EAI Endorsed Transactions
on AI and Robotics
Review Article **EAI.EU**

The Object Detection, Perspective and Obstacles In Robotic: A Review

Ge Xu[1], Amir Sohail Khan[2], Ata Jahangir Moshayedi[2,*], Xiaohong Zhang[2], Yang Shuxin[2]

[1]School of architecture and design, Jiangxi University of Science and Technology, No 86, Hongqi Ave, Ganzhou, Jiangxi,341000, China (972570068@qq.com).
[2]School of Information Engineering Jiangxi University Of Science And Technology, Ganzhou, China (mrsohail21@gmail.com, ajm@jxust.edu.cn, xiaohongzh@jxust.edu.cn, yangshuxin@jxust.edu.cn).

Abstract

A few years back, when the image processing hardware and software were created, it was limited, and most of the time, object detection would fail., But as with time, the advancement in technology has significantly changed the scenario. A lot of researchers worked in this field to carry out a solution through which they can detect objects in any field, especially in the robotic domain [1]. In today's world, with so much research in the field of deep learning, it is very easy to identify and detect any object using computer vision. This paper focuses on the various deep learning technologies and algorithms through which object detection can be done. A new and advanced deep learning technology known as salient object detection has been discussed. Also, the 3D object detection and the end-to-end approach for object detection are discussed. The existing methods of deep learning through which object detection can be done. The applications in which object detection is being used and the importance of object detection. It also reports; what the predecessors have done, what problems have been solved by them, how they solved these problems, the characteristics of the predecessors' methods and their future work.

Keywords: Deep Learning, Object Detection, Image Processing, Salient Object Detection (SOD), 3D object detection.

Received on 19 September 2022, accepted on 20 September 2022, published on 18 October 2022

DOI: 10.4108/airo.v1i1.2709

*Corresponding author. Email: ajm@jxust.edu.cn

1. Introduction

Object detection is the most important feature of computer vision and has attracted a lot more attention in recent years, especially because of its more attraction towards video analysis and a better understanding of images [2]. Object detection performs many computer vision tasks [3], such as image segmentation, image captioning, object tracking and many more. Object detection has various applications and has recently been used in different areas like weld failure detection or fault detection in pipelines and diagnosis [4], Corrosion Detection in pipelines [5] and license plate detection [6], defect detection of pipelines through robots [7] etc. Deep learning (DL) technologies and methods are mostly used [8], especially in computer vision. CNN (Convolutional Neural Network) is one of the most powerful deep learning technologies that can feature learning and transfer learning. In recent years, CNN has received a lot of interest in the computer vision community, making it an important aspect of success for object detection [9]. Imaging technology has recently advanced and improved greatly. Using image processing and machine learning for object detection or fault detection, an application was recently developed known as HPD, which helps to detect the fault of an object [10]. Some other specific object detection applications include face detection and recognition [11], text detection [12], pose detection [13], people counting [14], and wind tracking algorithms [15]. This paper is organized into 5 sections. Each section's contribution and detail are given below in the following way; Section I describes the different existing methods and algorithms that are used for any kind of object detection. The Challenges that occur during object detection are described in Section II. Section III explains the work of the different authors for object detection. The results and discussions, along with the table which compares the work of different authors for object detection, are explained in Section IV. While Section V provide the conclusion of the paper and future work that must be done.

Section I: Existing Methods and Algorithms used for Object Detection

Convolutional Neural Networks (CNNs) are the most popular algorithms used for object detection. These CNNs include R-CNN (Region-based Convolution Neural Network) [16], Faster R-CNN [17], Mask R-CNN [18], Mobile Net [19], SqueezeDet [20] and You Only Look Once (YOLO) [21]. Among the CNNs algorithms, YOLO is an advanced algorithm for object detection, and in some cases, different versions of YOLO algorithms are being used. Different versions of YOLO algorithms used in object detection can be YOLOv2 [22], YOLOv3 [23], YOLOv4 [24] and YOLOR [25]. Every algorithm used for object detection can detect an object in real-time and accurately.

Section II: CHALLENGES IN OBJECT DETECTION

Even with so much advanced technology, object detection may face some challenges in recognizing an object. These challenges are [26];

a. Change in angle: This is one of the biggest problems of object detection. Because if an object is viewed from one angle, that same object might look different from another. And sometimes, this gives inaccurate results. Hence this is one of the goals of object detectors, to recognize the object from a different angle. This problem mainly occurs in vehicle license plate detection [27].

b. Change in shape: Another difficulty in object detection is a change in shape or pose. For example, if the detector is trained only for a person standing, it may not be able to detect someone lying down or running.

c. Objects obscured: At times, objects can be obscured by other objects, which makes it difficult for the object detector to recognize or detect the object easily and sometimes gives wrong information.

d. Low Light: Lighting can make a big difference in detecting objects. The same object will look different at night or in bad lighting conditions.

e. Cluttered Background: Sometimes, there are many objects in an image, making it confusing for the object detector to recognize the object.

f. Different Varieties: Most of the time, the same object can vary in size and shape. Thus, it is up to computer vision to recognize them, no matter what changes are made in that object.

g. Change in Speed: This change in speed is one of the most challenging tasks. In today's world, the object detection algorithm should not only recognize an object accurately but also; it must be quick in the prediction to identify the objects that are in motion.

Section III: The basic point of a review

Extending FCN (Fully Convolutional Network) to 3D, author **Bo Li 2017**, designed a 3D full convolution network, which used the target as a 3D box in a point cloud to detect and locate; also it can be applied to 3D vehicle detection of the autonomous driving system. This paper presents the first 3D FCN (fully convolutional network) framework for end-to-end 3D object detection. This method can be extended to other target detection tasks on point clouds captured from motion by Kinect, stereo or monocular structures. Experiments on the Kitti dataset show that the performance of this method is significantly improved by 20% compared with the previous point cloud-based detection method [28]. The method of 3D object detection and pose estimation from a single image was presented by **Arsalan Mousavian et al. 2017**. This method uses the depth convolution neural network to get the 3D object attributes and combines the 2D object bounding box to get a complete 3D bounding box. Its characteristic is better than L2 consumption, the variance of 3D object dimension is relatively small, and it can predict many object types. Future work will study the advantages of using a single depth channel using stereo computing to enhance RGB image input and effectively use time information to explore 3D box estimation in video and predict the future object position and speed [29]. A depth complete convolution network model for salient object detection, suitable for other pixel vision tasks, was proposed by **Pingping Zhang et al. 2017**. The author introduces a new r-dropout (Regularized-dropout) after a specific convolution layer to promote probability training and reasoning to construct an uncertain set of internal feature units and proposes an effective hybrid up-sampling method to reduce the artefacts of the deconvolution operator in the decoder network, and explicitly force the network to learn the exact boundary of the significance detection. This method is also applicable in other depth convolution networks. This study aims to learn the uncertain convolutional features (UCF) and produce more accurate significance prediction. The feature of this method is that it can contain more uncertainty when facing the need for more accurate object boundary reasoning [30]. **Xinyi Zhou et al. 2017** summarized the data sets and neural networks related to deep learning for target detection. Common data sets include ImageNet, Pascal VOC (Visual Object Classes) and coco. Common neural networks include R-CNN, SPP

Net (Spatial Pyramid Pooling Network), and Fast R-CNN. Then the author processes the image data of a football match on the new VOC data set by a faster RCNN neural network. Through the recognition effect of the experiment, readers can realize the importance of deep learning technology and data [31]. The end-to-end object detection model based on CNN to predict the location of an Unmanned Aerial Vehicle (UAV) in the video frame was proposed by **Cemal Aker et al. 2017**. The trained network can distinguish and detect UAVs from birds. And the trained network has good generalization, high accuracy and recall value. Future work can further consider the time domain to improve performance [32]. The possibility of automatic analysis of camera trap images for expensive camera trap image analysis was presented by **Stefan Schneider et al. 2018**. The author trained two deep learning object detection classifiers, the first was faster R-CNN, and the second was YOLO V2.0 and compared their ability to recognize, quantify and locate animal species. In conclusion, the average accuracy of faster r-CNN is about 16% higher than that of Yolo v2.0. The future work is to make the data analysis of camera trap image automatic based on this data set and help humans understand the earth's ecosystem's population dynamics [33]. A new in-depth learning method known as (AfferceNet), which can detect multiple objects and their afferces from RGB images simultaneously, was proposed by **Thanh-Toan Do et al. 2018.** The advantage of this method was that its end-to-end architecture is 150 milliseconds per image. It shows that this method is suitable for real-time robot applications. The authors discussed three main components; the first was a sequence of decentralized layers, the second was a robust resizing strategy, and the third was a new loss function. These three play a key role in solving the problem of multiple providers in the detection task and improving the detection accuracy [34]. In another new research**, Yin Zhou et al. 2018** proposed a new end-to-end trainable deep architecture based on point cloud 3D detection known as voxel net. The feature of this method is to eliminate the need for artificial feature engineering for 3D point clouds. It combines feature extraction and bounding box prediction into a single, end-to-end trainable depth network, which can directly capture 3D shapes according to sparse 3D points and connect to RPN (Risk Priority Number) to generate inspection. This method is superior to the most advanced 3D detection method based on lidar. Future work will further extend the voxel net to combine lidar and image-based end-to-end 3D detection, improving detection and positioning accuracy [35]. A neural network known as the Attention Salinece Network (ASNet) was proposed by **Wenguang Wang et al. 2018**. The proposed network uses a fixed prediction to detect the protruding objects and takes the fixed graph as the derivation of the upper network layer. The protruding object detection is segmented by significance, and then under the guidance of the fixed graph, it is gradually optimized from top to bottom, showing a more refined object significance. The significance of this study is to provide an effective recursive mechanism for the sequence refinement of segmentation graphs to prove the importance of fixed graphs to readers, strengthen the relationship between significant target detection and fixed detection, and narrow the gap between them. The future work is to explore the basic principle of SOD (Salient Object Detection) from the perspective of fixed prediction and find a better loss function to improve the performance of the SOD model based on deep learning [36]. The research related to three-dimension authors **Charles R.Qi et al. 2018** explored three-dimensional target detection of RGB-d data in indoor and outdoor scenes. The author directly manipulates the original point cloud by pop-up RGB-d scanning and realizes the object location by combining the mature 2D object detector and 3D deep learning, thus obtaining a high recall rate. In addition, because of the direct learning in the original point cloud, this method can also output a very accurate 3D instance segmentation mask and 3D bounding box in the case of strong occlusion or for very sparse points. The feature of this method is that it is superior to the most advanced method and has real-time capability. Future work will improve the accuracy of attitude and size estimation in sparse point clouds and realize the recognition of multiple objects [37]. Detecting objects underwater through robots was proposed by **Fengqiang Xu et al. 2018.** The authors used novel methods known as Faster R-CNN and Kernalized Correlation Filter (KCF) algorithms. Through these methods, objects underwater can be detected. A total of 28,000 images were tested and trained in three categories. These categories were sea cucumber, sea urchin and scallop. The results obtained were quite good for the authors, achieving higher accuracy of 79.6% [38]. A novel filter framework which can solve any path of problems related to tracking was proposed by **Lei Sun et al. 2018.** The authors, after conducting experiments, stated that novel filter frameworks could realize and detect objects for control robot systems which are specially designed for motion parameter extraction. The experimental results showed very good and accurate results. But the camera used in their study is not calibrated, which is the authors' future work [39]. To deal with different tasks and challenges in object detection, authors **Guoling et al. 2018**, proposed a popular deep learning method known as CNN (Convolutional Neural Network), which makes the robot's services easier and easier to detect and identify any object. The authors also discussed the grabbing of objects and used the depth-first algorithm method. Experiments were performed to verify their proposed method, and the results were quite successful and accurate. For the object-grabbing tasks, 3D coordinates were used. The authors' future work is improving the structure of CNN for higher accuracy and methods for calculating the three-dimensional coordinate of any object to avoid problems and errors in grabbing objects [40]. Based on human bionics and Goodman model equations, author **Zhipeng Kuang et al. 2018**, designed a robot for daily use that can detect and track objects in motion or stability. The author conducted experiments based on stereo vision, through which the robot could track the object in a three-dimensional space using vision. From the experimental results, it can be concluded that the robot can detect objects both in motion and static and can have a high accuracy of detection to grab an object under the control of the bionic coordinated motion model algorithm [41]. For real-time object detection through robotic vision, **Jiazhen Guo et al. 2021**, proposed deep learning techniques and algorithms through which a robot can detect and identify an object. You Only Look Once (YOLO) model was used for quick and faster detection of objects. The main idea of this model was to create a real-time transmission structure between HoloLens and Ubuntu systems using the proposed YOLO model for faster object detection. HoloLens and Ubuntu rosbridge Application Programming Interface (API) were used for data transmission. For the YOLO model, two versions of YOLO were used. One was YOLOv3, and the other was YOLOv4. After conducting experiments, it was concluded that the YOLOv4 had the best speed and significantly higher object detection accuracy than the YOLOv3. In future work, these algorithms can be applied to the robot's arm to hold the objects [42] automatically.

Section IV: Application of image detection in robotics

The survey shows that the object detection task can apply in various missions. These object detectors can be used in embedded systems [43] for various tasks. These tasks can handle virtual power plants' security and transaction mechanisms [44]. Another application where image detection can be helpful is the school, college or university attendance system [45]. This system works on facial recognition [46]. Many Medical equipments, like heart rate monitoring, where the heart rate is monitored, is also an application of object detection and recognition in computer vision tasks. Games known as exergames for activity incrementation have used advanced machine learning technology combined with robotics and embedded systems [47]. The automatic detection of vehicles through their license plates is also an application of image detection. This method uses the modern and advanced deep learning process and algorithms [48], where license plates can be detected and recognized easily.

After so much development, robotic technology has finally made achievements in many fields. Robots can see and make decisions more like humans because of technological advancements like image processing [49], image detection, machine learning, AI and CV. Small robots are made for pipe inspection as well. These small robots go inside the pipelines and capture images of all the internal areas to find if there is any defect inside a pipeline. If any defect is detected, the robot will automatically signal the user to overcome that defect.

RESULTS AND DISCUSSION

Object detection is a vital task but also very challenging as well. It is challenging due to applications such as searching for images, auto-annotation, understanding the scene, and tracking down objects, especially objects in motion. Sometimes these moving objects can be very fast, making them difficult to detect. But this task is no longer difficult, as the advancements in deep learning make it easier for computer vision to detect objects even if they are in motion. Object detection has been applied in many fields, such as smart video surveillance, artificial intelligence (AI), military guidance, safety detection and robot navigation, and many medical and biological applications [50]. This review paper also focuses on 3D object detection. Depth convolution neural network and FCN (Fully Convolutional Network) are the methods used for 3D object detection, which are new and advanced methods in deep learning. CNN, Faster R-CNN, and YOLOv2 deep learning algorithms used for object detection have good accuracy. Faster R-CNN has high accuracy and faster recognition than other deep learning technologies. Object detection through robotics which is an advanced technology was also discussed in this paper. For better and faster detection through robots, the YOLOv4 algorithm can be used, which has a very faster speed of detection and relatively higher recognition accuracy than the rest of the deep learning algorithms. The paper also presented some new deep-learning technologies that can help object detection. Attention Salience Network **(**RSNet**)**and Regularized- dropout(R-dropout) are the new modern technologies that help detect the object. Three tables are made to discuss and compare the work of the authors.

Table 1 describes the SOD (Salient Object Detection) work of the authors. Table 2 describes the 3D object detection, while the end-to-end approach to object detection is discussed in Table 3.

Table 1. The object detection salient comparison

Reference Paper	Year	Result	Methods	Characteristic	Future work
[25]	2017	A new deep complete convolution network model	A new r-dropout (Regularized-dropout) is introduced to promote probability training and reasoning; the hybrid up-sampling method detects the precise boundary	It can contain more uncertainty and face the need for more accurate object boundary reasoning	-
[31]	2018	A new neural network: attention salience network (ASNET)	Using fixed prediction, the fixed graph is used as the derivation of the upper network layer, and the outstanding target detection is segmented by significance, and then it is gradually optimized from top to bottom under the guidance of the fixed graph	Using the importance of a fixed graph to strengthen the relationship between significant target detection and fixed detection	This paper discusses the basic principle of SOD from the perspective of fixed prediction and looks for a better loss function to improve the performance of the SOD model based on deep learning

As Table 1 shows the Salient Object Detection (SOD). ASNet and R-dropout are the methods that can be used for the SOD. The table shows that both ASNet and R-dropout are successful methods used for object detection.

Table 2. The 3D object detection comparison

Reference Paper	Year	Result	Methods	Characteristic	Future work
[23]	2017	A 3D full convolution network based on FCN can be applied to 3D vehicle detection of the autonomous driving system	Detect and locate the target as a 3D box in the point cloud	Compared with the previous point cloud-based detection method, the performance of this method is improved by 20%	-
[24]	2017	3D object detection and pose estimation from a single image	Using the depth convolution neural network to get the 3D object attributes, combined with a 2D object bounding box to get a new 3D bounding box	It is better than L2 consumption, and the variance of 3D object dimension is relatively small	Effective use of time information to explore 3D box estimation in video and predict object position and speed
[32]	2018	Achieve a high recall rate of RGB-d data 3D target detection in indoor and outdoor scenes	The original point cloud is operated by RGB-d scanning, and the object location is realized by combining the mature 2D object detector and 3D depth learning	Can adapt to strong occlusion or sparse point cloud	Improve recognition accuracy and realize the recognition of multiple objects

Table 2 shows 3D object detection, a modern-day technology. RGB-D, Depth Convolution Network, and FCN are used for 3D object detection. FCN has good detection accuracy and faster recognition than other neural networks.

Table 3. Shows End-to-End Approach comparison

Reference Paper	Year	Result	Methods	Characteristic	Future work
[27]	2017	The trained network can distinguish and detect UAVs from birds	The end-to-end object detection model based on CNN	the trained network has good generalization , high accuracy and recall value	Future work can take the time domain into account to further improve performance
[28]	2018	the average accuracy of faster r-CNN is about 16% higher than that of Yolo v2.0	Faster R-CNN and YOLOv2	--	to make the data analysis of camera trap image automatic based on this data set and help humans understand the population dynamics of the earth's ecosystem
[29]	2018	A new in-depth learning method (afforcenet) is proposed to detect multiple objects and their afforcements from RGB images simultaneously. Very suitable for real-time robot applications	Three components: a sequence of deconvolutional layers, a robust resizing strategy, and a new loss function	Its end-to-end architecture is 150 milliseconds per image	
[30]	2018	A new end-to-end trainable depth architecture (voxelnet) based on point cloud 3D detection is superior to the most advanced lidar 3D detection method	Capture 3D shape directly from sparse 3D points and connect to RPN to generate detection	Eliminate the need for a 3D point cloud for artificial feature engineering. It combines feature extraction and bounding box prediction into a single, end-to-end trainable depth network	Expand voxelnet to combine lidar and image-based end-to-end 3D detection with improving detection and positioning accuracy further

Table 3 discuss the end-to-end approach to object detection. CNN, Fast R-CNN and YOLOv2 deep learning methods are used for this approach. Among them, Fast R-CNN achieved higher accuracy and faster recognition results compared to CNN and YOLOv2. Figure 1 shows the deep learning working domain in various subjects.

The figure shows the object detection applications and methods used in those applications. Some object detection applications are; Robotics, Image detection and training, 3d object detection, Salient object detection and end-to-end approach. The object detection task almost uses CNN in most applications based on accuracy and has fast recognition speed in Deep Learning (DL) techniques and other neural networks.

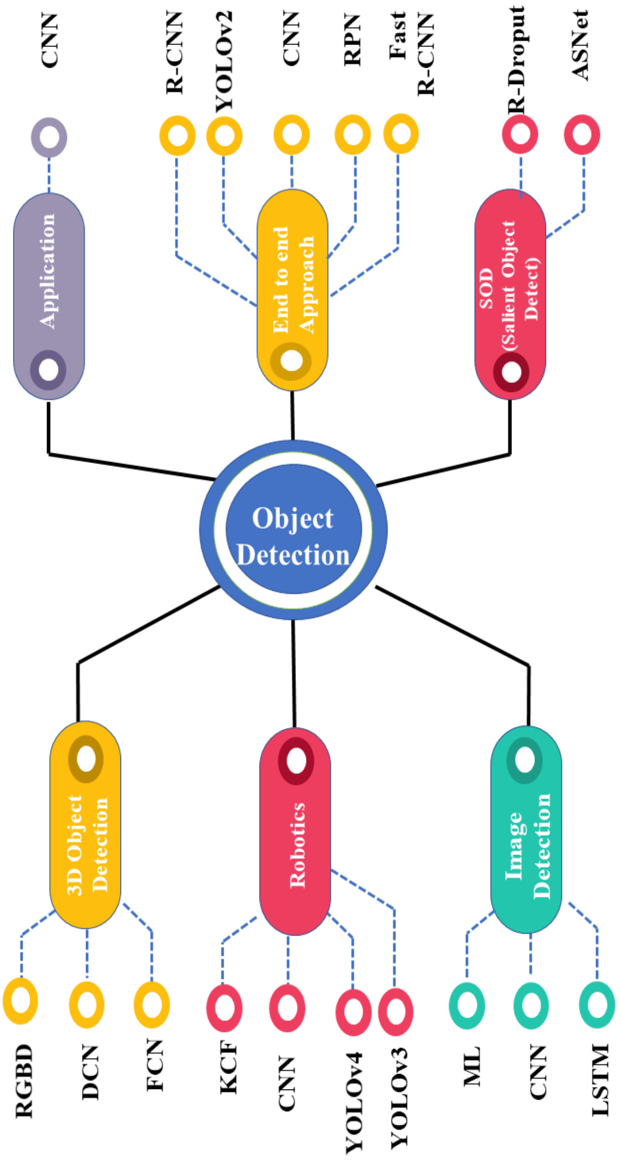

Figure 1. Object detection application Vs method

Section V: CONCLUSION

Object detection can help many security agencies, especially law enforcement security agencies, to avoid violations and problems. The main aim of this review was to find which deep learning method can be used for object detection and how it can be helpful to the human world. Some years ago, finding and classifying objects in a picture was hard and almost impossible. But today, with the help of computer vision and deep learning technology, it is quite easy to do these tasks and recognize any image. For object detection, many computer vision-based techniques and algorithms are used. CNN, R-CNN, Faster R-CNN, Mask R-CNN, YOLO (with different versions; V2, V3, V4 and R), MobileNet and SqueezeDet. These algorithms have a very high accuracy of recognition and detection, Especially the CNN, Faster R-CNN and YOLOv2. These Deep Learning methods and algorithms [51] make it easier for computer vision applications to recognize and detect any object, especially for detecting vehicle license plates.

Computer Vision plays a very key role and thus helps a lot of big industries these days. With the help of identifying objects in images or videos, they can improve many problems and difficulties they face. Smart systems equipped with computer vision face several difficulties, paying attention to which it only makes the system more accurate.

References

[1] A. J. Moshayedi and D. C. Gharpure, “Priority-based algorithm for plume tracking robot,” Proc. - ISPTS-1, 1st Int. Symp. Phys. Technol. Sensors, pp. 51–54, 2012, DOI: 10.1109/ISPTS.2012.6260876.

[2] K. Li and L. Cao, "A Review of Object Detection Techniques," 2020 5th International Conference on Electromechanical Control Technology and Transportation (ICECTT), 2020, pp. 385-390, DOI: 10.1109/ICECTT50890.2020.00091.

[3] A. J. Moshayedi, Z. Chen, L. Liao and S. Li, "Sunfa Ata Zuyan machine learning models for moon phase detection: algorithm prototype and performance comparison", TELKOMNIKA Telecomm. Compute. Electron. Control, vol. 20, no. 1, pp. 129-140, 2022. DOI: 10.12928/telkomnika.v20i1.22338

[4] J. Wang, S. Gao, B. Zhang and S. Lv, "A Survey of Fault Diagnosis Methods for White body Welding Production Line," 2019 CAA Symposium on Fault Detection, Supervision and Safety for Technical Processes (SAFEPROCESS), 2019, pp. 304-308, DOI: 10.1109/SAFEPROCESS45799.2019.9213399.

[5] M. S. A. Rahman and H. Hasbullah, "Early detection method of corrosion on buried steel gas pipeline using wireless sensor network," 2010 The 2nd International Conference on Computer and Automation Engineering (ICCAE), 2010, pp. 553-556, DOI: 10.1109/ICCAE.2010.5451824.

[6] A. Firasanti, T. E. Ramadhani, M. A. Bakri and E. A. Zaki Hamidi, "License Plate Detection Using OCR Method with Raspberry Pi," 2021 15th International Conference on Telecommunication Systems, Services, and Applications (TSSA), 2021, pp. 1-5, DOI: 10.1109/TSSA52866.2021.9768252.

[7] A. J. Moshayedi, S. S. Fard, L. Liao and S. A. Eftekhari, "Design and Development of Pipe Inspection Robot Meant for Resizable PipeLines", Int. J. Robot. Control, vol. 2, no. 1, pp. 25, 2019.

[8] W. Zhiqiang and L. Jun, "A review of object detection based on convolutional neural network," 2017 36th Chinese Control Conference (CCC), 2017, pp. 11104-11109, DOI: 10.23919/ChiCC.2017.8029130.

[9] A. J. . Moshayedi A. S. . Roy, A. Kolahdooz, and Y. . Shuxin, “Deep Learning Application Pros And Cons Over Algorithm”, EAI Endorsed Trans AI Robotics, vol. 1, no. 1, p. e7, Feb. 2022. DOI: 10.4108/airo.v1i.19

[10] A. J. Moshayedi, A. S. Khan, S. Yang and S. M. Zanjani, "Personal Image Classifier Based Handy Pipe Defect Recognizer (HPD): Design and Test," 2022 7th International Conference on Intelligent Computing and Signal Processing (ICSP), 2022, pp. 1721-1728, DOI: 10.1109/ICSP54964.2022.9778676.

[11] R. Xie, Q. Zhang, E. Yang and Q. Zhu, "A Method of Small Face Detection Based on CNN," 2019 4th International Conference on Computational Intelligence and Applications (ICCIA), 2019, pp. 78-82, DOI:10.1109/ICCIA.2019.00022.

[12] H. Xin, C. Ma and D. Li, "Comic Text Detection and Recognition Based on Deep Learning," 2021 3rd International Conference on Applied Machine Learning (ICAML), 2021, pp. 20-23, DOI: 10.1109/ICAML54311.2021.00012.

[13] T. Yasunaga, T. Oda, N. Saito, A. Hirata, K. Toyoshima and K. Katayama, "Object Detection and Pose Estimation Approaches for Soldering Danger Detection," 2021 IEEE 10th Global Conference on Consumer Electronics (GCCE), 2021, pp. 697-698, DOI: 10.1109/GCCE53005.2021.9621849.

[14] E. P. Myint and M. M. Sein, "People Detecting and Counting System," 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), 2021, pp. 289-290, DOI: 10.1109/LifeTech52111.2021.9391951.

[15] A. J. Moshayedi and D. C. Gharpure, “Development of position monitoring system for studying performance of wind tracking algorithms,” 7th Ger. Conf. Robot. Robot. 2012, vol. 32, pp. 161–164, 2012.

[16] W. Zhang, S. Wang, S. Thachan, J. Chen and Y. Qian, "Deconv R-CNN for Small Object Detection on Remote Sensing Images," IGARSS 2018 - 2018 IEEE International Geoscience and Remote Sensing Symposium, 2018, pp. 2483-2486, DOI: 10.1109/IGARSS.2018.8517436.

[17] C. Lee, H. J. Kim and K. W. Oh, "Comparison of faster R-CNN models for object detection," 2016 16th International Conference on Control, Automation and Systems (ICCAS), 2016, pp. 107-110, DOI: 10.1109/ICCAS.2016.7832305.

[18] D. Kumar and X. Zhang, "Improving More Instance Segmentation and Better Object Detection in Remote Sensing Imagery Based on Cascade Mask R-CNN," 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, 2021, pp. 4672-4675, DOI: 10.1109/IGARSS47720.2021.9554512.

[19] C. Gao, Y. Zhai and X. Guo, "Visual Object Detection and Tracking System Design based on MobileNet-SSD," 2021 7th International Conference on Computer and Communications (ICCC), 2021, pp. 589-593, DOI: 10.1109/ICCC54389.2021.9674450.

[20] B. Wu, A. Wan, F. Iandola, P. H. Jin and K. Keutzer, "SqueezeDet: Unified, Small, Low Power Fully Convolutional Neural Networks for Real-Time Object Detection for Autonomous Driving," 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 446-454, DOI: 10.1109/CVPRW.2017.60.

[21] A. Sarda, S. Dixit and A. Bhan, "Object Detection for Autonomous Driving using YOLO [You Only Look Once] algorithm," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 1370-1374, DOI:10.1109/ICICV50876.2021.9388577.

[22] J. Fan, J. Lee, I. Jung and Y. Lee, "Improvement of Object Detection Based on Faster R-CNN and YOLO," 2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC), 2021, pp. 1-4, DOI: 10.1109/ITC-CSCC52171.2021.9501480.

[23] S. T. Blue and M. Brindha, "Edge detection based boundary box construction algorithm for improving the precision of object detection in YOLOv3," 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2019, pp. 1-5, DOI: 10.1109/ICCCNT45670.2019.8944852.

[24] S. Ali, A. Siddique, H. F. Ateş and B. K. Güntürk, "Improved YOLOv4 for Aerial Object Detection," 2021 29th Signal Processing and Communications Applications Conference (SIU), 2021, pp. 1-4, DOI: 10.1109/SIU53274.2021.9478027.

[25] M. Sharma et al., "YOLOrs: Object Detection in Multimodal Remote Sensing Imagery," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 1497-1508, 2021, DOI: 10.1109/JSTARS.2020.3041316.

[26] S. Mane and S. Mangale, "Moving Object Detection and Tracking Using Convolutional Neural Networks," 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018, pp. 1809-1813, DOI: 10.1109/ICCONS.2018.8662921.

[27] B. Li, "3D fully convolutional network for vehicle detection in point cloud," 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2017, pp. 1513-1518, DOI: 10.1109/IROS.2017.8205955.

[28] A. J. Moshayedi, J. Li and L. Liao, "Simulation study and PID Tune of Automated Guided Vehicles (AGV)," 2021 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), 2021, pp. 1-7, DOI: 10.1109/CIVEMSA52099.2021.9493679.

[29] A. Mousavian, D. Anguelov, J. Flynn and J. Košecká, "3D Bounding Box Estimation Using Deep Learning and Geometry," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 5632-5640, DOI: 10.1109/CVPR.2017.597.

[30] P. Zhang, D. Wang, H. Lu, H. Wang and B. Yin, "Learning Uncertain Convolutional Features for Accurate Saliency Detection," 2017 IEEE International Conference on Computer Vision (ICCV), 2017, pp. 212-221, DOI: 10.1109/ICCV.2017.32.

[31] X. Zhou, W. Gong, W. Fu and F. Du, "Application of deep learning in object detection," 2017 IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS), 2017, pp. 631-634, DOI: 10.1109/ICIS.2017.7960069.

[32] C. Aker and S. Kalkan, "Using deep networks for drone detection," 2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2017, pp. 1-6, DOI: 10.1109/AVSS.2017.8078539.

[33] S. Schneider, G. W. Taylor and S. Kremer, "Deep Learning Object Detection Methods for Ecological Camera Trap Data," 2018 15th Conference on Computer and Robot Vision (CRV), 2018, pp. 321-328, DOI: 10.1109/CRV.2018.00052.

[34] T. -T. Do, A. Nguyen and I. Reid, "AffordanceNet: An End-to-End Deep Learning Approach for Object Affordance Detection," 2018 IEEE International Conference on Robotics and Automation (ICRA), 2018, pp. 5882-5889, DOI: 10.1109/ICRA.2018.8460902.

[35] Y. Zhou and O. Tuzel, "VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 4490-4499, DOI: 10.1109/CVPR.2018.00472.

[36] W. Wang, J. Shen, X. Dong and A. Borji, "Salient Object Detection Driven by Fixation Prediction," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 1711-1720, DOI: 10.1109/CVPR.2018.00184.

[37] C. R. Qi, W. Liu, C. Wu, H. Su and L. J. Guibas, "Frustum PointNets for 3D Object Detection from RGB-D Data," 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, pp. 918-927, DOI: 10.1109/CVPR.2018.00102.

[38] F. Xu et al., "Real-Time Detecting Method of Marine Small Object with Underwater Robot Vision," 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO), 2018, pp. 1-4, DOI: 10.1109/OCEANSKOBE.2018.8558804.

[39] L. Sun, S. -a. Wang, H. Chen and Y. Chen, "A novel object detection before tracking filter framework for assistive robot under global vision," 2018 IEEE 4th Information Technology and Mechatronics Engineering Conference (ITOEC), 2018, pp. 1011-1014, DOI: 10.1109/ITOEC.2018.8740736.

[40] G. Zhang, S. Jia, D. Zeng and Z. Zheng, "Object Detection and Grabbing Based on Machine Vision for Service Robot," 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON), 2018, pp. 89-94, DOI: 10.1109/IEMCON.2018.8615062.

[41] Z. Kuang, Y. Yang, T. Hau and Y. He, "Research on Biomimetic Coordination Action of Service Robot Based on Stereo Vision," 2018 2nd IEEE Advanced Information Management,Communicates,Electronic and Automation Control Conference (IMCEC), 2018, pp. 769-773, DOI: 10.1109/IMCEC.2018.8469748.

[42] J. Guo, P. Chen, Y. Jiang, H. Yokoi and S. Togo, "Real-time Object Detection with Deep Learning for Robot Vision on Mixed Reality Device," 2021 IEEE 3rd Global Conference on Life Sciences and Technologies (LifeTech), 2021, pp. 82-83, DOI: 10.1109/LifeTech52111.2021.9391811.

[43] A. J. Moshayedi, A. Kolahdooz, and L. Liefa, Unity in Embedded System Design and Robotics: A Step-by-step Guide, CRC Press, Routledge, 2022

[44] Zhang, X., Song, Z., Moshayedi, A.J. et al. Security scheduling and transaction mechanism of virtual power plants based on dual blockchains. J Cloud Comp 11, 4 (2022). DOI: 10.1186/s13677-021-00273-3

[45] Ata Jahangir Moshayedi, Atanu Shuvam Roy, Liefa liao, Hong Lan, Mehdi Gheisari,Aaqif Afzaal Abbasi and Seyed Mojtaba Hosseini Bamakan, "Automation Attendance Systems Approaches: A Practical Review BOHR

International Journal of Internet of Things Research2022, Vol. 1, No. 1, pp. 7–15, DOI: 10.54646/bijiotr.003

[46] Gheisari, M., Esnaashari, M. (2017). A survey to face recognition algorithms: advantageous and disadvantageous. Journal Modern Technology & Engineering, V. 2(1), pp. 57-65.

[47] Ata Jahangir Moshayedi, Amin Kolahdooz, Atanu Shuvam Roy, Seyyed Ali Latifi Rostami, and Xiaoyun Xie "Design and promotion of cost-effective IOT-based heart rate monitoring", Proc. SPIE 12303, International Conference on Cloud Computing, Internet of Things, and Computer Applications (CICA 2022), 123031N (28 July 2022); DOI: 10.1117/12.2642725.

[48] Alzubi J.A., Yaghoubi A., Gheisari M., Qin Y. (2018) Improve Heteroscedastic Discriminant Analysis by Using CBP Algorithm. In: Vaidya J., Li J. (eds) Algorithms and Architectures for Parallel Processing. ICA3PP 2018. Lecture Notes in Computer Science, vol 11335. Springer, Cham.

[49] A. J. Moshayedi, S. K. Sambo and A. Kolahdooz, "Design And Development of Cost-Effective Exergames For Activity Incrementation," 2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE), 2022, pp. 133-137, DOI: 10.1109/ICCECE54139.2022.9712844.

[50] Mukkamala Rohith Sri Sai, Sindhusha Rella, Sainagesh Veeravalli, “Object Detection and Identification” B.Tech Thesis, November 2019 [Online] Available at: https://www.researchgate.net/publication/337464355

[51] A. J. Moshayedi, A. Shuvam Roy, S. K. Sambo, Y. . Zhong, and L. Liao, “ Review On: The Service Robot Mathematical Model ”, EAI Endorsed Trans AI Robotics, vol. 1, no. 1, p. e8, Feb. 2022.DOI: 10.4108/airo.vli.20