

Object Detection For Blind People Using RNN Algorithm

^[1] Pavithra S, ^[2] Vaishnavi M, ^[3] Vidhya Kavi D, ^[4] Sowmya N, ^[5] M. Prakash

^[1] ^[2] ^[3] ^[4] Department of CSE, Er. Perumal Manimekalai College of Engineering, Hosur

^[5] Associate Professor, , Department of CSE, Er. Perumal Manimekalai College of Engineering, Hosur

Abstract: *Blind people face various challenges in their daily lives. Millions of people suffer from eye damage in one way or another. One of the most important nerves that is essential for a person to live a normal life is vision. Many people suffer from blindness and difficulty in walking around. This condition leads to the need for guidance or assistance in all the actions of a blind person. Blindness makes normal, professional, and social life difficult. Human vision has the amazing ability to store billions of images in the brain and to visualize images in comparison with previous images. Through the use of computerized visual aids, the quality of life of blind people can be improved. The RNN (Recurrent Neural Networks) algorithm was used to detect object in indoor and outdoor mode with high accuracy and processing speed. Input video to RNN, the object is found in the indoor environment such as person, seat, bench and outdoor area such as car, bus, pet, motorcycle etc. This will be done using Tensor Flow software and segregation of items as usual and harmful things will be done by taking Google site as a reference. Acquired image information is provided as voice output using Google speech synthesizer for visually impaired people to assist them in their movement. The app is easy to use and is integrated with speech integration so that the acquired object is transmitted to blind people as a voice output.*

I. INTRODUCTION

The ability to move from place to place is a vital part of daily life. Humans study the world around them, especially

in terms of sound and vision. It is a common belief that vision plays an important role, but many can have great difficulty in seeing the visual information they use, or when they use it. We find it easy to navigate in more familiar places without the idea of seeing. This can be mainly due to muscle memory. This can be done with examples such as going to the bathroom from your bedroom at midnight. But only a handful of people have experienced wandering in large, unfamiliar areas without the help of their eyes. Imagine trying to catch a train at a train station with your eyes closed while you are busy. However, those who are visually impaired walk independently every day. Their senses of smell and hearing are extremely sensitive, as they rely heavily on these senses. They also take time to feel the environment around them.

The enhanced app can detect objects in the user area. It can warn the user of obstacles in his path and in this way helps the user to navigate from one place to another and save him from any trip. It will also solve the problem of storing a special device. The reason it is so reliable is because it is developed in the Android operating system they are very common and are widely available almost everywhere. In fact, it is one of the most widely used operating systems. This makes the application easy to find.

OBJECTIVE

- The purpose of this project is to improve the application of blind people based on machine learning. It will eliminate the need for dedicated equipment and other wearables to help them see things as they move.
- This app will help them to overcome their difficulty in browsing which helps them to live their lives as a normal person.
- The aim of the project is to give blind users the ability to navigate in unfamiliar areas indoors and outdoors, using a friendly device through an object recognition system.

II. LITERATURE SURVEY

Another proposal during this paper focuses on providing data on the various forms of barriers before of the user, their size and distance from the user. MATLAB package is employed for signal process. Television camera is employed to record videos. Video process strategies area unit used subsequently. The output of this method provides not solely output by audio

format however conjointly vibration. The moving engine is connected to associate degree unhear able detector. The unhear able detector detects objects returning altogether directions and this causes the moving motor to vibrate. [1]

This system is making an attempt to induce the foremost out of the image. That's the essential essence of the system. It's a system during which the finders of associate degree N object area unit trained within the completely different N objects. Once a picture is shipped to the system, all the article receivers do their job. Once associate degree object is detected by a detector, it'll mark its boundary and write the name of the article. Once the method is complete for all N detectors, an image is displayed with all the tags. Moving the indicator over the article within the image shows the whole boundary of the article with its label nearby. This method is slower than different programs as a result of most object finder's work on one image. Performance is magnified by permitting quite one acquisition to figure seamlessly. [2]

Light distortion and hue distortion are addressed during this project through the utilization of shadow technique. In video, things area unit obtained by removing templates from the video. Infact this can be not the most effective approach. Applies once AN item is gift throughout the video. Compared to different options like SIFT, form options area unit higher accustomed notice objects in pictures. Therefore, during this project, native options area unit replaced by form options. [3]

This paper proposes a system in which two cameras are mounted on the glasses of a blind person. The proposed work has a wearable device and includes a blind rod and a hearing-based acquisition circuit. It uses an infrared sensor that uses infrared waves to scan a person's surroundings. The

system must be trained in the knowledge of the object. Feature removal is also part of the process [4]

III. EXISTING METHOD & PROPOSED METHOD

EXISTING METHOD:

- In the existing system mount the network end-to-end bluraid feature aggregation network(BFAN) feature to detect the video object. The proposed BFAN focuses on a blending process influenced by blurring that includes movement blurring and eliminating high-precision focus and increased mathematical calculations.
- In BFAN, we test the degree of blurring of each individual object as a compound weight. Notably, the background is usually flat which has a negative impact on the degree of object dimension.
- Therefore, we introduce a light detection network to reduce background distortion. An experiment conducted on the Image Net VID database shows that BFAN achieves a modern acquisition performance, 79.1% mAP, with a 3-point improvement compared to the video object base.
- In the present system, the acquired object information is displayed in the window only, not providing voice format output.

PROPOSED SYSTEM

- The proposed system captures the image through the continuous video stream rather than taking the picture of each object every time.
- The object detection and recognition have to be accurate and the detecting speed of the object should be high so that the navigation for the blind people would be easier. For such high-speed detection, RNN algorithm has been implemented.
- The first step is to capture the objects using the camera of this the application should get the camera permission of the device.
- Then the image is sub-segmented so that multiple regions can be formed from a single image. The algorithm then combines similar regions to form a larger region and finally produce the region of interest.
- The pre-trained recurrent neural network is trained again based on the number of classes that has to be detected then the region of interest is identified.
- Based on the region of interest, the objects and the backgrounds are classified. For each identified object in

the image, tighter bounding boxes are generated based on the linear regression model.

- The input video to the RNN, the object is detected in indoor environment like person, chair, bench and outdoor environments like car, bus, animal, motorcycle etc

IV. SYSTEM FUNCTION ARCHITECTURE DESIGN:

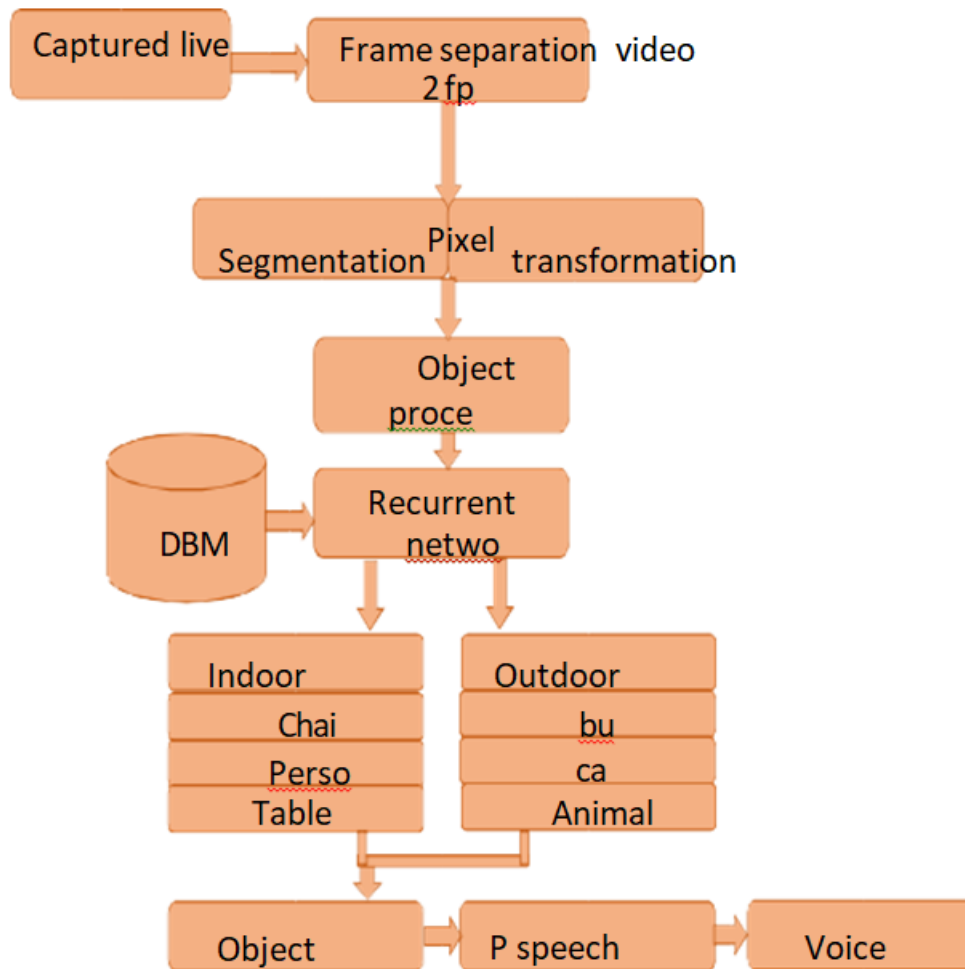


Fig.no:1 Architecture Design

V. MODULES

- Camera module
- Image segmentation module
- Database module
- Recognition module
- Audio module

STANDARD DESCRIPTION

1. Camera module: In this module video input can be captured using the camera. Video frame frames can be converted as a photo by removing the image from the video.
2. Image segmentation module: In image preprocessing module, the input extracted images can be sub- segmented so that multiple regions can be formed from a single image.

3. Database module: The model of trained images can be stored in data base management system for detection of objects in indoor and outdoor environments.
4. Recognition module: The Recurrent neural network algorithm is used for recognition of objects in indoor and outdoor environments by comparing with data base system.
5. Audio module: The recognized object images format can be converted as audio format based on google speech synthesizer and the audio output is easy for understanding blind and self-navigate without any persons help.

RECURRENT NEURAL NETWORK ALGORITHM (RNN)

- Recurrent Neural Network (RNN) is a type of Neural Network where the output from the previous step is given as input to the current step. In a normal neural network, all inputs and outputs are independent of each other, but in situations such as the need to predict the next word in the sentence, past words are needed and that is why it is necessary to remember previous words.

Thus, RNN came into existence, which resolved the issue with the help of a Hidden Layer. The main and most important feature of the RNN is the Hidden mode, which remembers certain information about the sequence.

- The RNN has a “memory” that remembers all the information about the content. It uses the same parameters for each input as it performs the same function in all inputs or hidden layers to generate output. This reduces the complexity of the parameters, unlike other sensory networks.

SEGMENTATION PROCESS

Segmentation divides an image into separate regions that contain each pixel with similar characteristics. In order to be meaningful and useful in image analysis and interpretation, regions must be closely related to the presented objects or features of interest. Meaningful separation is the first step from low-level image processing that converts a gray scale or a color image into one or more images into a high-definition image description depending on the features, objects, and scenes. The success of image analysis depends on the reliability of the separation, but accurate image separation is often the most challenging problem. Classification strategies are flexible or inconsistent with context. The latter does not matter the spatial relationship between the elements in the image and the group pixels together on the basis of a specific earth element, e.g., gray level or color. Contextual strategies additionally use these relationships, e.g., combine pixels with the same gray levels as well as areas that close the space.

COCO DATA SET

The COCO (Common Objects in Context) data set is for large quantity acquisition, classification, key point detection, and caption data set. The database contains 328K images.

Annotations: The database contains annotations for

- Object acquisition: bounding boxes and separating masks for each example have sections of 80 items,
- Captions: descriptions of the natural language of images (see MS COCO Captions),
- Key points detection: containing more than 200,000 images and events of 250,000 people labeled with key points (about 17 key points, such as left eye, nose, right hip, right ankle),
- Image classification - mask for each pixel class with 91 categories of objects, such as grass, wall, sky (see MS COCO Items),
- Panoptic: full-scale division, into 80 categories of objects (such as human, bicycle, elephant) and a small set of 91 objects (grass, sky, road),
- Crowded: more than 39,000 images and events of 56,000 people labeled with Dense Pose annotations - each label is defined by the model id and the drawing between the pixels of the human body and the 3D model image.

OBJECT DETECTION ALGORITHMS

Object detection is a technology related to computer vision and image processing related to the acquisition and acquisition of semantic objects of a particular category (such as people, buildings, or cars) in digital images and videos.

Neural Network Object Detection Algorithms:

Object discovery has always been an exciting challenge in the field of deep learning. It is mainly done through convolutional neural networks (CNNs), and especially regional-based CNNs (R-CNNs).

R-CNN (regional based CNNs):

Instead of working with a large number of regions or passing through each pixel and part of an image to search for an object, the R-CNN algorithm raises a number of boxes inside the inserted image and checks to see if any of these boxes contain any object. R-CNN then uses the selected search to extract these boxes from the image.

SPEECH SYNTHESIZERS

Speech synthesizers are text-to-speech systems used with computers. It could be a computer-generated card, a computer-connected computer box, or software that works with a computer audio card. Speech synthesizers are designed to incorporate all the phones and grammar rules, allowing words to be pronounced correctly even though compound words and words can cause problems because they often contain unusual spellings and spelling. Another artificial speech sounds like a robot; however, many new products sound almost human.

DBMS

- Database management software is software used to manage data. Example: MySQL, Oracle, etc. is a very popular commercial website used in various programs.
- DBMS provides a visual interface for performing various tasks such as creating a website, storing data on it, updating data, building a table on a website and much more.
- Provides security on the website. DBMS allows users the following functions:
- Data Definition: Used for the creation, modification, and deletion of definitions that define data processing on a website.
- Data Updates: Used for uploading, modifying, and deleting real data on a website.
- Data Recovery: Used for data retrieval on a website that can be used by applications for a variety of purposes.

VI. SYSTEM SOFTWARE

PYCHARM

The Community Edition is released under the Apache License, and there is also an educational version, as well as a Professional Edition with extra features (released under a subscription-funded proprietary license).

SQL Alchemy as Debugger:

You can set a breakpoint, pause the debugger and see the SQL representation of the user's SQL language code expression.

Git Visualization in Editor

When coding in Python, questions are common to the developer. You can easily check the last commitment on PyCharm as it has blue sections that can explain the difference between the last bond and the present.

Code Coverage in Editor

you can use .py files without PyCharm Editor and mark them as details for cover code elsewhere in the project tree, in the summary section etc.

VIII. CONCLUSION

The system has a new method of object detection and identification. It can be easily sold and made to benefit the visually impaired community. In our project, we have used a large website thanks to a pre-trained Neural Network model. The Single Shot Detection Mobile Net model is trained using a COCO database that contains approximately multiple images. From the analysis it was concluded that the system is extremely accurate in identifying objects in the indoor and outdoor environment with a 95% accurate result. Thus, the proposed system will assist visually impaired people in an effective way.

REFERENCES

- [1] arcbotics.com/products/sparki/parts/ultrasonic- rangefinder
- [2] Bin Jiang, Jiachen Yang, HoubingSong, "Wearable Vision Assistance System Based on Binocular Sensors for Visually Impaired Users," DOI:10.1109/IJOT.2018.2842229, IEEE INTERNET OF THINGS JOURNAL, 2018.
- [3] Computer vision guidance system for indoor navigation of visually impaired people, 2017 IEEE 8th International Conference on Intelligent Systems.
- [4] Daniyal Rajput, Faheem Ahmed, Habib Ahmed, Engr Zakir Ahmed Shaikh, Aamir Shamshad, "Smart Obstacle Detector for Blind Person," 2017.
- [5] Daniyal Rajput, Faheem Ahmed, Habib Ahmed, Engr Zakir Ahmed Shaikh, Aamir Shamshad, 2019, "Smart Obstacle Detector for Blind Person," Journal of Biomedical Engineering and Medical Imaging.
- [6] D. Dakopoulos, S. K. Boddhu, and N. G. Bourbakis. A 2d vibration array as an assistive device for visually impaired. Pages 930–937. IEEE Computer Society, 2017.
- [7] Hao Jiang, Thomas Gonnot, Won-Jae Yi and JafarSaniie, "Computer Vision and Text Recognition for Assisting Visually Impaired People using Android Smartphone," DOI: 978-1-5090-4767-3, IEEE, 2019.
- [8] <https://towardsdatascience.com/deep-learning-for-objectdetection-a-comprehensive-review-73930816d8d9> information on coco from www.tensorflow.org
- [9] JinqiangBai, ShiguoLian, Zhaoxiang Liu, Kai Wang, Dijun Liu, "Smart Guiding Glasses for Visually Impaired Package Management All installed packages are displayed with the appropriate visual representation. This includes a list of People in Indoor Environment," IEEE Transactions on Consumer Electronics, Vol. 63, No. 3, August 2017.
- [10] JinqiangBai, ShiguoLian, Zhaoxiang Liu, Kai Wang, DijunLiu, "Smart Guiding Glasses for Visually Impaired People in Indoor Environment," IEEE Transactions on Consumer Electronics, Vol. 63, No. 3, August 2018
- [11] Khushboo Khurana, Reetu Awasthi, "Techniques for Object Recognition in Images and Multi-Object Detection," 2018.
- [12] Khushboo Khurana, ReetuAwasthi, 2019, "Techniques for Object Recognition in Images and Multi-Object Detection," International Journal of Advanced Research in Computer Engineering & Technology.
- [13] Leo, M, Medioni, G.G., Trivedi, M.M., Kanade, T, Farinella, G.M, "Computer Vision for Assistive Technologies," CVIU. 2017, 154, 1–15.
- [14] L. A. Johnson and C. M. Higgins. A navigation aid for the blind using tactile-visual sensory substitution. Conf Proc IEEE Eng Med Biol Soc, 1:6289–6292, 2016.
- [15] Payal Panchal, Gaurav Prajapati, Savan Patel, Hinal Shah and Jitendra Nasriwala, "A Review on Object Detection and Tracking Methods," 2018.
- [16] Prof. Seema Udgirkar, Shivaji Sarokar, Sujit Gore, Dinesh Kakuste, Suraj Chaskar, 2019, "Object Detection System for Blind People," International Journal of Innovative Research in Computer and Communication Engineering.
- [17] TensorFlow Object Detection API tutorial [Tensorflow Object Detection API](https://github.com/tensorflow/models/tree/master/research/object_detection)
https://github.com/tensorflow/models/tree/master/research/object_detection
- [18] World Health Organization (WHO) - Visual impairment and blindness.