शोध पत्र/Research Paper

डबल-डीक्यूएन के एकीकरण के साथ पारंपरिक तंत्रिका तंत्र - द्विदिशिक दीर्घकालिक अल्पकालिक स्मृति का उपयोग करके गतिशील वस्तु पहचान और रोबोट ग्रास्पिंग

Dynamic Object Recognition and Robot Grasping using CNN-BiLSTM with Integration of Double-DQN

सौम्यश्री[1], गुरुमीत सिंह[2], हरजोत सिंह गिल[3], संजय सिंह राठौड़[4] एवं मणिकंवर सिंह[5]

Soumyashree[1], Gurmeet Singh[2], Harjot Singh Gill[3], Sanjay Singh Rathore[4] and Mani Kanwar Singh[5]

[1]Department of Mechatronics Engineering, Chandigarh University, India

[2]Associate Professor, Department of Mechatronics Engineering, Chandigarh University, India

[3]Professor, Department of Mechatronics Engineering, Chandigarh University, India

[4,5]Skill Department of Automotive Studies, Shri Vishwakarma Skill University, Haryana

[1]soumyashree1926@gmail.com, [2]gurmeet.mech@cumail.in, [3]harjot.gill@cumail.in,

[4]sanjay.singh@svsu.ac.in, [5]mani.singh@svsu.ac.in

https://doi.org/10.5281/zenodo.15564190

सारांश

यह अध्ययन CNN-BiLSTM तकनीक का उपयोग करते हुए एक नवीन वस्तु पहचान एल्गोरिथ्म का परिचय देता है। उन्नत संवेदी प्रौद्योगिकियों और गहन सुदृढीकरण सीखने के दृष्टिकोण के एकीकरण के माध्यम से, प्रस्तावित प्रणाली विविध वस्तुज्यामिति, आकार और सतह गुणों को समायोजित करने के लिए अपने पहचान मापदंडों को गतिशील रूप से समायोजित करती है। इस दृष्टिकोण में एक तंत्रिका नेटवर्क एजेंट को एक छवि विंडो से समझने के लिए प्रशिक्षण देना शामिल है, जिसमें एक पूर्व निर्धारित क्षेत्र के भीतर वस्तुओं पर ध्यान केंद्रित करने की आवश्यकता होती है। उच्च-आयामी और समय-श्रृंखला डेटा को प्रभावी ढंग से संभालने के लिए, डबल DQN ढांचे का उपयोग किया जाता है। प्रस्तावित दृष्टिकोण ऑब्जेक्टरिकग्निशन एल्गोरिथ्म को परिष्कृत सुदृढीकरण सीखने की तकनीकों, विशेष रूप से डबल-डीक्यूएन के साथ जोड़ता है, जो रोबोट को विभिन्न वातावरणों में वस्तुओं को गतिशील रूप से समझने और समझाने में सक्षम बनाता है। प्रस्तावित पद्धति की तुलना में वर्तमान दृष्टिकोणों की प्रभावशीलता का मूल्यांकन करने के लिए तुलनात्मक अनुसंधान करके परिणाम प्राप्त किए गए। क्यू-नेटवर्क को दो अलग-अलग गतियों, अर्थात् ''ट्विस्ट'' और ''पुश'' को निष्पादित करने के लिए प्रशिक्षित किया जाता है, इच्छित गति को एक निर्दिष्ट पैरामीटर (''पुश'' के लिए +1 और ''ट्विस्ट'' के लिए -1) के रूप में प्रदान किया जाता है। इसके अलावा, जब कार्य पैरामीटर-1 और +1 के बीच आता है तो नेटवर्क सीखी गई गतियों के बीच मध्यवर्ती गति उत्पन्न कर सकता है। निष्कर्षों से विभिन्न आइटम आकारों, आकृतियों और स्थितियों में सफलता दर और अनुकूलन क्षमता को समझने में उल्लेखनीय वृद्धि का पता चलता है। विशेष रूप से, अध्ययन में 90% की सफलता दर प्राप्त करते हुए, कार्य पूरा करने की दर में उल्लेखनीय वृद्धि दर्ज की गई है।

Abstract

This study introduces a novel object recognition algorithm employing the CNN-BiLSTM technique. Through the integration of advanced sensory technologies and deep reinforcement learning approaches, the proposed system dynamically adjusts its detection parameters to accommodate diverse object geometries, sizes, and surface properties. The approach entails training a neural network agent to discern, from an image window, which objects within a predetermined region warrant focused attention. To effectively handle high-dimensional and time-series data, a Double DQN framework is utilized. The proposed approach amalgamates object recognition algorithms with sophisticated reinforcement learning techniques, notably Double DQN, enabling robots to dynamically perceive and grasp items in varied environments. The outcomes were obtained by conducting comparative research to evaluate the effectiveness of current approaches in comparison to the proposed method. The Q-network is trained to execute two distinct motions, namely "twist" and "push," with the intended motion provided as an assigned parameter (+1 for "push" and -1 for "twist"). Moreover, the network can generate intermediate movements between the learned motions when the task parameter falls between -1 and +1. The findings reveal notable enhancements in grasping success rates and adaptability across diverse item sizes, shapes, and conditions. Notably, the study reports a significant increase in the task completion rate for grasping, achieving a success rate of 90%.

मुख्य शब्दः डबल-डीक्यूएन, सीएनएन-एलएसटीएम, रोबोट ग्रैस्पिंग, ऑब्जेक्ट डिटेक्शन।

Key Words: Double DQN, CNN-LSTM, Robot Grasping, Object Detection.

परिचय

ऑटोमेशन और रोबोटिक्स में पिछले कई वर्षों में बड़ी उपलब्धियाँ हासिल हुई हैं, खासकर रोबोटिकग्रास्पिंग के क्षेत्र में। रोबोटिक प्रणालियों की दक्षता और अनुकूलनशीलता को बढ़ाने की खोज ने अत्याधुनिक प्रौद्योगिकियों की खोज को जन्म दिया है, जिनमें से सुदृढीकरण सीखने के एल्गोरिथ्म एक महत्वपूर्ण शक्ति के रूप में उभरे हैं। यह शोध इस चुनौती के संभावित प्रभावी समाधान की जांच करता है: अनुकूली पहचान और सटीक पिक तंत्र के साथ रोबोटिक ग्रैस्पिंग को सुदृढीकरण सीखने के एल्गोरिथ्म द्वारा व्यावहारिक बनाया गया है[1]। गहरी शिक्षा के उपयोग के माध्यम से, और विशेष रूप से सुदृढीकरण के माध्यम से सीखने के माध्यम से, शोधकर्ता ऐसे संचालन को चुनना और रखना चाहते हैं जिसमें विनिर्माण प्रक्रियाओं, भंडारण सुविधाओं के परिवहन और उपचार के मूल्य के साथ संवेदनशील बातचीत शामिल हो, जिसमें अत्यधिक सटीकता की आवश्यकता होती है।

रोबोटिक्स अनुसंधान के भीतर प्राथमिक बाधाओं में से एक रोबोटिक ग्रैस्पिंग से संबंधित है, जो सटीकता और अनुकूलन क्षमता के साथ वस्तुओं में हेरफेर करने की मशीनों की क्षमता से संबंधित है। सरल यांत्रिक उपकरणों के साथ अपनी शुरुआत के बाद से इस क्षेत्र ने महत्वपूर्ण प्रगति की है, अब यह कृत्रिम बुद्धिमत्ता और सॉफ्टरोबोटिक्स में प्रगति के क्षेत्र में आगे बढ़ रहा है। रोबोटिक ग्रैस्पिंग की विकासवादी यात्रा को समझने से प्रौद्योगिकी में उन चुनौतियों और प्रगति के बारे में मूल्यवान दृष्टिकोण मिलते हैं जिनका शोधकर्ताओं ने वर्षो से सामना किया है।

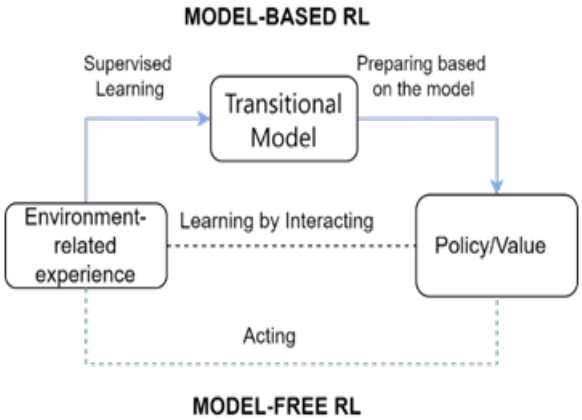
वस्तु हेरफेर में लचीलापन महत्वपूर्ण होता जा रहा है क्योंकि कंपनियां सटीकता और प्रभावशीलता के लिए प्रौद्योगिकी पर अधिक से अधिक निर्भर होती जा रही हैं। जीवन में परिस्थितियों की आंतरिक समृद्धि और अप्रत्याशितता कभी-कभी पारंपरिक स्वचालित समझ तकनीकों के लिए चुनौतियां पेश करती है। हमारा अध्ययन उन कठिनाइयों के अनुसार रोबोट में हेरफेर करने के पारंपरिक तरीकों पर पुनर्विचार करने के लिए आरएल पद्धति का उपयोग करता है। ये दृष्टिकोण रोबोटिक्स को ज्ञान प्राप्त करने और बदलते और अप्रत्याशित परिवेश में मौजूदा मनोरंजक तकनीकों को संशोधित करने में सक्षम बनाते हैं[2]। ये प्रयोग और विफलता की मूलभूत अवधारणाओं पर आधारित हैं। इस प्रयास में, केवल समझने की क्षमताओं से परे प्रगति की गई है, इसके बजाय गतिशील पहचान प्रणालियों के एकीकरण पर जोर दिया गया है।

रोबोटिक्स को समझने में विभिन्न प्रकार की जटिल समस्याएं शामिल हैं, जैसे लचीली प्रणालियों की आवश्यकता

जो वस्तुओं की एक विस्तृत श्रृंखला को संभाल सके, सेंसर की क्षमताओं पर बाधाएं, वास्तविक समय में धारणा, शोर और अनिश्चितता। मनोरंजक तकनीकों के लचीलेपन, सुरक्षा चिंताओं और अधिक उन्नत कार्यो के साथ सहज समन्वय से जटिलता और बढ़ जाती है। रोबोटिक्स प्रौद्योगिकी, कृत्रिम बुद्धिमत्ता और तंत्रिका नेटवर्क में विकास के माध्यम से, डेवलपर्स इन बाधाओं को दूर करने और विभिन्न स्थितियों में स्वचालित समझ की अनुकूलनशीलता और दक्षता में सुधार करने के लिए काम करते हैं।

रोबोटिक हैंडलिंग के तरीके: विभिन्न कारकों के अनुसार, दृष्टि-आधारित हेरफेर करने वाले रोबोट के लिए विभिन्न तरीके मौजूद हैं। विश्लेषणात्मक तरीके आदर्श पकड़ स्थिति निर्धारित करने के लिए वस्तु के रूप की जांच करते हैं। मशीन लर्निंग डेटा-संचालित पद्धतियों की नींव है, जो बेहतर प्रसंस्करण क्षमताओं, सांख्यिकीय तरीकों और डेटा पहुंच के परिणामस्वरूप काफी उन्नत हुई है। सूचना-संचालित तरीकों को परिष्कृत करने के लिए मॉडल-आधारित और मॉडल-मुक्तरण नीतियों का उपयोग किया जा सकता है।

मॉडल-आधारित आरएल एल्गोरिथ्म: मॉडल-आधारित रोबोटिक ग्रैस्पिंग प्रक्रियाओं को तीन अलग-अलग चरणों में विभाजित करना संभव है: ऑब्जेक्ट पोज का अनुमान लगाना, पकड़ने की स्थिति का निर्धारण करना, और आइटम का चयन करने के लिए मार्ग की योजना बनाना। इस दृष्टिकोण का उपयोग करते हुए, रोबोटिक भुजा विश्लेषण करती है और पकड़ने वाली क्रियाएं करती है, अक्सर सिमुलेशन या सांख्यिकीय तरीकों के माध्यम से पकड़ को अनुकूलित करती है। मॉडल-आधारित दृष्टिकोण के लिए उत्पाद और वातावरण प्रतिनिधित्व की सटीकता आवश्यक है, और भविष्यवाणी और वास्तविक परिस्थितियों के बीच कोई भी असमानता दृष्टिकोण की प्रभावशीलता पर प्रभाव डाल सकती है[3]।



चित्र 1. मॉडल-आधारित और मॉडल-मुक्त आरएल का संक्रमण

मॉडल-मुक्त आरएल एल्गोरिथ्म: मॉडल-आधारित तकनीकों के विपरीत, मॉडल-मुक्त दृष्टिकोण उन वस्तुओं के बारे में पिछली जानकारी पर निर्भर नहीं हो सकते हैं जिनसे निपटते हैं। गहन शिक्षण और सुदृढीकरण शिक्षण दो विधियाँ हैं जिनका उपयोग इस प्रकार के निर्देश को पूरा करने के लिए किया जा सकता है। अप्रत्याशित और अप्रशिक्षित संदर्भो में, जहां पहले से सटीक सिमुलेशन तैयार करना मुश्किल हो सकता है, मॉडल-मुक्त तकनीकें फायदेमंद होती हैं, लेकिन विभिन्न वस्तुओं और परिस्थितियों पर सामान्यीकरण के संदर्भ में अच्छा प्रदर्शन करने के लिए उन्हें बहुत सारी निर्देशात्मक जानकारी की आवश्यकता हो सकती है।

साहित्य समीक्षा

शोध से पता चलता है कि भौतिक यांत्रिकी की स्पष्ट रूप से भविष्यवाणी करने से एक ऐसी नीति बनती है जो हाथ से तैयार की गई गतिशीलता आधार रेखा और ''मूल्य-नेटवर्क'' दोनों से बेहतर प्रदर्शन करती है। उत्तरार्द्ध आमतौर पर सटीक मूल्य अनुमान उत्पन्न करने के लिए समान यांत्रिकी की अंतर्निहित भविष्यवाणी पर निर्भर करता है। ये निष्कर्ष नीति निर्माण प्रक्रिया में भौतिक यांत्रिकी को स्पष्ट रूप से एकीकृत करने की श्रेष्ठता पर जोर देते हैं[4]। कई समूहों के साथ प्रारंभिक वर्गीकरण के संबंध में समस्या का समाधान करने के लिए,[5] के शोधकर्ताओं ने (डीक्यूएन) तकनीक पर आधारित एक डीआरएल विधि का सुझाव दिया। किसी चीज को पकड़ना पकड़ने या उठाने की क्रिया है। मनोरंजक गतिविधि करते समय, तीन महत्वपूर्ण कारकों पर विचार किया जाना चाहिए: स्थान, वस्तु और परिवेश। प्रस्तावित ढांचे के प्रदर्शन का व्यापक मूल्यांकन करने के लिए सिमुलेशन और वास्तविक दुनिया की जांच के संयोजन के माध्यम से कठोर मूल्यांकन किया जाता है। सिम्युलेटेड परिदृश्यों का उपयोग ढांचे की जटिल कार्यक्षमताओं का पता लगाने के लिए किया जाता है, जो विभिन्न परिस्थितियों में इसके व्यवहार में अंतर्दृष्टि प्रदान करता है[6,7]। ग्रिप कार्य को बढ़ाने के लिए कई तरीके स्थापित किए जा रहे हैं;[8] में लेखक वस्तु का पता लगाने, सीखने और समझने के विकल्पों पर जोर देने के साथ वर्तमान विकास की समीक्षा प्रदान करते हैं। यह अनुभवजन्य सत्यापन भौतिक प्रणालियों की सटीक भविष्यवाणी और हेरफेर की आवश्यकता वाले

कार्यों के लिए नीतियों को अनुकूलित करने में भौतिक यांत्रिकी को स्पष्ट रूप से मॉडलिंग करने के महत्व को रेखांकित करता है[9,10]। उल्टे गतिकी का उपयोग करने वाली अन्य स्थापित विधियाँ भी मौजूद हैं, जिनमें[11,12] शामिल है, जो उल्टे गति की समाधानों के लिए ठोस व्यवस्था के संयुक्त-स्थान पथ प्रदान करने के लिए आर एल-आधारित दृष्टिकोण का सुझाव देता है। शोध का उद्देश्य व्युत्क्रम गतिशीलता का एक समाधान खोजना है जो एक स्वायत्त शाखा को एक गहरी पूर्वानुमानित नीति भिन्नता तकनीक[13] का उपयोग करके दृष्टि के अनुसार वस्तुओं को हथियाने का काम पूरा करने की अनुमति देता है। शिक्षाविदों द्वारा नई तकनीकों की जांच की जा रही है, जैसे सीखने के लिए तंत्रिका नेटवर्क जो ज्ञान का कुशलता पूर्वक उपयोग करते हैं[14]।

जबकि DQN कुछ परिदृश्यों में उत्कृष्ट प्रदर्शन प्रदर्शित करता है, यह उन नौकरियों के लिए अनुपयुक्त है जिनके लिए निरंतर प्रक्रिया स्थान की आवश्यकता होती है। DQN को पूर्वानुमानित नीति उन्नयन विधि" के साथ एकीकृत करके, DDPG इस समस्या से निजात पा लेता है। शोधकर्ताओं द्वारा[15] में कई ऑफ-पॉलिसी, मॉडल-मुक्त, डीआरएल विधियों का अनुभवजन्य मूल्यांकन प्रस्तुत किया गया है। मॉडल-मुक्तडीप-आरएल का उपयोग करके, शोध कर्ता यह पता लगाने और अध्ययन करने में सक्षम थे कि कैसे धक्का देने से कलाई और हाथों के लिए जगह बनाने के लिए भीड़ वाली वस्तुओं को पुनर्गठित करने में मदद मिल सकती है। यह स्व-पर्यवेक्षित डीआरएल के साथ पकड़ने और चलने के बीच तालमेल की खोज से संभव हुआ। जी एस एन के माध्यम से,[16] के शोधकर्ता एक ऐसा दृष्टिकोण सुझाते हैं जो एक साथ नियमों को हथियाने और व्यवस्थित करने को समझता है। यह एक रोबोटिक भुजा को सतह से उचित रूप से पैकेजों का चयन करने और उन्हें एक ऊंची सतह पर रखने की अनुमति देता है।

तालिका 1. साहित्य की समीक्षा

वर्ष	विधि का प्रयोग किया गया	कथित रोबोटिक गति	दक्षता दर (% में)	मॉडलमुफ्त (✓) / आधारित (-)	प्रसंग संख्या।
2024	डबलडीक्यूएन	उठाइए और जगह पर रखिए	100	✓	[1]
2024	डॉ एल	पकड़ने में	95	-	[2]
2024	डॉ एल	पकड़ने में	80	-	[3]
2023	डीआरएल (पीक्यूसीएन)	धकेलना, पकड़ना और रखना	88	✓	[4]
2023	डीआरएल (क्यू-लर्निंग)	पकड़ने में	93.8	✓	[5]
2023	आर एल	पकड़ने में	-	-	[6]
2023	आर एल	पकड़ें, और चुनें और रखें	100	-	[7]
2022	डेली	पकड़ने में	90	-	[8]
2021	डॉ एल	पकड़ने में	86-96	✓	[9]
2021	आर एल	मनोरंजक	-	✓	[10]
2021	डीआरएल, जेनेटिक एल्गोरिथ्म (जीए)	पकड़ने में	-	-	[11]
2020	डीआरएल (क्यू-मान)	पकड़-से-स्थान	90	✓	[12]
2020	आर एल	पकड़ने का कार्य	100	✓	[13]
2020	डेली	पकड़ना और हिलाना	-	-	[14]
2018	डबलडीक्यूएन	पकड़ने में	90	✓	[15]
2018	आरएल (क्यू-सीखना)	पकड़ना और धकेलना	70	✓	[16]

'पीक्यूसीएन-पिक्सेल-वारक्यू-मूल्यवान क्रिटिक नेटवर्क, रुजीए-जेनेटिक एल्गोरिथ्म, जीएसएन-स्टैकिंग नेटवर्क के लिए ग्रैस्पिंग

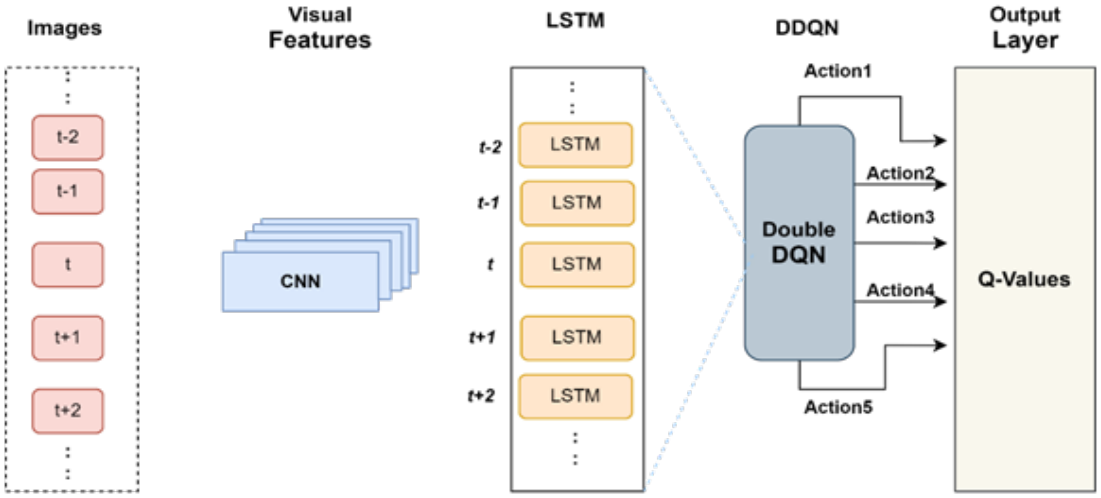
प्रक्रिया

सीएनएनः कन्वेन्शनल न्यूरल नेटवर्क (सीएनएन) कृत्रिम बुद्धिमत्ता के भीतर, विशेष रूप से कंप्यूटर दृष्टि के भीतर एक महत्वपूर्ण प्रगति के रूप में दृष्टिगोचर होता है। वे मशीनों द्वारा दृश्य डेटा की व्याख्या और समझ में एक परिवर्तनकारी सफलता का प्रतिनिधित्व करते हैं। सीएनएन के मूल में संकेंद्रित परतें होती हैं, जहां मिनट फिल्टर इनपुट छवियों को पार करते हैं, पिक्सेल के बीच स्थानिक संबंधों को बनाए रखते हुए सुविधाओं को निकालते हैं। इन परतों को आमतौर पर पूलिंग परतों द्वारा सफल बनाया जाता है, जो निकाली गई विशेषताओं का नमूना लेती है, जिससे कम्प्यूटेशनल जटिलता कम हो जाती है और ओवर फिटिंग कम हो जाती है। ReLU जैसे सक्रियण कार्यो द्वारा शुरू की गई गैर-रैखिकताएं, CNN को जटिल पैटर्न को समझने में सशक्त बनाती हैं। कंप्यूटर विजन के क्षेत्र में, फीचर निष्कर्षण और वर्गीकरण की खोज अनुसंधान का एक महत्वपूर्ण अवसर है। छवि प्रसंस्करण में पारंपरिक दृष्टिकोण आमतौर पर सांख्यिकीय नियमितताओं या मौजूदा ज्ञान से प्राप्तपूर्व-डिजाइन की गई सुविधाओं पर निर्भर करते हैं, जो मूल छवि डेटा को पूरी तरह और सटीक रूप से प्रस्तुत करने की उनकी क्षमता को सीमित करते हैं। कन्वेन्शनल न्यूरल नेटवर्क्स (सीएनएन) एक एंड-टू-एंड लर्निंग प्रतिमान प्रदान करते हैं जहां मापदंडों को ग्रेडिएंटडिसेंट के माध्यम से परिष्कृत किया जाता है।

BiLSTM: ऑब्जेक्ट डिटेक्शन के लिए बिडायरेक्शनल लॉन्ग शॉर्ट-टर्ममेमोरी (BiLSTM) नेटवर्क को नियोजित करने की प्रक्रिया छवियों और उनके संबंधित ऑब्जेक्ट बाउंडिंग बॉक्स वाले लेबल वाले डेटासेट के निर्माण से शुरू होती है। इसके बाद, इन छवियों से पदानुक्रमित विशेषताओं को निकालने के लिए पूर्व-प्रशिक्षित कन्वेन्शनल न्यूरलनेटवर्क (सीएनएन) मॉडल का उपयोग किया जाता है, जिन्हें फिर फीचरवैक्टर के अनुक्रम में एन्कोड किया जाता है। प्रासंगिक जानकारी को द्विदिशात्मक रूप से कैप्चर करने के लिए इन अनुक्रमों को BiLSTM परत में इनपुट किया जाता है। इसके बाद, वस्तुओं की उपस्थिति और उनके बाउंडिंग बॉक्स निर्देशांक की भविष्यवाणी करने के लिए नेटवर्क में एक डिटेक्शन हेड जोड़ा जाता है। पूरे नेटवर्क का प्रशिक्षण उचित हानि कार्यो का उपयोग करके शुरू से अंत तक आयोजित किया जाता है, जिसमें प्रदर्शन मूल्यांकन सटीक, रिकॉल और माध्य औसत परिशुद्धता (एमएपी) जैसे मैट्रिक्स का उपयोग करके किया जाता है। अस्थायी निर्भरता को पकड़ने में इस दृष्टिकोण की क्षमता के बावजूद, पारंपरिक सीएनएन-आधारित तरीकों के सापेक्ष इसकी दक्षता और सटीकता विशिष्टकार्य आवश्यकताओं के आधार पर भिन्न हो सकती है।

डबल डीक्यूएनः रोबोटिक ग्रैस्पिंग में डबल डीपक्यू-नेटवर्क (डबलडीक्यूएन) को शामिल करने से वस्तुओं को प्रभावी ढंग से पकड़ने मे रोबोट की दक्षता को बढ़ाने के लिए सुदृढीकरण सीखने के सिद्धांतों का उपयोग करना शामिल है। इस पद्धति के माध्यम से, रोबोट अपने पर्यावरणीय अवलोकनों का आकलन करके और पिछले अनुभवों का लाभ उठाकर इष्टतम समझ क्रियाओं को समझने की क्षमता प्राप्त करता है। प्रारंभिक चरण में पर्यावरणीय स्थिति को परिभाषित करना शामिल है, जिसमें ऑब्जेक्ट स्थिति, ग्रिपर स्थिति और सेंसर डेटा जैसे कैमरा छवियां या गहराई मान चित्र जैसे कारक शामिल हैं। इसके अलावा, रोबोट के लिए उपलब्ध क्रियाएं निर्दिष्ट की गई हैं, जिसमें ग्रिपर की स्थिति और अभिविन्यास में संभावित समायोजन शामिल हैं।

वास्तविक समय ग्रैस्पिंग निष्पादन के दौरान, प्रशिक्षित डबलडीक्यूएन एजेंट पर्यावरण की मौजूदा स्थिति का आकलन करके गतिशील रूप से कार्यो का चयन करता है, रोबोटिकआर्म या ग्रिपर को तदनुसार ग्रैस्पिंग गति को निष्पादित करने का निर्देश देता है। इसके बाद, एजेंट के मापदंडों को पुनरावृत्तीय रूप से समायोजित करने के लिए पर्यावरण से प्रतिक्रिया, जैसे समझने के प्रयासों (सफलता या विफलता) के परिणामों का लाभ उठाया जाता है। यह पुनरावृत्तीय प्रक्रिया समय के साथ एजेंट की समझने की रणनीतियों को परिष्कृत करती है, जिससे वस्तुओं को प्रभावी ढंग से पकड़ने में उसकी दक्षता बढ़ती है।



चित्र 2. ग्रैस्पिंग सिस्टम की कार्यप्रणाली का फ्लोचार्ट

मॉडल प्रशिक्षण और प्रयोग डिजाइन

डेटासंग्रहण: वर्तमान शोध में, प्रशिक्षण डेटा एकत्र करने के लिए मोशन इंटरपोलेशन नियंत्रण का उपयोग किया गया था। इस दृष्टिकोण में शुरुआत और अंत दोनों मुद्राओं का वर्णन शामिल है, जिन्हें फिर संयुक्त कोणों की गणना करके हाथ में हेरफेर कार्रवाई प्रदान करने के लिए प्रक्षेपित किया जाता है। जबकि रोबोटिक हाथों से पकड़ी गई गतिविधि को पूरा करने के लिए इंटरपोलेशन नियंत्रण उत्कृष्ट है, यह उंगलियों पर प्रारंभिक पकड़ने वाले स्थानों में दोषों के प्रति खराब अनुकूलन क्षमता दिखाता है और वस्तुओं पर अत्यधिक बल डाल सकता है, जिससे टूटने का खतरा बढ़ जाता है। तंत्रिका नेटवर्क और इंटरपोलेशन नियंत्रण के बीच एक तुलनात्मक विश्लेषण से इन सीमाओं का पता चला। इन-हैंड हेरफेर के लिए तंत्रिका नेटवर्क के प्रशिक्षण की प्रभावकारिता में सुधार करने के लिए, प्रयोगकर्ता जानबूझ कर प्रत्येक परीक्षण के लिए लक्ष्य वस्तु की प्रारंभिक पकड़ने की स्थिति को यादृच्छिक बनाता है। यह रणनीति सुनिश्चित करती है कि प्रशिक्षण डेटा में विभिन्न प्रारंभिक पकड़ स्थितियों से शुरू होने वाली गतियाँ शामिल हों।

दूसरों के मुकाबले अपनी पद्धति को बेंचमार्क करने के लिए, हम कागल डेटासेट का उपयोग करके अपनी वास्तुकला का मूल्यांकन करते हैं। इस डेटासेट में 228 अलग-अलग वस्तुओं वाली 864 छवियां शामिल हैं। प्रत्येक छवि में कई ग्रैस्प आयत होते हैं, जिन्हें समानांतर प्लेट ग्रिपर्स पर फोकस के साथ सफल (सकारात्मक) या असफल (नकारात्मक) के रूप में वर्गीकृत किया जाता है। कुल मिलाकर, डेटासेट में 7085 लेबल वाले ग्रैप्स हैं, जिनमें 5055 सकारात्मक उदाहरण और 2030 नकारात्मक उदाहरण शामिल हैं। डेटासेट के दो अलग-अलग विभाजन बनाए गए हैं:

<u>चित्रानुसार विभाजन:</u> इस विभाजन में डेटा सेट में प्रत्येक चित्र को यादृच्छिक रूप से पांच तहों में विभाजित किया गया है। यह विधि यह आकलन करने के लिए अच्छी तरह से काम करती है कि नेटवर्क विभिन्न अभिविन्यासों और स्थानों में पहले से देखी गई वस्तुओं को कितनी अच्छी तरह सामान्यीकृत करता है।

<u>वस्तु-वार विभाजन:</u> किसी निश्चित आइटम की सभी तस्वीरें एक ही सत्यापन सेट में एकत्रित की जाती हैं, और प्रत्येक ऑब्जेक्ट के सभी उदाहरण यादृच्छिक रूप से विभाजित होते हैं। यह तकनीक यह निर्धारित करने में सहायक है कि नेटवर्क उन वस्तुओं को कितने प्रभावी ढंग से सामान्यीकृत करता है जिन्हें उसने पहले नहीं देखा है।

प्रशिक्षण मॉडल: प्रशिक्षण के लिए, सीएनएन को आमतौर पर स्वतंत्र डेटा की आवश्यकता होती है, हालांकि एलएसटीएम नेटवर्क मेमोरी में लगातार संग्रहीत नमूनों से कनेक्शन का पता लगाने में काफी अच्छे हैं। स्थानांतरण शिक्षण तकनीक का उपयोग करते हुए, यह शोध इस संघर्ष को हल करने का प्रयास करता है। सीएनएन मॉडल को सबसे पहले उन डेटा पर प्रशिक्षित किया जाता है जो अव्यवस्थित हैं। फिर, बिना फेरबदल

किए गए नमूनों का उपयोग करके, सीखे गए मापदंडों को लोड किया जाता है और एलएसटीएम भाग को प्रशिक्षित करने के लिए तय किया जाता है। यह तकनीक सीएनएन और एलएसटीएम दोनों के फायदों का उपयोग करके अनुक्रमिक डेटा को संभालना आसान बनाती है। प्रशिक्षण के लिए डेटासेट से 7085 नमूनों का उपयोग किए जाने के बाद वर्तमान मॉडल के पैरामीटर रखे गए हैं। उसके बाद, नए डेटा सेट पर प्रशिक्षण जारी रखने के लिए इन संग्रहीत मापदंडों को फिर से आयात किया जाएगा। इस पद्धति का उपयोग करके, स्थिरता बनाए रखी जाती है और जब मॉडल को ताजा डेटा मिलता है तो वह पहले के प्रशिक्षण का लाभ उठा सकता है।

CNN & LSTM नेटवर्क का निर्माण और प्रशिक्षण Tensor Flow का उपयोग करके किया जाता है। प्रारंभ में, सीएनएन घटक अपनी पथ नियोजन क्षमता का आकलन करने के लिए प्रशिक्षण और मूल्यांकन से गुजरता है[20]। इसके बाद, सिमुलेशन अनुभाग में LSTM घटक का मूल्यांकन किया जाता है। यह अनुक्रमिक दृष्टिकोण कार्य के विभिन्न पहलुओं में नेटवर्क के प्रदर्शन के व्यापक मूल्यांकन की अनुमति देता है।

तालिका 2. विकास मैट्रिक्स की तुलना

तरीके	सीएनएन	आर-सीएनएन	सीएनएन-एलएसटीएम	सीएनएन-बीआईएलएसटीएम
एमएपीई%	7.5	5.87	2.34	1.76
आर$_2$	84	89.43	91.7	99.02

डीक्यूएन वास्तुकला

किसी दी गई अवलोकन योग्य स्थिति में, R_s एक रोबोट एक नीति के अनुसार अपनी कार्रवाई तय करता है π, अगले राज्य में संक्रमण $R_{s+1} = h(R_s, B_s)$ इस पुनरावृत्तीय प्रक्रिया में, एक अच्छी तरह से परिभाषित कार्यनीति का होना आवश्यक है। संचयी इनाम को इस प्रकार दर्शाया गया है. $P_t = \sum_{s=t} \delta\delta^{s-t} r_t$ सुदृढीकरण सीखने में, नीति के तहत क्रिया-मूल्य कार्य करता है π के रूप में दर्शाया गया है। $Q^{\pi}(r, b) = \varepsilon [Q_s|R_s, B_s, \pi]$ इष्टतम नीति निर्धारित करने में क्रिया-मूल्य फंक्शन का अनुमान लगाना शामिल है, जिसे बेलमेन समीकरण का उपयोग करके प्राप्त किया जा सकता है: $Q^{\pi}(r, b)$

$$Q_{s+1}(r, b) = \delta[a+\gamma \max P_t(r', b')] \quad (1)$$

यह अध्ययन डी आर एल आर्किटेक्चर का उपयोग करके आदर्श एक्शन-वैल्यू फंक्शन (Q*) का अनुमान लगाने के लिए सीएनएन का उपयोग करता है। पैरामीटर α द्वारा परिभाषित नेटवर्क, के एक गैर-रेखीय सन्निकटन के रूप में कार्य करता है, जिसे $Q(r,b;\alpha) = Q^*(r,b)$ के रूप में व्यक्त किया जाता है। इन मापदंडों की गणना अस्थायी अंतर त्रुटि (टीडी-त्रुटि) को कम करके बार-बार की जाती है।

$$\hat{\theta} = 2\arg\min \delta[a + \gamma \max Q(r',b';\theta) - Q(r,b;\theta)] \quad (2)$$

DQN में, अनुकूलन प्रक्रिया को अस्थायी अंतर त्रुटि (TD-err) को हानि फंक्शन के रूप में मानते हुए, एक प्रतिगमन कार्य में बदल दिया जाता है।

$$L(\theta) = [a+\gamma \max Q(r',b';\theta) - Q(r,b;\theta))]^2 \quad (3)$$

सीएनएन का उपयोग करके, छवियों से विशेषताएं पुनर्प्राप्त की जाती हैं। फीचर एक्सट्रैक्टर ResNet-18 की शुरुआती 9 परतों और दो अतिरिक्त सब सैंपलिंग कनवल्शनल मॉड्यूल से बना है। एक पूरी तरह से जुड़ी हुई पर तबाउंडिंग बॉक्स इनपुट को 512-आयामी वेक्टर में विस्तारित करती है। फिर इन दो इनपुटों को संयोजित किया जाता है, जिसमें Q का अनुमान लगाने के लिए दो पूरी तरह से जुड़ी हुई परतों का उपयोग किया जाता है।

डबल डीक्यूएन

कार्रवाई चयन और लक्ष्य मूल्य अनुमान दोनों पर अधिकतमीकरण ऑपरेशन को नियोजित करने के कारण मूल DQN विधि को अधिक अनुमान के मुद्दों का सामना करना पड़ता है। इसे संबोधित करने के लिए, डबल क्यू-लर्निंग को एक समाधान के रूप में पेश किया गया है। यह दृष्टिकोण अधिकतमीकरण ऑपरेटरों को दो नेटवर्कों में विभाजित करके अति मूल्यांकन त्रुटियों को कम करता है: Qtarget और Qeval. दोनों नेटवर्क की संरचना एकसमान हैं लेकिन पैरामीटर अलग-अलग हैं। Qtarget का उपयोग केवल क्रियाचयन के लिए किया जाता है:

$$d_t = \max Q_{evel}(R_s, b; \theta^-) \quad (4)$$

परिणाम

गतिशील वस्तु पहचान के लिए एक प्रभावी तकनीक के लिए एक कुशल व्यवहार प्रोटोकॉल स्थापित करने की आवश्यकता होती है जो रोबोट को तेजी से उपयुक्त अवलोकन परिप्रेक्ष्य प्राप्त करने की अनुमति देता है। इस चुनौती से निपटने के लिए तैयार डी आर एल ढांचा तैयार करके इसे हासिल किया जा सकता है। CNN और BiLSTM मॉडल को मिलाकर और उन्हें प्राथमिकता

वाले अनुभव रीप्ले के साथ प्रशिक्षित करके एक कुशल व्यवहार प्रोटोकॉल विकसित किया गया था। प्रत्येक चरण पर कार्यो की भविष्यवाणी का मार्ग दर्शन करना। तालिका 2 और चित्र 5 विभिन्न डेटा सेटों में सीएनएन-एलएसटीएम और सीएनएन-बीआईएल एसटीएम मॉडल के प्रदर्शन का तुलनात्मक विश्लेषण प्रस्तुत करते हैं, जो युग समय, भविष्यवाणी हानि और पैरामीटर्स जैसे मेट्रिक्स पर ध्यान केंद्रित करते हैं। एपोक टाइम प्रत्येक प्रशिक्षण पुनरावृत्ति को पूरा करने के लिए मॉडल के लिए आवश्यक अवधि को दर्शाता है। व्यावहारिक परिदृश्यों में, त्वरित कार्य प्रतिक्रिया और कुशल वास्तविक दुनिया के अनुप्रयोग के लिए प्रशिक्षण समय को कम करना महत्वपूर्ण है।

पकड़ने की प्रारंभिक स्थिति में लचीलापन

प्रारंभ में, डबलडीक्यूएन की प्रभावकारिता का मूल्यांकन मैनिपुलेटर गतियों से प्राप्त स्पर्शडेटा की प्रभावी ढंग से व्याख्या करने की क्षमता के संबंध में किया जाता है। इस जांच में सफलता की परिचालन परिभाषा वस्तु को गिरने के बिना स्थिरता बनाए रखना और अंतिम पकड़ स्थिति प्राप्त करना था। किसी वस्तु को धकेलते समय, कार्य तब सफल माना जाता था जब वह एक ही दिशा में इशारा करते हुए दोनों हाथों की युक्तियों तक पहुँच जाता था। घुमाने के अभ्यास के दौरान जब कोई वस्तु अंगूठे के केंद और तर्जनी के किनारे तक पहुंच जाती है, तो कार्य सफल माना जाता है। हर बार, हाथफीड बैक नियंत्रण प्रणाली के तहत काम करता है, जो डबल डीक्यूएन द्वारा इनपुट के रूप में वर्तमान सेंसर डेटा प्राप्त करने के परिणामस्वरूप उत्पन्न होने वाले संयुक्त कोणों का उपयोग करता है। डबल डीक्यूएन मॉडल ने प्रदर्शित किया कि ट्विस्टमोशन का सफल निष्पादन 40% और पुशमोशन 60% है। इसके विपरीत, डबल डीक्यूएन मॉडल ने 80% में ट्विस्टमोशन और 90 में पुशमोशन सफलता पूर्वक हासिल किया।

तालिका 3. प्रत्येक मैनिपुलेटर गति के लिए विभिन्न डिजाइनों द्वारा प्राप्त सटीकता प्रतिशत।

गति	सीएनएन	आर-सीएनएन	सीएनएन-एलएसटीएम	सीएनएन-बीआईएल एसटीएम
धकेलना	60%	0%	80%	80%
मोड़	40%	30%	90%	30%

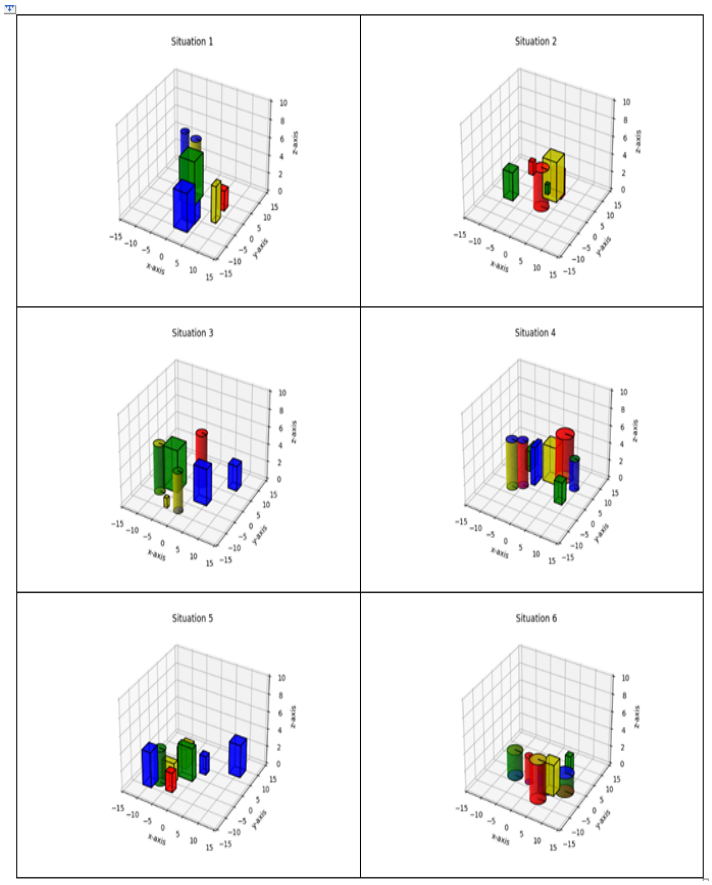
तालिका 3 से पता चलता है कि निर्धारित कार्य सूचना मूल्य की परवाह किए बिना, जो 0.5 की वृद्धि में-1 से 1 तक भिन्न था, प्रभावी इन-हैंड हेरफेर की उपलब्धि लगातार कम से कम 80% देखी गई थी। यह परिणाम पुष्टि करता है कि जिन कार्यो के लिए किसी पूर्व प्रशिक्षण की आवश्यकता नहीं है, उन्हें केवल कार्य पैरामीटर को बदल कर पूरा किया जा सकता है।

इसके अलावा, वस्तु की प्रारंभिक हथियाने की स्थिति के संबंध में प्रत्येक नेटवर्क का लचीलापन निर्धारित किया गया था। यह मूल्यांकन XYZ पोजिशनिंग के साथ चरण का उपयोग कर के आयोजित किया गया था, जिसने x-, y- और z-अक्ष के साथ 0.1 मिमी की वृद्धि में ऑब्जेक्ट की स्थिति के सटीक संशोधन की अनुमति दी थी।

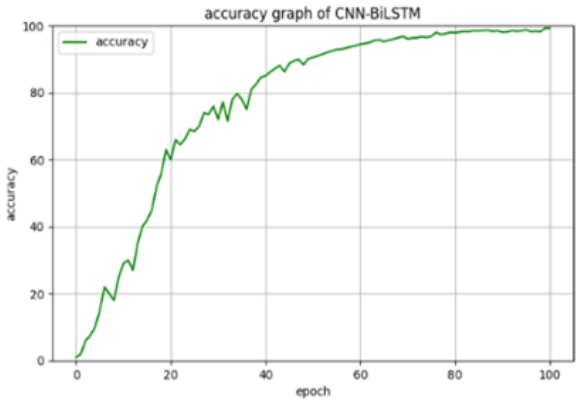
तालिका 4. प्रत्येक परिभाषित कार्य की सफलता दर पर जानकारी

काम	-1	-.6	0	.6	1
सफलता %	80%	80%	90%	90%	90%

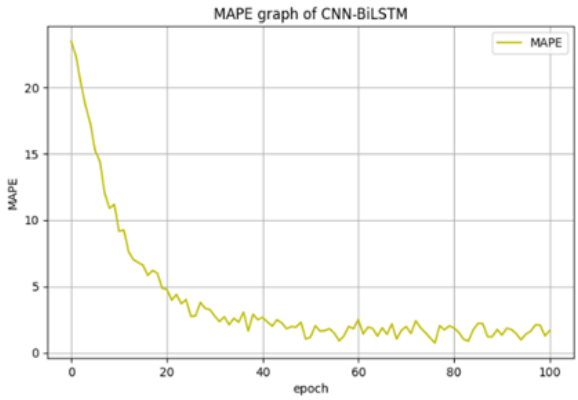
जांच में हेरफेर के परिणामों पर परिभाषित कार्य जानकारी को बदलने के प्रभाव का अध्ययन करने पर ध्यान केंद्रित किया गया। जैसे ही परिभाषित कार्य सूचनामान-1 (ट्विस्ट को इंगित करता है) से 1 (पुश को इंगित करता है) में स्थानांतरित हो गया, हेरफेर प्रक्रिया धीरे-धीरे अनुकूलित हो गई। उल्लेखनीय रूप से, जब कार्य सूचना मानकों 0 पर सेट किया गया था, तो एक मध्यवर्ती मुद्रा अंतिम पकड़ने की स्थिति के रूप में उभरी, जिससे हाथ को वस्तु को गिराए बिना सुरक्षित रूप से पकड़ने की अनुमति मिली।



चित्र 3. वस्तु की पकड़ने की स्थिति



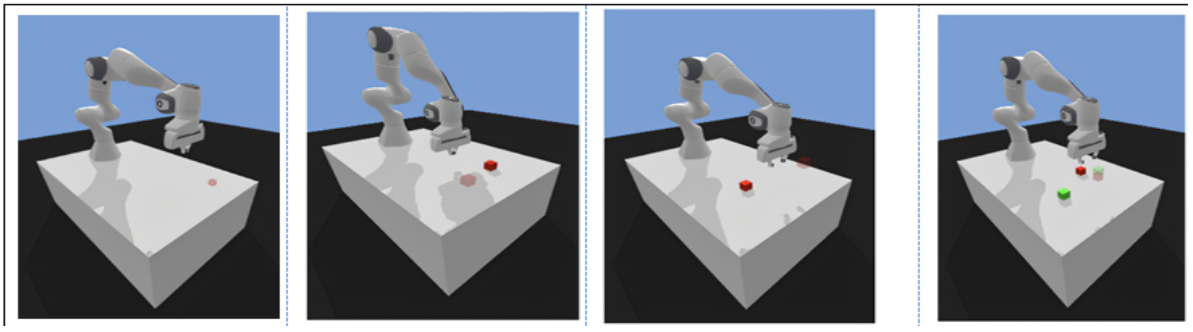
चित्र 4. CNN-BiLSTM का सटीकता ग्राफ



चित्र 5. CNN - BiLSTM का MAPE ग्राफ

तालिका 5. मौजूदा मॉडलों से तुलना

तरीके	सीएनएन	आर-सीएनएन	सीएनएन-एलएसटीएम	शुद्धता %	प्रस्तावित मॉडल
एमएपीई%	7.5	5.87	2.34	80	1.76
आर$_2$	84	89.43	91.7	90	99.02



(a) **पहुँचना** (b) **धकेलना** (c) **उठाइए और जगह पर रखिए** (d) **ढेर**

चित्र 6. लाल और हरी छायांकन लक्ष्य स्थिति से मेल खाती है

अवलोकन स्थान क्षण की स्थिति के अनुसार बदलता रहता है। प्रत्येक कार्य में ग्रिपर का स्थान और वेग देखा जाता है। इसके नियंत्रण का उपयोग करके ग्रिपर की दिशा बदलना संभव नहीं है। इस प्रकार, निम्नलिखित निर्देशांक इसकी संपूर्ण स्थिति को दर्शाते हैं। एक या अधिक वस्तुओं के साथ काम करते समय प्रत्येक वस्तु का स्थान, अभिविन्यास, रैखिक गति और घूर्णी गति सभी अवलोकन स्थान में शामिल होते हैं। इसके अलावा, ग्रिपर का उद्घाटन, या उसकी उंगलियों के बीच का स्थान, अवलोकन स्थान का एक घटक है जब यह एक बंद स्थिति तक सीमित नहीं होता है।

निष्कर्ष

यह शोध सक्रिय वस्तु पहचान की समस्या का समाधान करता है। इस मुद्दे को क्रमिक निर्णय लेने की प्रक्रिया के रूप में तैयार करके, हम एक उपन्यास गहन सुदृढीकरण सीखने की रूपरेखा का प्रस्ताव करते हैं। हमारा दृष्टिकोण विशेष रूप से कन्वेन्शनल न्यूरल नेटवर्क्स (सीएनएन) कोलॉन्ग शॉर्ट-टर्ममेमोरी (एलएसटीएम) नेटवर्क के साथ एकीकृत करता है और इष्टतम कार्रवाई नीति को कुशलतापूर्वक सीखने के लिए प्राथमिकता वाले अनुभव रीप्ले का उपयोग करता है। क्यू-नेटवर्क अध्ययन के अंतिम परिणामों में बहु-हाथ हेरफेर के लिए कार्य जानकारी शामिल है। डबल DQN को इन-हैंड हेरफेर कार्यो के लिए व्यापक स्पर्श और समय-श्रृंखला डेटा का प्रबंधन करने के लिए नियोजित किया गया था। घुमाव और धक्का देने वाली क्रियाओं की आवश्यकता वाली गतिविधियों के लिए कार्य जानकारी शामिल की गई थी। डीडीक्यूएन जटिल स्पर्श संबंधी जानकारी को संभालने के लिए पर्याप्त साबित हुआ, जिससे प्रारंभिक समझ की स्थिति को अपनाने में मजबूती आई। विशिष्ट कार्य जानकारी को शामिल करने से ऐसी जानकारी की कमी वाले परिदृश्यों की तुलना में इन-हैंड हेरफेर में सफलता दर अधिक हो गई। इसके अतिरिक्त, इस परिभाषित कार्य जानकारी का उपयोग पहले से प्रशिक्षित दो कार्यो के बीच स्थित नए कार्यो को बनाने के लिए किया गया था। अध्ययन दर्शाता है कि कार्य मापदंडों को संशोधित करके अंतिम समझ की स्थिति निर्धारित करना संभव है।

संदर्भ

1. H. Chen, W. Wan, M. Matsushita, T. Kotaka and K. Harada, "Robotic Test Tube, Rearrangement Using Combined Reinforcement Learning and Motion Planning," Jan. (2024), [Online]. Available: http://arxiv.org/abs/2401.09772.
2. Y. Hou, J. Li and I. M. Chen, "Self-Supervised Antipodal Grasp Learning With Fine-Grained Grasp Quality Feedback in Clutter," IEEE Transactions on Industrial Electronics, vol. 71, No. 4, pp. 3853–3861, Apr. (2024), doi: 10.1109/TIE.2023.3274854.
3. C. Acar, K. Binici, A. Tekirdag and Y. Wu, "Visual-Policy Learning Through Multi-Camera View to Single-Camera View Knowledge Distillation for Robot Manipulation Tasks," IEEE Robot Autom Lett, vol. 9, no. 1, pp. 691–698, Jan. (2024), doi: 10.1109/LRA.2023.3336245.
4. E. Okafor, M. Oyedeji, and M. Alfarraj, "Deep reinforcement learning with light-weight vision model for sequential robotic object sorting,"

Journal of King Saud University – Computer and Information Sciences, 101896, Jan. (2023), doi: 10.1016/j.jksuci.2023.101896.

5. F. Liu, F. Sun, B. Fang, X. Li, S. Sun and H. Liu, "Hybrid Robotic Grasping With a Soft Multimodal Gripper and a Deep Multistage Learning Scheme," IEEE Transactions on Robotics, Vol. 39, Issue 3 (2023), doi: 10.1109/TRO.2023.3238910.
6. D. Han, B. Mulyana, V. Stankovic, and S. Cheng, "A Survey on Deep Reinforcement Learning Algorithms for Robotic Manipulation," Sensors, vol. 23, no. 7. MDPI, Apr. 01, (2023). Doi: 10.3390/s23073762.
7. A. Lobbezoo and H. J. Kwon, "Simulated and Real Robotic Reach, Grasp, and Pick-andPlace Using Combined Reinforcement Learning and Traditional Controls," Robotics, vol. 12, no. 1, Feb. (2023), doi: 10.3390/robotics12010012.
8. M. H. Sayour, S. E. Kozhaya, and S. S. Saab, "Autonomous Robotic Manipulation: Real-Time, Deep-Learning Approach for Grasping of Unknown Objects," Journal of Robotics, vol. 25. (2022), doi: 10.1155/2022/2585656.
9. J. Ibarz, J. Tan, C. Finn, M. Kalakrishnan, P. Pastor, and S. Levine, "How to train your robot With deep reinforcement learning: lessons we have learned," International Journal of Robotics Research, vol. 40, no. 4–5, pp. 698–721, Apr. (2021), doi: 10.1177/0278364920987859.
10. J. Hua, L. Zeng, G. Li, and Z. Ju, "Learning for a robot: Deep reinforcement learning, Imitation learning, transfer learning," Sensors (Switzerland), vol. 21, no. 4. MDPI AG, pp. 1–21, Feb. 02, (2021), doi: 10.3390/s21041278.
11. P. Shukla, H. Kumar, and G. C. Nandi, "Robotic grasp manipulation using evolutionary Computing and deep reinforcement learning," Intell Serv Robot, vol. 14, no. 1, pp. 61–77, Mar. (2021), doi: 10.1007/s11370-020-00342-7.
12. M. Q. Mohammed, K. L. Chung, and C. S. Chyi, "Review of deep reinforcement learningbased object grasping: Techniques, open challenges, and recommendations," IEEE Access, vol. 8, pp. 178450–178481, (2020), doi: 10.1109/ACCESS.2020.3027923.
13. A. Shahid, L. Roveda, D. Piga, and F. Braghin, "Learning Continuous Control Actions for Robotic Grasping with Reinforcement Learning," in Conference Proceedings – IEEE International Conference on Systems, Man and Cybernetics, Institute of Electrical and Electronics Engineers Inc., Oct. 2020, pp. 4066–4072. Doi: 10.1109/SMC42975.2020.9282951.
14. M. Q. Mohammed, K. L. Chung, and C. S. Chyi, "Review of deep reinforcement learning based object grasping: Techniques, open challenges, and recommendations," IEEE Access, vol. 8, pp. 178450–178481, 2020, doi: 10.1109/ACCESS.2020.3027923.
15. IEEE International Conference on Robotics and Automation (ICRA) : May 21-25, (2018), Brisbane, Australia.
16. IEEE/RSJ International Conference on Intelligent Robots and Systems : Towards a Robotic Society : October, 1-5, (2018), Madrid, Spain, Madrid Municipal Conference Centre.

❒