VisuoSpeech Intelligence (VSI)

Team Members:

Mohamed Taher Mohamed

Eyad Wael Wagdy

Mark Edwar Shawky

Mohamed Gomaa Mohamed

Abdelrahman Khaled Abdalla

Project Idea

- **The project aims to develop an intelligent system capable of recognizing individuals through a camera, then listening to their speech in Arabic and converting it into written text.**
  After that, the system analyzes the spoken content using the Wikipedia knowledge base to generate an automatic, intelligent response that provides accurate information about the discussed topic.
- **This idea integrates vision and hearing into a single system that can interact with users in a natural, human-like way.**

Practical Applications of VSI (VisuoSpeech Intelligence)

- **Customer Service**
  VSI can monitor customer interactions through voice and facial expressions, understand their inquiries automatically, and respond with accurate information — reducing response time and improving customer satisfaction.
- **Voice-based Education**
  Used as a smart classroom assistant, VSI listens to students' Arabic questions and explains answers using Wikipedia, enabling interactive and self-directed learning experiences.
- **Healthcare and Emergency Rooms**
  In hospitals or clinics, VSI helps patients who cannot read or write communicate verbally with the system.
  It recognizes the patient's face, listens to their speech, and retrieves helpful medical information from Wikipedia — giving doctors a quick understanding of the case before examination.
- **Smart Help Desks**
  In universities, airports, or offices, VSI can answer visitors' spoken questions like *"Where is the registration office?"* using voice and facial recognition.
  It replies instantly with information retrieved from Wikipedia or an internal database — providing fast, human-like assistance without waiting for staff.

Tools & Technologies

- **TensorFlow → Gives the system AI power for face detection and recognition.**
- **OpenCV-Python → Handles the camera, captures video, and processes images.**
- **OpenCV-Contrib-Python → Adds extra tools for advanced face recognition.**
- **MTCNN → Detects human faces accurately from live video.**
- **SpeechRecognition → Listens to voice and converts speech into text.**
- **PyAudio → Connects the microphone for live audio input.**
- **Wikipedia → Provides smart answers using real online knowledge.**
- **Pandas → Manages and organizes speech and camera data in tables.**
- **Streamlit → Builds the interactive user interface for the VSI app.**
- **NumPy → Supports fast math and image operations in the background.**

Results & Impact

The system can automatically recognize speakers through the camera.

It analyzes voice input and converts Arabic speech into accurate text.

VSI uses Wikipedia to generate instant, informative answers.

Results such as voice activity and spoken sentence counts are displayed in an interactive visual interface.

Practical Applications

The system can be used in:

- **Smart offices and training centers**
- **Educational assistants**
- **Human-interactive robots**
- **Arabic voice-based customer service systems**

Final Statement

VSI represents a step toward combining visual and auditory intelligence into a single system that interacts with humans naturally, just like people do with each other.

Stop Deploy

Controls

Pause / Resume Camera

Show Speech Analysis

VisuoSpeech Intelligence (VSI)

CAMERA

ID 1

[3:11:04 AM] Said: من هو محمد صلاح

[3:11:04 AM] Wikipedia: محمد صلاح حامد محروس علي (مواليد 15 يونيو 1992) هو لاعب كرة قدم مصري، يلعب في مركز الجناح الأيمن مع نادي ليفربول في الدوري الإنجليزي الممتاز ومنتخب مصر. يُعد أحد أبرز اللاعبين العرب والأفارقة، حيث حصد العديد من الجوائز أبرزها جائزة أفضل لاعب في إنجلترا موسم 2017-18، و 2021-22، و 2024-25، و جائزة أفضل لاعب في إنجلترا من اتحاد كتاب كرة القدم موسم 2017-18، و 2021-22، و 2024-25، و جائزة هدف الموسم في الدوري الإنجليزي الممتاز لموسم 2017-18، و جائزة الحذاء الذهبي في الدوري الإنجليزي الممتاز (هداف الدوري الإنجليزي الممتاز) أعوام 2018، 2019 و2022، و جائزة أفضل لاعب أفريقي لعامي 2017 و 2018، جائزة أفضل لاعب أفريقي بواسطة بي بي سي لعامي 2017 و 2018، أصغر من قبل مجلة دايم الأمريكية في عام 2019 ضمن أكثر 100 شخصية تأثيرًا في العالم.

Stop Deploy

Controls

Pause / Resume Camera

Show Speech Analysis

Speech Analysis

Speech Count Table

	ID	Sentences Count
0	2	5
1	4	3
2	1	1
3	3	0

Speech Activity by ID

Live Feed and Speech Recognition