Perspectives in Communication, Embedded-Systems and Signal-Processing (PiCES) – An International Journal
ISSN: 2566-932X, Vol. 5, Issue 2, May 2021

A Comparative Study of Real-time Object Detection Systems for Navigation of the Visually Impaired

B R Karthik
Department of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India, br.karthik1@gmail.com

Manish N
Department of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India, manish.nagaraj12@gmail.com

Adithya Krishna V
Department of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India, aadithya794@gmail.com

Y V Sai Keerthana
Department of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India, keeru.y2000@gmail.com

Dr. S Prabhanjan
Department of Computer Science & Engineering, Jyothy Institute of Technology, Bangalore, India, hod.cse@jyothyit.ac.in

Archit Krishna K
Indian Institute of Technology, Kanpur, India, achitkk@iitk.ac.in

Varun C Shekar
Massey University, New Zealand, varun_shekar@yahoo.in

Abstract: ***The visually impaired face a plethora of problems. The primary problem they face is navigating from one place to another. The detection of obstacles in the user's proximity is another challenge that needs to be addressed. This paper provides a comparative study of various real-time image recognition and object detection methods that might help develop effective navigation systems for the visually impaired.***

Keywords: Real Time Object Detection; Machine Learning; Image Processing; Computer Vision

I. INTRODUCTION

Visual impairment is a major concern to the world with its rising nature. Recent estimates about the number of visually impaired people around the world are 285 million out of which 39 million are blind and 246 million suffer from low vision. One of the primary problems that the visually impaired face is navigation. To deal with this, a cane is the conventionally used aid. With advancements in assistive technologies, new tools are being developed which guide the visually impaired not only with mobility but also with a better understanding of the surroundings. With the likes of ever-growing technology, this industry has seen an all-time high in the recent past.

Artificial Intelligence is the reflection of human intelligence tasks by computers. Digital image processing defined as the use of algorithms to process data in images and extract data from them Artificial intelligence and Machine Learning algorithms of today are not just automating regular, monotonous, and tedious tasks, but also unlocking opportunities to people with disabilities and promising unique ways of experiencing the world. By blending these two technologies with Image Processing, critical problems can be solved. Powerful assistive solutions can be built which not only empowers people with disabilities but also enables them to look at life in a different perspective.

Object detection being one of the widely used applications in the field of computer vision, has been the popular area of research. Convolution neural network has made significant advancement in detection of objects and image classification.

This paper presents the study of various real-time object detection methods that have been developed. The results of the comparison will help the engineers to find out the most feasible method to be used to develop a real-time navigation system for the visually impaired.

II. LITERATURE REVIEW

The paper size is A4 (21cm x 29.7cm). The details of your work have to be written considering the defined specifications of the paper. Please do not add anything as header, footer or watermark to the manuscript that the authors submit to the journal for publication. The paper follows the two column format. The headings of sections have to centre aligned.

SqueezeDet: Unified, Small, Low Power Fully Convolutional Neural Networks for Real-Time Object Detection for Autonomous Driving[1] by Wu B, Iandola F, Jin PH, Keutzer K, proposed a system for Self-Driving cars. It uses Fully Convolutional Neural Networks for

Real-Time Object Detection. In this network, they used convolutional layers not only to extract feature maps but also as the output layer to compute bounding boxes and class probabilities. The detection pipeline of this model only contains a single layer of a neural network, thus it claims to be extremely fast.

Wang RJ, Li X, Ling CX.in Pelee: A Real-Time Object Detection System on Mobile Devices[2] proposed a system for Real-Time Object Detection System built for Mobile Devices. It is designed to meet strict constraints on memory and computational budget .

It uses similar architecture as DenseNet and has an efficient Convolutional Neural Network (CNN) architecture built with conventional convolution. Compared to other efficient architectures, Pelee has a great speed advantage and easy to be applied to the computer vision tasks.

A Fast Feature Extraction Algorithm for Image and Video Processing[3] by Abdulhussain, S. H., Rahman Ramli, A., Mahmmod, B. M., Iqbal Saripan, M., Al-Haddad, S. A. R., Baker, T., Jassim, W. A. This paper introduces a novel method to compute features from images or video frames. The captured features are used to represent the local visual content of images and video frames. It uses a content based video indexing and retrieval method. The proposed system was found to be more accurate and faster than any standard feature extraction techniques to the date. The proposed method is employed for shot boundary detection applications to detect transitions in video frames.

You only look once: Unified, real-time object detection[4] by Redmon J, Divvala S, Girshick R, Farhadi A, proposed a system which re-frames object detection as a single regression problem, straight from image pixels to bounding box coordinates and class probabilities. You only look once (YOLO) at an image to predict what objects are present and where they are and a result of this, it makes way for faster detection of objects compared to other existing methods like R-CNN and Fast R-CNN. The system can detect objects in real time at 25 fps and the fast version of the system can perform the same at 150 fps.

Digital Image Processing Techniques for Object Detection From Complex Background Image[5] by Hussin R, Juhari MR, Kang NW, Ismail RC, Kamarudin A. proposed a system that uses a selection of a mango in a mango tree. Firstly the image is passed through a color filter to eliminate the background image. After this, the image is converted to grayscale and then the Circular Hough Transform is applied to detect the circular objects.But, it fails to detect objects correctly in case of poor lighting conditions thus resulting in poor accuracy.

MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications[6] by Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam. This paper introduces a class of efficient models called MobileNets for mobile and embedded vision applications.It's based on a streamlined architecture that uses depth wise separable convolutions to build light weight deep neural networks. It can be deployed as an efficient as well as an effective base network in modern object detection devices. The model is trained on COCOtrain which does not include 8k minival images as well as on Faster-RCNN model evaluating 300 RPN proposal boxes per image and the results are comparable with huge difference in the usage of computational resources and model size.

Region-based Convolutional Networks for Accurate Object Detection and Segmentation[7] by Ross Girshick, Jeff Donahue, Trevor Darrell, Member and Jitendra Malik. This paper proposes an object detection and a segmentation algorithm which uses multi-layer conventional networks to compute highly particular, yet undeviating, features. Then these features are used to classify regions of images which can later be output as bounding boxes or segmentation masks. The proposed approach also has scalability as well as flexibility features with a number of object types.

Video Image Processing for Moving Object Detection and Segmentation using Background Subtraction[8] by Anaswara S Mohan, Resmi R. This paper proposes an algorithm which is based on object detection from the image background using the background subtraction method and segmentation using thresholding and edge detection. The proposed system detects objects through a series of 5 stages input, pre-processing the frame, subtraction of background morphological processing and finally object detection.The proposed method succeeded for detecting moving objects accurately in all video sequences and thus this technique can successfully extract moving objects from various sequences. The comparison is based on the PSNR(Peak Signal to Noise Ratio) values of the sequences and shows ample difference between the two methods and it is found that the background subtraction method is better compared to thresholding technique.

III. COMPARISON OF VARIOUS METHODS OF REAL TIME OBJECT DETECTION

Method	Input Dimension	mAP
SqueezeDet	1242x375	76.7
Pelee	304x304	76.4
Yolo	448x448	63.4
MobileNets	160x160	79.4
RegionBased	227x227	62.4

Table 1. Comparison of Various Methods of Real Time Object Detection

IV. CONCLUSION

Artificial Intelligence and Machine Learning are among the technologies growing at a rapid pace. These technologies are playing a pivotal role in the development of the IT sector. In this paper, we have tried to visit the existing real-time object detection methods to find out the best suitable method for the navigation of the blind.

Based on these findings, we propose to develop a real time navigation device for the visually impaired using machine learning and object detection models which can detect the obstacles accurately in front of the user and give accurate audio as well as haptic feedback to the user. The device will be built keeping in mind the comfort and portability for the user and will also be affordable to the common man while still being durable and reliable for them.

ACKNOWLEDGMENTS

The authors would like to thank the Department of Computer Science and Engineering at Jyothy Institute of Technology and Akansha Powertronix for providing us with unparalleled support without which we could not have completed this journal.

REFERENCES

[1] Wu B, Iandola F, Jin PH, Keutzer K. Squeezedet: Unified, small, low power fully convolutional neural networks for real-time object detection for autonomous driving. InProceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops 2017 (pp. 129-137).W.-K. Chen, Linear Networks and Systems. Belmont, CA, USA: Wadsworth, 1993, pp. 123–135.

[2] Wang RJ, Li X, Ling CX. Pelee: A real-time object detection system on mobile devices. arXiv preprint arXiv:1804.06882. 2018 Apr 18.

[3] Abdulhussain SH, Mahmmod BM, Saripan MI, Al-Haddad SA, Baker T, Flayyih WN, Jassim WA. A fast feature extraction algorithm for image and video processing. In2019 international joint conference on neural networks (IJCNN) 2019 Jul 14 (pp. 1-8). IEEE.

[4] Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: Unified, real-time object detection. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 779-788).

[5] Hussin R, Juhari MR, Kang NW, Ismail RC, Kamarudin A. Digital image processing techniques for object detection from complex background image. Procedia Engineering. 2012 Jan 1;41:340-4.

[6] Howard AG, Zhu M, Chen B, Kalenichenko D, Wang W, Weyand T, Andreetto M, Adam H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861. 2017 Apr 17.

[7] Girshick R, Donahue J, Darrell T, Malik J. Region-based convolutional networks for accurate object detection and segmentation. IEEE transactions on pattern analysis and machine intelligence. 2015 May 25;38(1):142-58.

[8] Mohan AS, Resmi R. Video image processing for moving object detection and segmentation using background subtraction. In2014 First International Conference on Computational Systems and Communications (ICCSC) 2014 Dec 17 (pp. 288-292). IEEE.