Tuijin Jishu/Journal of Propulsion Technology
ISSN: 1001-4055
Vol. 46 No. 2 (2025)

A Review on AI-Enhanced Adaptive Grasping for Industrial Robots

Sharan S[1], Mr. Deepak S[2]
Department of Mechanical Engineering, Ramaiah Institute of Technology, Bengaluru-54
Corresponding Author: Sharan S

Abstract:- The evolution of intelligent automation in industrial domains has catalysed the necessity for advanced robotic systems capable of exhibiting human-like dexterity and situational awareness. This study delineates an AI-enhanced adaptive grasping framework meticulously engineered for industrial robots operating in complex and unstructured environments. The proposed system synergises deep convolutional neural networks for precise object detection and pose estimation with reinforcement learning algorithms that iteratively refine grasping policies through experiential feedback. Furthermore, the integration of multimodal sensory data—encompassing visual, tactile, and proprioceptive inputs—enables the robot to dynamically modulate its grasping strategy in accordance with fluctuating object properties, such as geometry, texture, and weight distribution. Empirical evaluations conducted in a simulated industrial milieu reveal a substantial enhancement in grasp reliability, manipulation efficiency, and task generalisability when juxtaposed with conventional methodologies. The research underscores the transformative potential of artificial intelligence in augmenting the autonomy, flexibility, and operational robustness of robotic manipulators, thereby paving the way for more resilient and adaptive manufacturing ecosystems.

Keywords: *Adaptive Grasping, Industrial Robotics, Artificial Intelligence, Deep Learning, Reinforcement Learning, Multimodal Perception, Robotic Manipulation, Intelligent Automation*

1. Introduction

The advent of the Fourth Industrial Revolution has engendered a profound transformation in the operational paradigms of manufacturing and logistics industries, with a pronounced emphasis on the integration of intelligent automation technologies. Among the most critical components of such transformative systems are industrial robots, whose capacity to perform repetitive and labour-intensive tasks with high precision has rendered them indispensable in modern production environments. However, the conventional deployment of robotic manipulators has long been constrained by their reliance on deterministic programming, inflexible task parameters, and limited environmental adaptability. These limitations are particularly evident in grasping and manipulation tasks, wherein traditional systems often falter when presented with objects of unfamiliar shapes, variable orientations, or diverse material properties. In real-world scenarios, such as bin picking, assembly, and packaging, these inadequacies can lead to substantial inefficiencies, elevated error rates, and increased system downtime.

To overcome these challenges, recent advancements have witnessed the proliferation of artificial intelligence (AI)-driven methodologies aimed at endowing robots with human-like perceptual and cognitive capabilities. In particular, the amalgamation of deep learning, computer vision, and reinforcement learning has emerged as a promising avenue for enhancing robotic grasping through adaptive mechanisms. Adaptive grasping refers to a robot's ability to autonomously perceive its environment, interpret contextual cues, and modify its grasping strategy in real time to accommodate the specific characteristics of the target object. This level of responsiveness is made possible through the utilisation of convolutional neural networks (CNNs) for object recognition and pose estimation, as well as reinforcement learning frameworks that iteratively refine control policies based on sensory feedback and task outcomes.

Moreover, the incorporation of multimodal sensory systems—encompassing visual, tactile, and proprioceptive data streams—further augments the robot's situational awareness, enabling it to make more nuanced and contextually appropriate decisions. These sensors provide critical information pertaining to object geometry, surface texture, compliance, and weight distribution, all of which are indispensable for executing robust and stable grasps.

The AI-enhanced framework thereby transitions from a pre-scripted operational model to a learning-oriented paradigm wherein the robot acquires and continuously updates its manipulation competencies through experiential learning.

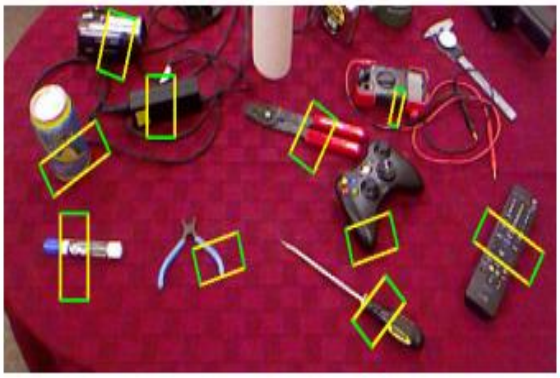


Fig. 1: Detecting Robotic Grasps

Fig. 1 illustrates implications of such innovations extend well beyond the confines of academic inquiry, offering tangible benefits for industrial stakeholders.

Enhanced grasping accuracy and reduced task execution time translate into heightened productivity, while the ability to handle a broad spectrum of objects mitigates the need for task-specific programming or mechanical redesign. Furthermore, adaptive systems can accommodate process variability and unexpected perturbations, thereby improving the overall resilience and agility of automated production lines.

This research, therefore, seeks to contribute to the growing body of knowledge surrounding AI-powered robotic systems by proposing and evaluating a novel adaptive grasping framework. Through rigorous experimental validation in a simulated industrial environment, the study aims to demonstrate the efficacy of the proposed approach in elevating the functional intelligence, operational flexibility, and real-world applicability of industrial robots.

There are a few other studies surveying AI-driven adaptive grasping system that enables industrial robots to intelligently perceive, learn, and adjust their grasping strategies. Honglak Lee[1] uses deep learning techniques to enhance robotic grasp detection by accurately identifying optimal grasping positions in complex visual scenes. The approach significantly improves robotic manipulation by integrating convolutional neural networks for precise spatial feature extraction. Suhas Kadalagere Sampath[2]human-like robotic manipulation, focusing on the capabilities of dexterous robotic hands to emulate nuanced human motor skills. Emphasis is placed on advancements in control strategies, tactile sensing, and adaptive learning for refined object interaction. Olaf Ronneberger[3] discussed precise biomedical image segmentation with limited training data.

It emphasises the model's encoder-decoder structure, enabling accurate delineation of complex anatomical features.Jianshu Zhou[4]robotic grasp detection, enabling machines to identify viable grasping points with high precision. The study underscores the efficacy of convolutional neural networks in interpreting visual data for autonomous manipulation tasks.Igor Zubrycki[5]a test setup for multi-finger gripper control, utilizing the Robot Operating System (ROS) to facilitate seamless communication and integration.

The system focuses on enhancing the precision and adaptability of robotic manipulation. Ran Qin[6]RGB-D grasp detection by leveraging depth-guided learning and cross-modal attention mechanisms to improve grasp prediction accuracy. The approach enhances the integration of visual and depth information. Mayada Abdalsalam [7] incorporates attention mechanisms and advanced U-Net architectures within generative grasping convolutional neural networks. The approach improves the model's ability to focus on critical regions, leading to more accurate and efficient grasp predictions in diverse environments. Douglas Morrison[8] is a real-time generative grasp synthesis framework that enables closed-loop robotic grasping through continuous feedback and dynamic adjustment.

It emphasises the integration of perception and control, allowing robots to adaptively refine grasp strategies in unstructured environments. Carlo Alberto Avizzano[9], the fusion of robotics with computer-integrated manufacturing to optimise industrial automation and operational precision. It accentuates the role of intelligent robotic systems in streamlining production processes and fostering agile, data-driven manufacturing environments. Zhen Xie[10]robotic tactile grasping to enable hyper-personalised pick-and-place operations along production lines. It emphasises the integration of tactile feedback and machine learning to adapt grasp strategies to diverse object geometries and user-specific requirements[11].

Together, these technologies enable adaptive, precise decision-making in dynamic defense and surveillance environments.

Although prior research exists in this domain, it is often narrowly concentrated on specific aspects of AI-enhanced adaptive grasping for industrial robots. This paper endeavors to address the existing research gap by adopting a more holistic perspective. The principal contributions of this study are delineated below:

1) Introduced deep learning methodologies to enhance robotic grasp detection by identifying optimal grasping configurations in visually complex scenarios, thereby advancing autonomous manipulation.

2) Reviewed advancements in human-like robotic manipulation using dexterous hands, with a focus on control strategies, tactile perception, and adaptive learning for emulating refined human motor skills.

3) Proposed a U-Net-based convolutional network for biomedical image segmentation, achieving high accuracy even with limited annotated data through an encoder-decoder framework.

4) Demonstrated the effectiveness of convolutional neural networks in robotic grasp detection, facilitating precise interpretation of visual input for autonomous grasp execution.

5) Developed a test bed for multi-finger gripper control using the Robot Operating System (ROS), enhancing system integration, precision, and real-time responsiveness in robotic tasks.

6) Employed depth-guided learning with cross-modal attention mechanisms to improve RGB-D grasp detection, enabling more robust perception and decision-making in cluttered environments.

7) Integrated attention mechanisms and enhanced U-Net architectures within generative grasping CNNs to improve focus on salient image regions, thus elevating grasp prediction accuracy.

8) Proposed a real-time generative grasp synthesis model that enables closed-loop robotic grasping, integrating perception and control for adaptive interaction in dynamic settings.

9) Investigated robotics within computer-integrated manufacturing, highlighting the potential of intelligent automation in increasing efficiency, precision, and agility in production workflows.

10) Explored data-driven tactile grasping for personalised pick-and-place tasks, leveraging tactile sensors and machine learning to tailor robotic responses to varying object geometries and user needs.

The review of this paper is organized as follows: Section 2 provides an overview of the algorithms utilized in the proposed system. Section 3 discusses prevailing security technologies relevant to autonomous surveillance. Section 4 presents the system architecture in detail. Section 5 explores the practical applications of the proposed approach. Finally, Section 6 concludes the study and outlines potential future work.

2. Overview of Algorithms

This section provides a concise overview of the core algorithms underpinning AI-enhanced adaptive grasping in industrial robotics. These algorithms are generally classified into model-based and data-driven approaches**,** each offering distinct advantages depending on the nature of the task and available data. Model-based methods rely on explicit physical and geometric models of the objects and the robotic system, offering precise control in structured environments. In contrast, data-driven techniques, particularly those based on machine learning and deep learning**,** learn grasping strategies directly from data, enabling adaptability to complex, unstructured, or dynamic environments. Among these**,** supervised learning requires labeled grasping outcomes, while reinforcement learning and unsupervised learning enable autonomous skill acquisition without extensive manual annotation. Deep learning architectures, such as convolutional and graph neural networks, excel in extracting high-level representations from sensory inputs, thereby facilitating robust and flexible grasp planning in diverse industrial scenarios.

2.1 Convolutional Neural Network (CNN)

Convolutional Neural Networks represent a distinguished subset of deep neural architectures, meticulously engineered to process and interpret data exhibiting a spatial or temporal structure, such as visual imagery[12]. Rooted in the principles of biological vision, CNNs emulate the human visual cortex by hierarchically learning features from input data through a series of structured layers[13].At their core, CNNs are composed of convolutional layers, non-linear activation functions, pooling (subsampling) layers, and fully connected layers[14].The convolutional layer employs a set of learnable kernels or filters that systematically traverse the input tensor, executing discrete convolutional operations to extract localised features. This mechanism can be mathematically expressed as:

$$Y(i,j) = (X * K)(i.j) = \sum_{m} \sum_{n} X(i+ m,j+n) .K(m, n) \text{----}(1)$$

where X signifies the input matrix, K denotes the convolutional kernel, and Y(i,j)) is the output feature map at spatial position (i,j).Following convolution, the incorporation of activation functions, such as the Rectified Linear Unit (ReLU), introduces non-linearity into the network, thereby augmenting its capacity to model intricate and non-trivial relationships. Pooling layers, commonly implemented via max-pooling or average-pooling strategies, perform spatial downsampling, thereby reducing dimensionality, mitigating overfitting, and enhancing translational invariance[15]. The concluding segment of a CNN involves fully connected layers, wherein the multidimensional feature maps are flattened and subjected to a dense neural structure to facilitate classification, detection, or regression tasks.CNNs are celebrated for their exceptional proficiency in hierarchical feature extraction, progressively discerning elementary edges, textures, and shapes, culminating in the recognition of complex, abstract patterns. This renders them quintessential in a myriad of domains, including but not limited to computer vision, medical diagnostics, autonomous navigation, and remote sensing[16].

2.2 YOLO (You Only Look Once)

The You Only Look Once (YOLO) algorithm represents a seminal advancement in the field of real-time object detection [17], renowned for its singular capability to perform object localization and classification in a single evaluation of the input image. Unlike traditional object detection paradigms that decompose detection into multiple stages—such as region proposal generation followed by classification—YOLO adopts a holistic approach by reframing object detection as a single regression problem. The image is partitioned into a fixed grid, and for each grid cell, the algorithm simultaneously predicts bounding box coordinates, objectness scores, and class probabilities, thereby dramatically reducing computational redundancy and latency.

YOLO's architecture is grounded in deep convolutional neural networks (CNNs), typically comprising a backbone feature extractor (e.g., Darknet) followed by detection heads that produce the final predictions. Each detection head outputs a vector comprising the bounding box parameters (x,y,w,h) a confidence score representing the probability of an object's presence, and conditional class probabilities. The combined loss function optimizes both localization and classification accuracy, utilizing mean squared error for coordinate regression and cross-entropy loss for classification. This unified loss formulation enables the network to be trained end-to-end on full images, facilitating global reasoning about object positions and interrelations.

One of YOLO's most distinguishing attributes is its extraordinary inference speed, rendering it particularly advantageous in time-sensitive applications such as autonomous vehicles, military surveillance, and aerial reconnaissance using drones. Moreover, the algorithm's ability to generalize from natural scenes to unseen contexts underscores its robustness and versatility. Subsequent versions, such as YOLOv4 through YOLOv8, have introduced architectural enhancements including spatial pyramid pooling, anchor box optimization, and transformer-based modules, thereby improving detection accuracy while preserving computational tractability. In essence, YOLO epitomizes the convergence of precision and performance in real-time vision systems.

2.3 Single Shot MultiBox Detector (SSD)

The Single Shot MultiBox Detector (SSD)epitomises a paradigm shift in contemporary object detection methodologies by unifying the processes of object classification and localisation into a singular, streamlined computational pipeline. Unlike antecedent region-based approaches, SSD[18] obviates the necessity for exhaustive region proposal networks, thereby mitigating latency and enhancing throughput. This architectural innovation is realised through the deployment of a fixed ensemble of default bounding boxes of diverse aspect ratios and spatial scales, applied uniformly across hierarchical convolutional feature maps. These multi-scale representations confer robustness in detecting objects of varying dimensions and spatial configurations, a critical requisite for real-time surveillance and autonomous navigation systems.

From a computational perspective, SSD is underpinned by a dual-objective optimization framework that concurrently minimizes both localization error and confidence loss. The algorithm employs a Smooth L1 loss function to regress the predicted bounding box coordinates towards their ground-truth counterparts, whilst a categorical softmax loss governs the accuracy of class predictions. The fusion of deep feature hierarchies with convolutional[19] predictors permits the network to extract semantically rich features at multiple abstraction levels. This end-to-end trainable architecture ensures that SSD maintains a judicious equilibrium between detection precision and inference speed, rendering it eminently suitable for embedded vision systems and mission-critical domains such as border surveillance, aerial reconnaissance, and autonomous robotic patrols.

2.4 Generative Grasping Convolutional Neural Network (GG-CNN)

The Generative Grasping Convolutional Neural Network (GG-CNN)**[20][21]** represents a pivotal advancement in the domain of robotic grasp detection, particularly due to its capacity to perform real-time inference directly from depth images. Unlike conventional grasp detection algorithms that rely on time-intensive sampling of grasp candidates followed by classification, GG-CNN introduces a more efficient, fully convolutional framework that enables the generation of dense pixel-wise grasp predictions. At the core of GG-CNN lies the principle of generative prediction, whereby the model does not merely identify isolated grasp candidates but instead produces a complete grasp map[22]. This map encapsulates three essential parameters for each pixel: grasp quality, grasp angle, and gripper width. The grasp quality score indicates the confidence or feasibility of executing a successful grasp at a given location, while the grasp angle denotes the orientation of the end-effector. The gripper width specifies the distance between gripper fingers required to enclose the object at that point.

2.5 RCB-D Fusion Algorithms: A Comprehensive Exposition

In the realm of computer vision, RGB-D fusion algorithms represent a significant paradigm for the integration of visual (RGB)[23] and depth (D) information to enhance the performance of a variety of perceptual tasks, including object recognition, scene reconstruction, semantic segmentation, and robotic navigation. The term "RGB-D" refers to image data that encapsulates both colour (Red, Green, Blue) and depth information, typically captured via depth sensors such as Microsoft Kinect, Intel Real Sense, or LIDAR-equipped cameras[24]. The fusion of these complementary modalities facilitates the extraction of richer and more discriminative features, yielding improved robustness under diverse lighting conditions and complex spatial configurations[25].

At the core of RGB-D fusion lies the multimodal integration of two distinct types of data: the colour image and the depth map in times where H and W denote the height and width of the image, respectively[26]. The process of fusion may be approached at various levels—namely, at the raw data level (early fusion), feature level (intermediate fusion), or decision level (late fusion). Among these, feature-level fusion is most widely adopted due to its balance between computational efficiency and representational fidelity[27].

Recent advancements in deep learning have spurred the development of sophisticated fusion networks, such as dual-stream CNNs, residual fusion networks, and Transformer-based architectures[28]. A dual-stream architecture comprises two parallel networks—each dedicated to one modality—converging at a fusion point, often followed by joint processing via shared layers[29]. Residual fusion strategies integrate the fused features with skip connections to preserve spatial information, while Transformer-based models leverage cross-modal attention to capture long-range dependencies between RGB and depth features[30].

2.6 Deep Q-Networks (DQN)

Deep Q-Networks (DQN) represent an advanced reinforcement learning algorithm that integrates Q-learning with deep neural networks, addressing challenges posed by large or continuous state spaces, such as those found in images or intricate environments. In classical Q-learning, the Q-value function is stored in a table, which is effective for small-scale problems but becomes inefficient for high-dimensional inputs[30].DQN leverages a deep neural network to approximate the Q-value function, enabling the agent to process complex state spaces. The network is trained via experience replay, where the agent stores its previous interactions in a replay buffer and samples random mini-batches to mitigate correlations and enhance learning stability. Additionally, DQN employs a target network, which is updated less frequently to further stabilise the training procedure[31].

Table 1:Comparison of Different Algorithms

Ref	Year	Title	Algorithm	Advantages	Drawbacks
[32]	2022	Golden wheel spider-inspired rolling robots for planetary exploration	Bio-inspired locomotion control algorithm	Enhances mobility over uneven extraterrestrial terrains through adaptive rolling motion.	Limited stability and control during rapid directional changes.
[33]	2024	Biomimetic soft-legged robotic locomotion, interactions and transitions in terrestrial, aquatic and multiple environments.	Central Pattern Generator (CPG)-based control algorithm.	Facilitates smooth and adaptable gait transitions across diverse environmental conditions.	Complex parameter tuning is required to maintain stability and efficiency in varying terrains.
[34]	2023	Digital twin-enabled grasp outcomes assessment for unknown objects using visual-tactile fusion perception	Visual-tactile fusion with deep learning-based digital twin modelling.	Enables precise grasp assessment for unfamiliar objects through multimodal sensory integration.	High computational complexity and data dependency hinder real-time deployment.
[35]	2023	A YOLO-NL object detector for real-time detection	YOLO-NL (You Only Look Once with Non-Local attention) object detection algorithm.	Achieves real-time object detection with enhanced contextual awareness via non-local features.	Increased computational overhead due to the integration of non-local attention mechanisms.
[36]	2024	Customizable 6 degrees of freedom grasping dataset and an interactive training method for graph convolutional network	Graph Convolutional Network (GCN)-based grasping algorithm.	Enables precise 6-DoF grasp prediction through spatial relationship learning in point cloud data.	Requires extensive annotated datasets and is sensitive to variations in object geometry.
[37]	2025	EfficientGrasp: A Unified Data-Efficient Learning to Grasp Method for Multi-fingered Robot Hands	Data-efficient deep reinforcement learning algorithm.	Achieves high grasping performance with minimal training data, enhancing learning efficiency.	May exhibit reduced generalisability when encountering highly novel object configurations.

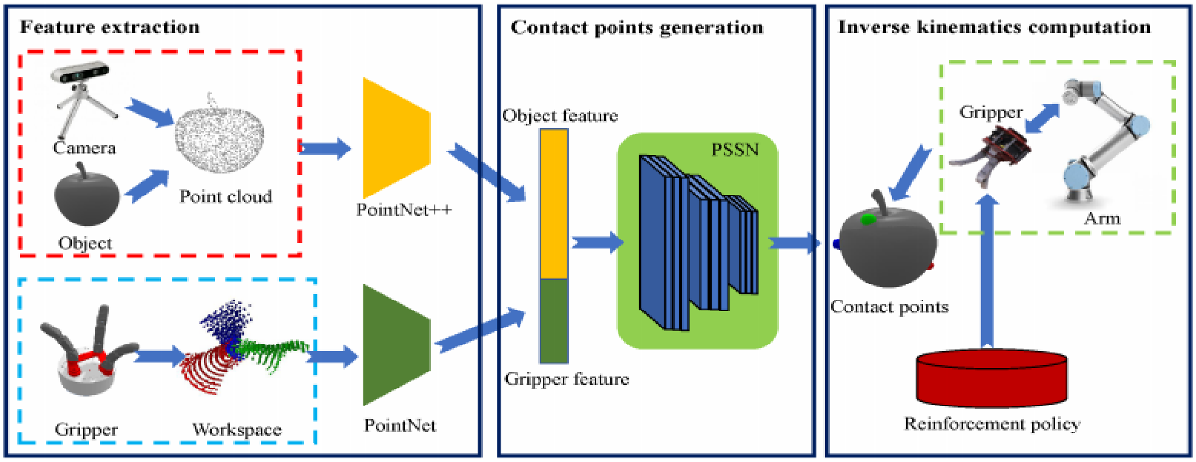


Fig. 2: The Flowchart of the Efficient Grasp framework comprises three principal phases: feature extraction, contact point generation, and inverse kinematics (IK) computation.

In the initial phase, feature extraction, salient features of both the target object and the gripper's workspace are extracted using the PointNet architecture.[38]These two sets of features are subsequently concatenated to form a unified representation, which is then input into a pre-trained Point-wise Semantic Segmentation Network (PSSN) to accurately identify potential contact points on the object's surface.

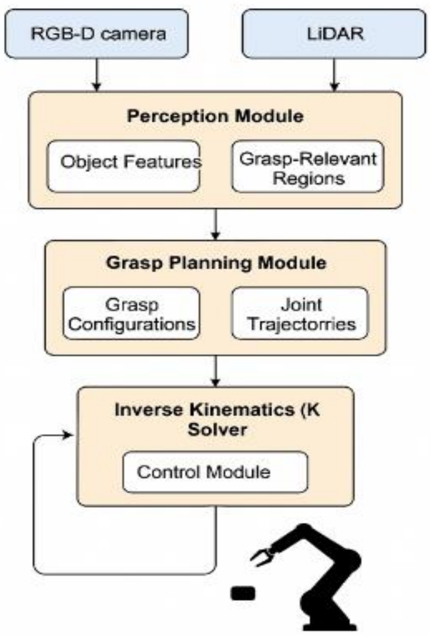
3. Prevailing Security Technologies

Numerous contemporary systems have been developed to advance AI-enhanced adaptive grasping in industrial robotics, aiming to facilitate precise and robust object manipulation within dynamic and unstructured environments. Among the most prominent is Dex-Net, a deep learning-based framework that employs a vast repository of synthetic data to train models for probabilistically robust grasp planning[39]. Similarly, GraspNet offers a comprehensive dataset and benchmark suite specifically designed to support deep grasp pose estimation and object interaction in cluttered scenes. In a parallel development, OpenAI's Dactyl utilizes reinforcement learning to endow a multi-fingered robotic hand with dexterous manipulation skills, enabling it to adapt to varying object geometries through simulation-to-reality transfer.

Industrial solutions such as ABB's YuMi and FANUC's collaborative robots seamlessly integrate artificial intelligence with high-resolution vision systems to achieve real-time object detection, spatial pose estimation, and adaptive grasping. These platforms frequently incorporate advanced neural architectures, such as convolutional neural networks (CNNs) and point cloud-based models like PointNet, in conjunction with reinforcement learning strategies to enhance adaptability and performance. Nonetheless, extant systems often encounter limitations[40].in terms of generalisability to novel objects, real-time computational efficiency, and seamless integration into complex, heterogeneous industrial ecosystems.

Several advanced systems, such as Dex-Net and GraspNet, exemplify the integration of deep learning for precise and adaptive grasp planning. Robotic platforms like OpenAI's Dactyl demonstrate dexterous manipulation through reinforcement learning and sim-to-real transfer[41].

4.



5. Architecture

Fig. 3: Architecture of AI-Enhanced Adaptive Grasping System for Industrial Robots

The diagram delineates a modular framework for robotic grasping, integrating perception, planning, and control mechanisms. Each constituent component contributes to a systematic and intelligent manipulation of objects within a robotic environment.

I Sensory Input Layer

The Sensory Input Layer comprises advanced sensing technologies that serve as the primary source of environmental data for the robotic system. The RGB-D camera captures both chromatic and depth information, thereby facilitating a comprehensive three-dimensional perception of the operational scene[42]. This dual-modality sensing capability enables the system to discern not only the visual appearance but also the spatial configuration of surrounding objects. Complementing this**,** LiDAR (Light Detection and Ranging) is employed as a precision-based sensor that utilises laser emissions to construct accurate spatial representations of the environment[43]. Together, these sensors provide the foundational input to the Perception Module, thereby enabling downstream processes such as object recognition, localisation, and grasp planning to be executed with high reliability and contextual awareness.

II Perception Module

The Perception Module is entrusted with the interpretation of raw sensory data to extract semantically meaningful features essential for effective robotic manipulation. It focuses primarily on identifying object features, which encompass the visual and geometric properties of an object, including its shape, dimensions, texture, and orientation[44].In addition, the module determines grasp-relevant regions**,** which are specific areas on the object's surface deemed optimal for grasping. These regions are selected based on multiple criteria such as mechanical stability, accessibility by the robotic end-effector, and the material composition of the object[45]. By extracting and synthesising this information, the module ensures that the subsequent grasp planning process is both data-driven and contextually informed, thereby enhancing the overall robustness and precision of the robotic grasping system[46].

III Grasp Planning Module

The Grasp Planning Module functions as a critical intermediary that translates perceptual insights into actionable motor commands. It synthesises the extracted data from the perception layer to formulate a coherent and feasible grasping strategy[47].Central to this process is the determination of grasp configurations, which define a set of permissible poses and orientations for the robotic end-effector to ensure a stable and secure interaction with the target object[48].In parallel, the module computes the joint trajectories**,** which represent a continuous sequence of joint angles required to transition the robotic arm smoothly from its initial configuration to the intended grasping posture. Through this dual-level planning of spatial alignment and kinematic movement, the module effectively facilitates the transformation of perceptual data into precise physical actions, thereby enabling robust and adaptive object manipulation.

IV Inverse Kinematics (IK) Solver and Control Module

The Inverse Kinematics (IK) Solver and Control Module constitutes the final operational stage in the robotic grasping pipeline, wherein planned motions are translated into physical execution. The Inverse Kinematics **Solver** serves as a mathematical framework that determines the precise joint angles necessary for a robotic manipulator to position its end-effector at a desired location and orientation in three-dimensional space[49].Once these joint parameters are computed, the Control Module orchestrates the execution of corresponding motor commands, ensuring the realisation of the planned trajectory with a high degree of precision, stability, and temporal responsiveness.

This stage culminates in the real-time actuation of the robot's mechanical components, thereby actualising the intended grasp and completing the perception-to-action loop with both accuracy and efficiency.This architecture exemplifies a cohesive and modular approach to robotic grasp synthesis, underpinned by robust perception, intelligent planning, and precise control. It embodies the principles of autonomy, adaptability, and efficiency in robotic manipulation.

6. Applications

1. Automated Assembly Lines
Robots can grasp parts of varying sizes, shapes, and materials with precision. Useful in electronics, automotive, and aerospace industries where parts often vary slightly.

2. Warehouse Automation & Logistics
Adaptive grasping enables robots to pick and place objects from bins or shelves. Essential in e-commerce and distribution centers (e.g., Amazon fulfillment).

3. Quality Inspection and Sorting
Robots can manipulate fragile or flexible items to inspect them without damage. Helps automated defect detection and sorting based on AI-identified criteria.

5. Recycling and Waste Management
Robots use AI to identify and grasp recyclable materials from mixed waste. Adaptive grasping handles irregular or deformed objects effectively.

6. Laboratory Automation
In research labs, robots can grasp pipettes, test tubes, or microplate**s**. Ensures delicate handling and precision, minimizing human error.

7. Conclusion

AI-enhanced adaptive grasping represents a significant leap in the capabilities of industrial robotics, facilitating the manipulation of a diverse range of objects with varying shapes, sizes, and material properties. Through the integration of advanced artificial intelligence techniques, including machine learning, computer vision, and real-time sensory feedback, these systems exhibit an unprecedented ability to dynamically adjust their grasping strategies in response to changing environmental conditions. This adaptability not only improves the accuracy and efficiency of robotic systems but also plays a crucial role in enhancing operational productivity while reducing the reliance on human intervention. As AI and sensor technologies continue to evolve, the potential of adaptive grasping to transform industrial processes and contribute to the broader vision of Industry 4.0 remains substantial. The ongoing development of these technologies is likely to drive further advancements in automation, creating new possibilities for their application across various sectors, including manufacturing, logistics, and material handling.

References

[1] Honglak Lee, Deep Learning for Detecting Robotic Grasps, 2013

[2] Suhas Kadalagere Sampath, Review on human-like robot manipulation using dexterous hands, 2022

[3] Olaf Ronneberger, U-Net: Convolutional Networks for Biomedical Image Segmentation, 2015

[4] Jianshu Zhou, A Soft-Robotic Gripper With Enhanced Object Adaptation and Grasping Reliability,2017

[5] Igor Zubrycki, Test setup for multi-finger gripper control based on robot operating system (ROS), 2013

[6] Ran Qin, RGB-D Grasp Detection via Depth Guided Learning with Cross-modal Attention, 2023

[7] Mayada Abdalsalam, Enhancing robotic grasping with attention mechanism and advanced UNet architectures in generative grasping convolutional neural networks, 2024

[8] Douglas Morrison, Closing the Loop for Robotic Grasping: A Real-time, Generative Grasp Synthesis Approach, 2018.

[9] Carlo Alberto Avizzano, Robotics and Computer-Integrated Manufacturing, 2024.

[10] Zhen Xie, DataDrivenRoboticTactileGraspingforHyperPersonalization Line Pick-and-Place, 2023.

[11] Darko Hercog, Simulation Model for Robotic Pick-Point Evaluation for 2-F Robotic Gripper, 2023.

[12] Marut Deo Sharma, Performance Comparison of Parallel and 3-Finger Gripper Using Human Hand Grasping Taxonomies, 2024.

[13] Zhikang Peng, The Enhanced Adaptive Grasping of a Soft Robotic Gripper Using Rigid Supports, 2024.

[14] Zhenjia Xu, AdaGrasp: Learning an Adaptive Gripper-Aware Grasping Policy, 2021.

[15] Xinquan Liang, Learning-based robotic grasping, 2023.

[16] Dongying Tian, Adaptive Motion Planning for Multi-fingered Functional Grasp via Force Feedback, 2024.

[17] Yang Yang, Attribute-Based Robotic Grasping with Data-Efficient Adaptation, 2025.

[18] Yitaek Kim, Robust Adaptive Safe Robotic Grasping with Tactile Sensing, 2024.

[19] Yilin Wang, Design of a spider inspired wheeled compliant leg for search mobile robots, 2024.

[20] Islam Mohamed Zaida, Virtual prototyping of vision-based tactile sensors design for robotic-assisted precision machining, 2024.

[21] Pengcheng Wei, Design of robot automatic navigation under computer intelligent algorithm and machine vision, 2022.

[22] Zhihua Chen, Control strategy of stable walking for a hexa pod wheel-leggedrobot,2022.

[23] Kang XU, High-adaption locomotion with stable robot body for planetary exploration robot carrying potential instruments on unstructured terrain. 2021.

[24] Xiaoyan Chen, Three-dimensional object detection and forward-looking control strategy for non-destructive grasp of thin-skinned fruits, 2024.

[25] Xudong Dong, A lightweight vehicle detection network model based on YOLOv5, 2022.

[26] Baojiang Li, Interactive learning for multi-finger dexterous hand: A model-free hierarchical deep reinforcement learning approach, 2024.

[27] Vitorino Biazi-Neto, FBG-based sensing system to improve tactile sensitivity of robotic manipulators working in unstructured environments, 2023.

[28] Peixing Li, Adaptive control algorithm for quadruped robots in unknown high-slope terrain, 2024.

[29] Yu Huang, A novel robotic grasping method for moving objects based on multi-agent deep reinforcement learning, 2024

[30] Zhihua Chen, Flexible gait transition for six-wheeled legged robot with unstructured terrains, 2024.

[31] Yuanxi Sun, Involute-arc-leg for Multi-legged Robot: High Stability and Low Energy Consumption, 2022.

[32] A. Western, Golden wheel spider-inspired rolling robots for planetary exploration, 2022

[33] Zhilin Yu, Biomimetic soft-legged robotic locomotion, interactions and transitions in terrestrial, aquatic and multiple environments, 2024.

[34] Zhuangzhuang Zhang, Digital twin-enabled grasp outcomes assessment for unknown objects using visual-tactile fusion perception, 2023.

[35] Yan Zhou, A YOLO-NL object detector for real-time detection, 2023.

[36] Wanhao Niu, Customizable 6 degrees of freedom grasping dataset and an interactive training method for graph convolutional network, 2024.

[37] Kelin Li, EfficientGrasp: A Unified Data-Efficient Learning to Grasp Method for Multi-fingered Robot Hands, 2025.

[38] Zhenning Zhou, Learning accurate and efficient three-finger grasp generation in clutter with an auto-annotated large-scale dataset, 2024.

[39] Francisco Yumbla, An Open-Source 3D Printed Three-Fingered Robotic Gripper for Adaptable and Effective Grasping, 2025.

[40] Weifei Hu, A grasps-generation-and-selection convolutional neural network for a digital twin of intelligent robotic grasping, 2022.

[41] Marwan Mohammed, Comprehensive Review on Reaching and Grasping of Objects in Robotics, 2021.

[42] Shuvo Kumar Paul, Object Detection and Pose Estimation from RGB and Depth Data for Real-time, Adaptive Robotic Grasping, 2021.

[43] Kento Kawaharazuka, Adaptive Robotic Tool-Tip Control Learning Considering Online Changes in Grasping State, 2024.

[44] Suhas Kadalagere Sampath, A Vision-Guided Deep Learning Framework for Dexterous Robotic Grasping Using Gaussian Processes and Transformers, 2022.

[45] Rhys Newbury, Deep Learning Approaches to Grasp Synthesis, 2022.

[46] Hanbo Zhang, Robotic Grasping from Classical to Modern: A Survey, 2024.

[47] Zhikang Peng, The Enhanced Adaptive Grasping of a Soft Robotic Gripper Using Rigid Supports,2024.

[48] Aiguo Chen, Adaptive Grasp Pose Optimization for Robotic Arms Using Low-Cost Depth Sensors in Complex Environments, 2025.

[49] Yang Yang, Attribute-Based Robotic Grasping with One-Grasp Adaptation, 2021.