https://doi.org/10.46335/IJIES.2022.7.5.9 **e-ISSN: 2456-3463**
Vol. 7, No. 5, 2022, PP. 49 -55

International Journal of Innovations in Engineering and Science, www.ijies.net

Application for the Voice Assistance of the Blind

E. Teja[1], Dr T Manikumar[2], Dr N Naveen Kumar[3]

[1] PG Research Scholar, [2] Assistant Professor, [3] Associate Professor
[1,2,3] Department of Computer Applications
Madanapalle Institute of Technology and Science, Madanapalle, Andhra Pradesh, India.

tejaevuri1@gmail.com

***Received on**: 15 April, 2022* ***Revised on**: 22 May, 2022,* ***Published on**: 24May, 2022*

Abstract - *This application enables consumers or users to gain knowledge about Obstacle detection and warning can improve the mobility as well as the safety of visually impaired people especially in unfamiliar environments. For this, firstly, obstacles are detected and localized and then the information of the obstacles will be sent to the visually impaired people by using different modalities such as voice. In this project, we present an assistive system for visually impaired people based on Tensor Flow object detection model and Google Speech's model.[1] This system consists of two main components: environment information acquisition and analysis and information representation The first component aims at capturing the environment by using TesnorFlow object detection model and analyzing it in order to detect the predefined obstacles for visually impaired people.[3]While the second component to tries to represent obstacles information under the form of speech to visually impaired person[1]*

Keywords*—Text-to-speech, Optical Character Recognition, Web Browsing. Navigation, Obstacle detection, handicapped aids, Vision defects, Visually Impaired, Android, Blind People.*

I -INTRODUCTION

Visual impairment can limit people's ability to perform everyday tasks and can affect their quality of life and ability to interact with the surrounding world. Blindness, the most severe form of visual impairment, can reduce people's ability to perform daily tasks, and move about unaided. Good quality rehabilitation allows people with different degrees of visual impairment to fully profit from life, achieve their goals and be active and productive in today's society. Ample efforts have been made to aid the blind by innovating and improving technologies. According to "The emerging ethics of human centric GPS tracking and monitoring", the main factors focused while tracking a person are privacy, accuracy and accessibility. By the introduction of a basic sensor, the provisions to the aid of blind people have remarkably increased. Many universities and companies like IBM have given a part of their focus to the development of aids for the blind people. Some of the popular ones are smart canes and obstacle sensors. Recent development includes self-driving cars and smart glasses. In the last 30 years, various other strides that have been developed are the text or speech software's and Smartphone apps.[4] Moreover, the systems that are being developed on robotics and artificial intelligence will be very advantageous to them as well. [5] One of the major factors in developing these technical aids is the compatibility with the user. He should not have trouble getting acquainted with the product. The features of the product should not be too difficult to use. Even the notification systems used to alert the blind should be comfortable and reliable. Another major factor is the cost of such products. Since they are already paying for treatments or other nursing cares, the price of the product should be in the range that is reasonable. Other feature of these products should be the durability. The users might not be able to charge the system.[2]
So, appropriate measure should be taken for it.

https://doi.org/10.46335/IJIES.2022.7.5.9 **e-ISSN: 2456-3463**
Vol. 7, No. 5, 2022, PP. 49 -55

International Journal of Innovations in Engineering and Science, www.ijies.net

In this system, an Android application will be developed for object detection, which will use the phone's camera to detect the objects in front of the user. The application will detect the objects using Tensor flow's object detection API and provide an audio message about the name and the location of the object to the user. The location includes the direction and the distance of the object with respect to the user. The audio message will be provided to the visually impaired user with the help of an audio device such as headphones or the phone's speaker. The system does not need any external camera, as it will be using the phone's camera to carry out the above-mentioned tasks. The purpose of this paper is to highlight how object detection, a technique based on computer vision, can help the visually impaired people to support independent travel by presenting an overview of application for object detection for visually impaired people, its modalities and their functionalities.[6]

II- LITERATURE REVIEW

Patrick Roth , Lori Stefano Petrucci , André Assimacopoulos, and Thierry Pun proposed an audio-haptic feedback mechanism for helping the blind visualize what is on the screen. of the Association of the Slovenian Blind and Visually Impaired Persons.

Simon Dobrišek (2003) published a survey paper that covers the evolution of architectures of the information retrieval systems used in building systems which aid the blind and the visually-impaired people.

Dr.Hersh Johnson's "Assistive Technology for Visually Impaired and Blind People" covers history of assistive technology for the blind along with block diagrams and approaches used for their implementation.[1]

Existing paper works and solutions related to GPS-GSM, and Obstacle detection reveal that, so far, all these technologies have been implemented individually but not integrated for the cause of the blind.

In sonar technology, sound propagations are used to navigate, communicate with or detect objects on or under the surface of the water, such as other vessels. This technology can be mainly divided into two types- passive and active. In passive sonar, when the vessels make sound, they listen to it. Active sonar is emitting pulses of sounds and listening for echoes. This is the mechanism we will be working with. This is primarily used for acoustic location tracking.

S. Maeyama, A. Ohya, and S. Yuta Positioning by tree detection sensor and dead reckoning for outdoor navigation of a mobile robot It proposes positioning method for outdoor navigation of a mobile robot, by fusing dead reckoning and the tree detection sensor which consists of sonar and vision. A street lined with trees is assumed to be the mobile robot's outdoor work space. [2]

B. Andò A smart multisensory approach to assist blind people in specific urban navigation tasks, Visual impairment is one of the major problems that people face. They need human assistance to do many activities in daily life. This paper describes the designation of the smart assistance device to help blind people by using multiple sensors. [3]

III- ARCHITECTURE AND WORKDONE

The flow for the project is given below(Fig1.1):

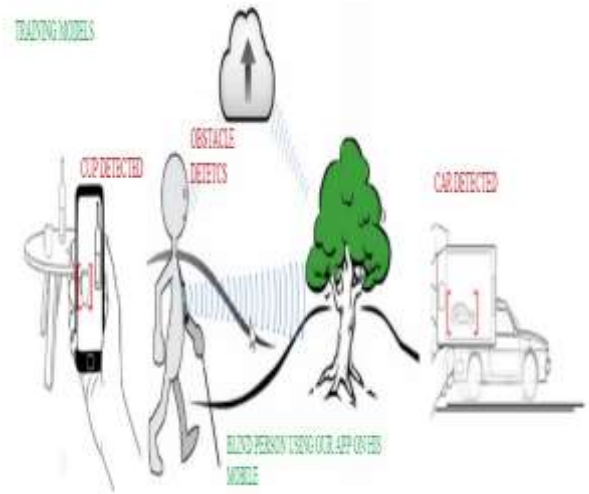

Fig1.1 Architecture of proposed system

Object Detection:

Object detection is the principal objective of this system. It consists of object classification and object localization. Object detection is the process of categorizing the object into different classes that were defined before. In other words, object classification assigns a label to an entire

https://doi.org/10.46335/IJIES.2022.7.5.9 e-ISSN: 2456-3463
Vol. 7, No. 5, 2022, PP. 49 -55

image. [6]That label is the name of the object present in that image. For instance, a computer is given an image of a cat and it will try to classify it and give the output as "Cat". It is easy for us to identify the objects present in any image, but for a computer, object classification is a tedious task. In object localization, the computer tries to isolate the object from the image by drawing a rectangular box around it which is also called Bounding Box. Thus, object detection is the combination of object classification and object localization in which we try to classify and isolate multiple objects present in the image. The output of this module will give us the name of the object and the coordinates of the bounding box. These coordinates of the bounding box will be used in the next modules to determine the direction of the object and to calculate the distance of the object from the user. **[**13]

Proposed System

Obstacle detection for visually impaired, we have developed an android application to detect obstacles on their way of blind peoples walk and telling them on which direction what objects are placed so that should help them in knowing things that they come across and to turn themselves on a right way. The difficulty of obstacle detection is how to locate and separate the obstacle from the complex background. Traditional computer vision algorithms cannot handle this problem very well due to the handcrafted designed features are vulnerable in complex background. In this article, we use deep convolution neural network (CNN) to detect obstacle in complex scene. The deep architecture of the CNN guarantees the features learned by the network are rich and effective for detecting the obstacle. The results show that the model achieved a good performance.[14][15]

Modules

Messages:

The user can check messages from our custom app. They can check in two forms inbox and sent. When User will select the system will speak out contents of it.[9]

Calls:

The user can check call logs from custom app. Also, user can use dialer. When User will select the contact the system will speak out as when last call received with time.[9][10]

User:

Blind people are really cursed to having that blindness. Visually impaired people can easily use this app. blind people can clicking the camera detection option; camera can detect the obstacle and speak the obstacle name by using Test to speech. So visually impaired people can know the obstacle behind him.[1][2]

Tensor Flow:

The Tensor Flow Object Detection API is an open-source framework built on top of Tensor Flow that makes it easy to construct, train and deploy object detection models. There are already pre-trained models in their framework which are referred to as Model Zoo. It includes a collection of pre-trained models trained on various datasets such as the COCO (Common Objects in Context) dataset, the KITTI dataset, and the Open Images Dataset. As you may see below there are various models available so what is different in these models. These various models have different architecture and thus provide different accuracies but there is a trade-off between speed of execution and the accuracy in placing bounding boxes.[2][3]

Tensor flow bundles together Machine Learning and Deep Learning models and algorithms. It uses Python as a convenient front-end and runs it efficiently in optimized C++. Tensor flow allows developers to create a graph of computations to perform. Each node in the graph represents a mathematical operation and each connection represents data. Hence, instead of dealing with low-details like figuring out proper ways to hitch the output of one function to the input of another, the developer can focus on the overall logic of the application. The deep learning artificial intelligence research team at Google, Google Brain, in the year 2015 developed Tensor Flow for Google's internal use. This Open-Source Software library is used by the research team to perform several important tasks.[2] [9][10].

CNN Algorithm:

A Convolution Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) 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 [10][1][2]

https://doi.org/10.46335/IJIES.2022.7.5.9 e-ISSN: 2456-3463
Vol. 7, No. 5, 2022, PP. 49 -55

While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics. The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area. [9][3]

Notes:

The user can add, make, edit and save notes from this custom notepad. It also includes custom made keypad. When user press the key, the system will speak out letting users know which key they are pressing. [10]

Text to Speech:

All the actions performed even the reminders are spoke. Since visually impaired people are able to see, this system is used to recite everything. It normally converts a normal text into speech.[11]

Navigation:

This system uses Google Map Api to show and speak out to the users their current position and also perform other tasks. It will be beneficial for them to navigate and walk by knowing their position.[7]

Web Browsing:

This system uses internet browser with help of speech to text and searches and speaks out the required result for the user. User enters whatever they want to search through speech recognition and the system will speak out the results for them. **[8].**

ARCHITECTURE:

Eclipse employs plug-ins in order to provide all of its functionality on top of (and including) the runtime system, in contrast to some other applications where functionality is typically hard coded. The runtime system of Eclipse is based on Equinox, an OSGi standard compliant implementation.

This plug-in mechanism is a lightweight software component try framework. In addition to allowing Eclipse to be extended using other programming languages such as C and Python, the plug-in framework allows Eclipse to work with typesetting languages like Latex, networking applications such as telnet, and database management systems. The plug-in architecture supports writing any desired extension to the environment, such as for configuration management. Java and CVS support is provided in the Eclipse SDK, with Subversion support provided by third-party plug-ins.[6]

With the exception of a small run-time kernel, everything in Eclipse is a plug-in. This means that every plug-in developed integrates with Eclipse in exactly the same way as other plug-ins; in this respect, all features are "created equal". Eclipse provides plug-ins for a wide variety of features, some of which are through third parties using both free and commercial models. Examples of plug-ins include a UML plug-in for Sequence and other UML diagrams, a plug-in for DB Explorer, and many others.[5][12]

Proposed System

Obstacle detection for visually impaired, we have developed an android application to detect obstacles on their way of blind peoples walk and telling them on which direction what objects are placed so that should help them in knowing things that they come across and to turn themselves on a right way. The difficulty of obstacle detection is how to locate and separate the obstacle from the complex background. Traditional computer vision algorithms cannot handle this problem very well due to the handcrafted designed features are vulnerable in complex background. In this article, we use deep convolution neural network (CNN) to detect obstacle in complex scene. The deep architecture of the CNN guarantees the features learned by the network are rich and effective for detecting the obstacle. The results show that the model achieved a good performance.[3][4]

Advantages

- Makes Independent

https://doi.org/10.46335/IJIES.2022.7.5.9 e-ISSN: 2456-3463
Vol. 7, No. 5, 2022, PP. 49 -55

- It is safe and ecofriendly application.
- Saves time and helps them to achieve daily tasks.

IV RESULT & DISCUSSION

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects. The task of the integration test is to check that components or software applications, e.g., components in a software system or – one step up – software applications at the company level – interact without error. [4]

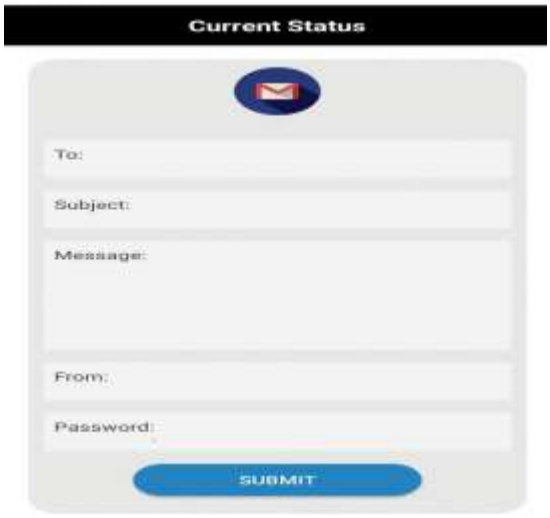

Fig: 4.1 Input Screen for the user

Fig: 4.2 Login Screen for the user

Fig: 4.3 Blind people interaction Output Screen

Fig: 4.4 Android outptut Screen

https://doi.org/10.46335/IJIES.2022.7.5.9 **e-ISSN: 2456-3463**
Vol. 7, No. 5, 2022, PP. 49 -55

International Journal of Innovations in Engineering and Science, www.ijies.net

CNN

When it comes to Machine Learning, Artificial Neural Networks perform really well. Artificial Neural Networks are used in various classification tasks like image, audio, words. Different types of Neural Networks are used for different purposes, for example for predicting the sequence of words we use Recurrent Neural Networks more precisely an LSTM, similarly for image classification we use Convolution Neural networks. In this blog, we are going to build a basic building block for CNN. Before diving into the Convolution Neural Network, let us first revisit some concepts of Neural Network. In a regular Neural Network there are three types of layers:

1. Input Layers: It's the layer in which we give input to our model. The number of neurons in this layer is equal to the total number of features in our data (number of pixels in the case of an image). [3][1]

2. Hidden Layer: The input from the Input layer is then feed into the hidden layer. There can be many hidden layers depending upon our model and data size. Each hidden layer can have different numbers of neurons which are generally greater than the number of features. The output from each layer is computed by matrix multiplication of output of the previous layer with learnable weights of that layer and then by the addition of learnable biases followed by activation function which makes the network nonlinear.

3.Output Layer: The output from the hidden layer is then fed into a logistic function like sigmoid or surtax which converts the output of each class into the probability score of each class.

Types of Layers:

Let's take an example by running a covnets on of image of dimension 32 x 32 x 3.

Input Layer: This layer holds the raw input of the image with width 32, height 32, and depth 3.

Convolution Layer: This layer computes the output volume by computing the dot product between all filters and image patches. Suppose we use a total of 12 filters for this layer we'll get output volume of dimension 32 x 32 x 12.

Activation Function Layer: This layer will apply an element-wise activation function to the output of the convolution layer. Some common activation functions are RELU: max (0, x), Sigmoid: 1/(1+e^-x), Tanh, Leaky RELU, etc. The volume remains unchanged hence output volume will have dimension 32 x 32 x 12.

Pool Layer: This layer is periodically inserted in the covnets and its main function is to reduce the size of volume which makes the computation fast reduces memory and also prevents over fitting. Two common types of pooling layers are max pooling and average pooling. If we use a max pool with 2 x 2 filters and stride 2, the resultant volume will be of dimension 16x16x12.[2][4]

V CONCLUSION

Our Application Concludes by including that the user can easily get to know obstacles on his way and can avoid them without the help from any other person. In this paper, we had proposed a much helpful voice assistant for Blind people. This system will be very easy to use. This application will run on Android operating System. This Voice assistant is very helpful towards the visually impaired people and makes very easy for them to use electronic gadgets with (text to speech) TTS technology, they will be able to interact more efficiently to the electronic system i.e. Mobile Phones, Tablets etc.[2][10][9]

VI FUTURE SCOPE

Further we can add up features like calculate distance between blind person or user and obstacle and telling to user on which direction he has to take moves. Further we can add up features like calculate distance between blind person or user and obstacle and telling to user on which direction he has to take moves. Navigation for the visually challenged using voice in structions and hepatic feedbacks can be added for improving the existing system. The system can be equipped with text to speech feature that can recognize the text in front of the visually impaired user and provide an audio output for the same. This will enable the user to understand the written text around him. The refined version of such a system can enable a visually impaired user to read a regular book without having to purchase an audio book.

• Navigation for the visually challenged using voice instructions and haptic feedbacks can be added for improving the existing system.

• Home automation using microcontrollers can be included in the future.[8]

https://doi.org/10.46335/IJIES.2022.7.5.9 **e-ISSN: 2456-3463**
Vol. 7, No. 5, 2022, PP. 49 -55

International Journal of Innovations in Engineering and Science, www.ijies.net

REFERENCES

[1] *"A navigational aid for persons with severe visual impairments: A project in progress," in Proc. Int. Conf. IEEE Eng. Med Biol. Soc., Cancun, Mexico, Sep. 17-21, 2003.*

[2] *S. Maeyama, A. Ohya, and S. Yuta, "Positioning by tree detection sensor and dead reckoning for outdoor navigation of a mobile robot, " in IEEE Int. Conf. Multisensory Fusion and Integration for Intelligent Systems '94, pp. 653-660, 1994.*

[3] *B. Andò, "A smart multisensory approach to assist blind people in specific urban navigation tasks, " IEEE Transactions Neural Systems and Rehabilitation Engineering, vol. 16, no. 6, pp. 592-594, Dec. 2008.*

[4] *J.K. Sung, Y.H. Kim, and H.K. In, "Development of an intelligent guide-stick for the blind" in IEEE Proc. Int. Conf. Robotics and Automation, vol. 4, pp. 3208-3213, 2001.*

[5] *J. Rudan and Z. Tuza. Using LMS-100 laser rangefinder for indoor metric map building. In Proceedings of 2010 IEEE International Symposium on Industrial Electronics (ISIE2010). IEEE ISIE, 2010.*

[6] *M. T. B. R. A. A. K. Y. Kevin Labuan, "A Wearable Portable Electronic Travel Aid for Blind," in International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), 2016 .*

[7] *V. Kaarthick Raja, A.Kartik, S.Prabhakaran "Voice Assistant for Visually Impaired People" in International Conference on Communication, Computing, and Internet of Things (IC3IoT).*

[8] *Arun Gopi, Shobhna Devi P, Sajini T, Bhadran V K," Implementation of Malayalam Text to Speech using Conca native based TTS for Android Platform", International Conference on Control Communication and Computing (ICCC), 2013*

[9] *Maya Ride, Michael Nahas, Milad Ghantous, "iSee: An Android Application for the Assistance of the Visually Impaired", Springer International Publication, 2014*

[10] *Sharaddha Buddha Khadilkar, Narendra Waghdarikar," Android Phone Controlled Voice Gesture and Touch Screen Operated Smart Wheel Chair", International Conference on Pervasive Computing (ICPC), 2015*

[11] *Prof. Rakhi Bharadwaj, Poonam Gupta, Pooja Jadhav, Bhagyashree Kadam, Amruta Kedari," Android based automated Wheel Chair", International Journal of Innovative Research in Computer and Communication Engineering, Vol 4, Issue 3, March 2016*

[12] *Dongmahn SEO, Suhyun KIM, Gyuwon SONG, Seung-Gil, "Speech-to-Text-based Life log system for smart phones", IEEE International Conference on Consumer Electronics (ICCE), 2014*

[13] *THOMAS, S. Natural Sounding Text-to-Speech Synthesis based on Syllable like Units, Department of Computer Science and Engineering, Indian Institute of Technology Madras, 2007*

[14] *EIKVIL, LINE. OCR Optical character Recognition, 1993*

[15] *Yu Zhong, T.V. Raman, Casey Burkhardt, FadiBiadsy and Jeffrey P. Bigham, "Just Speak: Enabling Universal Voice Control on Android*