iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 7 **Issue:** V **Month of publication:** May 2019

DOI: https://doi.org/10.22214/ijraset.2019.5345

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.177
Volume 7 Issue V, May 2019- Available at www.ijraset.com

Object Recognition for Visually Impaired Persons Using Smart phone

Sahil Vishwasrao[1], Kundan Dalvi[2], Prof. Namdev Kedare[3]
[1, 2]Student's, [3]Professor, Department of Information Technology, DPCOE, Pune University, India.

Abstract: In this project we describe main features of software modules developed for Android smartphones that are dedicated for the blind users. The main module can recognise and match scanned objects to a database of objects, e.g. food or medicine containers.
The two other modules are capable of detecting major colours and locate direction of the maximum brightness regions in the captured scenes.
The blind and the visually impaired people face diverse kinds of life challenges. The safety and independent mobility of these people is coming into danger. The system is based on a dedicated image recognition application running on an Android system Smartphone. Image recognition results are communicated to the blind user by means of pre-recorded verbal messages.
Keywords: Object Detection, Image Recognition.

I. INTRODUCTION

With the rapid development of artificial intelligence and mobile computing, modern technology has brought more convenient to the blind and visually impaired people. Lately, it's estimated that there are about 253 [1] million people with vision impairment. The group of visually impaired persons are undergoing an inconvenient daily life without the assistant from their family or friends. An effective method for guidance is using guide dogs.

However, the disadvantage of this method is that guide dogs need plenty of money and time to train and feed it. Also, the blind need an approach to get to know about the life outside their home and they desire to have access to internet and mobile services as normal one.

However, as for them, a lot of obstacles exist and there remains need improvement from society and technology, including lack of information resources for blind, inadequate infrastructure and lack of technical input. [2] But, according to our research, the rapid development of artificial intelligence and mobile computing technology can be an ideal solution to help the blind and visually impaired persons to perceive their surroundings.

To assist blind and visually impaired group, many solutions has been provided. Some solutions try to design a hardware system to provide some fundamental functions. For example, Mohamed Manoufali [3] et al. designed cane for blind with obstacle detection by ultrasonic sensor. And Siti Fauziah Toha [4] et al. also has the similar idea about assistance.

However, those solutions can't detect the objects around the user. Another type of solution for blind assistance is to provide guide and service to blind users.

Jiayin S. [5] et al. provides a construct of a guide device with some fundamental functions and the user need to press buttons to use those services but the user experience and functions are limited. In this paper, we presented" beEYE", an extendable system launched on Android phone to provide functions for visually impaired people. It has functions including messaging, describing the street view, navigating to certain place, etc. We have integrated those discrete function into a unified system with a voice interface provided to the blind. With our system, we hope to greatly improve their life.

II. GOALS AND OBJECTIVES

A. The aim of this project is to represent a method where a blind person can get information about any object through speech signal and Voice.
B. The project to build a system that converts visual input into audio signals which may lead to a practical application to help the blind or visually impaired navigate
C. Help blind people, to gain better knowledge about the front objects.
D. Alert blind person about the danger forward [digs, vehicles, etc].

©IJRASET: All Rights are Reserved

III. LITERATURE SURVEY

A. " Tang, H., and Beebe, D. J. An oral tactile interface for blind navigation.,2006. IEEE Trans Neural Syst Rehabil Eng. pp. 116123"

An oral tactile interface was designed and evaluated to provide directional cues through the tactile channel, which may be utilized by a blind traveller to obtain directional guidance in outdoor navigation. The device was implemented as a mouthpiece with a micro fabricated electro tactile display on top for tactile presentation onto the roof of the mouth and a tongue touch keypad at the bottom for simultaneous operation by the tongue. An experimental system allowed a user to communicate with a computer tactilely by using the oral interface.

Directional cues were presented to the user as line or arrow patterns with four moving directions (leftward, rightward, forward, and backward). Electro tactile presentation on the roof of the mouth was evaluated in experiments of threshold measurement and identification of directional cues. Experimental results from six human subjects showed that the roof of the mouth required stimulation intensities around 15 V for threshold sensation, and around 25-30 V for comfortable and well-perceived stimulation. Furthermore, identification of leftward or rightward movements was highly accurate while performance on forward or backward moving patterns was mixed and varied considerably among subjects.

B. " A. Quattoni, and A. Torralba. Recognizing Indoor Scenes. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2009."

Indoor scene recognition is a challenging open problem in high level vision. Most scene recognition models that work well for outdoor scenes perform poorly in the indoor domain.

The main difficulty is that while some indoor scenes (e.g. corridors) can be well characterized by global spatial properties, others (e.g., bookstores) are better characterized by the objects they contain. More generally, to address the indoor scenes recognition problem we need a model that can exploit local and global discriminative information. In this paper we propose a prototype-based model that can successfully combine both sources of information. To test our approach, we created a dataset of 67 indoor scenes categories (the largest available) covering a wide range of domains. The results show that our approach can significantly outperform a state of the art classifier for the task.

C. " Pedro F Felzenszwalb, Ross B Girshick, David McAllester, and Deva Ramanan. Object detection with discriminatively trained part-based models. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 32(9):16271645, 2010."

We describe an object detection system based on mixtures of multiscale deformable part models. Our system is able to represent highly variable object classes and achieves state-of-the-art results in the PASCAL object detection challenges. While deformable part models have become quite popular, their value had not been demonstrated on difficult benchmarks such as the PASCAL data sets.

Our system relies on new methods for discriminative training with partially labelled data. We combine a margin-sensitive approach for data-mining hard negative examples with a formalism we call latent SVM. A latent SVM is a reformulation of MI–SVM in terms of latent variables. A latent SVM is semi convex, and the training problem becomes convex once latent information is specified for the positive examples. This leads to an iterative training algorithm that alternates between fixing latent values for positive examples and optimizing the latent SVM objective function.

D. " Jizhong Xiao, Kevin Ramdath, Manor Iosilevish, Dharmdeo Sigh, and Anastasis Tsakas. A low cost outdoor assistive navigation system for blind people. In Industrial Electronics and Applications (ICIEA), 2013 8th IEEE Conference on, pages 828833. IEEE, 2013."

With over 39 million visually impaired people worldwide, the need for an assistive device that allows the blind user navigate freely is crucial. We have developed an off-line navigation device that uses 3-D sounds to provide navigation instructions to the user. Our device relays directional information to the user through special Audio Bone headphones, which use bone conduction technology. Sounds are recorded and can therefore be selected by the blind user. Navigation processing is handled by a Raspberry Pi. We are using a magnetic compass and gyroscope to calculate the direction that the user is facing. Route queries of the destination address are geocoded using the Geo-Coder-US module and passed to the MoNav module to generate a pedestrian route. Additional capabilities of the device include speech recognition and voice prompts for obtaining the user's desired destination address. The user can input the address by speaking into a microphone. The entire system is mounted to a pack that sits on the user's waist. It is very light and portable and it does not impede any of the user's senses while it is being used.

©IJRASET: All Rights are Reserved

E. " Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 580587, 2014."

Object detection performance, as measured on the canonical PASCAL VOC dataset, has plateaued in the last few years. The best-performing methods are complex ensemble systems that typically combine multiple low-level image features with high-level context. In this paper, we propose a simple and scalable detection algorithm that improves mean average precision (mAP) by more than 30 percentage relatives to the previous best result on VOC 2012 – achieving a mAP of 53.3 percentage. Our approach combines two key insights: (1) one can apply high-capacity convolutional neural networks (CNNs) to bottom-up region proposals in order to localize and segment objects and (2) when labelled 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. We also present experiments that provide insight into what the network learns, revealing a rich hierarchy of image features. Source code for the complete system is available at http://www.cs.berkeley.edu/ rbg/rcnn.

F. " Ross Girshick. Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision, pages 14401448, 2015."

This paper proposes a Fast Region-based Convolutional Network method (Fast R-CNN) for object detection. Fast R-CNN builds on previous work to efficiently classify object proposals using deep convolutional networks. Compared to previous work, Fast R-CNN employs several innovations to improve training and testing speed while also increasing detection accuracy. Fast R-CNN trains the very deep VGG16 network 9x faster than R-CNN, is 213x faster at test-time, and achieves a higher map on PASCAL VOC 2012. Compared to SPPnet, Fast R-CNN trains VGG16 3x faster, tests 10x faster, and is more accurate. Fast RCNN is implemented in Python and C++ (using Caffe) and is available under the open-source MIT License at https://github.com/rbgirshick/fastrcnn.Title and Author Details.

IV. PROBLEM DEfiNITION

To develop an android application, for blind people, that detects an image and analyse it from built in database and give vocal information and direction about the object.

V. SYSTEM ARCHITECTURE

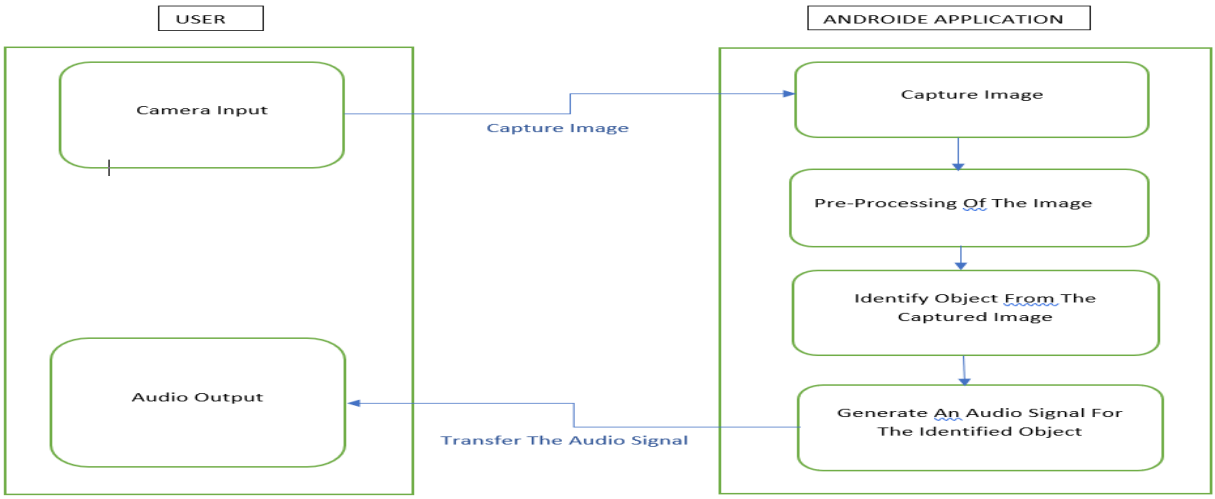


Fig. System Architecture

Modules

A. Frame Extraction

The application allows recognizing objects from images recorded by the camera of a mobile device. The object recognition algorithm should be insensitive to image registration parameters, i.e. scale, rotation and lighting conditions. Moreover, the recognized object should be robustly detected and localized in the image context (e.g. among other similar objects).

©IJRASET: All Rights are Reserved

B. Object Recognition

In this object is recognized by selecting the frame by a key frame algorithm for efficient frame selection, after that the object is converted into pixels. Then pixels is matched with the stored folder if the object recognized it will convert it to voice output, otherwise it takes the object as new one then it will be saved if recommended. The Object Recognition module provides a way to identify specific trained objects within the current image. Once the module is trained with sample template images it will identify those objects within the current image depending on the filtered parameters of confidence, size, rotation, etc

C. Feature Extraction

In this module we extract the feature from the image frame. In this module we do the following Edge Detection, Corner Detection, color Transformation and color classification and it will determine obstacle and image recognition result

D. TTS

After the object recognition, the text stored in the object below is converted into the voice output to the blind user. It is done by the text to speech application in android OS.

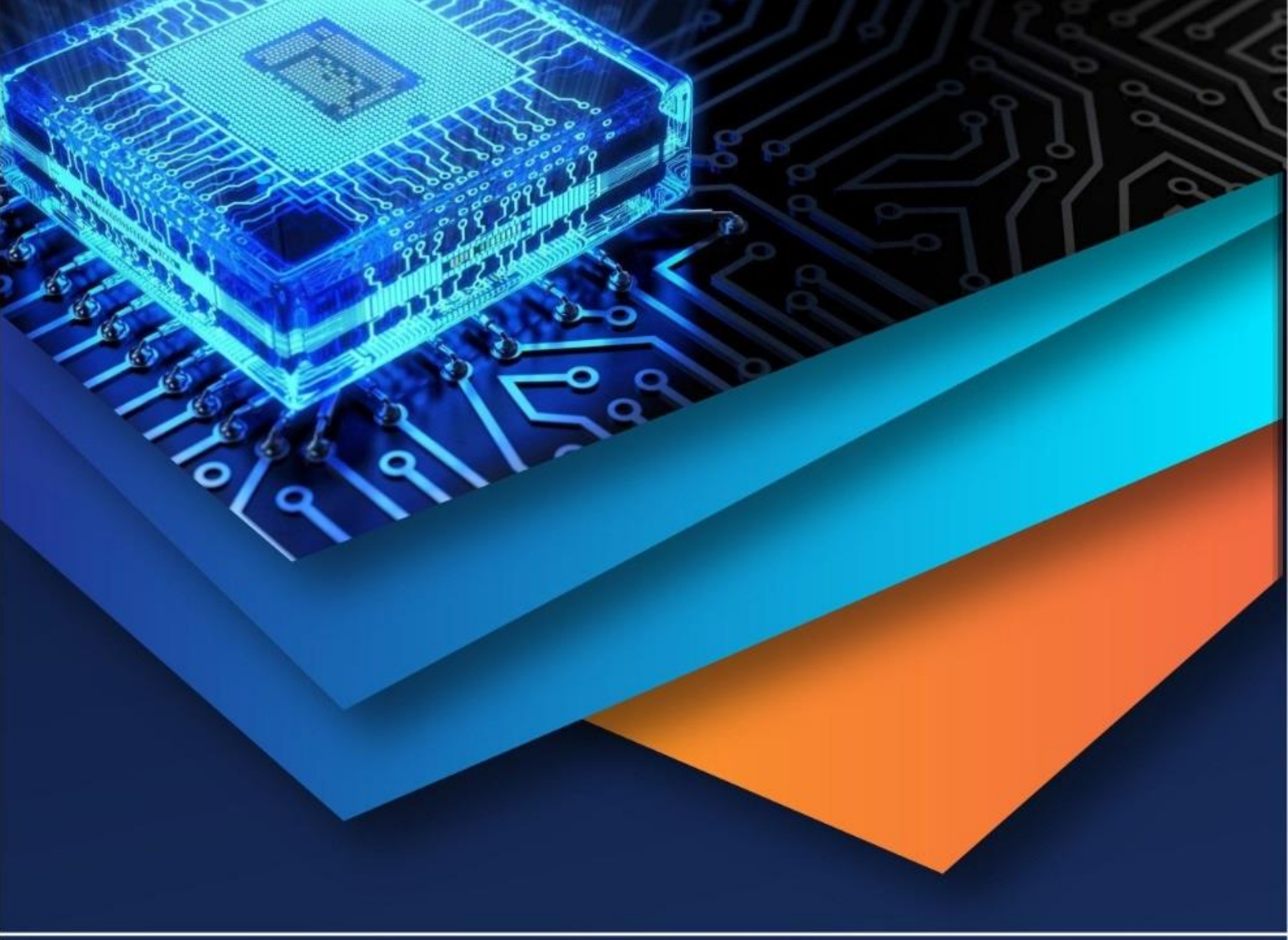
VI. CONCLUSION

This study designed a user-friendly guidance system for visually impaired people that involves a smartphone and server. When the system is in use, the smartphone continuously transmits images of the scene in front of the user to a server through 4G technology or a Wi-Fi network. Subsequently, the server performs the recognition process. The result is transmitted back to the smartphone, which provides the user with obstacle information through voice notations. In this prototype, we investigate the need from blind and visually impaired people we develop a blind visualization system that helps blind people better explore the surrounding environment. A portable and real time solution is provided in the work. Thus, we can conclude• The application is going to help the person in daily life with flexible environment. • Algorithms are efficiently executed and used as per the requirement. • As we are working on Social cause, application will fulfil the requirement of blind person. • In other words, technically we are donating the eye to blind person.

REFERENCES

[1] ” An oral tactile interface for blind navigation.”” Tang, H., and Beebe, D. J.”,” IEEE”, ”2006”
[2] ” Recognizing Indoor Scenes.”,” A. Quattoni, and A. Torralba”,” IEEE”,”2009”
[3] ” Object detection with discriminatively trained part-based models”, “Pedro F Felzenszwalb, Ross B Girshick, David McAllester, and Deva Ramanan”,” IEEE”,”2010”
[4] ” A low cost outdoor assistive navigation system for blind people”,” Jizhong Xiao, Kevin Ramdath, Manor Iosilevish, Dharmdeo Sigh, and Anastasis Tsakas”,” IEEE”,”2013”
[5] ”. Rich feature hierarchies for accurate object detection and semantic segmentation”, “Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik.”,” IEEE”,”2014”
[6] ” Fast r-cnn. In Proceedings of the IEEE International Conference on Computer Vision.”, Ross Girshick.”, ”2016.”
[7] “New Object Detection, Tracking, and Recognition Approaches for Video Surveillance”,” huai Zhang, Member, Chong Wang, Member, Shing-Chow Chan, Member, Xiguang Wei, and CheckHei Ho”,”2015”
[8] ” Computer vision guidance system for indoor navigation of visually impaired people IEEE 8th International Conference on Intelligent Systems”,” Kabalan Chaccour; Georges Badr”, ”2016”
[9] ” Being Aware of the World: Toward Using Social Media to Support the Blind with Navigation.”, “Joseph, S.L.; Xiao, J.; Zhang, X.; Chawda, B.; Narang, K.; Rajput, N.; Mehta, S.; Subramaniam, L.V.”,” IEEE Inter-national Conference”,”2015”

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)