Database Entry ID: 179

Title: Computer Vision and Speech Synthesis Based Scene Understanding for People with Impaired Vision

Name: Kyriakos Kyriakou

Reg No: 1905770

Location:

Vision impairment is a major health problem affecting millions of people around the globe. Scene Understanding is a challenging task for these people; hence, the project presents a mobile application ("BlindVision"), that uses Computer Vision and Speech Synthesis techniques to help them understand a scene ahead and assist them with their daily tasks.

Deep learning for Object Detection and Image Classification has been combined with Optical Character Recognition, to give an in-depth understanding. Different CNN object detection models have been tested and evaluated including YOLOv4, MobileNetV1 SSD, and EfficientDet Lite0/1/2/3, all trained with the COCO dataset, and Google's ML Kit was used for the image labelling and text recognition which uses the free Google Vision API. The app can recognize specific text, people, objects, and give audio explanation of the environment by using Text-To-Speech tools. Using Speech Recognition, the system can understand the user's commands by voice. Finally, system has built-in gestures to make the interaction with the user hands-free.

Testing and evaluation on the different object detection models showed that EfficientDet Lite0 achieves the overall best performance. The accuracy of the model on the device is 67%. Along with the image classification model, which achieves accuracy of 85%, and text recognition, the performance of the app while running, is 14.4 frames per second.

New features could be introduced, that will make the interaction with the users better, new object detection models could be evaluated, and lastly, system could make use of the Google's Cloud Vision APIs for further understanding.