

National Conference and Exhibition on Rural Innovations

DETECTING OBJECT AND TEXT- CONVERTING TO SPEECH

N. Padmapriya^a, G. R. Anantha Nivethan^b, V. Guru^c

^a Assistant Professor, Department of CSE, SSM Institute of Engineering and Technology, Dindigul, India.

^{b,c} UG Student, Department of CSE, SSM Institute of Engineering and Technology, Dindigul, India.

Corresponding Author Name & Email id: N. Padmapriya & priyapadma31@gmail.com

Abstract

According to estimates by the World Health Organization, about 285 million people suffer from some kind of visual disability, of whom 39 million are blind, resulting in 0.7% of the world population. Blindness is often used to describe some form of visual impairments or vision loss. There is a great dependency for the blind people navigation or walking in any unfamiliar area. They depend on any persons to help them or they use their natural senses such as touch or sound for identification or Navigation. Most learning and recognition of objects around us is accomplished using the eye. This device, the virtual smart glass assists them in their ways without the need of human help and help them walk independently. The device is installed on the glass frame and these glasses help to figure out the surroundings. The object detection technology is given as input to the microcontroller, to find out if there is any obstacle in front of it. If there is an obstacle, the device will produce an audio and alert the user. The object detection methodologies can be useful for detecting the object in their navigation path. We also going to recognize text captured from an video and convert it to speech. Text detection is achieved using the OpenCV software and open source Optical Character Recognition (OCR) tools Tesseract and Efficient and Accurate Scene Text Detector (EAST) based on Deep Learning techniques. The recognized text is further processed by Google's Text to Speech (gTTS) API to convert to an audible signal for the user. The recognized video is further processed by MobilenetV3 which is predefined deep learning algorithm in TensorFlow and predict the objects in front of them and converts that into a audio

Keywords---Deep Learning, Object Detection, Text Conversion.



ISBN 978-93-91347-59-8

40

Dr. D. Senthil Kumar, M.E., Ph.D., (NUS)
Principal

SSM Institute of Engineering and Technology
Kuttathupatti Village, Sindalagundu (Po),
Palani Road, Dindigul - 624 002.