Enhanced Yolo-V3 Deep Learning Algorithm for Real-time Object Detection – An Assistive Model for Visually Challenge People

S. Kiruthika devi[*] and Subalalitha C. N

Department of Computer Science and Engineering, SRM Institute of Science and Technology, Kattankulathur, Chennai

*Corresponding author

Abstract

Background: Visually Challenged persons confront a variety of hurdles in their daily lives while moving in real-time environment independently. They face difficulty in detecting different objects found in both indoor and outdoor surroundings even with the existing assistive devices. This issue stems from the inability of current assistive devices to detect and classify a wide range of multi-scale items against a complex background. The origin of Computer vision technology and deep learning algorithms found to be useful in the detection of objects in real-time environment [1].

With this motive, the deep learning algorithm YOLO-V3 a popular and fastest object identification algorithm [2] can be employed for object detection to assist the visually challenged people. In spite of the good object detection speed, the YOLO-V3 detect the objects with low accuracy, which may not be suitable for guiding the visually challenged people.

Objective: The objective of this paper is to propose an enhanced YOLO-v3 algorithm by adopting Dense-Net blocks, secondly the number and aspect ratio of anchor boxes are formulated mathematically instead of constant value, finally detections are done at four different scales instead of three scale levels to improve the performance.

Methodology: The adoption of Dense-Net blocks after every residual block in YOLO-V3 with darknet-53 as a back bone, will increase the feature extraction capability [3], to increase the accuracy of object detection. The generated anchor boxes will result in better detection of objects with maximal Intersection over Union score. The detection of objects at different scales [4], facilitate the model to detect even the tiny objects.

Result and discussion: To evaluate performance of our enhanced YOLO-V3, a set of experimental analysis are carried out on PASCAL VOC and KITTI data set. The proposed, enhanced yolo-v3 model have attained the mean Average Precision(map) of 87.53% and 85.26% on PASCAL VOC and KITTI data set respectively, which are 8.25% and 1.26% more accuracy than the original YOLOv3 model. Thus, our proposed model can be incorporated in assistive device for visually challenged people.

Future Work: As a future scope, aimed to reduce the computational cost without compromising an accuracy.

References

[1] Rachburee, Nachirat, and Wattana Punlumjeak. "An assistive model of obstacle detection based on deep learning: YOLOv3 for visually impaired people." International Journal of Electrical & Computer Engineering (2088-8708) 11.4 (2021).

[2] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

[3] Ju, Moran, et al. "The application of improved YOLO V3 in multi-scale target detection." Applied Sciences 9.18 (2019): 3775.

[4] Zhang, Xinliang, et al. "An improved YOLOv3 model based on skipping connections and spatial pyramid pooling." Systems Science & Control Engineering 9.sup1 (2021): 142-149.