ISSN: 2583-9888
DOI: doi.org/10.51219/JAIMLD/yijia/119

URF PUBLISHERS connect with research world

Journal of Artificial Intelligence, Machine Learning and Data Science
https://urfpublishers.com/journal/artificial-intelligence

Vol: 1 & Iss: 1

Research Article

Advancing Accessibility: The Transformative Role of Technology in Empowering Individuals with Visual Impairment

Amit Gupta[1]* and Yijia Zhao[2]

[1]Staff Engineer, San Jose, CA, USA

[2]Senior Member of Technical Staff, Palo Alto, CA, USA

Citation: Gupta A, Zhao Y. Advancing Accessibility: The Transformative Role of Technology in Empowering Individuals with Visual Impairment. *J Artif Intell Mach Learn & Data Sci 2020*, 1(1), 431-430. DOI: doi.org/10.51219/JAIMLD/yijia/119

Received: 01 October, 2020; **Accepted:** 18 October, 2020; **Published:** 20 October, 2020

*__Corresponding author:__ Amit Gupta, Staff Engineer, San Jose, CA, USA, E-mail: gupta25@gmail.com

ABSTRACT

Visual impairment presents formidable obstacles to mobility and autonomy for individuals with blindness or low vision. In this research paper, we delve into the innovative integration of computer vision technology to develop not only a mobile application but also an augmented reality (AR) and virtual reality (VR) headset application, with the goal of aiding individuals with visual impairments in navigating their surroundings with safety and independence. Our approach capitalizes on established object detection algorithms, notably YOLO v3, as we meticulously curated a bespoke dataset and conducted model training to enable accurate recognition of commonplace objects in various environments.

The mobile application, bolstered by custom algorithms and frameworks such as TensorFlow, Vision framework, ARKit, and ARCore, operates in real-time to identify objects and furnish users with auditory feedback detailing the distance from their location. This experimental endeavor serves as a testament to the transformative potential of technology, illustrating its capacity to augment accessibility and embolden individuals with visual impairments to traverse their surroundings with assuredness.

Furthermore, we delve into the expansive realm of possibilities where analogous technological innovations can be harnessed to accommodate the diverse needs of individuals grappling with an array of physical disabilities. By embracing such technologies, we aspire to foster inclusivity and foster independence across an extensive spectrum of scenarios, thereby empowering individuals with disabilities to partake fully in the fabric of society.

Keywords: Accessibility, Visual Impairment, Computer Vision, Object Detection, Mobile Application, Assistive Technology, Auditory Feedback, Navigation Assistance, Inclusivity, Physical Disabilities, Blind, Blindness, Artificial Intelligence, Machine Learning

1. Introduction

Visual impairment poses significant challenges to mobility, independence, and quality of life for a substantial portion of the global population. According to the National Library of Medicine and Global Blindness and Visual Impairment Data from 2015[1], an estimated 253 million people worldwide experience some form of visual impairment, with 36 million (14%) being completely blind and 217 million (86%) suffering from moderate to severe visual impairment. Notably, 89% of visually impaired individuals reside in low- and middle-income countries. Traditional mobility aids, such as canes and guide dogs, while offering assistance, may have limitations in providing real-time feedback about the surrounding environment.

This research endeavors to address these challenges by leveraging technological solutions to empower individuals with visual impairments to navigate their surroundings more effectively. Specifically, we investigate the development of a mobile application that employs computer vision algorithms to identify objects in the environment and provide auditory feedback on their distance from the user. Furthermore, we explore how similar technological solutions can be adapted to cater to the diverse needs of individuals with various physical disabilities, promoting inclusivity and independence across a range of scenarios.

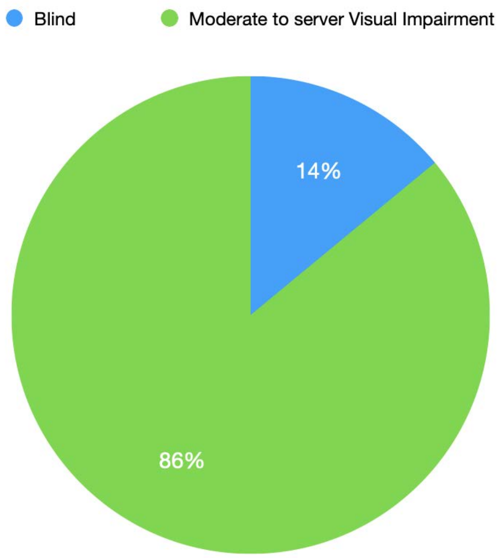

Figure 1: Blindness distribution vs visual impairment.

2. Literature Survey and Related Work

With the increasing computational power and storage capacity of mobile devices, coupled with the expanding speeds and coverage of mobile internet connectivity, unique opportunities arise for leveraging smartphones as versatile assistive devices. These rapidly advancing mobile technologies present a compelling platform to develop innovative solutions that can enhance accessibility and independence for individuals with disabilities or impairments[2].

Moreover, the widespread adoption and affordability of smartphones position them as an ideal medium for delivering assistive technologies to a global user base across socioeconomic contexts. Developing assistive solutions on the smartphone platform could circumvent the need for specialized, cost-prohibitive hardware, thereby promoting inclusivity and democratizing access to assistive aids. As these mobile technologies continue to evolve at an unprecedented pace, research efforts should be directed toward exploring their transformative potential in the assistive technology domain, ultimately empowering individuals with disabilities to lead more independent and enriched lives.

Recent advancements in computer vision and object detection algorithms[3] have paved the way for the development of camera-based assistive devices and systems, such as those leveraging mobile phones, augmented reality (AR), and virtual reality (VR) headsets[4] . These technological breakthroughs hold significant promise for creating innovative assistive solutions that can enhance accessibility and independence for individuals with visual impairments or other disabilities. By harnessing the capabilities of object detection models and computer vision techniques, these devices can analyze and interpret the user's surroundings in real-time, identifying obstacles, objects of interest, or navigational cues, and subsequently provide auditory, visual, or haptic feedback to the user. The capacity to accurately discern the objects present in their immediate environment, coupled with supplementary contextual information, proves invaluable for individuals grappling with blindness. (Brady, Morris, Zhong, White, & Bigham, 2013)[5].

In their study referenced as[6], the authors introduced a novel system prototype named Sharojan Bridge, designed to enhance communication efficacy among individuals with visual impairments. This wearable device is constructed with Internet of Things (IoT) components, aiming to facilitate seamless interaction for the blind community. Similarly, a navigation assistance system tailored for individuals with visual impairments, denoted as POSE, was meticulously designed and developed. This system incorporates a foot-mounted pedometer and a sensing package affixed to a white cane. The researchers devised an algorithm to construct maps for updating the individual's location and leveraged orthogonal elements in the building layout to serve as a 3D structural compass. Rigorous analysis and real-time evaluation of the system were conducted within authentic environmental settings[7].

Besides these literatures, there are multiple digital applications available on most popular Android and iOS platforms for accessibility.

In reference[8], a blind assistive and tracking embedded system was introduced, designed to aid visually impaired individuals in navigation while also offering the capability to track their movements and ensure security in case of loss. This system was realized through an Android application integrated with the Global Positioning System (GPS).

Envision AI[9]: Envision empowers people who are blind or have low vision to access everyday visual information for themselves and speak out about surroundings and objects.

TapTapSee[10]: TapTapSee stands as a mobile camera application meticulously crafted to cater to the needs of blind and visually impaired individuals, leveraging the CloudSight Image Recognition API. This innovative app harnesses the capabilities of the device's camera and VoiceOver functions to capture images of objects and vocalize their identification for the user. Within TapTapSee, users simply execute a double-tap on the device's screen to capture images of both two and three-dimensional objects from any angle, subsequently undergoing swift and precise analysis for identification. The device's VoiceOver functionality then articulates the identification audibly to the user, ensuring seamless accessibility and usability. (iOS and Android)

Lazarillo[11]: Lazarillo allows people with disabilities to explore their environments safely and efficiently with audible messages about their surroundings. As you walk, Lazarillo will announce places of interest around you, from streets and intersections to restaurants, shops, and transit areas.

Be my Eye[12]: Be My Eyes connects blind and low-vision users who want sighted assistance with volunteers and companies anywhere in the world, through live video

Lookout[13]: Built by Google, Lookout uses computer vision to assist people with low vision or blindness to get things done faster and more easily. Using a phone's camera, Lookout makes it easier to get more information about the world around the person and do daily tasks more efficiently like sorting mail, putting away groceries, and more.

3. Proposed Solution

Our research objective was to develop a prototype application showcasing the potential of modern technologies in assisting individuals with vision impairments to navigate and maintain awareness of their surroundings, even in the absence of a network connectivity. For this purpose, we selected the iPhone as our platform of choice and employed a combination of cutting-edge tools and techniques, including the YOLO object detection model, iOS vision framework, ARKit, and bespoke mathematical algorithms for proximity detection.

The workflow of our application is designed to cater specifically to individuals with low vision. Upon launching the application via voice command, tailored for mobile devices, users have the option to position the device conveniently, such as in a shirt pocket or attached to a lanyard with the back camera facing outward, allowing for hands-free operation. Subsequently, as users commence walking, the application leverages the iPhone's camera and integrated object detection machine learning model to identify objects within the camera frame and ascertain their distance from the device.

Initially, the application employs the YOLO 3 model, trained to recognize a diverse array of 80 different objects, encompassing items ranging from chairs to dining tables and people. Subsequent to object identification, the application proceeds to execute intricate mathematical calculations to determine the nearest object relative to the user's position. Once this proximity assessment is completed, the application communicates audibly, utilizing the device's speakers or potentially via integration with smart glasses, relaying to the user both the identity of the object detected and its corresponding distance.

This comprehensive workflow embodies our endeavor to harness advanced technologies to create a user-friendly and accessible solution for individuals with vision impairments. By seamlessly integrating object detection, distance computation, and auditory feedback functionalities, our prototype application aims to enhance the mobility and independence of users, enabling them to navigate and perceive their environment with heightened confidence and efficiency.

1. Our initial step is to employ an existing object detection framework such as YOLO v3 for its manifold advantages. Firstly, being open source, it facilitates accessibility and collaboration, aligning with the ethos of community-driven development. Moreover, YOLO v3 boasts remarkable speed, making it suitable for real-time applications necessitating swift object identification. Additionally, it supports a diverse array of various objects, rendering it highly versatile for various use cases. Notably, YOLO v3 furnishes essential functionalities such as bounding box prediction, class prediction, and prediction across different scales. Its underlying feature extractor, Darknet-53, further enhances its efficacy, ensuring qualitative results in object detection tasks.

2. Subsequently, our focus shifts to dataset creation and annotation, pivotal stages in the development of robust object detection models. To this end, we utilize tools like MakeML.app or IBM Cloud annotation, facilitating the systematic identification and annotation of new objects. This meticulous dataset creation process lays the groundwork for training accurate and reliable object detection models, crucial for subsequent stages of our prototype development.

3. Following dataset creation, we embark on the training phase using Darkflow, a Darknet translation compatible with TensorFlow. This stage involves the utilization of our meticulously curated dataset to train the object detection model, refining its ability to accurately identify and classify objects in diverse environments. Once the training process is completed, the resultant weights are converted into .ml models (iOS specific format) or TensorFlow Lite models, optimizing them for deployment on mobile devices.

4. The subsequent step entails integrating the trained object detection model into a mobile application, exemplified here with iOS as the target platform. This integration process involves the drag-and-drop placement of the model into the application, followed by recompilation to ensure seamless compatibility and functionality. Notably, APIs exist to facilitate dynamic recompilation of the model at runtime, enabling updates and enhancements without requiring manual intervention. This iterative process ensures that the deployed application remains adaptive and responsive to evolving user needs and technological advancements.

5. Finally, we implement the logic for object detection and distance calculation using custom algorithms and frameworks such as the Vision framework, ARKit, and Camera APIs. This entails considering various factors such as the height, area, and distance of objects from the camera to derive accurate distance estimates. The culmination of this process enables the application to provide real-time auditory feedback to users, conveying essential information about the nearest objects and their respective distances. This user-centric approach enhances accessibility and usability, empowering individuals with visual impairments to navigate their surroundings confidently and independently. Upon completion of the calculation, the application announces pertinent information to the user, providing contextual awareness and facilitating informed decision-making in navigation scenarios.

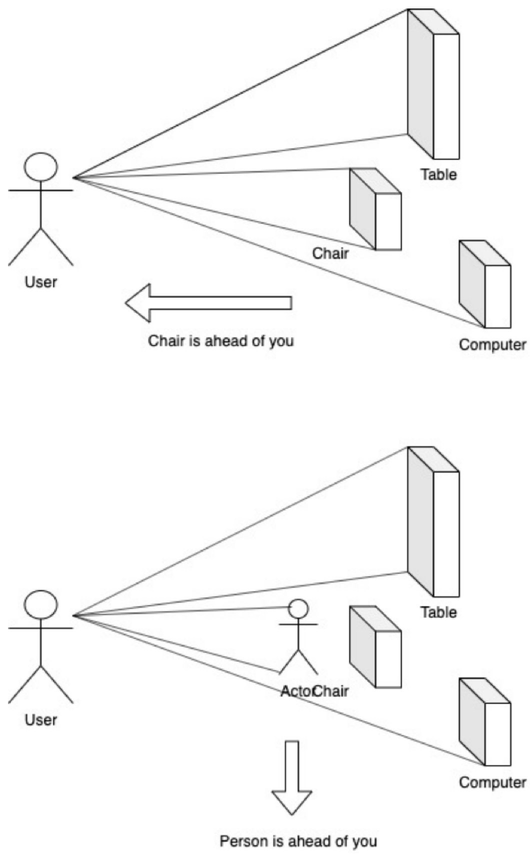

Figure 2: How prototype identify objects.

4. Technical Details

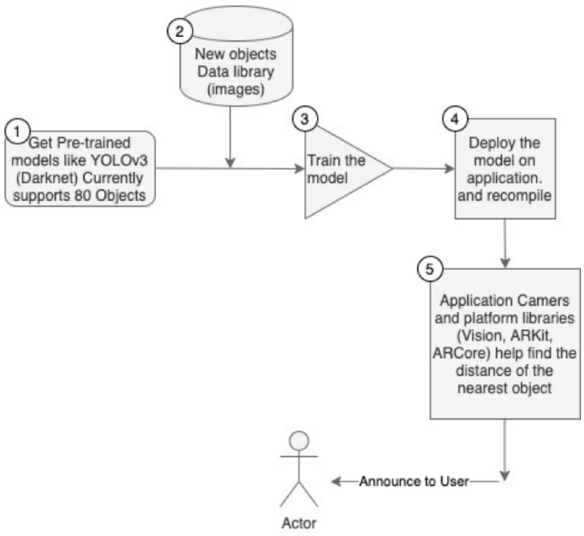

Figure 3: Technical flow of Solution.

5. Result

In our study, we conducted a demonstration of our prototype application on the iPhone, involving a group of participants, including individuals with visual impairments. Participants held the phone camera facing outward while navigating through a safe and controlled environment containing various objects such as tables, chairs, backpacks, and computers. Our prototype application successfully predicted and announced the objects along with their distances to the users with a decent level of accuracy.

During the demonstration, participants were able to utilize the application effectively, receiving real-time auditory feedback about their surroundings. The accuracy of object prediction and distance estimation garnered positive feedback from the participants, highlighting the potential utility of the application in enhancing navigation and spatial awareness for individuals with visual impairments.

Furthermore, we gathered valuable feedback from the participants regarding the adoption of compliances and suggestions for future enhancements. Participants expressed interest in the incorporation of additional objects to support both indoor and outdoor settings, emphasizing the importance of expanding the application's functionality to cater to diverse environments and user needs.

Overall, the results of our demonstration showcase the promising capabilities of our prototype application in assisting individuals with visual impairments in navigating their surroundings with accuracy and confidence. The feedback obtained from participants underscores the importance of ongoing refinement and expansion of the application to ensure its effectiveness and usability in various real-world scenarios.

6. Future Direction

1. **Advancement in Smartglasses, AR and VR Headsets**: The future of assistive technology for individuals with visual impairments lies in the advancement of augmented reality (AR) and virtual reality (VR) headsets. As these technologies continue to evolve, they have the potential to make blindness irrelevant by providing immersive environments and real-time guidance. AR headsets can overlay digital information onto the user's physical surroundings, offering contextual awareness and navigation assistance. Similarly, VR headsets can simulate realistic environments and scenarios, allowing users to practice navigation and orientation skills in a safe and controlled manner. Future research should focus on optimizing AR and VR technologies for individuals with visual impairments, ensuring seamless integration into everyday life.

2. **Impact of Compliance on Technology Adaptation:** The successful implementation of technology-assisted guidance for blind individuals is contingent upon various factors, including regulatory compliance and accessibility standards. Future research should explore how compliance requirements, such as the Americans with Disabilities Act (ADA) in the United States or similar regulations globally, influence the adoption and adaptation of assistive technologies. Understanding the regulatory landscape and addressing compliance challenges will be essential for ensuring equitable access to technology for individuals with visual impairments.

3. **Over-the-Air Delivery of ML Models:** As machine learning (ML) models play a crucial role in object recognition and distance estimation for assistive guidance applications, delivering updates to these models efficiently is imperative. Future research should focus on developing methods for delivering ML models over-the-air (OTA)[14], allowing for seamless updates and improvements without requiring manual intervention. OTA delivery mechanisms will enable developers to continually refine and enhance ML models based on real-world usage and feedback from users, ensuring optimal performance and accuracy over time.

4. **Expansion of Object Recognition Capabilities:** While existing assistive guidance applications can recognize common objects such as chairs and doors, further research is needed to expand their object recognition capabilities. Future studies should focus on training ML models to identify new and common objects encountered in everyday environments, such as traffic lights, road signs, and pedestrian crossings. By enhancing the object recognition capabilities of assistive guidance systems, individuals with visual impairments can navigate complex urban environments more confidently and independently.

The future of assistive technology for individuals with visual impairments holds immense promise, with advancements in AR and VR headsets, compliance considerations, OTA delivery of ML models, and expansion of object recognition capabilities paving the way for greater inclusivity and independence. Continued research and collaboration in these areas will be essential for realizing the full potential of technology in improving the lives of individuals with visual impairments.

7. Conclusion

In conclusion, our exploration of developing a prototype application for assisting individuals with visual impairments underscores the transformative potential of technology in enhancing accessibility and empowerment for individuals with physical disabilities. The utilization of existing object detection frameworks coupled with dataset creation and annotation processes, lays the foundation for training accurate and reliable object detection models. Through the integration of trained models into mobile applications and the implementation of custom algorithms for object detection and distance calculation, we have demonstrated the feasibility of providing real-time auditory feedback to users, facilitating informed navigation and enhancing spatial awareness.

This research journey exemplifies how technology can serve as a powerful catalyst for advancing inclusivity and independence for individuals with physical disabilities. Navigation assistance, communication aids, assistive learning tools, and health monitoring devices are just a few examples of the diverse array of interventions made possible through technology. By leveraging the transformative power of technology, we have the opportunity to create a more inclusive society where individuals with physical disabilities can thrive and participate fully in all aspects of life.

However, realizing the full potential of technology requires ongoing innovation, collaboration, and a commitment to inclusive design principles. As technology continues to evolve, it is essential to prioritize accessibility and ensure that technological solutions are designed with the needs of individuals with physical disabilities in mind. By fostering a culture of inclusivity and accessibility in technology development, we can break down barriers and create opportunities for individuals with physical disabilities to lead more independent and fulfilling lives.

Our research underscores the importance of leveraging technology to promote inclusivity and empowerment for individuals with physical disabilities. Through ongoing innovation and collaboration, we have the opportunity to harness the transformative power of technology to create a more inclusive society where all individuals, regardless of physical ability, can thrive and participate fully in all aspects of life.

8. References

1. Ackland P, Resnikoff S, Bourne R. World blindness and visual impairment: despite many successes, the problem is growing. Community Eye Health 2017;30: 71-73.
2. Mandl T, Punchoojit L, Hongwarittorrn N. Usability Studies on Mobile User Interface Design Patterns: A systematic literature review. Advances in Human-Computer Interaction 2017: 22.
3. Zhao Z-Q, Zheng P, Xu S-T, Wu X. Object detection with deep learning: A Review. IEEE Transactions on Neural Networks and Learning Systems 2019;30: 3212-3232.
4. Dharmale RD, Ingole PV. Text detection and recognition with speech output in mobile application for assistance to visually challenged person. IJSER 2015;6: 441-447.
5. Brady E, Morris MR, Zhong Y, Samuel W, Jeffrey B. Visual challenges in the everyday lives of blind people. CHI, 2013.
6. Rastogi R, Mittal S, Agarwal S, A novel approach for communication among Blind, Deaf and Dumb people. 2015 2nd International Conference on Computing for Sustainable Global Development 2015; 605-610.
7. Hesch JA, Roumeliotis SI. Design and Analysis of a Portable Indoor Localization Aid for the Visually Impaired. IJRR 2010;29: 1400-1415.
8. Tanveer SR, Hashem MMA, Hossain K. Android Assistant EyeMate for Blind and Blind Tracker. IEEE 2015 18th International Conference on Computer and Information Technology 2015.
9. Envision AI. Perceive Possibility. Envision.
10. TapTapSee
11. https://lazarillo.app/
12. https://www.bemyeyes.com/
13. 13. Look out
14. 14.https://medium.com/@miriamfriedel/delivering-ml-to-mobile-apps-with-skafos-and-aws-sagemaker-d565c884973e