EAI Endorsed Transactions
on Internet of Things
Research Article EAI.EU

Reimagining Accessibility: Leveraging Deep Learning in Smartphone Applications to Assist Visually Impaired People Indoor Object Distance Estimation

Talal Saleem[1,*] and V. Sivakumar[2]

[1,2]Faculty of Computing and Engineering Technology, Asia Pacific University of Technology and Innovation (APU), Kuala Lumpur, Malaysia

Abstract

Every other aspect of life is organized around the sight. A person with vision impairment suffers severely from independent mobility and quality of life. The proposed framework combines mobile deep learning with distance estimation algorithms to detect and classify indoor objects with estimated distance in real-time during indoor navigation. The user, wearing the device with a lanyard or holding it in a way that positions the camera forward, identifies in real-time surroundings indoor objects with estimated distance and voice commentary. Moreover, the mobile framework provides an estimated distance to the obstacles and suggests a safe navigational path through voice-guided feedback. By harnessing the power of deep learning in a mobile setting, this framework aims to improve the independence of visually impaired individuals by facilitating them a higher degree of independence in indoor navigation. This study's proposed mobile object detection and distance estimation framework achieved 99.75% accuracy. This research contributes by leveraging mobile deep learning with identifying objects in real-time, classification and distance estimation, a state-of-the-art approach to use the latest technologies to enhance indoor mobility challenges faced by visually impaired people.

Keywords: Artificial Intelligence, Distance Estimation, Mobile Deep Learning, Object Detection, Visual Impairment People, Indoor Navigation

Received on 30 December 2023, accepted on 26 June 2024, published on 03 July 2024

doi: 10.4108/eetiot.6501

1. Introduction

Deep learning is one of the latest technologies that is witnessing its golden era and is slowly becoming a leader in the domain of machine learning. Today we see a high rise in mobile usage. There are over 3.3 billion mobile phone users globally, and there is an estimate that in developing countries [1]. Artificial intelligence makes mobile more intelligent by making it learn and act automatically. Deep learning has acquired huge success in building intelligent applications. There are a large number of applications that are developed for the general end-users. However, in comparison, there is less work has been done for visually challenged individuals to have mobile on-device intelligent assistant for their carrying out their everyday life activities. Applications for vision-impaired or blind people have been almost ignored, particularly smartphone-based. In 2019, the World Health Organization estimated that 2.2 billion people worldwide have vision impairment or blindness [2]. Every other aspect of life is organized around the sight. It is difficult for visually challenged people to navigate new and unfamiliar places. Vision-impaired people accident occurs while walking when a collision happens with a person or a wall.

Running deep learning on mobile is challenging because of mobile device limitations of computational power, memory size, and battery power. It is challenging to run deep learning on-device locally on mobile to detect an object in real-time using a built-in camera with efficient performance and

*Corresponding author. Email: TP053459@mail.apu.edu.my

accuracy. This research study proposes a state-of-the-art approach smartphone-based on-device deep learning framework to assist visually challenged individuals with indoor navigation. The visually challenged people could wear it using a lanyard or holding a mobile while positioning the camera forward during indoor navigation to recognize real-time surrounding situations and notify objects with efficient performance and accuracy by taking the distance involved too into consideration. The proposed framework uses mobile-based deep learning techniques to safely detect while navigating and recognize the object of interest in real time. This framework could enable the voice command and describe detected classified object recognition as real-time speech commentary, thereby assisting the visually impaired efficiently in their indoor navigation. The proposed mobile deep learning bound approach alerts obstacle warnings with an estimated distance to the obstacle. Further, for the safe mobility of the visually challenged, this research study could guide safe navigation through voice-guided directions.

2. Related Work

Several research works have been proposed to assist visually impaired individuals.

Kay [3] proposed an audio device to assist visually impaired people in finding the object. The author experimented by taking a picture of an object using a webcam and describing it via an audio device. The proposed concept had limitations and was not intelligent to locate the object direction or while taking picture brightness of the object was required.

Durette et al. [4] proposed a concept to assist visually impaired people using computer vision techniques. A camera was attached forehead of the person, and the image captured from the video was translated via voice feedback sent to the headphone. The concept had limitations visually impaired person has to carry a heavy laptop in a backpack for processing.

Martinez et al. [5] proposed a concept work a computer webcam captures video image, and headphones translate the object into voice to assist visually challenged individuals. The proposed concept was not portable just presented an object detection concept for visually impaired people.

Tepelea et al. [6] proposed a portable sensors-based concept to assist visually impaired people with indoor navigation. The author combines RaspBerryPi, Arduino platform, and sensors. The proposed concept uses GPS to get coordinates to guide users for mobility and a camera to detect objects using a neural network.

Jabnoun et al. [7] proposed a work based on the classic feature detection SIFT algorithm to identify objects, and the author detected objects using object key points to assist blind people in identifying the object.

Zereen et al. [8] used the Kinect device to assist blind people in object detection during indoor walking. The Work based on a Kinect device camera and built-in sensor technology to detect objects. The object detection feedback to assist the blind was not presented in the Work.

Patel et al. [9] used a smart stick to assist visually challenged individuals with indoor mobility. The author connected with a regular stick a webcam to capture an image, an infrared sensor, and a power bank. RaspBerryPi process camera captured images using machine learning support vector machine algorithm. Headphones are used to alert visually impaired people.

Árvai [10] proposed a framework to support visually impaired indoor navigation. The author's Work used a webpage connected with a backend database of building information. The visually impaired person uses the mobile phone to browse webpage for building navigation.

Lin et al. [11] author's proposed Work used an RDGB camera to capture images and a laptop running deep learning GoogleLeNet Inception V3 model to process and give feedback via Bluetooth earphones. The proposed Work was using deep neural network architecture running on a powerful CPU computer.

Ooi et al. [12] proposed work that used the YOLO v3 algorithm to detect objects. On the forehead of the visually challenged person camera was placed, and object detection feedback was sent via headphones as voice. The concept had limitations visually impaired person has to carry a heavy laptop in a backpack.

Mahmud et al. [13] presented a concept model for indoor robots to assist visually impaired people to find common objects. The author placed cameras in the robot via voice assistance robot helps assist visually impaired people in finding common indoor objects.

Strbac et al. [14] calculate object distance using YOLO V3 and two cameras. The author used the stereoscopy two cameras technique to detect the same target object using the left and right camera and measure distance.

Karthika et al. [15] used deep learning and monocular vision. The author trained YOLO with the KITT dataset and achieved 2% improved object distance estimation accuracy compared to similar methods.

Sakic et al. [16] researcher used the camera for object detection and lidar for distance estimation. The author used YOLO to detect an object and used the method to map lidar points cloud with YOLO object detection and measure distance. The success of the lidar sensor-based distance estimation method depends on the YOLO object detection and classification.

Rahman et al. [17] used ultrasonic sensors to detect obstacles placed in front of the sensor. The system sounds buzz if an obstacle is found 3 meters away in the distance. In the past several pieces of research have been done to identify and recognize objects but do not estimate real-time video object distance. The stated research gaps form the basis of this proposed research study.

3. Methodology

In this study, we propose a smartphone-based on-device deep learning framework to assist the visually impaired in recognizing the object of interest and guide the estimated distance to the object. Our proposed mobile framework enables voice commands to describe detected classified object recognition as real-time speech commentary.

3.1 Proposed Work

Lidar sensors work the same principle as radar, but instead of using radio waves, lidar uses laser beam. Lidar does not classify objects, lidar is an expensive technology, and there is a high cost in adopting lidar for object distance estimation. The ultrasonic sensor measures the range between the object and the sensor using sound waves. The ultrasonic sensor calculates the elapsed time the sound wave takes to reflect from object to sensor to drive distance. Whereas ultrasonic sensors are low in cost. For the two approaches, ultrasonic and lidar sensors, there is no mobile that comes with an ultrasonic sensor. iPhone 12 pro and later only mobile model comes with a lidar sensor preinstalled. The sensor-based mobile object distance estimation approach depends on an angle at which the mobile sensor is held with minimum hand movement.

In computer vision, camera-based object distance estimation is achieved using stereo vision and monocular vision. The stereo vision-based object detection and distance estimation method use two cameras, and both cameras capture the same target object image. The object distance is calculated using the triangulation method. In computer vision, the distance estimation method uses monocular vision based on a single camera. The monocular vision-based range estimation method has less run-time due to single camera image processing is required in comparison to stereo vision, where two camera feed is required to estimate object distance. The stereo vision-based object distance estimation approach depends on two camera feeds, both cameras must be at the exact calibration alignment to estimate object distance, slight difference in camera angle calibration results in failure in object distance estimation. There is no assurance that a visually impaired or blind person will hold the phone steadily to estimate object distance. Based on this, our proposed mobile framework adopted a monocular vision approach for mobile deep learning-based object distance estimation. The monocular vision method depends on a single-camera feed, and smartphones come with a built-in at least one rare camera.

To calculate the distance from the mobile camera to the object to use the triangle similarity approach. To drive a distance to an object in an image using the method, the width of object W and D distance from a mobile camera, the width of the image in pixels P. To derive the length F of an object from the camera:

$$F = (D X P) / W$$

The equation helps us to calculate the distance to an object in real-time using a mobile camera. The width of the object W is driven from the pixel width of the object detection boundary box.

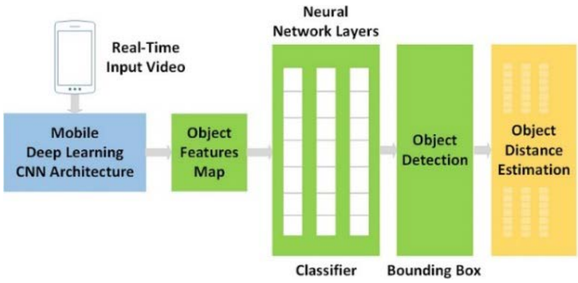

Figure 1. Mobile deep learning-based object distance estimation framework

Our proposed algorithm processes mobile real-time video and the distance to the object is estimated. Our proposed algorithm combines mobile deep learning with a monocular vision for object classification and object distance estimation, as illustrated in Figure 1. The deep learning module takes camera video input, extracts features, identifies the object, and finally, object coordinates are matched, and the distance to the object is estimated.

4. Implementation

4.1 Dataset

The proposed framework consists of two datasets. COCO dataset and our custom dataset. The COCO dataset contains more than 124,000 labelled largescale common objects. The COCO database contains 80 different classes with 83,000 training images and 41,000 Testing images. To prepare our custom dataset we used a short video of objects to transform video frames into images. The second method (no copyright) images taken from the internet. To improve accuracy, we have used a technique to crop the image and enlarge and focus the image object. We manually labelled each image using the annotation tool LabelImg. The custom dataset is split randomly, 80% selected data to train the model and the rest 20% data to test the model.

4.2 Target Objects

Many indoor objects visually challenged people would like to reach. In this study, we selected the following indoor objects along with distance estimation Couch, chair, refrigerator, and bed. The custom dataset contains indoors three object categories walking stick, door, and wardrobe.

4.3 Algorithm

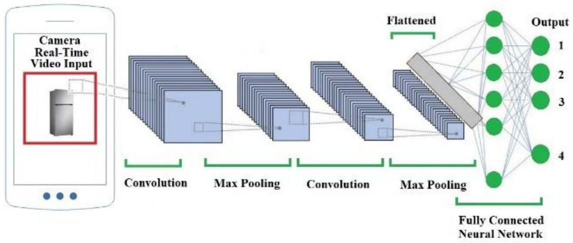

Figure 2. Mobile deep learning-based object detection

Our proposed framework is based on SSD MobileNet V2, which is one of the advanced real-time object detection and classification algorithms based on lightweight CNN architecture with efficient mobile accuracy with less computation cost. Figure 2 illustrates the algorithm model.

4.4 Custom Model Training

The custom dataset contains 460 images of each class walking stick, door and wardrobe. Total custom dataset images were 1,380 split randomly using Python scripts, with 80% of the total 1,080 training images and 20% testing 270 test images. The custom model was trained on the cloud colab host SSD_MobileNet_v2 model was retrained using TensorFlow object detection API and custom dataset, and RELU_6 activation function. Figure 3 shows 50,000 steps classification loss dropped at 0.001.

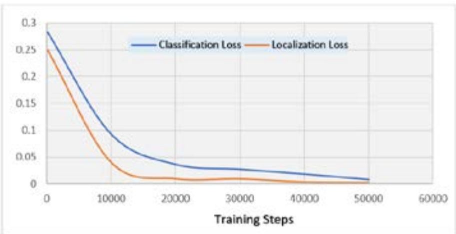

Figure 3. Custom model training at 50,000 steps

To use the TensorFlow-trained deep learning model on the mobile, first, we freeze the TensorFlow graph. TensorFlow frozen graph used in converting TensorFlow-Lite SSD custom model. To use a custom object detection trained model on mobile, the TensorFlow frozen graph has been exported as TensorFlow-Lite frozen graph. TensorFlow-Lite frozen graph is converted to TensorFlow-lite SSD custom mobile model. Figure 4 shows trained model sizes.

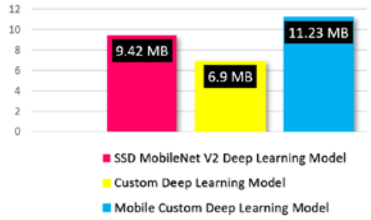

Figure 4. Trained model sizes

To deploy a custom object detection model on mobile, the SSD TensorFlow-lite model is saved with metadata and includes a custom object class map.

Custom Mobile deep learning model was deployed in mobile. Our proposed framework combines mobile deep learning with a monocular vision for object classification and object distance estimation as shown in Figure 5.

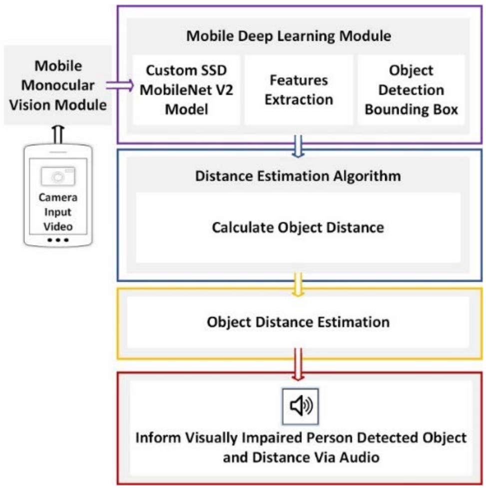

Figure 5. Mobile object detection and distance estimation.

Mobile application helps visually impaired people to know everyday indoor objects around them with the estimated distance that generate obstacles in their indoor path. The proposed framework includes the following functionalities.

4.5 Voice Commentary

The speech module guides visually impaired or blind people for safe in-door navigation, including distance to object along with object location directions to reach the object. In mobile navigation, the camera detects an obstacle in real-time, classifies and identifies objects class, and describes via voice visually impaired person object type and distance.

4.6 Object Distance Estimation

This module calculates the distance from the mobile camera to the object. The speech module pronounces voice feedback.

The distance calculation module keeps track of the range to the object once the distance is reached and notifies the end of the process. If the object is out of the detection reach of the mobile camera, voice alerts no object as an obstacle.

4.7 Recognition Objects in its Path

The module calculates the distance to the object, and text-to-voice API (Application Programming Interface) pronounces voice feedback. The distance calculation module keeps track of the range to the object once the distance is reached and notifies the end of the process. In case the object is out of the detection reach of the mobile camera, voice alerts no object.

4.8 Classify the objects and inform the visually impaired persons

To enable voice command to describe detected classified object recognition as real-time speech commentary. Our approach alerts obstacle warnings with an estimated distance to the obstacle. For safe mobility, this module guides safe navigation using voice-guided feedback.

5. Results & Discussion

We set up experiments for the mobile deep learning-based object detection framework locally on a mobile device to detect objects in real-time, classify with estimated distance and inform via voice feedback. Our experiment used a Samsung Galaxy A12 mobile device running the Android 12 operating system with 48 megapixels rare camera.

5.1 Object Detection and Distance Estimation (Model-1)

Experiment results show that at the distance of 100cm, 120cm, 150cm and 200cm in range, the mobile camera detected the object couch. Mobile deep learning-based android application for visually challenged people, object couch placed at the actual distance of 100 cm mobile application estimated distance 100 cm, object couch placed at the actual distance of 120 cm mobile application estimated distance 120 cm, object couch placed at the actual distance of 150 cm mobile application estimated distance 150 cm, similarly couch at the actual distance of 200 cm mobile application estimated distance as 201 cm. Figure 6 shows screenshots of mobile identifying object as couch, estimating object distance, warning alert to avoid obstacle and feedback to the user as a voice alert.

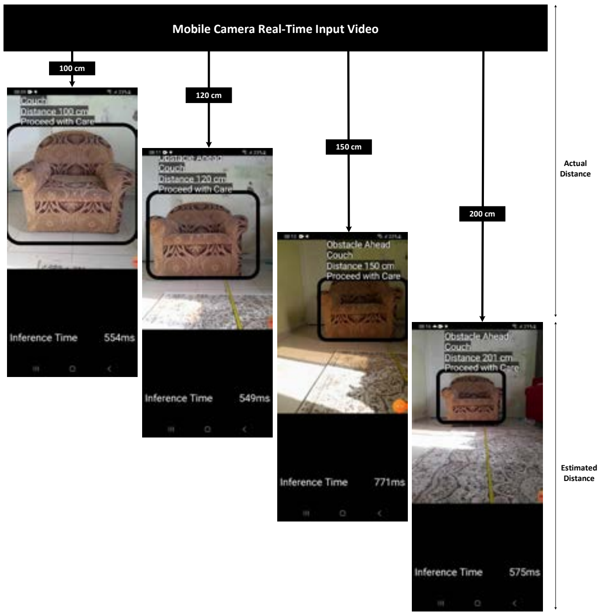

Figure 6. Couch Distance Estimation

Experiment results show that at the distance of 100cm, 120cm, 150cm and 200cm in range, the mobile camera detected the object chair. Mobile deep learning-based android application for visually challenged people object object chair placed at the actual distance of 100 cm mobile application estimated distance 100 cm, object chair placed at the actual distance of 120 cm mobile application estimated distance 120 cm, object chair placed at the actual distance of 150 cm mobile application estimated distance 152 cm, similarly chair at the actual distance of 200 cm mobile application estimated distance as 201 cm. Figure 7 shows screenshots of mobile application estimating object distance.

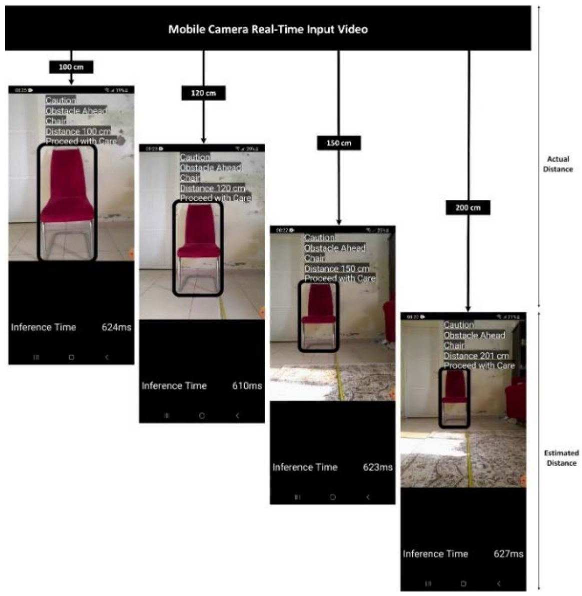

Figure 7. Chair Distance Estimation

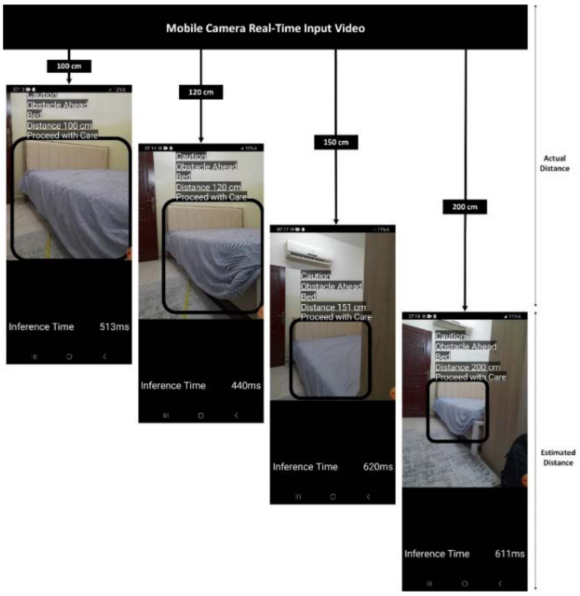

Figure 8. Bed Distance Estimation

Experiment results show that at the distance of 100cm, 120cm, 150cm and 200cm in range, the mobile camera detected the object bed. Mobile deep learning-based android application for visually challenged people object bed placed at the actual distance of 100 cm mobile application estimated distance 100 cm, object bed placed at the actual distance of 120 cm mobile application estimated distance 120 cm, object bed placed at the actual distance of 150 cm mobile application estimated distance 151 cm, similarly bed at the actual distance of 200 cm mobile application estimated distance as 200 cm. Figure 8 shows screenshots of mobile application estimating object distance.

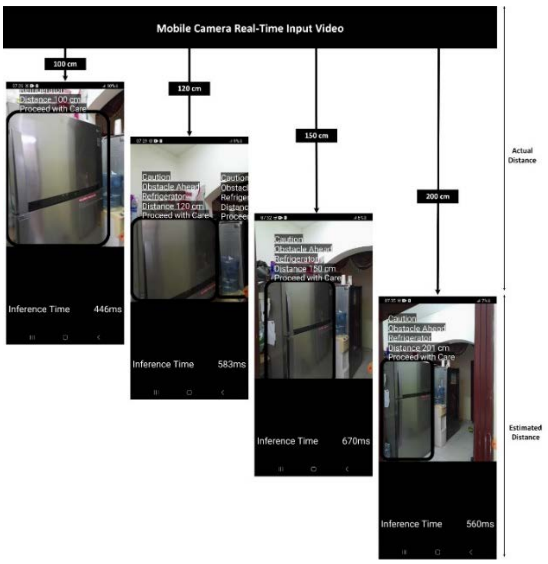

Figure 9. Refrigerator Distance Estimation

Experiment results show that at the distance of 100cm, 120cm, 150cm and 200cm in range, the mobile camera detected the object refrigerator. Mobile deep learning-based android application for visually challenged people object refrigerator placed at the actual distance of 100 cm mobile application estimated distance 100 cm, object refrigerator placed at the actual distance of 120 cm mobile application estimated distance 120 cm, object refrigerator placed at the actual distance of 150 cm mobile application estimated distance 150 cm, similarly refrigerator at the actual distance of 200 cm mobile application estimated distance as 201 cm. Figure 9 shows screenshots of mobile application estimating object distance.

Table 1. Model-1 Evaluation Results at 100 cm in Range

100 cm Range Object Distance Estimation Accuracy Evaluation Results				
	Couch	Chair	Bed	Refrigerator
Is Object Detected	Yes	Yes	Yes	Yes
Inference Time	554ms	624ms	513ms	446ms
Actual Distance	100 cm	100 cm	100 cm	100 cm
Estimated Distance	100 cm	100 cm	100 cm	100 cm
Distance Estimation Error (cm)	0 cm	0 cm	0 cm	0 cm
Obstacle or Not	Yes, An Obstacle	Yes, An Obstacle	Yes, An Obstacle	Yes, An Obstacle
Mobile Real-time Audio Output Notification	Caution Obstacle Ahead Couch Distance 100 cm Proceed with Care	Caution Obstacle Ahead Chair Distance 100 cm Proceed with Care	Caution Obstacle Ahead Bed Distance 100 cm Proceed with Care	Caution Obstacle Ahead Refrigerator Distance 100 cm Proceed with Care
Objects Accurately Detected	4/4			
Objects Not Detected	0/4			
Obstacles Accurately Identified	4/4			
Obstacles Not Identified	0/4			
Total Distance Estimation Accuracy	100 %			

We evaluate experiment-1 research results of large indoor objects detection for visually challenged people couch, chair, bed and refrigerator, a distance of 100 cm (refer to Table 1). At a distance of 100 cm, large indoor objects couch, chair, bed and refrigerator were classified and detected by drawing a bounding box around objects. According to experiment-1 research results at a distance of 100 cm, mobile applications could detect 4 out of 4 objects. At the distance of 100 cm mobile application calculated the distance to the couch with 100% accuracy. At the distance of 100 cm mobile application calculated the distance to the chair with 100% accuracy. At the distance of 100 cm mobile application calculated the distance to the bed with 100% accuracy. At the distance of 100 cm mobile application calculated the distance to the refrigerator with 100% accuracy. To evaluate the algorithm's performance, at a distance of 100 cm mobile application estimated the distance to the large objects with 100% accuracy.

At a distance of 120 cm, large indoor objects couch, chair, bed and refrigerator were classified and detected by drawing a bounding box around objects. According to experiment-1 research results at a distance of 120 cm, mobile applications could detect 4 out of 4 objects. At the distance of 120 cm mobile application calculated the distance to the couch with

100% accuracy. At the distance of 120 cm mobile application calculated the distance to the chair with 100% accuracy. At the distance of 120 cm mobile application calculated the distance to the bed with 100% accuracy. At the distance of 120 cm mobile application calculated the distance to the refrigerator with 100% accuracy. To evaluate the algorithm's performance, at a distance of 120 cm mobile application estimated the distance to the large objects with 100% accuracy. Table 2 presents 120 cm in range evaluation results.

Table 2. Model-1 Evaluation Results at 120 cm in Range

120 cm Range Object Distance Estimation Accuracy Evaluation Results				
	Couch	Chair	Bed	Refrigerator
Is Object Detected	Yes	Yes	Yes	Yes
Inference Time	549ms	610ms	440ms	583ms
Actual Distance	120 cm	120 cm	120 cm	120 cm
Estimated Distance	120 cm	120 cm	120 cm	120 cm
Distance Estimation Error (cm)	0 cm	0 cm	0 cm	0 cm
Obstacle or Not	Yes, An Obstacle	Yes, An Obstacle	Yes, An Obstacle	Yes, An Obstacle
Mobile Real-time Audio Output Notification	Caution Obstacle Ahead Couch Distance 120 cm Proceed with Care	Caution Obstacle Ahead Chair Distance 120 cm Proceed with Care	Caution Obstacle Ahead Bed Distance 120 cm Proceed with Care	Caution Obstacle Ahead Refrigerator Distance 120 cm Proceed with Care
Objects Accurately Detected	4/4			
Objects Not Detected	0/4			
Obstacles Accurately Identified	4/4			
Obstacles Not Identified	0/4			
Total Distance Estimation Accuracy	100 %			

According to experiment-1 research results at a distance of 150 cm, mobile applications could detect 4 out of 4 objects. At the distance of 150 cm mobile application calculated the distance to the couch with 100% accuracy. At the distance of 150 cm mobile application calculated the distance to the chair with 100% accuracy. At the distance of 150 cm mobile application calculated the distance to the bed with 99% accuracy. At the distance of 150 cm mobile application calculated the distance to the refrigerator with 100% accuracy. To evaluate the algorithm's performance, at a distance of 150 cm mobile application estimated the distance to the large objects with 99.75% accuracy. Table 3 presents 150 cm in range evaluation results.

Table 3. Model-1 Evaluation Results at 150 cm in Range

150 cm Range Object Distance Estimation Accuracy Evaluation Results				
	Couch	Chair	Bed	Refrigerator
Is Object Detected	Yes	Yes	Yes	Yes
Inference Time	771ms	623ms	620ms	670ms
Actual Distance	150 cm	150 cm	150 cm	150 cm
Estimated Distance	150 cm	150 cm	151 cm	150 cm
Distance Estimation Error (cm)	0 cm	0 cm	1 cm	0 cm
Obstacle or Not	Yes, An Obstacle	Yes, An Obstacle	Yes, An Obstacle	Yes, An Obstacle
Mobile Real-time Audio Output Notification	Caution Obstacle Ahead Couch Distance 150 cm Proceed with Care	Caution Obstacle Ahead Chair Distance 150 cm Proceed with Care	Caution Obstacle Ahead Bed Distance 151 cm Proceed with Care	Caution Obstacle Ahead Refrigerator Distance 150 cm Proceed with Care
Objects Accurately Detected	4/4			
Objects Not Detected	0/4			
Obstacles Accurately Identified	4/4			
Obstacles Not Identified	0/4			
Total Distance Estimation Accuracy	99.75 %			

According to experiment-1 research results at a distance of 200 cm, mobile applications could detect 4 out of 4 objects. At the distance of 200 cm mobile application calculated the distance to the couch with 99% accuracy. At the distance of 200 cm mobile application calculated the distance to the chair with 99% accuracy. At the distance of 200 cm mobile application calculated the distance to the bed with 100% accuracy. At the distance of 200 cm mobile application calculated the distance to the refrigerator with 99% accuracy. To evaluate the algorithm's performance, at a distance of 200 cm mobile application estimated the distance to the large objects with 99.25% accuracy. Table 4 presents 200 cm in range evaluation results.

Table 4. Model-1 Evaluation Results at 200 cm in Range

200 cm Range Object Distance Estimation Accuracy Evaluation Results				
	Couch	Chair	Bed	Refrigerator
Is Object Detected	Yes	Yes	Yes	Yes
Inference Time	575ms	627ms	611ms	560ms
Actual Distance	200 cm	200 cm	200 cm	200 cm
Estimated Distance	201 cm	201 cm	200 cm	201 cm
Distance Estimation Error (cm)	1 cm	1 cm	0 cm	1 cm
Obstacle or Not	Yes, An Obstacle	Yes, An Obstacle	Yes, An Obstacle	Yes, An Obstacle
Mobile Real-time Audio Output Notification	Caution Obstacle Ahead Couch Distance 201 cm Proceed with Care	Caution Obstacle Ahead Chair Distance 201 cm Proceed with Care	Caution Obstacle Ahead Bed Distance 200 cm Proceed with Care	Caution Obstacle Ahead Refrigerator Distance 201 cm Proceed with Care
Objects Accurately Detected	4/4			
Objects Not Detected	0/4			
Obstacles Accurately Identified	4/4			
Obstacles Not Identified	0/4			
Total Distance Estimation Accuracy	99.25 %			

5.2 Object Detection and Distance Estimation Custom (Model-2)

In experiment-2 custom deep learning model-2 was trained using a custom dataset. To detect custom indoor objects along with distance, the walking stick, door and cupboard. At the time of research, no dataset was available open free for download without copyright for object classes walking stick, cupboard and bedroom doors. We collected dataset images for this research, we labelled and annotated images and trained a custom object detection deep learning model using our novel dataset. During experiment-2, our research aim was to evaluate mobile deep learning-based custom indoor object detection and distance estimation for visually impaired people. First, we collect data on custom objects using a mobile camera. Then we analysis and evaluate collected data for indoor custom object detection and classification and distance estimation evaluation by comparing actual and estimated distance. Mobile applications real-time detect custom objects. Mobile application classifies objects with estimated distance and results in a voice as feedback. The mobile camera captures real-time video frames. The mobile deep learning module processes frames, and the distance estimation algorithm estimates the distance to the object, and the voice module provides speech output.

Mobile deep learning-based android application with video frame input size 300*300, our experiment recorded actual distance results of estimated distance, distance estimation error and distance estimation accuracy. Mobile deep learning-based android application for visually challenged people object walking stick placed at the actual distance of 100 cm mobile application estimated distance 100 cm, object walking stick placed at the actual distance of 120 cm mobile application estimated distance 121 cm, object walking stick placed at the actual distance of 150 cm mobile application estimated distance 149 cm, similarly walking stick at the actual distance of 200 cm mobile application estimated distance as 200 cm. Figure 10 shows screenshots of mobile application estimating object distance.

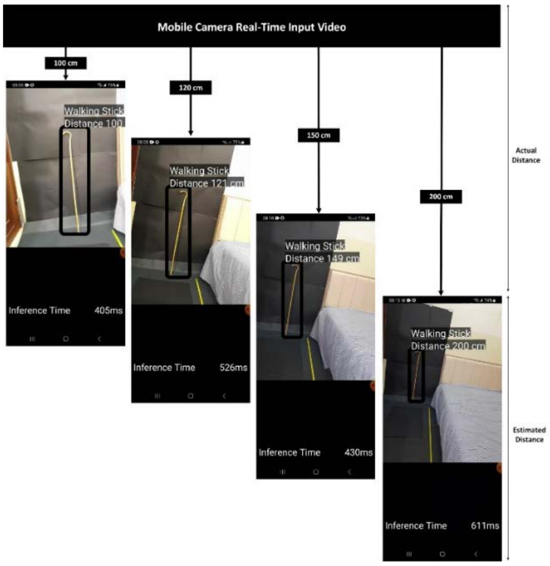

Figure 10. Walking Stick Distance Estimation

Mobile deep learning-based android application for visually challenged people object door placed at the actual distance of 100 cm mobile application estimated distance 100 cm, object door placed at the actual distance of 120 cm mobile application estimated distance 120 cm, object door placed at the actual distance of 150 cm mobile application estimated distance 151 cm, similarly door at the actual distance of 200 cm mobile application estimated distance as 200 cm. Figure 11 shows screenshots of mobile application estimating object distance.

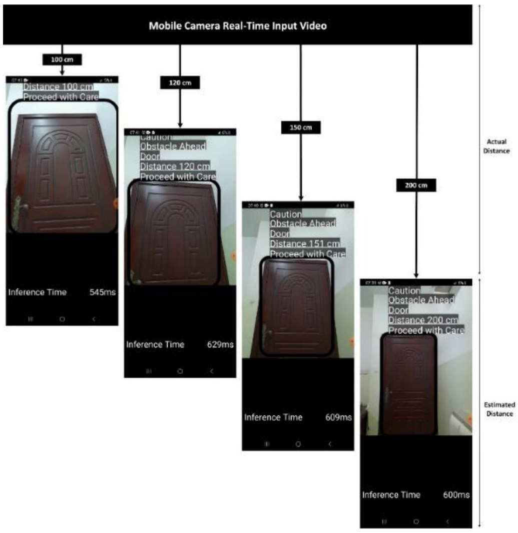

Figure 11. Door Distance Estimation

Mobile deep learning-based android application for visually challenged people object wardrobe placed at the actual distance of 100 cm mobile application estimated distance 100 cm, object wardrobe placed at the actual distance of 120 cm mobile application estimated distance 120 cm, object wardrobe placed at the actual distance of 150 cm mobile application estimated distance 150 cm, similarly wardrobe at the actual distance of 200 cm mobile application estimated distance as 200 cm. Figure 12 shows screenshots of mobile application estimating object distance.

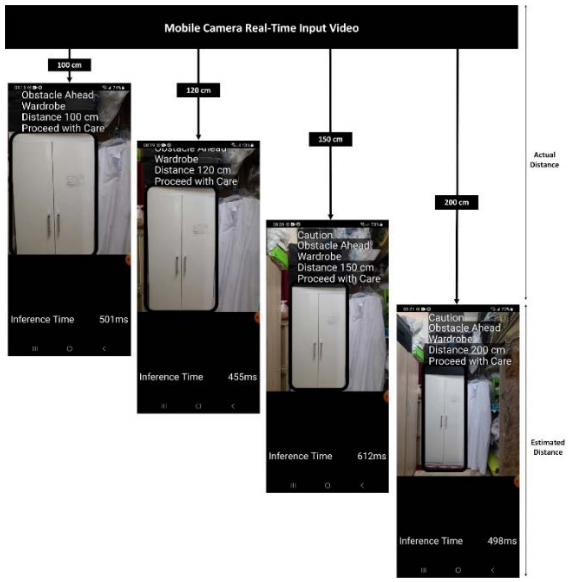

Figure 12. Wardrobe Distance Estimation

In experiment-2 we evaluated custom Model-2 objects detection and distance estimation for custom objects walking stick, door and wardrobe a distance of 100cm, 120cm, 150cm and 200cm. At a distance of 100 cm, custom indoor objects walking stick, door and wardrobe were classified and detected by drawing a bounding box around objects.

According to experiment-2 research results at a distance of 100 cm, mobile applications detect 3 out of 3 objects. At the actual distance of 100 cm mobile application calculated the distance to the walking stick 100 cm accuracy with 0 cm distance estimation error. At the actual distance of 100 cm mobile application calculated the distance to the door 100 cm accuracy with 0 cm distance estimation error. At the actual distance of 100 cm mobile application calculated the distance to the wardrobe 100 cm accuracy with 0 cm distance estimation error. Table 5 shows custom model-2 evaluation results at 100 cm in range.

Table 5. Custom Model (Model-2) Evaluation Results at 100 cm in Range

100 cm Range Object Distance Estimation Accuracy Evaluation Results			
	Walking Stick	Door	Wardrobe
Is Object Detected	Yes	Yes	Yes
Inference Time	405ms	545ms	501ms
Actual Distance	100 cm	100 cm	100 cm
Estimated Distance	100 cm	100 cm	100 cm
Distance Estimation Error (cm)	0 cm	0 cm	0 cm
Obstacle or Not	No Obstacle	Yes, An Obstacle	Yes, An Obstacle
Mobile Real-time Audio Output Notification	Walking Stick Distance 100 cm	Caution Obstacle Ahead Door Distance 100 cm Proceed with Care	Caution Obstacle Ahead Wardrobe Distance 100 cm Proceed with Care
Objects Accurately Detected	3/3		
Objects Not Detected	0/3		
Obstacles Accurately Identified	2/2		
Obstacles Not Identified	0/2		
Total Distance Estimation Accuracy	100 %		

Based on the finding from experiment-2 research results at a distance of 120 cm, mobile applications detect 3 out of 3 objects. At the actual distance of 120 cm mobile application calculated the distance to the walking stick with 121 cm with distance estimation error 1 cm. At the actual distance of 120 cm mobile application calculated the distance to the door with 120 cm with distance estimation error 0 cm. At the actual distance of 120 cm mobile application calculated the distance to the wardrobe with 120 cm with distance estimation error 0 cm. Table 6 shows custom model-2 evaluation results at 120 cm in range.

Table 6. Custom Model (Model-2) Evaluation Results at 120 cm in Range

120 cm Range Object Distance Estimation Accuracy Evaluation Results			
	Walking Stick	Door	Wardrobe
Is Object Detected	Yes	Yes	Yes
Inference Time	526ms	629ms	455ms
Actual Distance	120 cm	120 cm	120 cm
Estimated Distance	121 cm	120 cm	120 cm
Distance Estimation Error (cm)	1 cm	0 cm	0 cm
Obstacle or Not	No Obstacle	Yes, An Obstacle	Yes, An Obstacle
Mobile Real-time Audio Output Notification	Walking Stick Distance 121 cm	Caution Obstacle Ahead Door Distance 120 cm Proceed with Care	Caution Obstacle Ahead Wardrobe Distance 120 cm Proceed with Care
Objects Accurately Detected	3/3		
Objects Not Detected	0/3		
Obstacles Accurately Identified	2/2		
Obstacles Not Identified	0/2		
Total Distance Estimation Accuracy	99.66 %		

Based on the finding from experiment-2 research results at a distance of 150 cm, mobile applications detect 3 out of 3 objects. At the actual distance of 150 cm mobile application calculated the distance to the walking stick with 149 cm with distance estimation error 1 cm. At the actual distance of 150 cm mobile application calculated the distance to the door with 151 cm with distance estimation error 1 cm. At the actual distance of 150 cm mobile application calculated the distance to the wardrobe with 150 cm with distance estimation error 0 cm. Table 7 shows custom model-2 evaluation results at 150 cm in range.

Table 7. Custom Model (Model-2) Evaluation Results at 150 cm in Range

150 cm Range Object Distance Estimation Accuracy Evaluation Results			
	Walking Stick	Door	Wardrobe
Is Object Detected	Yes	Yes	Yes
Inference Time	430ms	609ms	612ms
Actual Distance	150 cm	150 cm	150 cm
Estimated Distance	149 cm	151 cm	150 cm
Distance Estimation Error (cm)	1 cm	1 cm	0 cm
Obstacle or Not	No Obstacle	Yes, An Obstacle	Yes, An Obstacle
Mobile Real-time Audio Output Notification	Walking Stick Distance 149 cm	Caution Obstacle Ahead Door Distance 151 cm Proceed with Care	Caution Obstacle Ahead Wardrobe Distance 150 cm Proceed with Care
Objects Accurately Detected	3/3		
Objects Not Detected	0/3		
Obstacles Accurately Identified	2/2		
Obstacles Not Identified	0/2		
Total Distance Estimation Accuracy	99.33 %		

According to experiment-2 research results at a distance of 200 cm, mobile applications detect 3 out of 3 objects. At the actual distance of 200 cm mobile application calculated the distance to the walking stick 200 cm accuracy with 0 cm distance estimation error. At the actual distance of 200 cm mobile application calculated the distance to the door 200 cm accuracy with 0 cm distance estimation error. At the actual distance of 200 cm mobile application calculated the distance to the wardrobe 200 cm accuracy with 0 cm distance estimation error. Table 8 shows custom model-2 evaluation results at 200 cm in range.

Table 8. Custom Model (Model-2) Evaluation Results at 200 cm in Range

200 cm Range Object Distance Estimation Accuracy Evaluation Results			
	Walking Stick	Door	Wardrobe
Is Object Detected	Yes	Yes	Yes
Inference Time	611ms	600ms	498ms
Actual Distance	200 cm	200 cm	200 cm
Estimated Distance	200 cm	200 cm	200 cm
Distance Estimation Error (cm)	0 cm	0 cm	0 cm
Obstacle or Not	No Obstacle	Yes, An Obstacle	Yes, An Obstacle
Mobile Real-time Audio Output Notification	Walking Stick Distance 200 cm	Caution Obstacle Ahead Door Distance 200 cm Proceed with Care	Caution Obstacle Ahead Wardrobe Distance 200 cm Proceed with Care

Objects Accurately Detected	3/3
Objects Not Detected	0/3
Obstacles Accurately Identified	2/2
Obstacles Not Identified	0/2
Total Distance Estimation Accuracy	100 %

Comparison of (Model-1) and (Model-2)

Table 9. Model-1 and Custom Model (Model-2) Evaluation Results Comparison

	Model-1 Summarized Evaluation Result (Couch, Chair, Bed, Refrigerator)		Model-2 Custom Model Summarized Evaluation Result Estimation Accuracy (Walking Stick, Door, Wardrobe)	
	Actual Objects/ Detected Objects	Distance Estimation Accuracy	Actual Objects/ Detected Objects	Distance Estimation Accuracy
100 cm	4/4	100 %	3/3	100%
120 cm	4/4	100 %	3/3	99.66%
150 cm	4/4	99.75 %	3/3	99.33%
200 cm	4/4	99.25 %	3/3	100%
Total	16/16		12/12	
	100%	99.75%	100%	99.75%

Table 9 shows the Model-1 and Custom Model (Model-2) Evaluation Results Comparison. Model-1 detected 16 out of 16 objects at the distance of 100cm, 120cm, 150cm and 200cm. Custom model-2 detected 12 out of 12 objects at the distance of 100cm, 120cm, 150cm and 200cm. MobileNet-V2-based Model-1 99.75% accurately estimated distance objects placed at 100cm, 120cm, 150cm and 200cm. Our proposed mobile deep learning custom model-2 99.75% accurately estimated distance objects placed at 100cm, 120cm, 150cm and 200cm. Our research finding shows the effectiveness of our proposed framework for mobile object detection and distance estimation. This is because our framework relies on the success of the deep learning model real-time object recognition to calculate object range. Table 10 shows the limitations of the existing Apple App Store and Google Play store-published object detection navigation Apps for visually challenged people.

Table 10. Limitations of the Existing Apple App Store and Google Play Store Navigations Apps for Object Detection to Assist Visually Impaired or Blind People.

	Limitations of the existing Apple app store and Google play store published navigations apps for the visually challenged peoples.				
	Object Detection	Voice Feedback	Obstacle Detection	Offline App	Distance Estimation
Seeing AI [18]	Preview Available	✓	✗	Limited Functions	✗
EnvisionAI [19]	✓	✓	✗	✗	✗
AipolyVision [20]	✓	✓	✗	✓	✗
TapTapSee [21]	Image upload to the cloud	✓	✗	✗	✗
Be My Eyes [22]	✗	✗	✗	✗	✗
Our Proposed Framework	✓	✓	✓	✓	✓

6. Conclusion

In this research, we propose a mobile deep learning-based object identification and range estimation framework locally on a mobile device to detect and classify objects in real-time using a built-in camera with voice commentary to assist visually challenged people in safe indoor navigation. Most published navigational apps for visually challenged people require a good internet connection. The existing systems that assist the visually challenged in navigation lack real-time raising alerts/ warnings in case of obstacle detection. In the past, several pieces of research have been done to identify and recognize objects in digital images or videos but do not estimate object distance. Results demonstrate that our custom dataset-trained mobile deep learning model effectively detects indoor objects for visually impaired people. Model-1 detected 16 out of 16 objects door, couch, chair, bed and refrigerator at 100cm, 120cm, 150cm and 200cm, with a distance estimation accuracy of 99.75%. Our custom dataset trained custom mobile deep learning model-2 detected 12 out of 12 objects, walking stick, wardrobe and refrigerator at distances of 100cm, 120cm, 150cm and 200cm with a distance estimation accuracy of 99.75%. The findings show that the mobile deep learning-based application effectively identifies indoor objects and estimates range with voice guidance to assist visually impaired people in safe indoor navigation. A dearth of research has been reported on deep learning models trained using the dataset, and the deep learning feature extraction technique classifies objects with distance. It is vital and imperative to fill this research gap in the research domain, and the stated research gap is the basis of this study's proposed framework.

References

[1] M. E. Gilbert, “Digitally Engaging Older Populations in Healthcare Requires New Practices to Be Effective,” *Gartner*, pp. 1–21, 2020.

[2] G. Qiao, H. Song, B. Prideaux, and S. (Sam) Huang, “The ‘unseen’ tourism: Travel experience of people with visual impairment,” *Ann. Tour. Res.*, vol. 99, p. 103542, 2023.

[3] L. Kay, “A sonar aid to enhance spatial perception of the blind: engineering design and evaluation,” *Radio Electron. Eng.*, pp. 605–627, 1974.

[4] B. Durette, N. Louveton, D. Alleysson, and J. Herault, “Visuo-auditory sensory substitution for mobility assistance: testing TheVIBE,” *Work. Comput. Vis. Appl. Vis. Impair. - Marseille, Fr.*, 2008.

[5] B. D. C. Martinez, O. O. V. Villegas, V. G. C. Sanchez, H. de J. O. Domínguez, and L. O. Maynez, “Visual perception substitution by the auditory sense,” *Int. Conf. Comput. Sci. Its Appl. Springer, Berlin, Heidelb.*, pp. 522–533, 2011.

[6] L. Tepelea, V. Tiponut, P. Szolgay, and A. Gacsadi, “Multicore portable system for assisting visually impaired people,” *Int. Work. Cell. Nanoscale Networks their Appl.*, pp. 3–4, 2014.

[7] H. Jabnoun, F. Benzarti, and H. Amiri, “Object detection and identification for blind people in video scene,” *15th Int. Conf. Intell. Syst. Des. Appl. - IEEE - 2015*, 2015.

[8] A. N. Zereen and S. Corraya, “Detecting real time object along with the moving direction for visually impaired people,” *ICECTE 2016 - 2nd Int. Conf. Electr. Comput. Telecommun. Eng.*, no. December, pp. 8–10, 2016.

[9] C. T. Patel, V. J. Mistry, L. S. Desai, and M. Y. K., “Multisensor-based object detection in indoor environment for visually impaired people,” *2018 Second Int. Conf. Intell. Comput. Control Syst. (ICICCS)-IEEE*, no. Iciccs, pp. 2018–2021, 2018.

[10] L. Árvai, “Mobile phone based indoor navigation system for blind and visually impaired people: VUK - Visionless supporting frameworK,” *Proc. 2018 19th Int. Carpathian Control Conf. ICCC 2018*, pp. 383–388, 2018.

[11] Y. Lin, K. Wang, W. Yi, and S. Lian, “Deep Learning based Wearable Assistive System for Visually Impaired People,” *Proc. IEEE/CVF Int. Conf. Comput. Vis. Work.*, 2019.

[12] S. Ooi, T. Okita, and M. Sano, “Study on A Navigation System for Visually Impaired Persons based on Egocentric Vision Using Deep Learning,” *ACM Int. Conf. Proceeding Ser.*, pp. 68–72, 2020.

[13] S. Mahmud, R. Haque Sourave, M. Islam, X. Lin, and J. H. Kim, “A vision based voice controlled indoor assistant robot for visually impaired people,” *IEMTRONICS 2020 - Int. IOT, Electron. Mechatronics Conf. Proc.*, 2020.

[14] B. Strbac, M. Gostovic, Z. Lukac, and D. Samardzija, “YOLO Multi-Camera Object Detection and Distance Estimation,” *2020 Zooming Innov. Consum. Technol. Conf. ZINC 2020*, pp. 26–30, 2020.

[15] K. Karthika, S. Adarsh, and K. I. Ramachandran, “Distance Estimation of Preceding Vehicle Based on Mono Vision Camera and Artificial Neural Networks,” *2020 11th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2020*, 2020.

[16] N. Sakic, M. Krunic, S. Stevic, and M. Dragojevic, “Camera-LIDAR Object Detection and Distance Estimation with Application in Collision Avoidance System,” *IEEE Int. Conf. Consum. Electron. - Berlin, ICCE-Berlin*, vol. 2020-Novem, 2020.

[17] M. M. Rahman, M. M. Islam, S. Ahmmed, and S. A. Khan, “Obstacle and Fall Detection to Guide the Visually Impaired People with Real Time Monitoring,” *SN Comput. Sci. - Springer*, vol. 1, no. 4, pp. 1–10, 2020.

[18] Microsoft, “Seeing AI,” *Microsoft Corporation*, 2021. [Online]. Available: https://apps.apple.com/us/app/seeing-ai/id999062298. [Accessed: 01-Feb-2021].

[19] Envision, “Envision AI,” *Envision Technologies B.V.*, 2021. [Online]. Available: https://apps.apple.com/us/app/envision-ai/id1268632314. [Accessed: 01-Feb-2021].

[20] Aipoly, “Aipoly Vision,” *Aipoly - V7 Ltd*, 2021. [Online]. Available: https://play.google.com/store/apps/details?id=com.aipoly.vision&hl=en&gl=US. [Accessed: 01-Feb-2021].

[21] TapTapSee, “TapTapSee,” *Cloudsight, Inc*, 2021. [Online]. Available: https://apps.apple.com/us/app/taptapsee/id567635020. [Accessed: 02-Feb-2021].

[22] BeMyEyes, “Be My Eyes,” *Be My Eyes*, 2021. [Online]. Available: https://apps.apple.com/us/app/be-my-eyes/id905177575. [Accessed: 02-Feb-2021].