# A Survey of Reinforcement Learning in Computer Vision for Object Detection and Tracking

www.surveyx.cn


## Abstract

This survey explores the integration of reinforcement learning (RL) with computer vision, emphasizing its transformative impact on tasks such as object detection and tracking. By addressing the limitations of traditional methods, RL enhances the adaptability and efficiency of visual systems, facilitating advancements in applications like autonomous vehicles, robotics, and surveillance. The synergy between RL and computer vision has significantly improved decision-making processes and operational efficiency across these domains. Deep reinforcement learning (DRL), leveraging deep neural networks (DNNs), plays a pivotal role in processing high-dimensional data, supporting the development of sophisticated visual perception systems. DNNs' ability to generalize across diverse datasets enhances the robustness of RL agents, enabling effective operation in dynamic environments. Despite these advancements, challenges in data efficiency, generalization, and model complexity persist, hindering the widespread adoption of RL in computer vision. Addressing these challenges through innovative exploration strategies and robust optimization techniques is crucial for future research. In conclusion, the integration of RL with computer vision has paved the way for intelligent systems capable of sophisticated visual perception and interaction, with promising potential for further advancements in various domains.




## 1 Introduction

### 1.1 Motivation for Integration

The integration of reinforcement learning (RL) with computer vision is driven by the need to overcome the limitations of traditional supervised learning, which relies heavily on extensive labeled datasets [1]. Additionally, RL's challenges in sparse reward environments hinder effective learning processes [2]. By leveraging visual inputs, RL enhances its ability to navigate and interpret complex environments, improving learning efficiency and adaptability.

A primary motivation for this integration is the enhancement of obstacle avoidance in dynamic environments, crucial for applications like autonomous driving and robotic navigation [3]. For instance, an Automatic Guided Vehicle (AGV) that autonomously navigates using onboard sensors such as RGB cameras and LiDAR illustrates the benefits of combining these technologies [4].

Moreover, bridging the gap between the RL and control communities is essential, as both share common goals and face methodological differences [5]. This convergence can lead to robust frameworks capable of addressing complex decision-making tasks, such as optimizing service delivery systems [6].

This integration also tackles sample complexity issues prevalent in visual navigation tasks, where deep RL methods often require extensive experience to learn from high-dimensional data [7]. By incorporating computer vision, RL can achieve more efficient exploration and policy learning, reducing reliance on large datasets and lengthy training periods.

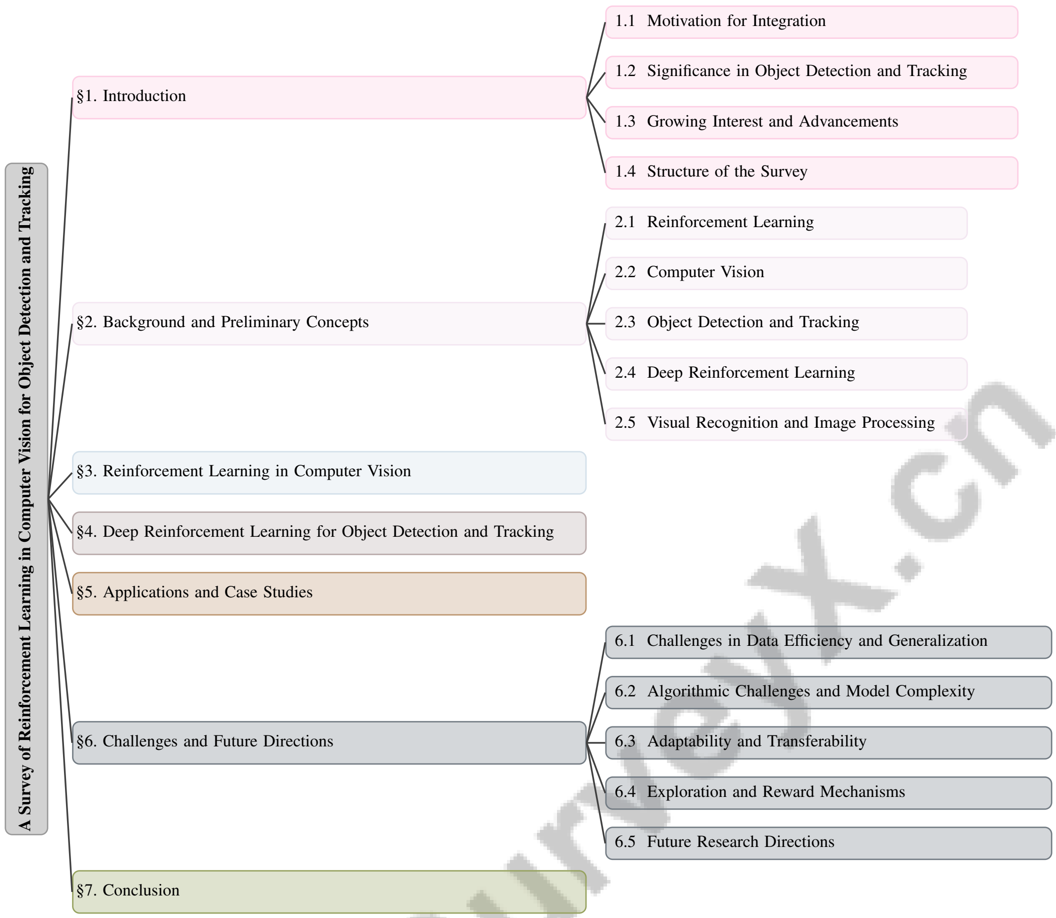



Figure 1: chapter structure



Furthermore, RL's integration with computer vision addresses complex multi-objective reinforcement learning (MORL) problems, enabling sophisticated decision-making in environments with competing objectives [8]. This capability is vital for systems that prioritize and balance multiple goals.

In robotic manipulation, particularly in tasks like cloth manipulation involving complex dynamics and high-dimensional state spaces, RL utilizing visual feedback shows promise in overcoming these challenges [9]. The limitations of existing RL methods in applications such as Atari games further underscore the need for innovative approaches that leverage modern computational capabilities [10].

## 1.2 Significance in Object Detection and Tracking

Reinforcement learning (RL) significantly enhances object detection and tracking through adaptive decision-making frameworks that manage the complexities of high-dimensional visual data. This integration facilitates the automation of imaging system designs by employing a context-free grammar (CFG) for systematic exploration of extensive configuration spaces—encompassing cameras, illumination sources, optical elements, and perception models—via reinforcement learning. This approach is particularly advantageous in partially observable environments, allowing for the joint optimization of imaging systems and task-specific perception models, thereby improving applications like depth estimation and autonomous vehicle camera rig design [11, 12, 13]. Traditional methods often falter in such conditions, making RL a valuable asset.

A key advantage of RL is its capacity to enhance learning efficiency, enabling agents to learn from limited interactions—especially beneficial in real-world scenarios where data collection is costly [14]. The integration of deep learning techniques into RL frameworks further refines decision-making across various domains, bolstering object detection and tracking capabilities [15]. Model-based RL



AI-generated, for reference only.

has demonstrated greater sample efficiency compared to model-free approaches, enhancing these capabilities [16].

RL's adaptability is amplified by its incorporation of contextual information into reward functions, significantly improving learning efficiency and enabling agents to adjust to dynamic environments [17]. This adaptability is crucial for tasks requiring generalization across varying scenarios. The use of off-policy learning methods and the Collect-and-Infer paradigm enhances learning from real robot experiences, vital for improving object detection and tracking in dynamic settings [18].

Integrating convolutional neural network (CNN) encoders with deep reinforcement learning (DRL) aims to enhance generalization and reduce training costs, particularly in socially contextual robotic navigation [19]. This integration exemplifies RL's potential to dynamically adapt recommendations based on real-time data, advancing intervention strategies in complex environments [20].

The emergence of complex RL applications has introduced challenges in analyzing and mitigating wireheading behaviors, highlighting the necessity for robust frameworks to address these complexities [21]. Nevertheless, advancements in RL methodologies position it as a pivotal technology in the evolution of intelligent systems, capable of operating effectively in intricate and dynamic environments. Notable methods such as DN-SARSA() excel in learning from delayed rewards and integrating continuous sensory-motor dynamics, making them suitable for real-world applications [22]. Applications like RL agents navigating race tracks to avoid obstacles further emphasize RL's role in enhancing navigation capabilities [23]. Incorporating state importance into reward learning, as proposed in Hindsight PRIOR, improves the alignment of learned rewards with human preferences [24].

Moreover, RL techniques adeptly handle non-differentiable objectives, proving valuable across various machine learning tasks, including computer vision [25]. The significance of RL in object detection and tracking is underscored by its ability to train autonomous driving agents in complex environments without extensive labeled datasets [1]. This survey emphasizes the importance of merging learning and control techniques to develop robust RL systems [5]. For example, AM-DQN employs a novel reward scheme and deep reinforcement learning to optimize taxi positioning dynamically, enhancing object detection and tracking efficiency in transportation contexts [6]. CityLearn exhibits notable improvements in sample efficiency, facilitating practical navigation algorithm deployment in real-world environments [7]. The Batch Asynchronous Advantage Actor Critic (BA3C) algorithm demonstrates significant advancements in training speeds and performance in Atari games, illustrating RL's potential impact on object detection and tracking [10].

### 1.3 Growing Interest and Advancements

The convergence of reinforcement learning (RL) and computer vision has attracted significant attention due to its potential to solve complex decision-making tasks across various domains. The unpredictable nature of environments such as warehouses necessitates the development of efficient models for autonomous navigation, where RL is pivotal [26]. In autonomous systems, recent advancements have showcased the efficacy of machine learning techniques in enhancing navigation and object detection capabilities, thereby improving operational efficiency in Unmanned Aerial Systems (UAS) [27].

The field of autonomous driving has experienced a surge in interest, integrating RL into motion planning and control systems to address the limitations of traditional approaches [28]. This growing interest is fueled by RL's success in achieving breakthroughs in various applications, such as AlphaGo and DeepStack, highlighting its transformative potential in gaming, robotics, and real-world systems [29].

Moreover, advancements in fine-tuning large vision-language models (VLMs) with RL have yielded significant performance improvements over traditional supervised learning methods [30]. This highlights deep RL's transformative impact as a key approach in artificial intelligence, particularly in gaming and robotics [15]. The capability of deep RL to address complex problems with minimal domain knowledge further underscores its potential across diverse applications [31].

Research has demonstrated the application of RL techniques to optimize non-differentiable objectives, leading to significant advancements in fields such as language modeling and robotics [25]. This



AI-generated, for reference only.

ability to tackle previously intractable problems has contributed to the growing interest in deep RL, particularly in domains like robotics and finance [32].

The emergence of eXplainable Reinforcement Learning (XRL) approaches, encompassing agent model-explaining, reward-explaining, state-explaining, and task-explaining methods, illustrates the expanding landscape of RL research [33]. These advancements underscore the increasing interest in merging RL with computer vision technologies, paving the way for innovative solutions to complex challenges across various industries.

### 1.4 Structure of the Survey

This survey provides a comprehensive analysis of the integration of reinforcement learning (RL) within computer vision, focusing on its applications in object detection and tracking. It explores how deep reinforcement learning, enhanced by deep learning advancements, transforms the interaction of autonomous systems with visual environments. The survey discusses key algorithms, such as deep Q-networks and policy optimization methods, while emphasizing the unique capabilities of deep neural networks in improving visual comprehension through RL [34, 35].

Beginning with an introduction to the motivation and significance of merging RL with computer vision, the survey highlights the growing interest and advancements in these fields. Subsequent sections delve into background concepts, offering definitions of core ideas such as reinforcement learning, computer vision, object detection, object tracking, deep reinforcement learning, visual recognition, and image processing.

The survey further explores RL's role in enhancing computer vision tasks, specifically object detection and tracking, by discussing the integration of RL algorithms and the application of hierarchical and deep reinforcement learning techniques. It examines how deep reinforcement learning integrates with object detection and tracking tasks, detailing the advantages of utilizing deep neural networks for managing complex visual data and decision-making.

Real-world case studies illustrate the application of RL to computer vision tasks, demonstrating its effectiveness in diverse fields such as robotics, where robots learn complex behaviors from human demonstrations; autonomous vehicles utilizing RL for navigation and decision-making; surveillance systems identifying anomalies in real-time; gaming achieving superhuman performance; simulation environments modeling complex agent interactions; healthcare optimizing treatment strategies; and adaptive systems enhancing user experiences through personalized interactions [36, 13, 37, 38]. The survey identifies current challenges in applying RL to computer vision and discusses future research directions, focusing on data efficiency, generalization, algorithmic challenges, model complexity, adaptability, transferability, exploration, and reward mechanisms.

The conclusion encapsulates the essential findings regarding RL's transformative role in computer vision, emphasizing its potential to enable fully autonomous systems capable of interpreting complex visual data. The discussion highlights the significance of ongoing research and development in deep reinforcement learning, particularly as it addresses challenges like real-time decision-making from high-dimensional inputs, and its applications across diverse fields, including robotics and interactive AI systems. This underscores the necessity for continued exploration and innovation in this rapidly advancing area of artificial intelligence [34, 35, 39, 40].The following sections are organized as shown in Figure 1.

## 2 Background and Preliminary Concepts

### 2.1 Reinforcement Learning

Reinforcement Learning (RL) is a machine learning framework where agents learn optimal decision-making by interacting with their environment to maximize cumulative rewards [41, 9]. This process is typically modeled as a Markov Decision Process (MDP), which provides a structured approach to decision-making in stochastic environments [10]. A critical aspect of RL is the exploration-exploitation trade-off, where agents must judiciously explore new actions and exploit known rewarding actions to achieve effective learning, particularly in environments with sparse rewards [42].

Deep Reinforcement Learning (DRL) extends RL by incorporating deep neural networks to handle high-dimensional sensory inputs, enabling applications in complex domains like robotic navigation and autonomous vehicles [10, 9]. DRL allows agents to autonomously discover effective strategies without prior system knowledge, though it requires substantial computational resources and extensive environmental interactions for policy development [9]. In scenarios with unknown dynamics, RL can be framed as an optimal control problem, suitable for continuous control tasks [43]. Techniques such as the Q-map model enhance policy learning by enabling generalization to unseen goals, improving exploration strategies and scalability [43].

Despite challenges like poor generalization, low sample efficiency, and interpretability issues, RL is rapidly advancing and finding applications across diverse domains, including natural language processing and healthcare. The integration of DRL with knowledge representation and reasoning (KRR) methods enhances RL systems' ability to create innovative solutions for complex decision-making scenarios by leveraging high-level domain knowledge [44, 45].

### 2.2 Computer Vision

Computer vision, an AI branch, enables machines to interpret visual data, encompassing tasks like image recognition, object detection, and scene understanding. These tasks extract meaningful information from images or video sequences, supporting decision-making in applications such as video violence recognition, autonomous driving, and robotic learning [46, 47, 36, 48, 11]. Integrating computer vision with RL has significantly advanced unmanned aerial systems (UAS), enhancing navigation and control through visual data interpretation [27]. By using computer vision algorithms to estimate UAV positions, robust RL controllers can perform complex maneuvers, such as landing under varying conditions [49].

Vision-language models (VLMs) illustrate computer vision's capabilities by merging visual recognition with language understanding for complex data interpretation tasks [30]. In robotics, the gym-gazebo2 toolkit integrates ROS 2 with the Gazebo simulator, facilitating the development of robotic environments reliant on computer vision for navigation and interaction [50]. This framework allows seamless visual data incorporation into RL pipelines, enabling agent training in simulated environments before real-world deployment.

Computer vision also plays a vital role in tasks such as cloth manipulation, where visual observations guide RL agents in performing precise actions [9]. By processing visual data, computer vision systems assist RL agents in executing complex manipulations, demonstrating the synergy between these technologies in addressing intricate challenges. The ongoing integration of DRL and other AI technologies significantly advances sectors like robotics, autonomous vehicles, and healthcare, improving diagnostics and patient care [27, 36, 37, 51].

### 2.3 Object Detection and Tracking

Object detection and tracking are core computer vision processes, enabling machines to perceive and interact with their environments. Object detection identifies and localizes objects within images or video frames, while object tracking monitors these objects' movement and trajectory across frames, facilitating behavior analysis over time [36, 52, 46, 53]. These tasks are crucial for applications like autonomous vehicles, surveillance systems, and robotics.

Integrating RL into object detection and tracking addresses traditional challenges by decomposing environments into manageable components, crucial for optimizing learning and generalization [54]. RL enables agents to learn adaptive policies that enhance detection and tracking performance in complex settings. In robotic environments, RL has been used to train agents in reaching target positions by learning appropriate motion trajectories, demonstrating potential for precise manipulation and control [55]. This capability is particularly advantageous in scenarios demanding high accuracy and repeatability, such as training robotic arms in simulated environments to achieve precise target points in 3D space [50].

The combination of RL with object detection and tracking techniques enhances systems' autonomy and efficiency in real-world environments. Leveraging RL's unique capabilities, these systems improve adaptability and robustness, making them suitable for diverse applications requiring precise perception. This adaptability is especially beneficial in complex environments, including robotics and recommendation systems, where agents learn from interactions and feedback to optimize performance over time. Advances in frameworks like RLInspect and Dopamine facilitate better assessment and benchmarking of RL models, ensuring effective navigation of challenging tasks and improved generalization across various scenarios [30, 51, 37, 38, 56].

### 2.4 Deep Reinforcement Learning

Deep Reinforcement Learning (DRL) integrates RL with deep neural networks, allowing the processing of high-dimensional data to tackle complex visual tasks [32]. This integration leverages deep learning's representational power to enhance RL's capabilities in challenging environments. DRL algorithms are categorized into model-free and model-based approaches, each employing distinct strategies for optimal policy learning [32]. Model-free methods, like Q-learning and policy gradients, learn policies directly from environmental interactions, while model-based methods construct an environment model for action simulation and planning.

A notable advancement in DRL's exploration capabilities is the Langevin Monte Carlo Least-Squares Value Iteration (LMC-LSVI), which samples the Q function from its posterior distribution, facilitating effective exploration and enhancing learning efficiency [42]. This exemplifies the innovative exploration strategies DRL can employ to navigate complex decision-making landscapes.

DRL's significance is further highlighted in multi-objective reinforcement learning (MORL), where it balances competing objectives using techniques like prioritized soft Q-decomposition. This method employs lexicographic decomposition of tasks, facilitating the reuse and adaptation of solutions to subtasks, enhancing DRL systems' adaptability and scalability in dynamic environments [44, 11, 8, 57]. By leveraging prioritized subtasks and a novel algorithm known as prioritized soft Q-decomposition (PSQD), this approach allows for efficient incremental learning without necessitating new environmental interactions, streamlining complex RL challenges [44].

In practical applications, DRL has proven effective in fields such as autonomous driving and robotic navigation, utilizing semantic segmentation and automatic curriculum learning to optimize agent training in real-world scenarios [32]. This capability to autonomously discover effective strategies without prior system dynamics knowledge is pivotal for advancing the autonomy and intelligence of systems across various domains.

### 2.5 Visual Recognition and Image Processing

Visual recognition and image processing are crucial components of computer vision, focusing on identifying, categorizing, and manipulating visual data to extract meaningful information. These processes are integral to developing intelligent systems capable of interacting with their environments. Visual recognition involves classifying and interpreting objects and patterns within images, enabling tasks such as face recognition, scene classification, and object detection. Image processing encompasses advanced techniques designed to enhance image quality, extract meaningful features, and support comprehensive visual data analysis. These techniques are vital in various applications, such as improving reinforcement learning performance in robotic tasks, optimizing semantic segmentation through active learning strategies, and automating imaging system design for specific tasks [46, 36, 12, 13, 53].

In RL contexts, visual recognition and image processing techniques provide agents with a robust understanding of their visual environments, enhancing decision-making capabilities. Visual grounding techniques, for example, create confidence maps that inform RL agents, improving their ability to interpret visual data and follow natural language instructions [58]. This integration enables the development of systems capable of navigating and interacting with complex environments, such as those encountered in autonomous driving and robotic navigation.

The application of RL to visual recognition tasks is exemplified by text-based games, which present varying complexity, state descriptions, and action possibilities. These games challenge RL agents' learning capabilities, providing a platform for testing and refining visual recognition algorithms [59]. Engaging with these environments allows RL agents to develop strategies for visual interpretation and decision-making, ultimately enhancing performance in real-world scenarios.

The synergy between visual recognition, image processing, and RL enhances the development of intelligent systems with advanced visual perception and interaction capabilities. This combination enables applications such as visual active search for identifying regions of interest in large geospatial areas, automating complex robotic tasks through human video demonstrations, and improving video violence recognition accuracy with transformer-based models. Moreover, deep reinforcement learning facilitates training autonomous agents that learn optimal behaviors from raw visual inputs, while tools like RLInspect enhance interpretability and reliability by providing insights into training processes [46, 36, 35, 38, 13]. As these technologies evolve, they promise to advance autonomous systems' capabilities across various applications, from robotics and surveillance to healthcare and beyond.

# 3 Reinforcement Learning in Computer Vision

The convergence of reinforcement learning (RL) and computer vision signifies a pivotal evolution in enhancing visual systems' capabilities. As computer vision advances, traditional algorithms often struggle with the complexities of dynamic environments, prompting the need for innovative strategies that enhance decision-making and efficiency. This section examines the integration of RL within computer vision tasks, showcasing how these methodologies overcome inherent limitations and foster the development of more sophisticated systems. Figure 2 illustrates this integration, highlighting key advancements and applications. The figure is divided into two main sections: the first focuses on the application of RL in computer vision tasks, emphasizing enhancements in visual systems and specific applications in robotics and computing. The second section delves into hierarchical and deep reinforcement learning applications, detailing frameworks and techniques that enhance system capabilities in managing complex visual tasks. Together, these elements underscore the transformative potential of RL in advancing computer vision technologies. The following subsection discusses specific RL applications in computer vision, highlighting practical implications and advancements achieved through this integration.

## 3.1 Integration of Reinforcement Learning in Computer Vision Tasks

Integrating reinforcement learning (RL) into computer vision tasks has led to sophisticated systems adept at processing complex visual data and enhancing decision-making. This integration mitigates traditional computer vision algorithms' inefficiencies by employing RL strategies to optimize performance in dynamic environments. A significant approach involves adapting many-goals updating using a single neural network, facilitating goal generalization without explicit definitions, thereby improving visual systems' adaptability [60].

In robotics, RL algorithms enhance navigation, as demonstrated by the CityLearn framework, which leverages visual place recognition and RL to refine navigation policies for mobile robots in diverse environments [7]. This framework exemplifies RL's potential to streamline decision-making by emphasizing relevant actions and improving adaptability to dynamic settings.

The use of convolutional neural networks (CNNs) in RL to generate Q-values and updates for multiple goals simultaneously overcomes traditional goal-conditioned networks' limitations, enhancing visual task performance [43]. In multi-objective settings, RL enables learning an approximate coverage set of policies, optimizing decision-making in complex environments.

The NAC algorithm further illustrates RL's potential to refine computer vision tasks by finding new policies that maximize expected returns while considering policy-switching costs, optimizing learning efficiency and adaptability [41]. Additionally, the TraKDis method integrates RL with visual feedback, bridging high-dimensional visual observations and low-dimensional state information, crucial for enhancing visual systems' interpretability and effectiveness [9].

In computing resource management, the BA3C algorithm applies RL principles to optimize training processes, effectively utilizing CPU resources and improving convergence rates, demonstrating efficiency gains achievable through RL integration [10]. The LMC-LSVI method employs Langevin Monte Carlo for noisy gradient descent updates, offering a principled exploration approach essential for navigating complex decision-making landscapes [42].

Integrating reinforcement learning into computer vision tasks has significantly advanced artificial intelligence by developing intelligent systems capable of autonomously learning and adapting to complex visual environments. This approach leverages deep learning techniques to tackle previously intractable problems, enabling robots to learn control policies directly from camera inputs in real-world settings and allowing software agents to interpret visual data for decision-making. Consequently,



AI-generated, for reference only.

these systems exhibit enhanced adaptability and efficiency, improving performance in dynamic and intricate scenarios [34, 35]. Through innovative exploration strategies, robust control policies, and distributed architectures, RL continues to push computer vision's boundaries, promising further advancements in diverse and dynamic settings.

### 3.2 Hierarchical and Deep Reinforcement Learning Applications

Hierarchical and deep reinforcement learning (DRL) applications in computer vision have significantly enhanced systems' capabilities to handle complex tasks by structuring decision-making into multiple layers. These applications leverage hierarchical frameworks to decompose intricate problems into simpler sub-tasks, facilitating more efficient learning and execution. A key advancement is pseudo-actions, allowing learning algorithms to utilize previously discarded data, enhancing performance in both continuous and discrete control tasks [61].

The Self-Attentional Credit Assignment for Transfer (Secret) framework exemplifies the integration of hierarchical and deep reinforcement learning by functioning without modifications to existing RL algorithms or agent architectures, making it highly adaptable to various applications [62]. This adaptability is crucial for deploying RL systems across diverse computer vision tasks, where transferring learned knowledge between contexts significantly enhances performance and efficiency.

The Policy Gradient from Demonstration and Curiosity (PGfDC) method introduces a dual reward formulation combining a demonstration term based on Jensen-Shannon divergence with a curiosity term from the agent's uncertainty [63]. This approach enables agents to balance exploiting known strategies with exploring novel actions, improving adaptability to dynamic and uncertain environments. The combination of hierarchical structures and deep learning techniques in PGfDC illustrates these methods' potential to optimize decision-making in complex visual tasks.

Overall, hierarchical and deep reinforcement learning applications in computer vision drive advancements by providing robust frameworks for managing complexity and enhancing adaptability. By utilizing a hierarchical task decomposition alongside deep neural networks' advanced representational capabilities, these applications are poised to significantly enhance intelligent systems' abilities to interpret and engage with visual data. This approach integrates diverse processing modules, including pre-trained networks for character recognition and symbolic transformation, coordinated by an RL controller to effectively tackle complex visual tasks. Moreover, incorporating interpretable decision-making strategies through techniques like Adversarial Inverse Reinforcement Learning ensures these systems perform efficiently and provide insights into their decision-making processes, fostering greater trust and usability in real-world applications [11, 64, 65, 66].

As shown in Figure 3, reinforcement learning, particularly its hierarchical and deep variants, has been making significant strides in computer vision, as illustrated by several compelling examples. A multi-head attention mechanism, depicted in one of the diagrams, represents a sophisticated neural network architecture instrumental in advancing natural language processing and machine learning. This mechanism utilizes multiple heads to focus on different parts of the input sequence, enhancing the model's ability to process complex data. Another example is a line graph showcasing the progression of games played over time, reflecting the dynamic nature of reinforcement learning algorithms as they improve through continuous interaction with their environment. This graph highlights these algorithms' competitive edge, with lines indicating performance metrics such as 'BEATING' and 'HIGHEST.' Lastly, a collage of images illustrates deep learning's practical applications in gaming and real-world scenarios, featuring scenes from classic video games like 'Pong' and other domains. Together, these examples underscore the versatility and potential of hierarchical and deep reinforcement learning in transforming computer vision and beyond [39, 67, 34].

## 4 Deep Reinforcement Learning for Object Detection and Tracking

The convergence of deep reinforcement learning (DRL) with object detection and tracking has emerged as a pivotal area of research, underscoring its potential to amplify intelligent systems' capabilities. Table 1 offers a detailed overview of the integration and advantages of deep reinforcement learning techniques in object detection and tracking, showcasing various methods and their contributions to enhancing visual task performance. Additionally, Table 3 provides a comprehensive comparison of different deep reinforcement learning methods, detailing their integration techniques,

| Category | Feature | Method |
|---|---|---|
| **Integration of Deep Reinforcement Learning Techniques** | Optimization Strategies | NAC[41], Q-map[43] |
| | Efficiency Enhancement | CL[7], BA3C[10] |
| | Model Compression | TD[9] |
| **Advantages of Deep Neural Networks** | Adaptive Optimization | RL-VLM[30] |

Table 1: This table provides a comprehensive summary of the integration of deep reinforcement learning (DRL) techniques in object detection and tracking, highlighting key methods and their associated features. It categorizes these methods based on optimization strategies, efficiency enhancements, and model compression, illustrating the diverse applications and advantages of DRL in visual tasks.

learning frameworks, and optimization strategies in the context of object detection and tracking. This section delves into the integration of DRL techniques and their contributions to enhancing visual task performance, followed by an exploration of specific advancements in detection and tracking frameworks.

## 4.1 Integration of Deep Reinforcement Learning Techniques

| **Method Name** | **Integration Techniques** | **Learning Frameworks** | **Optimization Strategies** |
|---|---|---|---|
| Q-map[43] | Convolutional Neural Networks | Gridworld Environments | All-goals Updates |
| CL[7] | Visual Place Recognition | Citylearn Framework | Multi-objective Optimization |
| TD[9] | Knowledge Distillation | Trakdis | Asynchronous Learning |
| BA3C[10] | - | - | Batching, Asynchronous Learning |
| NAC[41] | - | - | Batching, Asynchronous Learning |

Table 2: This table presents a comparative analysis of various deep reinforcement learning methods, highlighting their integration techniques, learning frameworks, and optimization strategies. The methods discussed include Q-map, CL, TD, BA3C, and NAC, each showcasing unique approaches to enhancing intelligent systems in complex visual data processing tasks.

The integration of deep reinforcement learning (DRL) techniques into object detection and tracking has significantly bolstered intelligent systems' capacity to manage complex visual data. Table 2 provides a comprehensive overview of different deep reinforcement learning methods, detailing their integration techniques, learning frameworks, and optimization strategies to contextualize their application in object detection and tracking. A key advancement is the deployment of convolutional neural networks (CNNs) to map input observations to Q-values for multiple goals, as demonstrated by Pardo et al. [43]. This approach facilitates simultaneous handling of multiple objectives within a unified framework, thereby enhancing detection and tracking task efficiency.

The CityLearn framework illustrates DRL's integration with visual place recognition, enabling sample-efficient policy learning in navigation algorithms [7]. By leveraging visual inputs, CityLearn improves navigation policy adaptability, allowing agents to function effectively in diverse dynamic environments.

TraKDis, a Transformer-based Knowledge Distillation approach, exemplifies DRL's application in visual tasks by transferring knowledge from a privileged agent to a vision-based agent [9]. This method is particularly effective in tasks like cloth manipulation, where visual feedback is critical for training agents in complex manipulations.

The BA3C method employs batching and asynchronous learning to optimize performance in visual tasks, highlighting DRL's potential to enhance training processes and convergence rates in object detection and tracking applications [10]. Additionally, Ma et al. propose a novel family of switching costs based on optimal transport principles, offering a flexible representation of switching dynamics in DRL, which is vital for adapting to the dynamic nature of detection and tracking tasks [41].

DRL techniques continue to advance the field by providing sophisticated methods that enhance system performance and adaptability. At the forefront of intelligent system development, DRL employs innovative architectures, exploration strategies, and multi-objective optimization techniques to improve visual data processing and interpretation. Its superiority in complex environments surpasses traditional algorithms, leading to applications in diverse fields such as healthcare and finance for tasks like data organization, scheduling, and natural language processing. Recent advancements have also yielded interpretable decision aids that translate opaque policies into comprehensible rules, significantly enhancing decision-making strategies across various applications [44, 11].

### 4.2 Advantages of Deep Neural Networks

Deep neural networks (DNNs) have revolutionized computer vision by offering advanced methodologies for processing intricate visual data, thus improving decision-making accuracy and efficiency. These networks utilize layered architectures to extract high-level features from raw inputs, facilitating the automatic extraction of relevant representations essential for tasks such as pattern recognition and object detection. The fusion of deep learning with reinforcement learning has further empowered agents to learn optimal actions through complex interactions with their environments, enabling applications ranging from superhuman performance in games to sophisticated robotic manipulation [64, 37, 40]. Their capability to autonomously learn hierarchical feature representations from raw data positions them as a critical component in advancing visual perception systems, particularly in high-dimensional data processing.

A primary advantage of DNNs in visual tasks is their capacity to generalize across diverse datasets, crucial for developing robust models that perform well in real-world scenarios. This generalization is largely attributed to deep architectures that enable the extraction of complex features and patterns often overlooked by traditional methods [30]. In reinforcement learning (RL), DNNs facilitate mapping high-dimensional visual inputs to actions or policies, thereby enhancing the adaptability and efficiency of RL agents in dynamic environments.

DNNs enhance the synergy between visual recognition and RL, enabling the creation of advanced systems capable of interpreting and interacting with their environments more effectively. This integration allows for improved decision-making processes in various applications, including robotics, gaming, and wildlife search scenarios. By leveraging DNNs, these systems can better utilize visual observations and learn from experiences, leading to nuanced and effective responses to dynamic surroundings [68, 37, 13, 38, 66]. In autonomous driving, for instance, DNNs process sensory data to inform decision-making processes, ultimately enhancing navigation and control strategies.

The scalability of DNNs significantly contributes to their widespread adoption, as they can be effectively trained on large datasets using modern computational resources. Techniques such as transfer learning and fine-tuning further enhance their applicability, enabling the efficient adaptation of pre-trained models to specific tasks or domains [30]. This flexibility is crucial for addressing the diverse challenges encountered in computer vision applications, ranging from object detection and tracking to scene understanding and beyond.

The advantages of DNNs, particularly their ability to autonomously learn and adapt strategies in complex environments, significantly enhance their effectiveness in processing intricate visual data and making informed decisions. This capability transforms the landscape of computer vision and extends its impact into diverse fields such as finance and symbolic reasoning, where traditional models often struggle to capture the nuances of dynamic data [64, 66, 69]. By leveraging powerful feature extraction capabilities and adaptability, DNNs continue to drive innovation in developing intelligent systems capable of sophisticated visual perception and interaction.

| Feature | Integration of Deep Reinforcement Learning Techniques | CityLearn | TraKDis |
|---|---|---|---|
| **Integration Technique** | Cnns For Q-values | Visual Place Recognition | Knowledge Distillation |
| **Learning Framework** | Unified Framework | Sample-efficient Policy | Transformer-based |
| **Optimization Strategy** | Multi-objective Optimization | Adaptability IN Navigation | Visual Feedback Training |

Table 3: This table presents a comparative analysis of various deep reinforcement learning techniques applied to object detection and tracking tasks. It highlights the integration techniques, learning frameworks, and optimization strategies employed by three distinct methods: CityLearn, TraKDis, and an approach utilizing CNNs for Q-value mapping. The comparison underscores the diverse strategies enhancing visual task performance in dynamic environments.

## 5 Applications and Case Studies

The exploration of reinforcement learning (RL) has significantly broadened its application scope across various domains, showcasing its versatility in addressing complex challenges. This section delves into specific applications of RL, starting with its transformative role in robotics.



AI-generated, for reference only.

### 5.1 Reinforcement Learning in Robotics

Reinforcement learning (RL) has become essential in developing intelligent robotic systems, enabling automation through adaptive decision-making. A key application of RL in robotics is robotic manipulation, where RL algorithms train robotic arms for intricate tasks like object grasping, utilizing visual feedback to handle objects of varying shapes and sizes in dynamic environments [9]. The integration of RL with computer vision has advanced autonomous navigation for mobile robots, as exemplified by the CityLearn framework, which employs RL to optimize navigation policies using visual place recognition, enhancing robots' efficiency in complex settings [7].

In unmanned aerial vehicles (UAVs), RL has been pivotal in developing control strategies for autonomous landing and obstacle avoidance, incorporating visual data to adapt to diverse environmental conditions. This adaptability is crucial for search and rescue missions, where UAVs navigate challenging terrains [68, 48, 13, 38, 70]. RL also enhances social navigation, allowing robots to learn appropriate behaviors in crowded environments, improving their utility in service-oriented tasks [19].

The application of RL in robotics continues to drive innovation, offering solutions to complex challenges across diverse domains. By leveraging RL advancements, robotic systems can execute complex tasks autonomously, crucial for deployment in commercial, civilian, and military applications. RL enhances the autonomy of unmanned aerial systems (UAS) and robotic platforms, facilitating improved safety and efficiency through real-time navigation and decision-making without human intervention [27, 71, 23, 72, 55].

### 5.2 Autonomous Vehicles and Navigation

Reinforcement learning (RL) has emerged as a transformative technology in autonomous vehicles and navigation systems, enhancing decision-making and control in complex driving environments. Integrating RL into autonomous driving systems enables vehicles to develop optimal driving strategies through environmental interactions, addressing complexities such as traffic rules, collision avoidance, and passenger comfort. Techniques like multi-objective deep reinforcement learning facilitate learning from diverse scenarios, improving data efficiency and enabling transfer learning [73, 1].

A significant challenge for autonomous vehicles is effectively operating in unexpected situations, where traditional rule-based systems struggle. The intersection of world models and anomaly detection is crucial in this context, enhancing the robustness of autonomous navigation systems [74]. RL's application in autonomous vehicles extends beyond anomaly detection, significantly enhancing motion planning, control, and decision-making. Recent advancements reveal two primary approaches: the traditional pipeline method, which utilizes hand-crafted modules for interpretability, and the end-to-end approach, which often outperforms despite challenges related to expert data and generalization [75, 48, 28, 74]. By learning from real-world interactions, RL optimizes these processes, reducing reliance on extensive labeled datasets and enabling adaptive driving policies.

Moreover, RL has advanced autonomous vehicles' capabilities in urban settings, where traffic scenarios demand sophisticated decision-making frameworks. Integrating visual data into the learning process allows RL agents to analyze and react to critical indicators, such as traffic signals and pedestrian crossings, enhancing decision-making for safe navigation [76, 13, 48, 38]. The integration of RL in autonomous vehicles and navigation systems significantly advances the field by addressing the intricate challenges of autonomous driving, paving the way for widespread adoption [73, 48, 28, 1].

### 5.3 Surveillance and Security

The integration of reinforcement learning (RL) into surveillance and security systems has enhanced monitoring and threat detection processes. Advanced RL algorithms enable these systems to autonomously adapt to changing environments, significantly improving their capabilities to identify and respond to security threats in real time. This adaptability is crucial for applications such as autonomous vehicles and unmanned aerial systems, where robust sensor data processing and response to cyber-physical attacks are essential for safety and operational efficiency [27, 75, 77, 33].

Reinforcement learning optimizes camera placement and coverage in surveillance systems, facilitating comprehensive monitoring while reducing blind spots. Techniques such as deep reinforcement learning and visual active search enable intelligent imaging system configuration, integrating components like illumination sources and sensors to enhance task-specific perception models [12, 13, 38].

Furthermore, RL techniques have improved threat detection algorithms, enabling systems to identify suspicious activities more accurately and swiftly through complex agent-environment interactions [51, 44, 37, 11, 38].

In security, RL has enabled the creation of autonomous patrol robots capable of independently navigating and monitoring extensive areas without human oversight. By leveraging sophisticated machine learning techniques, these robots enhance decision-making processes, adapting to complex environments and responding to potential threats [23, 27, 75]. The integration of RL into surveillance and security systems catalyzes significant advancements, providing innovative solutions to monitoring and threat detection challenges [51, 20, 29, 13, 77].

### 5.4 Gaming and Simulation Environments

Reinforcement learning (RL) has transformed gaming and simulation environments, serving as a testbed for developing advanced AI algorithms. Its application has led to significant advancements in AI agents' capabilities, enabling them to learn complex strategies and adapt to dynamic scenarios with minimal human intervention. A landmark achievement in this area is AlphaGo, which demonstrated RL's potential to surpass human expertise in strategic games by utilizing deep neural networks for game state evaluation and decision-making [29].

In gaming environments, RL agents optimize performance by balancing exploration and exploitation, discovering novel strategies, and enhancing decision-making through continuous interaction with the game environment. Simulation environments offer controlled settings for testing and evaluating RL algorithms, providing a safe and cost-effective experimentation platform [15, 32]. Moreover, integrating RL with simulation environments has facilitated developing adaptive systems that learn from past experiences, illustrated by RL applications in training autonomous agents to navigate intricate terrains and engage with dynamic elements within simulations [33, 23, 38, 59].

The application of RL in gaming and simulation environments continues to push AI research boundaries, offering fertile ground for exploring intelligent systems' capabilities. By leveraging RL, these environments enhance simulation realism and complexity while driving advancements in AI technologies across diverse fields [45, 33, 37, 25, 38].

### 5.5 Healthcare and Adaptive Systems

Reinforcement learning (RL) has shown promise in transforming healthcare and adaptive systems by providing innovative solutions to complex decision-making challenges. In healthcare, RL algorithms optimize treatment strategies by analyzing patient data and outcomes, personalizing care through tailored interventions, and enhancing diagnostic accuracy [44, 11, 77, 38]. By learning from extensive medical data, RL agents identify patterns and make informed decisions that improve patient outcomes, particularly valuable in managing chronic diseases requiring continuous treatment plan adjustments.

One notable application of RL in healthcare is developing adaptive treatment strategies tailored to individual patients. Leveraging RL allows healthcare providers to dynamically adjust treatment protocols to maximize therapeutic efficacy while minimizing adverse effects [11, 78, 20, 51]. In adaptive systems, RL enhances autonomous agents' performance in dynamic and uncertain environments, exemplified by the effectiveness of partially supervised reinforcement learning (PSVAS) in adapting search policies in real-world scenarios [13].

Moreover, integrating RL with adaptive systems has facilitated developing intelligent agents capable of learning from past experiences and improving over time, crucial for applications in robotics and autonomous vehicles [37, 33]. The application of RL in healthcare and adaptive systems continues to drive innovation, offering solutions to intricate challenges across various domains, enhancing personalized healthcare delivery and optimizing performance in complex, dynamic environments [44, 78, 51].

## 6 Challenges and Future Directions

### 6.1 Challenges in Data Efficiency and Generalization

Reinforcement learning (RL) faces significant obstacles in data efficiency and generalization, especially in complex and dynamic environments. The reliance on predefined goals and tabular representations poses a major challenge, as these methods are not scalable to high-dimensional state spaces, resulting in high sample complexity and extensive training requirements [60, 7]. The computational demand of parallel updates for each goal further limits scalability to small tabular cases [43]. Additionally, accurately modeling costs during policy switching, particularly in scenarios lacking online interaction, complicates the RL landscape [41]. Efficient exploration and generalization across slightly different contexts remain crucial, especially when precise state information is unavailable, as in cloth manipulation tasks [32, 9]. Asynchronous training inefficiencies can also lead to delays and reduced performance [10]. Innovative exploration strategies, like LMC-LSVI, which samples directly from the posterior distribution, have shown improved performance over traditional methods [42]. Advancing methodologies to optimize data utilization and exploration strategies is essential for enhancing data efficiency and generalization in RL, particularly as systems increasingly engage with complex environments. Recent developments in deep reinforcement learning (DRL) have demonstrated promise in learning adaptive strategies that surpass traditional algorithms. By utilizing frameworks like Dopamine and implementing robust exploration policies, researchers can better address challenges related to data collection, model learning, and deployment, ultimately enhancing RL applications across various domains, including healthcare and fintech [44, 29, 37].

### 6.2 Algorithmic Challenges and Model Complexity

Reinforcement learning (RL) is challenged by algorithmic complexities in model design and implementation. A primary issue with deep reinforcement learning (DRL) agents is high variance in training stability, complicating the search for optimal architectures through traditional neural architecture search methods [79]. This instability can hinder consistent performance across varying environments. In multi-agent scenarios, the computational complexity of nested conditioning presents significant challenges [80], making it difficult to scale RL algorithms effectively. The paradox-of-choice effect, where performance declines with an increasing number of options, can also impede decision-making efficiency [81]. Offline reinforcement learning methods struggle with generalization across diverse data compositions, often resulting in suboptimal performance [82]. This limitation emphasizes the need for algorithms that can leverage diverse datasets effectively. Furthermore, reliance on dimensionality reduction techniques can distort high-dimensional data visualizations, complicating the interpretation of model behaviors [38]. Challenges such as gradient overlapping and noise from entropy regularization significantly impact RL [83], necessitating the development of robust optimization techniques. In the context of the SRPMoE approach, overfitting in expert models can impair the decision-making capabilities of routers, complicating the algorithmic landscape [46]. Understanding which features are effectively learned by different algorithms and their influence on RL agent performance remains a significant challenge [84]. This highlights the need for transparent and interpretable models that can elucidate the decision-making processes of RL agents. Existing methods often struggle to guide agents in partially-observed environments, underscoring the complexity of the models required for such scenarios [85]. Addressing these algorithmic challenges is vital for enhancing RL system performance. By overcoming these obstacles, RL systems can better leverage diverse datasets and adapt to dynamic environments, improving decision-making capabilities across applications in healthcare, robotics, finance, and education. The evolution of RL algorithms from foundational methods like Q-learning to sophisticated approaches such as Proximal Policy Optimization (PPO) and offline reinforcement learning exemplifies this progress [40, 31, 15, 86, 77].

### 6.3 Adaptability and Transferability

The adaptability and transferability of reinforcement learning (RL) models are crucial for their effectiveness across diverse environments and tasks. A primary challenge in achieving adaptability is the complexity of RL algorithms, which often struggle to generalize learned behaviors to new contexts without extensive retraining [80]. This issue is compounded by high variance in training stability, which can hinder the development of robust models capable of adapting to varying conditions [79]. The transfer of knowledge between domains is also impeded by overfitting, particularly in offline reinforcement learning scenarios where models are trained on fixed datasets [82]. This can lead to suboptimal performance in environments differing from the training data, necessitating more flexible and generalizable learning strategies. Additionally, the paradox-of-choice effect, where numerous options can degrade decision-making performance, poses a challenge to the transferability of RL models [81]. This underscores the importance of developing algorithms that efficiently navigate complex decision spaces without cognitive overload. In multi-agent systems, the complexity of nested conditioning can hinder the transfer of learned behaviors, necessitating scalable and efficient knowledge transfer approaches [80]. The reliance on dimensionality reduction techniques can also distort high-dimensional data representations, complicating the transfer of learned features across tasks [38]. To address these challenges, there is a growing emphasis on developing RL algorithms that prioritize adaptability and transferability. Techniques such as hierarchical reinforcement learning and meta-learning present promising avenues for enhancing the generalization capabilities of RL models, enabling them to adapt more readily to new environments and tasks. By focusing on the diverse challenges inherent in deep reinforcement learning, researchers aim to create adaptable systems capable of navigating various real-world applications, from healthcare and finance to robotics and gaming. Frameworks like Dopamine facilitate the implementation of state-of-the-art RL agents, while tools like RLInspect enhance model assessment through interactive visual analytics, ultimately leading to more robust and reliable systems capable of continuous learning in dynamic environments [44, 38, 29, 37].

### 6.4 Exploration and Reward Mechanisms

Designing effective exploration and reward mechanisms in reinforcement learning (RL) is a significant challenge, particularly in environments with sparse extrinsic feedback. RL agents must balance exploration with the exploitation of known strategies to maximize cumulative rewards. In scenarios where extrinsic rewards are infrequent or absent, agents often rely on intrinsic motivation to guide their exploration, which can be difficult to model and implement effectively [63]. The Policy Gradient from Demonstration and Curiosity (PGfDC) method addresses these challenges by integrating a dual reward formulation, combining a demonstration term that evaluates the similarity between the agent's actions and expert demonstrations with a curiosity term derived from the agent's uncertainty about its environment [63]. This approach encourages exploration of novel actions while refining strategies based on expert knowledge, enhancing learning in sparse feedback environments. Moreover, reward mechanisms must mitigate the risk of suboptimal behavior stemming from poorly defined or misaligned rewards. This highlights the need for meticulously designed reward functions that encapsulate task objectives while minimizing unintended consequences. Recent advancements, such as STARC metrics for quantifying reward function differences and Differentiable Decision Trees for interpretable reward learning from human feedback, emphasize the importance of effective reward design [87, 88, 89, 51]. In multi-objective scenarios, balancing competing goals further complicates reward design, necessitating sophisticated methods to ensure appropriate prioritization. Addressing challenges in exploration and reward mechanisms is crucial for advancing RL capabilities. By implementing innovative strategies that promote effective exploration and align intrinsic rewards with desired outcomes, researchers can significantly enhance the adaptability and performance of RL agents in complex environments. This approach not only resolves the exploration-exploitation dilemma inherent in traditional RL methods but also leverages intrinsic motivation mechanisms, such as curiosity and novelty, to facilitate efficient exploration and learning. Techniques like intrinsic reward methods based on state entropy maximization and learning exploration bonuses from demonstrations can foster robust behaviors, enhancing performance across varied real-world applications, including education and autonomous systems [90, 91, 20, 72].

### 6.5 Future Research Directions

Future research in reinforcement learning (RL) for computer vision is set to explore several promising avenues to enhance the robustness, efficiency, and applicability of intelligent systems. Optimizing data efficiency and reducing model size are critical areas of focus, particularly in complex tasks like cloth manipulation, where RL agents can benefit from improved data utilization and streamlined models [9]. Additionally, applying RL techniques to more intricate manipulation tasks could advance the field by addressing challenges posed by high-dimensional state spaces and dynamic environments. The BA3C algorithm represents another opportunity for future research, with efforts aimed at optimizing its performance on larger CPU clusters and exploring its application on diverse hardware architectures

[10]. These enhancements could improve the scalability and efficiency of RL systems, enabling them to tackle more complex visual tasks. Extending the Langevin Monte Carlo Least-Squares Value Iteration (LMC-LSVI) method to more complex environments is another promising direction, focusing on improving the suboptimal dependence on the planning horizon in randomized algorithms [42]. Refining exploration strategies and optimizing planning processes can enhance the adaptability and performance of RL agents across diverse settings. The outlined research directions indicate a promising trajectory for integrating RL with computer vision, suggesting that advancements in deep reinforcement learning could lead to autonomous systems capable of sophisticated visual understanding and real-world interaction. These advancements include the ability to learn complex tasks directly from visual inputs, as seen in applications ranging from robotics to visual active search, and enhancing learning efficiency through innovative techniques like automatic curriculum learning and self-supervised visual correspondence. Collectively, these efforts underscore RL's potential to tackle previously intractable challenges in computer vision, paving the way for more effective and adaptable AI systems [47, 51, 35, 13, 34]. By pursuing these avenues, researchers aim to drive innovation across a wide range of applications, enhancing the robustness, efficiency, and adaptability of intelligent systems.

# 7 Conclusion

Reinforcement learning (RL) has profoundly influenced the field of computer vision, particularly in enhancing object detection and tracking capabilities. The integration of RL into visual systems has addressed many limitations of traditional methods, significantly improving adaptability and operational efficiency across domains like autonomous vehicles, robotics, and surveillance. This synergy has facilitated notable advancements in decision-making processes and system performance.

Deep reinforcement learning (DRL) plays a crucial role in managing the complexities of visual tasks, with deep neural networks (DNNs) being essential for processing high-dimensional data. These networks enhance visual perception systems by enabling effective generalization across diverse datasets, thereby bolstering the robustness of RL agents in dynamic environments. The ability of DNNs to handle complex visual inputs is critical to the development of sophisticated autonomous systems.

Nevertheless, challenges remain, particularly in data efficiency, generalization, and model complexity, which continue to impede the broader application of RL in computer vision. Addressing these challenges will require innovative strategies for exploration and optimization. Future research should focus on developing robust solutions to these issues, ensuring that RL can be more widely adopted and effectively utilized in complex visual environments.

## References

[1] Nayun Xu, Bowen Tan, and Bingyu Kong. Autonomous driving in reality with reinforcement learning and image translation, 2019.

[2] Gino Brunner, Manuel Fritsche, Oliver Richter, and Roger Wattenhofer. Using state predictions for value regularization in curiosity driven deep reinforcement learning, 2018.

[3] Enhanced method for reinforcement learning based dynamic obstacle avoidance by assessment of collision risk.

[4] Honghu Xue, Benedikt Hein, Mohamed Bakr, Georg Schildbach, Bengt Abel, and Elmar Rueckert. Using deep reinforcement learning with automatic curriculum learning for mapless navigation in intralogistics, 2022.

[5] Benjamin Recht. A tour of reinforcement learning: The view from continuous control. *Annual Review of Control, Robotics, and Autonomous Systems*, 2(1):253–279, 2019.

[6] Where to go: Agent guidance with deep reinforcement learning in a city-scale online ride-hailing service.

[7] Marvin Chancán and Michael Milford. Citylearn: Diverse real-world environments for sample-efficient navigation policy learning, 2020.

[8] Finn Rietz, Erik Schaffernicht, Stefan Heinrich, and Johannes Andreas Stork. Prioritized soft q-decomposition for lexicographic reinforcement learning, 2024.

[9] Wei Chen and Nicolas Rojas. Trakdis: A transformer-based knowledge distillation approach for visual reinforcement learning with application to cloth manipulation, 2024.

[10] Robert Adamski, Tomasz Grel, Maciej Klimek, and Henryk Michalewski. Atari games and intel processors, 2017.

[11] Julian Skirzyński, Frederic Becker, and Falk Lieder. Automatic discovery of interpretable planning strategies, 2021.

[12] Tzofi Klinghoffer, Kushagra Tiwary, Nikhil Behari, Bhavya Agrawalla, and Ramesh Raskar. Diser: Designing imaging systems with reinforcement learning, 2023.

[13] Anindya Sarkar, Nathan Jacobs, and Yevgeniy Vorobeychik. A partially supervised reinforcement learning framework for visual active search, 2023.

[14] Jessica Borja-Diaz, Oier Mees, Gabriel Kalweit, Lukas Hermann, Joschka Boedecker, and Wolfram Burgard. Affordance learning from play for sample-efficient policy learning, 2022.

[15] Shengbo Eben Li. Deep reinforcement learning. In *Reinforcement learning for sequential decision and optimal control*, pages 365–402. Springer, 2023.

[16] Benchmarking model-based reinfor.

[17] Peng Zhang, Yawen Huang, Bingzhang Hu, Shizheng Wang, Haoran Duan, Noura Al Moubayed, Yefeng Zheng, and Yang Long. Knowing the past to predict the future: Reinforcement virtual learning, 2022.

[18] Thomas Lampe, Abbas Abdolmaleki, Sarah Bechtle, Sandy H. Huang, Jost Tobias Springenberg, Michael Bloesch, Oliver Groth, Roland Hafner, Tim Hertweck, Michael Neunert, Markus Wulfmeier, Jingwei Zhang, Francesco Nori, Nicolas Heess, and Martin Riedmiller. Mastering stacking of diverse shapes with large-scale iterative reinforcement learning on real robots, 2023.

[19] Janderson Ferreira, Agostinho A. F. Júnior, Letícia Castro, Yves M. Galvão, Pablo Barros, and Bruno J. T. Fernandes. Analysis of social robotic navigation approaches: Cnn encoder and incremental learning as an alternative to deep reinforcement learning, 2020.

[20] Herkulaas Combrink, Vukosi Marivate, and Benjamin Rosman. Reinforcement learning in education: A multi-armed bandit approach, 2022.

[21] Vahid Behzadan, Roman V. Yampolskiy, and Arslan Munir. Emergence of addictive behaviors in reinforcement learning agents, 2018.

[22] Sohrob Kazerounian, Matthew Luciw, Mathis Richter, and Yulia Sandamirskaya. Autonomous reinforcement of behavioral sequences in neural dynamics, 2013.

[23] Meraj Mammadov. End-to-end lidar-driven reinforcement learning for autonomous racing, 2023.

[24] Mudit Verma and Katherine Metcalf. Hindsight priors for reward learning from human preferences, 2024.

[25] Bernhard Jaeger and Andreas Geiger. An invitation to deep reinforcement learning, 2024.

[26] Ismot Sadik Peyas, Zahid Hasan, Md. Rafat Rahman Tushar, Al Musabbir, Raisa Mehjabin Azni, and Shahnewaz Siddique. Autonomous warehouse robot using deep q-learning, 2022.

[27] Jithin Jagannath, Anu Jagannath, Sean Furman, and Tyler Gwin. Deep learning and reinforcement learning for autonomous unmanned aerial systems: Roadmap for theory to deployment, 2020.

[28] Fei Ye, Shen Zhang, Pin Wang, and Ching-Yao Chan. A survey of deep reinforcement learning algorithms for motion planning and control of autonomous vehicles, 2021.

[29] Yuxi Li. Reinforcement learning applications. *arXiv preprint arXiv:1908.06973*, 2019.

[30] Yuexiang Zhai, Hao Bai, Zipeng Lin, Jiayi Pan, Shengbang Tong, Yifei Zhou, Alane Suhr, Saining Xie, Yann LeCun, Yi Ma, and Sergey Levine. Fine-tuning large vision-language models as decision-making agents via reinforcement learning, 2024.

[31] Yuxi Li. Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*, 2017.

[32] Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G Bellemare, Joelle Pineau, et al. An introduction to deep reinforcement learning. *Foundations and Trends® in Machine Learning*, 11(3-4):219–354, 2018.

[33] Yunpeng Qing, Shunyu Liu, Jie Song, Huiqiong Wang, and Mingli Song. A survey on explainable reinforcement learning: Concepts, algorithms, challenges, 2023.

[34] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. A brief survey of deep reinforcement learning. *arXiv preprint arXiv:1708.05866*, 2017.

[35] Kai Arulkumaran, Marc Peter Deisenroth, Miles Brundage, and Anil Anthony Bharath. Deep reinforcement learning: A brief survey. *IEEE Signal Processing Magazine*, 34(6):26–38, 2017.

[36] Laura Smith, Nikita Dhawan, Marvin Zhang, Pieter Abbeel, and Sergey Levine. Avid: Learning multi-stage tasks via pixel-level translation of human videos, 2020.

[37] Pablo Samuel Castro, Subhodeep Moitra, Carles Gelada, Saurabh Kumar, and Marc G. Bellemare. Dopamine: A research framework for deep reinforcement learning, 2018.

[38] Geetansh Kalra, Divye Singh, and Justin Jose. Rlinspect: An interactive visual approach to assess reinforcement learning algorithm, 2024.

[39] Hao-nan Wang, Ning Liu, Yi-yun Zhang, Da-wei Feng, Feng Huang, Dong-sheng Li, and Yi-ming Zhang. Deep reinforcement learning: a survey. *Frontiers of Information Technology & Electronic Engineering*, 21(12):1726–1744, 2020.

[40] Seyed Sajad Mousavi, Michael Schukat, and Enda Howley. Deep reinforcement learning: an overview. In *Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016: Volume 2*, pages 426–440. Springer, 2018.

[41] Tao Ma, Xuzhi Yang, and Zoltan Szabo. To switch or not to switch? balanced policy switching in offline reinforcement learning, 2025.

[42] Haque Ishfaq, Qingfeng Lan, Pan Xu, A. Rupam Mahmood, Doina Precup, Anima Anandkumar, and Kamyar Azizzadenesheli. Provable and practical: Efficient exploration in reinforcement learning via langevin monte carlo, 2024.

[43] Fabio Pardo, Vitaly Levdik, and Petar Kormushev. Scaling all-goals updates in reinforcement learning using convolutional neural networks, 2020.

[44] Qingpeng Cai, Can Cui, Yiyuan Xiong, Wei Wang, Zhongle Xie, and Meihui Zhang. A survey on deep reinforcement learning for data processing and analytics, 2022.

[45] Chao Yu, Shicheng Ye, and Hankz Hankui Zhuo. Reinforcement learning with knowledge representation and reasoning: A brief survey, 2025.

[46] Hamid Mohammadi, Ehsan Nazerfard, and Tahereh Firoozi. Reinforcement learning-based mixture of vision transformers for video violence recognition, 2023.

[47] Lucas Manuelli, Yunzhu Li, Pete Florence, and Russ Tedrake. Keypoints into the future: Self-supervised correspondence in model-based reinforcement learning, 2020.

[48] Shahin Atakishiyev, Mohammad Salameh, Housam Babiker, and Randy Goebel. Explaining autonomous driving actions with visual question answering, 2023.

[49] Vishnu Saj, Bochan Lee, Dileep Kalathil, and Moble Benedict. Robust reinforcement learning algorithm for vision-based ship landing of uavs, 2022.

[50] Nestor Gonzalez Lopez, Yue Leire Erro Nuin, Elias Barba Moral, Lander Usategui San Juan, Alejandro Solano Rueda, Víctor Mayoral Vilches, and Risto Kojcev. gym-gazebo2, a toolkit for reinforcement learning using ros 2 and gazebo, 2019.

[51] Ryan Campbell and Junsang Yoon. Automatic curriculum learning with gradient reward signals, 2023.

[52] Lorenzo Mauro, Francesco Puja, Simone Grazioso, Valsamis Ntouskos, Marta Sanzari, Edoardo Alati, and Fiora Pirri. Visual search and recognition for robot task execution and monitoring, 2019.

[53] Arantxa Casanova, Pedro O. Pinheiro, Negar Rostamzadeh, and Christopher J. Pal. Reinforced active learning for image segmentation, 2020.

[54] Andrew Melnik, Sascha Fleer, Malte Schilling, and Helge Ritter. Modularization of end-to-end learning: Case study in arcade games, 2019.

[55] Márton Szép, Leander Lauenburg, Kevin Farkas, Xiyan Su, and Chuanlong Zang. Reinforcement learning for solving robotic reaching tasks in the neurorobotics platform, 2022.

[56] Chengpeng Li, Zhengyi Yang, Jizhi Zhang, Jiancan Wu, Dingxian Wang, Xiangnan He, and Xiang Wang. Model-enhanced contrastive reinforcement learning for sequential recommendation, 2023.

[57] Martha White. Unifying task specification in reinforcement learning, 2021.

[58] Haobin Jiang and Zongqing Lu. Visual grounding for object-level generalization in reinforcement learning, 2024.

[59] Mikuláš Zelinka. Using reinforcement learning to learn how to play text-based games, 2018.

[60] Vivek Veeriah, Junhyuk Oh, and Satinder Singh. Many-goals reinforcement learning, 2018.

[61] Taisei Hashimoto and Yoshimasa Tsuruoka. Utilizing skipped frames in action repeats via pseudo-actions, 2021.

[62] Johan Ferret, Raphaël Marinier, Matthieu Geist, and Olivier Pietquin. Self-attentional credit assignment for transfer in reinforcement learning, 2019.

[63] Jie Chen and Wenjun Xu. Policy gradient from demonstration and curiosity, 2020.

[64] Ludovic Denoyer and Patrick Gallinari. Deep sequential neural network, 2014.

[65] Sean Xie, Soroush Vosoughi, and Saeed Hassanpour. Towards interpretable deep reinforcement learning models via inverse reinforcement learning, 2024.

[66] Eric Crawford, Guillaume Rabusseau, and Joelle Pineau. Sequential coordination of deep models for learning visual arithmetic, 2018.

[67] Arun Nair, Praveen Srinivasan, Sam Blackwell, Cagdas Alcicek, Rory Fearon, Alessandro De Maria, Vedavyas Panneershelvam, Mustafa Suleyman, Charles Beattie, Stig Petersen, Shane Legg, Volodymyr Mnih, Koray Kavukcuoglu, and David Silver. Massively parallel methods for deep reinforcement learning, 2015.

[68] Misha Laskin, Kimin Lee, Adam Stooke, Lerrel Pinto, Pieter Abbeel, and Aravind Srinivas. Reinforcement learning with augmented data. *Advances in neural information processing systems*, 33:19884–19895, 2020.

[69] Sina Montazeri, Akram Mirzaeinia, and Amir Mirzaeinia. Cnn-drl with shuffled features in finance, 2024.

[70] Xiaobo Hu, Youfang Lin, Yue Liu, Jinwen Wang, Shuo Wang, Hehe Fan, and Kai Lv. A reliable representation with bidirectional transition model for visual reinforcement learning generalization, 2023.

[71] Henry Zhu, Justin Yu, Abhishek Gupta, Dhruv Shah, Kristian Hartikainen, Avi Singh, Vikash Kumar, and Sergey Levine. The ingredients of real-world robotic reinforcement learning, 2020.

[72] Mingqi Yuan. Intrinsically-motivated reinforcement learning: A brief introduction, 2022.

[73] Changjian Li and Krzysztof Czarnecki. Urban driving with multi-objective deep reinforcement learning. *arXiv preprint arXiv:1811.08586*, 2018.

[74] Daniel Bogdoll, Lukas Bosch, Tim Joseph, Helen Gremmelmaier, Yitian Yang, and J. Marius Zöllner. Exploring the potential of world models for anomaly detection in autonomous driving, 2023.

[75] Aidin Ferdowsi, Ursula Challita, Walid Saad, and Narayan B. Mandayam. Robust deep reinforcement learning for security and safety in autonomous vehicle systems, 2018.

[76] Maxime Bouton, Alireza Nakhaei, Kikuo Fujimura, and Mykel J. Kochenderfer. Safe reinforcement learning with scene decomposition for navigating complex urban environments, 2019.

[77] Mohamed-Amine Chadi and Hajar Mousannif. Understanding reinforcement learning algorithms: The progress from basic q-learning to proximal policy optimization, 2023.

[78] Charl Maree and Christian Omlin. Can interpretable reinforcement learning manage prosperity your way?, 2022.

[79] Yonggan Fu, Zhongzhi Yu, Yongan Zhang, and Yingyan Lin. Auto-agent-distiller: Towards efficient deep reinforcement learning agents via neural architecture search, 2020.

[80] David Tolpin and Tomer Dobkin. Bob and alice go to a bar: Reasoning about future with probabilistic programs, 2021.

[81] Eric Chalmers and Artur Luczak. Reinforcement learning with brain-inspired modulation can improve adaptation to environmental changes, 2022.

[82] Paria Rashidinejad, Banghua Zhu, Cong Ma, Jiantao Jiao, and Stuart Russell. Bridging offline reinforcement learning and imitation learning: A tale of pessimism, 2023.

[83] Federico A. Galatolo, Mario G. C. A. Cimino, and Gigliola Vaglini. Solving the scalarization issues of advantage-based reinforcement learning algorithms, 2021.

[84] Charline Le Lan, Stephen Tu, Mark Rowland, Anna Harutyunyan, Rishabh Agarwal, Marc G. Bellemare, and Will Dabney. Bootstrapped representations in reinforcement learning, 2023.

[85] Nicholas Rhinehart, Jenny Wang, Glen Berseth, John D. Co-Reyes, Danijar Hafner, Chelsea Finn, and Sergey Levine. Information is power: Intrinsic control via information capture, 2021.

[86] Sergey Levine, Aviral Kumar, George Tucker, and Justin Fu. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. *arXiv preprint arXiv:2005.01643*, 2020.

[87] Joar Skalse, Lucy Farnik, Sumeet Ramesh Motwani, Erik Jenner, Adam Gleave, and Alessandro Abate. Starc: A general framework for quantifying differences between reward functions, 2024.

[88] Akansha Kalra and Daniel S. Brown. Can differentiable decision trees enable interpretable reward learning from human feedback?, 2024.

[89] Pierre Sermanet, Kelvin Xu, and Sergey Levine. Unsupervised perceptual rewards for imitation learning, 2017.

[90] Yang Yu. Towards sample efficient reinforcement learning. In *IJCAI*, pages 5739–5743, 2018.

[91] Léonard Hussenot, Robert Dadashi, Matthieu Geist, and Olivier Pietquin. Show me the way: Intrinsic motivation from demonstrations, 2021.

**Disclaimer:**

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.

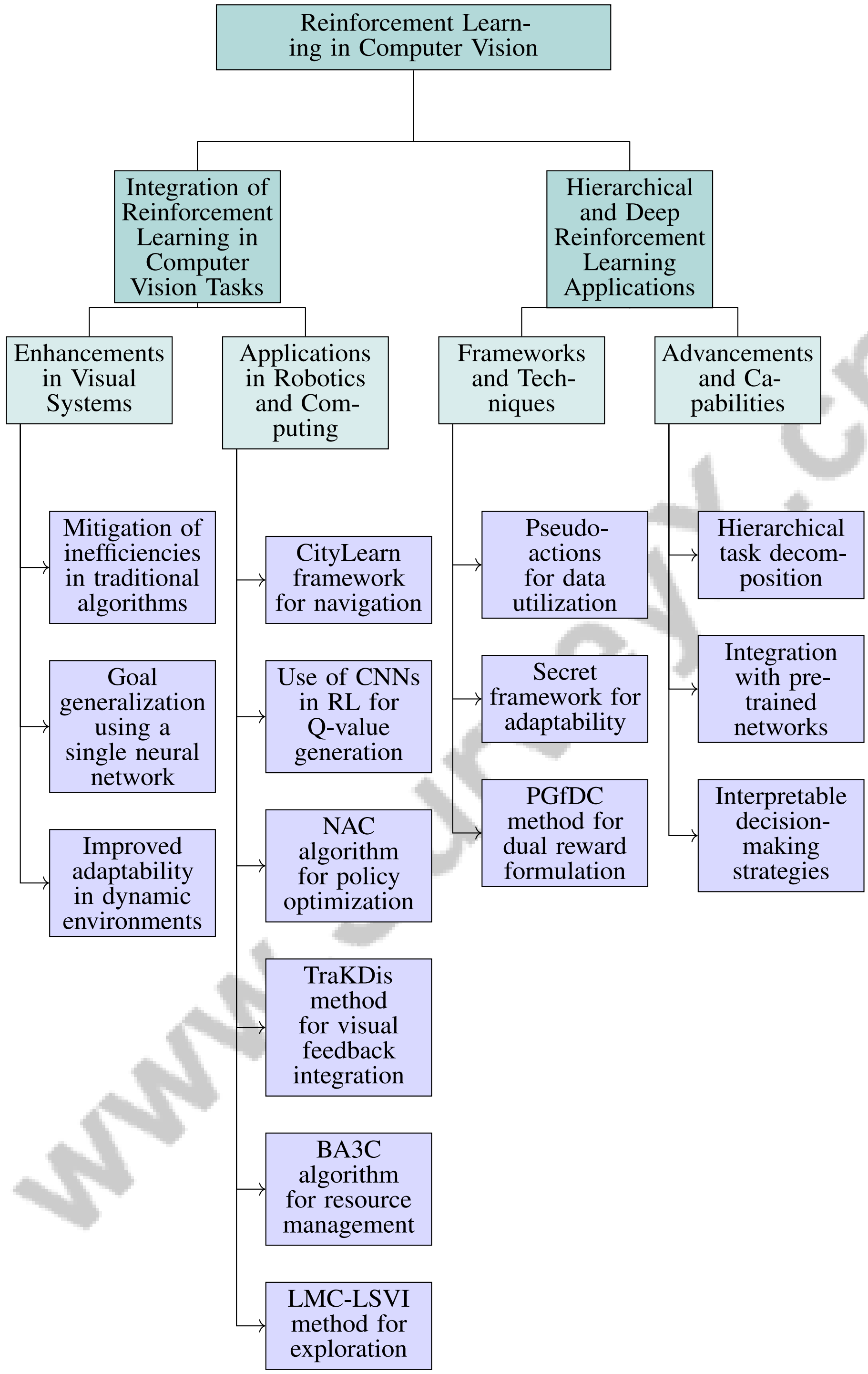
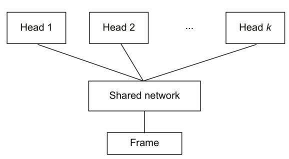
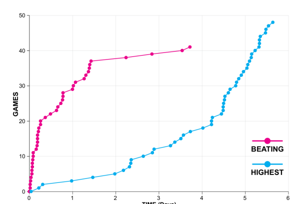


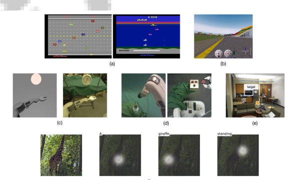

Figure 2: This figure illustrates the integration of reinforcement learning (RL) in computer vision, highlighting key advancements and applications. The first section focuses on the integration of RL in computer vision tasks, showcasing enhancements in visual systems and specific applications in robotics and computing. The second section delves into hierarchical and deep reinforcement learning applications, detailing frameworks and techniques that enhance system capabilities and advancements in managing complex visual tasks. Together, these elements underscore the transformative potential of RL in advancing computer vision technologies.




(a) A diagram illustrating a multi-head attention mechanism[39]




(b) Games Played Over Time[67]



(c) Gameplay and Real-World Applications of Deep Learning[34]

Figure 3: Examples of Hierarchical and Deep Reinforcement Learning Applications