

مروری بر کاربرد یادگیری تقویتی در رباتیک: پیشرفت ها ، چالش ها و افق های آینده

مجتبی عبدالمهی

گروه مهندسی کامپیوتر ، واحد تربت جام ، دانشگاه آزاد اسلامی، تربت جام ، ایران

حجت آزادروش

گروه مهندسی کامپیوتر ، واحد تربت جام ، دانشگاه آزاد اسلامی ، تربت جام ، ایران

چکیده

یادگیری تقویتی به عنوان یکی از پیشرفته ترین شاخه های هوش مصنوعی، نقش مهمی در ارتقاء عملکرد و هوشمندی سیستم های رباتیک ایفا کرده است. این الگوریتم ها با امکان تعامل مستمر ربات ها با محیط های پیچیده، به آنها اجازه می دهند تا از طریق تجربه مستقیم، تصمیم گیری های بهینه و خودکفا را بیاموزند. کاربردهای یادگیری تقویتی در رباتیک شامل ناوبری خودکار، دستکاری اشیاء، تعامل انسان با ربات و مدیریت سیستم های چندعامله می باشد که به طور قابل توجهی کارایی و انعطاف پذیری ربات ها را افزایش داده اند. با وجود این پیشرفت ها، چالش هایی همچون ناکارآمدی نمونه گیری، نیاز به منابع محاسباتی بالا، مسائل ایمنی و قابلیت اطمینان و مقیاس پذیری الگوریتم ها همچنان مانعی بر سر راه بهره برداری کامل از یادگیری تقویتی در رباتیک محسوب می شوند. یادگیری تقویتی در رباتیک به ویژه در زمینه های پیچیده ای همچون تعامل با محیط های دینامیک و غیرقابل پیش بینی، به ربات ها این امکان را می دهد که رفتارهای خود را به طور مداوم بهبود بخشند. این الگوریتم ها نه تنها به بهینه سازی فرآیندهای یادگیری در محیط های متغیر کمک کرده، بلکه باعث افزایش قابلیت تطبیق پذیری و یادگیری سریع تر ربات ها در مواجهه با موقعیت های جدید می شوند. این مقاله به بررسی کاربردهای یادگیری تقویتی در رباتیک، شناسایی چالش ها و محدودیت های موجود، ارزیابی مقایسه ای الگوریتم های مختلف یادگیری تقویتی و ارائه افق های پژوهشی آینده پرداخته است.

واژگان کلیدی: یادگیری تقویتی ، رباتیک ، هوش مصنوعی ، یادگیری ماشین.



مقدمه

تجزیه و تحلیل کاربردهای رباتیک نشان می‌دهد که تکنولوژی‌های نوین در این حوزه با موانع فنی و روانی مواجه هستند. تغییرات هوش مصنوعی در رباتیک که پیش‌تر چالش‌برانگیز بود، در سال‌های اخیر پیشرفت‌های چشمگیری داشته است. افزایش توسعه صنعتی و تقاضا برای کاربردهای خانگی، نشان‌دهنده رشد روزافزون علاقه به رباتیک است که با ظهور الگوریتم‌های نوین و افزایش مقالات پژوهشی تایید می‌شود. از جمله کاربردهای برجسته رباتیک می‌توان به انجام عملیات پزشکی، مونتاژ در خطوط تولید و کمک به بیماران و معلولین اشاره کرد (Al-Hamadani et al., ۲۰۲۴).

یادگیری تقویتی عمیق (Deep Reinforcement Learning - DRL) با ادغام ویژگی‌های یادگیری عمیق و تصمیم‌گیری بهینه، توانسته است کنترل‌های پیشرفته‌ای را ارائه دهد. در دهه گذشته، DRL در بسیاری از کاربردهایی که مستلزم درک ورودی با ابعاد بالا و تصمیم‌گیری بهینه یا نزدیک به بهینه هستند، پیشرفت‌های قابل توجهی داشته است. با این حال، هنوز چالش‌هایی مانند یادگیری با نمونه‌های محدود، پاداش‌های پراکنده و عوامل متعدد وجود دارد که محققان راه‌حل‌ها و نظریه‌های جدیدی برای غلبه بر این مشکلات ارائه کرده‌اند (Tang et al., ۲۰۲۴).

علاوه بر DRL، یادگیری تقویتی سلسله‌مراتبی (Hierarchical Reinforcement Learning - HRL)، یادگیری تقویتی چندعامله و یادگیری تقلیدی نیز از زیرشاخه‌های مهم یادگیری تقویتی به شمار می‌آیند که هر کدام به نوبه خود نقش مهمی در توسعه سیستم‌های رباتیک ایفا می‌کنند (Ramezani & Amir Atashgah, ۲۰۲۴). یادگیری تقویتی که از روان‌شناسی رفتارگرایی الهام گرفته است، بر رفتارهایی تمرکز دارد که ماشین باید برای بهینه‌سازی پاداش انجام دهد. این رویکرد در زمینه‌های متنوعی مانند نظریه بازی‌ها، نظریه کنترل، تحقیق در عملیات و هوش‌زدحامی به کار گرفته می‌شود (Sekkat et al., ۲۰۲۴).

در مقایسه با روش‌های سنتی، یادگیری تقویتی نیازی به داشتن اطلاعات کامل درباره فرآیند تصمیم‌گیری ندارد و می‌تواند بر روی فرآیندهای مارکوفی بزرگ‌تری کار کند که روش‌های سنتی در آن ناکارآمد هستند. همچنین، بر خلاف یادگیری با نظارت که به زوج‌های صحیح ورودی و خروجی نیاز دارد، یادگیری تقویتی بر کارایی زنده تمرکز دارد و نیازمند تعادل مناسبی بین اکتشاف و بهره‌برداری است (Dong et al., ۲۰۲۴).

هدف یادگیری تقویتی به عنوان زیرشاخه‌ای از یادگیری ماشین، بهینه‌سازی تصمیم‌گیری با استفاده از نمونه‌های تعامل یک عامل با محیط و بازخوردهای تاخیری است. برخلاف یادگیری نظارت شده سنتی که معمولاً بر سیگنال‌های پاداش یکباره و نظارت شده تکیه دارد، RL مشکلات تصمیم‌گیری متوالی را با بازخوردهای نمونه‌گیری، ارزیابی و تأخیر به طور همزمان حل می‌کند. این ویژگی، تکنیک‌های RL را کاندیدای مناسبی برای توسعه راه‌حل‌های قدرتمند در حوزه‌های مختلف مراقبت‌های بهداشتی می‌کند، جایی که تصمیم‌های تشخیص یا رژیم‌های درمانی معمولاً با دوره‌های طولانی و بازخوردهای تاخیری مشخص می‌شوند (Nguyen et al., ۲۰۲۴).

در این مقاله، به بررسی جامع کاربردهای یادگیری تقویتی در رباتیک، شناسایی چالش‌ها و محدودیت‌های موجود و ارائه افق‌های آینده این حوزه پرداخته می‌شود. موارد اصلی مورد بررسی عبارتند از:

- معرفی مبانی و اصول یادگیری تقویتی
- تحلیل کاربردهای مختلف یادگیری تقویتی در رباتیک
- مقایسه الگوریتم‌های مختلف یادگیری تقویتی بر اساس معیارهای کلیدی
- شناسایی و تحلیل چالش‌ها و محدودیت‌های یادگیری تقویتی در رباتیک
- ارائه افق‌های آینده پژوهشی برای یادگیری تقویتی در رباتیک

ساختار این مقاله به شرح زیر است: در بخش ۲، پیشینه تحقیق بررسی می‌شود. بخش ۳ به معرفی ساختار و عملکرد یادگیری تقویتی اختصاص دارد. در بخش ۴، کاربردهای مختلف یادگیری تقویتی در رباتیک مورد تحلیل قرار می‌گیرند. بخش ۵ به بررسی چالش‌های

موجود در استفاده از یادگیری تقویتی در رباتیک می‌پردازد. در بخش ۶، الگوریتم‌های مختلف یادگیری تقویتی در رباتیک به صورت مقایسه‌ای ارزیابی می‌شوند. نهایتاً، در بخش ۷، افق‌های آینده یادگیری تقویتی در رباتیک مورد بحث قرار می‌گیرند.

پیشینه تحقیق

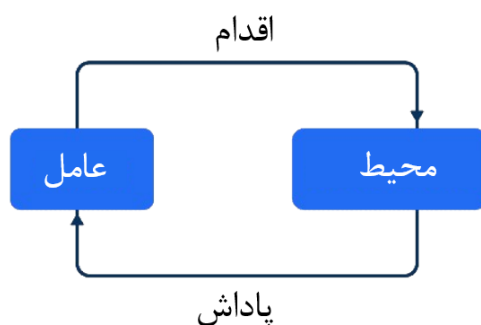
یادگیری تقویتی به‌عنوان یکی از شاخه‌های مهم یادگیری ماشین، در دهه‌های اخیر توجه گسترده‌ای را در حوزه رباتیک به خود جلب کرده است. مطالعات اولیه در این زمینه بیشتر بر توسعه الگوریتم‌های ساده RL برای کنترل ربات‌های پایه‌ای متمرکز بودند (Kaelbling et al., ۱۹۹۶). با پیشرفت‌های فناوری و افزایش قدرت محاسباتی، الگوریتم‌های پیچیده‌تری مانند یادگیری تقویتی عمیق معرفی شدند که امکان یادگیری از داده‌های بزرگ و پیچیده را فراهم کردند (Mnih et al., ۲۰۱۵). این پیشرفت‌ها منجر به توانمندی‌های پیشرفته‌تر ربات‌ها در انجام وظایف پیچیده‌ای مانند مانورهای پیچیده در محیط‌های نامطمئن و انجام وظایف چند مرحله‌ای شده‌اند (Levine et al., ۲۰۱۶).

کاربردهای یادگیری تقویتی در رباتیک گسترده و متنوع است. به عنوان مثال، ربات‌های خودران از یادگیری تقویتی برای بهبود مسیریابی و اجتناب از موانع در محیط‌های پیچیده استفاده می‌کنند (Wang et al., ۲۰۲۴). همچنین، در ربات‌های خدمت‌رسان، یادگیری تقویتی به منظور بهبود تعامل با انسان‌ها و ارائه خدمات بهینه به کار می‌رود (Chen et al., ۲۰۲۴). ترکیب RL با تکنیک‌های یادگیری عمیق منجر به پیشرفت‌های قابل توجهی در توانمندی‌های ربات‌ها شده است که امکان همکاری و تعامل بین چندین ربات را در سیستم‌های چندعامله فراهم کرده است (Mehta et al., ۲۰۲۴).

با وجود این پیشرفت‌ها، یادگیری تقویتی در رباتیک با چالش‌های متعددی مواجه است. یکی از مهم‌ترین این چالش‌ها نیاز به تعداد زیادی نمونه آموزشی برای آموزش مدل‌های یادگیری تقویتی است که در محیط‌های واقعی رباتیک پرهزینه و دشوار می‌باشد (Abouelyazid, ۲۰۲۴). علاوه بر این، مسائل مربوط به ایمنی و اطمینان از عملکرد صحیح ربات‌ها در شرایط غیرقابل پیش‌بینی از دیگر چالش‌های اساسی محسوب می‌شوند. برای غلبه بر این مشکلات، محققان به توسعه الگوریتم‌های یادگیری تقویتی ایمن و بهینه‌سازی فرآیندهای یادگیری با کاهش نیاز به داده‌های آموزشی پرداخته‌اند (He et al., ۲۰۲۴).

ساختار و عملکرد یادگیری تقویتی

همانطور که در شکل (۱) مشاهده می‌شود، یادگیری تقویتی را می‌توان به‌عنوان یک حلقه متشکل از اجزای اصلی زیر در نظر گرفت: عامل، محیط، اقدام، پاداش.



شکل (۱): اجزای اصلی یادگیری تقویتی

- عامل: تصمیم‌گیرنده‌ای که بر اساس مشاهدات خود اقداماتی را انجام می‌دهد.
- محیط: سیستم خارجی که عامل در آن عمل می‌کند.



- اقدام : تصمیم یا انتخابی که عامل در پاسخ به یک حالت اتخاذ می کند.
 - پاداش: سیگنال بازخوردی که کیفیت یا مطلوبیت عمل عامل را ارزیابی می کند.
- عامل با استفاده از استراتژی (Policy) یا رویکردی که برای انتخاب اقدامات بر اساس مشاهدات به کار می گیرد، عمل می کند. پس از انجام هر اقدام، عامل بازخوردی در قالب پاداش یا جریمه دریافت می کند. هدف RL یادگیری یک خطمشی بهینه است که پاداش های تجمعی را در طول زمان به حداکثر می رساند. عامل از طریق فرآیند آزمون و خطا محیط را کاوش می کند، اقداماتی را بر اساس وضعیت فعلی انجام می دهد و بازخورد دریافت می کند. این بازخورد برای به روزرسانی خطمشی و اتخاذ تصمیم های بهتر در آینده استفاده می شود (Milani et al., ۲۰۲۴).
- الگوریتم های RL اغلب از توابع ارزش (Value Functions) برای تخمین پاداش ها یا مقادیر مورد انتظار مرتبط با حالات و اقدامات مختلف بهره می برند. این توابع به عامل امکان می دهند تا توانایی های تصمیم گیری خود را یاد بگیرد و بهبود بخشد. با هر تعامل و یادگیری مکرر، خطمشی عامل به تدریج به سمت یک راه حل بهینه همگرا می شود و به رفتار هوشمندانه و سازگار در محیط های پیچیده و پویا می انجامد (Ernst & Louette, ۲۰۲۴).

کاربردهای یادگیری تقویتی در رباتیک

یادگیری تقویتی در سال های اخیر به عنوان یکی از روش های پیشرفته در حوزه رباتیک شناخته شده است. این الگوریتم ها با امکان یادگیری از طریق تعامل مداوم با محیط، توانسته اند ربات ها را قادر سازند تا وظایف پیچیده را به صورت خودکار و بهینه انجام دهند. در ادامه به برخی از مهم ترین کاربردهای یادگیری تقویتی در رباتیک پرداخته می شود:

- **ناوبری خودکار و مسیریابی**
یادگیری تقویتی به ربات ها امکان می دهد تا مسیرهای بهینه را در محیط های پیچیده و پویا شناسایی و طی کنند. با استفاده از RL، ربات ها می توانند موانع را شناسایی کرده و از طریق تجربه های گذشته، تصمیمات بهتری برای حرکت در محیط اتخاذ نمایند. این کاربرد به ویژه در ربات های خدماتی و صنعتی که نیاز به حرکت در فضاهای نامنظم دارند، بسیار موثر است (Zhu & Zhang, ۲۰۲۱).
- **دستکاری و تعامل با اشیاء**
در حوزه دستکاری و تعامل با اشیاء، یادگیری تقویتی به ربات ها اجازه می دهد تا عملیات پیچیده ای مانند گرفتن، حمل و نقل و چرخش اشیاء را با دقت و کارایی بالا انجام دهند. این الگوریتم ها با یادگیری از تجربیات گذشته، می توانند مهارت های دستکاری خود را بهبود بخشند و در محیط های متغیر عملکرد بهتری از خود نشان دهند (Elguea-Aguinaco et al., ۲۰۲۳).
- **تعامل انسان با ربات**
یادگیری تقویتی نقش مهمی در بهبود تعامل بین انسان و ربات ایفا می کند. با استفاده از یادگیری تقویتی، ربات ها می توانند رفتارهای انسان محور را یاد گرفته و به صورت طبیعی تری با کاربران تعامل داشته باشند. این کاربرد در ربات های خدمت رسان و ربات های همیار برای افراد دارای معلولیت به کار می رود (Xie et al., ۲۰۲۴).

چالش های یادگیری تقویتی در رباتیک

اگرچه یادگیری تقویتی به عنوان یکی از روش های پیشرفته و موثر در بهبود عملکرد ربات ها شناخته شده است، اما همچنان با چالش ها و محدودیت های متعددی مواجه است که مانع از بهره برداری کامل از پتانسیل های آن در حوزه رباتیک می شود. در ادامه به بررسی مهم ترین چالش ها و محدودیت های موجود پرداخته می شود:

- **ایمنی و قابلیت اطمینان**
استفاده از یادگیری تقویتی در محیط های واقعی رباتیک مستلزم تضمین ایمنی و قابلیت اطمینان سیستم ها است. الگوریتم های یادگیری تقویتی معمولاً از طریق فرآیند آزمون و خطا یاد می گیرند که در برخی موارد می تواند منجر به اعمال خطرناک یا غیرمنتظره شود.



بنابراین، توسعه الگوریتم‌های یادگیری تقویتی ایمن که بتوانند در شرایط غیرقابل پیش‌بینی به صورت مطمئن عمل کنند، از اهمیت بالایی برخوردار است (Retzlaff et al., ۲۰۲۴).

• انتقال یادگیری

یکی دیگر از چالش‌های اساسی یادگیری تقویتی در رباتیک، انتقال دانش از یک وظیفه به وظیفه دیگر است. در بسیاری از موارد، مدل‌های یادگیری تقویتی که برای یک وظیفه خاص آموزش دیده‌اند، نمی‌توانند به راحتی به وظایف جدید و متفاوت اعمال شوند. تحقیق در زمینه یادگیری انتقالی (Transfer Learning) و یادگیری چند وظیفه‌ای (Multi-task Learning) می‌تواند به بهبود قابلیت‌های عمومی مدل‌های یادگیری تقویتی کمک کند (Souza et al., ۲۰۲۴).

• منابع محاسباتی

الگوریتم‌های یادگیری تقویتی به‌ویژه آن‌هایی که با یادگیری عمیق ترکیب شده‌اند، نیازمند منابع محاسباتی بالایی هستند. پردازش‌های پیچیده و نیاز به حافظه و توان محاسباتی قوی می‌تواند محدودیت‌هایی را در پیاده‌سازی یادگیری تقویتی در سیستم‌های رباتیک ایجاد کند، به ویژه در ربات‌های کوچک با منابع محدود (Zhou et al., ۲۰۲۴).

• مقیاس پذیری و پیچیدگی محیط‌ها

محیط‌های واقعی رباتیک اغلب پیچیده و دینامیک هستند، که این امر مقیاس‌پذیری الگوریتم‌های RL را به چالش می‌کشد. الگوریتم‌های RL باید توانایی مقابله با تغییرات سریع در محیط و تعامل با عوامل متعدد را داشته باشند. تحقیقات بیشتر در زمینه الگوریتم‌های مقیاس‌پذیر و مقاوم در برابر تغییرات محیطی برای پیشبرد RL در رباتیک ضروری است (Calderon-Cordova et al., ۲۰۲۴).

مقایسه الگوریتم‌های یادگیری تقویتی در رباتیک

با وجود پیشرفت‌های چشمگیر در یادگیری تقویتی، انتخاب الگوریتم مناسب برای کاربردهای مختلف رباتیک همچنان یکی از چالش‌های مهم پژوهشگران محسوب می‌شود. در جدول (۱) چندین الگوریتم اصلی یادگیری تقویتی مورد استفاده در رباتیک، مورد مقایسه قرار گرفته است.

جدول (۱): مقایسه الگوریتم‌های یادگیری تقویتی در رباتیک

الگوریتم	کارایی نمونه گیری	نیازمندی‌های محاسباتی	کاربردهای خاص	مزایا	معایب
Q-Learning	پایین	پایین	کنترل ساده ربات‌ها، مسیریابی	سادگی پیاده‌سازی، عدم نیاز به مدل محیط	مقیاس پذیری پایین، مشکل در فضای حالت‌های بزرگ
Deep Q-Network (DQN)	متوسط	بالا	بازی‌های پیچیده، کنترل ربات‌های پیچیده	توانایی کار با فضای حالت‌های بزرگ، یادگیری عمیق	نیاز به داده‌های زیاد، حساسیت به تنظیمات پارامترها
Policy Gradient Methods	بالا	بالا	ربات‌های همیار، تعامل انسان-ربات	بهبود مستقیم خط مشی	حساسیت به نرخ یادگیری
Actor-Critic Method	بالا	بالا	ربات‌های چندعامله، کنترل تطبیقی	پایداری بهتر	پیچیدگی بیشتر نسبت به الگوریتم‌های ساده‌تر
Proximal Policy Optimization	بالا	بالا	کنترل ربات‌های پیشرفته، سیستم‌های چندعامله	پایداری بالا در آموزش، کارایی بالا در استفاده	ساده‌تر از الگوریتم‌های Actor-Critic
Deep Deterministic Policy Gradient	بالا	بالا	ربات‌های اتوماتیک و کنترل پیوسته	مناسب برای فضای عمل پیوسته، یادگیری از طریق تجربه	مناسب برای مسائل کنترل پیوسته، کارایی بالا

افق های آینده یادگیری تقویتی در رباتیک

تحقیقات آینده در زمینه یادگیری تقویتی در رباتیک می تواند به بهینه سازی عملکرد ربات ها ، افزایش توانایی های تطبیقی و ارتقای ایمنی سیستم های رباتیک کمک شایانی کند. در ادامه ، به بررسی چندین جهت گیری کلیدی در افق های آینده یادگیری تقویتی در رباتیک پرداخته می شود که می تواند نقش مهمی در پیشبرد این حوزه ایفا نماید.

• بهبود الگوریتم های یادگیری تقویتی

یکی از زمینه های اصلی تحقیقاتی آینده، بهبود و توسعه الگوریتم های یادگیری تقویتی است. هدف اصلی در این راستا افزایش کارایی نمونه گیری، کاهش نیاز به داده های آموزشی و افزایش سرعت همگرایی الگوریتم ها می باشد. الگوریتم های پیشرفته تر مانند Proximal Policy Optimization (PPO) و Soft Actor-Critic (SAC) در حال حاضر در حال توسعه هستند تا بتوانند عملکرد بهتری در محیط های پیچیده تر ارائه دهند (Kamuni et al.).

• ترکیب یادگیری تقویتی با سایر تکنیک های یادگیری ماشین

ترکیب یادگیری تقویتی با تکنیک های دیگر یادگیری ماشین مانند یادگیری انتقالی (Transfer Learning) و یادگیری نیمه نظارتی (Semi-Supervised Learning) می تواند به بهبود عملکرد و کارایی الگوریتم های یادگیری تقویتی کمک کند. این ترکیب ها می توانند به ربات ها امکان دهند تا از دانش قبلی خود بهره برداری کرده و سریع تر با محیط های جدید سازگار شوند (Rane et al., ۲۰۲۴).

• مقیاس پذیری الگوریتم های یادگیری تقویتی

با افزایش پیچیدگی محیط های رباتیک، مقیاس پذیری الگوریتم های یادگیری تقویتی به یک چالش مهم تبدیل شده است. الگوریتم های یادگیری تقویتی باید توانایی مقابله با تغییرات سریع در محیط و تعامل با عوامل متعدد را داشته باشند. تحقیقات بیشتر در زمینه الگوریتم های مقیاس پذیر و مقاوم در برابر تغییرات محیطی برای پیشبرد یادگیری تقویتی در رباتیک ضروری است (Zhang et al., ۲۰۲۴).

نتیجه گیری

یادگیری تقویتی به عنوان یکی از پیشرفته ترین شاخه های یادگیری ماشین، نقش برجسته ای در تحول حوزه رباتیک ایفا کرده است. این الگوریتم ها با فراهم آوردن امکان یادگیری از طریق تعامل مداوم با محیط، توانسته اند قابلیت های ربات ها را در انجام وظایف پیچیده، هماهنگی و همکاری چندعامله، و تعامل موثر با انسان ها به طور قابل توجهی ارتقاء دهند. کاربردهای گسترده یادگیری تقویتی در رباتیک، از ناوبری خودکار و مسیریابی تا دستکاری اشیاء و کنترل تطبیقی، نشان دهنده پتانسیل بالای این روش در بهبود کارایی و هوشمندی سیستم های رباتیک است.

این مقاله به بررسی کاربردهای یادگیری تقویتی در رباتیک، شناسایی چالش ها و محدودیت های موجود، ارزیابی مقایسه ای الگوریتم های مختلف یادگیری تقویتی و ارائه افق های پژوهشی آینده پرداخته است.

افق های آینده یادگیری تقویتی در رباتیک شامل بهبود و توسعه الگوریتم های پیشرفته تر، افزایش ایمنی و قابلیت اطمینان سیستم ها، ترکیب یادگیری تقویتی با سایر تکنیک های یادگیری ماشین و توسعه یادگیری تقویتی ایمن می باشد. همچنین، تحقیقات بیشتر در زمینه یادگیری انتقالی، مقیاس پذیری الگوریتم ها، و استفاده از یادگیری تقویتی در سیستم های چندعامله و ربات های پزشکی و خدماتی می تواند به رفع چالش های موجود کمک کرده و امکان بهره برداری بهینه تر از پتانسیل های یادگیری تقویتی را فراهم آورد. این جهت گیری ها نشان دهنده پتانسیل بالای یادگیری تقویتی در ارتقاء قابلیت ها و عملکرد ربات ها هستند و مسیرهای جدیدی را برای تحقیقات آتی در این حوزه باز می کنند.



منابع

- Abouelyazid, M. (۲۰۲۴). Reinforcement Learning-based Approaches for Improving Safety and Trust in Robot-to-Robot and Human-Robot Interaction. *Advances in Urban Resilience and Sustainable City Design*, ۱۶(۰۲), ۱۸-۲۹.
- Al-Hamadani, M. N., Fadhel, M. A., Alzubaidi, L., & Harangi, B. (۲۰۲۴). Reinforcement Learning Algorithms and Applications in Healthcare and Robotics: A Comprehensive and Systematic Review. *Sensors*, ۲۴(۸), ۲۴۶۱.
- Calderon-Cordova, C., Sarango, R., Castillo, D., & Lakshminarayanan, V. (۲۰۲۴). A deep reinforcement learning framework for control of robotic manipulators in simulated environments. *IEEE Access*.
- Chen, M., Huang, Y., Wang, W., Zhang, Y., Xu, L., & Pan, Z. (۲۰۲۴). Model inductive bias enhanced deep reinforcement learning for robot navigation in crowded environments. *Complex & Intelligent Systems*, ۱۰(۵), ۶۹۶۵-۶۹۸۲.
- Dong, Y., Zhao, X., Wang, S., & Huang, X. (۲۰۲۴). Reachability Verification Based Reliability Assessment for Deep Reinforcement Learning Controlled Robotics and Autonomous Systems. *IEEE Robotics and Automation Letters*.
- Elguea-Aguinaco, Í., Serrano-Muñoz, A., Chrysostomou, D., Inziarte-Hidalgo, I., Bøgh, S., & Arana-Arexolaleiba, N. (۲۰۲۳). A review on reinforcement learning for contact-rich robotic manipulation tasks. *Robotics and Computer-Integrated Manufacturing*, ۸۱, ۱۰۲۵۱۷.
- Ernst, D., & Louette, A. (۲۰۲۴). Introduction to reinforcement learning. *Feuerriegel, S., Hartmann, J., Janiesch, C., and Zschech, P.*, ۱۱۱-۱۲۶.
- He, X., Hu, Z., Yang, H., & Lv, C. (۲۰۲۴). Personalized robotic control via constrained multi-objective reinforcement learning. *Neurocomputing*, ۵۶۵, ۱۲۶۹۸۶.
- Kaelbling, L. P., Littman, M. L., & Moore, A. W. (۱۹۹۶). Reinforcement learning: A survey. *Journal of artificial intelligence research*, ۴, ۲۳۷-۲۸۵.
- Kamuni, N., Dodda, S., Vuppapapati, V. S. M., Arlagadda, J. S., & Vemasani, P. Advancements in Reinforcement Learning Techniques for Robotics. *Journal of Basic Science and Engineering*, ۱۹, ۱۰۱-۱۱۱.
- Levine, S., Finn, C., Darrell, T., & Abbeel, P. (۲۰۱۶). End-to-end training of deep visuomotor policies. *Journal of Machine Learning Research*, ۱۷(۳۹), ۱-۴۰.
- Mehta, S. A., Habibi, S., & Losey, D. P. (۲۰۲۴). Waypoint-Based Reinforcement Learning for Robot Manipulation Tasks. *arXiv preprint arXiv:۲۴۰۳.۱۳۲۸۱*.
- Milani, S., Topin, N., Veloso, M., & Fang, F. (۲۰۲۴). Explainable reinforcement learning: A survey and comparative review. *ACM Computing Surveys*, ۵۶(۷), ۱-۳۶.
- Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., & Ostrovski, G. (۲۰۱۵). Human-level control through deep reinforcement learning. *nature*, ۵۱۸(۷۵۴۰), ۵۲۹-۵۳۳.
- Nguyen, H., Kozuno, T., Beltran-Hernandez, C. C., & Hamaya, M. (۲۰۲۴). Symmetry-aware Reinforcement Learning for Robotic Assembly under Partial Observability with a Soft Wrist. *arXiv preprint arXiv:۲۴۰۲.۱۸۰۰۲*.
- Ramezani, M., & Amiri Atashgah, M. (۲۰۲۴). Energy-Aware Hierarchical Reinforcement Learning Based on the Predictive Energy Consumption Algorithm for Search and Rescue Aerial Robots in Unknown Environments. *Drones*, ۸(۷), ۲۸۳.
- Rane, N., Mallick, S., Kaya, O., & Rane, J. (۲۰۲۴). Techniques and optimization algorithms in machine learning: A review. *Applied Machine Learning and Deep Learning: Architectures and Techniques*, ۳۹-۵۸.
- Retzlaff, C. O., Das, S., Wayllace, C., Mousavi, P., Afshari, M., Yang, T., Saranti, A., Angersschmid, A., Taylor, M. E., & Holzinger, A. (۲۰۲۴). Human-in-the-loop reinforcement learning: A survey and position on requirements, challenges, and opportunities. *Journal of artificial intelligence research*, ۷۹, ۳۵۹-۴۱۵.
- Sekkat, H., Moutik, O., Ourabah, L., ElKari, B., Chaibi, Y., & Tchakoucht, T. A. (۲۰۲۴). Review of reinforcement learning for robotic grasping: Analysis and recommendations. *Statistics, Optimization & Information Computing*, ۱۲(۲), ۵۷۱-۶۰۱.
- Souza, G. K. B., Santos, S. O. S., Ottoni, A. L. C., Oliveira, M. S., Oliveira, D. C. R., & Nepomuceno, E. G. (۲۰۲۴). Transfer reinforcement learning for combinatorial optimization problems. *Algorithms*, ۱۷(۲), ۸۷.
- Tang, C., Abbatematteo, B., Hu, J., Chandra, R., Martín-Martín, R., & Stone, P. (۲۰۲۴). Deep reinforcement learning for robotics: A survey of real-world successes. *Annual Review of Control, Robotics, and Autonomous Systems*, ۸.
- Wang, Z., Yan, H., Wang, Z., Xu, Z., Wu, Z., & Wang, Y. (۲۰۲۴). Research on autonomous robots navigation based on reinforcement learning. *۲۰۲۴ ۳rd International Conference on Robotics, Artificial Intelligence and Intelligent Control (RAIIC)*.
- Xie, Z., Lu, L., Wang, H., Su, B., Liu, Y., & Xu, X. (۲۰۲۴). Improving workers' musculoskeletal health during human-robot collaboration through reinforcement learning. *Human Factors*, ۶۶(۶), ۱۷۵۴-۱۷۶۹.
- Zhang, Z., Jiang, X., Yang, Z., Ma, S., Chen, J., & Sun, W. (۲۰۲۴). Scalable Multi-Robot Task Allocation Using Graph Deep Reinforcement Learning with Graph Normalization. *Electronics*, ۱۳(۸), ۱۵۶۱.
- Zhou, G., Tian, W., Buyya, R., Xue, R., & Song, L. (۲۰۲۴). Deep reinforcement learning-based methods for resource scheduling in cloud computing: A review and future directions. *Artificial Intelligence Review*, ۵۷(۵), ۱۲۴.
- Zhu, K., & Zhang, T. (۲۰۲۱). Deep reinforcement learning based mobile robot navigation: A review. *Tsinghua Science and Technology*, ۲۶(۵), ۶۷۴-۶۹۱.



A Review of the Applications of Reinforcement Learning in Robotics: Advancements, Challenges, and Future Perspectives

Mojtaba Abdollahi

Department of Computer Engineering, Torbat Jam Branch, Islamic Azad University, Torbat Jam, Iran

Hojjat Azadravesh

Department of Computer Engineering, Torbat Jam Branch, Islamic Azad University, Torbat Jam, Iran

Abstract

Reinforcement learning, as one of the most advanced branches of artificial intelligence, plays a crucial role in enhancing the performance and intelligence of robotic systems. These algorithms enable continuous interaction between robots and complex environments, allowing them to learn optimal and autonomous decision-making through direct experience. Applications of reinforcement learning in robotics include autonomous navigation, object manipulation, human-robot interaction, and multi-agent system management, significantly improving the efficiency and flexibility of robots. Despite these advancements, challenges such as inefficient sampling, high computational resource requirements, safety and reliability concerns, and algorithm scalability continue to impede the full exploitation of reinforcement learning in robotics. Particularly in complex domains like interaction with dynamic and unpredictable environments, reinforcement learning enables robots to continuously improve their behaviors. These algorithms not only aid in optimizing learning processes in changing environments but also enhance the adaptability and faster learning of robots when faced with new situations. This paper examines the applications of reinforcement learning in robotics, identifies existing challenges and limitations, provides a comparative evaluation of different reinforcement learning algorithms, and presents future research directions.

Keywords: Reinforcement Learning, Robotics, Artificial Intelligence, Machine Learning.