

Assistance System for Visually Impaired using AI

Ms. Kavya. S
Scholar
Yenepoya Institute of Technology,
Moodbidri

Ms. Swathi
Scholar
Yenepoya Institute of Technology,
Moodbidri

Mrs. Mimitha Shetty
Assistant Professor, Dept. of CSE
Yenepoya Institute of Technology,
Moodbidri

Abstract— in today's advanced hi-tech world, the need of independent living is recognized in case of visually impaired people who are facing main problem of social restrictiveness. Due to lack of necessary information in the surrounding environment visually impaired people face problems and are at disadvantage since visual information is what they lack the most. With the help of the advanced technology, the visually impaired can be supported. The idea is implemented through Android mobile app that focuses on voice assistant, image recognition, currency recognition, e-book, chat bot etc. The app is capable to assist using voice command to recognize objects in the surrounding, do text analysis to recognize the text in the hard copy document. It may be the effective way blind people will interact with other people and may help blind people independent life.

I. INTRODUCTION

Visually impaired are the ones who are completely or partially blind. According to an estimate made by the World Health Organization(WHO) 285 million of the population suffer from visual impairment and 39 people were blind and approximately 3% of all the ages in a nation are visually impaired.

The leading causes of blindness are cataract, trachoma, glaucoma and also deficiency like Vitamin A, onchocerciasis, and leprosy. People who are visually impaired suffer a lot and face great challenges in their day to day life for instance finding their way and directions and to places which they do not visit often.

Visual aids such as glasses cannot be a cure or improvement for blindness. While medicines fail to restore the sight of the visually impaired, assistive technologies help them to carry on with their everyday life and also improves their quality of life. A blind person cannot feel that emotion what a person who can see the world feels. Billions of people around the entire face this visibility problem and this is a black dot. With the help Artificial Intelligence and Machine Learning we have aimed at removing this black dot.

The severe consequences visual impairment presents on certain capabilities related to visual function:

- i. Day to day activities(requires average distance vision)
- ii. Communicating, reading, writing(requires a precise vision)
- iii. Estimating area and displacement(requires far vision)

- iv. Extended care of optical observation is needed for tracking activities.

Speech synthesis is used in the existing system for the visually impaired to read e-books using a mobile application which converts the document/soft copy of the books to the speech using natural language and Text-to-Speech [1].

The problem with the existing system is that it works for single language (English) only, and is not compatible with any other languages apart from English.

It cannot be used offline which means that it needs to be connected to the internet for any kind of feedback or responses.

The visually impaired are assisted by voice commands. The existing system using Artificial System not only assist the visually impaired through voice commands but also does image recognition of the photographs clicked or uses camera to recognize the objects and describes them in audio and using a chat bot light and friendly conversations are made.

II. RELATED WORK

An object recognition system for visually impaired people [2]. This system will be a boon to the visually impaired person as well as society. It helps in the detecting the direction of maximum brightness and major colors. A hybrid algorithm is proposed for object recognition in which the Artificial Neural Networks and Euclidean Distance measures are used in combination.

The video is captured with the help of camera and will be categorized into several frames, all the frames are then compared with the previous frames [3] and the data will be stored and response will be given on the basis of stored data about an object.

The first step is edge detection and it is one of the preprocessing steps involved in the process. The boundary between two homogeneous regions may be defined as an Edge. Important points found can be collected from the edges of an image which were detected. To figure out the edges canny operator is used in this project. By computing the derivative of the image intensity function, an edge is typically founded [4].

The user has to select a file and the selected file should be in either .ppt or .doc file format. The file is converted into pdf file format when it is selected and then the text present in the file format is reconstructed as a

collection of words [5]. The information collected from the image is filtered and shown on the output screen.

The technique proposed is based on key points extraction and matching in video. The numbers of frames and database objects are compared and made to detect the object in each frames present [6]. The audio containing the information about the object is activated when an object is found.

Three software tools are developed for Android Smart phones image processing module, a module which detects colors, object as well as light source detector. The algorithm works on the images taken with a flash which is automatic with possibly very small resolutions. The HIS (Hue Saturation Intensity) [7] conversion takes place for the RGB color images.

An image reading tool which read the text present in image. A very important work to be checked in the image is the property of an image and the text. The text present in the picture or image can be of many types; it can be of scene text or Graphics text which depends on the image source as like whether the image is a machine generated image or it is captured from a camera. The color of the image is taken and it gets converted into Grayscale image [8]. To clear the challenges created by noise and uneven lighting in image preprocessing of image is done.

The interest point is found out with the help of local feature extraction method for which a feature vector and descriptor is computed. SIFT the earliest schema [9] for feature extraction. It helps in the representation of that image as a collection of interest points which is not variant to image transformation and partial to illumination changes. The method is based on Gabor filters [10] for texture characterization which is used for variety of texture segmentation and classification tasks.

III. PROPOSED SYSTEM

A. Problem Description

Blind people come across a number of challenges in everyday life from reading a book to walk on the street. Many tools are available to meet the challenges faced by them, but they are not sufficient. The most essentials thing a human can have is vision and it plays a very essential role in the life of a person either a person can see or not. A visually challenged people need an assistant to carry on work on daily basis. In this paper we have discussed the challenges faced by blind people and tried to provide a satisfactory solution to them for working everyday life.

B. System Architecture

Cloud APIs of Google are an integral of Google Cloud Platform, allowing user to easily add the power of everything from storage access to ML based image analysis to the user application. All the Cloud APIs works a simple JSON REST interface and are called directly or via the client libraries. chat-bot client are mainly used in Google Cloud API for the vision as well as Dialog-Flow which is shown in "Fig. 1".

It is used for recognizing the speech as well as translating the textual documents. A Web-Hook is a HTTP callback: a HTTP POST that happens when some request is

made a direct notification is sent by means of HTTP POST. When request is made a web application running WebHook triggers a message.

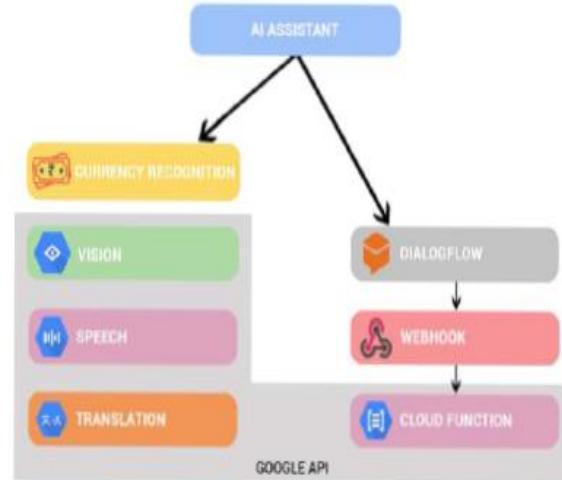


Fig. 1 System Architecture

C. Modules

We all know the importance of vision. A person who has lost vision feels incomplete. A lot of challenges have to be faced when he/she has no vision. The system proposed handles the difficulties faced by the visually challenged people. Some of the problems faced by the visually challenged people are tried to be solved with the help of Artificial Intelligence [11] and Machine learning. Natural and rich conversational experiences are building using the dialog-flow platform. A chat-bot is assisted with Google Assistant, which helps the visually impaired to talk and get the desired response which is shown in "Fig. 2".

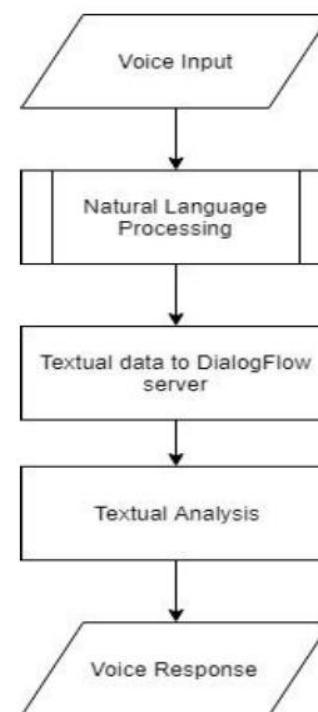


Fig. 2: Flowchart of Chat Bot

Chat-bot takes input from the user in the form of voice or text and responds back with satisfactory reply. Through voice or text a person can ask the chat-bot to guide the person to know about a place or if he/she has to visit the place, then the chat guides the map with voice assistant.

The chat-bot is trained by using Dialog-Flow platform so that a person can interact with it. Keywords that are given by the user are used, and are compared with training data, and response is given back accordingly.

Cloud Vision API which encapsulates powerful machine learning provides The REST API that is used to analyze the image that is captured. Pictures are grouped rapidly into many classifications.(e.g., "Taj Mahal", "Deer", "Footwear") identifies singular protests and face inside pictures and finds and reads printed words, which is contained inside the pictures and is shown in "Fig. 3".

This stage is used to differentiate the picture captured by mobile camera. After examining the picture, the insights about the picture is reacted back.

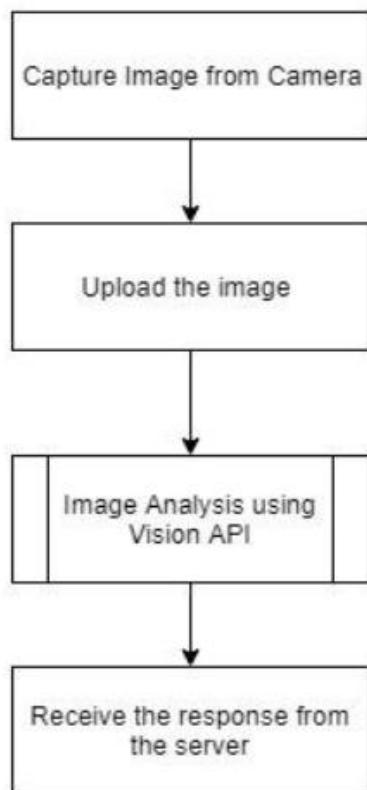


Fig. 3 Flowchart for working of Vision API

As like on the off chance that somebody clicks a picture of Taj Mahal, Google Vision API processes the picture and reacts back with the coveted information in either voice or printed information. API the captured images can be categorized in following classes as like Landmark Detection, Logo Detection, Explicit Content Detection, Image Properties, Label Detection and Document Text Detection in Google Vision.

With the application of neural system an effective model can be developed in a very simple way to use an API. To change the voice input to text the Cloud Speech API helps

the designer. 100 plus variations are provided in API and dialects to help the customer all over the world. You can decipher the content of clients can be deciphered directing to an application's mic, and it is controlled by voice command, or interpret sound documents, among many other cases. The Google Cloud Speech API is used in the proposed framework to change over the voice contribution of the client into literary information and printed information can be sent to the informing android application such as Message application, WhatsApp and so on.

The ID3 (Iterative Dichotomiser 3) algorithm was developed by Ross Quinlan which is used to produce a decision tree from dataset as given in algorithm 1. As a part of machine learning and natural language preparing spaces ID3 is the antecedent to the C4.5 that is being utilized.

```

id3(exp, characteristics)
  Begin
    /*exp are the training examples and the characteristics
    is a list of attributes that will be tested by the trained
    decision tree. It returns a tree that classifies the given
    exp. This algorithm uses to build the AI chat bot to give
    appropriate reply to the user based on the key words
    present in the question and those key words are searched in
    the trained dataset to produce the appropriate output in the
    sound format */
    n = dTNode(exp)
    // handle target characteristics with characteristics
    labels
    dict = sumUpExp(exp, tAttr)
    for key in dict:
      Begin
        if dict [key] == total no. of exp
          n.label = key
          return n
      End
      //test for no. of exp to avoid overfitting
      if characteristics are empty or no. of exp < min per
      branch:
        n.label = most common value in exp
        return n
      bestAtr = characteristics with the most infoGain
      n.decision = bestAtr
      for each possible value v of bestAtr:
        Begin
          subset = the subset of exp that have value V for
          bestAtr
          if subset is not empty:
            n.addBranch(id3(subset,tAttr,attrbestA))
        End
        return n
      End
    Algorithm 1: Decision Tree Algorithm
  
```

As a part of DialogFlow the accompanying algorithm is utilized to settle on a choice as indicated by the client voice summon and the reaction is classified based on enter component in a sentence, then the key elements are

searched using decision tree algorithm in the trained dataset. Information gain IG (A) is the measure of the distinction in entropy from before to after the set S is part on an attribute A which is given in Algorithm 2.

```

infoGain(exp, characteristics, eOfSet)
Begin
  g = eOfSet
  for value in characteristics  Values(exp,
characteristics):
    Begin
      sub = subset(exp, characteristics, value)
      gain-=(no. in sub)/(total no. of exp)* entropy(sub)
    End
    return g
End

```

Algorithm 2: Information Gain

IV. RESULTS AND DISCUSSION



Fig. 4 Chat Bot GUI

The output that is obtained from the android application is shown in "Fig. 4. The Output shown in the chat-bot is according to the training given to the machine with the help of Dialog-Flow and the responses according to the input given to the machine.

In "Fig. 5" Google Cloud Vision API shows the data in the output screen, after the analysis of the Image captured by the mobile camera. Thus, the analysis of the image is done based on the image captured and the JSON data responds back to the app with different parameters of the image and with the help of confidence score and also based on the retrieved JSON data the app responds with a appropriate result.

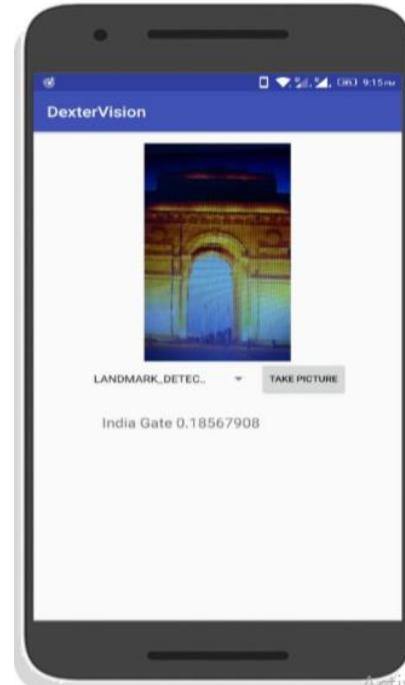


Fig. 5 Image Recognition GUI



Fig. 6 Textual Recognition GUI

In “Fig. 6” the working of the chat-bot is demonstrated and the discussion between the client and application is shown. All the results shown in “Fig. 6” are as per the training given to the machine through Dialog-Flow platform. The machine needs to be trained so that it captures the input keyword and responds with desired output.

V. PERFORMANCE ANALYSIS

The comparison of features between different research papers according to the accuracy of features are shown in “table 1”. The comparison is done between object recognition, landmark detection, textual analysis, interactive, voice input, response and multilingual properties. From the comparison table we can conclude that the proposed system acquires more features than others do.

Features	M 1	M 2	M 3	M 4
Object Recognition	✓	✓	✓	✗
Landmark Detection	✓	✗	✗	✗
Textual Analysis	✓	✗	✗	✓
Interactive	✓	✗	✗	✓
Voice Input & Response	✓	✗	✓	✗
Multilingual	✗	✗	✗	✗

Table 1: Features comparison;

M 1: Personal AI Assistant for Visually Impaired (Proposed); M 2: Visual-Pal: A Mobile App for Object Recognition for the Visually Impaired; M 3: Image Recognition for Visually Impaired People by Sound; M 4: Character Detection and Recognition System for Visually Impaired

The notations used in the graph are as follows; M1: Personal AI Assistant for Visually Impaired; M2: Visual-Pal: A Mobile App for Object Recognition for the Visually Impaired; M3: Image Recognition for Visually Impaired People by Sound; M4: Character Detection and Recognition System for Visually Impaired People.

In “Fig. 7” the accuracy appearance is shown which has a higher percentile. The above bar graph shows the accuracy comparison between our paper and other papers. It is also shown in the chart that our proposed system has more accuracy with respect to other proposed system. Our proposed system M1 has accuracy in between 80 to 90, which is the maximum accuracy among all.

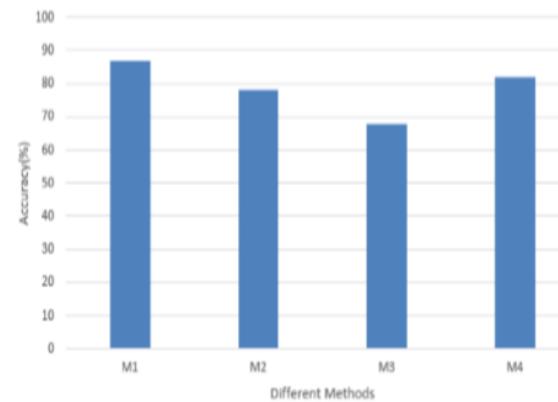


Fig. 7 Accuracy

In “Fig. 8” the sensitivity of M2 is shown which has a higher percentile above all. The chart shows that our proposed system has little less sensitivity with respect to M2 and higher with respect to other proposed system. M1 has sensitivity above 90, which is the second maximum sensitivity among all as shown in the proposed system.

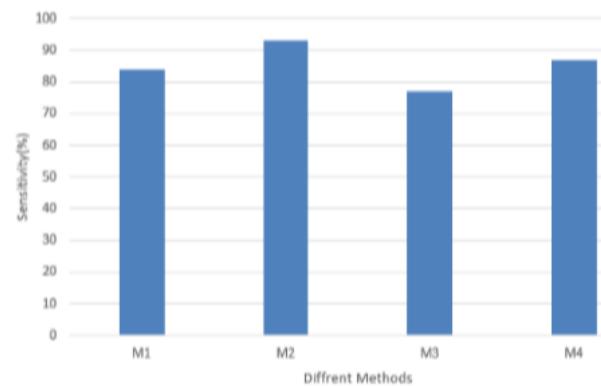


Fig. 8 Sensitivity

In “Fig. 9” the feasibility of M2 is shown which has a higher percentile above all. The chart shows that our proposed system has more optimal solution with respect to other proposed system. The proposed system M1 has feasibility above 90, which is the maximum among all.

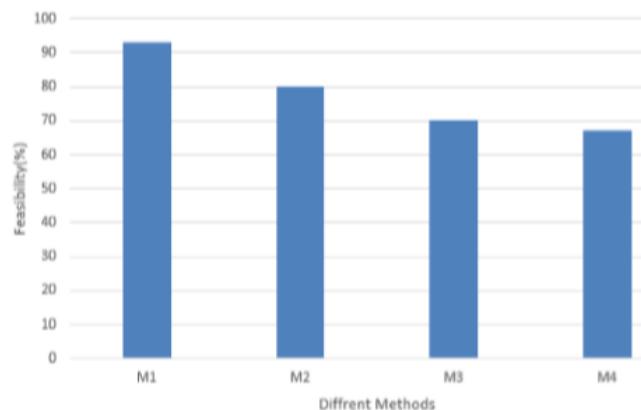


Fig. 9 Feasibility

VI. CONCLUSION

The technologies like artificial intelligence and machine learning plays a vital role in the development of the IT sector. We have made use of these technologies for the visually impaired people so that they too can lead a normal and independent life like other people. The friendly chat bot helps the visually challenged to recognize the objects and surroundings. Currency recognition helps in easy payment. Text recognition helps in reading and analyzing text. The development of the proposed system if is completed, it can serve the visually challenged people with a better assistant.

The proposed system could be applied in multilingual application in the coming days so that a person will be able to use their application in their own language without any trouble. In addition, our proposed system can be deployed with the IoT.

In future the proposed system will be able interpret the textual description in a much better way. The Image recognition can be enhanced with much more details about the image captured through the camera. Enhancement to this system can be done by adding the features of currency recognition [12]. The existing methodology for image and currency recognition can be done with more accuracy.

ACKNOWLEDGMENT

I wish to express my sincere gratitude to Prof. Guruprasad, HOD Department of Computer Science & Engineering Yenepoya Institute of Technology, Moodbidri, for providing me an opportunity to do my paper work on "*Assistance system for visually impaired using AI*". I sincerely thank my guide , Prof. Mimitha Shetty Department of Information Science & Engineering for her valuable help and guidance in carrying out this paper work.

REFERENCES

- [1] Aatisha Cyril, Shubham Melvin Felix, L. Mary Gladence, "Text Reader for Blind: Text-To-Speech", International Journal of Pure and Applied Mathematics Volume 117 No. 21, 119-125, 2017.
- [2] Shagufta Md.Rafique Bagwan, Prof. L.J.Sankpal," VisualPal: A Mobile App for Object Recognition for the Visually Impaired", IEEE International Conference on Computer, Communication and Control (IC4-2015).
- [3] Hanen Jabboun, Faouzi Benzarti, Hamid Amiri, "Object recognition for blind people based on features extraction", IEEE IPAS'14: International Image Processing Applications and Systems Conference 2014.
- [4] K. Matusiak, P. Skulimowski and P. Strumiáão," Object recognition in a mobile phone application for visually impaired users", Lodz University of Technology, Lodz, Poland.
- [5] Shahed Anzarus Sabab, Md. Hamjajul Ashmafee, "Blind Reader: An IntelligentAssistant for Blind", 19th International Conference on Computer and Information Technology, December 18-20, 2016, North South University, Dhaka, Bangladesh.
- [6] Hanen Jabboun, F. Aouzi Benzarti, Hamid Amiri , "Object Detection and Identification for Blind People in Video Scene", Universite de Tunis El Manar, Ecole Nationale d'Ingenieur de Tunis 1002, Tunis Le Belvedere, Tunisie.
- [7] K.Gopala Krishnan, C.M.Porkodi, K.Kanimozhi, "Image Recognition For Visually Impaired People By Sound", International conference on Communication and Signal Processing, April 3-5, 2013, India.
- [8] Akhilesh A. Panchal, Shrugal Varde, M.S. Panse," Character Detection and Recognition System for Visually Impaired People", IEEE International Conference on Recent Trends in Electronics Information Communication Technology, May 20-21, 2016, India.
- [9] Nada N. Saeed, Mohammed A.-M. Salem, Alaa Khamis," AndroidBased Object Recognition for the Visually Impaired", German University in Cairo, Ain Shams University.
- [10] Vincent Gaudissart, Silvio Ferreira, Celine Thillou, Bernard Gosselin, "Mobile Reading Assistant for Blind People", SPECOM'2004: 9th Conference Speech and Computer St. Petersburg, Russia September 2022, 2004.
- [11] N.G.Bourbakis, D. Kavraki, " An Intelligent Assistant for Navigation of Visually Impaired People", 2011.
- [12] Noura A. Semary, Sondos M. Fadl, Magda S. Essa, Ahmed F. Gad," Currency Recognition System for Visually Impaired: Egyptian Banknote as a Study Case", Menoufia, Egypt, 2015