International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

Smart Blind Kit for Visually Impaired People

Prof. Poonam Pawar[1], Rajeshwari Bhirud[2], Karan Shinde[3], Snehal Gawade[4], Reshma Gholap[5]

[1]*Lecturer, Department of Information Technology, Sinhgad Academy of Engineering, Pune, India*
[2,3,4,5] *Students, Department of Information Technology, Sinhgad Academy of Engineering, Pune, India*

ABSTRACT - *Visually impaired person/people have difficulty to interact and feel their environment. It is challenging for physical movement for visually impaired person, because it can become tricky to distinguish where he is, and how to get where he wants to go from one place to another. Researchers have spent the decades to develop an intelligent and smart stick to assist and a alert system for to alert visually impaired persons from obstacles and give information about their location. In this paper we will discuss about creating a smart kit system for assisting blind people. The smart kit comes as a proposed solution to enable visually impaired people to find difficulties in detecting obstacles and dangers in front of them during walking and to identify the world around. The system is designed to act like an artificial vision and alarm unit. The alarm system notifies the visually impaired person through beeps which is assigned to a particular action. The designed system consist of hardware and software part; hardware detects the slippery area, potholes on the road or path where the user is walking and the objects that comes in contact with the stick through ultrasonic sensor, infrared sensor and water sensor; software uses various algorithms to processes images for face recognition, to detect the text through image processing and notify the user through speech.*

***Key Words*: Visually impaired, face recognition, text-to-speech, text through image processing, text to speech.**

1. INTRODUCTION

Smart Blind Kit is an innovative kit consisting of smart camera module and smart blind stick designed for visually disabled people. We here propose an advanced blind kit that allows visually challenged people to navigate with ease, recognize known people and read text using advanced technology. The blind kit is integrated with ultrasonic sensor along for sensing obstacle, camera to detect and recognize people as well as capture images of text to be converted into speech.

Our proposed project first uses ultrasonic sensors to detect obstacles ahead using ultrasonic waves. On sensing obstacles the sensor passes this data to the micro-controller. The micro-controller then processes this data and calculates if the obstacle is close enough. If the obstacle is not that close the circuit does nothing. If the obstacle is close the micro-controller sends a signal to sound. It also detects and sounds of a different objects if it detects and alerts the blind.

The system has buzzer which states the detected obstacle that is 1 beep for water, 2 beeps for object and 3 beeps for pothole. It consist of camera used to provide real time face recognition and alert the individual if any known person is present nearby, on button click a photo gets captured from the camera to further process the image extract text and provide voice regarding the textual contents of the image.

1.1 PROJECT PURPOSE

The purpose of this project is that blind person can used smart phone easily by using different voice commands and by using kit which consists of stick and goggles is integrated with smart phone will detect if any obstacle in user path, notify if there is known person nearby or in front, Convert text from an image to speech to stimulate reading ability. In this system blind person can speak out commands and smart application will response them in sound. Smart phone is integrated with kit so user can find out obstacle and can avoid any problem which may happen with itself.

1.2 PROJECT SCOPE

- The Smart Blind Kit will help the blind person by providing more convenient means of life.
- Proposed system help blind to move safely and detect obstacles in their path, get notified if there are known people around him and stimulate reading ability.
- In case obstacles/known people were detected it will send a voice messages to the user.
- When picture is clicked the text from image is extracted and converted to speech so blind person can hear it.

1.3 WORKING PRINCIPLE

The purpose of this system is to develop an application for blind person. Main aim is to sense the obstacle in their path and detect it, detect known people nearby and notify and read text and provide speech so that visually impaired people travel safely without depending on others.

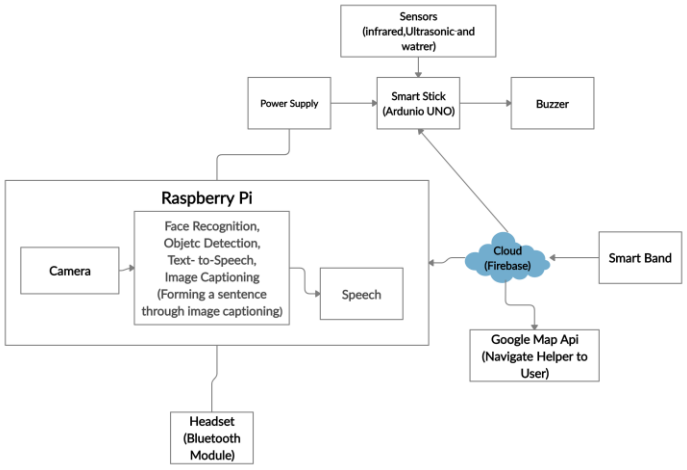

Figure 1: Block Diagram of Proposed System

The proposed system helps user notify the person through face recognition algorithm, to detect object in front of the user. Also, notifies user through beeps if the floor is slippery, pot holes in his paths and object in the path while walking.

2. SYSTEM REQUIREMENT

A. Software Requirements:-

- Arduino Ide
- Python

B. Hardware Requirement:-

- Arduino Microcontroller
- Ultrasonic Sensor
- Buzzer
- Bluetooth
- Power Supply
- GPS
- Camera Module
- Raspberry pi
- Infrared sensor
- Water sensor

3. IMPLEMENTATION

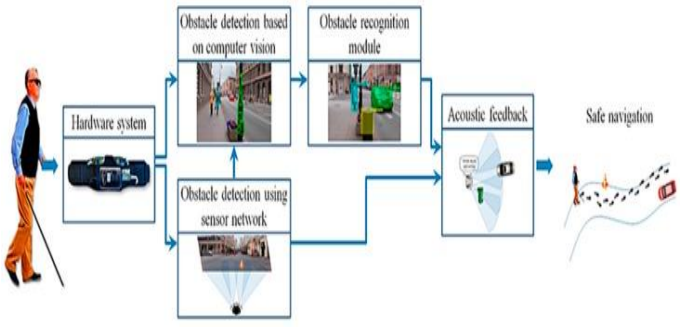

Figure 2: Implementation of proposed system

The above diagram tells the implementation of our system as follows:

- Ultrasonic sensor detects the obstacle near to it.
- Indoor location not easy to find.
- Text in image should be clear for good results.
- Internet connectivity required.
- The system should have data of known person.

The above operations are integrated in one single smart band. Smart Band contains 6 buttons which on being pressed will perform certain operation specified. Each particular button performs a single operation. Following information will provide more detailed description of buttons and their respective operations:

1) On pressing this button, you will get connected to the smart stick and obstacle detection can be carried out using the stick. Our proposed project first uses ultrasonic sensor to detect obstacles ahead using ultrasonic waves. On sensing obstacles the sensor passes this data to the micro-controller. The micro-controller then processes this data and sends a signal to sound. It calculates if the obstacle is close enough. If the obstacle is not that close the circuit does nothing. If the obstacle is close then the micro-controller produce sounds for different objects detected and alerts the blind. The system has buzzer which states the detected obstacle that is 1 beep for water, 2 beeps for object and 3 beeps for pothole.

2) On pressing this button, you will be able to capture the image of the person who is in front of you and recognize whether he/she is a known person and get notified accordingly. Here face recognition is used.

CNN Algorithm For face Detection:

1: Find all faces in current image of video stream - we train our model to recognize faces using Histogram of Oriented Gradients (HOG algorithm) we convert our image to gray-scale, Then we'll look at every single pixel in our image one at a time. We want to look at the pixels for every single pixel that directly surrounds it. We figure out how dark the current pixel is compared to the pixels directly surrounding it. Then we draw an arrow showing in which direction the image is getting darker: process for every single pixel in the image, you end up with every pixel being replaced by an arrow. These arrows are called gradients and they show the flow from light to dark across the entire image we'll break up the image into small squares of 16 x 16 pixels each. We count up how many gradients point in each major direction in each square. Then replace that square in the image with the arrow directions that were the strongest. The end result is turn the original image into a very simple representation that captures the basic structure of a face in a simple way. To find faces in this HOG image, all we have to do is find the part of our image that looks the most similar to a known HOG pattern that was extracted from a bunch of other training faces:

2 : Posing and Projecting Faces - to deal with the problem that faces turned different directions look totally different to a computer, try to warp each picture so that the eyes and lips are always in the

sample place in the image we use algorithm face landmark estimation to do this. we come up with 68 specific points (called landmarks) that exist on every face Then will train our model on a algorithm to be able to find these 68 specific points on any face We are using basic image transformations like rotation and scale that preserve parallel lines (called affine transformations), so no matter how the face is turned, we are able to center the eyes and mouth are in roughly the same position in the image.

3: Encoding Faces - we train a Deep Convolution Neural Network to generate 128 measurements for each face. The training process works by looking at 3 face images at a time: -Load a training face image of a known person -Load another picture of the same known person -Load a picture of a totally different person i.e we perform triplet training. After repeating this step millions of times for millions of images of thousands of different people, the neural network learns to reliably generate 128 measurements for each person. Any ten different pictures of the same person should give roughly the same measurements. So all we need to do ourselves is run our face images through their pre-trained network to get the 128 measurements for each face.

4: Finding the person's name from the encoding - We train a classifier that can take in the measurements from a new test image and tell which known person is the closest match. Running this classifier takes milliseconds. The result of the classifier is the name of the person!

3) On pressing this button, you can click the picture of the newspaper, article and get text to speech conversion of those words present in the captured image which increases the reading ability.

Algorithm for Image-text-Speech using Python-tessereact OCR tool

1: Capture Image

2: Convert to gray scale

3: Extract text from image: We use Python-tessereact which is an optical character recognition (OCR) tool for python. It will recognize text and "read" the text embedded in images.

4: Speak the extracted text: We use e-speak tool for text-to-speech.

4) On pressing this button, you can capture the image in front of you and get it captioned accordingly which makes it easy for you to understand your surroundings clearly.

CNN Algorithm for image captioning:

Image Captioning is the process of generating textual description of an image. It uses both Natural Language Processing and Computer Vision to generate the captions. The data set will be in the form [image → captions]. The data set consists of input images and their corresponding output captions.

ENCODER: The Convolutional Neural Network(CNN) can be thought of as an encoder. The input image is given to CNN to extract the features. The last hidden state of the CNN is connected to the Decoder.

DECODER: The Decoder is a Recurrent Neural Network(RNN) which does language modelling up to the word level. The first time step receives the encoded output from the encoder and also the vector.

TRAINING: The output from the last hidden state of the CNN(Encoder) is given to the first time step of the decoder. We set x1 = vector and the desired label y1 = first word in the sequence. Analogously, we set x2 =word vector of the first word and expect the network to predict the second word. Finally, on the last step, xT = last word, the target label yT = token. During training, the correct input is given to the decoder at every time-step, even if the decoder made a mistake before.

TESTING: The image representation is provided to the first time step of the decoder. Set x1 = vector and compute the distribution over the first word y1. We sample a word from the distribution (or pick the argmax), set its embedding vector as x2, and repeat this process until the token is generated. During Testing, the output of the decoder at time t is fed back and becomes the input of the decoder at time t+1.

5) On pressing this button, the blind person can detect objects and the same can be conveyed to him in a speech format.

YOLO object detector algorithm for object detection:

We use YOLO Object Detector which was first introduced in 2015 and is capable of super real-time object detection, obtaining 45 FPS on a GPU. YOLO is able to achieve such a large number of object detections by performing joint training for both object detection and classification. Using joint training the authors trained YOLO9000 simultaneously on both the ImageNet classification dataset and COCO detection dataset.

The result is a YOLO model, called YOLO9000, that can predict detections for object classes that don't have labeled detection data. We use OpenCV's DNN function cv2.dnn.readNetFromDarknet to load this model from disk We initialize list consists of:

1 - Boxes: Our bounding boxes around the object.

2 - Confidences: The confidence value that YOLO assigns to an object. Lower confidence values indicate that the object might not be what the network thinks it is. Remember from our command line arguments above that we'll filter out objects that don't meet the 0.5 threshold.

3 - ClassIDs: The detected object's class label. We populate layerOutputs by: 1 - Loop over each of the layerOutputs 2 - Loop over each --detection -- in -- output 3 - Extract the -- classID -- confidence

4 - Use the confidence to filter out weak detections. After filtering unwanted detections:

1 - Scale bounding box coordinates so we can display them properly on our original image

2 - Extract coordinates and dimensions of the bounding box. YOLO returns bounding box coordinates in the form: (centerX, centerY, width, and height)

3 - Use this information to derive the top-left (x, y)-coordinates of the bounding box

4 -Update the boxes, confidences, and classIDs lists Then we apply Non-Maxima Suppression: Applying non-maxima suppression suppresses significantly overlapping bounding boxes, keeping only the most confident ones. NMS also ensures that we do not have any redundant or extraneous bounding boxes. (We use OpenCV's built-in DNN module implementation of NMS, we perform non-maxima suppression.

All that is required is that we submit our bounding boxes, confidences, as well as both our confidence threshold and NMS threshold.) Speaking name of the detected object: Assuming at least one detection exists; we proceed to loop over idxs determined by non-maxima suppression. Then, we simply draw the bounding box and text on image using our random class colors. Finally, we display our resulting image and speak name of the detected object. We again make use of espeak tool for text to speech.

6) On pressing this button, the blind person can share his/her live location to its registered contact numbers or relatives which can then use Google Map API to navigate and reach the blind person to help them.

The below figure Google Map API implemented to provide route assistance for blind person's relatives to reach out to him in case of any problem and help him out,

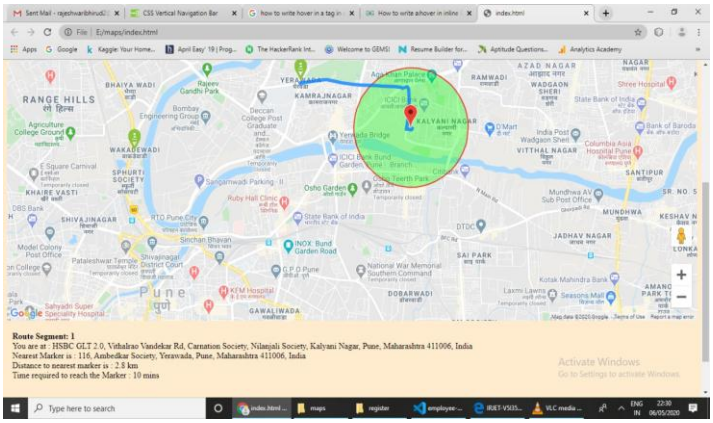

Figure 3: Output of Google API Implementation

4. TECHNOLOGY USED:

1) gTTS

Proposed system uses gTTS (Google Text-to-Speech), which is a Python library and CLI tool to interface with Google Translates text-to-speech API. Write spoken mp3 data to a file, a file-like object (byte-string) for further audio manipulation, or stdout. Or simply pre-generate Google Translate TTS request URLs to feed to an external program.

2) Pyttsx

Proposed system makes use of pyttsx (Python Text-to-speech) which is a cross-platform text to speech library which is platform independent. The major advantage of using this library for text-to-speech conversion is that it works offline.

pyttsx is a Python package that supports common text-to-speech engines on Mac OS X, Windows, and Linux.

3) OCR

Proposed system uses Optical Character Recognition (OCR) which is a system that provides a full alphanumeric character recognition on an image. The system allows extracting text from an image, to convert it later into an editable file.

Some open source libraries for OCR such as Tesseract, Gocr, JavaOCR, and Ocrad. Some of most popular on the list is Tesseract, Tesseract is an open source OCR engine combined with the processing library of Leptonic image can read a wide variety of image formats and turn them into text. As well, it has good support from the community, it has wrappers for different languages and it has good results among others.

4) pytesseract

Proposed system uses Python-tesseract which is an optical character recognition (OCR) tool for python. It will recognize text and "read" the text embedded in images.

Python-tesseract is a wrapper for Google's Tesseract-OCR Engine, it is also useful as a stand-alone invocation script to tesseract, as it can read all image types supported by the Pillow and Leptonic imaging libraries, including jpeg, png, gif, bmp, tiff, and others. Additionally, if we use it as a script, Python-tesseract will print the recognized text instead of writing it to a file.

5) OpenCV

Proposed system uses OpenCV (Open Source Computer Vision Library) which is a library of programming functions mainly aimed at real-time computer vision. It is library used for Image Processing and mainly used to do all the operation related to Images.

OpenCV application areas include:

- Facial recognition system
- Gesture recognition

- Human–computer interaction (HCI)
- Mobile robotics
- Motion understanding
- Object identification
- Motion tracking
- Augmented reality

6) CNN

Proposed System uses Convolutional Neural Network (ConvNet/CNN) which is a Deep Learning algorithm which can take in an input image and assign importance to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms, while in primitive methods filters are hand-engineered with enough training Convent's have the ability to learn these filters/characteristics.

7) YOLO: Real Object Detection algorithm

Proposed system makes use of You only look once (YOLO) which is a system for detecting objects on the Pascal VOC 2012 dataset. It can detect the 20 Pascal object classes:
•person
•bird, cat, cow, dog, horse, sheep
•areophane, bicycle, boat, bus, car, motorbike, train
•bottle, chair, dining table, potted plant, sofa, tv/monitor

8) Image Captioning

Proposed system uses Image Captioning which is the process of generating textual description of an image and it uses both Natural Language Processing and Computer Vision to generate the captions. The data set will be in the form [image → captions]. The data set consists of input images with their corresponding output captions. Image captioning is an application of one to many RNN's.

A recurrent neural network (RNN) is a class of artificial neural networks where connections between nodes form a directed graph along a temporal sequence and this allows it to exhibit temporal dynamic behavior. Derived from feedforward neural networks, RNNs can use their internal state (memory) to process variable length sequences of inputs. This makes them applicable to tasks such as unsegmented, connected handwriting recognition or speech recognition.

9) Google Map API

The Google Maps Platform is a set of APIs and SDKs that allows developers to embed Google Maps into mobile apps and web pages, or to retrieve data from Google Maps.

5. CONCLUSIONS

The main objective behind project is to design one system which becomes helpful to the visually impaired people by providing a kit, which detect obstacle in front of the person or recognize a person in front of them, due to these feature, one smart equipment for visually impaired people for walking on the road or surrounding environment through it we can give them some sense of vision. Tracking of the user's current location is also implemented for the security or safety purpose of the blind person. The purpose of this system is to develop an application for blind person. Main aim is to sense the obstacle in their path and detect it, detect known people nearby and notify and read text and provide speech so that visually impaired people travel safely without depending on others. Providing information about battery status, current time, calling and text messaging via audio signal(speech).

ACKNOWLEDGEMENT

I would like to express my deepest appreciation to all those who provided me the possibility to complete this report. A special gratitude I give to our project guide, [Mrs. P. Y. Pawar], who helped me in deciding my topic and contents and whose contribution in stimulating suggestions and encouragement, helped me to coordinate my research work especially in writing this report.
I would also thank Mr. Abhay Adapanawar [HOD of Information Technology Department] for providing healthy environment and facilities in the department.

Thanks to all the colleagues for their extended support and valuable guidance. I would like to be grateful to all my friends for their consistent support, help and guidance.

I convey my heartfelt affection to all those people who helped and supported me during the course for completion of my project report.

REFERENCES

[1] Ashraf Anwar, Sultan Aljahdali –"A Smart Stick for Assisting Blind People" IOSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661,p-ISSN: 2278-8727, Volume 19, Issue 3, Ver. II (May.-June. 2017), PP 86-90

[2] Abhishek Gupta, Abhishek Yadav, Bhaskar Sharma, Amrit Pal-"Smart Guiding System for Blind: Obstacle Detection and Real-time Assistance via GPS" International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 03 | Mar-2018 p-ISSN: 2395-0072

[3] Mrs. M. Deepthi, P. Sowmya, N. Naga Mounika, K. Pavani Sai, N. Sripriya, N. Balasaida-"IOT Based Smart Stick with Voice Module" International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor : 6.887 Volume 6 Issue II

[4] Asha G. Hagargund, Sharsha Vanria Thota, Mitadru Bera, Eram Fatima Shaik "Image to Speech Conversion for Visually Impaired" International Journal of Latest Research in Engineering and Technology (IJLRET) ISSN: 2454-5031 Volume 03 - Issue 06 || June 2017 || PP. 09-15

[5] R.Dhanuja1, F.Farhana, G.Savitha "SMART BLIND STICK USING ARDUINO" International Research Journal of Engineering and Technology (IRJET) Volume: 05 Issue: 03 | Mar-2018

[6] Bogdan MOCANU1, Ruxandra TAPU1, Titus ZAHARIA, "Design of a CNN Face Recognition System" Dedicated to Blinds Bogdan MOCANU1,2, Ruxandra TAPU1,2, Titus ZAHARIA1 , Member, IEEE 1Institute Mines- Télécom /Télécom SudParis, ARTEMIS

Department, UMR CNRS 8145 MAP5 and 5157 SAMOVAR, Evry, France

[7] Mr.Rajesh M., Ms. Bindhu K. Rajan Ajay Roy, Almaria Thomas K, "TEXT RECOGNITION AND FACE DETECTION AID FOR VISUALLY IMPAIRED PERSON USING RASPBERRY PI" 2017 International Conference on circuits Power and Computing Technologies [ICCPCT]

[8] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5375851/#sec3-sensors-17-00565

[9] https://en.wikipedia.org/wiki/Canny_edge_detector