International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 08 Issue: 04 | Apr 2021 www.irjet.net p-ISSN: 2395-0072

An Intelligent Assistive System for Visually Impaired People

Prof. Vrushali V. Kondhalkar[1], Ayesha R. Shaikh[2], Yogesh A. Chaudhari[3], Omkar R. Bhandar[4], Gaurav B. Katkar[5]

[1]*Prof. Vrushali V. Kondhalkar, Dept. of Computer Engineering, JSCOE, Pune, Maharashtra, India*
[2]*Ayesha R. Shaikh, Dept. of Computer Engineering, JSCOE, Pune, Maharashtra, India*
[3]*Yogesh A. Chaudhari, Dept. of Computer Engineering, JSCOE, Pune, Maharashtra, India*
[4]*Omkar R. Bhandar, Dept. of Computer Engineering, JSCOE, Pune, Maharashtra, India*
[5]*Gaurav B. Katkar, Dept. of Computer Engineering, JSCOE, Pune, Maharashtra, India*

Abstract – *According to WHO (World Health Organization) over 280 million people worldwide live with visual impairment which prevents them from basic access to written or visual material. They often need to identify objects around them. How do these people see or recognize objects and read? The only way is by touching and listening. That is the major problem faced by visually impaired people, to detect the object without touching it and in this pandemic situation with COVID like deadly virus around, touching any object directly with bare hands is not safe. In past decades, many efforts have been done in developing several devices or techniques to assist the visually impaired. Many of those are too costly or heavy or doesn't help the visually impaired much. So, the goal of the proposed system is to develop an application that will help them to recognize and identify various objects and currencies and understand image/ scene text in real-time using object, text and currency detection and recognition algorithms and assist them hence leading to improve independence and ease of its user.*

Key Words: Visually Impaired, Object Detection, Object Recognition, Text Recognition, Currency Recognition

1. INTRODUCTION

The term Visual Impairment as experts describe refers to total or partial vision loss which cannot be fixed even with any corrective means such as eyeglasses or contact lenses or even eye surgery. Vision Loss is most often accompanied by loss of independence and the person needs some sort of assistance to carry out day to day activities. They often need to identify objects around them, from a pen to clothing items to a person, read from a note to a bill to books. Hence, it is necessary to develop an intelligent system that will assist visually impaired people. An intelligent system that helps in object, text and currency detection and recognition which then provides output in form of audio/ speech.

2. OBJECT DETECTION

Object Detection is a computer vision technology that works with artificial intelligence and machine learning. This is the technology where you can find multiple objects present in front of the person in a single frame. Object detection has various problem-solving capabilities. This technique helps visually impaired people to recognize the object present in front of them. A self-reliable robot requires an object detection technique to detect objects present in front of it.

There are two aspects of object detection:

1) Object Detection can be done on still images -In this method, the captured images get processed and the object present in the image shows to the user. It is most reliable and accurate than the live tracking method.
2) Object Detection by using live camera preview – In this method, the live preview of the camera continuously shows the name of the detected objects those are coming in front of the camera.

3. TEXT DETECTION AND RECOGNITION

Text is one of the most expressive and widely used means of communications. Text present in the images, signboards, social media posts or any document has a lot of potential of conveying information about the said image or post. Retrieval of textual data from such images can be of significant importance. Retrieved text from the images or video frames can be converted to audio output which can be very useful for visually impaired people.

Text detection and recognition is a task of extracting text data in a machine-readable format from still images, scenes, video frames and live video stream. Text Detection is the process of detecting text regions in the input image and labelling them. Whereas Text recognition is to recognize the text found in the picture and making it available in a machine-readable format. This machine-readable data further can be fed to text to speech conversion method to get audio as an output. This audio output can be provided to the visually impaired person.

4. CURRENCY DETECTION AND RECOGNITION

Currency recognition is the identification of various denomination of currency. As money plays an important role in our daily lives as a medium for any type of transaction to have goods and services. It is very difficult for the visually

impaired to recognize the currency. The tactile marks on the surface of banknotes vanish eventually after continuous use. Hence, real-time detection and recognition of notes is necessary for the visually impaired.

A currency recognition system can be very useful to help blind and visually impaired people.

5. LITERATURE REVIEW

5.1 Object Detection Using Convolutional Neural Networks

This model uses Convolutional Neural Networks (CNN) to detect objects in the environment [1]. In this model, two states of Convolutional Neural Networks (CNN) are compared to find the best technique to detect the object on the basis of reliability, accuracy and speed. This model uses Single Shot Multi-Box Detector (SSD) with MobileNetV1 as the first method and Faster Region-based Convolutional Neural Network (Faster-RCNN) with InceptionV2 as the second method.

The developer's team of this model train the image set for SSD with MobileNetV1 and Faster-RCNN with InceptionV2 with the help of the TensorFlow model. TensorFlow is an open-source library designed for high-performance computation and has various applications.

As it was a comparative study between two techniques, they have shown the comparative performance of these techniques. It shows that the first technique that is the SSD with MobileNetV1 has high detection speed but low accuracy compared with Faster-RCNN with InceptionV2 that has a low speed of detection of objects but more accuracy to detect an object.

5.2 NETRA: Smart Hand Gloves Comprises Obstacle Detection, Object Identification & OCR text to speech converter for blinds

This paper describes the development of smart hand gloves which helps to detect objects and text on the objects, and convert that text and object name to speech i.e., audio format so that visually impaired persons can easily recognize the object present in front of them [2]. These hand gloves consist of Arduino Processor, Ultrasonic Sensors and software like MATLAB and Arduino Software.

This model is totally hardware-based as it involves Arduino Processor and ultrasonic sensors. These gloves are mainly useful for visually impaired persons. The working of the gloves is divided into three parts,

1) Detection of an obstacle

2) Image and text to speech

3) Identification of object

The developer team uses two software to build and implements this glove. MATLAB and Arduino Software are used to implement this glove. For image processing, MATLAB is used while for programming of Ultrasonic Sensors Arduino Software is used.

5.3 Multiple Real-time object identification using Single shot Multi-Box detection

This model implements the object detection technique using Single Shot Multi-Box Detection (SSD) with the help of TensorFlow object detection API [3]. Here, TensorFlow object detection API is used to train and deploy the object models.

In this proposed model, they have used Single Shot Multi-Box Detection (SSD) with a lightweight network model called as Mobile Net. This combination of techniques increases the accuracy of the model to detect the object present in the surroundings.

5.4 Automatic Scene Text Detection and Recognition System for Visually Impaired People

Akhilesh Panchal and the team introduces Automatic Scene Text Detection and Recognition System which will be embedded into the hardware to provide travel aid for visually impaired people [5]. The system is based on image processing, pattern matching and Computer vision techniques, which detects and recognizes text from surrounding and then gives speech output. Image processing and text detection can be very challenging due to poor quality of image, blurring, scene complexity, font variation, uneven lighting and multilingual text etc. There are various algorithms that work on the individual challenge. Authors have proposed an automatic system, combining various algorithms and techniques which results in better and accurate outputs.

In the proposed system, an input image goes through 4 stages.

1) Pre-processing
2) Processing
3) Post-processing
4) Text to speech conversion

Standard median filter (SMF), histogram equalization method, Wiener filter is used for pre-processing of input image/ scene. Binarization takes place in the second stage. In the post-processing stage syntactic approach is used to extract the characters and then the template matching

method is applied for character recognition. Finally, the recognized text is converted into audio output.

5.5 YOLO-v3 Based Currency Detection and Recognition System for Visually Impaired Persons

Rakesh Chandra Joshi and the team proposes the YOLO-v3 CNN model-based standalone banknote detection and recognition system [6]. The proposed system uses deep learning to detect banknote in real-time. The authors first created a vast dataset consisting of various images of banknotes. This dataset is divided into training, validation and testing set. YOLO network converts the problem of detection to a problem of regression. Multi-scale prediction is used for the detection of the targeted objects. Camera input is first pre-processed and then it is given to the trained model. The model trained on banknotes gives output for recognized banknotes and generate the label and bounding box. The label, text output is converted to speech.

5.6 Summary of Literature Review

Table -1: Summary of Literature Review

Title of Paper	Techniques Used	Results
1) Object Detection Using Convolutional Neural Networks [1]	Convolutional Neural Networks Single Shot Multi-Box detector (SSD) with MobileNetV1 and a Faster Region-based Convolutional Neural Network (Faster-RCNN) with InceptionV2	1)SSD with MobileNetV1 has high detection speed but low accuracy rate. 2)Faster-RCNN with InceptionV2 has high accuracy but low detection speed.
2)NETRA: Smart Hand Gloves Comprises Obstacle Detection, Object Identification & OCR text to speech converter for blinds [2]	Arduino Processor with Arduino Software and MATLAB for text to speech conversion.	The smart gloves effectively detect the object and convert text to speech. Cost of the gloves is about 30$.
3) Multiple Real-time object identification using Single shot Multi-Box detection [3]	Single Shot Multi-Box Detector with MobileNet	Performance of this model is about the 60.6% as a result.
4)Real-time Object Detection for Visually Impaired with Optimal Combination of Scores [4]	YOLO (You Only Look Once) Algorithm, SSD (Single Shot Detector)	System gives better results when YOLO and SSD integrate with the help of Proportional Conflict Resolving (PCR-6) Rules.
5) Automatic Scene Text Detection and Recognition System for Visually Impaired People [5]	Computer Vision, Pattern Matching, Standard median filter (SMF), Histogram equalization method, Wiener filter	Combining different techniques for text detection and extraction results into better and improved outcomes
6) YOLO-v3 Based Currency Detection and Recognition System for Visually Impaired Persons [6]	Deep Learning, YOLO-v3 CNN	Detection accuracy – 95.71%, Recognition accuracy – 100%

6. CONCLUSIONS

In this paper, we have proposed an intelligent, portable assistance system to assist visually impaired people. A robust application which is integration of techniques like object, text, and currency detection and recognition. The application is mobile based because today almost everyone uses smart phones. It is easy to carry and user friendly unlike other assistive systems and devices. There are many standalone assistive systems and devices which work on an

individual challenge faced by the blind people. But maintaining and carrying all those heavy and costly devices together is very hard, here the user gets overwhelmed and confused. Hence, the main focus is to incorporate all the techniques and technologies under one roof which promotes ease of use, independence and usability resulting in improved quality of life.

REFERENCES

[1] Reagan Galvez, Argel Bandala, Elmer Dadios, Ryan Rhay Vicerra, Jose Martin Maningo, Object Detection Using Convolutional Neural Networks, Proceedings of TENCON 2018 - 2018 IEEE Region 10 Conference (Jeju, Korea, 28-31 October 2018)

[2] N. K. Srivastava and Satyam Satyam Singh, NETRA: Smart Hand Gloves Comprises Obstacle Detection, Object Identification & OCR text to speech converter for blinds, 2018 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)

[3] Kanimozhi S., Gayathri G., Mala T., Multiple Real-time object identification using Single shot Multi-Box detection, Second International Conference on Computational Intelligence in Data Science (ICCIDS-2019)

[4] Siddharth Sagar Nijhawan, Shubham Bhardwaj, Aditi Kumar, Geeta Nijhawan, Real-time Object Detection for Visually Impaired with Optimal Combination of Scores, 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom)

[5] Akhilesh Panchal, Shrugal Varde, Dr. M. S. Panse, Automatic Scene Text Detection and Recognition System for Visually Impaired People, International Journal for Research in Emerging Science and Technology, Volume-3, issue-6, jun-2016

[6] Rakesh Chandra Joshi, Saumya Yadav, Malay Kishore Dutta, YOLO-v3 Based Currency Detection and Recognition System for Visually Impaired Persons, 2020 International Conference on Contemporary Computing and Applications (IC3A) Dr. A.P.J. Abdul Kalam Technical University, Lucknow. Feb 05-07, 2020