International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 03 | Mar 2020 www.irjet.net p-ISSN: 2395-0072

Smart Blind Stick using Image Processing

Bhushan Chaudhari[1], Priyanka Dorugade[2], Hrushikesh Konde[3], Rehan Kondkari[4], Prof. Deepti S. Lawand[5]

[1,2,3,4]BE Student, Dept. of Computer Engg. Pillai college of Engineering, New Panvel, Maharashtra, India
[5]Professor, Dept. of Computer Engg. Pillai college of Engineering, New Panvel, Maharashtra, India

---

Abstract - *There are millions of blind people in this world who always need some help. These visually impaired people find it challenging to travel outside their homes independently. The Smart Blind Stick that we will design helps the blind society by providing a better and more convenient means of life by moving around independently. The stick consists of ultrasonic sensors, one camera, and an earphone/speaker. Using a network of ultrasonic sensors, this system can detect obstacles around the users up to 400 cm in their direction, i.e. forward, left and right. For further processing of data these ultrasonic sensors are attached to the raspberry pi. The algorithm running in raspberry pi determines the distance from the obstacle that it informs the user by triggering the buzzer and by illustrating the environment through the camera's captured image. The camera is used for object recognition, and the image obtained through the camera will be captioned and presented to the user in the form of audio. This audio will tell what that image is and what should be done if it is an obstacle; thus, working as a virtual eye for blind people.*

Key Words: Blind Stick, Image Processing, Object Detection, Raspberry Pi, Image Captioning, Recurrent Neural Network (RNN), Convolutional Neural Network (CNN), Camera

1. INTRODUCTION

Living in a world of touch-button life, there are millions of blind people in this world who always need some help; this makes them feel low. These visually impaired people find it challenging to travel outside their homes independently. The Smart Blind Stick that we build can support the blind community by providing a better and simpler way of life by moving independently. So, the idea is simple, unlike the traditional stick, they need to carry a smart blind stick, which will help them to some extent by avoiding the obstacles around their way while walking or going out, which may be caused by accident. The stick will be having sensors and cameras to find the objects and give feedback alert messages to the user to avoid unnecessary accidents. The stick is very similar to those of the traditional, but it can save them from accidents as well as save their lives.

2. LITERATURE SURVEY

A. Visual Assistance for Blind using Image Processing:

In this paper, the authors have given idea about device which uses ultrasonic sensors to pick any object within a 180-degree path, it will send the information to the microcontroller, including the bearing and distance from the blind user. Microcontroller will issue command to camera to take a photograph of the object, then via image identification the image information will be conveyed to the user via microphone/a pair of headphones. Hence, it makes easy to manage. It helps user to complete their day to day activities easily.[1]

B. Using Artificial Tokens to Control Languages for Multilingual Image Caption Generation: In this paper, the authors have given brief about the system for Image captioning using artificial tokens to translate the computer vision into a caption in multiple languages such as English, Japanese and Chinese (Mandarin).The images are given to the Convolution Neural Network (CNN) to extract the features of the given images, and then the feature is converted to a sentence using Recurrent Neural Network (RNN). The benefit of multilingual model makes it open for multiple users.[2]

C. Object Detection for Visually Impaired Using Raspberry Pi and Ultrasonic Sensors: In this paper, the authors have described the use of ultrasonic sensors integrating it with Raspberry Pi for detecting the obstacles. The ultrasonic starts sending the signals with minimum delay and after that, the signal returns as an echo to the sensor receiver and the Raspberry Pi estimates the time it takes to get the signal back from the sensor. Using the time taken the distance of the object is calculated and converted into speech and feedback is given through speakers or headphones. The blind person will be stopped from colliding with any object.[3]

D. Show and Tell: A Neural Image Caption Generator: In this paper, the authors have planned a Natural Image Captioning System (NIC), an end-to-end neural network architecture that can produce a logical interpretation of an image instantly in plain English. NIC is dependent on a convolution neural network encoding an image into a compact representation, followed by an RNN producing a related expression.[4]

E. A Design Review of Smart Stick for the Blind Equipped with Obstacle Detection and Identification using Artificial Intelligence: The purpose of this paper is to be visual assistance of a visually impaired person. It takes speech of a user via microphone and search for it with the help of image processing algorithm and try to locate it and give instructions to the user via headphones to reach the destination.t is a wearable device which consists of a microphone, headphone and Pi camera which is connected to Raspberry Pi to process the images and video captured by the Pi camera. So, if a user speaks, the device will listen and try to process the image and find the position or place and also guide the user to reach out.[5]

F. Blind Guider – A Smart Blind Stick: The paper emphasizes object recognition more for object detection and recognition. It consists of an ultrasonic sensor node and object detection device. The device uses ultrasonic sensors for detecting wet flour and staircases. It detects the objects in front of the user using the camera with the help of object detection and machine learning algorithms to produce the feedback sound via headphones about the object detected.[6]

G. Electronic travel Aid for Visually Impaired People based on Computer Vision and Sensor Nodes using Raspberry Pi: The objective of this paper is to build a device using an ultrasonic sensor, which will give the distance from an obstacle and also vibrate a vibrator if the object is too near. Also, by using direction identification, obstacle detection, path recognition and navigation system will take the users to its destination. Navigation is performed with the help of GPS and maps.[7]

H. Artificial Vision For The Blind Using I-Cane Electronice aid for blind people: The proposed plan in this paper is to use different components in a single device to provide more convenience to the user. It as an image processing camera for detecting the obstacles and giving an alert in the form of audio. In the case of people standing near them, it also detects the traffic and tells them about the traffic via audio. It also uses different sensors like the ultrasonic sensor is used for detecting the staircases and pits at the same time temperature sensor is used for detecting fire.[8]

3. PROPOSED WORK

The proposed model attempts to design a system whereby obstacles are detected in front of the user using sensors and where obstacles are currently using a camera. The proposed system will perform image captioning on the scene captured through the camera using a Convolutional Neural Network (CNN) model and Recurrent Neural Network (RNN) model. This system will be using a Raspberry Pi, Ultrasonic Sensors, Camera Module, Buzzer, and an Audio Device. The Raspberry Pi triggers the ultrasonic sensor to start sending the burst signal. Here all the sensors start sensing at the same time, so it does not cause any delay between any sensor. If there is an obstacle detected, then the sensor sends the echo signal to the Raspberry Pi. Then the Raspberry Pi calculates the difference between the time of transmitting the burst signal and receiving the echo signal, using this time the Raspberry Pi calculates the distance between the user and the obstacle by using Distance=(Speed of Sound*Time)/2. Then it checks whether the distance of the object is less than the specified limit. If neither of the sensors is below the minimum distance, the whole process starts again. If any sensor is less than the specified threshold distance, it tracks the direction of the obstacle and uses a camera to image the direction of that point. This image is then fed into the Convolutional Neural Network (CNN) model, which extracts the features from the image and detects the objects in it. The output of the Convolutional Neural Network (CNN) model is then passed to the Recurrent Neural Network (RNN) model which is used to frame the sentence, i.e. generate the caption using the objects detected by Convolutional Neural Network (CNN) model. This caption created by the Recurrent Neural Network (RNN) model is transferred to the Text-to-Speech API, which translates the text output, i.e. the image caption, into speech, and, in the last stage, sends this message to a visually impaired person via an earphone or speaker. The process starts again and continues until the machine is shut down after the message is sent.

3.1 SYSTEM ARCHITECTURE

The system architecture is given in Figure 1. Each block is described in this Section.

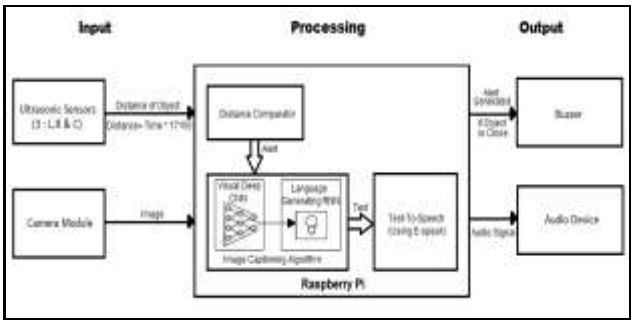

Fig. 1 Proposed System Architecture

A. Raspberry Pi

It acts as the heart of the system as it controls, processes and generates all the inputs and outputs. The Raspberry Pi used in this system will be running on Raspbian Lite OS. It receives an echo signal from the ultrasonic sensors which trigger it to take further actions, which is to check whether the obstacle is there or not. It generates an immediate alert using the buzzer. It also generates a caption for the image captured by the camera and later converts that caption into a speech that is played through an audio device.

B. Ultrasonic Sensor

The ultrasonic sensor is one of the main components of the system. It is used to detect the obstacles that come in front of the user. This system will be using three ultrasonic sensors for center, left and right directions, respectively. It sends an echo signal to the Raspberry Pi so that it can decide further actions.

C. Camera Module

The camera module used in this system can act as an eye for the visually impaired person. Each time the ultrasonic sensor detects an obstacle it captures the picture. The picture is sent to Raspberry Pi, so that it can process the image and generate the caption for that image.

D. Buzzer

The buzzer is used in this system for an immediate alert. When the ultrasonic sensor detects some obstacle, it is triggered.

E. Audio Device

The audio device is the one that conveys the image caption to the user in the form of audio. It receives an audio signal from Raspberry Pi once the caption for the image is successfully converted into an audio format using a text to speech algorithm.

4. REQUIREMENT ANALYSIS

The implementation details are given in this section.

4.1 Software

In this section, software which will be needed to achieve our aim for developing Smart Blind Stick are given.

A. Raspbian OS

Raspbian is a computer OS developed for Raspberry Pi. It is highly optimized for the Raspberry Pi line's low-performance ARM CPUs.[9]

B. Python

Python is a high level, general-purpose-programming language. It mainly focuses on object-oriented approach aims to help programmer write clear logical code for small and large-scale project.[10]

4.2 Hardware

To complete our aim of developing a smart blind stick, the following hardware will be used.

A. Raspberry Pi

It is a development board and can be considered as a single-board computer that works on the Linux operating system. It has a clock speed of 1.4 GHz and 2-4 GB of Ram with wireless LAN and Bluetooth facility.

B. Ultrasonic sensor (HC-SR04)

HC-SR04 sensor is a 4-pin module. It is an Ultrasonic sensor which works with the help of a transmitter and Receiver embedded in it. It ranges from 2cm to 4m with a measuring angle of 15□

C. PiCamera Module

Pi camera module which is widely used to take high-definition photographs, as well as high-definition video. It has a resolution of 5 megapixels.

D. Buzzer

A buzzer is an audio signal generating device. This buzzer is mostly used for making different sounds like beeps, tones and alerts. This one is petite but loud!

E. Audio Device

The term "audio device" refers to any device that attaches to a system for the purpose of playing a different kind of sound, such as music or speech. Audio devices include loudspeakers, earphones or headphones.

4.3 DATASETS

The system will need some inputs to generate a model for the proposed system; for this purpose, we are using the following datasets:

A. MS COCO

MS COCO stands for Microsoft COCO, Common Objects in Context, it is an object detection dataset with 80 number of classes, 80000 number of training images and 40000 number of validation images. The goal of this dataset is used to determine the state of object recognition by placing different questions on object detection. This can be accomplished by gathering in a natural context various complex everyday scene consisting of common objects from the environment. Object is labeled by optimization per instance to help locate objects precisely.[11]

B. Flickr30k

The Flickr30k dataset is a definitive source for representation of phrase-based images. This paper introduces Flickr30k Entities, which increases the 158k captions from which the single image is given different captions.[12]

Table 1 Dataset Used for Proposed System

Dataset Name	Images		
	Train	Valid	Test
MSCOCO	1,18,287	5,000	40,670
Flickr30k	28,000	1,000	1,000

CONCLUSION

The blind population may step in or out without any help while utilizing this smart blind stick as the stick has various functions including image processing, obstacle detection and text to speech. Furthermore, our system does not need any input from the user, as it consists of ultrasonic sensors and camera modules that sense obstacles around the user and provide feedback to the user with the assistance of an audio device connected to the module. The system continually operates so that the visually disabled can get updates on the challenges at any point on the way. Consequently, with the aid of our project, blind people could perform their tasks comfortably, so they might even substitute the regular stick with a smart stick such that it can be used in everyday life.

ACKNOWLEDGEMENT

We would love to convey our gratefulness and sincere thanks to our supervisor Prof. Deepti Lawand for the encouragement and support throughout the project. We deeply express our profound gratitude to our Head of the Department Dr. Sharvari Govilkar and our Principal Dr. Sandeep M. Joshi, for providing us with all the facilities required and allowing us to present our work

REFERENCES

[1] Deepthi Jain B, Shwetha M Thakur, K V Suresh, "Visual Assistance for Blind using Image Processing" in IEEE (April 2018)

[2] Satoshi Tsutsui, David Crandall, "Using Artificial Tokens to Control Languages for Multilingual Image Caption Generation" in arXiv:1706.06275v1 (June 2017)

[3] Ayush Wattal, Ashutosh Ojha, Manoj kumar, "Obstacle Detection for Visually Impaired Using Raspberry Pi and Ultrasonic Sensors" in National Conference on Product Design (NCPD 2016), July 2016

[4] Oriol Vinyals, Alexander Toshev, Samy Bengio, Dumitru Erhan, "Show and tell: A neural image caption generator" in arXiv:1411.4555 (June 2015)

[5] Balu N. Ilag and Yogesh Athave, "A Design review of Smart Stick for the Blind Equipped with Obstacle Detection and Identification using Artificial Intelligence" in International Journal of Computer Applications, Volume 182 – No. 49 (April 2019)

[6] Akshay Santosh Jadhav, Pooja Chandrakant Jadhav, Vrushikesh Madhukar Patil, "Blind Guider – A Smart Blind Stick" in International Journal of Engineering Research in Computer Science and Engineering, Volume 5, Issue 4 (April 2018)

[7] Ram Tirlangi and Ch. Ravi Sankar, "Electronic Travel Aid for Visually Impaired People based on Computer Vision and Sensor Nodes using Raspberry Pi" in Indian Journal of Science and Technology, Vol 9(47) (December 2016)

[8] Divyank Yarravarapu, Tina Manghnani, "Artificial Vision For The Blind Using I-Cane Electrnoice aid for blind people", The Journal of Electronics and Communication Engineering (IOSR-JECE) Volume 11, Issue 6, Ver. IV (Dec .2016)

[9] E Super Maker, "Raspbian", Internet: https://en.wikipedia.org/wiki/Raspbian, 24 October 2019, [October 25, 2019]

[10] Akeosnhaoe, "Python (programming language)", Internet: https://en.wikipedia.org/wiki/Python_(programming_language), 28 October 2019, [October 4, 2019]

[11] Tsung-Yi Lin, Michael Maire, Serge Belongie, Lubomir Bourdev, Ross Girshick, James Hays, Pietro Perona, Deva Ramanan, C. Lawrence Zitnick, Piotr Dollár "Microsoft COCO: Common Objects in Context" in arXiv:1405.0312v3 [cs.CV] (Feb 2015) [September 10, 2019]

[12] Bryan A. Plummer, Liwei Wang, Chris M. Cervantes, Juan C. Caicedo, JuliaHockenmaier, Svetlana Lazebnik, "Flickr30k Entities: Collecting Region-to-Phrase Correspondences for Richer Image-to-Sentence Models" in arXiv:1505.04870v4 [cs.CV] 19 Sep 2016 [September 10, 2019]