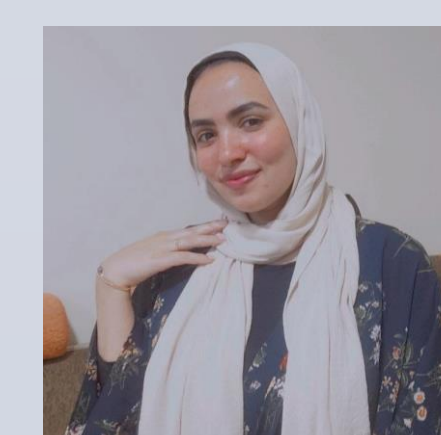
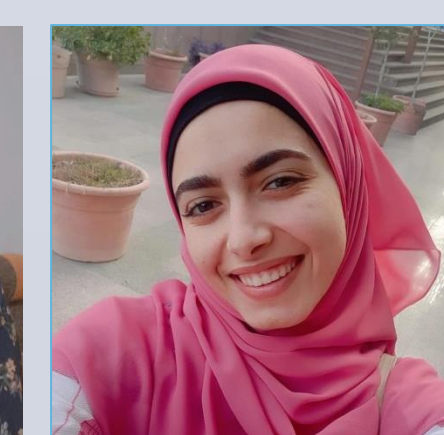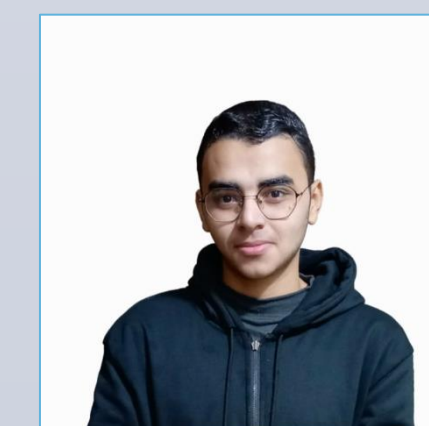## Team Members

Ahmed Nasser   Abdulrahman Dessoky   Abdulrahman Badr

Esraa Salem   Amany Karam   Aya Ragab

Amgad Salem   Asem Saber