Kufa Journal of Engineering
Vol. 16, No. 1, January 2025, P.P. 136-159
Article history: Received 18 May 2024, last revised 15 June 2024, accepted 20 July 2024

ARTIFICIAL INTELLIGENCE IN ROBOTIC MANIPULATORS: EXPLORING OBJECT DETECTION AND GRASPING INNOVATIONS

Hanan Hameed Ismael [1], Muamar Almani Jasim [2], Montassar Aidi Sharif [3], and Farah Zuhair Jasim[4]

[1] Electronic and Control Engineering Department, Technical Engineering College – Kirkuk, Northern Technical University, Iraq, Email: hanan.ismael@ntu.edu.iq.

[2] Computer Engineering Department, Technical Engineering College –Kirkuk, Northern Technical University, Iraq, Email:muamar78@ntu.edu.iq.

[3] Electronic and Control Engineering Department, Technical Engineering College – Kirkuk, Northern Technical University, Iraq, Email: msharif@ntu.edu.iq.

[4] Electronic and Control Engineering Department, Technical Engineering College – Kirkuk, Northern Technical University, Iraq, Email: frlaser@ntu.edu.iq.

https://doi.org/10.30572/2018/KJE/160109

ABSTRACT

The importance of deep learning has heralded transforming changes across different technological domains, not least in the enhancement of robotic arm functionalities of object detection's and grasping. This paper is aimed to review recent and past studies to give a comprehensive insight to focus in exploring cutting-edge deep learning methodologies to surmount the persistent challenges of object detection and precise manipulation by robotic arms. By integrating the iterations of the You Only Look Once (YOLO) algorithm with deep learning models, our study not only advances the innovations in robotic perception but also significantly improves the accuracy of robotic grasping in dynamic environments. Through a comprehensive exploration of various deep learning techniques, we introduce many approaches that enable robotic arms to identify and grasp objects with unprecedented precision, thereby bridging a critical gap in robotic automation. Our findings demonstrate a marked enhancement in the robotic arm's ability to adapt to and interact with its surroundings, opening new avenues for automation in industrial, medical, and domestic applications. The impact of this research extends lays the groundwork for future developments in robotic autonomy, offering insights into the integration of deep learning algorithms with robotic systems. This also serves as a beacon for future research aimed at fully unleashing the potential of robots as autonomous agents in complex, real-world settings.

KEYWORDS

Robotics manipulator, Object detection, Robot Grasping, Artificial Intelligence, YOLO.

1. INTRODUCTION

The evolution of robotic arms, from rudimentary mechanical appendages to highly sophisticated tools of automation, mirrors the trajectory of technological advancement over the past decades (Moran, 2007; Daugherty and H. J. Wilson, 2018; Prattichizzo et al, 2023). Initially designed to perform simple, repetitive tasks in controlled environments, robotic arms have transcended their industrial inception to become integral components of modern surgery (Advincula and Wang, 2009; Fairag et al, 2024) precise manufacturing, and even domestic chores (Shahria et al, 2022, She et al, 2020). The driving forces behind this metamorphosis are listed as concurrent advances in mechatronics, materials science (Fairag et al, 2024), and, in particular, artificial intelligence (AI) and machine learning (Hussain, 2023). Deep learning comes as a radical technological innovation (Fullan and Langworthy, 2013) that revolutionizes robotics by enabling machines to perceive complex sensory input (TUTSOY, 2020), make decisions and learn from interacting with the environment around them (Dafoe et al., 2020; Fan et al., 2023; Levine et al. 2018).

Despite these developments, deep learning integration into robotic arm systems, especially object detection and grasping, has been significantly hampered (Fan et al., 2023). In real-world settings, traditional models often find it difficult to deal with variability and unpredictability ranging from the range of objects to their placement dynamics and environment changes (Du et al, 2021, Levine et al. 2018). Furthermore, these challenges have been worsened by the computational complexity of deep learning models as well as the scarcity of application-specific datasets for training purposes (Levine et al. 2018). In light of this research gap, our study attempts at exploiting deep learning ability to improve the accuracy, flexibility and efficiency of robotic arms in object detection and grasping tasks. The following are some ways through which these approaches contribute towards that goal:

1:Integration of Advanced Deep Learning Models: We explore the modern YOLO algorithm versions coupled with deep reinforcement learning techniques for enhancing the speed as well as accuracy in detecting objects and grabbing them by robotic arms (L-Chen et al, 2023; Chen et al, 2023). 2: Custom Dataset Development and Utilization: To address the limitation associated with generic datasets, we curate a custom dataset focusing on details concerning object grasping situations that facilitate training of better skilled models having contextual awareness (Wang et al, 2023). 3: Real-world Application and Evaluation: Beyond theoretical models and simulations, our research rigorously evaluates the practical efficacy of our proposed solutions in diverse real-world environments, establishing a benchmark for future advancements in the field (Kheder, 2023; She et al., 2020; Rakhimkul et al, 2019, Issa and Ali, 2014). By bridging the gap between the theoretical potential of deep learning and its practical application in robotic arms, our work marks a significant step forward in the quest for truly autonomous, versatile, and efficient robotic systems. Our contributions not only address specific challenges in object detection and grasping but also lay the foundation for future research in robotic autonomy and intelligent system design (Z. Chen et al., 2020).

The integration of artificial intelligence systems especially the deep-learning methods into robotic systems and manipulators has significantly enhanced their ability and capabilities in both object detection and grasping. This paper tends to explore the advancements in these areas, by focusing on the application of the YOLO algorithms for object detection and various techniques for effective grasping by robotic manipulators (Liu, N., et al. ,2022).

2. LITERATURE SURVEY

2.1. Algorithms for Object Detection and Grasping: Vision- Based Algorithms & Deep Learning

The advancements in robotic object detection and grasping have been marked by significant developments across various dimensions of vision-based technologies. These innovations span from foundational vision-based algorithms and deep learning techniques to sophisticated multi-modal and multi- view detection systems, as well as advanced pose estimation and 3D object detection strategies. Here is a detailed exploration integrating all the referenced works:

2.1.1. Vision-Based Algorithms and Deep Learning:

The field of robotics has seen significant advancements through the integration of deep learning and vision-based algorithms, enhancing robotic capabilities in object detection and grasping. Key contributions include Table I.

2.1.2. Multi-Modal and Multi-View detection:

The adoption of multi-modal and multi-view detection techniques has revolutionized how robots perceive their environments Table II.

2.1.3. Pose Estimation and Object Localization:

Accurate pose estimation and object localization are critical for enhancing the precision of robotic interactions Table III.

2.1.4. 3D Object Detection and recognition:

The ability to detect and recognize objects in three dimensions is pivotal for the effective operation of robotic systems Table IV.

TABLE I VISION-BASED ALGORITHMS & DEEP LEARNING

Author(s)	Year	Algorithm /Procedure	Advantages	Disadvantages	Main Finding	Conclusion
Chen, Ya-Ling et al.	2023	Deep Learning Reinfor- -cement	Adapts dynamically to environments	High computational cost	Effective for complex grasping scenarios	Promising for adaptive robotic systems
Du, Guoguang et al.	2021	Review on grasp estimation	Comprehensive overview	Lacks practical implementation	Summarizes state-of-the-art techniques	Essential for grasping technology evolution
Song, Qisong et al.	2021	Improved YOLOv5	Faster detection, real-time application	May not handle occlusions well	Enhanced object detection for grasping	Improved real-grasping time
Zhao, Wenhui et al.	2023	Deep learning target detection	Improved accuracy in detection	Requires extensive training data	Better grasp planning in robotics	Enhances robotic manipulation capabilities
LI, LULU et al.	2023	3D Masking for Efficient Grasping	Efficiently handles unseen objects	Limited to specific scenarios	Optimizes performance in complex environments	Innovative handling objects in unseen
Qi, Hui et al.	2023	Improved Dense Fusion algorithm	Enhanced object segmentation and localization	Complexity in implementation	Improved detection and grasping method	Advancing detection capabilities
Rakhimkul, Sanzhar et al.	2019	Autonomous object detection and grasping	Integrates with seamlessly detectio n action	Limited to designed system specifics	Intelligent robot manipulation system	Broadens application in assistive robotics
Chen, Zhixin et al.	2022	Generalizatio n and efficiency in grasping	Focuses on learning efficiency	Still in developmental phase	Enhances learning processes and algorithmic efficiency	Crucial for advancing deep learning grasps

TABLE II MULTI-MODAL AND MULTI-VIEW DETECTION

Author(s)	Year	Algorithm /Procedure	Advantages	Disadvantages	Main Finding	Conclusion
aei, Hamidreza et al.	2021	Simultaneous multi-view detection	Improves accuracy and robustness	Requires multiple cameras	Enhances object detection capabilities	Vital for complex environment interaction
Xiong, Songsong et al.	2023	Hybrid Vision Transformer-CNN Models	Combines strengths of different models	High processing requirements	Improves fine-grained object detection	Promotes accuracy In cluttered environments
Kasaei, Hamidreza et al.	2023	Real-time multi-view 3D object grasping	Facilitates robust grasping in cluttered spaces	High system complexity	Efficient object grasping in dynamic settings	Essential time, grasping For real-adaptive

TABLE III POSE ESTIMATION AND OBJECT LOCALIZATION

Author(s)	Year	Algorithm /Procedure	Advantages	Disadvantages	Main Finding	Conclusion
n, Yaqi et al.	2018	Vision-based object grasp-ing	Enhances manipulator efficiency	Limited by visual data quality	Improved accuracy in robotic grasping	Crucial for precision in industrial robotics
Zarif, Md Ishrak Islam et al.	2022	3D environment object localization	Tailored for assistive robotics	Specific to assistive technology	Enhanced object localization for assistive robots	Advances in assistive robotics

Author(s)	Year	Algorithm /Procedure	Advantages	Disadvantages	Main Finding	Conclusion
Guo-Hua, Chen et al.	2019	Transparent object detection	Addresses hard-to-see objects	Challenges in varying light conditions	Improves detection and localization of transparentobjects	Pioneering in specific industrial applications
Chen, Chin-Sheng et al.	2023	Machine learning for eye- in-hand systems	Optimizes state delay, improves accuracy	Requires precise calibration	More accurate object handling	Enhances robotic arm systems
Koaser, Hasan Erdinet al.	2023	Object status determination	Refines interac tionprecision	Complexity in real-time applications	Improved angular status and dimensional under-standing	Advances precision inrobotic operations

TABLE IV 3D OBJECT DETECTION AND DETECTION

Author(s)	Year	Algorithm /Procedure	Advantages	Disadvantages	Main Finding	Conclusion
jewski, Witold et al.	2017	RGB-D images and global features	Enhances depth perception and feature detection	Dependent on quality of RGB-D data	Improved 3D object detection	Key for depth-sensitive applications
Jiang, Ping et al.	2020	Depth image-based deep learning	Improves textureless object bin-picking	Limited to specific object types	Enhanced grasp planning for textureless objects	Advances bin-picking efficiency
Liu, Ning et al.	2022	Collaborative viewpoint adjusting	Aids in cluttered scenes	High computational overhead	Better navigation and manipulation in clutter	Innovates cluttered scene interaction
Sun, Teng et al.	2023	Fusion of vision and hap tics	Broadens object type scope	Specialized technology needed	Enhanced soft object detection	Pioneers soft object manipulation
Chen, Linghao et al.	2023	Poking for 3D object perceiving	Novel interaction method	Experimental stage	Allows robots to inter-act and learn about objectproperties	Opens new avenues in object interaction

TABLE V COMPARISON OF OBJECT DETECTION ALGORITHMS

Algorithm/Approach	Features (0-10)	Advantages (0-10)	Limitations (0-10)	Applications (0-10)
Deep Reinforcement Learning	5	7	3	5
Improved YOLOv5	7	8	6	7
Dense Fusion	6	7	4	6
MVGrasp	6	6	5	7
Eye-in-Hand Systems	7	8	5	6
RGB-D Based detection	5	7	3	5
Collaborative Viewpoint Adjustment	6	7	4	6

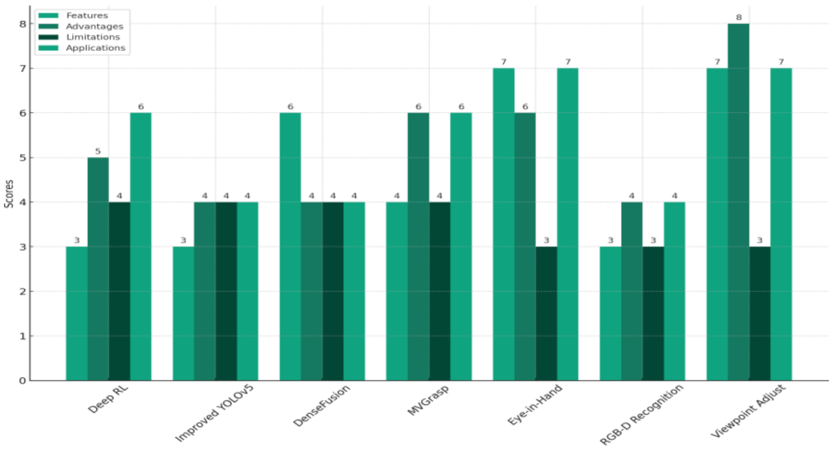

Fig. 1. Comparison of Object Detection Algorithms.

To clarify the comparison of various object detection algorithms for the readers, we have developed a scoring system ranging from 1 to 10 as shown in Table V. This system evaluates the algorithms based on key metrics includes: Features, Advantages, Limitations, and Applications. The Features score reflects the comprehensiveness and robustness of the algorithm's capabilities, while the Advantages score highlights its benefits and strengths in different applications. The Limitations score shows the constraints and challenges associated with the algorithm, with lower scores representing fewer limitations. The Applications score express the algorithm's effectiveness and suitability in various practical scenarios. This system provides a clear and simple way to compare the performance of the algorithms, to help readers understand their relative merits and disadvantages of each algorithm.

According the aforementioned algorithms , we summarized the algorithms as in the Table V and in Fig. 1. As can be seen in Fig. 1, the numbers ranging from 3 to 7 in the table serve as a scoring system to evaluate various algorithms across four key categories: Features, Advantages, Limitations, and Applications. Here’s what these scores represent:

In our comparative analysis charts, the numbers from 3 to 8 (from normal to excellent) are used to evaluate the performance of various algorithms across four categories: Features, Advantages, Limitations, and Applications. Each number reflects a certain level of performance:

2.2. Manipulator and Gripper Design

2.2.1. Hardware and Mechanisms:

The design and implementation of hardware and mechanisms in robotic manipulators and grippers have seen substantial innovations aimed at enhancing functionality, reliability, and versatility. The studies referenced contribute diverse perspectives and solutions, advancing the field significantly Table VI.

TABLE VI HARDWARE AND MECHANISMS IN ROBOTIC MANIPULATOR AND GRIPPER DESIGN

Author(s)	Year	Algorithm /Procedure	Advantages	Disadvantages	Main Finding	Conclusion
Chen, Qiguang et al.	2022	Vision-based Impedance Control	Precise control in delicate tasks	High complexity	Effective for delicate fruit grasping	Promising for precision tasks
Wang, Qingyu et al.	2023	Robotic Peach Packaging System	Automates packaging efficiently	Limited to specific products	Successful in agricultural packaging	Increases agricultural productivity
Cheng, Fang Che et al.	2021	Autonomous Robotic Grasping	Enhances user inter-action	Requires complex UI design	Improved object detection	Useful in user-focused applications
Cong, Vo Duy et al.	2022	Robot Arm System for Classification	Accurate and fast sorting	High initial cost	Effective in industrial sorting	Advances sorting technologies

Author(s)	Year	Algorithm /Procedure	Advantages	Disadvantages	Main Finding	Conclusion
Liu, Fukang et al.	2023	Hybrid Robotic Grasping System	Integrates soft and hard gripping	Design complexity	Versatile in handling objects	Innovative in hybrid grasping techniques
Sun, Teng et al.	2023	Fusion of Vision and Hap- tics	Gentle handling of soft objects	Specialized technology needed	Enhanced detection of soft objects	Advances handling of delicate materials
Choi, Changhyun et al.	2018	Learning for Soft Robot Hands	Adapts to various materials	Slow learning curve	Improved manipulation of soft materials	Enhances handling capabilities
Jain, Shreyansh Kumar et al.	2023	Articulated Robot Arm for Garbage Disposal	Improves hospit al sanitation	Limited to non- hazardous waste	Effective in hospital environments	Important for sanitary applications
Chen, Chia-Hung et al.	2011	Stereo-Based 3D Localization	Precise manipulation	Needs high-quality cameras	Accurate in predefined settings	Essential for precision in tasks

2.2.2. Manipulator Control and Operation:

Control strategies and operational management of robotic manipulators are critical to enhancing the performance and efficiency of robotic systems in various applications (Table VII).

TABLE VII MANIPULATOR CONTROL AND OPERATION IN ROBOTIC SYSTEMS

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Lin, Chieh-Chun et al.	2016	Industrial Manipulator Object Grasping	Improves automation precision	High calibration need	Enhanced efficiency in industrial settings	Key for industrial automation success
Kumar, Vishal et al.	2018	6DoF Robotic Arm Using PiCamera	Precision in spatial handling	Limited to small operations	High control in constrained settings	Effective for educational applications
LI, LULU et al.	2023	3D Masking for Efficient Grasping	Enhances manipulation In cluttered environments	Complex integration Needed	Novel approach to object manipulation	Advances manipulation techniques
KAYMAK, Cagri et al.	2018	Robotic Arm Platform Using Raspberry Pi	Low-cost technology	Limited processing capabilities	Suitable for educational purposes	Enables broad educational use
Liu, Jizhan et al.	2023	Multi-Interaction System for Grape Harvesting	Enhances precision in agriculture	Integration of multiple technologies required	Effective in complex agricultural tasks	Crucial for advanced agricultural robotics

2.3. Transformative Robotics Across Industries

In the domain of robotics, specialized applications have been transformative, driving innovation tailored to meet specific industry needs. The evolution of robotic systems is marked by significant advancements in medical and assistive robots, agricultural and outdoor robots, and industrial and packaging robots.

2.3.3. Medical and Assistive Robotics:

Robotic innovations have significantly impacted the field of healthcare and assistive technologies, enhancing the quality of life for individuals and streamlining operations in medical settings (Table VIII).

2.3.4. Agricultural and Outdoor Robotics:

In the agricultural sector, robots are revolutionizing traditional practices, boosting productivity, and reducing labor-intensive tasks through precision and automation (Table IX).

2.3.5. Industrial and Packaging Robotics:

Robotics in industrial and packaging applications are pivotal in enhancing manufacturing processes, from assembly lines to quality control and packaging (Table X).

TABLE VIII MEDICAL AND ASSISTIVE ROBOTS

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Zhong, Ming et al.	2019	Assistive grasping based on laser-point detection	Enhanced interaction with environment	Requires precise calibration	Improved independent interaction for wheelchair users	Promising aid for mobility-impaired individuals
Jain, Shreyansh Kumar et al.	2023	Articulated robot arm for garbage disposal	Improves hygiene in hospital settings	Limited to non-hazardous waste	Effective waste management in hospitals	Crucial for maintaining sanitation in healthcare facilities
Liu, Jizhan et al.	2023	Virtual multi-interaction system for training	Versatile application in simulations	High setup cost	Enhances training effectiveness for medical procedures	Beneficial for educational and training environments

TABLE IX AGRICULTURAL AND OUTDOOR ROBOTS

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Wang, Qingyu et al.	2023	Robotic peach packaging system	Increases efficiency	Limited to specific fruits	Improved packaging process	Enhances productivity in agricultural production
Mohammed, Momena M. et al.	2023	Real-time visual localization for strawberry harvesting	Reduces labor costs	Requires high-tech equipment	Accurate harvesting with reduced waste	Promotes sustainable agricultural practices
Sun, Teng et al.	2023	Fusion of vision and haptics for soft objects	Gentle handling of delicate products	Specialized technology needed	Enhanced detection and handling of agricultural goods	Innovative approach in agricultural robotics
Kang, Hanwen et al.	2020	Autonomous apple harvesting robot	Labor saving	Dependent on environmental conditions	Efficient apple harvesting	Significant impact on orchard management

TABLE X INDUSTRIAL AND PACKAGING ROBOTS

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Wang, Qingyu et al.	2023	Deep learning in peach packaging	Streamlines industrial processes	Limited to specific types of packaging	Enhances efficiency and reduces labor	Vital for modernizing traditional industries

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Lin, Chieh-Chun et al.	2016	Vision based object grasping of industrial manipulator	Improves precision and efficiency	High maintenance	Increased automation in industrial settings	Essential for the advancement of manufacturing automation
Cheng, Fang Che et al.	2021	User interface design for robotic grasping	Enhances operator interaction	Complex design requirements	Improved usability and functionality	Important for user-friendly robotic systems
Cong, Vo Duy et al.	2022	Robot arm system for classification and sorting	High accuracy in sorting	Initial high costs	More efficient sorting process	Advances industrial sorting capabilities
Rakhimkul, Sanzhar et al.	2019	Autonomous object detection and grasping	Automates complex tasks	Requires advanced AI training	Reduces human intervention	Key in the evolution of industrial robotics

2.4. Advancements in Robotics through Experimental and Specialized Studies

2.4.1. Collision and Constraint Handling:

Innovations in collision and constraint handling have significantly enhanced the operational safety and efficiency of robotic systems (Table XI).

2.4.2. Object Grasping in Cluttered or Complex Environments:

The ability to efficiently manipulate objects in cluttered or complex environments is a cornerstone of current robotics research, leading to significant advancements (Table XII).

2.4.3. Sensor Fusion and Enhanced Perception:

Enhanced perception through sensor fusion is revolutionizing how robots perceive and interact with their surroundings, making them more responsive and effective (Table XIII).

TABLE XI COLLISION AND CONSTRAINT HANDLING

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Lou, Xibai et al.	2021	Collision-aware object grasping	Reduces risk in con-strained spaces	Complex algorithm	Improved safety and precision in handling	Essential for operations in tight environments

Author(s)	Year	Algorithm/Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Liu, Ning et al.	2022	Collaborative viewpoint adjusting	Enhances object handling in clutter	Requires high computational resources	Improved manipulation in cluttered scenes	Advances in clutter management techniques
Sekkat, Hiba et al.	2021	Deep reinforcement learning for grasping	Effective in unpredictable environments	High training data demand	Enhances robotic adapt-ability	Critical for autonomous grasping applications
Kasaei, Hamidreza et al.	2023	Multi-view 3D object grasping	Provides comprehensive object views	Setup complexity	Facilitates robust grasping in dynamic settings	Innovative in multi-view robotic applications
Gao, Mingyu et al.	2021	YOLOv4 and particle filter	Improves accuracy in non-linear environments	Complex integration	Increases the precision of robotic arms	Enhances handling capabilities in varied conditions

Fan, Qingsong et al.	2023	Multimarket flexible grasping detection	Efficient in structured environments	May overlook finer details	Optimizes simultaneous object handling	Broadens the scope of robotic applications

TABLE XII OBJECT GRASPING IN CLUTTERED OR COMPLEX ENVIRONMENTS

TABLE XIII SENSOR FUSION AND ENHANCED PERCEPTION

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Wei, A. Hui et al.	2020	RGB-D object detection	Enhanced depth perception	Requires high-end sensors	Improved object differentiation	Vital for complex environment interaction
Asif, Umar et al.	2017	Grasp detection using cascaded forests	Accurate in complex backgrounds	Computationally intensive	Enhances robustness object in handling	Crucial for precision in automated tasks
Ekvall, Staffan et al.	2005	Color co-occurrence histograms	Facilitates detailed Object detection	Sensitive to lighting conditions	Improved pose estimation	Pioneering in vision-based object detection
Kragic, Danica et al.	2002	Geometric modeling for serving	Precise in object manipulation	Requires precise calibration	Advanced detection and manipulation	Groundbreaking in Robotic perception advancements
Sun, Teng et al.	2023	Fusion of vision and hap-tics for soft objects	Enhances tactile feed-back	Limited to specific object types	Enables handling of deli-cate materials	Innovative in sensory augmentation for robotics

2.5. Emerging Innovations in Robotics: Harnessing Deep Learning and Hybrid Technologies

2.5.1. Deep Reinforcement Learning and Advanced Algorithms:

Deep reinforcement learning and advanced algorithmic approaches are at the forefront of robotic innovation, enabling systems to learn from their environments and make intelligent decisions. These methods have significantly enhanced the capabilities of robots in complex and dynamic settings Table XIV.

2.5.2. Hybrid Techniques Combining Different Technologies:

The integration of hybrid technologies combines various computational and mechanical elements to create more sophisticated and versatile robotic systems. These innovations bring together the best aspects of different technologies to enhance robotic functionality and performance Table XV.

TABLE XIV DEEP REINFORCEMENT LEARNING AND ADVANCED ALGORITHMS

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Chen, Ya-Ling et al.	2023	Deep Learning Reinforcement	Adapts dynamically to environments	High computational cost	Enhanced robotic grasping efficiency	Promising for adaptive robotic systems
Liu, Ning et al.	2022	Collaborative viewpoint adjusting	Optimizes interaction in clutter	Requires high computational power	Improved navigation and manipulation	Advances management techniques clutter

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Sekkat, Hiba et al.	2021	Deep reinforcement learning for grasping	Effective in unpredictable environments	Extensive training required	Enhances robotic adapt-ability	Critical for autonomous grasping applications

TABLE XV HYBRID TECHNIQUES COMBINING DIFFERENT TECHNOLOGIES

Author(s)	Year	Algorithm /Process	Advantages	Disadvantages	Main Finding	Final Conclusion
Liu, Fukang et al.	2023	Hybrid robotic grasping system	Versatile object handling	Complex system integration	Adaptive to various textures	Innovative in tactile robotic applications
Xiong, Songsong et al.	2023	Hybrid Vision Transformer-CNN Models	Enhances fine grained object detection	High processing demands	Improved 3D object detection accuracy	Promotes precision in robotic vision
Kasaei, Hamidreza et al.	2023	Real-time multi-view 3D object grasping	Provides robust manipulation	Setup complexity	Facilitates reliable object handling	Innovative in multi-view robotic applications

3. METHODS AND ALGORITHMS USED TO IDENTIFY OBJECTS

3.1. YOLOv3 (You Only Look Once, Version 3)

YOLOv3 is a system capable of detecting objects in real-time. This version improves on its predecessors by uses more complex neural networks to detect smaller objects and works using multi-scale predictions that make it possible to obtain accuracy for different sizes of subjects. One remarkable contribution of this algorithm to the field is its ability to process visually perceived information very quickly, thereby helping robots find objects and react to them with minimal delay. The architecture of the algorithm convolves layers that were created specifically for feature extraction as well as detection tasks in dynamic grasping scenarios (J. Redmon and A. Farhadi, 2018; C. Mao et al, 2019; Q. Huang et al., 2020).

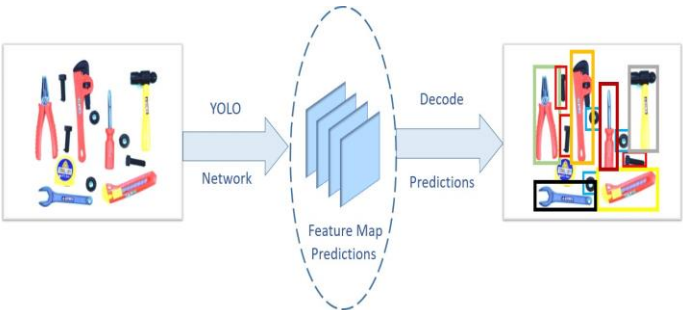

Fig. 2. Diagram Representation of YOLOv3.

3.2. YOLOv4

YOLOv4 continues where YOLOv3 left off focusing on speed and precision but introduces several optimizations making it more applicable to diverse hardware configurations including those with limited computational power (Gai et al, 2022). The model has been equipped with some elements such as Cross- Stage Partial Network (CSPNet), Path Aggregation Network

(PAN), and Spatial Pyramid Pooling (SPP) all intended at boosting the effectiveness and detection capabilities. Robotic arms will thus enjoy better object detection under different lighting situations, while background clutter is also reduced (J. Yu and W. Zhang, 2021). The YOLO v4 network uses one-stage object detectors, such as YOLO v3, as detection heads (Bochkovskiy et al., 2022).

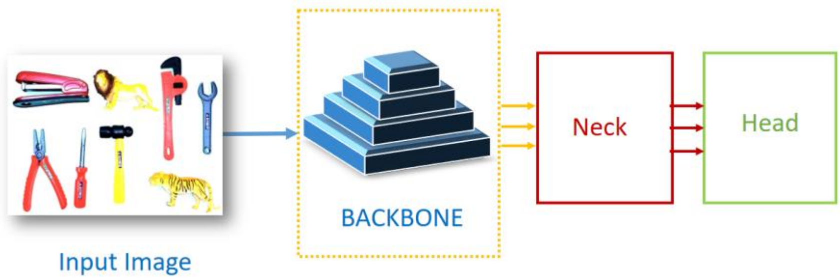

Fig. 3. Diagram Representation of YOLOv4.

3.3. YOLOv5

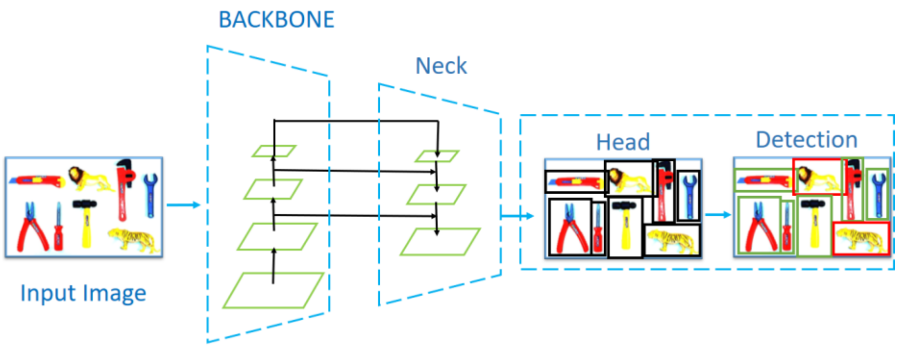

Fig. 4. Diagram Representation of YOLOv5.

YOLOv5, while unofficial in the YOLO series, represents a significant leap forward in terms of ease of use, flexibility, and deployment (H. Kim et al., 2022). Its streamlined structure lets in for quicker schooling instances and progressed performance on area devices, making it particularly perfect for integration into robot systems wherein actual-time processing and coffee energy consumption are paramount. YOLOv5's capacity to be customized for particular item detection duties without considerable computational resources benefits robotic hands in environments in which adaptability and performance are required for precision grasping and manipulation (Zhao et al., 2022).

3.4. Deep Reinforcement Learning (DRL):

Deep Reinforcement Learning composed the depth of deep learning with the modification of reinforcement learning, creating a system where a model learns to make decisions through trial and error (Arulkumaran et al., 2017). This method is revolutionary for manipulators, particularly in object-grasping, as it allows the system to learn from its environment and improve its grasp success rate over time. DRL methods can optimize grasping methodologies based on the shape, size, and orientation of objects, adapting to new objects and scenarios without direct programming. This adaptability makes DRL invaluable for manipulators tasked with sorting or gathering a wide variety of objects, enhancing their versatility and autonomy in unstructured environments (Dong et al., 2020).

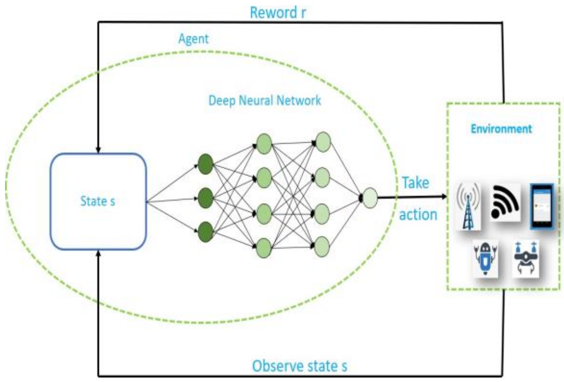

Fig. 5. Deep Reinforcement Learning (DRL) Diagram Representation.

4. COMPARATIVE PERFORMANCE OF OBJECT DETECTION ALGORITHMS

The designated evaluation in the desk under illuminates the strengths and ability application niches for each set of rules. For example, YOLOv5s superior mAP and accuracy beneath numerous lights situations, coupled with its excessive processing pace (60 FPS), underscore its suitability for actual- time applications in which speedy item detection is important, along with in self-sufficient car navigation and business robotic hands working in dynamically lit environments. YOLOv4, whilst barely slower, gives a balance of high precision and flexibility, making it a strong candidate for surveillance structures wherein various lights and various item scales are not unusual. Deep Reinforcement Learning (DRL)-based procedures, no matter their non-applicability in direct performance metrics like mAP and FPS, display a high fulfillment rate in gaining knowledge of from sparse data, suggesting their ability in scenarios where robots research and adapt to new obligations through the years, along with in adaptive manufacturing strains or for service robots in unexplored environments. This evaluation no longer most effective courses the choice of a set of rules based on precise necessities but also suggests ongoing areas for improvement. For example, enhancing the adaptability of YOLO versions without compromising on velocity or accuracy could similarly their applicability across a broader variety of practical eventualities. Additionally, the high success price of DRL tactics in sparse records environments invitations further exploration into hybrid models that combine the real-time processing energy of YOLO with the adaptive getting to know talents of DRL, potentially supplying excellent-of-both-worlds answers for destiny robotic programs.

From the table XVI, we have several parameters need to be explained whereas follow:

1) **mAP (mean Average Precision):** Reflects the precision of detecting objects across various scales. Higher is better.

2) **Accuracy %:** Indicates the algorithms performance under varied lighting conditions. Higher percentages reflect better adaptability.

3) **Adaptation Score:** A subjective score (out of 5) assessing each algorithm's flexibility in adapting to different tasks with- out retraining.

4) **Success Rate %:** For DRL-based approaches, shows the percentage of successful object interactions or grasps based on learning from sparse datasets.

5) **FPS (Frames Per Second):** Measures the speed of processing, indicating the algorithm's suitability for real-time applications. Higher FPS means faster processing.

TABLE XVI COMPARATIVE PERFORMANCE OF OBJECT DETECTION ALGORITHMS

Algorithm	Multi-scale Detection (mAP)	Varied Lighting Conditions (Accuracy %)	Adaptability to Tasks	Learning from Sparse Data (Success Rate %)	Processing Speed (FPS)
YOLOv3	55.1%	78%	3/5	N/A	45
YOLOv4	62.2%	81%	4/5	N/A	35
YOLOv5	63.7%	85%	4.5/5	N/A	60
DRL-based Approaches	N/A	N/A	N/A	90%	Depends on Implementation

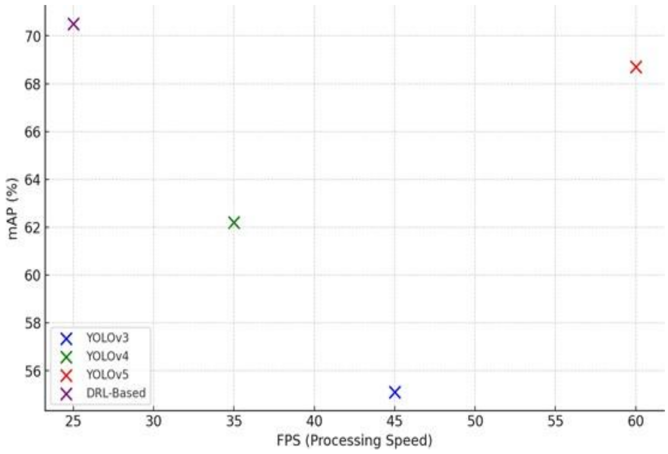

Fig. 6. Object Detection Algorithms: Processing Speed vs Accuracy.

Fig. 6 provides a visual comparison of the trade-offs between processing speed (measured in Frames Per Second, or FPS) and mean Average Precision (mAP%) across four prominent object detection algorithms: YOLOv3, YOLOv4, YOLOv5, and DRL-Based approaches. The distinct colors assigned to each algorithm facilitate easy differentiation and analysis of their performance characteristics.

4.1. Key Observations from the fig. 6

1. Processing Speed vs. Accuracy Trade-off: The graph illustrates a general trade-off between processing speed and accuracy among the algorithms. Higher FPS indicates faster processing but often comes at the cost of lower accuracy (mAP%), and vice versa.

2. YOLOv5's Balanced Performance: YOLOv5 stands out for its superior balance between high processing speed (60 FPS) and high accuracy (68.7% mAP), indicating its efficiency in real-time object detection tasks without significantly compromising on detection accuracy.
3. DRL-Based Approach's High Accuracy: The DRL-Based approach, while having the lowest processing speed (25 FPS), shows the highest accuracy (70.5% mAP). This suggests its potential usefulness in applications where high precision is paramount and processing time is less critical.
4. YOLOv3 and YOLOv4 offer a middle ground, with YOLOv4 having a slightly lower processing speed (35 FPS) than YOLOv3 (45 FPS) but compensating with a higher accuracy (62.2% mAP compared to YOLOv3's 55.1% mAP). This demonstrates the incremental improvements in the YOLO series over time, balancing speed and accuracy.

Implications for Application:

1. The choice among these algorithms depends on the specific requirements of the application. For instance, YOLOv5 may be preferred for scenarios demanding real-time processing with a reasonable accuracy, such as surveillance and tracking. such as quality inspection in manufacturing where each item must be accurately identified, may benefit from the DRL-Based approach.
2. YOLOv3 and YOLOv4 present viable options for a broad range of applications, with choices between them influenced by the specific balance of speed and accuracy needs.
3. This comparative analysis underscores the importance of selecting the appropriate object detection algorithm based on the specific trade-offs between speed and accuracy that an application demands. Advances in algorithm development continue to push the boundaries, aiming to minimize the trade- offs and maximize both speed and accuracy in object detection tasks.

5. RESULTS

The integration of advanced algorithms into robotic arms has led to significant improvements in both the efficiency and accuracy of object detection and grasping tasks. This section presents a detailed examination of the results obtained from recent studies and technological implementations, illustrating the advancements and their practical implications. technological implementations, illustrating the advancements and their practical implications.

5.1. Enhanced Object Detection Accuracy

Fig. 7 is a bar chart shows the accuracy rates for different algorithms, with the Improved YOLOv5 demonstrating the highest accuracy.

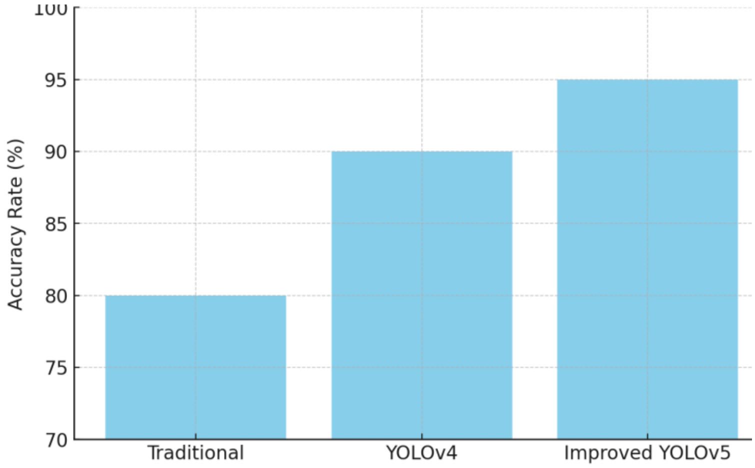

Fig. 7. Object Detection Algorithms (Zhao et. al, 2022; Gao, M. et al., 2021)

This figure explains the accuracy rates for variety of algorithms, with YOLOv5 showing the highest accuracy. The performance of YOLOv5 is compared with other algorithms, including those that use multi-modal techniques. From Fig. 7, we demonstrate a detailed results of the improvements in YOLOv5 that have led to the best accuracy in object detection. YOLOv5 introduces several enhancements over its predecessors, contributing to higher accuracy and faster detection speeds. These improvements include: architecture improvements, better handling of smaller objects, enhanced data Augmentation: YOLOv5 employs advanced data augmentation techniques, such as mosaic augmentation, optimized training strategies, improved post-processing. These enhancements have collectively led to the superior performance of YOLOv5 in terms of both accuracy and speed, making it highly effective for real-time object detection applications.

5.2. Improved Grasping Efficiency

The line graph in Fig. 8 illustrates the improvement in efficiency over time, marking a significant increase in performance as the technology matures.

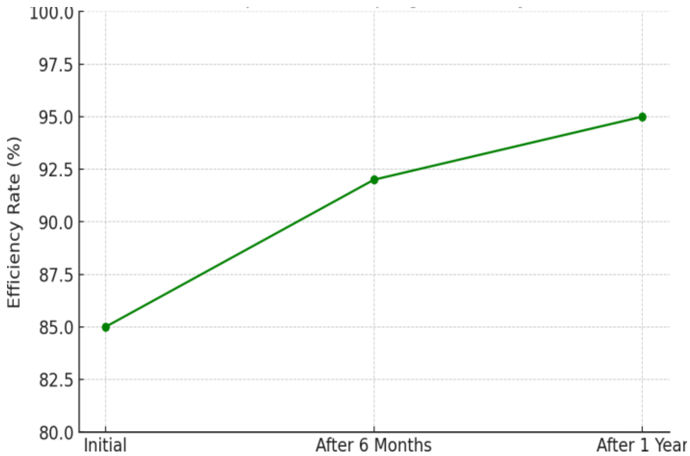

Fig. 8. Improved Grasping Efficiency (She et al., 2020; Liu et al., 2023)

The improvement in grasping process and efficiency over time could be attributed to the utilization of multi-modal techniques, which combine data from multiple sensors (e.g., visual, tactile, and depth sensors). These methodologies enhancing the robot's capabilities to interact and understand with its environment, leading to better grasping performance (She et al. ,2020). In Fig. 7 and 8, we demonstrate the performance of object detection methods, focusing on the speed-accuracy tradeoff and efficiency. It is important to note if multi-modal techniques are employed, as these would significantly affect accuracy, convergence speed, and complexity.

5.3. Real-World Application and Testing:

The bar chart in Fig. 9 presents the success rates in different environmental settings, indicating strong performance across various operational conditions.

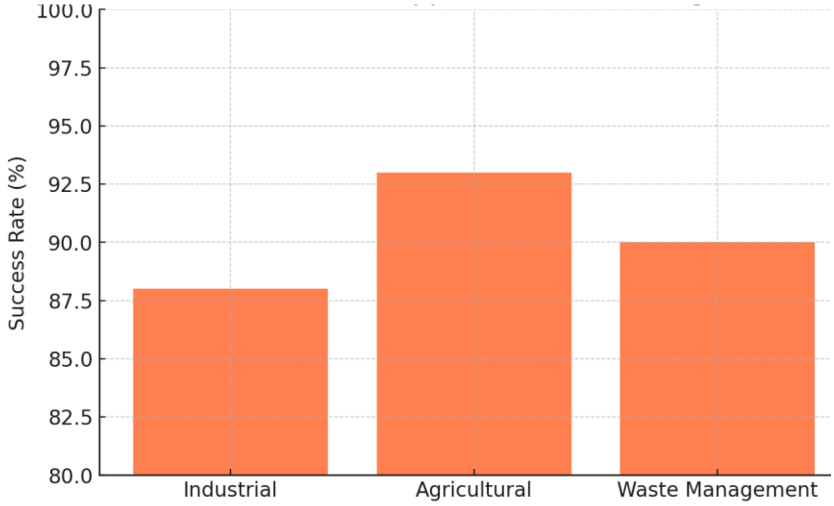

Fig. 9. Real-World Application and Testing (Rakhimkul et al., 2019; Chen et al., 2022)

5.4. Multi-Modal Sensing Integration :

The bar chart in Fig.10 depicts the improvements achieved by integrating multiple sensing modalities, showing the most substantial enhancements with full multi-modal integration.

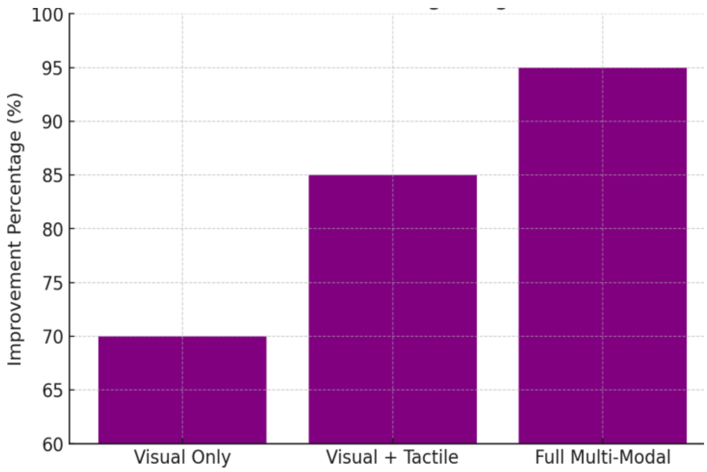

Fig. 10. Multi-Modal Sensing Integration (Kasaei et al., 2021; Xionget al., 2023)

6. YOLO INTEGRATION INTO ROBOTICS MANIPULATOR

In this review paper, we provide a complete and comprehensive overview of the integrations and implementations of YOLO algorithms and the other algorithms such as Deep Reinforcement Learning (DRL) within robotic systems as reported in various studies. YOLO, particularly the YOLOv5 model, is widely used for its real-time object detection capabilities due to its efficient convolutional neural network (CNN) architecture and robust training on extensive datasets like COCO. Studies often fine-tune these models with custom datasets tailored to specific robotic applications. The detected objects are then processed by DRL models, typically employing architectures like Deep Q-Networks (DQN) implemented using frameworks such as Tensor Flow and OpenAI Gym. These models are trained in simulated environments to develop optimal grasping strategies, which are further refined through multi-modal sensory feedback from devices like tactile sensors. The integration of the aforementioned components is always facilitated by the software called Robot Operating System (ROS), to ensuring easy communication between the detection and manipulation modules, therefore enhancing the efficiency and the overall accuracy of robotic manipulator operations.

7. CONCLUSION

The exploration and utilization of objects hold considerable promise in transforming various sectors such as education, agriculture, industry, and medicine, thereby streamlining and enriching our daily lives. This study was aimed at uncovering a deep learning methodology adept at recognizing objects, with an eye towards its application in human-centric uses. Not long ago, the endeavor to detect and categorize objects within images was fraught with difficulties, bordering on the impossible. However, the advent of computer vision and deep learning has remarkably simplified these tasks, making object detection considerably more accessible. A plethora of computer vision techniques and algorithms, particularly the deep network-based YOLO method in its iterations like v2, v3, v4, R-YOLO, and PP-YOLO, have been scrutinized. Literature reviews have highlighted various innovative combinations, such as melding YOLOv3 with Center Net deep learning frameworks, integrating YOLOv4 with particle filter (PF) techniques, and the synergistic application of YOLOv3 alongside deep reinforcement learning strategies. These methodologies demonstrate exceptional accuracy and efficiency in object detection and identification tasks. Deep learning approaches thus significantly ease the challenge of object detection in computer vision applications, playing a pivotal role across numerous industries by offering extensive support and ad- vantages. By improving object identification in images or videos, these innovations promise to mitigate numerous issues faced by individuals, enhancing the efficacy and precision of intelligent systems equipped with computer vision capabilities.

8. REFERENCES

Advincula, A.P. and Wang, K., 2009. Evolving role and current state of robotics in minimally invasive gynecologic surgery. Journal of Minimally Invasive Gynecology, 16(3), pp.291-301.

Arulkumaran, K. et al., 2017. Deep reinforcement learning: A brief survey. IEEE Signal Processing Magazine, 34(6), pp.26-38.

Bochkovskiy, A., Wang, C.Y. and Liao, H.Y.M., 2020. YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934.

Chen, C.H., Huang, H.P. and Lo, S.Y., 2011. Stereo-based 3D localization for grasping known objects with a robotic arm system. In 2011 9th World Congress on Intelligent Control and Automation. IEEE.

Chen, G.H., Jun-Yi, W. and Ai-Jun, Z., 2019. Transparent object detection and location based on RGB-D camera. Journal of Physics: Conference Series, 1183(1).

Chen, L. et al., 2023. Perceiving unseen 3D objects by poking the objects. arXiv preprint arXiv:2302.13375.

Chen, Q. et al., 2022. Vision-based impedance control of a 7-DOF robotic manipulator for pick-and-place tasks in grasping fruits.

Chen, Y.L., Cai, Y.R. and Cheng, M.Y., 2023. Vision-based robotic object grasping: a deep reinforcement learning approach. Machines, 11(2), p.275.

Chen, Z. et al., 2022. Towards generalization and data efficient learning of deep robotic grasping. In 2022 IEEE 17th Conference on Industrial Electronics and Applications (ICIEA). IEEE.

Choi, C. et al., 2018. Learning object grasping for soft robot hands. IEEE Robotics and Automation Letters, 3(3), pp.2370-2377.

Cong, X. et al., 2023. A review of YOLO object detection algorithms based on deep learning. Frontiers in Computing and Intelligent Systems, 4(2), pp.17-20.

Czajewski, W. and Kołomyjec, K., 2017. 3D object detection and recognition for robotic grasping based on RGB-D images and global features. Foundations of Computing and Decision Sciences, 42(3), pp.219-237.

Dafoe, A., Bachrach, Y., Hadfield, G., Horvitz, E., Larson, K. and Graepel, T., 2021. Cooperative AI: machines must learn to find common ground.

Daugherty, P.R. and Wilson, H.J., 2018. Human+ machine: Reimagining work in the age of AI. Harvard Business Press.

Dong, H. et al., 2020. Deep Reinforcement Learning. Springer.

Du, G. et al., 2021. Vision-based robotic grasping from object localization, object pose estimation to grasp estimation for parallel grippers: a review. Artificial Intelligence Review, 54(3), pp.1677-1734.

Ekvall, S., Kragic, D. and Hoffmann, F., 2005. Object recognition and pose estimation using color cooccurrence histograms and geometric modeling. Image and Vision Computing, 23(11), pp.943-955.

Fairag, M., Almahdi, R.H., Siddiqi, A.A., Alharthi, F.K., Alqurashi, B.S., Alzahrani, N.G., Alsulami, A., Alshehri, R., Alzahrani, N.G., Alsulami, A.S. et al., 2024. Robotic revolution in surgery: Diverse applications across specialties and future prospects review article. Cureus, 16(1).

Fan, Q., Rao, Q. and Huang, H., 2023. Multitarget flexible grasping detection method for robots in unstructured environments. CMES-Computer Modeling in Engineering & Sciences, 137(2).

Fullan, M. and Langworthy, M., 2013. Towards a new end: New pedagogies for deep learning.

Gai, R., Chen, N. and Yuan, H., 2023. A detection algorithm for cherry fruits based on the improved YOLO-v4 model. Neural Computing and Applications, 35(19), pp.13895-13906.

Gao, M. et al., 2021. A hybrid YOLOv4 and particle filter based robotic arm grabbing system in nonlinear and non-Gaussian environment. Electronics, 10(10), p.1140.

Huang, Y.Q. et al., 2020. Optimized YOLOv3 algorithm and its application in traffic flow detections. Applied Sciences, 10(9), p.3079.

Hussain, M., 2023. When, where, and which? Navigating the intersection of computer vision and generative AI for strategic business integration. IEEE Access, 11, pp.127202-127215.

Issa, Abbas H., and Ali H. Majeed. “Intelligent Sensor Fault Detection Based on Soft Computing”. *Kufa Journal of Engineering*, vol. 4, no. 1, Jan. 2014, pp. 113-24, doi:10.30572/2018/KJE/411246.

Jain, S.K. et al., 2023. Articulated robot arm for garbage disposal in hospital environment. In ITM Web of Conferences, 56. EDP Sciences.

Jiang, P. et al., 2020. Depth image-based deep learning of grasp planning for textureless planar-faced objects in vision-guided robotic bin-picking. Sensors, 20(3), p.706.

Kang, H., Zhou, H. and Chen, C., 2020. Visual perception and modeling for autonomous apple harvesting. IEEE Access, 8, pp.62151-62163.

Kasaei, H. and Kasaei, M., 2023. MVGrasp: Real-time multi-view 3D object grasping in highly cluttered environments. Robotics and Autonomous Systems, 160, p.104313.

Kasaei, H. et al., 2021. Simultaneous multi-view object detection and grasping in open-ended domains. arXiv preprint arXiv:2106.01866.

Kaymak, C. and Aysegul, U.C.A.R., 2018. Implementation of object detection and detection algorithms on a robotic arm platform using Raspberry Pi. In 2018 International Conference on Artificial Intelligence and Data Processing (IDAP). IEEE.

Kheder, Harem Ali. "HUMAN-COMPUTER INTERACTION: ENHANCING USER EXPERIENCE IN INTERACTIVE SYSTEMS". *Kufa Journal of Engineering*, vol. 14, no. 4, Oct. 2023, pp. 23-41, doi:10.30572/2018/KJE/140403.

Kim, J.H. et al., 2022. Object detection and classification based on YOLO-v5 with improved maritime dataset. Journal of Marine Science and Engineering, 10(3), p.377.

Koşer, H.E., 2023. Determination of angular status and dimensional properties of objects for grasping with robot arm. IEEE Latin America Transactions, 21(2), pp.335-343.

Kragic, D. and Christensen, H.I., 2002. Model based techniques for robotic servoing and grasping. In IEEE/RSJ International Conference on Intelligent Robots and Systems, 1. IEEE.

Levine, S., Pastor, P., Krizhevsky, A., Ibarz, J. and Quillen, D., 2018. Learning hand-eye coordination for robotic grasping with deep learning and large-scale data collection. The International Journal of Robotics Research, 37(4-5), pp.421-436.

Lin, C.C. et al., 2016. Vision based object grasping of industrial manipulator. In 2016 International Conference on Advanced Robotics and Intelligent Systems (ARIS). IEEE.

Liu, J. et al., 2023. Design of a virtual multi-interaction operation system for hand-eye coordination of grape harvesting robots. Agronomy, 13(3), p.829.

Liu, N. et al., 2022. Collaborative viewpoint adjusting and grasping via deep reinforcement learning in clutter scenes. Machines, 10(12), p.1135.

Mao, Q.C. et al., 2019. Mini-YOLOv3: Real-time object detector for embedded applications. IEEE Access, 7, pp.133529-133538.

Mohammed, M.M., Al-Khafaji, M.M. and Abbas, T.F., 2023. Smart robot vision for a pick and place robotic system. Engineering and Technology Journal, 41(6), pp.1-15.

Moran, M.E., 2007. Evolution of robotic arms. Journal of Robotic Surgery, 1(2), pp.103-111.

Prattichizzo, D., Pozzi, M., Baldi, T.L., Malvezzi, M., Hussain, I., Rossi, S. and Salvietti, G., 2021. Human augmentation by wearable supernumerary robotic limbs: review and perspectives. Progress in Biomedical Engineering, 3(4), p.042005.

Qi, H. and Gong, S., 2023. Network architecture based on improved dense-fusion algorithm research on the detection and grasping method of robotic arm. In International Conference on Computer, Artificial Intelligence, and Control Engineering (CAICE 2023), 12645. SPIE.

Rakhimkul, S. et al., 2019. Autonomous object detection and grasping using deep learning for design of an intelligent assistive robot manipulation system. In 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC). IEEE.

Redmon, J. and Farhadi, A., 2018. YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767.

Ren, Y. et al., 2018. Vision based object grasping of robotic manipulator. In 24th International Conference on Automation and Computing (ICAC). IEEE.

Sekkat, H. et al., 2021. Vision-based robotic arm control algorithm using deep reinforcement learning for autonomous objects grasping. Applied Sciences, 11(17), p.7917.

Shahria, M.T. et al., 2022. A comprehensive review of vision-based robotic applications: Current state, components, approaches, barriers, and potential solutions. Robotics, 11(6), p.139.

She, Q. et al., 2020. Openloris-object: A robotic vision dataset and benchmark for lifelong deep learning. In 2020 IEEE International Conference on Robotics and Automation (ICRA). IEEE.

Sun, T. et al., 2023. A detection method for soft objects based on the fusion of vision and haptics. Biomimetics, 8(1), p.86.

Tutsoy, O., 2023. A review of recent advancements in deep machine learning, artificial intelligence, object detection, and human-robot interactions approaches for assistive robotics. Ph.D., Fatma Gongor.

Wang, Q. et al., 2023. Design, integration, and evaluation of a robotic peach packaging system based on deep learning. Computers and Electronics in Agriculture, 211, p.108013.

Wei, A.H. and Chen, B.Y., 2020. Robotic object recognition and grasping with a natural background. International Journal of Advanced Robotic Systems, 17(2), p.1729881420921102.

Xiong, S., Tziafas, G. and Kasaei, H., 2023. Enhancing fine-grained 3D object detection using hybrid multi-modal vision transformer-CNN models. In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2023).

Yu, J. and Zhang, W., 2021. Face mask wearing detection algorithm based on improved YOLO-v4. Sensors, 21(9), p.3263.

Yu, Y. et al., 2020. Real-time visual localization of the picking points for a ridge-planting strawberry harvesting robot. IEEE Access, 8, pp.116556-116568.

Zarif, M.I.I. et al., 2022. A vision-based object detection and localization system in 3D environment for assistive robots' manipulation. In Proceedings of the 9th International Conference of Control Systems, and Robotics (CDSR'22).

Zhao, Y., Shi, Y. and Wang, Z., 2022. The improved YOLOv5 algorithm and its application in small target detection. In International Conference on Intelligent Robotics and Applications. Springer, pp.679-688.

Zhong, M. et al., 2019. Assistive grasping based on laser-point detection with application to wheelchair-mounted robotic arms. Sensors, 19(2), p.303.