ID 13

Computer Vision for Object Detection in Assistive Technologies: A Comparative Review

M Marapperuma[1#] and BVKI Vidanage[2]

[1]Department of Computer Engineering, Faculty of Computing, General Sir John Kotelawala Defence University
[2]Department of Computer Science, Faculty of Computing, General Sir John Kotelawala Defence University

[#]39-bce-0006@kdu.ac.lk

Abstract

This paper aims to identify the major object detection techniques performed on assistive technologies for visually impaired individuals, with a particular focus on the shift from conventional computer vision methodologies to current deep learning frameworks. The most popular object detection models under discussion are single-stage detectors: Single Shot Detector (SSD) and You Only Look Once (YOLO), and two-stage detectors: Faster R-CNN and RetinaNet, all evaluated in terms of performance in per-pixel accuracy on the COCO dataset, expressed through mean Average Precision (mAP) and inference time on GPU. The proposed models are discussed in the context of their practical applicability for assistive purposes by considering problems such as small object recognition, time constraints, fluctuations in the environment, energy demands, limitations of devices, and interfaces. Observations show that single-stage detectors such as the SSD and YOLO can provide faster inference time, which is ideal for real-time application at 22ms and 29ms, though with lower mAP performance than two-stage detectors in research with 23.2 and 33.0, respectively. The two-stage detectors such as Faster R-CNN and RetinaNet are more accurate with 36.2% and 37.8% mAP respectively, but have higher inference times of 200ms and 73ms for real-time assistive tasks. The study also highlights the issues in small and occluded object detection, that can be detected in diverse lighting and weather conditions, and power and hardware constraints of wearable technology. These problems must be addressed by optimizing the model and enhancing the hardware and software of assistive technologies for the visually impaired.

Keywords: *Computer vision, Object detection, SSD, YOLO, Faster R-CNN*