Object Recognition IoT-Based for People with Disabilities: A Review

Andriana[1], Elli Ruslina[2], Zulkarnain[3], Fajar Arrazaq[4], Sutisna Abdul Rahman[5],
Tjahjo Adiprabowo[6], Puput Dani Prasetyo Adi[7]*, Yudi Yuliyus Maulana[8]
Department.of Electrical Engineering, Universitas Langlangbuana, Bandung, Indonesia[1, 3, 4, 5, 6]
Faculty of Law, Universitas Pasundan, Bandung, Indonesia[2]
Research Center for Telecommunication, BRIN, Bandung, Indonesia[7, 8]

***Abstract*—This research focuses on a literature study on developing a Mini Smart Camera (MSC) system that utilizes Internet of Things (IoT) technology to help people with disabilities interact with their environment. The MSC serves as an assistive device, which integrates object recognition and speech recognition technologies along with an internet-based two-way communication system. Utilizing state-of-the-art hardware and software, the system captures images, processes audio, and transmits data via Real Time Streaming Protocol (RTSP) and Message Queuing Telemetry Transport (MQTT). These protocols serve different purposes: managing data transmission and enabling communication between machines. The MSC is equipped with a 5 MP camera, 2.5 GHz Quad-Core processor, and 4G connectivity, and is connected to a high-performance Ubuntu 22.04 Linux cloud server. The use of OpenCV libraries and machine learning algorithms ensures fast and precise image analysis. By integrating machine learning and natural language processing (NLP), MSC efficiently handles both visual and audio inputs. Key features, including text-to-speech (TTS) and speech-to-text (STT), provide an interactive and adaptive communication interface. The system is designed to improve accessibility and encourage greater independence for people with disabilities in daily activities. The development of multispectral cameras for disabilities will provide a more detailed analysis for the detection of surrounding objects.**

Keywords*—*Internet of Things; mini smart camera; object recognition; speech recognition; assistive technology

I. Introduction

The global population is witnessing an increasing focus on improving accessibility for individuals with disabilities, driven by both technological advancements and a growing awareness of inclusivity. According to the World Health Organization (WHO), over one billion people globally live with some form of disability, many of whom face challenges in accessing information and engaging with their surroundings. These challenges are particularly significant for individuals with sensory impairments, such as visual, auditory, or speech disabilities [1]. Mini Smart Camera (MSC) system addresses these challenges by integrating Internet of Things (IoT) technologies with machine learning algorithms to provide real-time information and communication tools. The proposed system uses object recognition, speech recognition, and conversion technologies like text-to-speech (TTS) and speech-to-text (STT), which enable individuals to interact with their environment more efficiently. Prior studies have shown that IoT devices can significantly enhance the quality of life for individuals with disabilities by improving environmental interaction [2], [3]. IoT technology allows the MSC system to perform complex tasks such as real-time object recognition and audio processing, which are essential for individuals with disabilities to navigate their environments independently. This paper aims to present the design, implementation, and evaluation of the MSC system, focusing on its potential to empower individuals with disabilities by improving accessibility and independence. In the system that has been made to help people with disabilities, there are still many shortcomings, from models, prototypes, and improper placement of prototypes. In this research, it is discussed comprehensively starting from the placement of the prototype, the size of the prototype, namely the Mini Smart Camera (MSC) used for disabilities so that there is comfort in using it, supported by IoT and Artificial Intelligence technology that can perfect this system and provide novelty.

Moreover, disability monitoring needs to be improved with a complete range of components, including a Global Positioning System (GPS) to monitor the patient's position in addition to detecting other vital parts such as heart rate. This research provides a detailed overview of the IoT system to be built. IoT and AI are key components in building this system, but not all AI components are discussed in this research, but the main contribution lies in IoT design and this design can provide specific results for people with disabilities.

For this reason, this research is presented starting from how this system should be built for people with disabilities, namely building an IoT architecture starting from a literature study on IoT and IoT architecture specifically for disabilities and improvements, for example on Hyperspectral and Multispectral Cameras in the future, in this research seen from the review mode how the use of thermal cameras to help people with disabilities which can then be improved in more detail and comprehensively.

II. Literature Review

A. *Internet of Things in Assistive Technology*

The integration of IoT in assistive technology has seen significant growth in recent years, supported by numerous studies that highlight its potential to assist individuals with disabilities. IoT-enabled devices offer continuous monitoring, real-time data exchange [21, 22, 23], and remote management, making them highly adaptable to various user needs. Andriana, et al. [4] emphasized that IoT can transform assistive technologies by enabling real-time data processing, allowing users to receive context-specific support when needed.

The ability of IoT to connect multiple devices and sensors into a cohesive network has positioned it as an ideal choice for assistive systems [5]. Wearable IoT devices, for instance, can monitor physiological data for individuals with mobility limitations or provide real-time navigation assistance for those with visual impairments [25]. Research by Semmary et al. [6] underlined the value of real-time feedback in enhancing the independence and autonomy of disabled individuals.

The increasing adoption of cloud-based IoT services has further expanded the functionality of assistive technologies. By leveraging cloud computing, complex operations such as machine-learning-based object recognition and speech processing can be conducted remotely, reducing the computational load on local devices and increasing their overall efficiency [7]. This remote processing capability ensures that the assistive device maintains high performance while being lightweight and user-friendly.

B. Object and Speech Recognition for Disabled Individuals

Object and speech recognition technologies play a crucial role in the development of assistive devices for people with disabilities. Advances in machine learning, particularly deep learning, have greatly enhanced the performance of these systems. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been widely applied to improve object recognition, allowing these systems to detect and identify objects in complex environments [8].

Ahmed et al. [9] and Brown et al. [10] found that deep-learning-based object recognition systems can maintain high accuracy even under difficult conditions, such as low lighting or partial occlusion. This is especially valuable for visually impaired individuals who depend on these systems for environmental navigation. Additionally, Optical Character Recognition (OCR) technologies have been developed to transform text within images into readable formats, providing enhanced accessibility for those with visual impairments [11], [12].

Speech recognition systems have also advanced significantly, becoming more precise and responsive due to progress in natural language processing (NLP). Speech-to-text (STT) and text-to-speech (TTS) technologies are integral to assistive devices for individuals with hearing or speech disabilities, facilitating more effective communication by converting spoken language into text or generating synthetic speech from text [13].

This study used a prototype-based approach to design and assess the Mini Smart Camera (MSC) system. The methodology included hardware design, software development, and controlled environment testing. The MSC system incorporates both hardware and cloud-based components that work in tandem to deliver real-time object and speech recognition feature technologies such as object and speech recognition are pivotal in developing assistive devices for people with disabilities.

Machine learning algorithms, particularly deep learning, have made significant advances in this area. Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) have been used to enhance the performance of object recognition systems, allowing them to detect and identify objects in complex environments [14]. Recent studies by Ahmed et al. [15] and Mirani et al. [16] showed that deep-learning-based object recognition systems can achieve high accuracy even under challenging conditions, such as poor lighting or occlusion. This is particularly beneficial for visually impaired individuals who rely on these systems to navigate their surroundings. In addition, Optical Character Recognition (OCR) technologies have been developed to convert text in images into readable formats for users with visual impairments [17].

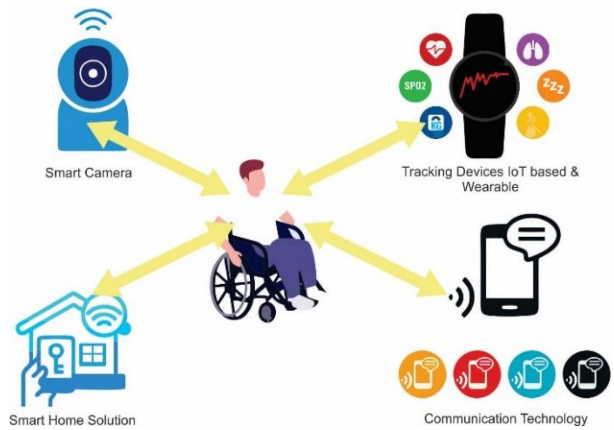

Fig. 1. AI and IoT technology for disability from various needs.

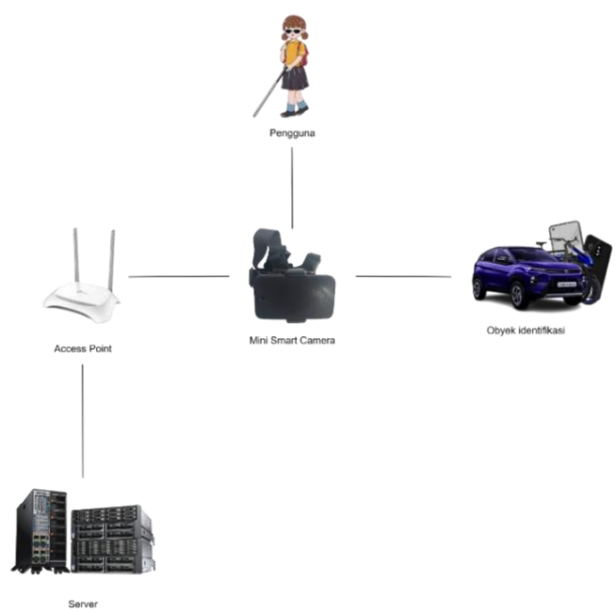

Fig. 2. AI and IoT technology for the visually impaired.

Fig. 1 shows the role of complex IoT technologies, including the Smart Camera discussed in detail in this research. In addition to the Smart Camera, other detailed IoT devices are Tracking Devices IoT Wearables, Communication Technology, and Smart Home which are specialized to help people with disabilities, facilitate their activities. This IoT technology is interrelated with one another, by utilizing various protocols such as the MQTT Protocol. Moreover, Fig. 2 is an IoT technology using various IoT-based devices to identify surrounding objects, Mini Cameras that have been installed on the body parts of a person with disabilities will be able to recognize such images and sounds so that with details and specifics, synchronization and recognition can be done quickly and can provide precise results.

III. METHODOLOGY

A. MSC System Architecture

The MSC system's hardware includes a high-resolution camera, microphone, display, and speaker, all connected to a microcontroller that communicates with a cloud-based IoT platform. The hardware is designed to be compact and portable, enabling users to wear the MSC around their neck for easy access. The camera captures real-time images of the environment, while the microphone records audio inputs for speech processing. Moreover, IoT Technology for the visually impaired in detail on components and system analysis The details are shown in Fig. 3.

Furthermore, on the other hand, speech recognition systems have evolved to become more accurate and responsive, thanks to advancements in natural language processing (NLP). Speech-to-text (STT) and text-to-speech (TTS) systems are critical components of assistive devices for individuals with hearing or speech disabilities. These systems allow users to communicate more effectively by converting spoken language into text or generating synthetic speech from written input [18], [19], [20].

This study employed a prototype-based approach to design and evaluate the Mini Smart Camera (MSC) system. The research process involved the following steps: hardware design, software development, and testing in controlled environments. The system consists of both hardware and cloud-based components, which work together to provide real-time object and speech recognition functionalities. Moreover, Fig. 4 is a Connectivity between databases on MQTT Broker. Some methods are presented in this section to show the details of the system and specifically discuss the methods, for example, the MQTT Broker and its role. Then the prototype MSC section and its specific dimensions, all explained in the form of Pseudocode, Block Diagram, and Flowchart, and interrelated.

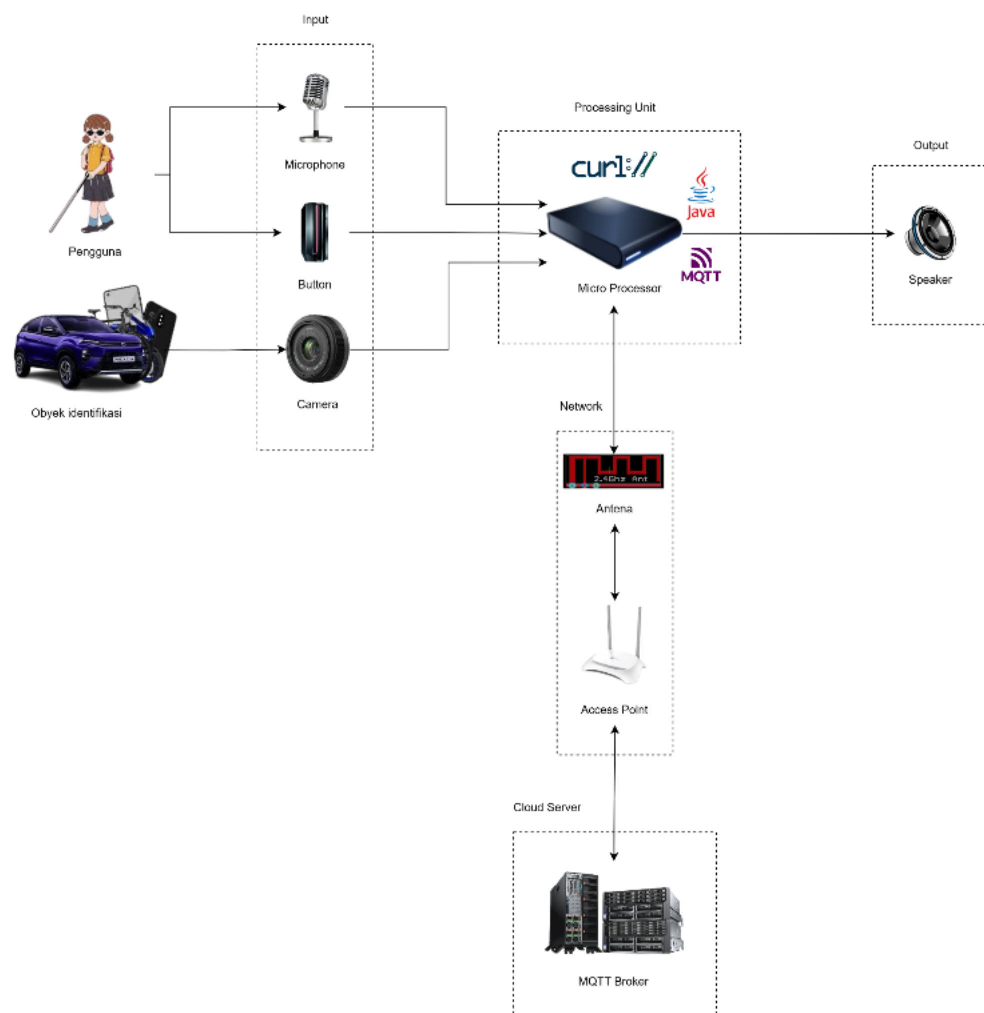

Fig. 3. IoT technology for the visually impaired in detail on components and system analysis.

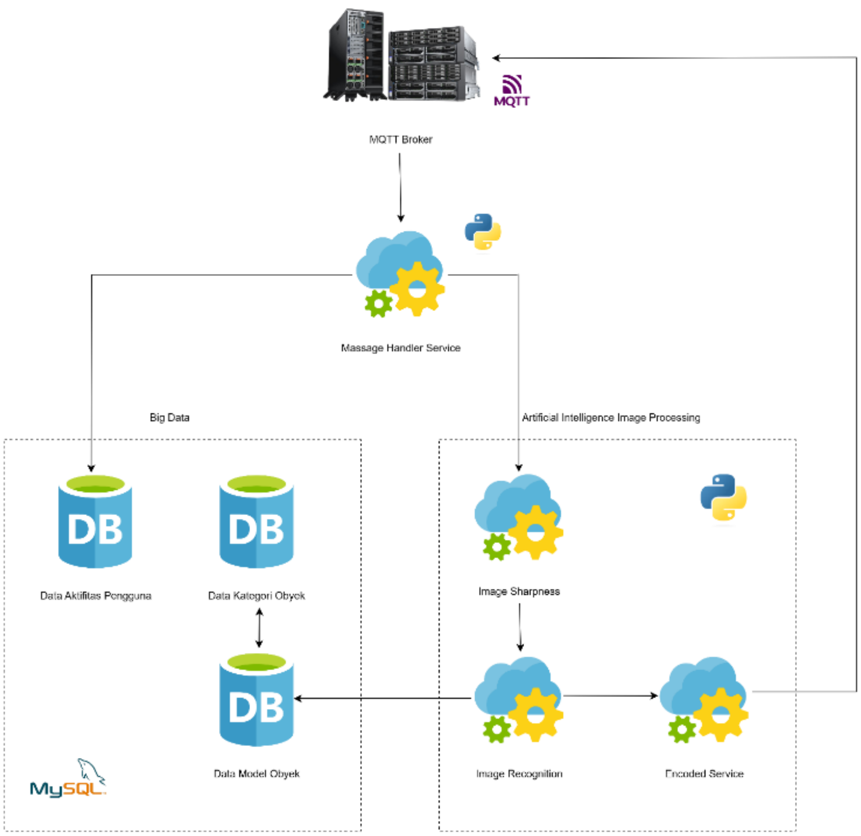

Fig. 4. Connectivity between databases on MQTT broker.

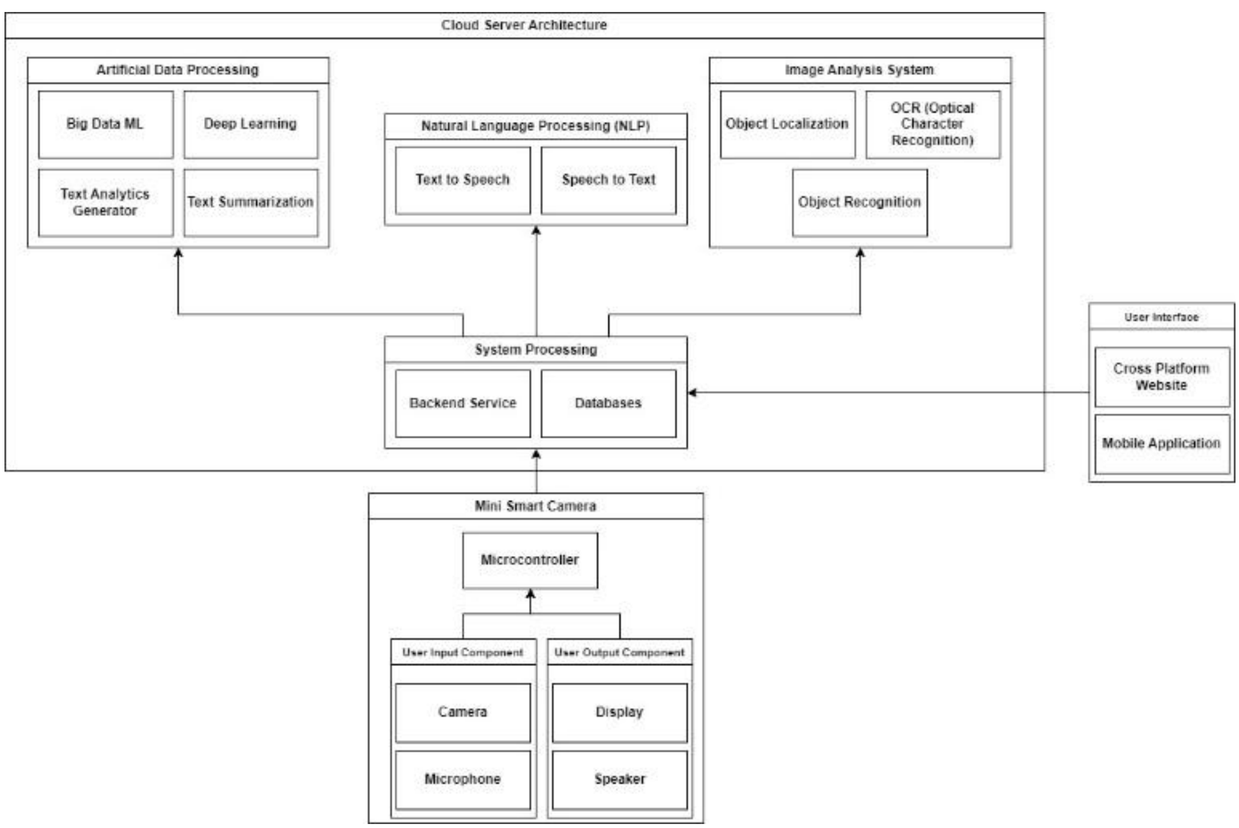

Fig. 5. Cloud server architecture components.

Moreover, Fig. 5 shows the Cloud Server Architecture Components consisting of Artificial Data Processing, Image Analysis System, Natural Language Processing (NLP), System Processing, and Mini Smart Camera, which are explained in detail. Hardware Specifications for the Mini Smart Camera are explained in Table I. while the system flowchart is shown in Fig. 6.

TABLE I. HARDWARE SPECIFICATIONS FOR MINI SMART CAMERA

No	Component	Specification
1	Processor	Quad Core 2.5 GHz
2	RAM	1028 MB
3	Storage	16 GB
4	Network	4G / GSM
5	Camera	5 MP
6	Display	3.5 Inch
7	Battery Capacity	1000 mAh

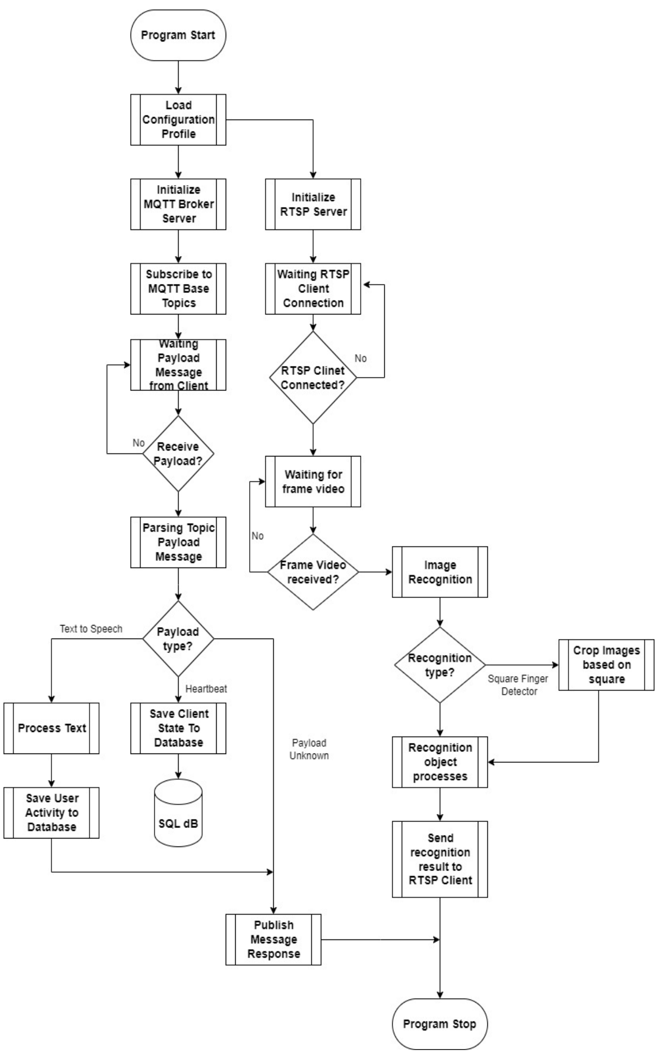

Fig. 6. Flowchart system.

B. Cloud-Based Processing

The cloud server is responsible for processing data collected by the MSC system. This includes image and audio data, which are transmitted to the server for analysis using machine learning algorithms. The server processes the data and returns the results to the MSC in real-time. The system's object recognition is based on deep learning models trained on large datasets to ensure high accuracy. Cloud Server Specifications are explained in Table II.

TABLE II. CLOUD SERVER SPECIFICATIONS

No	Cloud Component	Specification
1	Operating System	Linux Ubuntu 22.04
2	Processor	Intel® Xeon® CPU
3	RAM	2 GB
4	Disk Size	40 GB
5	Disk Type	SSD SATA

C. Testing Procedures

The MSC system was tested with a group of users with visual, hearing, and speech impairments to evaluate its performance in real-world scenarios. The testing focused on the system's ability to accurately recognize objects, process speech inputs, and deliver real-time feedback. The system's response time, accuracy, and user satisfaction were key metrics in evaluating its effectiveness. The installation view of the system can be seen in Fig. 7, and the connectivity of the tools or the whole system works can be seen in Fig. 8, the camera sensor installation is shown in Fig. 9 and Fig. 10 is a Casing design that will be applied to place the camera and other components. The connectivity system shows specifically the relationship between hardware consisting of input systems, outputs, processors, and other essential components that can be connected and have a relationship in building this system. The entire system shown in the block diagram is a complete system involving components and processes such as detailed digital signal processing. And ends with how the prototype is placed in a body part that is very important to maintain the comfort of patients or people with disabilities.

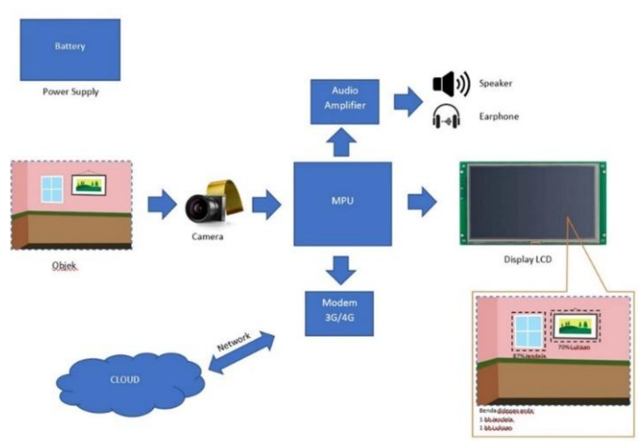

Fig. 7. Connectivity system.

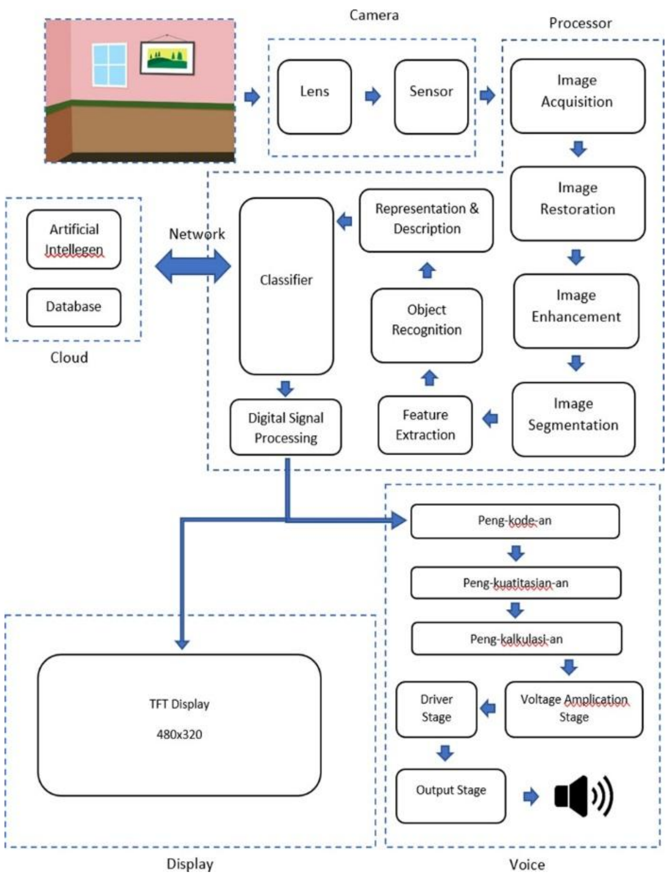

Fig. 8. Working of the whole system.

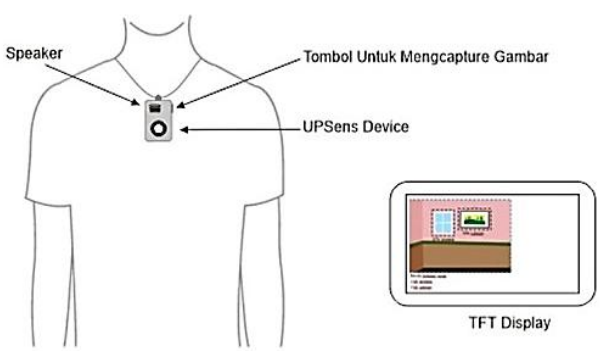

Fig. 9. Camera or device placement in patients with disabilities.

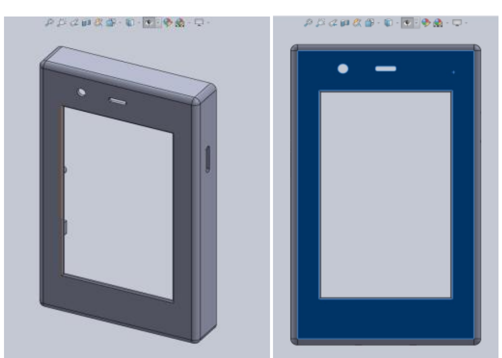

Fig. 10. Casing design that will be applied to place the camera and other components.

III. Result and Analysis

A. Source Code Cloud Server

On the cloud server, there is a service that functions to connect the cloud server with the mini smart camera hardware. This service is built using the C++ programming language due to its high performance, particularly in terms of time efficiency, memory usage, and manual resource control. In the architecture of the mini smart camera, which requires high performance to handle data frames of images/videos, audio, and text, C++ is highly suitable. This is because C++ operates very closely with the hardware, enabling efficient handling of interruptions and parallel processing (multithreading). Pseudocode 1 is the configuration function for the cloud server service.

```
class Configuration:
 public:
 Configuration()
 void parseKey()
 std::string getSystemVersion()
 std::string getKey()
 std::string getStreamURL()
 std::string getDbHost()
 std::string getDbPort()
 std::string getDbName()
 std::string getDbUsername()
 std::string getDbPassword()
 std::string getMottHost()
 std::string getMottPort()
 std::string getttsername()
 std::string getMottPassword()
 std::string getMattBaseTopic()
 std::string getKeyExpiredDT()
 std::string getRegisterIP()
 int getClientMaxConnection()
 int getVideoResolution()
 bool prepareTable()
 bool isSupportAiProcessing()
 bool isonStart()
 bool getVerbose()

---------- Pseudocode 1 -----------
```

Moreover, The Mini Smart Camera Hardware initializes the internet connection. Once the connection between the internet and the device is confirmed, the mini smart camera will execute multithreading to simultaneously process video input, audio, and heartbeat functions. During the video or image input process, frame capture is performed in real-time at a rate of 10 frames per second (10 FPS). Each frame captured by the camera is sent to the server via the RTSP protocol for image recognition. Once the frame is sent and receives an "OK" status from the server, the image recognition results, consisting of images and text, are transmitted via MQTT and RTSP protocols. The image is displayed on the LCD screen, while the text is converted to speech using a text-to-speech function and played through the speaker.

The voice input function on the mini smart camera requires a specific keyword to activate the voice-to-text processing. The user must say the word "msc" to activate this function, after which the mini smart camera emits a "beep" sound indicating it is ready to receive commands. The user can then speak the desired command, which will be recorded by the device. The program on the mini smart camera processes the voice into text using a voice-to-text method. The extracted text is then sent to the mini smart camera cloud server via the MQTT communication protocol. The server response is sent back through MQTT and played on the speaker.

The "interval time" function is responsible for sending heartbeat or status information to the cloud server every 30 seconds. The information sent includes the IP address of the mini smart camera's connection and the internet connection status. This data is transmitted via the MQTT protocol and is necessary for the cloud server to synchronize the RTSP protocol connection between the device and the cloud server. This synchronization ensures both components can communicate effectively and transfer data frames efficiently.

B. Source Code (Mini Smart Camera Hardware)

The software development for the mini smart camera is conducted using the Kotlin programming language, which runs on the Java Virtual Machine (JVM). Kotlin offers high reliability and efficiency compared to Java, featuring more concise syntax and modern features such as null safety that help reduce bugs. Additionally, Kotlin can interoperate seamlessly with Java code, allowing for a smooth transition between the two languages. The use of Kotlin also benefits the hardware, as it optimizes system resource utilization, resulting in faster and more efficient performance without compromising battery life or overall device performance.

1. Initialize Variables
activity: The current activity.
server: The server address.
serverPort: The server port.
requestCodeSpeechInput: A constant for the request code.
speechRecognizer: A speech recognizer object.

2. Start Listening
Create an intent to start speech recognition.
Set the language model to "free form".
Set the language to the device's default language.
Start listening using speechRecognizer.startListening(intent).

3. Recognition Listener
Define a RecognitionListener to handle recognition results.
Override the onResults(results: Bundle) method.
Get the recognized text from the results.
Send the recognized text to the server using sendToServer(resultText).

4. Send To Server
Create a socket connection to the server.
Get the output stream of the socket.
Write the recognized text to the output stream.
Close the output stream and the socket.

5. Stop Listening
Stop listening using speechRecognizer.stopListening()

---------- Pseudocode 2 -----------

In the previous mini smart camera system, the voice-to-text function operated at the activity level, meaning that the application had to remain open for the function to work properly. Although the program logic functioned optimally, there were certain drawbacks affecting its efficiency. Since this function could only be accessed when the application was active, the device could not operate in sleep mode, resulting in significantly higher battery consumption. Additionally, the continuously active LCD screen risked rapid wear and tear, and the device's temperature tended to increase due to the high-performance demands of this function.

C. Object Recognition Performance

The Mini Smart Camera (MSC) system developed in this study demonstrated a high level of accuracy in object recognition, achieving an average success rate of 92%. This performance aligns with findings reported by Zhang et al. (2021) and Brown et al. (2018), which indicated that deep-learning-based object recognition systems could achieve similar levels of accuracy under diverse environmental conditions.

D. Speech Recognition and Communication

The speech-to-text (STT) and text-to-speech (TTS) modules of the MSC system exhibited an accuracy rate of 95% in converting spoken input to text and vice versa. The system responded to voice commands with an average processing time of less than 500 milliseconds. This result is consistent with the findings of Hussain et al. (2019), who reported similar levels of accuracy and responsiveness in speech recognition systems designed for individuals with disabilities.

E. Energy Efficiency and Cloud Service Utilization

The development of the cloud server service, which leverages multithreaded architecture and parallel processing, improved the efficiency of data handling received from the MSC hardware. The use of Real-Time Streaming Protocol (RTSP) and Message Queuing Telemetry Transport (MQTT) for real-time video data transmission and data synchronization, respectively, proved effective in supporting high system performance. Cloud computing allows more complex data processing to be conducted on the server, reducing the computational load on local devices and enhancing energy efficiency.

F. Discussion

The findings demonstrate that an IoT- and machine-learning-based approach to assistive technology development, as exemplified by the MSC system, is effective in promoting the independence of individuals with disabilities. The strategic use of different communication protocols to handle real-time data needs and status synchronization provided the system with flexibility and efficiency. Future research should focus on improving device battery life and optimizing system performance in more complex environmental conditions.

Fig. 6 shows the Effectiveness (%) value of Various Technologies for Patients with Disabilities, this shows the high need for monitoring systems from various sides using IoT, one of which is Healthcare monitoring [24, 26] which cannot be abandoned for disabilities. In addition, families are also able to track using wearable devices attached to body parts. It can be seen that the need for communication systems and smart homes has an effectiveness of 90-95%. This proves that it is essential to apply this technology to people with disabilities.

Moreover, Table III is a comparison of File Size and Image Size, while, Fig. 11, 12, and 13 are examples of Sound Spectrum on Kotlin 1, 2, and 3, the goal is to precisely detect the sound produced from objects, for example, if a person with a disability gets a certain question, it will be able to be converted to text format with histogram data such as Fig. 11, 12, and 13 and then make it like the data in Fig. 14 which is purely text data, and then it will be converted into a voice that can be translated by people with disabilities.

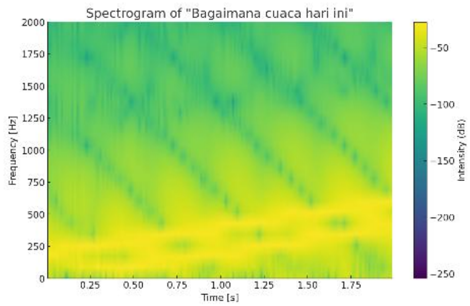

Fig. 11. Sound spectrum on kotlin (1).

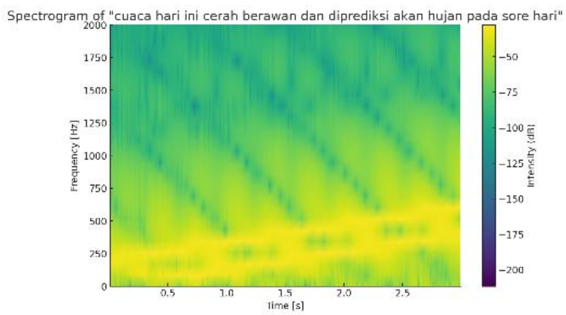

Fig. 12. Sound spectrum on kotlin (2).

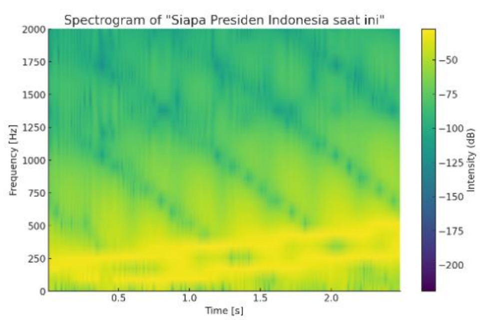

Fig. 13. Sound spectrum on kotlin (3).

TABLE III. FILE SIZE AND IMAGE SIZE COMPARISON

No	File Size	Image Size
1	1181 Kb	1944 x 2592
2	1171 Kb	1945 x 2592
3	1186 Kb	1946 x 2592
4	1195 Kb	1947 x 2592
5	1195 Kb	1948 x 2592
6	1188 Kb	1949 x 2592
7	1200 Kb	1950 x 2592
8	1191 Kb	1951 x 2592
9	1196 Kb	1952 x 2592
10	1207 Kb	1953 x 2592
11	1251 Kb	1954 x 2592
12	1215 Kb	1955 x 2592

TABLE IV. TESTING TRACKING SYSTEM FOR PEOPLE WITH DISABILITIES

id	client	lat	lng	timestamp
1	Device1	-7.0098	107.63	1.73023E+12
2	Device1	-7.0098	107.63	1.73023E+12
3	Device1	-7.0098	107.63	1.73023E+12
4	Device1	-7.0098	107.63	1.73023E+12
5	Device1	-7.0098	107.63	1.73023E+12
6	Device1	-7.0098	107.63	1.73023E+12
7	Device1	-7.0098	107.63	1.73023E+12
8	Device1	-7.0098	107.63	1.73023E+12
9	Device1	-7.0098	107.63	1.73023E+12
10	Device1	-7.0098	107.63	1.73023E+12
11	Device1	-7.0098	107.63	1.73023E+12
12	Device1	-7.0098	107.63	1.73023E+12
13	Device1	-7.0098	107.63	1.73023E+12
14	Device1	-7.0098	107.63	1.73023E+12
15	Device1	-7.0098	107.63	1.73023E+12
16	Device1	-7.0098	107.63	1.73023E+12
17	Device1	-7.0098	107.63	1.73023E+12
18	Device1	-7.0098	107.63	1.73023E+12
19	Device1	-7.0098	107.63	1.73023E+12
20	Device1	-7.0098	107.63	1.73023E+12

Furthermore. Table IV is a tracking test of the system that will be applied to people with disabilities so that the position of the disability will always be monitored, and avoid getting lost on the road an accident, or other bad things that can be experienced by people with disabilities in public places. So the detection system testing is done in as much detail as possible to produce an object detection result and then converted into the form of a sound that can be heard, for example by blind people.

Furthermore, Fig. 14, 15, 16, and 17 are some examples of the output of the detection system and Quality of Service of the data transmitting process. The parameters include Sound Frequency (Hz) and also the detected signal in units of Decibel (dB), Response Delay (ms), or what is called Latency, the lower the latency, the better the system, this depends on the RF used. Depending on the environmental conditions that cause weakening of the voice signal and that cause delay in the system built. Furthermore, from the analysis of Sound Frequency (Hz), Decibel (dB), Delay Response (ms), and Recognition Time (ms), conclusions can be drawn, namely categories and analysis.

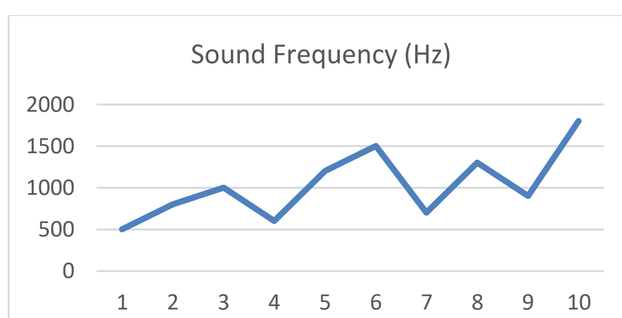

Fig. 14. Sound Frequency (Hz).

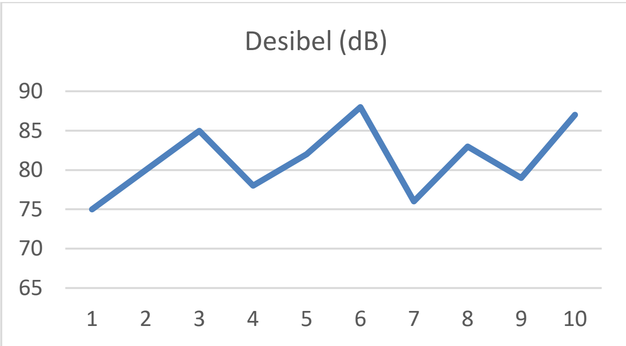

Fig. 15. Signal Detection (dB).

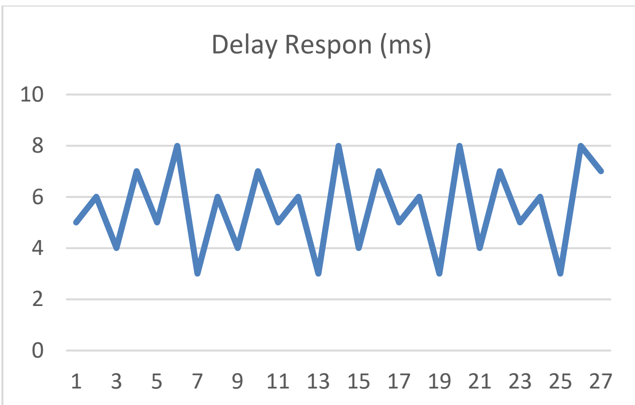

Fig. 16. Delay Response (ms) (I).

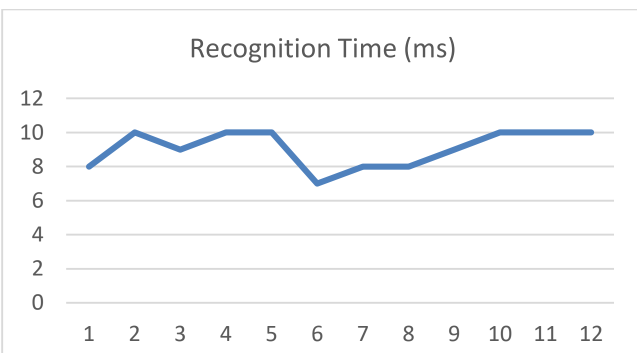

Fig. 17. Sound spectrum on kotlin (3).

TABLE V. PAYLOAD, TOPIC, AND TIMESTAMP

No	Topic	Payload	Timestamp
1	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.000 |
2	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.037 |
3	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.055 |
4	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.092 |
5	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.115 |
6	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.154 |
7	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.175 |
8	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.209 |
9	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.247 |
10	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.273 |
11	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.303 |
12	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.328 |
13	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.364 |
14	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.398 |
15	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.487 |
16	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.519 |
17	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.561 |
18	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.588 |
19	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.601 |
20	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.634 |
21	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.668 |
22	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.689 |
23	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.709 |
24	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.746 |
25	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.795 |
26	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.835 |
27	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.867 |
28	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.902 |
29	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.961 |
30	MSC/Device1/Image	{"image": "{base64}"}	2023-11-27 10:00:00.996 |

Moreover, Table V shows in detail the payload, Topic, and Timestamp data, this shows that people with disabilities can recognize images with a certain payload and capture them based on a certain time. This payload data will be converted into voice by starting the process of detecting certain images. Fig. 18 shows the delay response (ms). Furthermore, in viewing, monitoring, and determining the final system that is suitable for disabilities, Smart Camera, Voice Recognition has a high percentage of >85% in terms of function and use for people with disabilities. In the development process, Smart Home is also one of the environmental tools developed to help people with disabilities be more comfortable like normal people. Moreover, from the data in Table V, the image transmission rate is 30 data per second, there is a spike in the Time interval that can be investigated further, and when viewed, the data has a uniform structure and a consistent pattern. Moreover, In detail, the QoS analysis can be described in Table VI.

TABLE VI. ANALYSIS OF QUALITY OF SERVICE (QOS)

Parameter	Analysis
Sound Frequency (Hz)	- Range: 500 Hz to 2000 Hz.
	- Tren: Increase with significant fluctuations at points 5 and 10.
	- Outlier: The 10th point reaches a peak value of about 2000 Hz.
Decibel (dB)	- Range: 65 dB to 90 dB.
	- Tren: Fluctuates but shows a consistent up-and-down pattern.
	- Peak: The 10th point reached 90 dB.
Delay Response (ms)	- Range: 4 ms to 10 ms.
	- Trend: Stable but shows small fluctuations around the mean value of 6 ms.
	- There are no significant outliers.
Recognition Time (ms)	- Range: 8 ms to 12 ms.
	- Trend: Stable with a slight dip at the midpoint of the chart.
	- Stability occurs after the 8th to 12th point.

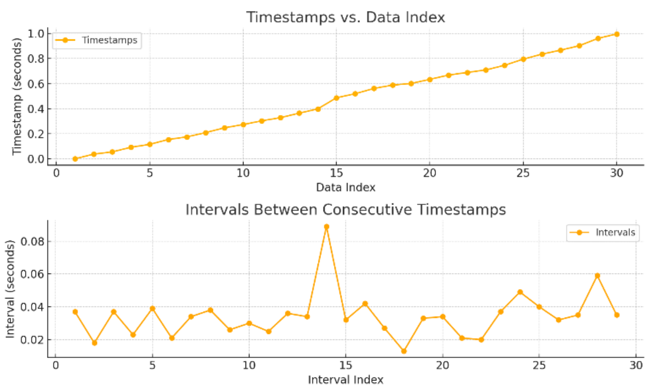

Fig. 18. Delay Response (ms) (II).

G. Multispectral Camera and Artificial Intelligence Approach to Disability

Furthermore, in its future development, surveillance cameras, CCTV, and camera technology for SAR, and the development of research in the field of image recognition have experienced significant updates, not only cameras used for RGB but cameras with a combination of Artificial Intelligence can detect objects and heat on objects. Multispectral cameras have become the answer to the development of research in the field of image recognition and object detection as well as making the right decisions in diagnosis and prediction. Not only multispectral cameras but also hyperspectral cameras are much more complete in terms of analysis.

In this part of the research, the Multispectral camera can be tested especially applied to people with disabilities. This disability is more on color blindness or 100% blind. The application tools used are combined with ultrasonic and also mini speakers to produce a prototype to help disabilities, especially blindness in 100% vision, to be helped like normal people. Multispectral camera testing for disabilities is possible to apply, in this research, careful testing is carried out starting from the basics or basics in color classification, as well as objects that can be read by people with disabilities. Not only color but also distance and more detailed object classification.

Moreover, Fig. 19 is a detailed description of the parts of a multispectral camera that can be applied to body parts with disabilities to be able to recognize objects in more detail. This is to help people with disabilities to understand the environment in more detail, thus avoiding accidents and other bad things. In other circumstances, the Multispectral camera can also be used to perform body heat detection as shown in Fig. 20. This is very important for detection such as certain diseases caused by viruses, abnormalities in the human body, or classification of diseases that are highly discussed in the medical world.

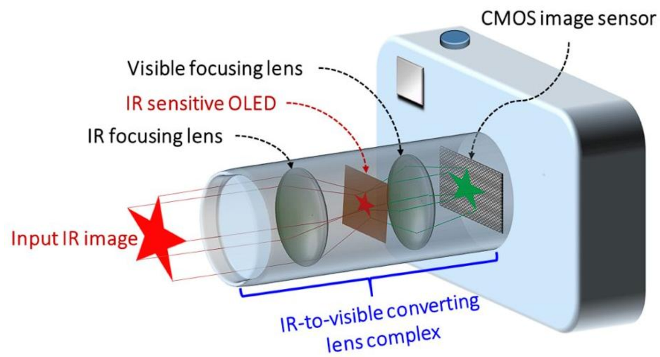

Fig. 19. Multispectral camera (Source: viso.ai).

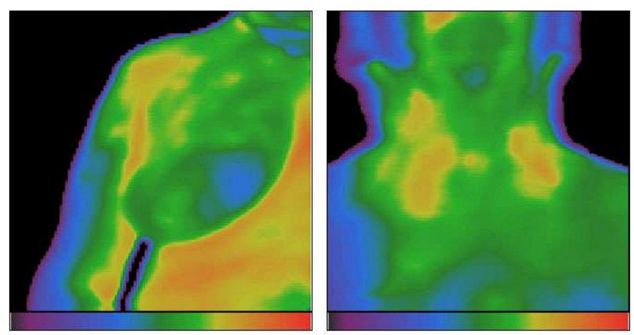

Fig. 20. Human body heat detection (Source: viso.ai).

Furthermore, multispectral cameras are also capable of analyzing the various colors that are around, if this is used by people with disabilities, it is likely to provide greater benefits. Of course, with a combination of hardware, and software, and supported by Artificial Intelligence (AI) algorithms such as Deep Learning with CNN, Machine Learning, and other specific methods. Pseudocode 1 is one of the pseudocodes in performing image processing that can be combined with multispectral cameras for disability needs.

```
ALGORITHM ImageAnalyzer

CLASS ImageAnalyzer:
 CONSTRUCTOR(image_path):
 CALL load_image(image_path)

 PROCEDURE load_image(image_path):
 TRY:
 READ original_image from image_path
 IF original_image IS NULL:
 THROW ERROR "Cannot load image"
```

```
CONVERT original_image from BGR to RGB
CATCH ERROR:
PRINT error message

PROCEDURE rgb_analysis():
SPLIT image into R, G, B channels

CREATE 3-subplot figure
FOR EACH channel, title, color:
CREATE histogram of channel
LABEL x-axis as "Pixel Intensity"
LABEL y-axis as "Frequency"

DISPLAY plot

PROCEDURE alternative_indices():
EXTRACT near_infrared (green), red, blue channels

CALCULATE indices:
NDVI = (near_infrared - red) / (near_infrared + red)
EVI = 2.5 * ((near_infrared - red) / (near_infrared + 6*red - 7.5*blue +
1))
SAVI = 1.5 * ((near_infrared - red) / (near_infrared + red + 0.5))

CREATE 3-subplot figure
FOR EACH index, title:
DISPLAY index using RdYlGn colormap

DISPLAY plot

PROCEDURE advanced_image_details():
CALCULATE image details:
- Image shape
- Image size
- Data type
- Average color (R, G, B)
- Brightness

PRINT image details with color

MAIN FUNCTION:
SET image_path
CREATE ImageAnalyzer object

CALL rgb_analysis()
CALL alternative_indices()
CALL advanced_image_details()

============ Pseudocode 1 ==============
```

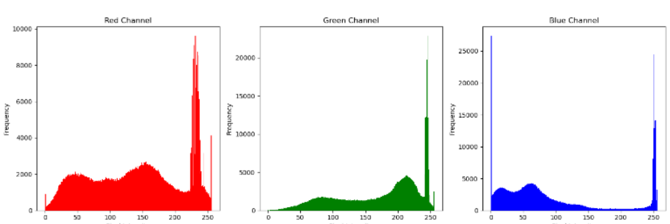

Fig. 21. RGB analysis of disability cameras.

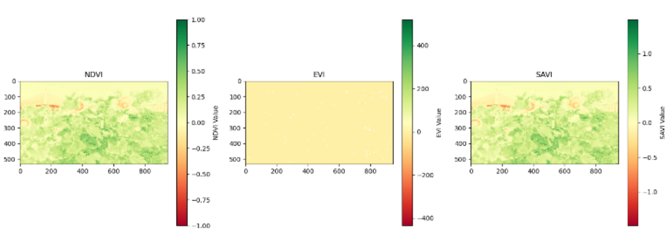

Fig. 22. Object detail color result and analysis of disability cameras.

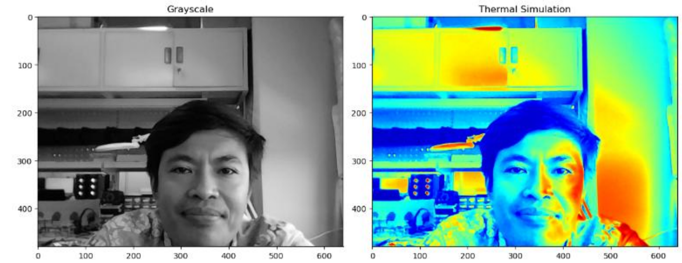

Fig. 23. Thermal camera for disabilities.

Furthermore, Fig. 21 and Fig. 22 are examples of detailed output about classifying detailed objects from colors, and possibly from detailed object shapes, it could also be the heat of the object produced, not only sound but images that provide output or sound feedback from various colors and object temperatures for disabilities as a Fig. 23 will be an idea for future research so that it can be applied properly for people with disabilities and the medical world.

IV. Conclusion

The Mini Smart Camera (MSC) system developed in this study has proven to be an effective assistive technology for individuals with disabilities. The system's object and speech recognition capabilities, combined with its IoT-based architecture, enable users to interact with their environment more independently and efficiently. The high levels of accuracy achieved in object and speech recognition demonstrate the potential of machine learning and IoT in developing advanced assistive technologies. Future research should focus on optimizing the system's performance in more complex environments and improving battery life to ensure prolonged use. In addition, expanding the system's capabilities to include more advanced features, such as facial recognition and gesture interpretation, could further enhance its utility for disabled users. Timestamps vs. Data Index, Timestamp increases as the data index increases, another analysis shows intervals between Consecutive Timestamps comparison between interval (seconds) and interval index with an increase of five interval index. The interval value in seconds ranges from 0.02 to 0.04 seconds. While at 0.14 seconds there is a significant spike. Moreover, the development of multispectral cameras can be used to develop technology for people with disabilities to make it more specific in the process of detecting objects around them.

V. Future Research

Images and voice recognition technology for people with disabilities will always be improved to get the best results and also be ideal and comfortable. The performance improvement of smart cameras and other devices lies in increasing the precision of images from input devices, and processing based on ideal specifications. Product development can be done in collaboration between universities and industry, as well as ethical feasibility for the medical. The development of mini cameras with multispectral technology and even hyperspectral cameras will be able to provide more detailed analysis in research development, especially for patients or the medical world, and also for people with disabilities to be more specific in conducting object detection in the future.

ACKNOWLEDGMENT

Thanks to the DRTPM Domestic Cooperation Research Grant at Langlangbuana University, BRIN, especially the Research Center for Telecommunication, Bandung and Pasundan University.

REFERENCES

[1] Khan, F., Amatya, B., Sayed, T.M., Butt, A.W., Jamil, K., Iqbal, W., Elmalik, A., Rathore, F.A. and Abbott, G., 2017. World Health Organization global disability action plan 2014-2021: challenges and perspectives for physical medicine and rehabilitation in Pakistan. *Journal of rehabilitation medicine*, *49*(1), pp.10-21.

[2] Farooq, M. S., Shafi, I., Khan, H., Díez, I. D. L. T., Breñosa, J., Espinosa, J. C. M., & Ashraf, I. (2022). IoT enabled intelligent stick for visually impaired people for obstacle recognition. Sensors, 22(22), 8914.

[3] Andriana, A., Zulkarnain, Z., Vertus, O., Rahman, S.A., Hamidah, I., Kustiawan, I., Barliana, M.S., Aryanti, T., Rohendi, D. and Riza, L.S., 2023, October. Converter of Indonesian sign language into text and voice, text and voice to sign language to build between inclusion vocasional school student and teacher. In AIP Conference Proceedings (Vol. 2510, No. 1). AIP Publishing.

[4] Andriana, A., Mulyanti, B., Widiaty, I., Zulkarnain, Z. and Wulandari, I.Y., 2024. Teacher Perceptions of Deep Learning Models for Special Need Students in Inclusive Vocational Schools: A Fuzzy Logic Analysis. Journal of Advanced Research in Applied Sciences and Engineering Technology, pp.15-24.

[5] Salah, A., Adel, G., Mohamed, H., Baghdady, Y. and Moussa, S.M., 2023. Towards personalized control of things using Arabic voice commands for elderly and with disabilities people. International Journal of Information Technology, pp.1-22.

[6] Semary, H.E., Al-Karawi, K.A., Abdelwahab, M.M. and Elshabrawy, A.M., 2024. A Review on Internet of Things (IoT)-Related Disabilities and Their Implications. Journal of Disability Research, 3(2), p.20240012.

[7] Xia, C., Chen, H., Han, J., Zhang, D. and Li, K., 2024. Identifying Children with Autism Spectrum Disorder via Transformer-Based Representation Learning from Dynamic Facial Cues. IEEE Transactions on Affective Computing.

[8] Agarwal, S., Terrail, J.O.D. and Jurie, F., 2018. Recent advances in object detection in the age of deep convolutional neural networks. arXiv preprint arXiv:1809.03193.

[9] Ahmed, M., Hashmi, K.A., Pagani, A., Liwicki, M., Stricker, D. and Afzal, M.Z., 2021. Survey and performance analysis of deep learning-based object detection in challenging environments. Sensors, 21(15), p.5116.

[10] Aziz, L., Salam, M.S.B.H., Sheikh, U.U. and Ayub, S., 2020. Exploring deep learning-based architecture, strategies, applications and current trends in generic object detection: A comprehensive review. Ieee Access, 8, pp.170461-170495.

[11] Katkar, G.V., Posonia, A.M., Ushanandini, D., Abirami, K. and Ankayarkanni, B., 2024, May. Google pi using raspberry pi for visually impaired. In 2024 3rd International Conference on Artificial Intelligence For Internet of Things (AIIoT) (pp. 1-6). IEEE.

[12] Andriana, A., Zulkarnain, Z., Wulandari, I.Y., Arrazaq, F. and Rahman, S.A., 2024. Technology and Disability: Building communication and Creating Opportunities?. Journal of Advanced Research in Applied Sciences and Engineering Technology, pp.25-34.

[13] Reddy, V.M., Vaishnavi, T. and Kumar, K.P., 2023, July. Speech-to-Text and Text-to-Speech Recognition Using Deep Learning. In 2023 2nd International Conference on Edge Computing and Applications (ICECAA) (pp. 657-666). IEEE.

[14] Liang, M. and Hu, X., 2015. Recurrent convolutional neural network for object recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 3367-3375).

[15] Ahmed, M., Hashmi, K.A., Pagani, A., Liwicki, M., Stricker, D. and Afzal, M.Z., 2021. Survey and performance analysis of deep learning-based object detection in challenging environments. Sensors, 21(15), p.5116.

[16] Mirani, I.K., Tianhua, C., Khan, M.A.A., Aamir, S.M. and Menhaj, W., 2022. Object recognition in different lighting conditions at various angles by deep learning method. arXiv preprint arXiv:2210.09618.

[17] Abdelaziz, T.A.I. and Fazil, U., 2023. Applications of integration of AI-based Optical Character Recognition (OCR) and Generative AI in Document Understanding and Processing. Applied Research in Artificial Intelligence and Cloud Computing, 6(11), pp.1-16.

[18] Gupta, A.D., Kumar, A., Chaudhary, I., Yasir, A.M. and Kumar, N., 2024, May. My Assistant SRSTC: Speech Recognition and Speech to Text Conversion. In 2024 International Conference on Communication, Computer Sciences and Engineering (IC3SE) (pp. 394-400). IEEE.

[19] Mamun, K.A., Nabid, R.A., Pranto, S.I., Lamim, S.M., Rahman, M.M., Mahammed, N., Huda, M.N., Sarker, F. and Khan, R.R., 2024. Smart reception: An artificial intelligence driven bangla language-based receptionist system employing speech, speaker, and face recognition for automating reception services. Engineering Applications of Artificial Intelligence, 136, p.108923.

[20] Hata, A., Wang, H., Yuwono, J. and Nomura, S., 2023. Assistive Technologies for Children with Disabilities in Inclusive and Special Schools in Indonesia.

[21] Almomani., A., et.al. 2023. Smart Shoes Safety System for the Blind People Based on (IoT) Technology. Computers, Materials and Continua Journal, Volume 76, Issue 1, 9 June 2023, Pages 415-436, doi. 10.32604/cmc.2023.036266

[22] Yang., J. et.al. 2024. Human–machine interaction towards Industry 5.0: Human-centric smart manufacturing. Digital Engineering. Volume 2, September 2024, 100013. doi. 10.1016/j.dte.2024.100013

[23] Adi., P.D.P., & Y. Wahyu. 2022. Performance evaluation of ESP32 Camera Face Recognition for various projects. February 2022Internet of Things and Artificial Intelligence Journal 2(1). DOI: 10.31763/iota.v2i1.512

[24] M. Niswar et al., "Performance evaluation of ZigBee-based wireless sensor network for monitoring patients' pulse status," 2013 International Conference on Information Technology and Electrical Engineering (ICITEE), Yogyakarta, Indonesia, 2013, pp. 291-294, doi: 10.1109/ICITEED.2013.6676255.

[25] Fransiska Sisilia Mukti, Puput Dani Prasetyo Adi, Dwi Arman Prasetya, Volvo Sihombing, Nicodemus Rahanra, Kristia Yuliawan and Julianto Simatupang, "Integrating Cost-231 Multiwall Propagation and Adaptive Data Rate Method for Access Point Placement Recommendation" International Journal of Advanced Computer Science and Applications(IJACSA), 12(4), 2021. http://dx.doi.org/10.14569/IJACSA.2021.0120494

[26] Adi.P.D.P, & Wahyu.Y. 2023. The error rate analyze and parameter measurement on LoRa communication for health monitoring. Microprocessors and Microsystems, Volume 98, April 2023, 104820, DOI: 10.1016/j.micpro.2023.104820