OBJECT DETECTION AND RECOGNITION FRAMEWORK FOR THE VISUALLY-IMPAIRED

A V SAI RADHIKA[1], BANDA SINDHURA[2], GOLI SNIGDHA REDDY[3], P SHRAVANI[4]

ASSISTANT PROFESSOR[1], UG SCHOALR [2, 3 &4]

DEPARTMENT OF CSE, BHOJ REDDY ENGINEERING COLLEGE FOR WOMEN, VINAY NAGAR, HYDERABAD-59

SCOPE

Object detection and recognition framework for the visually-impaired" is a deep learning based application, This framework uses the SSD algorithm which will create bounding boxes around the objects in the frame and classify them into given categories. Lastly application will produce an audio output of the object detected**.**

ABSTRACT

Object detection plays a very important role in many applications such as image retrieval, surveillance, robot navigation, wayfinding, etc. In this thesis, we propose different approaches to detect indoor signage, stairs and pedestrians. In the first chapter we introduce some related work in this field. In the second chapter, we introduced a new method to detect the indoor signage to help blind people find their destination in unfamiliar environments. Our method first extracts the attended areas by using a saliency map. Then the signage is detected in the attended areas by using bipartite graph matching. The proposed method can handle multiple signage detection. Experimental results on our collected indoor signage dataset demonstrate the effectiveness and efficiency of our proposed method. Furthermore, saliency maps could eliminate the interference information and improve the accuracy of the detection results. In the third chapter, we present a novel camera-based approach to automatically detect and recognize restroom signage from surrounding environments. Our method first extracts the attended areas which may content signage based on shape detection. Then, Scale-Invariant Feature Transform (SIFT) is applied to extract local features in the detected attended areas. Finally, signage is detected and recognized as the regions with the SIFT matching scores larger than a threshold. The proposed method can handle multiple signage detection. Experimental results on our collected restroom signage dataset demonstrate the effectiveness and efficiency of our proposed method. In the fourth chapter, we develop a new framework to detect and recognize stairs and pedestrian crosswalks using a RGBD camera. Since both stairs and pedestrian crosswalks are featured by a group of parallel lines, we first apply Hough transform to extract the concurrent parallel lines based on the RGB channels. Then, the Depth channel is employed to further recognize pedestrian crosswalks, upstairs, and downstairs using support vector machine (SVM) classifiers. Furthermore, we estimate the distance between the camera and stairs for the blind users. The detection and recognition results on our collected dataset demonstrate that the effectiveness and efficiency of our proposed framework

Keywords: Blind people, Navigation and wayfinding, Camera, Signage detection and recognition, Independent travel"

INTRODUCTION

Based on the study of the World Health Organization (WHO), there were about 161 million visually impaired people around the world in 2002, about 2.6% of the total population. Among these statistics, 124 million had low vision and 37 million were blind [1]. Robust and efficient indoor object detection can help people with severe vision impairment to independently access unfamiliar indoor environments and avoid dangers [2] . Context information (including signage and other visual information) plays an important role in navigation and wayfinding for unsighted persons. As shown in Figure 1, signage is particularly important for discriminating between similar objects in indoor environments such as elevators, bathrooms, exits, and office doors

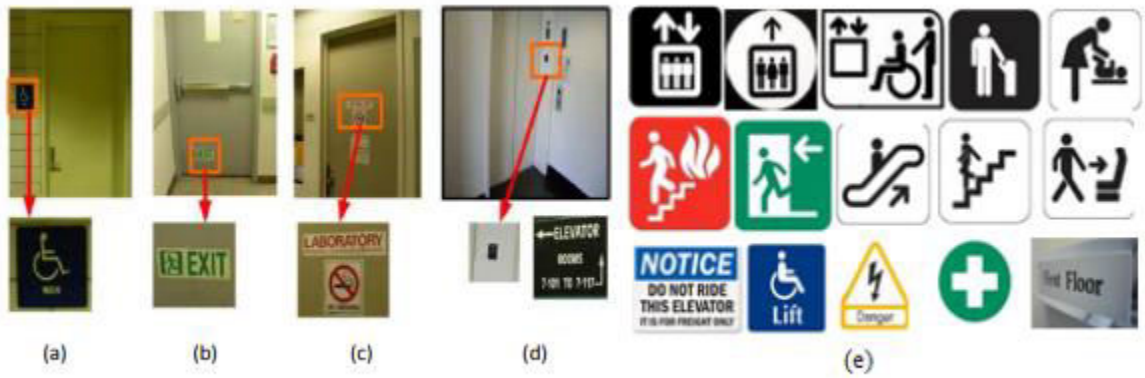


Figure 1. Typical indoor signage: (a) a bathroom, (b) an exit, (c) a laboratory, (d) an elevator, (e) stairs

Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of a priori objects of a certain class (such as faces, signs, buildings, etc) in digital images and videos captured by cameras [3]. Camera-based indoor signage detection is a challenging problem due to the following factors: 1) large variations of appearance and design (shape, color, texture, etc.) of signage in different buildings; and 2) large variations in the camera view and image resolution of signage due to changes in position and distance between the blind user with wearable cameras and the targeted signage. Some examples of camera used in our experiment are shown in Figure 2

Figure 2. Wearable camera used in our test procedures

Object detection and recognition is a fundamental component for scene understanding. The human visual system is powerful, selective, robust, and fast [4]. It is not only very selective, which allows us to distinguish among very similar objects, such as the faces of identical twins, but also robust enough to classify same category objects with large variances (e.g. changes of position, scale, rotation, illumination, color, occlusion and many other properties). Research shows that the human visual system can discriminate among at least tens of thousands of different object categories [5]. Object recognition processes in the human visual system are also very fast: it can take as little as 100 to 200 ms [6][7][8]. However, it is extremely difficult to build robust and selective computer vision algorithms for object recognition which can handle very similar objects or objects with large variations. For example, state-of-the-art object detection methods require hundreds or thousands of training examples and very long durations to learn visual models of one object category [9][10][11].

RELATED WORK

1) Real-Time Objects Recognition Approach for Assisting Blind People: In this paper, two cameras placed on blind person's glasses, GPS free service, and ultrasonic sensors are employed to provide information about the surrounding environment. Object detection is used to find objects in the real world such as faces, bicycles, chairs, doors, or tables that are common in the scenes of a blind. Here, GPS service is used to create groups of objects based on their locations, and the sensor detects an obstacle at a medium to long distance. The descriptor of the Speeded-Up Robust Features (SURF) method is optimized to perform the recognition. The use of two cameras on glasses can be sophisticated. [2]

2) Wearable Object Detection System for the Blind: In this paper, the RFID device is designed as a support for the blind for the disclosure of objects; especially, it is developed for searching the medicines in a cabinet at home. This device can provide information about the distance of a defined object, how near or far it is and simplifies the search. For identifying the medicines, the device can provide the user with an acoustic signal to find the desired product as soon as possible. The measure of the distance, in particular the movement of the antenna respect to tag, is made using the RSSI value. This application uses the RSSI (Received Strength Signal Indicator) value, measuring the power of the received signal of the tag. [3]

3) Smart Obstacle Detector for Blind Person: Another system proposed in this paper focuses on giving information about what are the different types of obstacles in front of the user, their size, and their distance from the user. MATLAB Software is used for signal processing. The camcorder is used for recording videos. Video processing methods are used after that. The output of this system not only gives output in audio format but also vibration. A vibrating motor has been connected with an ultrasonic sensor. The ultrasonic sensor detects objects coming in its range and this makes the vibrating motor vibrate. Use of Camcorder, a stick with an Ultrasonic sensor makes this system bulky and dependent on the stick

EXISTING SYSTEM

The Existing System is navigating manually.Navigating around is one of the biggest challenges visually impaired people face. It is difficult for them to travel independently as they cannot analyze the position of the objects and the

people surrounding them. The white cane is one of the most common aids for the visually impaired people. Though it is helpful in navigating around, it does not inform the user about the various obstacles until they are very close to them

DISADVANTAGES

- Increased interference from the public wanting to assist
- Cane travel can be more cumbersome and not as fluid. A cane gets stuck in cracks and you get a poke in the stomach
- Weather negatively impacts cane travelers.
- Folded cane requires large area for storing or large bag for carrying while folded.

PROPOSED SYSTEM

- Object detection is the principle objective of this system. It consists of object classification and object localization.
- The application is build to recognize or detect some household objects like chair, bed, refrigerator, laptops, etc and some outdoor objects like vehicles, people, etc.
- The application will use mobile phone camera to scan the surrounding in real time and take the frames from the ongoing video.

ADVANTAGES

- Detects objects quickly and specifies the object type which helps to avoid obstacles.
- Can be used at any place and any time irrespective of the weather conditions and locations.
- Informs the user about the object without getting close to the object.
- Decreases the public interference wanting to assit

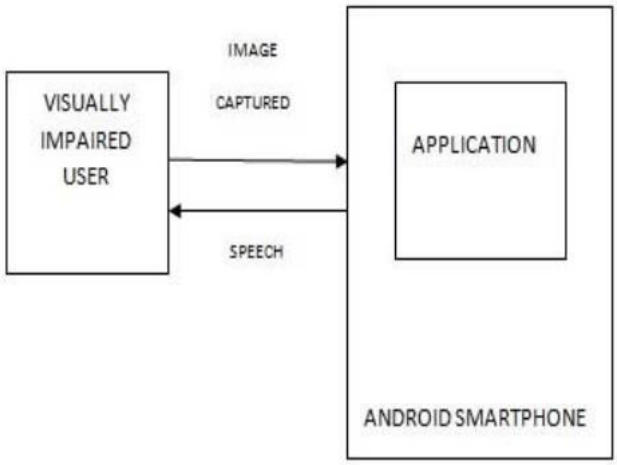
IMPLEMENTATION

Data Preparation and Preprocessing

- **Dataset**
- **Data Preparation**
- **Data Labeling**
- **Train-Test Split**
- **Model Training**

- **Object Detection**

ARCHITECTURE



ALGORITHMS

SSD(Single shot detector):

SSD Algorithm is used for the object recognition as well as detection. Also this algorithm gives nearly accurate results for real time object detection and is proven to be faster than other relative algorithms.

- SSD takes only one shot to detect multiple objects present in an image using multibox.
- There exist 2 sections for an object detection system using SSD algorithm.

    1. One is to extract the feature maps
    2. Second is to apply filters of convolution to detect objects.

CONVOLUTIONAL NEURAL NETWORK(CNN) :

- CNN is a neural network with multiple layers. It is a class of Deep Neural Network that is mostly used for the purpose of analysis of visual images.
- CNN contains one layer for input, called as input layer, next it has one or more middle layers which are called as convolutional layers, next the network consists of one or more fully connected layers and lastly the network is completed by adding an output layer in the respective order.

It works as follows:

- A convoluted process is applied to the incoming data by the convolutional layers after which the outcome that is generated is passed to the next layer.

- This convolution layer has reaction similar to the human's neuron to vision process.

FUTURE SCOPE The future perspective of this project is to increase the object recognition rate which can be achieved by using the TensorFlow library and to provide an exact distance measurement between the people and object. However, for developing an application that involves many objects that are fast-moving, you should instead consider faster hardware. Further, we can implement face recognition and text recognition in the same system. Thus, making the system compatible overall.

CONCLUSION In recent years, some solutions have been devised to help blind or visually impaired in recognizing objects in their environment but they are not efficient. Our purpose is to provide a robust and comfortable system for the blind to recognize their surrounding objects. Our advanced system uses a USB camera to seize real-time images in front of the users. The machine learning and feature extraction technique used here is YOLO. The YOLO framework trades with object detection by choosing the entire image in a single instance, and splits the image into grids, then predicts the bounding box coordinates and class probabilities for these boxes. The biggest advantage of sing YOLO is its excellent speed – it's incredibly fast and YOLO also understands generalized object representation. This system will make visually impaired virtually visible also it innovatively uses the text-to-speech technology which provides audio descriptions of their surroundings and helps them to travel with self-confidence. The proposed system is mobile, robust, and efficient. Also, it creates a virtual environment and this system provides a sense of assurance as it voices the name of the object recognized.

REFERENCES

Book,

[1] Peter Harrington, "Machine Learning in Action, pp."by Manning Publications-1st edition.

Papers,

[2] Zraqou, Jamal & Alkhadour, Wissam & Siam, Mohammad. (2017). "Real-Time Objects Recognition Approach for Assisting Blind People." Multimedia Systems Department, Electrical Engineering Department, Isra University, Amman-Jordan Accepted 30 Jan 2017, Available online 31 Jan 2017, Vol.7, No.1

[3] A. Dionisi, E. Sardini and M. Serpelloni, "Wearable object detection system for the blind," 2012 IEEE International Instrumentation and Measurement Technology Conference Proceedings, Graz, 2012, pp. 1255-1258, doi: 10.1109/I2MTC.2012.6229180.

[4] Daniyal, Daniyal & Ahmed, Faheem & Ahmed, Habib & Shaikh, Engr & Shamshad, Aamir. (2014). "Smart Obstacle Detector for Blind Person." Journal of Biomedical Engineering and Medical Imaging. 1. 31-40. 10.14738/jbemi.13.245.

[5] Christian Szegedy Alexander Toshev Dumitru Erhan, "Deep Neural Networks for Object Detection."

[6] N.Saranya, M.Nandinipriya, U.Priya,"Real Time Object Detection for Blind People",Bannari Amman Institute of Technology, Sathyamangalam, Erode.(India).

[7] Rui (Forest) Jiang,Qian Lin,Shuhui Qu,"Let Blind People See: RealTime Visual Recognition with Results Converted to 3D Audio",Stanford,2018. Website,

[8] https://www.who.int/news-room/detail/08-10-2019-who-launchesfirst-world-report-on-vision

[9] https://www.analyticsvidhya.com/blog/2018/12/practical-guideobject-detection-yolo-framewor-python/

[10] http://datahacker.rs/face-detection-opencv-images/

[11] https://towardsdatascience.com/yolo-v3-object-detection53fb7d3bfe6b

[12] https://medium.com/@jonathan_hui/real-time-object-detection-withyolo-yolov2-28b1b93e2088

[13] https://opencv.org/about/ [14] https://github.com/TheNsBhasin/DNN_Object_Detection