SMART ASSISTIVE DEVICE FOR SPECIALLY ABLED PEOPLE USING YOLO ALGORITHM

Mrs.SRI DEVI K[1], AADHANA NEYA V[2], SRIMATHIMEENAL K V[3], SHREEN TAJ N [4], SIBI SHREE M.[5]

1Assistant Professor,2, 3, 4,5 UG Students,

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

SRI SHAKTHI INSTITUTE OF ENGINEERING AND TECHNOLOGY,

COIMBATORE, INDIA.

Abstract

The proposed invention has a solution for deaf and dumb people those who are facing problems in communication among themselves as well as with the normal people. A clever supportive system for especially abled individuals is akin to possessing intelligent aid leveraging technology to simplify the lives of those with particular requirements. It encompasses items such as gadgets and detectors that are operable through voice commands or applications customized to specific needs. This system is crafted to offer additional assistance and elevate the everyday experiences of individuals confronting distinct challenges. This paper explores the symbiotic relationship between Artificial Intelligence (AI) and data-driven decision making, as embodied in the discipline of Machine Learning (ML). ML serves as the conduit through which computers assimilate data, infer patterns, and generate predictive insights, thereby facilitating informed decision making. By elucidating the foundational principles and methodologies of ML, this paper aims to delineate its transformative potential in various domains. Through a nuanced examination of ML's role in analyzing data and making predictions, this paper underscores its pivotal significance in contemporary computational paradigms. Here we have used YOLO algorithm for image detection.

Keywords-Deaf, Dumb, YOLO, Machine Learning, Artificial Intelligence, Data Analysis, Predictive Modeling, Decision Making

INTRODUCTION

This Project is about creating an effective system for Deaf and dumb people. The main purpose of this project is that the specially abled people can communicate and interact with normal people. This system interprets gestures, touch and sound. We are mainly aiming to build a friendly and inclusive environment that everyone can communicate and interact with each other. In a global landscape were around 9.1 billion individuals struggling with deafness and muteness, the dynamics of daily communication present formidable obstacles for both the mainstream and disabled communities.

The intricacies of sign language, characterized by elaborate patterns rooted in body language and arm movements, introduce complexities in interactions for those who are deaf and vocally impaired. The challenge extends beyond the acquisition and utilization of improvised sign language; the paramount concern lies in their capacity to engage with the lager society, given that not everyone possesses an understand of this specialized mode of communication. As a result, president communication gaps endure among the deaf and mute, leading to potential misunderstanding in the conveyance of messages through sign language. In response to the communication challenges encounters by individuals with visual and vocal impairments, an innovative application has been developed. This solution is designed to meet the specific needs of those who are deaf and mute. For individuals facing visual impairment, the application utilizes tesseract software to transform images into spoken words, thereby providing an auditory representation of textual information. Deaf individuals benefit from receiving spoken message in a text format as soon as they are uttered, eliminating the necessity for sign language interpretation. Additionally, mute individuals can convey messages through text, avoiding the complexities of sign language and these messages are delivered using the espeak functionality. This application presents practical measures to surmount the distinctive communication obstacles encountered by individuals with visual and vocal disabilities.

LITERATURE SURVEY

Recent literature has reported very few works on the assistive device for specially abled people due to the miscommunication problems between the people. YOLO algorithm has been employed in image detection. The state of art algorithm tries to detect the sign gestures from the whole input images. Incase if the signs are not clear or shown in a proper way then these algorithm fail to detect the gestures further. In order to avoid such cases, a YOLO algorithm can be used to detection the sign gestures from the region of Interest (ROI). The main algorithm of this is to communicate with the deaf and dumb people. [1] Sharana Basavaraj B et al. proposed a paper that uses Tesseract Software to allow people to understand words and the e-speak will read out loud the vocally impaired people and convey the messages through the text. This idea of the paper is very much portable and small in its size and it is less in weight. But the major drawback of this invention is it is quite expensive.[2] Sharana Basavaraj B et al. has suggested the concepts of Machine Learning Techniques such as Convolution Neural Network and Recurrent Neural Network to help the deaf and dumb people. It helps the people in a variety of context in a more natural way and its Accuracy is less. [3] B.S.Anirudh Bhargav et al. has proposed a technique for a blind person to read a text: a text is converted to speech(TTS) . It also provides a deaf people to read a text by speech to text (STT) conversion technology. Tesseract OCR (online character recognition) is used for blind, the dumb people can communicate through the text and gestures. This aims at providing single application that is needed for deaf, dumb and blind people. The device is more portable and compact but the major drawback is the time delay.[4]This paper uses Tessaract and Online Character Recognition (OCR) to help the deaf and dumb. It is a compact device that can work stand-alone but the accuracy of the image recognition is low. [5]

S.Kumuda et al. has used LabVIEW Software to build the algorithm for deaf, dumb and blind. This idea of paper is stand-alone device and it is portable. But it is not cost efficient. [6]Dhaya Sindhu Battina et al. has proposed a AI voice based smart device to guide the deaf people, here for this model they have used SLR techniques use Sequence-based Machine learning. The model is portable and compact but the drawback is it doesn't works on all languages of people.

The contribution of this paper is:

- Recognizing the speech words and converting it in to sign gestures.
- Recognizing the sign gestures into text(complete word).

After this section the paper is structured as follows. The proposed technique for Assistive device is introduced in section II. The working of the model is explained in Section III. Section IV explains about the experiment results achieved by this model and the summarization of the work along with it the future scope are detailed in Section IV.

II. PROPOSED METHODOLOGY

By considering the two ways conversion:

- Voice recognition to sign language
- Sign gestures to text format

The sensor output is converted into a diagram form using Raspberry pi and after processing the Data audio output is generated using gTTs.

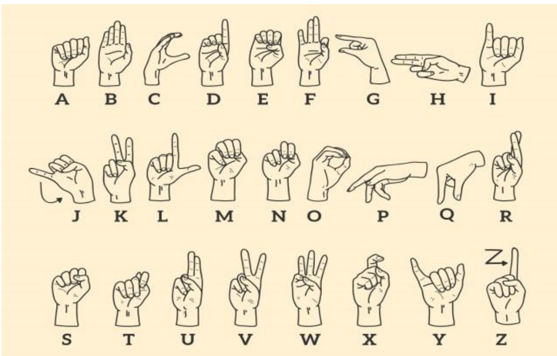

Fig 1: Alphabetical Sign languages

For each combination of hand gestures made by the person, a character/sentence presented by the system is transferred to the display screen. Raspberry pi uses Open CV, media pipe and Google tensor flow to process the input data and translate the sign gestures into text.
For the Sign to text communication YOLO (You Only Look Once) algorithm is used. This is a Real-time Object detection developed in 2015.It is a single-stage object detector that uses a Convolutional Neural Network(CNN) to predict the clear cases and class probabilities of object in input images. YOLO has been developed in several versions: such as YOLOv1, YOLOv2, YOLOv3, YOLOv4, YOLOv5, YOLOv6, and YOLOv7. Each version has been built on top of the previous version with enhance features such as improved accuracy, faster processing, and better handling of small objects.

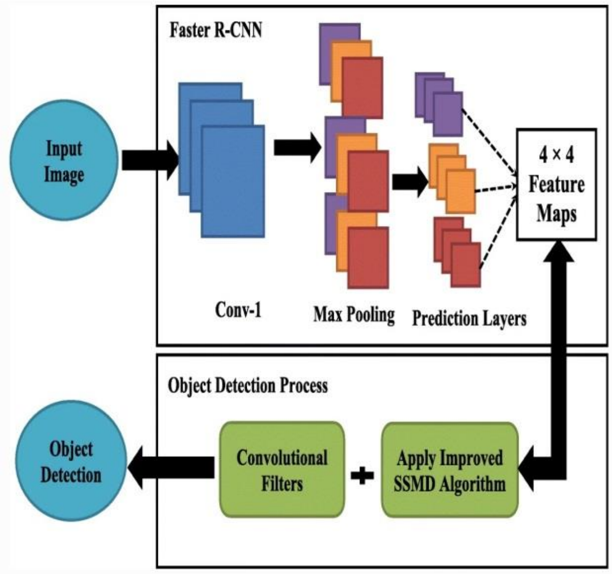

Fig 2: GENERALIZED WORKING OF YOLO ALGORITHM

Here's a simplified breakdown of the YOLO (You Only Look Once) process:
Feature Extraction: The input image is analysed by a Convolutional Neural Network (CNN) to identify important features.

Prediction: These features are used to make predictions about the objects in the image, including their class (what they are) and bounding box coordinates (where they are).

Grid Division: The image is divided into a grid of cells, with each cell responsible for predicting objects within its boundaries.

Output Generation: For each cell, the model produces predictions about potential objects, including their classes and bounding boxes.

Non-Max Suppression: To avoid redundancy, a technique called non-max suppression is applied to filter out overlapping bounding boxes, keeping only the most likely ones.

Final Results: The remaining bounding boxes, along with their associated class labels, are considered the final output, representing the objects detected in the image.
The version used in this paper is YOLOv7, the latest version of the YOLO algorithm, brings several improvements over previous versions. One major enhancement is the introduction of anchor boxes. These anchor boxes, which come in different shapes, help the algorithm detect objects of various sizes and shapes more accurately. By using nine anchor boxes, YOLO v7 can detect a wider range of objects, reducing the number of false detections. Another important improvement in YOLO v7 is the use of a new loss function called "focal loss." Unlike the traditional loss function used in earlier versions, focal loss focuses more on difficult examples, such as small objects, by adjusting the weight of the loss. This adjustment helps the algorithm pay more attention to challenging cases, ultimately improving its performance.

III. WORKING OF THE PROPOSED MODEL

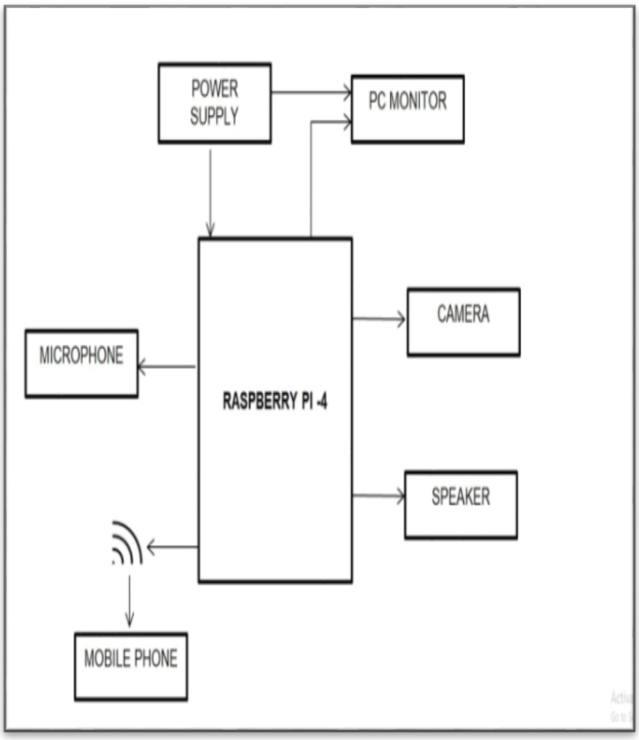

Fig 3: BLOCK DIAGRAM OF THE PROPOSED MODEL

Our project centres on the Raspberry Pi 4 platform, focusing on the conversion of voice and images. The core algorithm employed is the YOLO (You Only Look Once) algorithm within the domain of machine learning.

- **Preparation:** Assemble the Raspberry Pi device, which includes a 3-output pin jack, a headphone jack, a microphone, and a camera.

- **Connection Setup:** Use a C type cable to connect the Raspberry Pi to an adapter with less than 5 watts of power.

- **Network Configuration:** Establish connections between the Raspberry Pi and a mobile phone using a hotspot and Wi-Fi. Similarly, connect a laptop to the same mobile phone network.

- **Remote Access:** Launch the REAL VNC VIEWER application on the laptop and connect to the Raspberry Pi using its IP address.

- **Execution**: Open a terminal window on the laptop and enter the command **"sudo python3 audiotext.py"** to initiate audio-to-text conversion.

- **Audio Input:** Speak into the microphone connected to the Raspberry Pi, and observe the conversion of speech into symbols, aiding differently abled individuals in understanding spoken content.

- **Sign Language Translation:** In the terminal, execute the command **"sudo python3 signaudio.py"** to activate sign language interpretation.
- **Visual Input:** Upon execution, a new window displays live camera feed capturing sign language gestures. These gestures are translated into text and transmitted to a mobile device via the Blynk application.

- **Additional Functionality:** Repeat the aforementioned steps to enable voice output corresponding to sign language gestures captured by the camera.

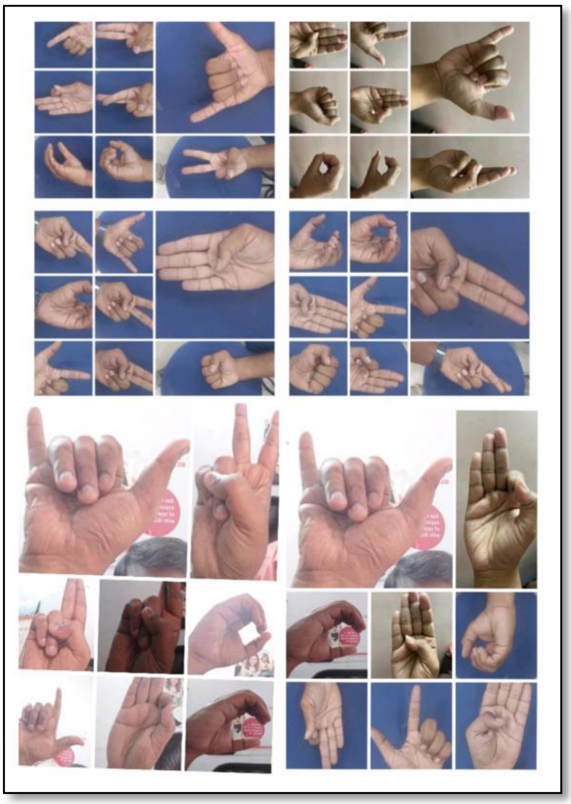

Fig 4: DATABASE COLLECTION FOR SIGN GESTURES

The sign gestures are loaded as the fig shown above; these are some examples of the images loaded in the software.

From this paper we could understand how we can communicate with specially abled people and these signs could help us in some emergency situations. We have parted a dataset into 80% as a testing dataset and 20% as a training dataset.

IV. EXPERIMENTAL RESULTS

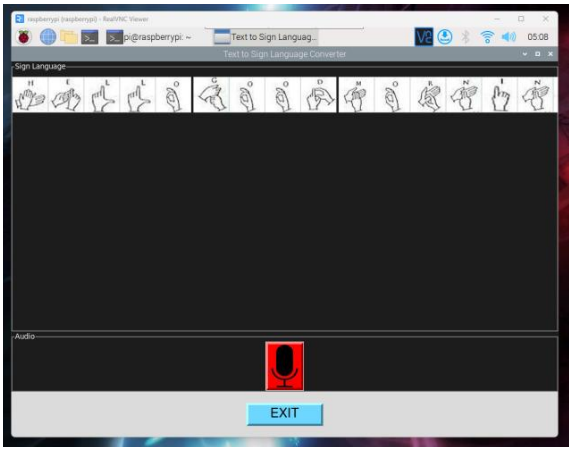

Fig 5: OUTPUT OF VOICE SIMULATION

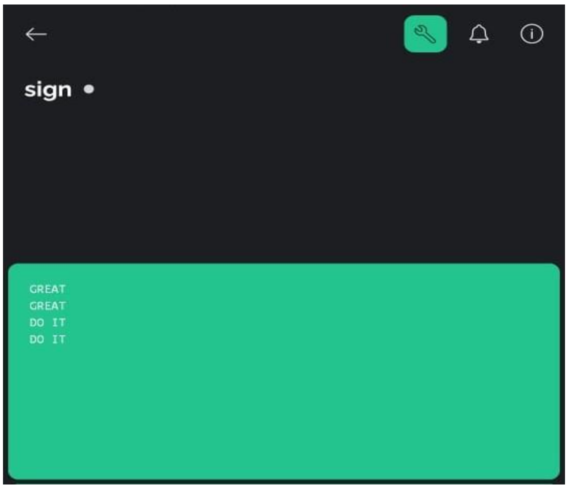

Fig 6: OUTPUT OF SIGN SIMULATION

The prototype models have designed for deaf and dumb people into a single compact device.

o Using this device a person can communicate and transfer the message as per his/her ability and desire.

o The proposed system supports real time communication which makes it more efficient.

o Constructing a system to make blind and visually impaired persons self-dependent.

o Providing efficient and unique capability in accessing private documents or text.

CONCLUSION

In brief, this project has devised a customizable system tailored to meet the needs of individuals with physical disabilities, with the overarching objective of improving their quality of life. The project's primary aim was to create a portable and wireless solution capable of translating sign language gestures into text, thereby enabling effective communication for individuals with physical limitations. By converting sign language into universally comprehensible text, the project seeks to bridge communication gaps between individuals with physical impairments and the broader community. The future trajectory of this model involves extending its utility to assist individuals with visual impairments by leveraging the headphone jack connection. Additionally, there is a plan to further streamline the product, aiming to develop it into a standalone solution for enhanced ease of use and accessibility.

REFERENCES

[1] Assistive System for deaf, dumb and blind Sharana Basavaraj B, Anusha T J, D Bhavani, Nandini K Assistant Professor, UG Students, Department of Electronics and Communication Engineering, RYMEC, Bellary, INDIA.

[2] ML Algorithms For Deaf, Dump, and Blind Assistive Device P THANUSH Dept. Electronics and Communication Engineering, R.V. College Of Engineering, Bangalore, India. Nagaraj bhat Dept. Electronics and Communication Engineering, R.V. College Of Engineering, Bangalore, India

[3] Assistive Device for Deaf, Dumb and Blind B. S. Anirudh Bhargav, N. Gowtham, P. Pandethi Pavan, A. P. Vinay Kumar, Y. N. Santhosh Student, Dept. of Information Science and Engineering, Sai Vidya Institute of Technology, Bangalore, India Assistant Professor, Dept. of Information Science & Engg., Sai Vidya Institute of Technology, Bangalore, India

[4] Smart Communication Frame work for deaf, dumb and blind people using Raspbery Pi Murthy S V, Gagana Venkatesh , Bhumika , Chaithra M N Dept. of CSE, SJCIT, Chickballapur, India

[5] Smart Assistant for Deaf and Dumb Using Flexible Resistive Sensor Implemented on Labview Platform S.Kumuda Dept.of Electronics & Instrumentation Engineering B.M.S College of Engineering Bangalore, India Preethi K Mane Dept.of Electronics & Instrumentation Engineering B.M.S College of Engineering Bangalore, India

[6] Innovative study of an AI voice based smart Device to assist deaf people in understanding and responding to their body language Dhaya Sindhu Battina Lakshmisri S Sr. Data Automation Engineer at Key Bank & Department of Information Technology Sr. Data Scientist at The Walt Disney & Department of Information Technology

[7] Raspberry Pi based Sign to Speech Conversion System for Mute Community G.Nirosha1 , Dr. Velmani Ramasamy2 1 PG Scholar, Department of ECE, Siddartha Institute of Technology and Sciences, Hyderabad, Telangana,India 2 Associate Professor and Head, Department of ECE, Siddartha Institute of Technology and Sciences, Hyderabad, Telangana, India

[8] Gesture Language Translator Using Raspberry Pi Supriya Pawar1 , Apoorva Bamgude2 , Sayali Kamthe3 , Arundhati Patil4 , Prof. Rajesh Barapte5 1, 2, 3, 4, 5Electronics and Telecommunication Engineering, Keystone School Of Engineering, Pune, India

[9] Sign Language Recognition And Speech Conversion
Using Raspberrypi
Ramasuri Appalanaidu CH1 , Nambolu Sai Ramya2 , Killada Sumanjali3
K Venkata Lakshmi4 , Kinthali Gayatri5
2,3,4,5 Student , B.Tech (Information Technology) Assistant Professor,Information Technology, Vignan's Institute of Engineering for Women, Visakhapatnam, Andhra Pradesh, India

*[10] Sign language voice convertor design using Raspberry pi for impaired individuals Serhat Küçükdermenci *,1 1Department of Electrical and Electronics Engineering, Faculty of Engineering, Balıkesir University, 10463, Balikesir, Turkey.*

[11] Voice aassisted Real-time object detetction using YOLOv4 algorithm for visual challenged ZAEEM NAZIR Ph.D. Scholar, Department of Computer Science, Superior University, Lahore, Pakistan. MUHAMMAD WASEEM IQBAL. Ph.D., Associate Professor Department of Software Engineering, Superior University, Lahore, Pakistan. KHALID HAMID Ph.D. Scholar, Department of Computer Science, Superior University, Lahore, Pakistan and Lecturer at NCBA & E University East Canal Campus Lahore. Corresponding Author HAFIZ ABDUL BASIT MUHAMMAD.Ph.D. Scholar, Department of Computer Science, Superior University, Lahore, Pakistan and Lecturer at Minhaj University Lahore. M. ASHRAF NAZIR Ph.D. Scholar, Department of Computer Science, Superior University, Lahore, Pakistan and Lecturer at GC University, Lahore. QURRA-TUL-ANN Government College University, Lahore, Pakistan. NAZIM HUSSAIN Lecturer, Department of Computer Science, Government College University, Lahore, Pakistan.