Modeling Real Time Object Identification Performance Evaluation with YOLO

*Ashish Joshi[1], Dr. Ankur Dumka[2], Aditya Raturi[3]
Santosh Kumar[4] and Isha Pant[5]
[1,3,4]*Security Consultant, BOSCH, Bangalore*
[2]*Associate Professor, Women Institute of Technology, Uttarakhand*
[5]*Assistant Professor, THDC Institute of Hydropower Engineering and Technology, Uttarakhand*

[1]*a.joshicse1986@gmail.com,*[2]*ankurdumka2@gmail.com,*
[3]*addy3177@gmail.com,*[4]*santosh51623z@gmail.com,*
[5]*ishapant324@gmail.com,*

Abstract

Object identification in pictures, videos, and signal processing is not a big frontier and has been around for a few years. Though object recognition in static pictures has proven highly promising in terms of specific items such as facial recognition systems, illness diagnosis, and so on, it has been tricky when it comes to video processing and real-time image processing. Traditional object identification techniques are coupled with machine learning methodologies to improve algorithms' speed and accuracy. The paper focuses on real-time object recognition systems and how recent advances in the realm of object recognition and identification have been made while bearing in mind the real-time scenarios for recognizing varied objects.The paper discusses the current implementations illustrated by varied authors using far more widely used algorithms including YOLO for real-time visualizations.

Keywords: *Object recognition, YOLO, Machine learning, RCNN, faster RCNN.*

1. Introduction

Object detection and identification is a key field of computer vision that tries to categorise or recognise distinct things in images based on particular traits.With the help of machine learning/ deep CNN (convolution neural network) object detection and identification has improved to a large extent. Object detection (also known as object identification) is a crucial subdomain of Computer Vision because tasks like detection[28], recognition, & localization have pervasive implementation in the real scenarios. Object detection and identification comprises of vast set of predefined mathematical approaches used largely used in Medical sciences, remote sensing, agriculture, etc. these mathematical approaches are combined with deep learning/machine learning approaches to produce the result faster and accurate.

Machine Learning algorithms have enormously transformed many of approaches used in image processing, video processing, signal processing, as well as industrial and corporate processes[21]. Machine Learning proposes several techniques for extracting knowledge from data and turning it into goals that provide meaningful information. Machine Learning has created new possibilities and have drastically enhanced certain capabilities in a variety of areas, including health, transport, education, market research, statistical analysis, space exploration, and many more. Learning approaches together combined with computer vision techniques have proven to be very promising in field of self-driven cars for example towards image acquisition, real time analysis and interpretation for checking different procedures related to diagnosis, therapy etc. not only medical sciences active learning or deep learning approaches are widely used for facial recognition systems[23], but motion also capturing while analysing for live feed data, object classification segmentation detection etc.

2. Related Work

Object identification is a computer vision technology that allows you to recognise items may be person place or thing in photos or movies[22]. Object recognition is indeed a leading result of deep learning and machine learning algorithms.We can easily spot individuals, objects, settings, and visual features what people do naturally: obtain a level of comprehension of the world around them. The current synopsis tries to incorporate strategies such as to enhance the current recognition systems in terms of accuracy and preciseness. knowing that current approaches might incorporate semi structured data or unstructured the project works to overcome the challenges in mechanisms[15]. Object recognition is a fundamental feature of self-driving automobiles, allowing them to recognise a traffic light or distinguish between a pedestrian and a lamppost. It's also useful for disease detection in bio imaging, industrial inspection, and robotic vision, among other things. Active learning having combined with such approaches helps improve data efficiency for supervised learning[17]. Current strategies have already incorporated relevant technologies, their strategies will be part of implementation strategies in our study.

2.1. Methodology Adopted

Object detection classified into 2 forms: one-stage objectdetectors and two-stage objectdetectors. Deep learning-based object detectors, in particular, extract features from the input picture or video frame.

Two segments are solved by an object detector:
Task 1: Out of a given input image, find an arbitrary number of items (possibly even zero).
Task 2: Identify each object and determine its size using a bounding box.
We can divide those activities into two stages to make the process easier. Other techniques (single-stage detectors) integrate both jobs into one step to attain high performance in the presence of accuracy[16]. The two Stages Deep features are utilised in two-stage object detectors to suggest rough object areas before being used for object candidate classification and bounding box regression.

1. Proposal of an object region using traditional Computer Vision algorithms or deep networks, followed by.
2. Object categorization using extracted features

2.2. Algorithms in Existence

CNN: Using a CNN to classify the location of an object within distinctive regions of interest from an image would be a misleading way to addressing the problem of yield span variable due to the effects of an object in a specified image[14], but spatial locations within the picture and divergent aspect ratios are quite daunting in this case.

RCNN: To overcome the problem of picking a large number of regions, Ross Girshick et al. proposed a method called region suggestions, in which we use selective search to choose roughly 2000 areas from a picture. As a consequence, instead of attempting to identify a huge number of areas, you may now focus on just 2000. These 2000 region concepts were found using a selective search approach. However, there are several drawbacks[20]. To train the model, you had to categorise 2000 area suggestions for each image, which might take a long time. Because each test image takes around 47 seconds, it can't be done in real time.

Fast R-CNN: R-CNN algorithm is equivalent to this technique. We feed the CNN the input picture instead of the region recommendations to generate a feature map[18]. We take entire region of suggestions from the convolutional feature map, wrap them into squares, and then reassemble those into a set size using a RoI-pooling layer so they may be passed into linked layers[25].

Faster R-CNN: A convolutional neural network is fed an image, which produces a convolutional feature map[24]. Instead of using a selective search approach on the feature map to identify the area suggestions, a separate network is used to forecast them.

3. YOLO Framework

YOLO — You Only Look Once: Unlike previous object detection algorithms, which repurposed classifier to work more effectively, YOLO contends the use of an edge neural network that predicts boundary boxes & class probabilities in one go[26,27]. The most useful feature of using YOLO is its incredible speed – it can process 45 frames a second. YOLO is indeed aware of generalised image representation.YOLO achieves cutting-edge results by taking a diverse approach to perform detection, outperforming other context of real - world detection techniques by a wide edge[29]. YOLO developed a super robust and efficient feature selection method, which revolutionised object detection exploration in computer vision[19].YOLO has undergone a radical transformation ever since it was proposed in 2015, with over 5 variants and being quoted over 16 thousand times.YOLO does have a wide range of applications, with thousands of use instances, including automated driving, motion detection, & smart surveillance analytics.The image is splitted into SxS grids with m bounding boxes in every grid, which is how YOLO works.The network generates a class probability & bounding box offset values for every bounding box[30].To locate the item in the image, bounding boxes with a class probability greater than a threshold level are picked and utilized. YOLO is 45 frames per second faster orders of magnitude fast than conventional object detection algorithms. A comparative survey is illustrated as follows in Table 1.

Table 1. Comparative analysis of current reseach in selected domain

Title	Example	Published year	Journal Name
Oriented Bounding Boxes for Small and Freely Rotated Objects	The suggested method uses classification to learn all of the crucial data, with the added feature of enabling oriented bounding box recognition without any additional processing [2].	2022	IEEE
MSFYOLO: Feature fusion-based detection for small objects	This research indicates that combining global and local information can assist in detecting the expressiveness of small objects under various lighting conditions [1].	2022	IEEE
Object Detection Using Deep Learning Methods in Traffic Scenarios	The authors focus on changing traffic conditions [3], including weather difficulties for object detections, as well as applying deep learning algorithms and covering open research areas.	2021	ACM

Towards Open World Object Detection	Open set classification, Open word object detection, Known and unknown object [5] classifications are discussed by the authors.	2021	IEEE
End-to-End Object Detection with Fully Convolutional N/W.	3D max filtering and end to end detector mechanism used to achieve results faster [4].	2021	IEEE
Inception Recurrent Convolutional Neural Network for Object Recognition	A combined approach of inception based networks and recurrent networks in DCNN architecture. There is certain increase in recognition system as in previous cases.[6]	2021	Springer
Vehicle Detection and Speed Tracking Based on Object Detection	Uses YOLOv2, YOLOv3 for object detection in real time and for faster access, the main idea is to detect and classify the different cars.[7]	2019	Springer ICBMA 2019
Convolution NN to Perform Object Detection Identification in Visual Large-Scale Data	Proposes deep convolution neural network with combined approach of classification and regression to improve speed and accuracy.[10]	2021	Big Data
Joint weak saliency attention aware for person re-identification	Removes the various noises in the images background and various features are again remapped with persons image to re identify the persons [12]	2021	Elsevier
Object Detec. and Pose Estimation from RGB & Depth Data for Realtime, AdaptRobotic Grasping	Enables real time object detection mostly focused on robotic arm movement and gasping a specified object.[13]	2021	Springer
Analysis of Pruned Neural N/W for Underwater Object Detection	Authors have taken underwater object like specified species of fishes YOLOv2 algorithm has been used.[9]	2021	Springer
Hybrid Mamdani Fuzzy Rules C-NN for Analys. & Identification of Animal Image	Mamdani Fuzzy Rules for Edge Detection used to train CNN models in Moving images and 2 datasets are used more faster classification results are obtained.[8]	2021	MPDI
Novel automatic scorpion-detection & recognition sys. based on ML techniques	Florent features and object detection with features strategies are discussed with respect to scorpion detection[11]	2021	Machine Learning: Science and Tech.

Conclusion

Using deep learning-based object classification algorithm in challenging environments has resulted in noticeable improvements due to rapid computational breakthroughs. This article gives an overview of the relevant approaches that have aimed to determine objects under complex scenarios using deep neural networks, as well as an observational performance monitoring of presently used algorithms. In this proposed article, the study looks at real-time object recognition systems and how recent advances in the field of object recognition and identification have indeed been made while retaining real-time presets in mind. The intended study describes current real-time visualisation implementations as conveyed by numerous authors using markedly more broadly used technics such as YOLO.

References

[1] *Z. Song, Y. Zhang, Y. Liu, K. Yang and M. Sun, "MSFYOLO: Feature fusion-based detection for small objects," in IEEE Latin America Transactions, vol. 20, no. 5,* ***(2022),*** *pp. 823-830, doi: 10.1109/TLA.2022.9693567.*

[2] *M. Zand, A. Etemad and M. Greenspan, "Oriented Bounding Boxes for Small and Freely Rotated Objects," in IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1-15,* ***(2022),*** *Art no. 4701715, doi: 10.1109/TGRS.2021.3076050.*

[3] *Azzedine Boukerche and Zhijun Hou. 2021. Object Detection Using Deep Learning Methods in Traffic Scenarios. ACM Comput. Surv. 54, 2, Article 30* ***(2022)****, 35 pages. DOI:https://doi.org/10.1145/3434398*

[4] *J. Wang, L. Song, Z. Li, H. Sun, J. Sun and N. Zheng, "End-to-End Object Detection with Fully Convolutional Network," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),****(2022)****,pp. 15844-15853.*

[5] *K. J. Joseph, S. Khan, F. S. Khan and V. N. Balasubramanian, "Towards Open World Object Detection," 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR),****(2022)****, pp. 5826-5836, doi: 10.1109/CVPR46437.2021.00577.*

[6] *Alom, M.Z., Hasan, M., Yakopcic, C. et al. Inception recurrent convolutional neural network for object recognition. Machine Vision and Applications* ***32,*** *28* ***(2021)****.*

[7] *Song, H., Liang, H., Li, H. et al. Vision-based vehicle detection and counting system using deep learning in highway scenes. Eur. Transp. Res. Rev.* ***11,*** *51* ***(2019).***

[8] *Mohammed, H.R.; Hussain, Z.M. Hybrid Mamdani Fuzzy Rules and Convolutional Neural Networks for Analysis and Identification of Animal Images. Computation* ***(2021).***

[9] *Ayob, A.F., Khairuddin, K., Mustafah, Y.M., Salisa, A.R., Kadir, K.* ***(2021)****, "Analysis of Pruned Neural Networks (MobileNetV2-YOLO v2) for Underwater Object Detection." In: , et al. Proceedings of the 11th National Technical Seminar on Unmanned System Technology 2019.*

[10] *Ayachi R, Said Y, Atri M. A Convolutional Neural Network to Perform Object Detection and Identification in Visual Large-Scale Data. Big Data.* ***(2021).*** *Feb;9(1):41-52.: 32991200.*

[11] *Francisco L Giambelluca et al* ***(2021).*** *Mach. Learn.: Sci. Technol.* ***2*** *025018*

[12] *Xin Ning, Ke Gong, Weijun Li, Liping Zhang, "JWSAA: Joint weak saliency and attention aware for person re-identification, Neurocomputing",Volume 453, https://doi.org/10.1016/j.neucom.2020.05.106* ***(2021).***

[13] *Paul, S.K., Chowdhury, M.T., Nicolescu, M., Nicolescu, M., Feil-Seifer, D. "Object Detection and Pose Estimation from RGB and Depth Data for Real-Time, Adaptive Robotic Grasping." In: Arabnia, H.R., Deligiannidis, L., Shouno, H., Tinetti, F.G., Tran, QN. (eds) Advances in Computer Vision and Computational Biology. Transactions on Computational Science and Computational Intelligence. Springer, Cham, **(2021**). https://doi.org/10.1007/978-3-030-71051-4_10*

[14] *S. P. K. Reddy and G. Kandasamy, "Cusp Pixel Labelling Model for Objects Outline Using R-CNN," in IEEE Access, vol. 10, pp. 8883-8890, **(2021)**, doi: 10.1109/ACCESS.2021.3139896.*

[15] *S. Mane and S. Mangale, "Moving Object Detection and Tracking Using Convolutional Neural Networks," Second International Conference on Intelligent Computing and Control Systems (ICICCS), **(2018)**, pp. 1809-1813.*

[16] *Y Chen, X. Yang, B. Zhong, S. Pan, D. Chen and H. Zhang, "Cnn tracker: Online discriminative object tracking via deep convolutional neural network" in Applied Soft Computing, **(2016)**.*

[17] *R. Feris, A. Datta, S. Pankanti and M. T. Sun, "Boosting object detection performance in crowded surveillance videos", IEEE Workshop on Applications of Computer Vision, pp. 427-432, **(2013)**.*

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

[19] *Bilel Benjdira and Taha Khursheed, "Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv 3", International Conference on Unmanned Vehicle Systems-Oman (UVS), **(2019)**.*

[20] *S. B. Mane and S. Vhanale, "Real time obstacle detection for mobile robot navigation using stereo vision", 2016 International Conference on Computing Analytics and Security Trends (CAST), pp. 637-642,**(2016)**.*

[21] *Z. Ozkan, E. Bayhan, M. Namdar and A. Basgumus, "Object Detection and Recognition of Unmanned Aerial Vehicles Using Raspberry Pi Platform," 2021 5th International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), pp. 467-472, **(2021)**.*

[22] *A. Saini and M. Biswas, "Object Detection in Underwater Image by Detecting Edges using Adaptive Thresholding," 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 628-632, **(2019)**.*

[23] *M. Mahendru and S. K. Dubey, "Real Time Object Detection with Audio Feedback using Yolo vs. Yolo_v3," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 734-740, **(2021)**.*

[24] *Z. Zhao, Q. Zheng, P. Xu, S. T and X. Wu, "Object detection with deep learning: A review", IEEE transactions on neural networks and learning systems, vol. 30, no. 11, pp. 3212-3232,**(2021)**.*

[25] *X. Wang, A. Shrivastava and A. Gupta, "A-fast-rcnn: Hard positive generation via adversary for object detection", Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2606-2615, **(2017)**.*

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

[27] *J. Redmon and A. Farhadi, "YOLO9000: better faster stronger", Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7263-7271,**(2017)**.*

[28] *J. Du, "Understanding of Object Detection Based on CNN Family and YOLO" in Journal of Physics: Conference Series, IOP Publishin, g, vol. 1004, no. 1, pp. 012029, April* ***(2019)***.
[29] *A. P. Jana and A. Biswas, "YOLO based Detection and Classification of Objects in video records", 3rd IEEE International Conference on Recent Trends in Electronics Information & Communication Technology (RTEICT), pp. 2448-2452,* ***(2018)***.
[30] *Y. Lu, L. Zhang and W. Xie, "YOLO-compact: An Efficient YOLO Network for Single Category Real-time Object Detection," 2020 Chinese Control and Decision Conference (CCDC), pp. 1931-1936,* ***(2020)****, doi: 10.1109/CCDC49329.2020.9164580.*