RESEARCH ARTICLE OPEN ACCESS

Third Eye – An App for Blind Assistance Using Machine Learning

Muhammed Shihab V.K [1], Malavika Dileep [2], K.M. Muhammed Rameez [3],
Ashish Soni [4], Sheena Kurian K [5]
Department of Computer Science & Engineering
KMEA Engineering College,
Ernakulam, Kerala, India

ABSTRACT
Third Eye focuses on helping the visually impaired people to detect objects which are used in day to day life and to recognize all people who are trained by the system. Machine learning is used for object detection and face recognition. Text on detected objects are read using Optical Character Recognition technology. When the name of a trained object is given as voice input by the visually impaired, Third Eye will search for the object in the vicinity and inform whether the object is present or not. OCR is also used for recognizing sign board, product names, etc. Third Eye also has facility to inform the caretaker who registered the visually impaired person, about his location using Google maps. Mobile phone camera is used for capturing images. Third Eye is implemented as an android application which will be paired with a Bluetooth headphone. The feedback to visually impaired user is in the form of speech.

Keywords :- Back propagation, Deep Learning, Geo-fencing, Neural network, Optical character recognizer

I. INTRODUCTION

Within the last decade, technology has flourished indefinitely. In every country, a small proportion of population are disabled in one way or other. Regular activities commonly performed by normal humans such as safe navigation in a novel indoor/outdoor environment, independent shopping or simply reaching a desired destination become highly challenging for visually impaired people [1].

Humans are getting busier each day and does not have time even for their families. Today in this independent world, it is important that we should do our day-to-day activities independently. For the visually impaired it is not always possible to be independent. The ability of people who are visually poor or have significant visual impairments to read printed text and product packages will enhance independent living and foster economic and social self sufficiency [2,3]. With advanced technologies like machine learning and image processing, we can help visually impaired people in managing their life easier.

With rapid development in artificial intelligence, machine learning, mobile computing and other modern technological domains, there are many innovative researches taking place to help visually challenged persons. Deep learning is a subdivision of machine learning algorithms, which identify interesting patterns. An artificial neural network is inspired by biological nervous system and process information. Convolutional Neural Network (CNN) is a special type of neural network that works just like the normal neural network, which initially has a convolution layer [4]. CNN is one of the most popular technique with high accuracy for image classification. Back propagation algorithm is a supervised learning algorithm that has multiple layers and is used for training neural network. The number of smartphone users are increasing each day. Now it is possible, even for poor ones to have a smartphone at an affordable price. Development of an application that can run on smartphones can indeed help many, as they are easily accessible.

Third Eye is an android application that provide assistance for visually challenged persons. Machine learning is used as the primary case. This system can provide safe navigation facility to the blind user. Furthermore, it provides provision for detecting objects as well as facility for recognizing friends and relatives. Using Third Eye visually challenged people can read text on handheld objects, sign boards, etc. An android smartphone is used which is paired with a Bluetooth headphone. The output is fed as an audio to the visually impaired user.

II. EXISTING SYSTEM

In today's world there exists many applications to help the visually impaired people, ranging from infrared enabled blind stick to wearable bands. These devices help blinds for safe navigation and enable them to live like normal humans. Most of these devices are embedded and hence it need to be carried with utmost care. Moreover, poor visually impaired people cannot afford these devices due to this high cost.

The SmartVision system introduced in [5] detects sidewalks border and objects which are in front of visually impaired user. Canny edge detection algorithm is used and

objects are identified using a zero crossing approach. In [6], it presents a camera based system that helps visually impaired in reading text written on hand held objects and converts it into an audio output. Here image is captured using a camera and object region is detected. OCR is used for detecting text and the detected text is compared with template and is converted to audio output.

Mobile device based framework named Mobile Vision, as proposed in [7] is a prototype used for identifying landmarks in an environment and guide visually impaired people towards it. In [8] it describes a method using computer vision that helps blind in reading text and identifying their surroundings using k-Nearest neighbor classifier. In [9] it proposes a system that provides voice assistance for blind including fall detection, safety care, providing warnings and daily information broadcasting. Object detection algorithm based on Mask R-CNN is used. [10] is based on computer vision algorithms and deep convolutional neural network that is designed to help visually impaired people for their safe navigation in outside environment. It provides warnings to blind with a headphone and also read sign boards.

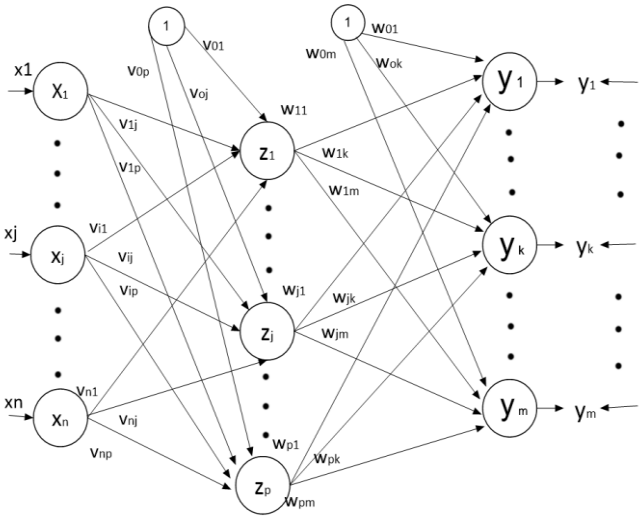

Fig. 1 Back Propagation Network

III. PROPOSED SYSTEM

The Third Eye is implemented as an android application that is paired with a Bluetooth headphone. The phone camera works as the eye of the blind. This framework is mainly based on deep learning and convolutional neural network (CNN) using back propagation algorithm. Back propagation is a supervised learning algorithm that is used for training neural network. It is a multilayer, feed-forward neural network that consist of an input layer, hidden layer and also an output layer and have biases.

Third Eye mainly focuses on helping the visually impaired people. It provides assistance to blind users like identification of family members and friends, detecting objects which are used in their daily life, detection of text on hand held objects and sign boards, searching an object, tracking their current location, etc. The system mainly consist

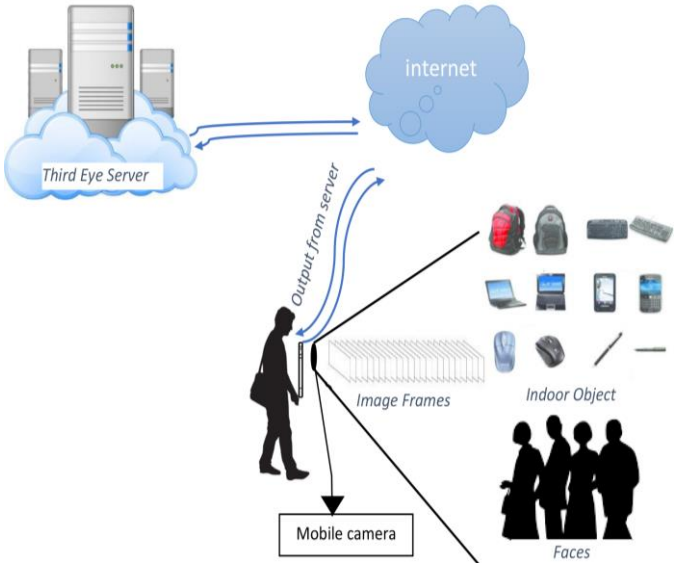

Fig. 2 Proposed System

of three modules out of which two modules are web interface and the main module is based on android. To compose the program Third Eye utilizes Python and TensorFlow. SQL is used for creating the required databases.

Tensorflow provides different functionalities for implementing Deep Learning Models. Tensorflow is an open source deep learning structure that builds up the granular control of designers on each neuron [11]. Tensorflow is used for object detection. It creates models for provided dataset and also there is a large pre-trained dataset available in Tensorflow. The three modules in Third Eye can be briefly described as follows:

A. Blind user

This is the most important module that is purely based on android. Android Studio is used for creating this module. Android Studio is an official IDE for Google's Android operating system designed especially for android development. Android SDK provides several features like identifying location using Google maps, text to speech conversion and speech to text conversions, etc., with more accuracy. Android is connected to python via REST API.

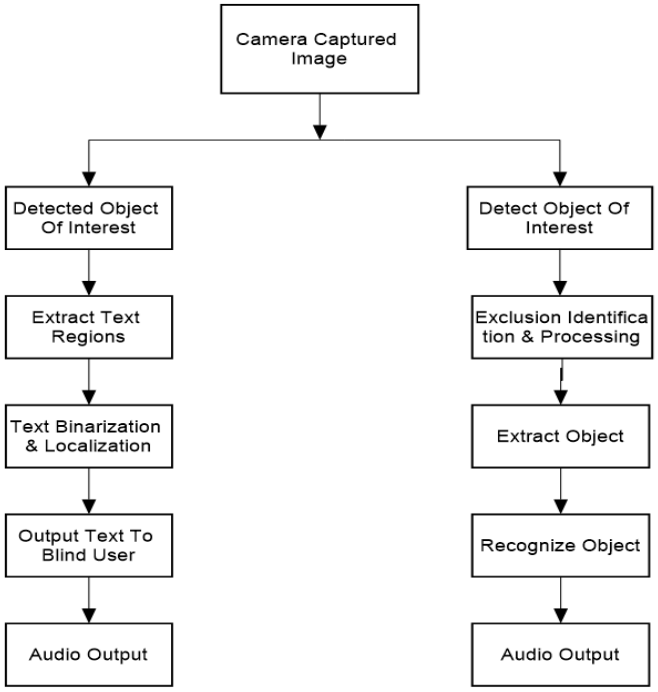

Fig. 3 Flow chart of proposed system

Blind user module has many sub-modules which are as follows:

1) ***Optical Character Recognition (OCR):***

Third Eye uses Optical Character Recognition (OCR) for detecting text such as text on sign boards, handheld objects, etc. It is the commonly and widely used technology for converting captured images to text. Here the text on captured images is detected with the help of a camera. Mobile phone cameras are used for capturing images and from these images, text are detected. Detected text are then converted to speech and provided as voice assistance to blind. This can make independent shopping possible for blinds.

2) ***Object Detection:***

Tensorflow is used for object detection. This is also used for face recognition. It has many worked in libraries and high accuracy compared to other techniques. Although, it recognizes objects and images accurately, it requires high specification system for fast detection. COCO dataset is a purely tensorflow based dataset that contains about 90 pre-trained objects.

3) ***Convert to Voice and Play:***

The detected objects as well as detected text must be converted to audio output for blind. This is then fed as an audio output via headphone. This functionality is performed by android.

4) ***Search for an Object:***

When the name of a trained object is given as voice input by the visually impaired, Third Eye will search for the object in the vicinity and inform him. This is also implemented using android.

5) ***Request for location via voice:***

Third Eye provides a facility for blinds to identify their current location. When the blind requests his current location as voice input, with the help of GPS, location is detected with the help of Google maps and output is given in the form of speech.

B. Caretaker

This module is implemented as the web interface. Caretaker is responsible for taking care of blind. Caretaker first registers himself with his email id as username and a unique password. He then registers blinds. There can be more than one blind under a single caretaker. IMEI number of mobile phone is used as a unique identification for each blind. Blind is logged in to the android application by his caretaker. Once logged in, the application runs in the background of the android device.

Caretaker has access to information about all the blind he registers. He can track the current location of blind with the help of Google maps. Third eye provides a facility to alert to caretaker whenever a blind leaves a particular area. This is done with the help of a technique known as geo-fencing. A geo-fence is a virtual perimeter for a real-world geographic area [12]. This could be dynamically generated also. In addition to track location, latest image captured by mobile camera is stored in the server for finding the exact location of blind. Furthermore it is the duty of caretaker to train the system with images of family members and friends. When a new blind is registered, caretaker trains the system with the images of close relatives and family

friends, so that the system can recognize them later. In short caretaker manages the activities of the blind.

C. Admin

Final module in Third Eye is an admin module. This is also a web interface. Admin controls the entire system. This module has access to all the information about caretaker as well as blinds registered under each caretaker. He also has the facility to track the blinds. He can view all the caretakers.

IV. CONCLUSION

Third Eye proposes an android application that can assist blind in doing their day to day activities like safe navigation, independent shopping, identification and recognition of people and objects, detecting text, etc. This is a low cost application that can be affordable. This is because the system operates with the help of smartphones. Advancement in machine learning and deep learning has made possible for visually impaired people to live like normal people. OCR technology is used for detecting text and tensorflow is used for detecting objects and recognizing people. Detected objects and text are converted to speech and is provided as an audio output to blind via a Bluetooth headphone.

REFERENCES

[1] R. Tapu, B. Mocanu and E. Tapu, "A survey on wearable devices used to assist the visual impaired user navigation in outdoor environments," 2014 11th International Symposium on Electronics and Telecommunications (ISETC), Timisoara, 2014, pp. 1-4.

[2] N. Giudice and G. Legge, Blind navigation and the role of technology, in The Engineering Handbook of Smart Technology for Aging, Disability, and Independence, A. A. Helal, M. Mokhtari, and B. Abdulrazak, Eds. Hoboken, NJ, USA: Wiley, 2008.

[3] B. Epshtein, E. Ofek, and Y. Wexler, Detecting text in natural scenes with stroke width transform, in Proc. Comput. Vision Pattern Recognit., 2010, pp. 29632970.

[4] Sudharshan Duth P and Swathi Raj, "Object Recognition in Images using Convolutional Neural Network," 2nd International Conference on Inventive Systems and Control (ICISC), 2018 IEEE.

[5] S. Cloix, V. Weiss, G. Bologna, T. Pun and D. Hasler, "Obstacle and planar object detection using sparse 3D information for a smart walker," International Conference on Computer Vision Theory and Applications (VISAPP), 2014, pp. 292-298.

[6] Samruddhi Deshpande and Revati Shriram, "Real Time Text Detection and Recognition on Hand Held Objects to Assist Blind People," International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), 2016 IEEE.

[7] R. Manduchi, Mobile vision as assistive technology for the blind: An experimental study. In. ICCHP' 2012.

[8] Pranab Gajanan Bhat, Deepak Kumar Rout, Badri Narayan Subudhi, and T. Veerakumar, "Vision Sensory Substitution to Aid the Blind in Reading and Object Recognition," Fourth International Conference on Image Information Processing (ICIIP), 2017 IEEE.

[9] Runze Chen, Zhanhong Tian, Hailun Liu, Fang Zhao, Shuai Zhang, Haobo Liu, "Construction of a Voice Driven Life Assistant System for Visually Impaired People," International Conference on Artificial Intelligence and Big Data (ICAIBD), 2018 IEEE.

[10] Bogdan Mocanu, Ruxandra Tapu and Titus Zaharia, "Seeing without sight – An automatic cognition system dedicated to blind and visually impaired people," 2018 IEEE.

[11] https://en.wikipedia.org/wiki/TensorFlow [Accessed : 2019-02-15]

[12] https://en.wikipedia.org/wiki/Geo-fence [Accessed : 2019-02-15]