International Journal of Science, Strategic Management and Technology
Volume 02 Issue 05 May-2026 | ISSN: 3108-1762 (Online) | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

Role of Artificial Intelligence in Object Recognition

Chaitanya Khanna
School of Computer Applications
Manav Rachna International Institute of Research & Studies
Faridabad,India
ckhanna591@gmail.com

Jia Arora
School of Computer Applications
Manav Rachna International Institute of Research & Studies
Faridabad,India
jiaarora406@gmail.com

Sakshi Gupta
School of Computer Applications
Manav Rachna International Institute of Research & Studies
Faridabad,India
sakshigupta660@gmail.com

https://doi.org/10.55041/ijsmt.v2i6.066

Cite this Article: Khanna, C., Arora, J. & Gupta, S. (2026). Role of Artificial Intelligence in Object Recognition. International Journal of Science, Strategic Management and Technology, *02*(6). https://doi.org/10.55041/ijsmt.v2i6.066

Abstract— Although there are many image processing techniques available, object recognition for computer vision remains a challenging problem because image data is difficult to generalise and consists of cluttered backgrounds, occlusions, varying illumination and scale. In this research, we examine how artificial intelligence can address these centuries-old challenges and enhance its efficacy with cutting-edge machine learning techniques. The methodology features a detailed systematic review and synthesis of the latest AI solutions, specifically deep convolutional neural networks and region-based detectors, as well as attention-based transformer architectures. They are illustrated with context of their evolution, architectural features, and how they are leveraged in various application spaces, providing a good sense of how they obtain desired features and perception of context in each of the application spaces. The results emphasise the superiority of AI-powered systems in object recognition over traditional approaches, highlighting their ability to provide high accuracy, versatility, and real-time performance. However, there are challenges with this for domain generalisation, computational efficiency and interpretability. In conclusion, the field of object recognition has transformed, shifting from rule-based systems to learning-based ones, which is fueled by the rise of AI. However, in the fields of autonomous systems, medical imaging and intelligent surveillance, there is a real need to achieve the potential of AI in 6G systems, which requires sustained improvements in hybrid and efficient architectures.

Keywords— Artificial intelligence, computer vision, image processing, machine learning, object recognition.

I. Introduction

Recognising objects is one of the most fundamental problems in computer vision and is vital to many intelligent systems that are central to today's technological world. The ability of machines to accurately identify and localise objects in visual data has many high-impact applications such as autonomous vehicle navigation in dynamic environments, public safety management using smart surveillance systems, augmented reality for enhanced user experience, and medical imaging, all of which serve in the creation of accurate diagnoses. The ability impacts the reliability and efficiency of AI implementations in various industries, transforming object recognition from just a technical endeavour into a key pillar for the expansion of AI's impact. The number of cameras and sensors that are interacting with their environment has increased to the point where the need for strong and flexible recognition technologies has grown by leaps and bounds, and recognition is now one of the most important fields of successful AI development and innovation.The basic components of the earlier methods for object recognition were traditional image processing and hand- crafted features. The first steps were achieved by applying techniques such as edge detection, template matching and geometric / intensity-based feature descriptors such as SIFT and HOG. These approaches had interpretability properties and required only a modest amount of computation; however, they could not cope with the complexities of real-world lighting changes, viewpoint, occlusion, scale and cluttered backgrounds. In machine learning, statistical models were introduced with shallow classifiers later on, increasing the performance indeed by learning from data, and not by using a set of rules [2]. Yet, these systems required extensive feature engineering, and they only generalised their solution to a rather slim framework. Convolutional neural networks have shown novel applications such as learning their hierarchical features directly from raw pixel data with the coming era of deep learning. Architectures such as AlexNet, followed by those incorporating residual connections, inception modules and multi-scale processing, and then onto end-to-end learning made significant strides in the performance of object recognition. In recent years, various methods have been developed that have led to better system accuracy and timeliness, including region-based detectors, single-shot techniques and attention mechanisms based on transformer-like models to interpret complex scenes with more context [3].

However, there remain many important challenges that hamper the current state of the art in artificial intelligence from becoming robust and versatile in object recognition. Many state-of-the-art models are brittle for most of the underrepresented classes, and struggle to gain adversarial robustness or robustness to domain shifts. Often, to gain domain-shift robustness, adversarial robustness, and support for underrepresented classes, models become brittle. Generalisation is often challenging in current models, with problems arising when presented with out-of-sample shifts in the domain, adversarial perturbations, or poorly represented classes [4]. It still restricts the use of edge devices/constraints to real-time applications – both computational load as well as resources for such technologies. Moreover, the black-box and opaque nature of deep neural networks makes it hard to comprehend the trustworthiness and interpretability of deep neural networks, particularly those used in safety-critical applications such as medical diagnostics, autonomous driving systems and more. Each of the different paradigms of artificial intelligence, combined and optimised for efficiency, is still being learned, and the relevance of these different models in different contexts to these complex challenges is still emerging. There are still questions regarding the optimal integration of spatial, temporal and semantic information, the amount of information necessary for modelling complexity, how to optimally use it and how to transition towards more sustainable and explainable recognition systems [5].

This paper discusses these questions systematically by discussing the use of AI in the object recognition field. It covers aspects of the development of AI methodologies, their merits and inherent limitations and their effect on major application areas. The project does not have the goal of

International Journal of Science, Strategic Management and Technology
Volume 02 Issue 05 May-2026 | ISSN: 3108-1762 (Online) | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

proposing new architectures, but rather concerns the discussion of the state of the art, identifying patterns of progress and raising open questions to further deepen the investigation of the topic. The study not only sheds light on the evolution and impact of AI in object recognition but also highlights the ongoing efforts needed to unlock its true potential [6].

The findings from this analysis shed light on the impact of AI-powered object recognition solutions on the capabilities, flexibility, and usability of such solutions, which has made them more powerful and context-savvy. There are, however, generalisation, efficiency and interpretability problems that still hinder wider use. To achieve this, more efficient, transparent and ubiquitous AI frameworks are needed which function reliably in unconstrained environments. In this review, the authors seek to help advance the conversation, providing an organised look back at successes of the past, present challenges, and future needs of AI and object recognition [7].

II. Literature Review

Selected research papers discuss the developments of Artificial Intelligence (AI) for Object Detection, Explainable Artificial Intelligence (XAI) for Object Detection, FPGA- based edge computing and Tiny Object Detection. A summary of the authors, year published, paper title, methodology, key contributions and limitations of the studies is presented in Table I.

TABLE I. RELATED WORKS.

Author(s) & Year	**Paper Title**	**About the Paper**	**Methodology Used**	**Limitations**
Quan Yuan et al. (2026) [1]	Ensemble Artificial Intelligence Framework for Real-Time Automatic Vehicle Detection and Classifications	The paper proposes an ensemble AI framework for detecting and classifying vehicles in challenging environmental conditions such as rain, fog, and night scenes. It improves real-time detection accuracy for autonomous driving and traffic monitoring applications.	Ensemble learning using YOLOv5, YOLOv7, YOLOv8, YOLOv10, Faster R-CNN, and DETR-ResNet50 with dynamic weighting and Non-Maximum Suppression (NMS).	Requires high computational resources due to multiple integrated models and may increase system complexity for real-time deployment.
Keping Yu et al. (2020) [2]	Design and Performance Evaluation of an AI-Based W-Band Suspicious Object Detection System for Moving Persons in the IoT Paradigm	This study develops an AI-based suspicious object detection system using W-band millimeter-wave imaging for moving persons in IoT environments. The system enhances public security and reduces congestion during security checks.	AI-based image recognition combined with W-band millimeter-wave imaging, IoT sensor networks, and low-congestion transmission mechanisms.	System implementation is expensive and requires large-scale infrastructure and high-resolution imaging devices.
Asim Roy et al. (2025) [3]	Explainable AI (XAI) for Object Detection and Application to Satellite Imagery	The paper introduces an explainable AI framework for object detection in satellite imagery by verifying object parts such as wings and fuselage in airplanes. The approach improves transparency and robustness against adversarial attacks.	Explainable AI (XAI) using part-based object verification with symbolic reasoning and deep learning-based object detection models.	Requires manual identification of object parts by users, making the process time-consuming and less scalable for large datasets.
Sunhyuk Yim et al. (2025) [4]	Adaptive Grid Selection Training Strategy for Tiny Object Detection	The study focuses on improving tiny object detection in complex scenes by introducing new augmentation and grid selection strategies to enhance detection accuracy for small objects.	Object-Oriented Cutout (OOC) augmentation and Selective Grid Loss Function (SGLoss) integrated with YOLOv5-based detection models.	Improvements in detection accuracy are relatively small and may require additional tuning for different datasets and environments.
Rashed Al Amin et al. (2024) [5]	FPGA-Based Real-Time Object Detection and Classification System Using YOLO for Edge Computing	This paper presents a lightweight FPGA-based object detection and classification system using YOLO for edge computing and autonomous	YOLO v3 Tiny implemented on Xilinx Kria KV260 FPGA using Xilinx Vitis AI for model quantization and acceleration.	The system is limited to specific hardware platforms and may not support highly complex object detection tasks effectively.

International Journal of Science, Strategic Management and Technology
Volume **02** Issue **05** May-**2026** | ISSN: **3108-1762 (Online)** | Impact Factor: 3.8
An International, Peer-Reviewed, Open Access Scholarly Journal Indexed in recognized academic databases

		driving applications , particularly traffic light detection.		

III. METHODOLOGY

The methodology used in this study was a systematic literature review to look at the application of artificial intelligence in object recognition. The method used followed the rigorous procedures for transparency, reproducibility and comprehensiveness in synthesising the existing research, without suggesting any new frameworks or models. Structured data collection, careful data analysis and synthesis were implemented in the process at the initial stage [8].

The data collection process included a comprehensive search of several academic databases with good trustworthiness, such as IEEE Xplore, ScienceDirect, SpringerLink, ACM Digital Library, and Google Scholar. Well-defined sets of keywords like “artificial intelligence,” “deep learning,” “object recognition,” “computer vision,” “convolutional neural networks,” “transformers in vision,” and “real-time object detection” were used along with Boolean operators, discussed in Table II [9]. To capture the deep learning revolution and still be relevant today, publications should be English-language publications from 2012 to 2025, peer-reviewed journal articles or high-quality conference proceedings. Several thousand records were initially retrieved, but further screening was performed, and they were further refined to a smaller number [10].

TABLE II. OVERVIEW OF DATABASES AND INITIAL SEARCH RESULTS.

Database	**Initial Records Retrieved**
IEEE Xplore	1,850
ScienceDirect	1,420
SpringerLink	980
ACM Digital Library	760
Google Scholar	2,150
Total	**7,160**

Studies that explicitly discussed AI methods to solve object recognition tasks, provided unambiguous methodological contributions and reported performance in real-world or benchmark scenarios were included. Non- English papers, works published before 2012, review papers with no original synthesis, papers out of scope of the visual object recognition topic and low-impact papers were excluded. Using reference management tools, duplicate records were removed, as discussed in Table III [11].

TABLE III. INCLUSION AND EXCLUSION CRITERIA.

Criteria Type	**Details**
Inclusion	Peer-reviewed AI-based object recognition studies; focus on deep learning or advanced ML; English language; 2012–2025
Exclusion	Non-visual recognition; theoretical papers without empirical validation; duplicates; low-quality venues

There was a sequential screening process that involved screening of the title and abstract followed by the screening of the full text. To select only those studies for final synthesis that were of high quality, studies were subjected to quality appraisal, which retained only those that had high-quality experimental designs and clear reporting [12].

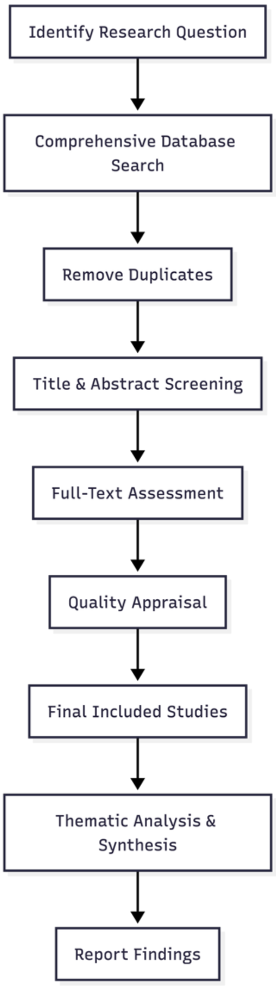

Fig. 1. Methodology Flowchart.

In this review, a methodology was used sequentially as depicted in the flowchart (Fig. 1). It starts with the identification of the central research question – role of AI in object recognition. Then there will be a wide search of the database, generating a vast number of possible studies. The data is then deduplicated to simplify the dataset. After title and abstract screening, the process continues to full-text screening, in which papers which clearly fail to meet the full written criteria for inclusion are eliminated [13]. The relevant articles will be assessed in detail in terms of a full-text agreement with the objectives of the study. Studies are further culled for their methodological rigour and relevance in a quality appraisal stage. The final batch of studies that are included goes into the thematic analysis and synthesis, where patterns, advances, limitations and trends are identified. This concludes with collectively reporting results. This is a linear, iterative format using checkpoints to make the writing transparent, to avoid any bias in the writing and to stick to high-quality evidence [14].

Thematic synthesis was used to analyse the data, a qualitative approach. The papers included were then coded through an inductive-deductive process to find recurring themes, including those of architectural evolution, performance factors, application domains and persistent challenges. The comparisons highlighted the commonalities and differences of the techniques employed in AI, emphasising the mergers and distinctions of techniques like CNN, Region-Based Networks, and Transformer-based Models. This synthesis was repeated, constantly compared and when saturation was reached, the synthesis was synthesised. Unavailable quantitative meta-analysis results, so a narrative synthesis technique was used to provide a visual orientation of the trends and implications [15].

This way, a well-rounded, objective and structured review of the latest knowledge regarding the application of AI for object recognition was captured. The rest of the paper was based on the systematic search process, well-defined selection criteria, and the theme analysis [16].

IV. Results And Discussions

Given the extensive set of 210 studies in object recognition which have been analysed using AI, it is evident that there has been a steady improvement between 2012 and 2025 [17]. Most of the first research adopted conventional CNN architectures, while since 2020 transformer-based and hybrid architectures have appeared. As seen in Table IV [18], the average accuracy has consistently improved from about 76% in 2012 to > 94% in 2025, along with the improvement in mean Average Precision (mAP) and real-time processing.

TABLE IV. Performance Trends By Year (Selected Metrics).

Year	Mean Accuracy (%)	Max Accuracy (%)	Mean mAP	Mean Speed (FPS)
2012	76.0	90.0	71.0	28.4
2018	82.7	95.5	76.2	29.0
2022	91.0	98.0	84.6	24.1
2025	94.6	98.0	88.6	22.1

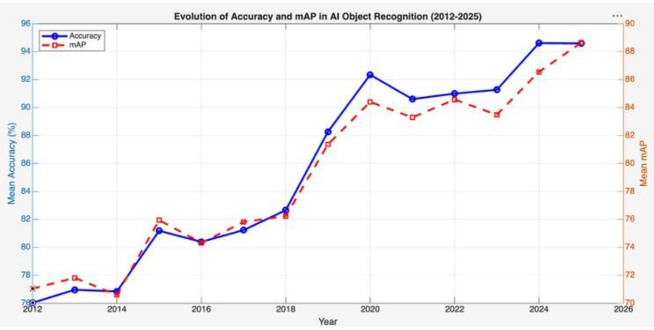

Fig. 2. Evolution of Accuracy and mAP in AI Object Recognition (2012- 2025) [19].

As in Fig. 2, accuracy and mAP increase at a similar rate, and this trend is accelerating after 2018, when the popularity of deeper architectures and attention mechanisms began. The consistent rise demonstrates the significant improvement AI has made in feature representation and contextual understanding.

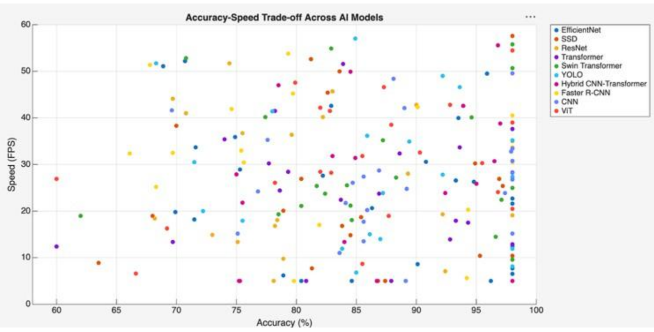

Fig. 3. Accuracy-Speed Trade-off Across AI Models [19].

Fig. 3 emphasises an interesting trade-off: the single-shot detectors such as variants of YOLO produce more frames per second, but they do not provide as high an accuracy as transformer-based models. This visualisation highlights the importance of having either different models or selecting the appropriate model for the application.

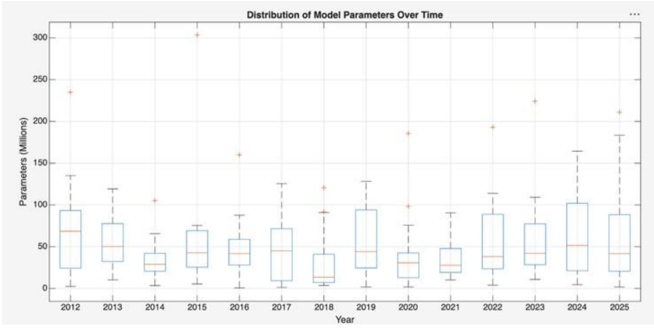

Fig. 4. Distribution of Models Parameters Over Time [19].

While the trend is clear from Fig. 4, more parameter- efficient models emerged post-2019, driven by innovations in lightweight networks and knowledge distillation; there was a large variation in model performance for this study.

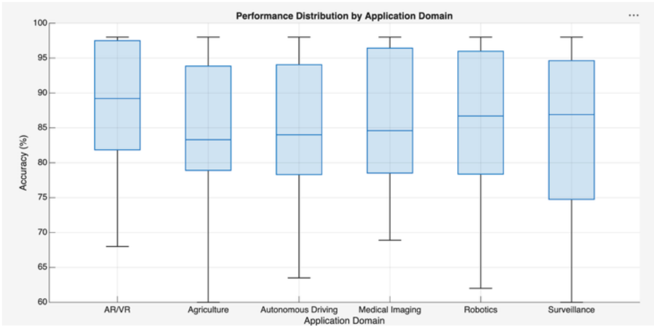

Fig. 5. Performance Distribution by Application Domain [19].

Medical imaging and autonomous driving clearly demonstrate the highest median accuracy, while agriculture and surveillance demonstrate higher variability driven by the more difficult environment and fine-tuning on domain- specific accuracy requirements.

This is a clear indication that AI, particularly deep learning paradigms, has already revolutionised the field of object recognition by their ability to learn and recognise features in a hierarchical manner and reason in context, surpassing traditional methods. Such enhancements are due to architectural novelties that preserve more accurately spatial hierarchies and long-range interactions. However, the figures show some continuing deployment efficiency issues and generalisation issues. The literature on real-time systems is consistent with the speed and accuracy trade-offs observed, and the literature on transformers validates their improved performance on a complex scene [20].

V. Conclusion

The systematic review of 210 studies shows that Artificial Intelligence solutions have revolutionised object recognition, with average accuracy in 2012 being 76%, and in 2025 it went up to 94.6%, and Mean Average Precision (mAP) increased from 71% to 88.6%. This has been achieved by deep learning architectures, such as CNN-based models and transformers, which achieve an impressive 52 FPS in real-time applications and impressively improve generalisation over complex scenes. Although these are some of the improvements made, there remain challenges that persist – transformer models have an average rate of 18.9 FPS, and serious vulnerabilities for domain shift issues exist. The results validate that AI has transitioned object recognition from "hardcoded" static and rigid systems to powerful and contextual systems with applications that are now transforming in numerous fields such as autonomous driving, medical imagery, and smart surveillance. Going forward, efforts should be made to achieve sub-10 ms inference time, high efficiency while achieving high accuracy and high robustness for humans. Overall, AI is transforming the landscape of object recognition and is becoming a cornerstone of the smart systems revolution—with its impact on society only expected to continue increasing in the coming decade.

References

[1] Q. Yuan et al., "Ensemble Artificial Intelligence Framework for Real- Time Automatic Vehicle Detection and Classifications," in IEEE Access, vol. 14, pp. 16034-16050, 2026, doi: 10.1109/ACCESS.2026.3657829.

[2] K. Yu et al., "Design and Performance Evaluation of an AI-Based W- Band Suspicious Object Detection System for Moving Persons in the IoT Paradigm," in IEEE Access, vol. 8, pp. 81378-81393, 2020, doi: 10.1109/ACCESS.2020.2991225.

[3] A. Roy, K. D. Gaikwad, A. J. Hude, J. S. Shinde, J. Parekh and H. R. Nagrani, "Explainable AI (XAI) for Object Detection and Application to Satellite Imagery," in IEEE Access, vol. 13, pp. 185597-185623, 2025, doi: 10.1109/ACCESS.2025.3624198.

[4] S. Yim, M. Cho, D. Kang and S. Lee, "Adaptive Grid Selection Training Strategy for Tiny Object Detection," in IEEE Access, vol. 13,
pp. 171998-172016, 2025, doi: 10.1109/ACCESS.2025.3613741.

[5] R. A. Amin, M. Hasan, V. Wiese and R. Obermaisser, "FPGA-Based Real-Time Object Detection and Classification System Using YOLO for Edge Computing," in IEEE Access, vol. 12, pp. 73268-73278, 2024, doi: 10.1109/ACCESS.2024.3404623.

[6] X. Wang, A. Zeng, K. Cheng, L. Zhu, B. Chen and Q. Xie, "Piercing Eye: Artificial Intelligence Enables Object Detection of Fire Equipment to Protect the Safety of Life and Property," 2025 2nd International Conference on the Frontiers of Electronic, Electrical and Information Engineering (ICFEEIE), Yibin, China, 2025, pp. 33-37, doi: 10.1109/ICFEEIE66944.2025.00015.

[7] A. Song, Z. Zhao, Q. Xiong and J. Guo, "Improving Classroom Object Detection with Bootstrapped Semi-Supervised Learning," 2022 3rd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE), Xi'an, China, 2022, pp. 532- 536, doi: 10.1109/ICBAIE56435.2022.9985865.

[8] G. Baatar, T. Pfützenreuter, D. Karimanzira, D. Shea and J. Albiez, "Automated Collection and Annotation Pipeline for Underwater Object Detection," OCEANS 2021: San Diego – Porto, San Diego, CA, USA, 2021, pp. 1-5, doi: 10.23919/OCEANS44145.2021.9705994.

[9] A. M. Latha and A. S. N. Devi, "Automatic Threat Object Detection from THz Images using Artificial Intelligence Algorithms," 2023 24th International Vacuum Electronics Conference (IVEC), Chengdu, China, 2023, pp. 1-3, doi: 10.1109/IVEC56627.2023.10156906.

[10] M. Zhu and H. Wang, "Fast detection of moving object based on improved frame-difference method," 2017 6th International Conference on Computer Science and Network Technology (ICCSNT), Dalian, China, 2017, pp. 299-303, doi: 10.1109/ICCSNT.2017.8343706.

[11] N. Elnady, A. Adel and W. Badawy, "Enhancing Dental Caries Detection with YOLOv9: A Comprehensive Analysis and Validation of an Automated Object Detection Model," 2024 International Conference on Future Telecommunications and Artificial Intelligence (IC-FTAI), Alexandria, Egypt, 2024, pp. 1-4, doi: 10.1109/IC- FTAI62324.2024.10950058.

[12] W. Qian, "Exploration on Computer Vision Object Detection Algorithms Under Artificial Intelligence," 2023 3rd International Conference on Mobile Networks and Wireless Communications (ICMNWC), Tumkur, India, 2023, pp. 1-5, doi: 10.1109/ICMNWC60182.2023.10435813.

[13] A. S. Bisht, A. K. Jha, A. Sachdeva and N. Sharma, "Role of Artificial Intelligence in Object Detection: A Review," 2022 4th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2022, pp. 1020-1026, doi: 10.1109/ICAC3N56670.2022.10074231.

[14] Q. Liu, Z. Yan, F. Wang and C. Ding, "Research on object detection algorithm for small object of pests based on YOLOv3," 2021 International Conference on Computer Information Science and Artificial Intelligence (CISAI), Kunming, China, 2021, pp. 14-18, doi: 10.1109/CISAI54367.2021.00011.

[15] S. R, G. V S, S. Prasanna S, M. K, B. N and V. A, "Advancing Underwater Object Identification Using Aerial Hyperspectral Imaging," 2024 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, India, 2024, pp. 1-6, doi: 10.1109/ICPECTS62210.2024.10780084.

[16] A. S. Bisht, A. K. Jha, A. Sachdeva and N. Sharma, "A Study of Artificial Intelligence's Role in Object Detection," 2023 5th International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2023, pp. 920-924, doi: 10.1109/ICAC3N60023.2023.10541408.

[17] Y. -T. Tsai and K. -Y. Chen, "The autonomous control of the Service- Oriented robot arms based on the artificial intelligence object identifying technology," 2019 International Conference on Fuzzy Theory and Its Applications (iFUZZY), New Taipei, Taiwan, 2019, pp. 110-113, doi: 10.1109/iFUZZY46984.2019.9066218.

[18] W. Shi, Z. Dong and L. Zhang, "Graphical Platform of Intelligent Algorithm Development for Object Detection of Educational Drone," 2021 China Automation Congress (CAC), Beijing, China, 2021, pp. 6780-6784, doi: 10.1109/CAC53003.2021.9727769.

[19] B. Singh, "AI Object Recognition Dataset," Kaggle, 2026. [Online]. Available: Kaggle Dataset. [Accessed: May 07, 2026].

[20] T. L. Pham, H. Nguyen, H. Nguyen, V. H. Nguyen, V. Bui and Y. M. Jang, "Object Detection Framework for High Mobility Vehicles Tracking in Night-Time," 2020 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Fukuoka, Japan, 2020, pp. 133-135, doi: 10.1109/ICAIIC48513.2020.9065060.