IJNRD.ORG ISSN : 2456-4184
INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD) | IJNRD.ORG
An International Open Access, Peer-reviewed, Refereed Journal

A Survey on Object Detection Techniques in Image

[1]Jyoti Kumari
[1]Student
[1]Department of Computer Science & Information Technology,
[1]Mahatma Gandhi Central University, Motihari, Bihar, India

Abstract : Comparatively speaking to humans, computers struggle with the process of object detection and recognition. An object detection system finds objects of the real world present either in a digital image or a video, where the object can belong to any class of objects namely humans, cars, etc. This paper presents a review of the various techniques that are used to detect an object, localise an object, categorise an object, extract features in images and videos. The comments are drawn based on the studied literature. This paper is suitable for the researchers who are the beginners in this domain. Decision making tips are included with extensive discussion of the advantage & disadvantage of each technique. The survey presented in this paper can be useful as a quick overview and a beginner's guide for the object detection field. Based on the study presented here, researchers can choose a specific object detection problems for better accuracy in object detection.

IndexTerms – Object Detection, Convolutional Neural Network, Advantage & Disadvantage, Deep learning.

1 INTRODUCTION

Object detection and recognition is a difficult task for computers compared to humans. It involves identifying main objects and understanding the relationships between them and other objects. There are several research results in this area, including methods like HOG and linear SVM, DPM, RCNN, and YOLO. YOLO is a CNN-based end-to-end model that takes the whole image as input and divides it into grids to detect objects. Object recognition and detection are challenging tasks in AI, and convolution neural networks (CNNs) have greatly improved computer vision. Object detection has numerous applications, and research in this area is progressing from single object recognition to multi-object recognition. Traditional algorithms required knowledge of image data and feature detection, but CNNs are self-adaptive and can approximate any function with arbitrary accuracy. Object detection is a challenging problem in computer vision, which involves identifying individual objects in an image & localizing objects in an image. Early models were slow and inaccurate, but the introduction of convolutional neural networks and deep learning revolutionized the field. Object detection is a crucial computer vision task that identifies objects in digital images, such as humans, animals, cars, or buildings. It aims to develop computational models that can determine the location of objects, which is essential for various computer vision applications. While a survey of object detection in image is focused specifically on the task of localizing objects in images.

Image Classification

In computer vision, image classification analyzes a picture and makes an object prediction, whereas object detection not only makes an item prediction but also determines the object's location in terms of bounding boxes. An object detection model would also provide us with the position of the pool, whereas when we create a swimming pool classifier, we take an input image and predict whether it contains a pool.

Image classification Object detection

Figure 1 Difference between classification and object detection

In order to serve as an example, let's assume that a picture contains no more than one class and one object. The output of an object recognition model should therefore include: Probablity that there is an object, Height of the bounding box, Width of the bounding box Horizontal coordinate of the center point of the bounding box, Vertical coordinate of the center point of the bounding box.One convention for specifying output is this one. The idea—to report the object's probability and location—may be modified for different models and implementations, but it always remains the same.

1.1History

Object-based image recognition is a subset of image recognition that involves identifying and segmenting objects within an image. Here is a brief history of object-based image recognition:

In the 1980s, image recognition technology was primarily rule-based and focused on identifying specific features of an object, such as shape or color.

In the 1990s, researchers began to focus on more sophisticated techniques, such as template matching and neural networks, to improve object recognition accuracy.

In the 2000s, advances in computer vision and machine learning led to the development of more sophisticated algorithms for object-based image recognition, including clustering and segmentation techniques.

In 2001, David Lowe introduced the Scale-Invariant Feature Transform (SIFT), which is a powerful technique for detecting and describing local features in images that is invariant to scale, rotation, and illumination changes.

In 2006, a team of researchers introduced the Bag-of-Features (BoF) model, which is a method for representing an image as a histogram of visual words. This technique is widely used in object recognition and classification.

In 2014, the deep learning revolution began, and convolutional neural networks (CNNs) quickly became the dominant approach to object-based image recognition. CNNs are particularly well-suited to object recognition because they can learn to detect hierarchical features automatically from raw pixel data.

In recent years, object-based image recognition has continued to evolve, with researchers developing more sophisticated deep learning techniques, such as recurrent neural networks (RNNs) and attention mechanisms, to improve object recognition accuracy.

Object-based image recognition has come a long way since its early days in the 1980s, and it continues to be an active area of research with many exciting developments on the horizon.

Modern techniques can be divided into two categories: one-stage procedures and two-stage methods.

- Examples of one-stage algorithms that favor inference speed are YOLO, SSD, and RetinaNet.
- Two-stage approaches prioritize accurate detection, and Faster R-CNN, Mask R-CNN, and Cascade R-CNN are three examples of such models.

1.2Application

Object detection is a popular computer vision technique that involves detecting and locating objects within an image or video. It has a wide range of applications in various fields, including:

Autonomous vehicles: Object detection is essential for autonomous vehicles to detect and recognize objects such as other vehicles, pedestrians, traffic signs, and traffic lights.

Security and surveillance: Object detection is used in security systems and surveillance cameras to detect and track suspicious activity, intruders, and anomalies.

Medical diagnosis: Object detection can be used to detect and locate abnormalities in medical images such as X-rays, CT scans, and MRI scans.

Retail: Object detection is used in retail for product recognition, inventory management, and customer tracking.

Robotics: Object detection is used in robotics to enable robots to detect and interact with objects in their environment.

Agriculture: Object detection is used in precision agriculture for crop monitoring, yield estimation, and pest detection.

Sports analysis: Object detection can be used in sports analysis to track the movements of athletes and analyze their performance.

Environmental monitoring: Object detection is used in environmental monitoring to detect and track wildlife, monitor deforestation, and track changes in land use.

Augmented reality: Object detection is used in augmented reality applications to detect and track real-world objects and overlay virtual objects on top of them.

Object detection has many applications in different fields, and its potential uses continue to grow as computer vision technology advances.

1.3Approaches
There are several approaches to object detection in images, but some of the most common ones are:

Region-based Convolutional Neural Networks (R-CNN): This approach involves first proposing regions in the image that could contain objects using algorithms like Selective Search or Edge Boxes. These regions are then passed through a CNN to classify and refine the object boundaries.

Single Shot Detection (SSD): This is a real-time object detection method that uses a single deep neural network to predict object classes and locations directly from the image.

You Only Look Once (YOLO): This is another real-time object detection method that divides the image into a grid and makes predictions for each cell. The predictions include the class of the object and its bounding box.

Faster R-CNN: This is an extension of the R-CNN method that uses the establishment of a Region Proposal Network (RPN) to suggest areas of interest in the image.

Mask R-CNN: This is an extension of the Faster R-CNN method that also predicts object masks in addition to class labels and bounding boxes.

These are just a few examples of the many approaches to object detection in images. The choice of method depends on the specific requirements of the application and the available computing resources.

Advantages and Disadvantages of object detection in image
Object detection in images has several advantages and disadvantages, including:

Advantages:
Automated and Efficient: Object detection in images is an automated process that can be done quickly and efficiently with the use of computer algorithms.

Accurate: With the advances in deep learning techniques, object detection models can achieve high accuracy in detecting and localizing objects in images.

Versatility: Object detection can be used in a wide range of applications such as surveillance, autonomous driving, medical diagnosis, and many more.

Real-Time Detection: Object detection models can be optimized for real-time detection, which is essential in many applications, such as self-driving cars.

Disadvantages:
Data Requirements: Object detection models require large amounts of data to train, which can be expensive and time-consuming to collect and annotate.

Complexity: Object detection models are often complex and require significant computational resources, which can make them difficult to implement on low-powered devices.

Variability: The performance of object detection models can be affected by factors such as lighting conditions, object occlusion, and camera perspective.
Vulnerability: Object detection models are vulnerable to adversarial attacks, where subtle modifications to the input image can cause the model to misclassify or fail to detect objects.

Privacy Concerns: Object detection technology can raise privacy concerns, particularly in applications such as surveillance, where it can be used to monitor individuals without their consent or knowledge.

The advantages of object detection in images outweigh the disadvantages, and the technology is increasingly being used in a wide range of applications, with ongoing research to address the limitations and challenges.

2.Literature Review
The use of selective search to generate candidate regions and the combination of CNNs and SVMs for classification and regression were major breakthroughs in the development of deep learning models for object detection. However, RCNN has some limitations, such as its slow speed due to the need to extract features for each candidate region independently. This led to the development of subsequent models that addressed these limitations [1]

The use of graph models reduces the matching space, making it more efficient to perform the recognition and localization tasks.Furthermore, the passage discusses the use of contextual information in OD, which can be extracted from images using visual features like color and texture, saliency, and object co-occurrences. Lastly, the passage describes various methods for OD, such as

boosting, watershed-based transformation, object descriptor, clustering, color processing, Haar-like features, histogram features, genetic programming, and deformable models. Each method has its advantages and limitations depending on the nature of the objects, image quality, occlusion, clutter, and other factors[2].

The authors also describe various visualization techniques that can help to interpret the output of object detection models. Finally, the paper concludes with a discussion of the current challenges and future research directions in the field. The authors identify several open research questions, such as developing more efficient and accurate object detection algorithms, improving the robustness of models to different environmental conditions, and exploring the potential of deep learning-based approaches for multi-object tracking and video analysis.[3].

OpenCV is a set of libraries in Python that can be used for image processing, video capture, and analysis. The data set used for the object detection task is a collection of labelled images of all the objects that are to be identified. The data set is usually divided into three subsets, namely the training dataset, the validation dataset, and the testing dataset. The training dataset is used to train the model, the validation dataset is used for validation purposes, and the testing dataset is used to test the performance of the model[4].

The author compares the performance of different object detection techniques on various datasets, such as PASCAL VOC, MS COCO, and ImageNet, and provides a detailed analysis of the results. Furthermore, the author discusses the challenges and future directions in object detection research, such as improving the accuracy of detection, reducing computational complexity, handling large-scale datasets, and addressing domain-specific challenges, such as detecting objects in aerial or underwater images.[5].

They highlight the key challenges faced in object detection such as occlusion, scale variation, and cluttered background. The authors also provide an overview of the popular datasets used for training and evaluation of object detection models such as COCO, PASCAL VOC, and ImageNet. The paper then delves into the details of various deep learning based object detection models such as R-CNN, Fast R-CNN, Faster R-CNN, YOLO, SSD, and Retina Net. The authors discuss the key features, strengths, and weaknesses of each of these models. They also compare these models in terms of accuracy, speed, and memory requirements. Furthermore, the authors provide an insight into the recent advancements in object detection such as one-stage detectors, anchor-free detectors, and transformers-based detectors. They also discuss the challenges faced by these models and highlight the potential areas for future research.[6].

They describe the architecture of popular deep learning models for object detection, such as R-CNN, Fast R-CNN, and Faster R-CNN. They also discuss one-stage object detection models, such as YOLO and SSD, which are faster but less accurate than two-stage models. The authors then discuss some recent advances in object detection, such as object detection in videos, object detection with semantic segmentation, and object detection with attention mechanisms. They also briefly touch upon some other related tasks, such as object tracking and object recognition. [7].

The authors then provide an overview of the various datasets used for training and testing deep learning models for underwater object detection. In addition, the authors discuss the future directions for research in this field, including the need for more extensive datasets, the development of more robust deep learning architectures, and the exploration of multi-modal sensor fusion techniques. Overall, "Deep Learning on Underwater Marine Object Detection: A Survey" is a well-written and comprehensive literature survey that provides an excellent overview of the recent advancements in this field. The article is a valuable resource for researchers and practitioners interested in underwater marine object detection and deep learning techniques[8].

The paper also discusses various CNN-based object detection models, including the Faster R-CNN, You Only Look Once (YOLO), and Single Shot MultiBox Detector (SSD) models. The authors compare these models in terms of speed, accuracy, and efficiency. The authors then provide a detailed explanation of the Faster R-CNN model and its components, including the Region Proposal Network and the Fast R-CNN. They also provide experimental results to demonstrate the effectiveness of the Faster R-CNN model in object detection tasks. The paper concludes by summarizing the advantages of CNN-based object detection models over traditional object detection methods. The authors highlight the ability of CNNs to learn features automatically, the high accuracy achieved by these models, and the ability to process images in real-time[9].

The paper also presents a review of some of the recent advances in deep learning-based object detection, including the use of convolutional neural networks (CNNs), region proposal networks (RPNs), and multi-scale feature fusion. The authors then discuss the various datasets used for on-road object detection, including KITTI, Cityscapes, and ApolloScape. They describe the characteristics of each dataset and how they have been used to evaluate the performance of object detection algorithms. The paper concludes by discussing some of the challenges and future directions for on-road object detection, including the need for real-time performance, robustness to environmental changes, and the integration of object detection with other tasks such as semantic segmentation and path planning. [10].

Here we have some images that shows objects in images, Some of objects.

Figure 2 Bike[4]

Figure 3 Train[4]

Figure 4 Car[4]

The paper titled "Comparative Study of Object Detection Algorithms," authored by Nikhil Yadav and Utkarsh Binay, presents an in-depth analysis and comparison of various object detection algorithms. The primary focus of the study is to evaluate and contrast the performance of different algorithms designed for object detection tasks in computer vision.The authors begin by introducing the significance of object detection in computer vision applications and highlight its relevance in areas such as autonomous driving, surveillance, and image analysis. They acknowledge the rapid advancements in the field and the emergence of various object detection methods tailored to different requirements.Throughout the paper, the authors systematically review and compare several object detection algorithms, possibly including popular ones like CNN-based methods, R-CNN, Fast R-CNN, Faster R-CNN, YOLO, and SSD. They outline the key features, characteristics, and operational mechanisms of each algorithm. Moreover, the authors delve into the specific strengths and limitations associated with each approach[11].

Table 1 Detailed comparison of existing surveys and reviews on object detection

Year	Topic	Remarks
2018	Research of Image Main Objects Detection Algorithm Based on Deep Learning	This paper likely presents a specific algorithmic approach for detecting main objects in images using deep learning techniques.
2017	A Review and an Approach for Object Detection in Images	This paper might offer insights into existing object detection techniques and suggest a novel approach for addressing the problem.

2020	Deep Learning in Object Detection: a Review	This review likely provides an overview of various deep learning methods applied to object detection, discussing their strengths and weaknesses.
2020	REAL TIME OBJECT DETECTION WITH DEEP LEARNING	This paper may present an approach to achieve real-time object detection using deep learning models.
2012	Survey of The Problem of Object Detection In Real Images	This survey likely examines the challenges and existing solutions related to object detection in real-world images.
2022	A survey of modern deep learning based object detection models	This survey could provide an overview of the state-of-the-art deep learning models and techniques used for object detection.
2013	A Survey on Approaches of Object Detection	This survey might cover a wide range of approaches and methods used for object detection.
2017	Deep Learning on Underwater Marine Object Detection: A Survey	This survey could explore the challenges and advancements in using deep learning for detecting objects underwater.
2018	Object Detection Using Convolutional Neural Networks	This paper likely investigates the use of convolutional neural networks for object detection tasks.
2016	On-road object detection using Deep Neural Network	This paper may present an approach for detecting objects on roads using deep neural networks.
2017	Comparative Study of Object Detection Algorithms	This paper could compare and analyze different object detection algorithms to understand their strengths and weaknesses.

2 Methodology

The OD system basically comprises of two main phases namely: the learning phase and the testing phase that shows the normal working of the OD system. Learning phase is mainly meant for the classifier so that it recognises the objects present in the image that is given as input to the system. Learning phase can be further classified as learning through training and learning through validation.The main purpose of the testing phase is to decide whether an object is present in the image that is given to the system as input and if yes then to which object class does it belongs to. Here the image is searched for an object by various searching techniques like the sliding window technique, and according to the output of the searching mechanism, a decision is made on the object class.

3.1 Problem Defination

The problem definition of object detection in images is to develop an algorithm or system that can automatically identify and locate objects of interest within an image.The system should also be able to accurately distinguish between objects and background noise or clutter in the image.Object detection in images is a challenging problem because images can vary in lighting conditions, object size and orientation, clutter, and occlusion. The solution to this problem requires developing sophisticated computer vision algorithms that can learn to detect objects under varying conditions and generalize to new, unseen images.Object detection in images is a critical task in many applications, including self-driving cars, surveillance, robotics, medical imaging, and more. As such, it is an active area of research in computer vision and machine learning.

A **dog** is standing on a hardwood floor A group of **people** sitting on a boat in the water

Figure 5 Images can vary in lighting conditions

3.2 Objective of Thesis

Object detection is the process of identifying and locating specific objects within an image using computer vision techniques. The main objectives of object detection in image are as follows:

Object recognition: The primary objective of object detection is to accurately recognize the objects present in an image. Object recognition involves identifying the type of object, such as a person, car, animal, or any other specific *object.*

Object localization: Object detection also aims to locate the objects within an image accurately. This involves identifying the position and size of the object within the image.

Object classification: Object detection can also be used to classify objects into different categories or classes. This involves assigning a label to each object in the image based on its type or category.

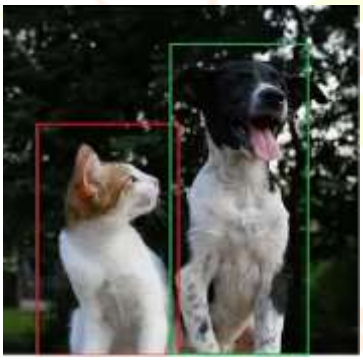
Object tracking: Object detection can also be used to track objects across different frames of a video. This involves identifying the objects in each frame and tracking their movement across frames.

Object segmentation: Object detection can also be used to segment the objects in an image from the background. This involves separating the objects from the surrounding environment and creating a mask for each object.

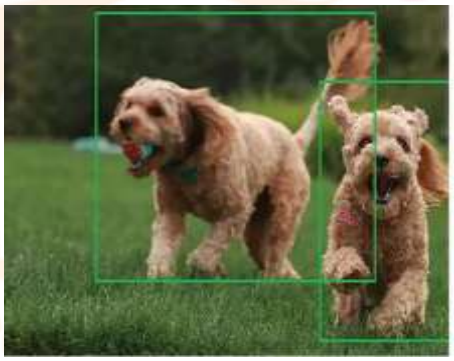
Image belongs to Dog Class

Image belongs to Cat Class

Figure 6 Object Classification (Assigning label of object)



cat detected on the image dog detected on the image



dog detected on the image dog detected on the image

Figure 7 Object Detection

Overall, the main objective of object detection in image is to provide accurate and reliable information about the objects present in an image to enable more advanced computer vision tasks like scene understanding, video analysis, and autonomous driving.

Table 2 Object Detection(Methods, Uses, Features, Region Proposal, Speed, Accuracy, single/Multi Stage)

Model	Uses	Features	Region Proposal	Speed	Accuracy	Single/Multi-Stage
CNN	Feature extraction for object detection and other computer vision tasks.	Feature extraction using convolutional layers	None	Variable	Variable	Single
R-CNN	Introduced the idea of region proposals for object detection.	Region proposals, CNN-based features, refinement	External	Slow	Moderate	Multi
Fast R-CNN	Improves efficiency of R-CNN by sharing	RoI pooling, shared	External	Moderate	Moderate	Multi

	computation and using RoI pooling.	computation				
Faster R-CNN	Enhances efficiency by integrating region proposal generation with the network.	Integrated RPN, shared computation	Integrated	Fast	High	Multi
YOLO	Real-time object detection in video streams and images.	Grid-based prediction, bounding box regression	None	Very Fast	Good	Single
SSD	Balances accuracy and speed for real-time object detection.	Multi-scale prediction, bounding box regression	None	Fast	Good	Single

4.Future Work

An overview of related work is presented in response to the proposed research work, which mainly includes modern object detection in state-of-the-art method.

We will extend a work, proposed work in which suitable neural network based technique will be incorporated for identifying the object in image. The performance of employed technique will also be compared with its contemporary works.

5 Conclusion

This paper presents the review of the various methods for detecting objects in images as well as in videos. The process of Object Detection is classified into state-of-the-art method using multi stage & single stage This paper provides the details of the existing approaches which is used as the base for development of the approach. This paper is useful for the study purpose as well as for the new researchers who want to explore the Object Detection research area. The advantages and disadvantages of the works are discussed and helpful in choosing a suitable technique Thus, the paper presents a concise summary of the state-of-the-art techniques in object detection for upcoming researchers. Object detection is still one of the most essential deep learning and computer vision applications to date. Now, this paper presents Methods, Uses, Features, Region Proposal, Speed, Accuracy, single/Multi Stage of CNN, Fast R-CNN, Faster R-CNN, YOLO, and SSD .This paper also presents the literatures remarks of each topic. In the future, a lot more successful algorithms and libraries for object detection still await us.

ACKNOWLEDGMENT

I have made this dissertation report on the topic "**A Survey on Object Detection Techniques in Image**". I have also tried my best in this dissertation to explain all the related detail. I would like to express my sincere gratitude towards my supervisor **Dr. Atul Tripathi, Department of CS & IT**, for providing excellent guidance, encouragement, inspiration, and constant as well as timely support throughout this ongoing dissertation work.

REFERENCES

[1] L. Yu, X. Chen, and S. Zhou, "Research of image main objects detection algorithm based on deep learning," in *2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)*, IEEE, 2018, pp. 70–75.

[2] K. U. Sharma and N. V Thakur, "A review and an approach for object detection in images," *Int J Comput Vis Robot*, vol. 7, no. 1–2, pp. 196–237, 2017.

[3] K. L. Masita, A. N. Hasan, and T. Shongwe, "Deep learning in object detection: A review," in *2020 International Conference on Artificial Intelligence, Big Data, Computing and Data Communication Systems (icABCD)*, IEEE, 2020, pp. 1–11.

[4] D. Pavan, "REAL TIME OBJECT DETECTION USING DEEP LEARNING," 2020.

[5] D. K. Prasad, "Survey of the problem of object detection in real images," *International Journal of Image Processing (IJIP)*, vol. 6, no. 6, p. 441, 2012.

[6] S. S. A. Zaidi, M. S. Ansari, A. Aslam, N. Kanwal, M. Asghar, and B. Lee, "A survey of modern deep learning based object detection models," *Digit Signal Process*, p. 103514, 2022.

[7] S. Shantaiya, K. Verma, and K. Mehta, "A survey on approaches of object detection," *Int J Comput Appl*, vol. 65, no. 18, 2013.

[8] M. Moniruzzaman, S. M. S. Islam, M. Bennamoun, and P. Lavery, "Deep learning on underwater marine object detection: A survey," in *Advanced Concepts for Intelligent Vision Systems: 18th International Conference, ACIVS 2017, Antwerp, Belgium, September 18-21, 2017, Proceedings 18*, Springer, 2017, pp. 150–160.

[9] R. L. Galvez, A. A. Bandala, E. P. Dadios, R. R. P. Vicerra, and J. M. Z. Maningo, "Object detection using convolutional neural networks," in *TENCON 2018-2018 IEEE Region 10 Conference*, IEEE, 2018, pp. 2023–2027.

[10] H. Kim, Y. Lee, B. Yim, E. Park, and H. Kim, "On-road object detection using deep neural network," in *2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia)*, IEEE, 2016, pp. 1–4.

[11] N. Yadav and U. Binay, "Comparative Study of Object Detection Algorithms," *International Research Journal of Engineering and Technology*, 2017, [Online]. Available: www.irjet.net.