IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

AI Powered Glasses Sync-in With Android Application For Visually Impaired Person

Authors: Anirudh Malik, Ashwin Punyani, Nishit Kumar, Prithvi Bhadra Co-Author: Prof. Dr. Debirupa Hore

Abstract

This dissertation presents a smart glass which coordinates with an android application to perform object detection and help the visually impaired. It's an innovative symbiotic relationship of an android application and a smart glass, in which they would interact with each other to give out a prediction of objects in front of the visually impaired person hereby helping her/him to overcome this disability and feel a bit more comfortable and familiar with the surroundings without enabling human touch. Libraries like OpenCV, Tesseract and Yolo were inculcated to perform and give out precise results. Glaucoma, macular degeneration, and diabetic retinopathy are some of the visual conditions which can take more advantage with the same.

Keywords: *Object detection, smart glasses, wearable technology*

I. INTRODUCTION

Considering all the technology developments today, it gives us immense pleasure to go ahead with object detection and intelligent machines. These technical advancements since their innovation have always headed towards the development and ease of life for humans. Surrounded by all types of technologies, we can conclude how comfortable having different machines makes us. However, humans are still trying to make new progressions in technology by innovations and making the innovation affordable to the people. Like this methodology, we have tried to make our product 'smart glasses' as innovative and as affordable as possible to give these types of glasses a little more reach to ordinary people. In this paper, one can find a discrete understanding of our product, using an app exclusively created by us to connect to the glasses through a Bluetooth module. The paper also consists of the essential use of object detection using APIs. One also see our vision of congregating all the hardware and the interconnection between the hardware and the software. While working on the project, we have had the fortune of having a hands-on experience exploring multiple types of design that align with our requirement of space for the hardware but concurrently keeping in mind about keeping the design as elegant as possible. We have also exploited the wide range of applications and benefits of how wearable technology could be a new trend in future.

II. A REVIEW OF RELATED WORKS

Recent times have been fruitful for the visually impaired since many studies and enhancement in technology have happened to help them live a better and more comfortable life. Listed below are a few such successful attempts:

1.1 The camera app by Microsoft

It is just another VR glasses with functionalities such as massive 14X zoom, voice commands to the Iris Vision Assistant to enable a hands-free experience. With OCR (Optical Character Recognition), it can read the text to the reader as in audio format.

1.2 The Iris vision by Samsung

It's just another VR glasses with functionalities such as massive 14X zoom, voice commands to the Iris Vision Assistant to enable a hands-free experience. With OCR it can read text to the reader as in audio format.

1.3 MyEye2 by orcam.com

A detachable camera to the spectacles being innovative in design, myeye2 offers colour detection, automatic page detection gesture reader commanded by pointing fingers, face recognition, a bar code reader, and product identifier.

III. AI GLASSES FOR VISUALLY IMPAIRED PERSON

AI-powered glasses coordinate with the libraries loaded in the android application in the phone, thereby creating a system to roll out results in audio format and bring down the cost of the product since everyone already owns an established smartphone.

1.1 Arduino Nano

This particular microcontroller board designed by Arduino has perks like small and compact size with enormous applications

Figure 1. Arduino Nano

In the Bluetooth module, the ESP camera would interact with Nano, and as a result of this, it will interact with the android application for image processing and object detection. Programming for the same would be done by Arduino IDE software which is compatible with the Nano

1.2 ESP camera module

This particular microcontroller board designed by Arduino has perks like small and compact size with enormous applications. In the Bluetooth module, the ESP camera would interact with Nano, and as a result of this, it will interact with the android application for image processing and object detection. Programming for the same would be done by Arduino IDE software which is compatible with the Nano. It is a full-featured microcontroller that also has an integrated video camera and microSD card socket. It is accurate, distinct and affordable for advanced applications like Image

tracking and recognition. Arduino IDE can program the same.

1.3 Power supply

Figure 3. Battey

3.7V 300 Mah Battery which is rechargeable.

1.4 Mobile Android Application

The Designed application by our Engineers in flutter SDK works on the support language of Dart Programming language. Google backs flutter, so it has more capability like inbuilt widgets and a more well-designed SDK architecture. We have also used the Flutter Bluetooth package to connect the device as Arduino Bluetooth as a slave and Phone Bluetooth as a master to receive data and produce text to speech commands. The device as Arduino Bluetooth as slave and Phone Bluetooth as a master to receive data and produce text to speech commands.

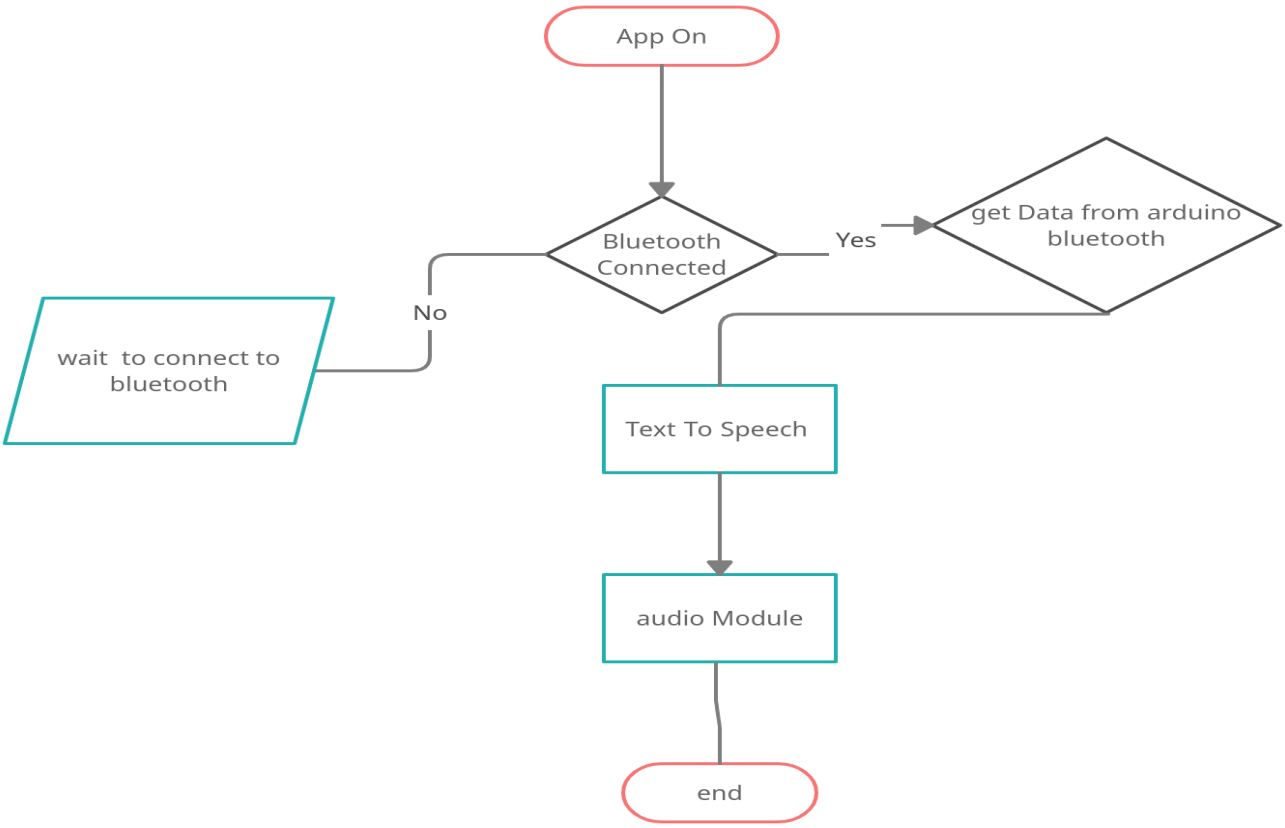

Figure 4. System design

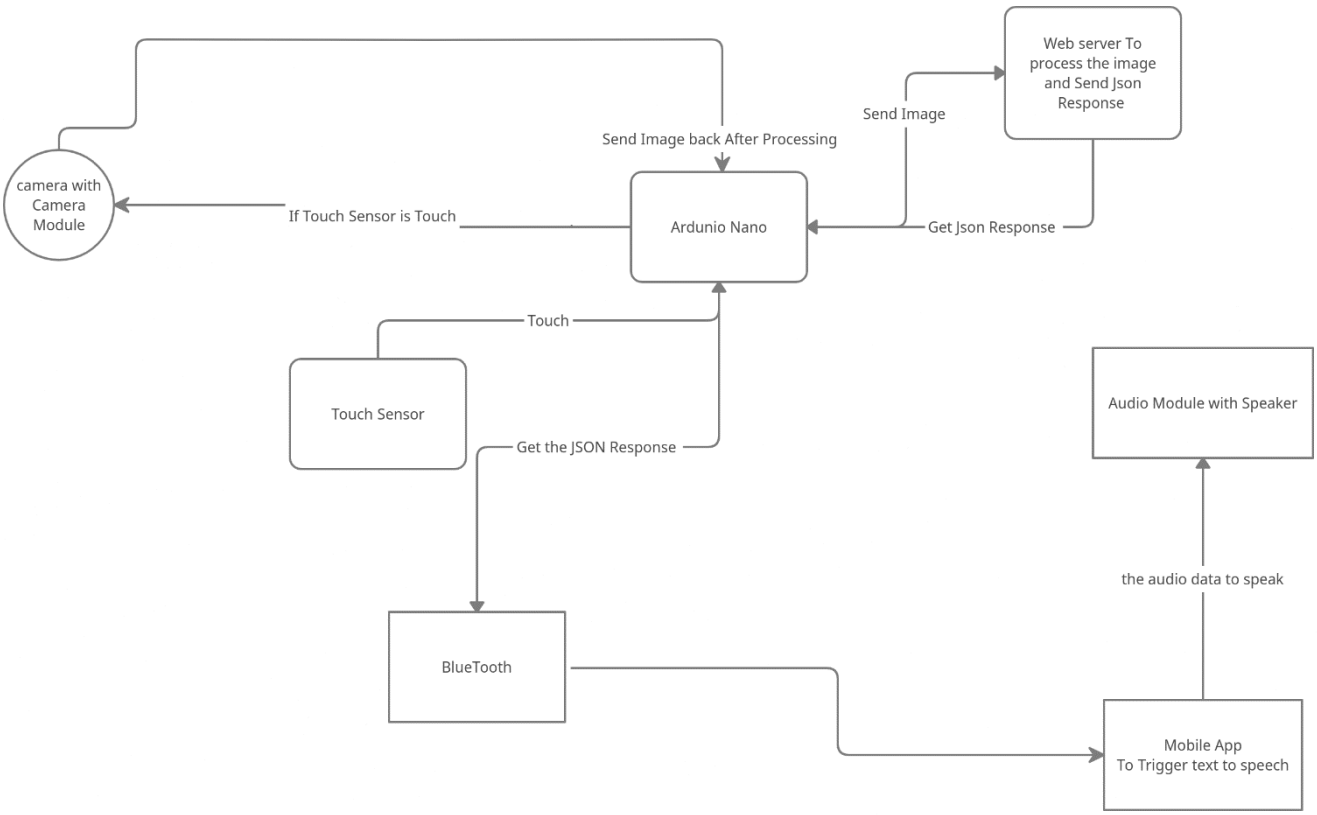

Figure 5. Workflow diagram

1.5 APPROACH

The following procedures explicitly express how the smart glasses would perform specific tasks and how exactly the results are acquired to fulfil the project's requirements.

1.5.1 Face Recognition

A dataset of known/recognized faces is accumulated. Interpreting this dataset in our program calculation of the Eigenfaces from the dataset (Training set), keeping solely m pictures that correspond to highest eigenvalues. These m pictures outline the face area. Calculation of corresponding distribution in m-dimensional weight area for every renowned individual by protruding their face pictures in the face area. Loading any image format (.bmp, .jpg, .png) from given supply. Determining the faces available within the pictures exploitation the Cascade Classifier in OpenCV. Face Recognizer technology is adopted for face recognition. Calculation of a set of weights supported the input image. Therefore, the m eigenfaces by protruding the input image onto every Eigen's face, validating if the image is understood or unknown by checking to ascertain if the image is sufficiently on the point of the face area. Text to Speech conversion exploitation renowned description of person and output through the headphone.

1.5.2 Object recognition

Preparation of a dataset of known/recognized objects and then the dataset was read in the program. Loading any image format (.bmp, .jpg, .png) from a given supply. Then objects were found available within the pictures exploitation Edge detection, line detection, and pattern detection algorithms. Extracted objects square measure compared with the recognized dataset. Call is predicated on this comparison if the body is recognized or not. Once the body is recognized its individual description is extracted. Text to Speech conversion exploitation renowned description of object and output is provided through headphones.

1.5.3 Text-To-Speech

Text-To-Speech Loaded image of any format (BMP, jpg, png) from the given supply. Conversion of the image to grayscale and binaries exploitation the edge price. Finding image options like resolution and inversion. So conversion of the photo into a straightened image for the additional process. Lines detection and removal process takes place. This step is needed to boost page layout analysis, attain higher recognition quality for underlined text, notice tables, etc. Page layout analysis. During this step, we tend to try to spot the text zones available within the image. So solely that portion is employed for recognition, and the remainder of the region is neglected—detection of text lines and words. Here we tend to conjointly ought to look out of various font sizes and tiny areas between words. Recognition of characters. This is often the best algorithm; a picture of each character should be transformed into an acceptable character code. Saving results to choose the output format, such as searchable PDF, DOC, RTF, and TXT.Text to Speech conversion exploitation Google API and output through headphones.API's and output through headphones.

V. RESULT AND DISCUSSION

The images were captured by the camera and transferred to the application to run through the modules. Thereby the results were calculated and presented via earphones by the application on the phone to the user. Here are some results of the same:

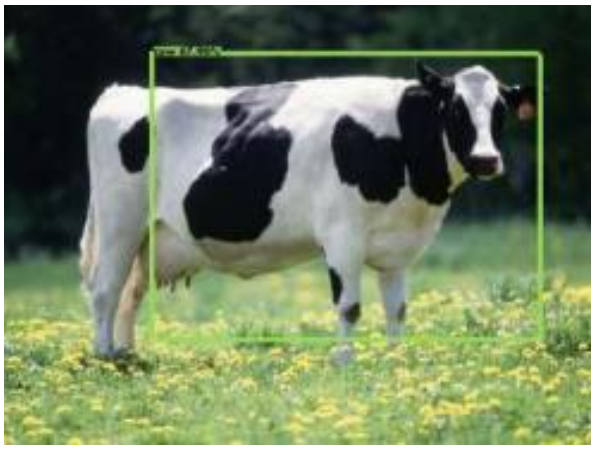

Figure A: COW

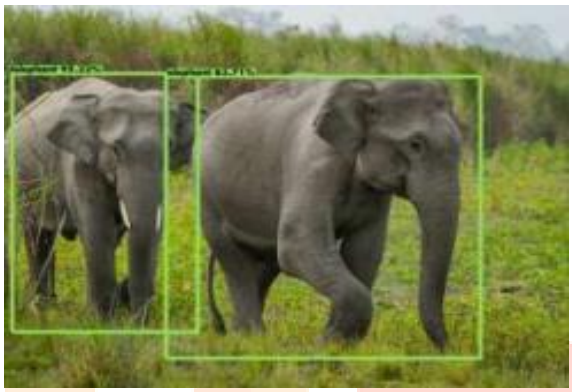

Figure B: Elephant

Figure C: Group of People

Object	Accuracy	Result
Cow	97.96%	Cow detected
Elephant	95.22%	Elephant detected
Faces	86.77%	Face recognized

Table 1: Result

System deficiencies	Remediation
One drawback of this system is it is comparatively slow.	By using another faster algorithm and highly efficient processor, speed can be improved.
Limited objects detection possible due to Arudino nano limited efficiency.	The highly efficient processor can solve this problem.

Table 2: System deficiencies and remediation

VI. CONCLUSION

The results given in this report emphasizes the possibility of using ML models and algorithms for object detection with the help of a mobile application. With the advancement in specifications of smart phones, the trained models are feasible for use in a real-time environment with ease without any delays. A diversified training data could increase the performance and accuracy, as more images from the environment typical to the end use case could increase the performance of the less complex models when detecting multiple notes, as these models might have been limited by the quality of the training data. The proof-of-concept version can be used in a production environment. The delay in responses and speed of results could be enhanced using a better processor and a better server. The object detection procedure was carried out decently well. Areas of improvement could be in enhancing the processor for multiple object detection.

REFERENCES

[1] ‘Steven Richardson and Thomas Abrams . ‘Assistive spectacles: A vision for the future’ Springer Nature Limited 2020

[2] Feng Lan, Guangtao Zhai and Wei Lin, "Lightweight smart glass system with audio aid for visually impaired people", TENCON IEEE Region 10 Conference, 2015.

[3] Bing Xiao, Hang Wong and Kwen L Yeung "Penta-band Dual-fed Smart Glasses IoT Antenna", IEEE , 2020 14th European Conference on Antennas and Propagation (EuCAP) [4] https://ieeexplore.ieee.org/abstract/document/8117194

[5] ASM Iftekhar Anam, Sahinur Alam and Md Yeasin, "A dyadic conversation aid using google glass for people who are blind or visually impaired", Mobile Computing Applications and Services(MobiCASE)6th International Conference, 2015. [6] https://www.tensorflow.org/tutorials

[6] Yuhang Zhao, Elizabeth Kupferstein, Hathaitorn Renirun, Riah Findlater and Shiri Azenkot "The Effectiveness of Visual and Audio Wayfinding Guidance on Smartglasses for People with Low Vision" , CHI '20: Proceedings of the 2020 CHI Conference on Human Factors in Computing SystemsApril 2020