**ABSTRACT**

Computer vision deals with how computers can be made to gain high-level understanding from [digital images](https://en.wikipedia.org/wiki/Digital_image) or [videos](https://en.wikipedia.org/wiki/Video). It seeks to automate tasks that the [human visual system](https://en.wikipedia.org/wiki/Human_visual_system) can do. Humans glance at an image and instantly know what objects are in the image, where they are, and how they interact. An estimate of 285 million people are visually impaired worldwide, stated by WHO. The proposed Blind Sight-Object Detection with Voice Feedback is a computer vision-based application that leverages state of the art object detection techniques. These are employed to detect objects in the vicinity. You Only Look Once (YOLO): Unified, Real-Time ObjectDetection a new approach to object detection is deployed in this proposed work. Yolo has 75 Convolutional Neural Network (CNN). Image classiﬁcation techniques are used to identify the features of the image and categorize them into their appropriate class. The COCO dataset used in this project consists of around 123,287 hand labelled images classiﬁed into 80 categories. This wide set of data is used to describe spatial relationships between objects and their location in the environment. The text description of the recognised object will be sent to the Google Text-to-Speech API using the gTTS package. Voice feedback on the 1st frame of each second will be scheduled as an output to help the visually impaired hear what they cannot see.