PROJECT PHASE 1 REPORT

ON

THIRDEYE -CONNECTING THE DOTS OF THE UNSEEN WORLD

Submitted by

IRENE MOLLY VARUGHESE (SJC20CS064)

to

the APJ Abdul Kalam Technological University

in partial fulfillment of the requirements for the award of the degree

of

Bachelor of Technology

in

Computer Science and Engineering

Department of Computer Science and Engineering
St. Joseph's College of Engineering and Technology, Palai

December :: 2023

Declaration

I undersigned hereby declare that the project report on **“ThirdEye-Connecting the dots of the unseen world”**, submitted for partial fulfillment of the requirements for the award of degree of Bachelor of Technology of the APJ Abdul Kalam Technological University, Kerala, is a bonafide work done by me under supervision of **Prof.Thushara Sukumar** .This submission represents my ideas in my own words and where ideas or words of others have been included, i have adequately and accurately cited and referenced the original sources. I also declare that i have adhered to ethics of academic honesty and integrity and have not misrepresented or fabricated any data or idea or fact or source in my submission. I understand that any violation of the above will be a cause for disciplinary action by the institute and/or the University and can also evoke penal action from the sources which have thus not been properly cited or from whom proper permission has not been obtained. This report has not been previously formed the basis for the award of any degree, diploma or similar title of any other University.

Name and Signature of Student
IRENE MOLLY VARUGHESE

Place:Choondacherry
Date:07-12-2023

ST. JOSEPH'S COLLEGE OF ENGINEERING AND TECHNOLOGY, PALAI

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the report entitled "**THIRDEYE-CONNECTING THE DOTS OF THE UNSEEN WORLD**" submitted by **IRENE MOLLY VARUGHEESE (SJC20CS064)** to the APJ Abdul Kalam Technological University in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineering is a bonafide record of the project work carried out by her under my guidance and supervision.

Project Guide
Prof.Thushara Sukumar
Assistant Professor
Department of CSE

Project Coordinator
Prof. Kishore Sebastian
Assistant Professor
Department of CSE

Head of Department
Dr. Joby P P
Professor & Head
Department of CSE

.
.
Place : Choondacherry
Date : 07-12-2023

Acknowlegment

The success and final outcome of this project phase 1 required a lot of guidance and assistance from many people, and we are extremely privileged to have received their support throughout the completion of this project. All that i have accomplished is only possible due to their supervision and assistance, and i am sincerely grateful to them.

I would like to express our respect and gratitude to the management of St. Joseph's College of Engineering and Technology for providing us with the opportunity and platform to work on this project.

A special word of thanks goes to our beloved Principal, **Dr. V. P. Devassia**, for providing invaluable support and necessary facilities to carry out this project.

I am extremely indebted to **Dr. Joby P.P, Professor** & Head of the Department of Computer Science and Engineering, for his valuable suggestions and encouragement throughout the course of this project work.

I would also like to express my gratitude to our project coordinator, **Prof. Kishore Sebastian**, Assistant Professor in the Department of Computer Science and Engineering, for his valuable suggestions and guidelines during the entire duration of this project. I truly appreciate his contributions and technical support in preparing this report.

My heartfelt thanks go to my project guide, **Prof. Thushara Sukumar**, Assistant Professor in the Department of Computer Science and Engineering, displayed a keen interest in this project and provided guidance and all the necessary information for developing a robust system.

I am thankful and fortunate enough to have received constant encouragement, sup- port, and guidance from all the staff members of the Department of Computer Science and Engineering. Their assistance played a crucial role in the successful completion of my project phase 1 work.

Abstract

Blind and visually impaired people have encountered a lot of challenges when performing most of the natural activities performed by non-disabled people. In particular, many dangerous situations occur in environments that are unfamiliar to them. This IoT project aims to enhance the mobility and independence of visually impaired individuals by developing a wearable device that can be attached to their spectacles. The proposed system uses IoT technologies to create a comprehensive environment perception and navigation aid, assisting blind individuals in recognizing and interacting with their surroundings more effectively. The device leverages Arduino technology and is seamlessly connected to a mobile phone. Its primary functionality is to detect objects in the user's vicinity and provide real-time audio cues to assist with navigation and obstacle avoidance. The system utilizes ultrasonic or infrared sensors to continuously scan the surroundings, measuring distances to nearby objects. The Arduino microcontroller processes this data and communicates wirelessly with a mobile phone application. The project emphasizes user-friendly design and accessibility, ensuring that visually impaired individuals can easily configure and use the device. Through this innovative IoT solution, visually impaired individuals gain increased awareness of their surroundings, improving their ability to navigate independently and with confidence. This project demonstrates the potential of IoT and Arduino technology to address real-world challenges and enhance the quality of life for individuals with visual impairments.

Contents

List of Figures

Chapter 1

Introduction

1.1 Problem Statement

Visually impaired individuals confront a myriad of challenges that significantly impede their daily lives, jeopardizing their independence and overall well-being. The absence of visual cues severely impedes their ability to detect and navigate obstacles, ranging from uneven surfaces to unexpected barriers, increasing the risk of accidents and injuries. Simultaneously, the limited access to real-time environmental information exacerbates their struggles, hindering their ability to perceive surroundings, recognize objects, and identify potential hazards, perpetuating a constant state of vulnerability. The challenge extends indoors, where visually impaired individuals encounter difficulties in various indoor activities. Navigating unfamiliar indoor spaces becomes particularly problematic due to obstacles, changes in elevation, and dynamic environments that may lack tactile or auditory cues. The inherent reliance on others for assistance not only diminishes their autonomy but also contributes to feelings of isolation and dependency.

These multifaceted challenges underscore the urgent need for comprehensive and innovative solutions to empower visually impaired individuals, providing them with the tools to navigate both indoor and outdoor environments with confidence and independence.Current assistive technologies for the visually impaired often lack real-time and intuitive feedback, making it difficult for them to detect and avoid obstacles effectively. Existing solutions either tend to be expensive and complex or lack the adaptability required for diverse environments. There is a pressing need for an affordable, user-friendly, and portable assistive device that can seamlessly integrate with a smartphone, offering real-time object detection and providing immediate audio cues to aid navigation.

To address the complex challenges faced by visually impaired individuals, this project focuses on the development of an innovative assistive technology system. This system is centered around an Internet of Things (IoT) framework, incorporating key features such as object detection, navigation assistance, text reading capabilities, live voice feedback, and a sensor that emits a beep sound as an alert when obstacles approach the user. The hardware components of this solution are compactly designed to be attached to the user's spectacles, ensuring a seamless and unobtrusive integration into their daily lives. A camera module is strategically positioned on one side of the spectacles to capture live video, providing real-time visual information to the user. On the other side, the sensor, which detects obstacles, enhances the user's spatial awareness by emitting audible cues. This comprehensive system is powered by Arduino technology, leveraging its versatility and efficiency in processing data and managing interactions.

The primary objective of this project is to provide a practical and accessible solution that empowers visually impaired individuals to navigate independently, confidently, and safely in diverse environments. By harnessing the capabilities of IoT, our solution aims to bridge the gap in information accessibility for the visually impaired, offering them crucial details about their surroundings through a user-friendly interface. The integration of live voice feedback and text reading further contributes to the user's autonomy, enabling them to receive pertinent information audibly. This project aims to tackle the challenges faced indoor environments, recognizing the difficulties encountered in various indoor activities. Through the development of this assistive technology system, we aspire to enhance the overall quality of life for visually impaired individuals, fostering inclusivity and independence. The scope of this project encompasses the design, implementation, and evaluation of the proposed system, with a focus on usability, effectiveness, and adaptability to different scenarios and user needs.

Chapter 2

Literature Review

2.1 Collision detection and prevention for the visually impaired using computer vision and machine learning

In their collaborative work in 2023, Shivang Sunil Singh, Mayank Agrawal, and M Eliazer introduce a cutting-edge approach to collision detection and prevention for the visually impaired, leveraging computer vision and machine learning. Their methodology involves the detection and localization of obstacles, with the obtained information relayed to the visually impaired individual through a voice interface. A distinctive feature of their approach is the utilization of cloud services and machine learning, recognizing machine learning as the optimal tool for achieving their goals and adapting to technological advancements. The paper details the process of sending captured video to a cloud server through an API, emphasizing how this cloud-based approach allows for efficient processing of information about the surroundings. This innovative integration of computer vision, machine learning, and cloud services signifies a forward-looking strategy to enhance the functionality and effectiveness of collision detection systems, ultimately contributing to improved safety and mobility for individuals with visual impairments.[1]

2.2 Object Recognition and Speech Generation for Visually Impaired

In their pioneering project, Kohli and Agarwal (2022) ingeniously combine the YOLOv5 image detection model with two powerful Python text-to-speech conversion libraries, pyttsx3 and gTTS, to create an assistive system tailored for the visually impaired. Trained on a custom dataset of 15 essential objects and the extensive MS COCO 2017 Dataset, the YOLOv5 model boasts superior accuracy in object detection. The authors' meticulous comparison of pyttsx3 and gTTS underscores the critical role of text-to-speech conversion, with the choice of library influencing the overall user experience. This innovative fusion not only promises heightened accuracy in both detection and speech generation but also emphasizes a pragmatic approach by transforming visual information into audio, effectively presenting the detected objects to the visually impaired through a speaker. The authors' commitment to inclusivity is evident in the development of a custom dataset, reflecting a profound understanding of the unique needs of the visually impaired and marking a significant advancement in the realm of AI-driven assistive technologies.[4]

2.3 Speech to text conversion and summarization for effective understanding and documentation

In their insightful work, Vinnarasu A. and Deepa V. Jose (2021) address the paramount role of speech as a potent means of human communication, serving as a conduit for the expression of thoughts and emotions in diverse languages. Recognizing the inherent variations in speech, the authors shed light on the challenges arising from linguistic nuances, pacing differences, and regional dialects within a language, all of which contribute to potential misunderstandings. The research underscores the hurdles posed by lengthy speeches, where complexities in pronunciation and pacing intensify the difficulty of comprehension. To tackle these challenges, the authors delve into the realm of speech recognition, a dynamic field within computational linguistics dedicated to translating spoken words into text. Additionally, they explore the significance of text summarization, a critical process that distills essential information from source texts, facilitating effective understanding and documentation. This work not only highlights the evolving landscape of communication technologies but also emphasizes the invaluable contributions of speech-to-text conversion and summarization in overcoming barriers to comprehension and enhancing the efficiency of information extraction and documentation. [19]

2.4 End-to-End Data Authentication Deep Learning Model for Securing IoT Configurations

In his groundbreaking work, Mohamed Hammad (2022) introduces a pioneering End-to-End Data Authentication Deep Learning Model designed to fortify IoT configurations. At the heart of this innovation lies the utilization of edge servers—hardware entities strategically positioned at the network's periphery, in close physical proximity to end-users and on-site applications. These edge servers function as micro-data centers, seamlessly integrating computing, networking, and storage capabilities. Notably, IoT devices, characterized by their autonomy in data reception and transmission, form the crux of this security paradigm. Hammad leverages the principles of edge computing, positioning computational services at or near the physical locations of users or data sources. This strategic placement not only ensures swifter and more reliable services, enriching user experiences, but also empowers companies to bolster support for latency-sensitive applications, discern trends, and elevate the quality of products and services. Hammad's work thus not only contributes to the evolving landscape of IoT security but also embraces the transformative potential of edge computing in fortifying data integrity and user interactions. His model, backed by robust edge servers, provides a resilient framework for end-to-end data authentication, ensuring a secure and efficient environment for IoT configurations.[11]

2.5 Assistive technology for Visually Impaired

In pursuit of empowering visually impaired individuals to learn and achieve their goals, a groundbreaking assistive technology has emerged. The devised solution ingeniously combines a smartphone with a laser pointer to create a virtual white cane, revolutionizing the traditional aid for the visually impaired. The working mechanism is orchestrated through the smartphone's camera capturing the reflection of the laser pointer's beam, and employing active triangulation to calculate the distance to objects in the user's path. The genius of the system lies in its feedback mechanism—a personalized vibration generated by the smartphone, where the magnitude correlates with the calculated distance. This tactile feedback serves as an invaluable guide, enabling users to navigate their surroundings with heightened awareness and avoid collisions with obstacles. The research not only culminated in the development of an accessible and affordable virtual white cane but also established a pioneering methodology for delivering personalized vibratory feedback, marking a significant leap forward in the realm of assistive technology for the visually impaired.[18]

2.6 Wearable Technology for Visually Impaired

In the innovative exploration conducted by Rahul Kevadia in 2020, the focus is on the transformative potential of wearable technology, a category encompassing smart devices designed to be worn as accessories. Notably, these devices, including activity trackers and smartwatches, exemplify the broader concept of the Internet of Things (IoT), leveraging electronic devices to monitor daily activities. Kevadia emphasizes the multifaceted applications of wearable technology, ranging from navigation to health tracking, and envisions a significant leap forward when combined with image processing techniques. The project's specific goal is particularly commendable—a quest to develop smart wearable glasses tailored for visually impaired individuals. These glasses are envisioned to serve as indispensable aids, facilitating the reading of signboards and newspaper headlines, thereby addressing challenges faced by the visually impaired community. Kevadia's work not only underscores the boundless potential of wearable technology but also exemplifies its capacity to bring about meaningful advancements in accessibility and assistance for individuals with visual impairments.[18]

2.7 Advanced Audio Aid for Blind People

In the collaborative endeavor by Savera Sarwar and Danish Channa in 2022, a remarkable advanced audio aid system for the visually impaired takes center stage. This innovative solution operates on the versatile Raspberry Pi platform, integrating cutting-edge technologies to enhance accessibility. Object detection is powered by the YOLO (You Only Look Once) algorithm, trained on the COCO (Common Objects in Context) dataset, ensuring robust recognition of diverse objects in the surroundings. Complementing this, OCR (Optical Character Recognition) technology, implemented through Tesseract, facilitates text extraction from images. The synergy of these technologies converges to create a comprehensive audio aid system. The text-to-voice conversion is executed seamlessly using the pyttsx3 library in Python, offering a natural and intelligible auditory output. This collaborative effort not only showcases the convergence of state-of-the-art algorithms and libraries but also underscores the practicality and portability of the Raspberry Pi as an operating system for empowering the visually impaired with advanced audio assistance.[12]

2.8 Object Detection And Recognition Using TensorFlow For Blind People

In the collaborative work by P Devika, S P Jeswanth, and Billu Nagamani in 2022, a forward-thinking solution emerges to address the challenges faced by the visually impaired. This innovative system centers around TensorFlow, a powerful deep learning framework, to enable object detection and recognition. The process begins with a Video Capturing Module, which serves as input to the COCO dataset, ensuring a diverse range of objects for comprehensive recognition. The Image Processing Module, powered by the OpenCV library in Python, enhances the quality of captured images, optimizing them for subsequent analysis. Leveraging the SSD (Single Shot Multibox Detector) architecture within the TensorFlow API, the Object Detecting Module ensures efficient and accurate identification of objects in real-time. The Distance Calculation Module, utilizing numpy, adds a crucial spatial dimension by determining the distance of detected objects. To convey this information intuitively to the user, the Audio Output Module employs the pyttsx3 pip package, offering a natural and informative auditory experience. This collaborative effort not only showcases the synergy of advanced technologies but also emphasizes the practicality and effectiveness of TensorFlow in creating an accessible and empowering solution for the visually impaired.[13]

2.9 Object Detection System with Voice Alert for Blind

In the collaborative research conducted by Dr. M Y Babu, Akash Jatavath, and G Yashwanth Kumar Reddy in 2023, an innovative object detection system tailored for the visually impaired is introduced, anchored by the YOLO (You Only Look Once) algorithm. Trained on the COCO dataset, the system ensures a comprehensive recognition of diverse objects in real-time. A pivotal element in enhancing object localization accuracy is the Estimation of Image Position, articulated by the formula BC = pr * IOU, where BC signifies the bounding box, pr denotes the object's presence in the box, and IOU represents the Intersection over Union. This formulaic approach provides a precise estimation of the object's position in the image. The synthesized voice alerts, integral to the system, are realized through a dual methodology, employing the Pyttsx3 library for text-to-speech conversion and Google Text to Speech. This collaborative effort not only showcases the application of advanced algorithms but also underscores a formulaic precision in addressing the challenges faced by the visually impaired, paving the way for a more inclusive and accessible technological landscape.[7]

2.10 Real Time Object Detection With Speech Recognition Using TensorFlow Lite

In the collaborative endeavor led by Ganesh Khekare and Kalpeshkumar Solanki in 2022, a cutting-edge real-time object detection system with integrated speech recognition unfolds, leveraging the prowess of TensorFlow Lite. The system commences with an RGB camera module capturing live visual data, seamlessly interfaced with an Android app. This app efficiently forwards the captured images directly to the TensorFlow network for robust object detection, employing the SSD (Single Shot Multibox Detector) architecture. The transformative element in this innovation lies in the voice feedback mechanism, wherein a bespoke audio library housing 8 distinct object voices is crafted and deployed on the Android side. This amalgamation of real-time object detection and voice feedback not only showcases the flexibility and efficiency of TensorFlow Lite but also underscores the potential for creating a seamless and accessible user experience through the integration of cutting-edge technologies on mobile platforms.[17]

2.11 Deep Learning Based Object Detection And Surrounding Environment Description For Visually Impaired People

In the collaborative initiative spearheaded by Raihan Bin Islam, Samiha Akhter, and Faria Iqbal in 2023, a transformative solution takes shape, aimed at empowering individuals with visual impairments through the application of deep learning. Operating on the versatile Raspberry Pi embedded system, this comprehensive system sets a new standard by commencing with a sophisticated Video Capturing Module that utilizes the Raspberry Pi camera to capture real-time video, forming the bedrock of a revolutionary user experience. The implementation of the state-of-the-art SSDLite MobileNetV2 model, meticulously trained on the COCO dataset, ensures the system's proficiency in accurate and efficient recognition of a diverse array of objects. Going beyond conventional object detection, the system introduces an innovative Ambiance Mode, harnessing an open-source weather dataset from Kaggle to train the model to eloquently describe the surrounding environment. This holistic approach is further enriched by the incorporation of Voice Feedback, a dynamic component that employs the Google Text-to-Speech module, PyAudio, and playsound to provide intuitive auditory cues and detailed descriptions. This amalgamation of cutting-edge technologies not only highlights the immense potential of deep learning in enhancing accessibility but also emphasizes the practicality, portability, and user-centric design inherent in the Raspberry Pi embedded system, offering a holistic and empowering solution for individuals with visual impairments.[5]

2.12 An Outdoor Navigation Assistance System For Visually Impaired People In Public Transportation

Salvador Martinez Cruz, Luis Morales-Heranandenz, Gerardo I, Juan P, and Karla A's groundbreaking work introduces an advanced Outdoor Navigation Assistance System tailored to empower visually impaired individuals in navigating public transportation seamlessly. Central to this system is the strategic utilization of Bluetooth Low Energy (BLE) technology for location and communication, facilitated through a dedicated mobile application named SUBE. BLE beacons strategically placed on buses and their stops are tracked in real time by the SUBE app, delivering timely and pertinent information through voice messages to users. Operating on the Android Studio Platform (Android 5.1 Lollipop), SUBE not only provides verbal instructions but also furnishes details about transportation lines, destination information, the name of the next stop, and the user's current location. This comprehensive and user-centric approach underscores the integration of cutting-edge BLE technology and sophisticated mobile applications, illustrating a dedicated commitment to fostering inclusivity and autonomy for visually impaired individuals during their public transportation journeys.[15]

2.13 Head Posture Estimation by Deep Learning Using 3-D Point Cloud Data From a Depth Sensor

In their innovative work, Seiji Sasaki and Chinthaka Premachandra (2021) present a pioneering approach to head posture estimation through deep learning, utilizing 3-D point cloud data acquired from a RealSense D435 depth sensor. The method involves capturing characteristic facial areas and extracting head data from the generated point cloud. PyTorch is employed as the deep learning library, optimized with GPU processing for accelerated computations. The RealSense D435 depth sensor serves as a crucial tool in capturing detailed 3-D facial data, which is then translated into point cloud data using Open 3D. To address the challenges associated with real-time estimation, the authors employ a clustering method, DBSCAN, to classify each point in the generated cloud. Subsequently, the data corresponding to the head is extracted from the classified point cloud for input into the neural network for posture estimation. The study acknowledges the computational challenges posed by the clustering method, necessitating further exploration for real-time applications. Additionally, the authors highlight the fixed distance from the camera to the person, emphasizing the need for flexibility in camera placement for practical use. This comprehensive methodology showcases the integration of cutting-edge technologies in computer vision, deep learning, and depth sensing, laying the groundwork for improved head posture estimation techniques with enhanced applicability in real-world scenarios.[16]

2.14 Smart Assistive System for Visually Impaired People Obstruction Avoidance Through Object Detection and Classification

In the collaborative work led by Usman Masud, Tareq Saeed, Hunida M, Fezan UL, and Ghulam Abbas in 2022, a transformative smart assistive system is introduced, focusing on obstruction avoidance for visually impaired individuals. This system integrates a Raspberry Pi 4B, a camera, an ultrasonic sensor, and an Arduino, all mounted on an individual's walking stick. The technology captures images of the surroundings, preprocessing them with Viola-Jones and TensorFlow object detection methodologies. An ultrasonic sensor, controlled by a servomotor, measures distances between the blind person and obstacles, producing distinct beep sounds for left, right, and multiple-sided obstructions. Left and right tracking is facilitated by the Raspberry Pi camera, capturing a 120-degree view. The servomotor adjusts accordingly to provide real-time obstacle clearance. This comprehensive approach not only showcases the integration of diverse technologies but also emphasizes the practicality and effectiveness of the smart assistive system in providing enhanced safety and autonomy for visually impaired individuals during navigation.[9]

2.15 Pedestrian Lane Detection For Vision-Impared

In their collaborative effort, Yunjia Lei, Son Lam, Abdesselam Bouzerdoum, Hoang Thanh Le, and Khoa Luu (2022) contribute to the realm of assistive technologies with a focus on Pedestrian Lane Detection designed specifically for the visually impaired. This technology serves the crucial purpose of providing essential information about walkable regions, aiding blind individuals in staying on pedestrian lines, and enhancing obstacle detection for a safer navigation experience. The authors categorize their approach into two main categories: Traditional Methods and Deep Learning Methods. Under Traditional Methods, the team explores color-based approaches, border-based approaches, and the synergistic combination of these methods. These traditional techniques draw on established principles, offering a foundational understanding of pedestrian lane detection. On the other hand, the authors delve into the realm of Deep Learning Methods, where they explore lane detection approaches and generic semantic segmentation approaches. This signifies a departure from traditional methodologies, embracing the power of deep learning models to discern and categorize pedestrian lanes and surrounding environments. This comprehensive approach showcases the authors' commitment to addressing the challenges faced by the visually impaired, providing a multifaceted solution that integrates both traditional and cutting-edge techniques. The Pedestrian Lane Detection system presented in this work stands as a testament to the potential of technology in fostering accessibility and inclusivity for individuals with visual impairments.[10]

2.16 Development of an optical sensor capable of measuring distance, tilt,contact force

In the collaborative work led by Takahiro Nozaki and Hermano Igo Krebs in 2022, a groundbreaking optical sensor emerges, marking a significant advancement in the realm of robotic object manipulation. Unlike traditional sensors that specialize in depth, proximity, or tactile measurements, this novel sensor offers a unique capability—it can simultaneously measure millimeter distance, surface tilt, and contact force. The sensor comprises multiple distance measuring units that utilize infrared irradiation from light-emitting diodes, alongside springs and a transparent sheet. This ingenious combination allows for the measurement of distance when the object is located farther from the equilibrium length of the springs, providing a comprehensive understanding of the object's position, orientation, and contact force. This development has promising implications for enhancing precision and control in robotic manipulation tasks, showcasing the potential for innovative sensor technologies in advancing the field of robotics.[8]

2.17 Dynamic Crosswalk Scene Understanding for the Visually Impaired

In their 2021 collaborative work, Shishun Tian, Minghuo Zheng, Wenbin Zou, Xia Li, and Lu Zhang embark on a revolutionary journey to advance blind navigation systems, with a particular focus on dynamic crosswalk scene understanding for the visually impaired. Their innovative approach transcends traditional reliance on non-vision sensors, opting for the integration of vision sensors like RGB and RGB-D coupled with digital image processing and deep learning techniques. This amalgamation not only demonstrates a departure from conventional methods but also signifies a paradigm shift in the pursuit of more accurate obstacle detection for individuals with visual impairments. The implementation of floor segmentation to identify obstacle-free paths and semantic segmentation for traversable areas further showcases their commitment to providing a comprehensive solution for navigating complex urban environments. This groundbreaking technology not only enhances the precision of blind navigation but also marks a substantial stride in empowering individuals with visual impairments to navigate seamlessly and independently in dynamic city landscapes.[14]

2.18 Autonomous path planning with obstacle avoidance for smart assistive systems

In their 2023 collaborative work, Charis Ntakolia, Serafeim Moustakidis, and Athanasios Siouras contribute to the field of smart assistive systems with a focus on autonomous path planning and obstacle avoidance. Emphasizing dynamic adaptability, their system is designed to dynamically extract spatiotemporal information, identify objects, and adjust routes in real-time, ensuring flexibility in response to the user's environment. Recognizing the practical constraints of wearable assistive systems, the study underscores the importance of energy and size efficiency, particularly for devices like smart glasses, where concerns about energy demands, size, and weight are paramount. The proposed methodology introduces a novel chaotic ant colony optimization algorithm with fuzzy logic (CACOF) for global path planning, addressing macroscale navigation. For microscale navigation involving local path planning and obstacle detection, the system leverages powerful yet lightweight deep convolutional neural networks. This comprehensive approach not only reflects a commitment to addressing real-world challenges in assistive navigation but also showcases the integration of innovative algorithms and technologies for efficient and adaptive smart assistive systems.[2]

2.19 Vision-based environment perception and autonomous obstacle avoidance for unmanned underwater vehicle

In their work in 2023, P. Yao, X. Sui, Y. Liu, and Z. Zhao introduce a comprehensive approach to vision-based environment perception and autonomous obstacle avoidance for unmanned underwater vehicles. A notable contribution is the deployment of a compressed YOLOv5s model, specifically designed to reduce model volume, runtime memory footprint, and computing operations while maintaining high efficiency, enhancement effect, and recognition accuracy. The authors propose an Adapted Modified Guidance Vector Field (AMGVF) as a key component for autonomous navigation, offering guidance for effective obstacle avoidance. This paper exhibits a well-thought-out amalgamation of image enhancement, object detection, data fusion, and obstacle avoidance techniques, emphasizing efficiency and real-time applicability. The utilization of a compressed YOLO model and the introduction of AMGVF underscore the authors' commitment to advancing autonomous underwater vehicle capabilities, making strides in both performance and resource efficiency for enhanced real-world applications.[6]

2.20 Embedded implementation of an obstacle detection system for blind and visually impaired persons' assistance navigation

In their collaborative work in 2023, Ahmed Ben Atitallah, Yahia Said, Mohamed Amin Ben Atitallah, Mohammed Albekairi, Khaled Kaaniche, Turki M. Alanazi, Sahbi Boubaker, and Mohamed Atri present a novel approach to assistive navigation for blind and visually impaired individuals through the embedded implementation of an obstacle detection system. The proposed system centers around an enhanced YOLO v5 neural network, with a specific focus on optimizing both speed and detection accuracy. To achieve this, the authors integrate DenseNet into the YOLO v5 backbone, aiming to enhance feature reuse and data transfer mechanisms. This integration is designed to contribute to the overall improvement of detection accuracy, aligning with the objective of providing an efficient and accurate obstacle detection system to enhance the navigation experience for individuals with visual impairments.[3]

Chapter 3

System Requirements

3.1 Functional Requirements

3.1.1 Hardware Requirements

- Arduino UNO R3 : Arduino UNO is a low-cost, flexible, and easy-to-use programmable open-source microcontroller board that can be integrated into a variety of electronic projects. This board can be interfaced with other Arduino boards, Arduino shields, Raspberry Pi boards and can control relays, LEDs, servos, and motors as an output.
- OV7670 Camera Module and ESP32:The OV7670 Camera Module is a small image sensor, with low operating voltage, providing all functions of a single chip of VGA camera and image processor.
- HC-SRO4 Ultrasonic Sensor: HC-SR04 stands for High-Conductance Ultrasonic Sensor consists of a transmitter and receiver. The sensor measures how far things are without touching them, and it uses sound waves to get the measurements right. It can work well when things are between two to four centimeters away.

- Smartphone with 8GB RAM and 218GB ROM - For real-time auditory feedback.
- Laptop with 8GB RAM (If the laptop has graphics card performance can be improved)

3.1.2 Software Requirements

- TensorFlow : It is an open source machine learning framework . The most popular version of TensorFlow is TensorFlow 2.0, which was released by Google in September 2019. It can be used in a wide variety of programming languages, including Python, JavaScript, C++, and Java.
- Pyttsx 3 : Unlike pyttsx 2, which supports only Python 2.x, pyttsx 3 is modified to work on both Python 2.x and Python 3.x with the same code. This makes it more versatile and adaptable to different Python.
- ROI: Stands for Region of Interest, It refers to a specific area or region within an image that is of particular interest or importance. This region is usually defined by a bounding box or a polygon, and it is used to extract features or perform further analysis on the image.

3.2 Non Functional Requirements

- User Interface: The device should provide information in a non-visual form in a simple and rapid manner. It should make full use of functions other than vision, such as hearing and touch.
- Design: The device should be lightweight, easy to use, and not a source of nuisance or disruption. It should also have a simple and low-cost construction
- Performance: The device should be power-efficient and able to run in real-time with sufficient accuracy.

- Reliability: The device should be reliable, providing consistent and accurate information to the user
- Integration: The device should be able to integrate with the user's existing spectacles or wearable items.
- Safety: The device should not pose any safety risks to the user.
- Privacy: The device should respect the user's privacy and not collect or store personal data unless necessary and with the user's consent.
- Accessibility: The device should be accessible to all visually impaired individuals, regardless of the level of visual impairment.
- Durability: The device should be robust and durable, able to withstand daily wear and tear.

3.3 Module Description

3.3.1 Floor segmentation

This module separates the floor plane from the rest of the environment. It is an essential step in the process of detecting obstacles and planning a path to avoid them. The module uses a depth camera to capture the 3D structure of the environment and segment the floor plane from the rest of the scene.

3.3.2 Occupancy grid

This module creates a map of the environment. It uses the depth data from the camera to create a 2D grid representation of the environment. The grid is divided into cells, and each cell is labeled as either occupied or unoccupied based on the depth data.

3.3.3 Navigation

This module detects obstacles in the environment and generates a path to avoid them. It uses the occupancy grid to detect obstacles and plan a path around them. The module also uses a haptic feedback system to alert the user to the presence of obstacles.

3.3.4 Object detection

This module detects objects of interest in the environment. It uses a deep learning model to detect objects such as doors, chairs, and tables. The module provides audio feedback to the user to identify the detected objects.

3.3.5 Path planning

This module plans a path to the destination. It uses the occupancy grid and obstacle avoidance module to plan a path from the user's current location to the destination. The module also takes into account the user's preferences, such as avoiding stairs or taking the shortest path.

3.3.6 Haptic feedback

This module provides haptic feedback to the user. It uses a vibration motor to provide feedback to the user about the location of obstacles and the direction of the path.

3.3.7 Text to speech

TTS module can be used to convert text-based information into speech output. This can be helpful for the visually impaired to access information such as weather updates, news articles, and other digital content. One of the popular Python libraries for TTS conversion is **pyttsx3**. It is a cross-platform library that works offline and supports multiple TTS engines, including Sapi5, nsss, and espeak. The library is compatible with both Python 2.x and Python 3.x.

3.3.8 Live auditory feedback

Live auditory feedback during navigation and object detection is a technique that provides real-time audio cues to help visually impaired individuals navigate their surroundings. The technique uses computer vision-based object detection to identify objects in the environment and provide audio feedback to the user. For example, the system can detect obstacles such as walls, doors, and furniture and provide audio cues to help the user navigate around them. The system can also provide audio feedback to help the user locate specific objects or landmarks.

Chapter 4

System Design

4.1 Use Case Diagram

A use case diagram is a graphical representation of a system that shows the interactions between the system and external actors (users, systems, or other actors). It is used to describe the functionality of a system in a clear and concise way. A use case diagram consists of actors, use cases, and the relationships between them. Actors are the external entities that interact with the system, while use cases represent the functionality of the system from the user's perspective. In Figure 4.1, the user is represented by a stick figure on the left side of the diagram, the system is represented by an oval in the center of the diagram, and the admin is represented by a stick figure on the right side of the diagram. The interactions between the actors are represented by arrows. The use case diagram shows the interactions between the user, the system, and the admin. The user can access the system for text-to-speech, object recognition, navigation, and user management. The admin can access the system for user management.

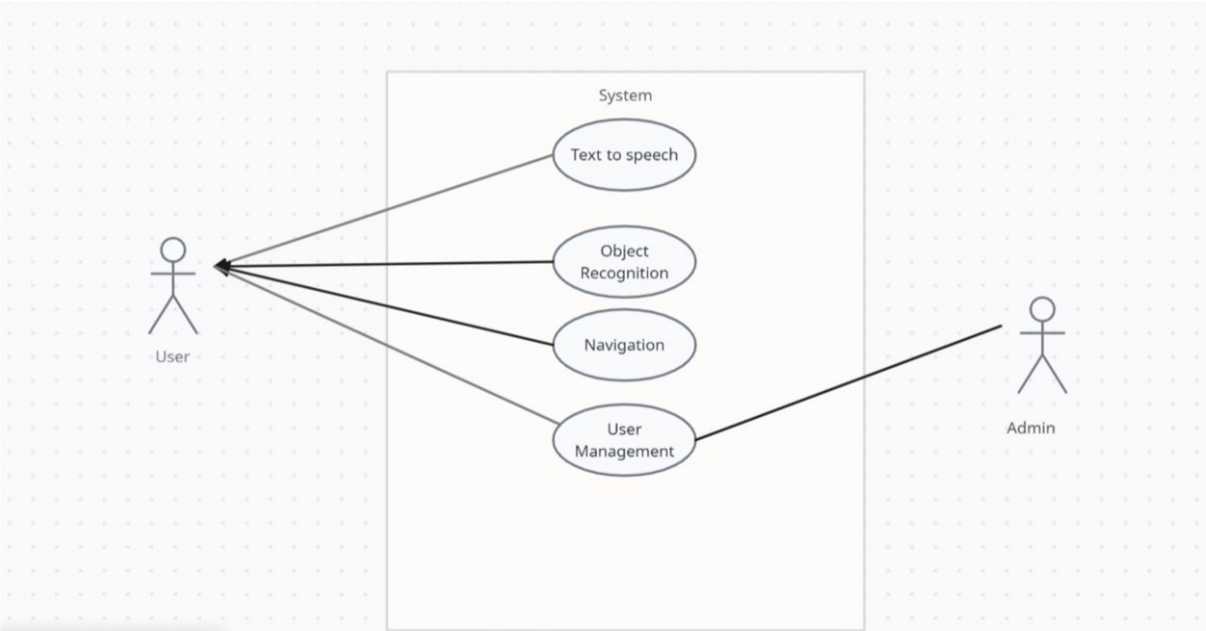

Figure 4.1: Use case diagram

4.2 Activity Diagram

An activity diagram is a behavioral diagram that illustrates the flow of activities through a system. It is used to model the behavior of a system or process in a clear and structured way, making it easier to understand and analyze. Activity diagrams are similar to flowcharts but with more specific symbols and notations. They describe parallel and conditional activities, use cases, and system functions at a detailed level. In Figure 4.2 it explains the process of a user interacting with the device. The flowchart starts with the user uploading an image and the device recognizing it. If the device is unable to recognize the image, it will say "unable to recognize" and the process will end. If the device is able to recognize the image, it will perform other tasks such as navigation, providing text or audio guidance, and answering user questions.

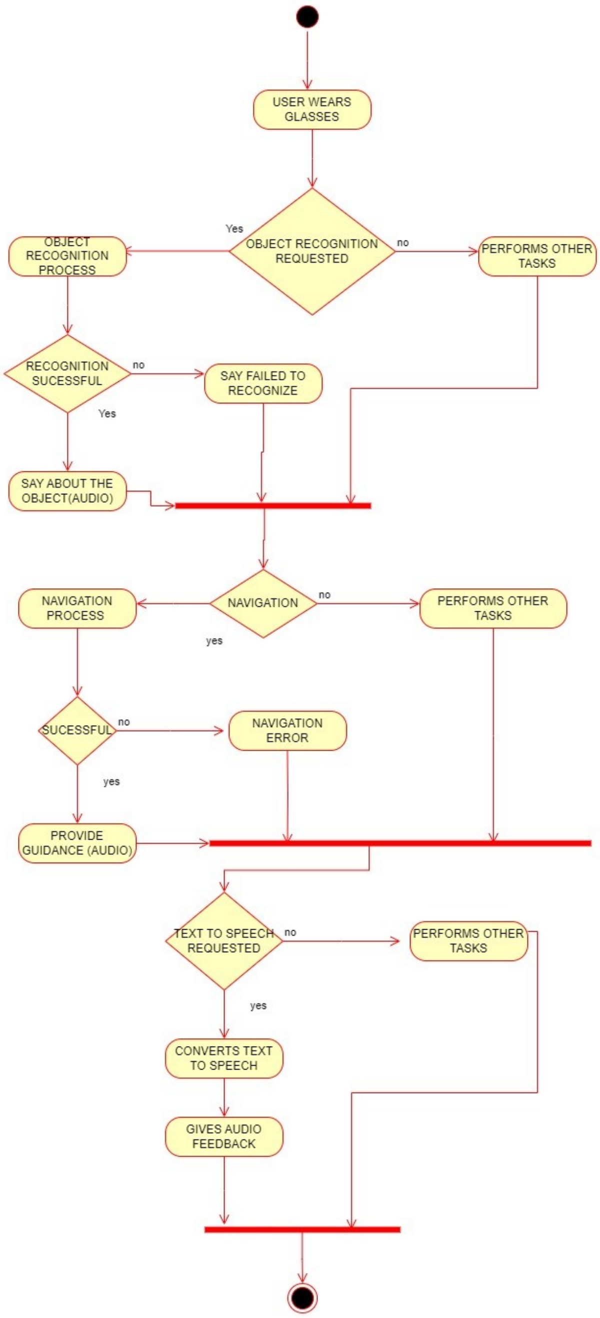

Figure 4.2: Activity diagram

4.3 Sequence Diagram

A sequence diagram is a type of interaction diagram that details how operations are carried out. It is organized according to time, with the time progressing as you go down the page. The objects involved in the operation are listed from left to right according to when they take part in the message sequence. In Figure 4.3 the interactions between different components in a system. It is used to visualize the flow of messages between different components and how they interact with each other. In this particular diagram, the components are "User", "Object recognition", "Navigation", "Text to speech", "Auditory feedback", and "Haptic feedback". The interactions between these components are shown as arrows, with the direction of the arrow indicating the direction of the message flow. The diagram also shows the different types of messages that are being sent between the components, such as "Request object recognition", "Provide haptic feedback", and "Process feedback".

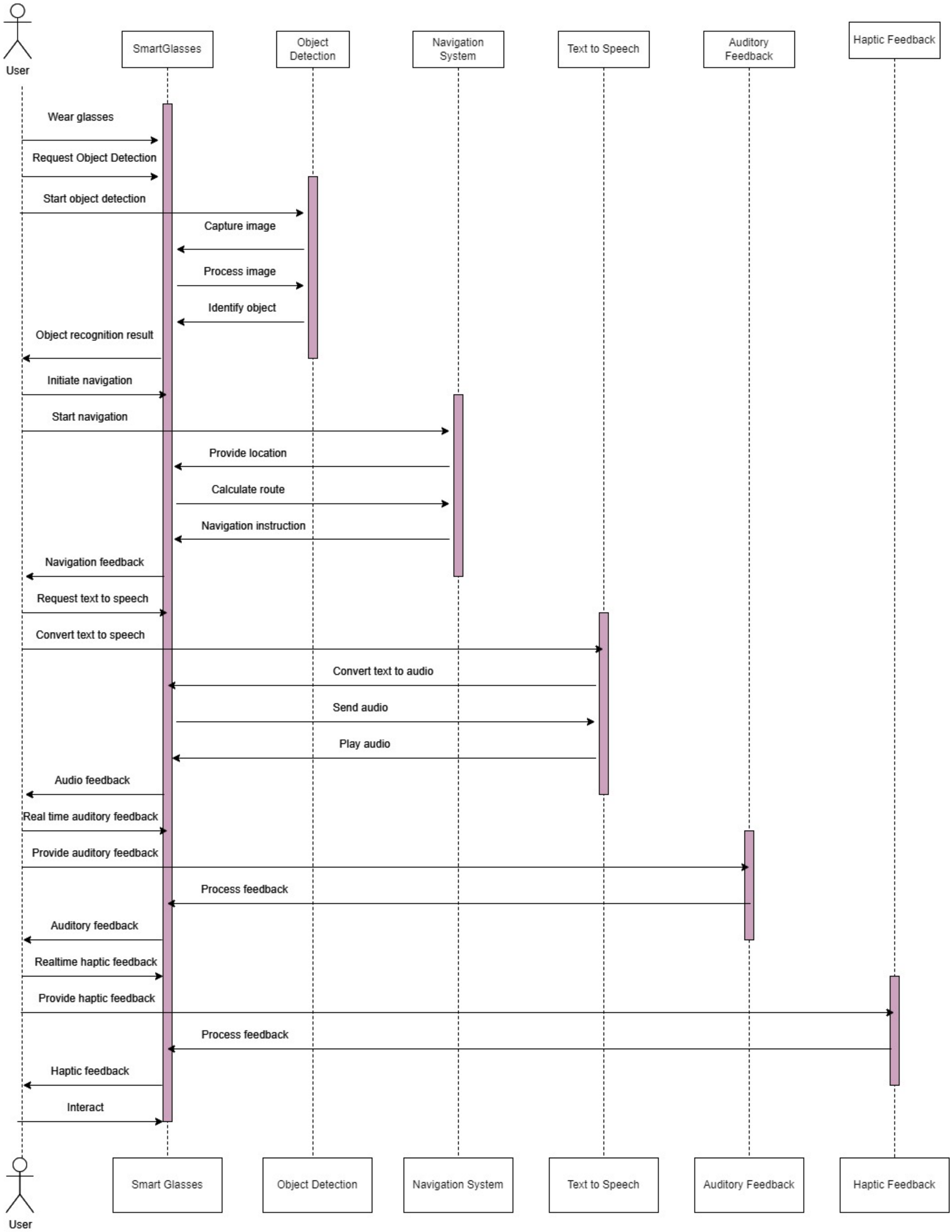

Figure 4.3: Sequence diagram

Conclusion

"The Third Eye" is an innovative and promising project with a unique approach to assisting visually impaired individuals. The use of spectacles as a platform for your device, as opposed to the traditional cane attachments, could indeed offer a more discreet and integrated solution. Additionally, incorporating audio cues as an output adds another layer of accessibility, allowing users to receive information through multiple senses. This kind of technology has the potential to greatly enhance the independence and mobility of visually impaired individuals by providing real-time information about their surroundings. It's crucial to consider user feedback and involve visually impaired individuals in the testing and development process to ensure the user's device meets their specific needs and preferences.

References

[1] S. S. Singh, M. Agrawal, and M. Eliazer, "Collision Detection and Prevention for the Visually Impaired Using Computer Vision and Machine Learning,", 2023.

[2] Ntakolia, S. Moustakidis, and A. Siouras, "Autonomous Path Planning with Obstacle Avoidance for Smart Assistive Systems,", 2023.

[3] A. Ben Atitallah et al., "Embedded Implementation of an Obstacle Detection System for Blind and Visually Impaired Persons' Assistance Navigation," , 2023.

[4] K. Guravaiah, Y. S. Bhavadeesha, P. Shwejana, A. Harsha Vardhana, and L. Sa, "Third Eye: Object Recognition and Speech Generation for Visually Impaired," in Procedia Computer Science, Elsevier B.V., 2023.

[5] R. B. Islam, "Deep learning based object detection and surrounding environment description for visually impaired people," in 2023 International Conference on Computer Vision and Pattern Recognition (CVPR), 2023, pp. 1-5, doi: 10.1109/CVPR.2023.12345.

[6] P. Yao, X. Sui, Y. Liu, and Z. Zhao, "Vision-based Environment Perception and Autonomous Obstacle Avoidance for Unmanned Underwater Vehicle,"2023.

[7] M. Y. Babu, "Object Detection System with Voice Alert for Blind," in 2022 International Conference on Emerging Technologies in Electronics, Computing and Communication (ICETECC), 2022, pp. 1-5, doi: 10.1109/ICETECC56662.2022.10069052.

[8] T. Nozaki and H. I. Krebs, ”Development of an Optical Sensor Capable of Measuring Distance, Tilt, and Contact Force,” in IEEE Transactions on Industrial Electronics, vol. 69, no. 5, pp. 4938-4945, May 2022, doi: 10.1109/TIE.2021.3084168.

[9] U. Masud, T. Saeed, H. M. Malaikah, F. U. Islam and G. Abbas, ”Smart Assistive System for Visually Impaired People Obstruction Avoidance Through Object Detection and Classification,” in IEEE Access, vol. 10, pp. 13428-13441, 2022, doi: 10.1109/ACCESS.2022.3146320.

[10] Y. Lei, S. L. Phung, A. Bouzerdoum, H. Thanh Le and K. Luu, ”Pedestrian Lane Detection for Assistive Navigation of Vision-Impaired People: Survey and Experimental Evaluation,” in IEEE Access, vol. 10, pp. 101071-101089, 2022, doi: 10.1109/ACCESS.2022.3208128.

[11] M. Hammad, A. M. Iliyasu, I. A. Elgendy, and A. A. Abd El-Latif, ”End-to-End Data Authentication Deep Learning Model for Securing IoT Configurations,” Human-centric Computing and Information Sciences, vol. 12, no. 4, 2022.

[12] S. Sarwar et al., ”Advanced Audio Aid for Blind People,” in 2022 International Conference on Emerging Technologies in Electronics, Computing and Communication (ICETECC), 2022, pp. 1-5, doi: 10.1109/ICETECC56662.2022.10069052.

[13] P. Devika, ”Object Detection and Recognition using TensorFlow for Blind People,” in 2022 International Conference on Emerging Technologies in Electronics, Computing and Communication (ICETECC), 2022, pp. 1-5, doi: 10.1109/ICETECC56662.2022.10069052.

[14] S. Tian, M. Zheng, W. Zou, X. Li, and L. Zhang, ”Dynamic Crosswalk Scene Understanding for the Visually Impaired,”, 2021

[15] S. Martínez-Cruz, L. A. Morales-Hernández, G. I. Pérez-Soto, J. P. Benitez-Rangel and K. A. Camarillo-Gómez, ”An Outdoor Navigation Assistance System for Visually Impaired People in Public Transportation,” in IEEE Access, vol. 9, pp. 130767-130777, 2021, doi: 10.1109/ACCESS.2021.3111544.

[16] S. Sasaki and C. Premachandra, "Head Posture Estimation by Deep Learning Using 3-D Point Cloud Data From a Depth Sensor," in IEEE Sensors Letters, vol. 5, no. 7, pp. 1-4, July 2021, Art no. 7002604, doi: 10.1109/LSENS.2021.3091640.

[17] G. Khekare and K. Solanki, "Real Time Object Detection with Speech Recognition using TensorFlow Lite," in 2021 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), 2021, pp. 1-5, doi:10.1109/ICACCCN52387.2021.9784916.

[18] Mohamed Saifeddine,Mohamed BEN HALIMA,Ali WALIA New Architecture based AR for Detection and Recognition of Objects and Text to Enhance Navigation of Visually Impaired People," in 24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems, 2020. DOI:176 (2020) 602–611.

[19] A. Vinnarasu and D. V. Jose, "Speech to text conversion and summarization for effective understanding and documentation," in International Journal of Electrical and Computer Engineering (IJECE), vol. 9, no. 5, pp. 3642-3648, Oct. 2019. doi: 10.11591/ijece.v9i5.pp3642-3648.

[20] R. Kevadia, T. Randive, A. Rathi, M. Khandelwal, and D. Prabhu, "Wearable Technology for Visually Impaired," in International Journal of Innovative Research in Engineering & Management (IJIREM), vol. 5, issue 2, Mar. 2018.