UNIVERSITY OF ENGINEERING & MANAGEMENT
Good Education, Good Jobs

IEM CENTRE OF EXCELLENCE FOR DATA SCIENCE

INSTITUTE OF ENGINEERING & MANAGEMENT
Good Education, Good Jobs

COMPUTER VISION HACKATHON

ACCESSIBILITY THEME

PROJECT ANDHADHUN

Team Name - camel_case

Team Members -

Arnab Sengupta (Lead)
Sagnik Basak
Sachindra Kumar Singh
Anidipta Pal
Ritesh Das
Tamojit Das

Problem Statement

The problem statement is to develop computer vision applications to assist individuals with disabilities, particularly the visually and hearing impaired. Solutions include real-time object recognition for enhanced navigation and automated sign language interpretation for improved communication. These tools aim to promote independence, accessibility, and inclusivity in everyday life.

Why is this problem important?

- **Inclusivity**: Individuals with disabilities often face barriers in navigating everyday tasks. Developing technologies like object recognition or sign language translation fosters independence and autonomy.
- **Technological Gap**: Although advancements in AI are widespread, accessibility technology has not kept pace with the needs of disabled communities. Closing this gap can improve the quality of life for millions of people.
- **Social and Economic Impact**: Enhancing accessibility promotes equal opportunities in education, employment, and social interactions, allowing disabled individuals to fully participate in society.

Who is affected by this problem?

- **Visually Impaired Individuals**: They struggle with tasks like recognizing objects, reading text, or navigating unfamiliar environments. Real-time object recognition and guidance systems can significantly help.
- **Hearing-Impaired Individuals**: Communicating with those who don’t understand sign language is a significant challenge. Automated sign language interpretation can bridge communication gaps.

Solution Overview

Our solution enhances the lives of individuals with disabilities through two key components: *real-time object recognition* for the visually impaired and *automated sign language interpretation* for the hearing impaired.

1. **Real-time Object Recognition :**
   - This system uses advanced computer vision to identify and describe objects, providing audio feedback via wearable tech to assist navigation and interaction.
   - Impact: Empowers visually impaired users to navigate confidently and independently.
2. **Automated Sign Language Interpretation :**
   - Employing machine learning, this tool translates sign language into spoken or written text in real-time, facilitating communication between hearing-impaired individuals and others.
   - Impact: Bridges communication gaps and fosters inclusivity.

Expected Impact

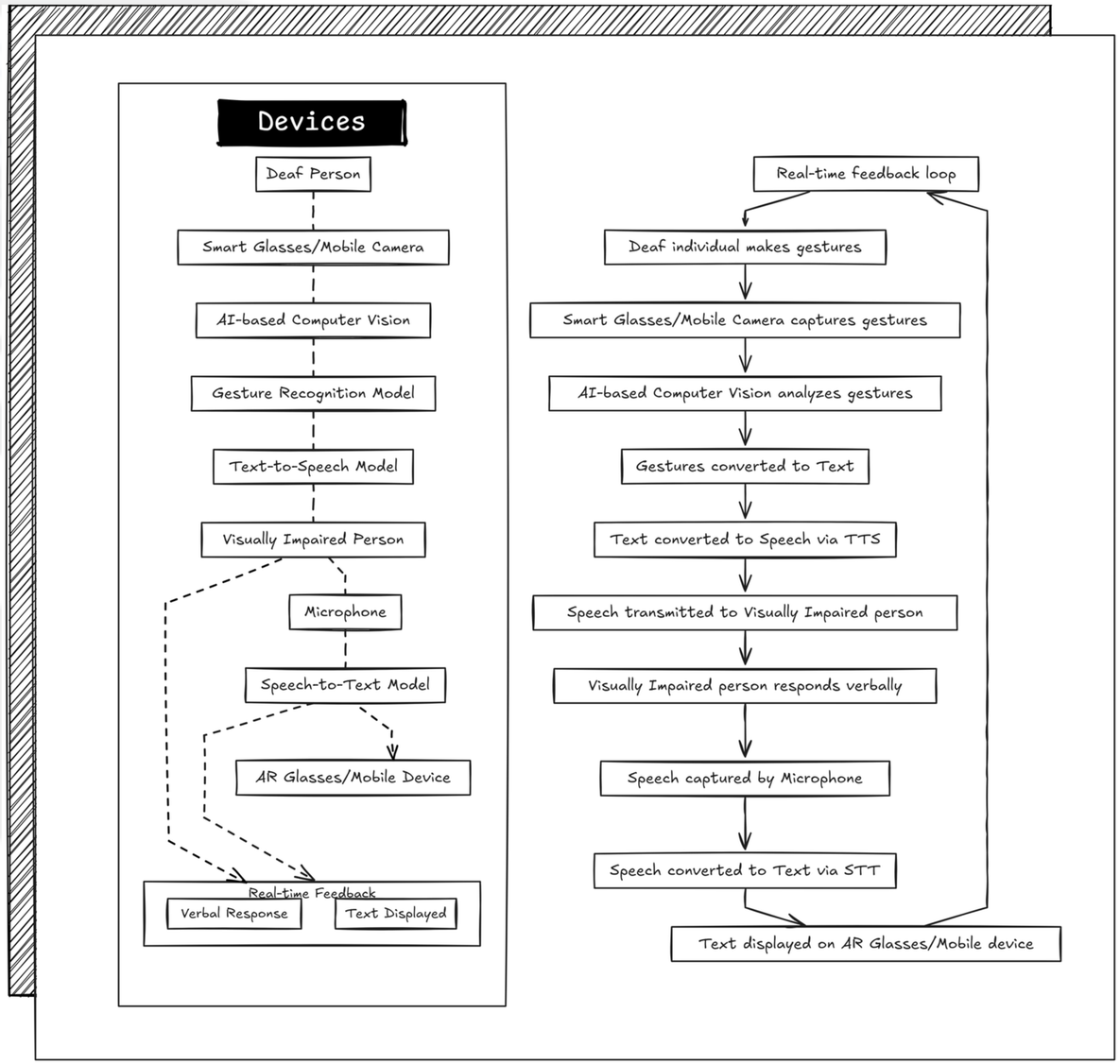
- **Increased Autonomy**: Enhances independence and quality of life.
- **Greater Accessibility:** Promotes innovation in assistive technologies.
- **Social Equity:** Facilitates equal opportunities in education and employment.
- **Awareness and Empathy:** Raises awareness of the challenges faced by individuals with disabilities.

Tools and Technology Used

- **TensorFlow & PyTorch**: Frameworks for developing machine learning models for object recognition and sign language interpretation.
- **OpenCV & Mediapipe**: Libraries for real-time computer vision tasks, enabling object detection and tracking.
- **Google Cloud Vision API**: Provides robust image analysis capabilities for object recognition and scene understanding.
- **Kinect and LiDAR Sensors**: Used for depth perception and spatial awareness in navigation aids for the visually impaired.
- **Android/iOS SDKs**: Facilitates mobile application development for real-time communication and navigation solutions.
- **AR Glasses (Microsoft HoloLens)**: Integrates augmented reality for enhanced feedback and user guidance in real-world environments.
- **WebRTC**: Enables real-time communication for live sign language interpretation, bridging gaps between users and non-signers.

Architecture/Workflow

- **Mediums/Devices Proposed/Use Cases:**
  - Deaf Person: Communicates through gestures.
  - Visually Impaired Person: Communicates verbally.
  - Smart Glasses/Mobile Camera: Captures gestures.
  - AR Glasses/Mobile Device: Displays text and provides audio feedback.
- **Communication Workflow:**
  1. **From Deaf to Visually Impaired**:
     - *Gesture Capture*: Gestures by deaf person captured by smart glasses/mobile camera.
     - *Gesture to Text*: AI-based computer vision and gesture recognition model convert gestures to text.
     - *Text to Speech*: Text converted to speech via TTS and transmitted to the visually impaired person.
  2. **From Visually Impaired to Deaf**:
     - *Speech Capture*: Verbal response captured by microphone.
     - *Speech to Text*: Speech converted to text via STT.
     - *Text Display*: Text displayed on AR glasses/mobile device for the deaf person.
- **Real-time Feedback Loops:**
  - Ensures accuracy in gesture recognition and speech translation, allowing for immediate conversational adjustments.



Feasibility & Viability

Feasibility:

- **Technical Readiness**: The system has been optimized through model pruning and data augmentation, making it suitable for real-time use on mobile and AR devices.
- **Infrastructure:** Existing cloud and mobile hardware infrastructure ensures seamless data processing and integration, making deployment technically feasible.
- **User Interface**: The user-centered design process has resulted in an intuitive and user-friendly interface, increasing initial acceptance and ease of use.

Viability:

- **Scalability**: The system is built to scale across various industries, including healthcare, education, and assistive technologies, offering flexibility for widespread adoption.
- **Continuous Improvement**: Regular data updates and feedback from users will enable retraining and refinement of AI models, ensuring the system stays up to date and effective over time.
- **Adoption and Integration**: Collaborating with healthcare providers, educational institutions, and tech firms will expand the system's reach and encourage industry-wide adoption.
- **Sustainability:** Adapting to advancements in AI, AR, and mobile technologies while addressing ethical and privacy concerns will be crucial for long-term success and trust.

Future Work

- **Enhanced AI Models:**
  - **Goal**: Increase accuracy and responsiveness.
  - **Approach**: Implement state-of-the-art neural networks and transformer-based models for more precise gesture recognition and speech synthesis.
  - **Impact**: Improved real-time communication with reduced errors, enhancing the user experience.
- **Language Support:**
  - **Goal**: Make the system accessible on a global scale.
  - **Approach**: Expand system capabilities to support multiple languages and dialects, including region-specific variations.
  - **Impact**: Enables broader adoption and usability across different cultural and linguistic backgrounds.
- **Augmented Reality Enhancements:**
  - **Goal**: Enrich interaction with intuitive features.
  - **Approach**: Integrate augmented reality technologies to provide real-time, contextual prompts and visual aids during interactions.
  - **Impact**: Facilitates smoother and more natural conversations, aiding comprehension and response times.
- **User Personalization:**
  - **Goal**: Tailor the system to individual user needs.
  - **Approach**: Develop adaptive AI models that learn and adjust to the unique communication styles, gestures, and preferences of each user.
  - **Impact**: Personalized user experience increases effectiveness, comfort, and satisfaction, encouraging longer-term use.

SWOT Analysis

<-- Strength -->

1. Innovative Communication Bridge: The project effectively bridges communication gaps between deaf and visually impaired individuals, a solution not widely available in existing assistive technologies.
2. Real-time Interaction: The system's real-time feedback loop allows for immediate responses, enhancing the natural flow of conversation between users.
3. Multi-Modal Integration: Combining AI-based computer vision, gesture recognition, TTS, and STT provides a comprehensive approach to communication, making the system versatile and effective.

<-- Weaknesses -->

1.Accuracy Limitations: Variability in gesture recognition and speech-to-text accuracy, especially in challenging conditions (e.g., noisy environments, low light), can impact the system's effectiveness.
2.High Computational Demand: The AI models, particularly for gesture recognition and TTS, require significant computational resources, which may limit performance on lower-end devices.
3.Dependence on Devices: The reliance on AR glasses or mobile devices for operation could be a barrier for users without access to such technology or those who are not tech-savvy.

<-- Opportunities -->

1.Market Expansion: There is significant potential for market expansion, particularly in sectors like healthcare, education, and public services.
2.Technology Enhancement: Advances in AI, such as improved gesture recognition models and speech processing, could be integrated to enhance system performance and accuracy.
3.Personalization Features: Developing personalized AI models that learn and adapt to individual users' gestures and speech patterns can greatly improve the user experience.

<--Threats-->

1.User Adaptability and Accessibility: Adapting to AR devices can be difficult for non-tech-savvy users, and high costs may hinder access for low-income groups.
2.Data Privacy, Security, and Maintenance: Ensuring privacy and security of real-time data is critical, and maintaining AI models requires continuous updates and resources.
3.Environmental and Operational Constraints: Performance can be impacted by factors like lighting, noise, and background activity, affecting gesture and speech recognition accuracy.

Links:

Github Link :

https://github.com/SagnikBasak04/Andhadhun

CONCLUSION

- **Empowerment Through Technology:** Real-time object recognition and automated sign language interpretation enhance autonomy for visually and hearing-impaired users.
- **Innovative Solutions:** Leverages advanced computer vision and machine learning to address accessibility challenges in daily life.
- **Increased Independence:** Users gain confidence in navigation and communication, significantly improving quality of life.
- **Closing the Technological Gap:** Addresses the lag in assistive technology advancements, promoting broader accessibility.
- **Social Equity Enhancement:** Facilitates equal opportunities in education and employment for individuals with disabilities.
- **Raising Awareness:** Promotes understanding of the challenges faced by disabled individuals, fostering a more inclusive community.