DESIGN AND IMPLEMENTATION OF SMART STICK AID FOR VISUALLY IMPAIRED PEOPLE USING YOLO V3 ALGORITHM.

Deepshikha Koul[1], Falguni Sawant[2], Chhayal Mahajan[3] ,

Saurabh Salade[4], Prof. Sonali Patil[5]

[1]Dr. D Y Patil Institute of Technology, Pimpri, Pune.

---***---

Abstract - *This paper proposes Raspberry Pi based Smart Stick Aid for visually-impaired people, which helps a person by detecting the obstacles using different sensors and Raspberry Pi. Using speaker, the blind person knows about the circumstances & present condition of the path where they are walking. The issues of blind people can be addressed with the intervention of technology. The methodology employs the Internet of Things (IoT) paradigm to provide a medium between the blind and the environment. It uses sensor like ultrasonic sensor and moisture sensor to detect obstacles in front of a person. It is based on ultrasonic sensor for distance measurement. Also, there is voice playback which helps to mention a direction of obstacles around a visually Impaired person by the help of sensors. We have used WiFi in stick to locate the visually Impaired people.*

Key Words: Raspberry Pi, obstacles detection, Panic button, Sensors.

1. INTRODUCTION

According to WHO (World Health Organization) there are around eighteen million blind individuals in the world. Visual defect seriously affects the quality of life of these population and limit daily activities particularly using cash for transactions. Due To impairment, they're taken advantage of and might be cheated by anyone. Hence, it is necessary to design a system which can guide them through various problems of their life. Hence, we are using object recognition methods using image processing, audio Support that are all embedded in one that will help blind people to recognize obstacles in front of them without any external help.

The main characteristics for the stick to be useful to every visually impaired person is for it to be fast and economically feasible. The obstacles in the indoors hinder the way of the blind. This aid alerts the user about various obstacles through a vocal sound from a speaker on the stick. This stick can detect wet surfaces and raise a vibratory response to the user.

The blind stick is integrated with ultrasonic sensor along with moisture sensor . We are using ultrasonic sensor to detect obstacles ahead using ultrasonic waves, the sensor passes this information to the raspberry pi. The microprocessor then does processing of this data and calculates if the obstacle is close nough. If obstacle is not present close the circuit doesnot do anything,but if the obstacle is close the microprocessor sends a signal which informs the blind person in sound format.

2. LITERATURE SURVEY

Ross Girshick[1] propose an easy and scalable detection algorithm that improves mean average precision (mAP) by more than 30% relative to the previous best result on VOC 2012—achieving a mAP of 53.3%. Our method combines two key insights:(A) one can apply high-capacity convolutional neural networks (CNNs) to bottom-up region proposals in order to localize and segment objects and (B) when labeled training data is scarce, supervised pre-training for an auxiliary task, followed by domain-specific fine-tuning, yields a significant performance boost. Since we combine region proposals with CNNs, we call our method R-CNN: Regions with CNN features.

Xiang Wang[2] proposed a novel edge preserving and multi-scale contextual neural network for salient object detection. The proposed framework is aiming to address two limits of the existing CNN based methods. First, region-based CNN methods lack sufficient context to accurately locate salient objects since they deal with each region independently. Second, pixel-based CNN methods suffer from blurry boundaries due to the presence of convolutional and pooling layers.We first propose an end-to-end edge-preserved neural network based on Fast R-CNN framework (named RegionNet) to efficiently generate saliency map with sharp object boundaries. The proposed framework achieves both clear detection boundary and multiscale contextual robustness simultaneously for the first time, and thus achieves an optimized performance. Experiments on six RGB and two RGB-D benchmark datasets demonstrate that the proposed method achieves state-of-the-art performance.

Rithika H1[3] The main problem in communication is language bias between the communicators. This device basically can be used by people who do not know English and want it to be translated to their native language. The novelty component of this research work is the speech output which is available in 53 different languages translated from English. This paper is based on a prototype which helps user to hear the contents of the text images in the desired language. It involves extraction of text from the image and converting the text to translated speech in the user desired language. This is done with Raspberry Pi and a camera module by using the concepts of Tesseract OCR [optical character recognition] engine, Google Speech API [application program interface] which is the Text to speech engine and the Microsoft translator. This relieves the travelers as they can use this device to hear the English text in their own desired language. It can also be used by the visually impaired. This device helps users to hear the images being read in their desired language

3. METHODOLOGY

- **YOLO V3**

The YOLOv3 method considers object detection as a regression problem. It directly predicts convolution neural Electronics 2020, 9, 537 3 of 11 network. It completely eliminates region proposal generation and feature resampling, and encapsulates all stages in a single network in order to form a true end-to-end detection system.

The YOLOv3 method divides the input image into S × S small grid cells. If the center of an object falls into a grid cell, the grid cell is responsible for detecting the object. Each grid cell predicts the position information of B bounding boxes and computes the objectness scores corresponding to these bounding boxes. Each objectness score can be obtained as follows:

$$C_i^j = P_{i,j}(Object) * IOU_{pred}^{truth}$$

whereby C j i is the objectness score of the jth bounding box in the ith grid cell. Pi,j(Object) is merely a function of the object. The IOUtruth pred represents the intersection over union (IOU) between the predicted box and ground truth box. The YOLOv3 method uses binary cross-entropy of predicted objectness scores and truth objectness scores as one part of loss function.

$$E_1 = \sum_{i=0}^{S^2} \sum_{j=0}^{B} W_{ij}^{obj} [\hat{C}_i^j \log(C_i^j) - (1 - \hat{C}_i^j) \log(1 - C_i^j)]$$

It can be expressed as follows: whereby S 2 is the number of grid cells of the image, and B is the number of bounding boxes. The C j i and Cˆ j i are the predicted abjectness score and truth abjectness score, respectively.

The position of each bounding box is based on four predictions:tx,ty, tw,th , on the assumption that (cx,

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 05 | May - 2022 Impact Factor: 7.185 ISSN: 2582-3930

cy) is the offset of the grid cell from the top left corner of the image. The center position of final predicted bounding boxes is offset from the top left corner of the image by (bx, by). Those are computed as follows:

$$b_x = \sigma(t_x) + c_x$$
$$b_y = \sigma(t_y) + c_y$$

whereby σ() is a sigmoid function. The width and height of the predicted bounding box are calculated thus:

$$b_w = p_w e^{t_w}$$
$$b_h = p_h e^{t_h}$$

whereby pw, ph are the width and height of the bounding box prior. They are obtained by dimensional clustering.

$$E_2 = \sum_{i=0}^{S^2}\sum_{j=0}^{B} W_{ij}^{obj}[(\sigma(t_x)_i^j - \sigma(\hat{t}_x)_i^j)^2 + (\sigma(t_y)_i^j - \sigma(\hat{t}_y)_i^j)^2] + \sum_{i=0}^{S^2}\sum_{j=0}^{B} W_{ij}^{obj}[((t_w)_i^j - (\hat{t}_w)_i^j)^2 + ((t_h)_i^j - (\hat{t}_h)_i^j)^2]$$

The ground truth box consists of four parameters (gx, gy, gw and gh), which correspond to the predicted parameters bx, by, tw and th , respectively. Based on (3) and (4), the truth values of ˆtx, ˆty, ˆtw and ˆth .

$$\sigma(\hat{t}_x) = g_x - c_x$$
$$\sigma(\hat{t}_y) = g_y - c_y$$
$$\hat{t}_w = \log(g_w / p_w)$$
$$\hat{t}_h = \log(g_h / p_h)$$

The YOLO V3 method the square error of coordinate prediction as one part of lost function.

4.IMPLEMENTATION:-

The proposed system provides improvements to the existing system design. It tries to make the existing system more efficient, convenient and user- friendly. The implementation of the proposed design of the stick requires following hardware components:

1. Ultrasonic sensors
2. Moisture sensor
3. Raspberry Pi

1. Ultrasonic Sensor:

In the proposed system we use an ultrasonic sensor. ultrasonic sensor detects the obstacles in a range of about 50 - 100 cm. The distance of the obstacle is determined based on the delay between the emission of sound and the arrival of an echo. The distance of the obstacle is obtained using:

Distance = (time*speed of sound in air)/2

Where, time is the time taken by ultrasonic waves to travel and Speed of sound in air is 340m/s.

We divide the product of time and speed by 2 because the time is the total time it took to reach the obstacle and return.

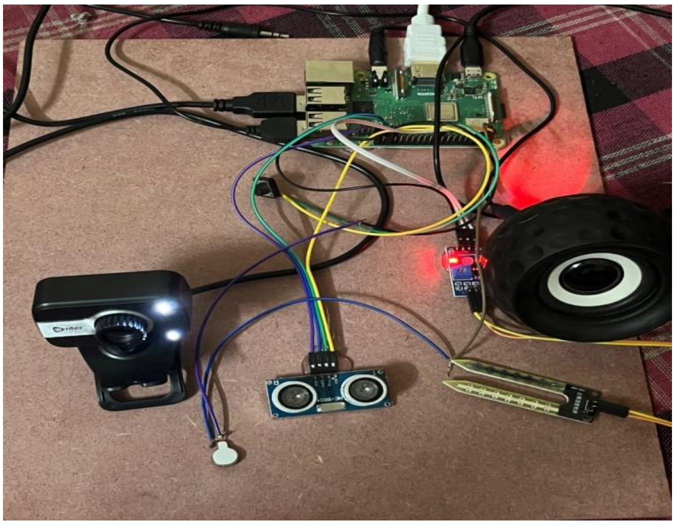

Fig 1: Working of ultrasonic sensor.

2. Moisture Sensor:

Water sensor are used to detect water presence. Our objective is to detect water and send alert to the person . So we used a costless alternative. the moisture sensor was put in the forward leg of stick to show the degree of moisture of earth soil to prevent immersing in the mud. In addition, the four legs stick must be used because all the sensors must read the values properly by making the stick in 90° angle on the earth. Beep sound is produced by the speaker to warn the blind person about the puddle.

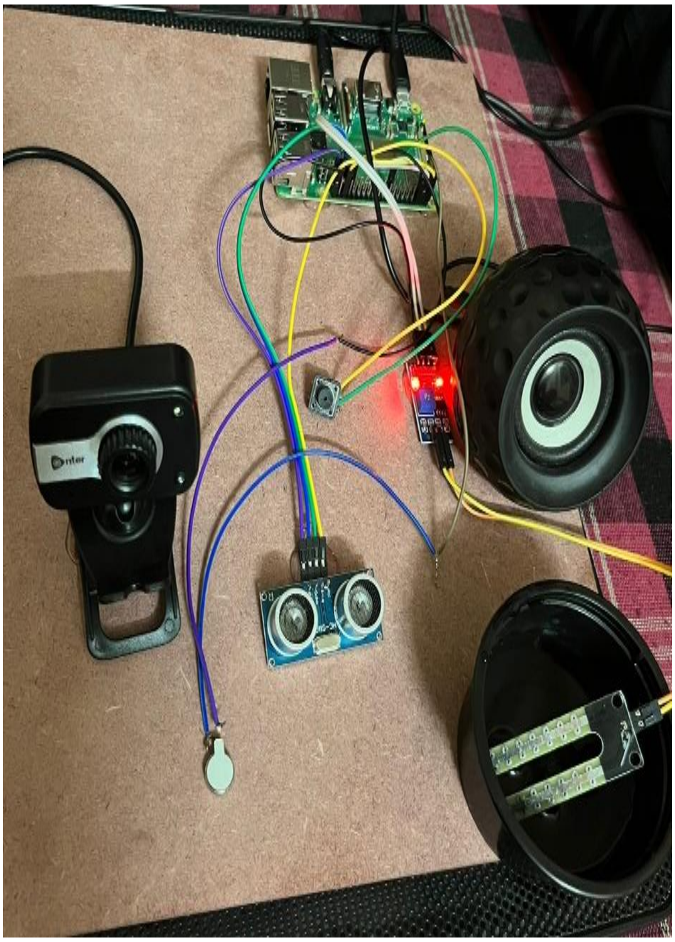

Fig 2: Working of moisture sensor

3. Raspberry Pi:

The Raspberry Pi 3 is featured with a **quad-core 64-bit Broadcom BCM2837 ARM Cortex-A53 SoC processor running at 1.2 GHz**, that makes it 50% more powerful than the Pi 2. New Raspberry Pi 3 can be used for office applications and web browsing.

Fig 3: Raspberry Pi 3B+

5. RESULTS:

Testing of moisture sensor:-

Sr.	Test case description	Test data	Expected result	Actual result	Pass/fail
1.	To detect water/moisture from sensor.	Water in bowl	Water detected.	Water detected and message " high,0",displayed on screen and conveyed through speaker.	Pass.

Table 1:Moisture sensor test result.

Testing of Ultrasonic sensor:-

Sr no	Test case description	Test data Distance	Expected Result	Actual Result	Pass/ Fail
1.	To measure distance of obstacle from sensor	10.0 cm	Obstacle Is detected	Obstacle Is detected	Pass
2.	To measure distance of obstacle from sensor	12.5 cm	Obstacle Is detected	Obstacle Is detected	Pass
3.	To measure distance of obstacle from sensor	14.6 cm	Obstacle Is detected	Obstacle Is detected	Pass
4.	To measure distance of obstacle from sensor	3.7 cm	Obstacle is detected	Obstacle is detected	Pass
5.	To measure distance of obstacle from sensor	502.0cm	Obstacle is not detected	Obstacle is not detected	Pass

Table 2: Ultrasonic sensor Test Result.

6. CONCLUSION:-

It is worth mentioning at this point that the aim of this study which is the design and implementation of a smart blind stick aid achieved. The main objective is to provide an aid to visually impaired people to navigate from one place to another using real time video streaming technology and conventional sensor-based technology. It can be improved by using VLSI to design to design PCB unit.

7. REFERENCES:-

[1] Ross Girshick1 Jeff Donahue1,2 Trevor Darrell1,2 Jitendra Malik1 1UC Berkeley and 2 ICSI Rich feature hierarchies for accurate object detection and semantic segmentation

[2] Xiang Wang , Huimin Ma , Member IEEE, Xiaozhi Chen, and Shaodi You Edge Preserving and Multi-Scale Contextual Neural Network for Salient Object Detection.

[3] Rithika.H1 , B. Nithya santhoshi Image Text To Speech Conversion In The Desired Language By Translating With Raspberry Pi

[4] Prof. Vaibhav V. Mainkar Raspberry pi based Intelligent Reader for Visually Impaired Persons

[5] Dr. I S Akila A Text Reader for the Visually Impaired using Raspberry Pi