Team Name - InSight

Title - Vision For Blind

We are 2nd Year Undergraduates from IIT(ISM) Dhanbad.

Members

Priyansh Sharma, Ishal Garg, Pratham Pahariya, Aditya Kumar Garg, Bhavesh Sharma.

Problem : It is estimated that around 285 million people in the world are visually impaired, out of which 39 million are completely blind. While in India, nearly 40 million people including 1.6 million children, are blind or visually impaired due to uncorrected refractive error. These people face many difficulties while performing their daily life tasks. They are prone to road accidents and are not aware of any danger that might be present around them. Sometimes they find it very difficult to understand their nearby surrounding situations and require someone else to help them. Also as they are blind they are not able to read and understand any type of written text, bill, currency denomination, and other such things. They also face difficulties in reaching their destination properly and at times they detour from their actual path.

Idea: Since the problems faced by visually impaired people are numerous and most current available solutions are either too obsolete or too expensive to adopt. Hence, the idea behind this project was to come up with smart AI-based solutions to help visually impaired people lead normal lives without spending much money, so we have come up with an idea of making Visual Support Glasses that will help visually impaired people live a better life, mainly we have focused the glasses on two core areas which are reading and navigation, as these are the most common and impactful things where visually impaired people struggle the most. These glasses will give an audio description of the things around them, assist them in reading, and also help in their other daily life work. So the basic aim of our product is to make these people's lives easier by assisting them in their day-to-day tasks and ultimately reducing their dependence on others. All these will be done using Computer Vision and deployed on an android based mobile application which will be connected to the Visual Support Glasses through a wifi module.

Features:

All these features will come with multi-language support. Navigation: The smart glasses will tell the user by audible commands on how to reach their destination, it

will tell them either to move straight or to turn back, to turn left/right, and also how many steps they should take until they reach a specific location.

Capture & Read: The smart glasses can be able to read and identify the words on a piece of paper or a signboard in real-time, and so it will basically work like an audiobook transmitting the words to the person along with it the ebooks could also be converted to audio format.

Explore Nearby Environment: Device will identify the nearby objects, and classify them in the speed in which it is approaching (very high speed, slow, etc.) so as to provide safety to the visually impaired person.

Color Recognition: It will detect the color of various objects like clothes, etc, and also the color of traffic lights so as to reduce the number of accidents (This feature can be used by color blind people too).

Currency Identifier and Counter - It can detect the currency denomination and can also count them and tell the total value.

Face Recognition - It will detect the gender, age, and mood of the person with whom the user is interacting.

Technology Used: Software (Image processing) The below-given models(supervised learning model) will work on the image/video given captured through the camera. The prediction will be made and the message will be sent in audio/text format when required.

Computer Language: Python(for deep learning) and Java/Kotlin (for android application) Deep Learning Framework: TensorflowLite/PyTorch Image processing library: OpenCV Android Development IDE: Android Studio Others - Google Text to Speech and Speech to Text, Google Maps, MapBox, Solidworks.

1.Optical Character Recognition(OCR) - This model will help to read the text. It will take a picture of what we want to read and convert it into text which it will recite or display according to the user's choice.

2.Object Recognition - We will make a real-time object detection model using deep learning and OpenCV using YOLO Algorithm. This model will enable the user to identify different kinds of objects in front of him. It will also give an alert message if there is an object which is coming on the user's path. We will also include the speed calculating and distance of the object from the camera features in our model. This will help us in alerting the user about speeding vehicles and too close objects in the vicinity of the user.

3.Face Recognition - It will be an integrated model of the YOLO Algorithm for face detection and a simple CNN for facial characteristics. This model will specifically help in detecting people's faces. It will give information regarding the gender, age, and mood of the person.

4.Color Recognition - It will help us in segmenting the color in the image. For example, In the case of a traffic signal, this model will help us in specifying the color of the traffic light. A simple CNN/U-net will help us in getting the task done.

5.Currency Recognition - A simple CNN model with help of OpenCV and Pytorch which help in recognizing the currency type and value. So that it helps the user in an easy and confident way for exchanging currency.

Hardware integration: Visual Support Device which will be worn by Visually impaired people consists of a Camera module, which captures high-resolution visuals, and with the help of ESP8266 Wifi module, it will continuously transport the visuals to our mobile application which will then process the image and send it to the deep learning model which will classify the text/image and then send an audio message which will reach the user through earphones. Our support device will also provide them with navigation support. The device consists of a microphone that will take a vocal input, send it to the mobile application for processing, find a refined path for the person, and will tell the vocal command to the person for navigation. For this navigation purpose, we will use a GPS module.

Deployment on the Mobile App: All the above-mentioned models will be deployed on devices with the help of TensorFlow Lite, which will basically optimize and convert the deep learning models into a flat buffer file, making them easier and convenient to deploy. All the above-mentioned ML models will be converted to APIs through FastAPI and then it will deployed on mobile app.

Further Additions:

If the time permits, we are planning to add certain features in our app that will also help people with other disabilities like hearing and speaking disabilities. For a person with a hearing disability, we will be adding a feature that will convert the voice of some other person into text which can be easily read and understood by that person. In the case of speaking disability, we will add a feature that will detect the sign language of the person with a disability through the camera and then convert it to text or speech so that it can be easily understood by some other person who doesn’t understand sign language, this will minimize the communication problems between the disabled and normal persons