# Deep Learning-Based Object Detection: Advances, Techniques, and Applications


Arshdeep Yadav
Apex Institute of Technology-AIML
Chandigarh University, Mohali,Punjab,India
22BAI70007@cuchd.in

Mayank Chauhan
Apex Institute of Technology-AIML
Chandigarh University, Mohali,Punjab,India
22BAI70050@cuchd.in

Devanshu Arora
Apex Institute of Technology-AIML
Chandigarh University, Mohali,Punjab,India
22BAI71425@cuchd.in

Gokula Krishnan G V
Apex Institute of Technology-AIML
Chandigarh University, Mohali,Punjab,India
22BAI70054@cuchd.in

Mr.Jaswinder Singh
Apex Institute of Technology-AIML
Chandigarh University, Mohali,Punjab,India
jassi724@gmail.com



***Abstract*— Deep learning has significantly improved object recognition by overcoming the limitations of traditional techniques. This paper examines modern deep learning frameworks specifically designed for object recognition, evaluating and comparing their performance using deep neural networks. It emphasizes the importance of employing object recognition techniques tailored to specific categories like faces, buildings, and plants. The study discusses the advantages and disadvantages of current methods and suggests potential future directions for research in object recognition.**



**Keywords— computer vision, deep learning, yolo, object detection, python,**


## INTRODUCTION

The rapid progress of AI-powered assistants, driven by advancements in machine learning, speech recognition, and natural language understanding, has the potential to improve various aspects of our lives. However, we must also address pressing challenges such as traffic accidents, a leading cause of death worldwide, and medical misdiagnosis. Between now and 2030, 500 million individuals will either die or sustain permanent injuries in traffic accidents, resulting in a financial loss exceeding $25 trillion[1]. Object detection, a crucial task in computer vision, holds great promise in addressing these diverse challenges by identifying and detecting specific visual objects in digital images and videos.

Despite significant investment, the development of fully autonomous vehicles has not yet reached the desired level of functionality, indicating the need for more advanced object detection capabilities. To make object detection technology accessible to the general public, further research is required to reduce associated costs. Additionally, integrating object detection into various domains, such as road safety and medical science, can prevent accidents and save lives.

This research aims to deepen our understanding of object detection through deep learning, which could lead to discoveries and solutions to these challenges. By exploring the intricacies of this field, the study has the potential to uncover new insights and pave the way for enhanced comprehension and advancements. Object detection holds great promise in addressing these issues, but further research is necessary to overcome current limitations, reduce costs, and integrate these technologies into diverse domains.

### *Identification of Problem*

Deep learning-based object recognition systems for the transportation safety and medical diagnostics industries could help solve the challenges and issues that still face our technologically evolved world. In order to comprehend and assess obstacle detection, this might be accomplished by further investigating and understanding object detection. The many everyday traffic accidents that cause injuries and monetary losses, the incorrect diagnoses made during medical testing that result in fatalities and disabilities, and the obstacles that impede the development of driverless cars despite large investments are just a few of the issues raised. The ability of object detection, especially in the context of deep learning, could offer novel solutions to these problems. The study report seeks to deepen knowledge of object detection and its uses in other sectors where object detection can be implemented.

## BACKGROUND STUDY

### *Timeline of the Reported Problem*

In many domains, including computer vision and artificial intelligence, object detection is critical. CNN and deep learning play a key role in this process. In industries like traffic safety and medical imaging, where the growing usage of these technologies produces a large volume of medical pictures, object detection is essential. Nonetheless, deciphering and evaluating these pictures may result in misunderstandings, inaccurate diagnosis, and unsuitable care, which may cause patient death or irreversible impairments. According to CNN Health, misdiagnosis results in 424,000 people having lifelong disability and 371,000 deaths per year[2]. Technological developments in medical imaging object recognition may help to reduce these hazards by giving medical personnel fast, accurate help when analyzing medical pictures, which might enhance patient outcomes and diagnostic precision.

***Researching object detection is imperative for several reasons:***

The objective of developing autonomous vehicle technology has driven significant investments in technology and research, particularly in object detection. However, the unpredictable nature of data underscores the need for further advancements in this field. Notably, object detection technology is expensive, necessitating further research to reduce costs and enhance accessibility for a broader audience. Moreover, object detection technology holds great potential in improving road safety and saving lives by significantly reducing accidents. Nevertheless, more research is required to optimize the effectiveness of these strategies before integrating them into other domains[5]. This deep learning research project holds immense significance as it will illuminate new concepts and drive rapid progress in this transformative domain.

***Why Use Deep Learning for Object Detection?***

Deep learning, a subset of machine learning, is a powerful tool that enhances prediction analysis by employing deep neural networks (DNNs) that mimic the decision-making processes of the human brain. These models are trained with large datasets to detect and classify phenomena, recognize patterns and relationships, evaluate possible outcomes, and provide predictions and judgments. Deep learning can involve both descriptive and predictive analysis, depending on the task at hand. Descriptive analysis involves exploring and understanding data patterns, structures, and relationships, such as in image recognition tasks. Deep learning models can extract features, identify clusters, and visualize data in meaningful ways, contributing to descriptive analytics. In contrast, predictive analysis uses data to make predictions or decisions, identifying patterns and relationships in data. In summary, deep learning is a powerful tool that can support descriptive analysis by uncovering insights and patterns in data.

***Advantages Of Using Deep Learning in Object Detection:***

- Deep Neural Networks (DNN): Applications employing deep learning (DL) benefit from training neural networks rather than traditional programming methods, reducing the need for extensive expert analysis and fine-tuning. Additionally, they leverage the vast reservoir of video data prevalent in contemporary systems.
- High Accuracy: Deep learning models, particularly convolutional neural networks (CNNs), have demonstrated superior performance in object detection tasks. They can learn intricate patterns and features from data, leading to highly accurate detection results.
- Flexibility: Deep learning models can be trained to detect a wide range of objects in various environments and conditions. These models can adapt to variations in lighting, background, scale, and orientation, making them suitable for diverse applications.
- End-to-End Learning: Deep learning enables end-to-end learning, where the model learns to extract features and perform detection tasks directly from raw data like images or video frames. This eliminates the need for manual feature extraction, simplifying the development process.

Given the numerous benefits of deep learning, including its high accuracy, adaptability, scalability, and efficiency in processing large datasets, we have chosen to prioritize the use of deep learning techniques for object detection in this project.

***Generic Or Traditional Object Detection Technique***

Traditional techniques like SIFT, SURF, and BRIEF were used for object detection before the rise of deep learning in computer vision. These methods involved feature extraction, where small, informative image patches were identified using algorithms like edge detection, corner detection, or threshold segmentation[4]. These features were then used to create a representation for each object class, often referred to as a bag-of-words. If a significant number of features matched those in a particular bag-of-words, the image was classified as containing the corresponding object. However, the manual selection of significant features for individual images became a challenge, and as the number of object classes increased, the process became more complex. Computer vision (CV) engineers were tasked with determining which features effectively characterize various object classes through iterative experimentation and optimizing numerous parameters. Deep learning introduces the concept of end-to-end learning, where a machine learns directly from annotated datasets containing images and their corresponding object classes. Deep neural networks (DNNs) deliver superior performance compared to traditional algorithms, but there are trade-offs in terms of computational requirements and training duration.

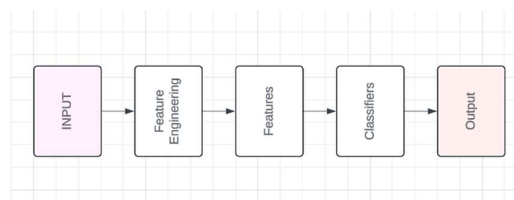



Figure 1

***Existing Models for Object Detection***

Region-based object detection algorithms are widely used for identifying objects in images. These models usually start by finding regions within the image and then perform classification on those regions. For example, R-CNN uses selective search to extract regions of interest, followed by regression and classification of the image. Another prominent model, Faster R-CNN, includes a region proposal network to generate regions of interest by refining predefined boxes. It then uses a region-based prediction network to detect objects within these regions.

***A. Region-Based Convolutional Neural Networks(R-CNN)***

R-CNN utilizes a strategy where numerous bounding boxes, referred to as regions, are proposed within an image to identify whether any of them contain an object. These regions are extracted using a technique called selective search, which generates approximately 2000 region proposals. These proposals are then cropped and resized to conform to the input requirements of a CNN feature extractor. The CNN extracts a multi-dimensional feature vector from each region, which is then used for class prediction by multiple classifiers. Each class is associated with a support vector machine (SVM) for object occurrence classification based on the feature vector[7]. Furthermore, a linear regressor predicts four offset values to improve the precision of the chosen bounding boxes and minimize localization errors.

The R-CNN operates through three primary steps:

- Utilize selective search to scan the input image and propose approximately 2000 candidate boxes, potentially containing objects.
- Apply a CNN to each candidate box for feature extraction.
- Transmit the CNN's results to an SVM for classification and to a linear regressor to refine the bounding box coordinates of the detected objects.

Despite its intuitive workflow, R-CNN is known for its slow processing speed.

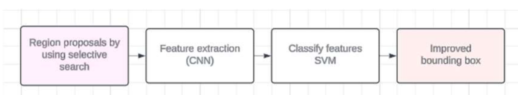



Fig2: R-CNN

### *B. Fast Region Based Convolutional Network (Fast R-CNN)*

The Fast Region-Based Convolutional Network (Fast R-CNN) represents a significant advancement over the original R-CNN. It streamlines object detection by directly inputting the entire image into a Convolutional Neural Network (CNN). Fast R-CNN employs a Region of Interest (RoI) pooling layer to efficiently extract features for each proposed region, leading to a significant reduction in computational costs. By integrating region proposal generation and feature extraction into a single network, Fast R-CNN achieves efficient object detection with faster processing compared to its predecessor.

### *C. Faster Region-based Convolutional Network (Faster R-CNN)*

In contrast to previous models that rely on a slow and time-consuming process called selective search to identify potential object regions, Faster R-CNN introduces a novel concept known as Region Proposal Networks (RPN). This groundbreaking innovation allows the network to directly generate region proposals, thereby eliminating the need for selective search. With Faster R-CNN, the network first generates region proposals and then proceeds with box prediction, resulting in a streamlined and significantly more efficient overall process. Faster R-CNN enhances the efficiency of object detection by introducing a Region Proposal Network (RPN) that shares convolutional layers with the detection network. This setup enables simultaneous object identification and region proposal generation. The RPN generates region proposals, which are then refined and classified by the detection network. By integrating these two processes, Faster R-CNN achieves both improved speed and accuracy in object detection tasks.

Despite numerous advancements, the Faster R-CNN technology still falls short of achieving real-time object detection capabilities. Detecting smaller and more intricate objects further exacerbates this perpetual challenge.

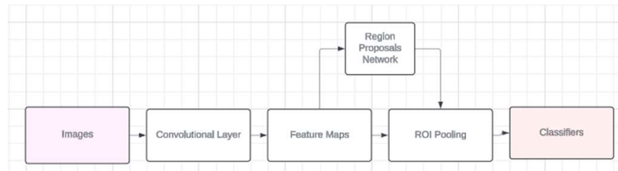



Fig3: Faster R-CNN

## DESIGN FLOW / PROCESS

### *A. Evaluation & Selection of Specifications / Features*

Within the realm of deep learning-based object detection, the judicious selection of appropriate specifications and features assumes paramount importance in the relentless quest for achieving optimal performance. This intricate process entails a meticulous evaluation of the stringent requirements of the target application and the distinctive characteristics inherent in the available data

**1. Application Requirements:** Embark on the initial juncture of establishing an object detection system by meticulously outlining its indispensable application requirements. Encompassing a comprehensive array of factors, these prerequisites may embrace:

a. Object classes: Precisely discern and categorize the diverse array of objects that necessitate detection and analysis, encompassing vehicles of various types, pedestrians traversing the cityscape, traffic signs conveying pertinent information, and medical abnormalities requiring prompt attention.
b. Accuracy requirements: Ascertain the appropriate level of precision for object detection, which may fluctuate in accordance with the mission-critical nature of the application (for instance, the medical realm demands an exceptionally high degree of accuracy).
c. Real-time performance: Critically evaluate whether the object detection system necessitates real-time functionality or if batch processing adequately satisfies the requirements. Consider scenarios where immediate response and low latency are paramount, as well as instances where

accumulated data analysis suffices.

d. Hardware constraints: Take into account the computational resources accessible for model training and inference, such as the cutting-edge capabilities of advanced GPUs or the inherent constraints imposed by edge devices.

**2. Data Characteristics:** The subsequent phase entails a meticulous examination of the inherent qualities embedded within the readily accessible data, which will serve as a guiding beacon in the discernment of suitable features and the precise delineation of model specifications[3].

2.1 Image Characteristics

a. Resolution: Scrutinize the resolution levels of the images, since resolutions that are notably higher may necessitate more extensive computational resources, yet they can simultaneously furnish superior accuracy.
b. Variability: Critically evaluate the spectrum of variability within the image data, encompassing diverse lighting conditions, intricate backgrounds, and multifaceted object orientations. This rigorous assessment ensures the model's exceptional ability to generalize effectively across a broad range of scenarios.
c. Object Sizes: It is crucial to ascertain the spectrum of object sizes represented within the images, taking into account that objects of diminished dimensions may require specialized methodologies or heightened input resolution.

2.2 Annotation Quality

a. Bounding box accuracy: Critically evaluate the accuracy and precision of the bounding box annotations, as even moderately imprecise annotations can adversely affect downstream model performance and potentially compromise the integrity of the entire machine learning pipeline.
b. Class Imbalance: Vigilantly identify any class imbalances lurking within the meticulously annotated data, where certain object classes may be woefully underrepresented in comparison to their counterparts, leading to potential biases in the model's predictions.

**3. Feature Selection:** Based on the intricacies of the application requirements and the distinctive characteristics of the data, meticulously select the most appropriate features for the object detection model. These features may encompass:

a. **Input size:** Ascertain the most suitable input size for the model, maintaining a delicate equilibrium between achieving high accuracy and accommodating the computational demands of the system, ensuring optimal performance without sacrificing efficiency.
b. **Feature extractors:** For encoding visually pertinent information, meticulously select feature extractors, such as intricately intricate convolutional layers or proficiently pre-trained networks, that exhibit exceptional aptitude in capturing and representing the underlying patterns and structures within visual data.
c. **Localization and classification components:** Meticulously select the most appropriate architectures for localizing objects, such as intricate region proposal networks that exhibit exceptional performance in handling complex object shapes or precisely aligned anchor boxes that ensure pinpoint accuracy in object localization. Furthermore, employ cutting-edge classification algorithms to categorize the detected objects with utmost precision, leveraging deep learning models that have been rigorously trained on extensive datasets to achieve state-of-the-art classification accuracy.

**4. Model Selection:** Upon meticulously assessing the intricate specifications and advanced features, judiciously select the most appropriate deep learning architecture for object detection, ensuring optimal performance and accuracy. Prominent options encompass Faster R-CNN, renowned for its remarkable precision, YOLO, distinguished by its exceptional speed, and SSD, characterized by its efficient computational requirements. Each architecture possesses distinct strengths and entails specific trade-offs in terms of accuracy, speed, and computational complexity, necessitating careful consideration to align with the specific project objectives.

**5. Validation and Refinement:** Subject the chosen specifications and attributes to rigorous validation by meticulously training and thoroughly evaluating the model against a comprehensive validation dataset. Analyze and refine specifications and features with meticulous attention, taking into account critical factors such as accuracy, inference time, and efficient resource utilization.

Through meticulous scrutiny of the application's intricate requirements, discerning examination of data characteristics, and comprehensive assessment of available features, you can judiciously select the most optimal specifications and model architecture for your state-of-the-art deep learning-based object detection system, thereby establishing a robust and reliable foundation for the subsequent design and implementation phases.

### B. *Design Constraints*

In the design and implementation of a deep learning-based object detection system, various constraints need to be considered to ensure optimal performance, efficiency, and practical deployment. These constraints can significantly impact the model architecture, training strategies, and overall system design. It is crucial to identify and address these constraints early in the development process to avoid potential bottlenecks or limitations later on.

**1. Computational Resource Constraints:** AI models need a lot of processing power for training and inference, especially those made for object detection. GPUs are essential for fast training and effective inference because of their large memory, powerful computation, and complex architecture. The undisputed AI champs, GPUs carry out the intricate calculations that enable deep learning's amazing achievements. But CPUs are also crucial for data pre- and post-processing, among other crucial operations that are the cornerstone of every effective AI system. These CPUs' many cores, fast clock speeds, and large cache sizes can all have a big impact on system performance, especially when inference is being done on devices with limited resources. Because every millisecond saved might lead to new opportunities, GPUs and CPUs are crucial parts of the AI system.

**2. Memory and Storage Constraints:** Deep learning models and datasets often require extensive memory and storage due to their large sizes and intricate structures. These models can vary significantly in terms of the number of parameters, precision, and architectural complexity. Techniques such as compression or specialized hardware may be necessary to facilitate efficient deployment on resource-constrained devices or in challenging edge scenarios. Furthermore, the substantial datasets utilized for object detection tasks also strain memory and storage resources. To address this challenge, effective data management strategies become crucial, including distributed storage solutions or data sharding approaches. These methods involve splitting the dataset into manageable chunks and leveraging multiple storage systems to enable efficient data processing and retrieval.

**3. Real-Time Performance Constraints:** Strict latency constraints require autonomous cars and surveillance systems to perform in real time or near-real time. These constraints affect the selection of inference techniques and model design, necessitating a thorough assessment of the trade-offs between accuracy and computational efficiency. In real-time applications, inference latency—the time it takes to process a single input sample—must be considered. Striking a balance between computational efficiency and model accuracy may be necessary to achieve minimal latency. High throughput requirements, such as processing multiple input samples simultaneously, influence advanced model parallelization techniques, batching strategies, and hardware accelerators[8].

**4. Deployment and Integration Constraints:** Platform deployment for both software and hardware is essential for a smooth system integration. The chosen model architecture must be flexible, including specialized optimizations, quantization methods, and libraries, to guarantee compatibility. This is especially crucial for cloud platforms, embedded systems, and edge devices. Along with seamlessly integrating with current systems, the object detection system needs to handle issues with software interfaces, data formats, and communication protocols. In order to guarantee the system's operation and compatibility with the intended hardware and software environments, meticulous design and implementation are required.

Through the proactive identification and meticulous resolution of these design constraints at an early stage of the development process, you can confidently ensure that the cutting-edge deep learning-based object detection system seamlessly aligns with the meticulously defined performance objectives, efficiency benchmarks, and practical deployment necessities. This comprehensive and proactive approach effectively sets the stage for a triumphant and enduring implementation.

### *C. Design Flow*

The design flow for a deep learning-based object detection system outlines the sequential steps and processes involved in the development and implementation of the system. This structured approach ensures a systematic and organized workflow, facilitating effective collaboration, reproducibility, and maintainability of the project.

**1. Data Acquisition and Preparation**

a) Collect and curate a diverse dataset of images relevant to the target application domain.
b) Annotate the images with bounding boxes around the objects of interest, ensuring accurate and consistent labeling.
c) Preprocess the data by resizing, normalizing, and applying data augmentation techniques as needed.
d) Split the dataset into training, validation, and testing subsets[9].

**2. Model Architecture Design**

a) Select an appropriate deep learning architecture for object detection, considering factors such as accuracy, speed, and computational requirements.
b) Design the backbone network for feature extraction, leveraging transfer learning from pre-trained models if applicable.
c) Incorporate a region proposal network (RPN) or similar mechanism for generating object proposals.
d) Design the components for object classification and bounding box regression.
e) Include additional modules or components as needed (e.g., attention mechanisms, feature pyramid networks) to enhance performance.

**3. Training Setup**

a) Define the loss function combining components for object classification and bounding box regression.
b) Select an optimization algorithm and determine the hyperparameters (e.g., learning rate, batch size, etc.).
c) Implement regularization techniques (e.g., L1/L2 regularization, dropout, batch normalization) to prevent overfitting.

d) Set up the training schedule, including the number of epochs, learning rate decay strategies, and checkpoint saving.
e) Configure the training environment, including hardware setup (e.g., GPU configuration) and software dependencies.

### 4. Model Training and Evaluation

a) Train the model on the prepared dataset, monitoring the training progress using appropriate evaluation metrics.
b) Evaluate the model's performance on the validation set, analyzing metrics such as precision, recall, average precision (AP), and mean average precision (mAP).
c) Fine-tune the model architecture, hyperparameters, or training strategies based on the evaluation results.
d) Repeat the training and evaluation process until the desired performance is achieved[11].

### 5. Model Optimization and Deployment

a) Optimize the model for deployment, considering techniques such as quantization, pruning, or knowledge distillation to reduce computational requirements.
b) Integrate the optimized model into the target application or system, ensuring compatibility and efficient inference.
c) Implement a monitoring and maintenance strategy to track the model's performance in the production environment and update it as needed.

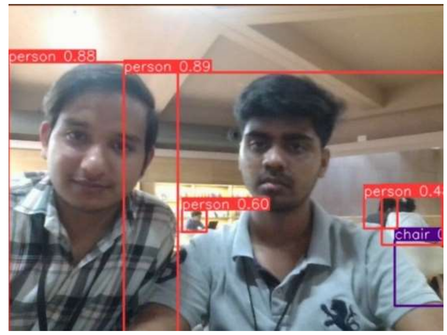
By following this structured design flow, you can ensure a systematic and organized approach to developing a deep learning-based object detection system, facilitating collaboration, reproducibility, and maintainability throughout the project lifecycle.

## RESULTS

This section of our research presents the successful development of a deep learning-based object detection model. This model can detect and identify objects in front of a camera, which has potential applications in the transportation safety and medical diagnostics industries. By overcoming the challenges and issues that still face our technologically advanced world, this research also deepens our understanding of object detection through deep learning by further investigating and comprehending object recognition. This study expands our knowledge of object detection and its uses in various sectors where it can be implemented.

Furthermore, we gained insights into various contemporary machine learning models, including their advantages and disadvantages. We also studied traditional computer vision technologies and compared them to deep learning, which allowed us to recognize the benefits of performing object detection through deep learning.

Through this research, we were able to address various issues that require further investigation to overcome current limitations, reduce costs, and integrate this object detection technology into diverse domains




These two Objects are detected as person with an Accuracy of detection of object in this case as 0.8 out of 1.


The Object present in this case are detected as chairs with an overall Accuracy of 0.85 out of 1.

## FUTURE SCOPE

The future prospects of this research paper on object detection using deep learning are exceptionally promising. This research paper endeavors to construct an object detection model, harnessing the immense power of deep learning, that possesses the capability of detecting objects with highly impressive precision. Furthermore, this research paper significantly enhances our comprehension of object detection, deep learning, and other sophisticated machine learning models advantages and potential shortcomings. Additionally, this research paper contrasts Traditinal computer vision techniques with the cutting-edge paradigm of object detection through deep learning highlighting their strengths and limitations. Through this research, we effectively illuminated challenges that necessitate further investigation with substantially reducing costs, and seamlessly integrating this object detection technology into a diverse spectrum of domains.

Furthermore, future research can focus on improving the accuracy and efficiency of object detection using deep learning. This includes advancements in model architecture and the use of enhanced datasets for diversified training of models, which will ultimately enhance the accuracy of the models. Additionally, this research can aim to provide domain-specific solutions so that object detection is not limited to one domain. This will increase the accuracy in specific domains such as healthcare, agriculture, manufacturing, and security. By using domain-specific solutions, it can easily address unique challenges and requirements within each domain.

## CONCLUSION

This study has successfully developed a deep learning-powered object detection model for identifying and detecting objects, contributing to our comprehension of this field within our technologically advanced society. Through thorough investigation, we have gained deeper insights into object detection utilizing deep learning techniques, broadening our understanding of its applications across diverse sectors.

Moreover, our exploration encompassed various modern machine learning models, analyzing their respective strengths and weaknesses. We conducted a comparative study between traditional computer vision methodologies and deep learning approaches, discerning the advantages of employing deep learning for object detection tasks.

In Conclusion of this research, we have achieved the creation of a robust object detection model based on deep learning principles. Additionally, our comprehension of object detection, deep learning, and alternative machine learning models has been enriched. Furthermore, we have acquired the ability to discern and contrast traditional computer vision techniques with deep learning-based object detection methods.

## REFERENCES


[1]. D. Zhong-Qiu Zhao, Member, IEEE, Peng Zheng, Shou-tao Xu, and Xindong Wu," Object Detection with DeepLearning: A Review" Fellow, IEEE, IN 16 APR 2019.

[2]. D. K. Vaishnavi a, G. Pranay Reddy a, T. Balaram Reddy a, N. Ch. Srimannarayana Iyengar a and Subhani Shaik a "Real-time Object Detection Using Deep Learning".

[3]. D.Ajeet Ram Pathak'* , Manjusha Pandeya, Siddharth Rautaraya, "Application of Deep Learning for Object Detection" Procedia Computer Science 132(2018) 1706-1717

[4]. D. Jun Deng,a, Xiaojing Xuan,b, Weifeng Wang, Zhao Li, Hanwen Yao, ZhiqiangWang "A review of research on object detection based on deep learning", Journal of Physics: Conference Series.

[5]. Andreas Geiger, Philip Lenz, Christoph Stiller and Raquel Urtasun, Vision meets Robotics: The KITTI Dataset, International Journal of Robotics Research (IJRR), 2013.

[6]. Diagnostic errors linked to nearly 800,000 deaths or cases of permanent disability in US each year, study estimates, By Deidre McPhillips, CNN.

[7]. P.Devaki, S.Shivavarsha, G.Bala Kowsalya, M.Manjupavithraa, E.A. Vima " Real-Time Object Detection using Deep Learning and Open CV", International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-12S, October 2019.

[8]. Chinthakindi Kiran Kumar and Kirti Rawal 2022 J. Phys.: Conf. Ser. 2327 012012" A Brief Study on Object Detection and Tracking".

[9]. Introduction to object detection with deep learning https://www.superannotate.com/blog/object-detection-with-deep-learning

[10]. Object Detection in 2024: The Definitive Guide Read more at: https://viso.ai/deep-learning/object-detection/

[11]. Deep Learning-Based Object Detection Techniques for Remote Sensing Images: A Survey https://www.mdpi.com/2072-4292/14/10/2385

[12]. Getting Started with Object Detection Using Deep Learning https://in.mathworks.com/help/vision/ug/getting-started-with-object-detection-using-deep-learning.html

[13]. A comprehensive and systematic look up into deep learning-based object detection techniques: A review https://www.sciencedirect.com/science/article/abs/pii/S1574013720304019