Mothusi

**Introduction**

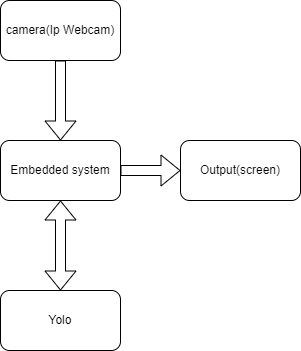
Recent advancements in embedded systems, AI, and IoT have the potential to greatly improve the quality of life for individuals with disabilities. This project focuses on creating **Mothusi**, an AI-powered assistive smart assistant designed to support people with disabilities by recognizing objects in their environment and responding to voice commands with real-time feedback. **Mothusi** is intended to assist individuals with visual impairments or motor challenges, allowing them to engage with their surroundings using a camera, microphone, and speaker. The system will leverage pre-trained machine learning models for object recognition and speech processing, ensuring it remains efficient and responsive for real-time assistance.

Milestone 1

System Requirements:

1. Functional:
   1. The system shall be able to capture pictures from video in real-time.
   2. The system shall be able to classify objects within the images.
   3. The system shall be able to capture voice to initialize or disable command mode and object detection mode.
   4. The system shall be able to process commands and give feedback through speakers in real-time.
   5. The system shall allow seamless integration of mobile I/O devices.
2. Non-Functional:
   1. The system shall be able to detect in real-time with minimal delay in voice commands.
   2. The system will accurately detect wide range of objects in various environments and be robust to different accents and noise levels.
   3. The system shall require minimal setup making it easy to use
   4. The system shall be scalable
   5. The system shall be reliable
   6. The system shall be portable
   7. The system shall process data locally and not storing or transmitting personal data for security.

Initial Design:



Narration:

IP webcam app will be installed on the user’s phone, which will turn it into a network camera which will be used to send the taken video to the embedded system by sharing its server’s Ip address with it, hence the real time detection of objects. The YOLO model will then detect different object images and give feedback to the embedded system which will give output of an image to the screen, in this milestone.