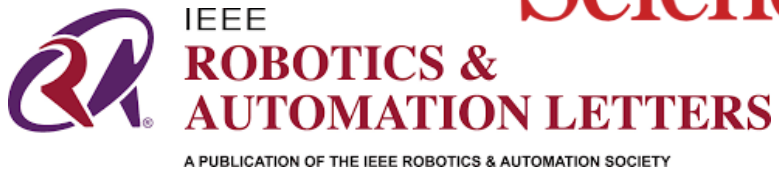




DEEPRob

Discussion 5
 How to Read Deep Learning Papers
 University of Michigan | Department of Robotics

 **OCTOBER 1 - 5, 2023**
 IEEE/RSJ International Conference on
 Intelligent Robots and Systems



Science Robotics





Agenda

- The importance of reading papers
- How to approach research papers in deep learning
- Discussion of AlexNet, PoseCNN and NeRF

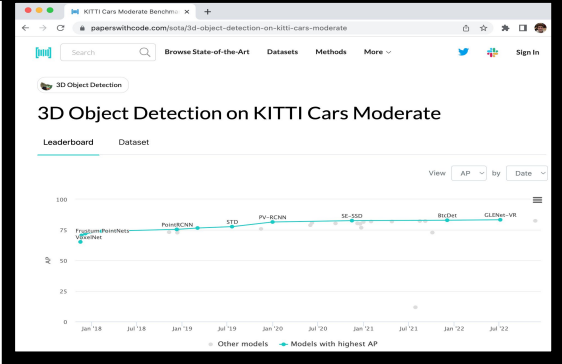
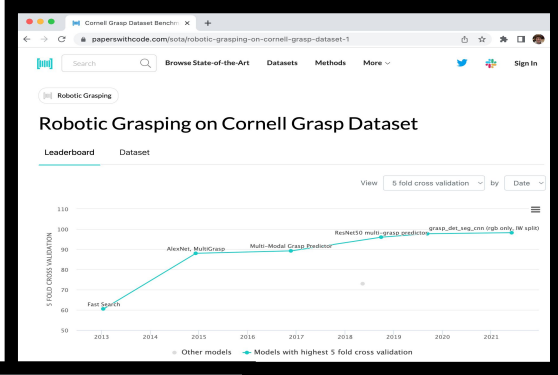
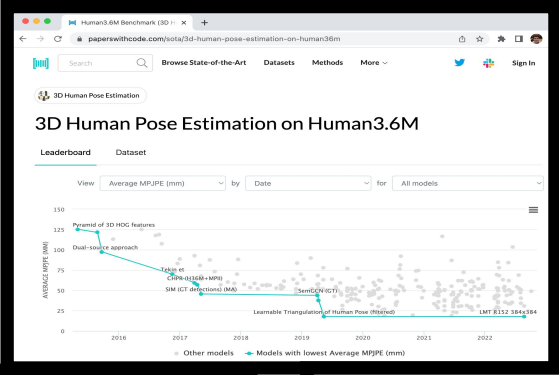
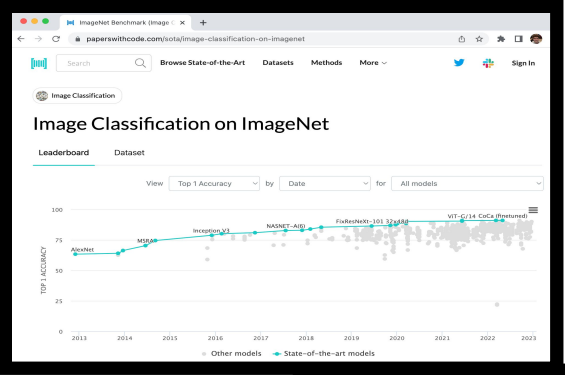


Reading Papers is an Important Skill

- Applied Side
 - Practitioners want state of the art performance
 - Look to academia for what exists and how it can be replicated
- Research Side
 - Understand the field as a way to find ideas for contributing
 - New datasets, techniques, methods defined by research community



State of the Art is Always Changing



Datasets

7,656 machine learning datasets

Filter by Modality: Images (2214), Texts (2048), Videos (707), Audio (458)

Filter by Task: Question Answering (333), Semantic Segmentation (250)

7656 dataset results

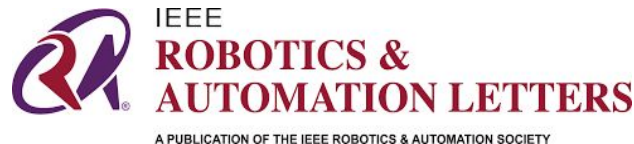
- CIFAR-10**: The CIFAR-10 dataset (Canadian Institute for Advanced Research, 10 classes) is a subset of the Tiny Images dataset and consists of 60000 32x32 color images. T...
- ImageNet**: The ImageNet dataset contains 14,197,122 annotated images according to the WordNet hierarchy. Since 2010 the dataset is used in the ImageNet Large Scale...
- COCO (Microsoft Common Objects in Context)**: The MS COCO (Microsoft Common Objects in Context) dataset is a large-scale object detection, segmentation, key-point detection, and captioning dataset. T...



Where to Look for Deep Learning Papers in Robotics?



Science Robotics





Where to Look for Deep Learning Papers in Robotics?

The screenshot shows the arXiv.org website interface. At the top, it says "Cornell University" and "We gratefully acknowledge support from the Simons Foundation and University of Michigan". The arXiv logo is prominent. A search bar is located on the right with a "Log" button. Below the search bar, there's a "Subject search and browse:" section with a dropdown menu set to "Physics" and a "Search" button. There are also buttons for "Form Interface" and "Catchup". A "News" section follows, with a link to "robots beware". The "Physics" section lists various sub-fields like Astrophysics, Condensed Matter, and High Energy Physics. A "COVID-19 Quick Links" box highlights SARS-CoV-2 preprints from arXiv and medRxiv, with an important disclaimer that these are not peer-reviewed.



Publishing Never Stops

Cornell University We gratefully acknowledge support from the Simons Foundation and member institutions.

arXiv > cs > cs.RO Search... All Fields Search
Help | Advanced Search

Robotics

Authors and titles for recent submissions

- Tue, 6 Feb 2024
- Mon, 5 Feb 2024
- Fri, 2 Feb 2024
- Thu, 1 Feb 2024
- Wed, 31 Jan 2024

[total of 162 entries: 1-25 | 26-50 | 51-75 | 76-100 | ... | 151-162]
[showing 25 entries per page: fewer | more | all]

Tue, 6 Feb 2024 (showing first 25 of 44 entries)

[1] [arXiv:2402.03156 \[pdf, other\]](#)
DogSurf: Quadruped Robot Capable of GRU-based Surface Recognition for Blind Person Navigation
Artem Bazhenov, Vladimir Berman, Sergei Satsevich, Olga Shalopanova, Miguel Altamirano Cabrera, Artem Lykov, Dzmityry Tsetserukou
Comments: This paper has been accepted for publication at the HRI2024 conference
Subjects: **Robotics (cs.RO)**; Machine Learning (cs.LG)

[2] [arXiv:2402.03135 \[pdf, other\]](#)
GPU-Accelerated 3D Polygon Visibility Volumes for Synergistic Perception and Navigation
Andrew Willis, Collin Hague, Artur Wolek, Kevin Brink
Subjects: **Robotics (cs.RO)**; Computational Geometry (cs.CG); Computer Vision and Pattern Recognition (cs.CV)

[3] [arXiv:2402.03125 \[pdf, other\]](#)
Shape Manipulation of Bevel-Tip Needles for Prostate Biopsy Procedures: A Comparison of Two Resolved-Rate Controllers
Yanzhou Wang, Lidia Al-Zogbi, Jiawei Liu, Lauren Shepard, Ahmed Ghazi, Junichi Tokuda, Simon Leonard, Axel Krieger, Iulian Iordachita



How Research a Topic

- Related Works
- Demo Videos
- Git Repos
- YouTube Explainers



Example

LLM + Robotics



Example: Vox Poser

<https://arxiv.org/abs/2307.05973>

VoxPoser: Composable 3D Value Maps for Robotic Manipulation with Language Models

Wenlong Huang¹, Chen Wang¹, Ruohan Zhang¹, Yunzhu Li^{1,2}, Jiajun Wu¹, Li Fei-Fei¹
¹Stanford University ²University of Illinois Urbana-Champaign

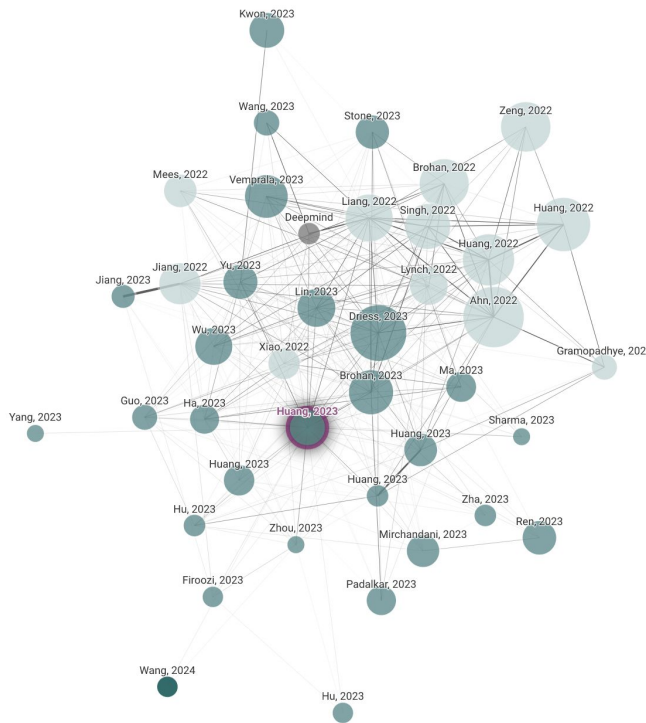
Abstract: Large language models (LLMs) are shown to possess a wealth of actionable knowledge that can be extracted for robot manipulation in the form of reasoning and planning. Despite the progress, most still rely on pre-defined motion primitives to carry out the physical interactions with the environment, which remains a major bottleneck. In this work, we aim to synthesize robot trajectories, i.e., a dense sequence of 6-DoF end-effector waypoints, for a large variety of manipulation tasks given an *open-set of instructions* and an *open-set of objects*. We achieve this by first observing that LLMs excel at inferring affordances and constraints given a free-form language instruction. More importantly, by leveraging their code-writing capabilities, they can interact with a vision-language model (VLM) to compose 3D value maps to ground the knowledge into the observation space of the agent. The composed value maps are then used in a model-based planning framework to *zero-shot* synthesize closed-loop robot trajectories with robustness to dynamic perturbations. We further demonstrate how the proposed framework can benefit from online experiences by efficiently learning a dynamics model for scenes that involve contact-rich interactions. We present a large-scale study of the proposed method in both simulated and real-robot environments, showcasing the ability to perform a large variety of everyday manipulation tasks specified in free-form natural language. Videos and code at [voxposer.github.io](https://github.com/voxposer).

Keywords: Manipulation, Large Language Models, Model-based Planning





Related Works





Videos and Git Repo

<https://voxposer.github.io/>





How to Read Deep Learning Research Papers?

Everyone develops their own style over time



Questions I Consider When Reading a Paper

What is the primary field and subfield of the work?



Questions I Consider When Reading a Paper

What is the primary field and subfield of the work?

What problem are the authors trying to address?



Questions I Consider When Reading a Paper

What is the primary field and subfield of the work?

What problem are the authors trying to address?

What are the primary claims and contributions?



Questions I Consider When Reading a Paper

What is the primary field and subfield of the work?

What problem are the authors trying to address?

What are the primary claims and contributions?

What are the key results?



Questions I Consider When Reading a Paper

What is the primary field and subfield of the work?

What problem are the authors trying to address?

What are the primary claims and contributions?

What are the key results?

What progress have other researchers made on this problem?



Questions I Consider When Reading a Paper

What is the primary field and subfield of the work?

What problem are the authors trying to address?

What are the primary claims and contributions?

What are the key results?

What progress have other researchers made on this problem?

How were these results achieved?



Questions I Consider When Reading a Paper

What is the primary field and subfield of the work?

What problem are the authors trying to address?

What are the primary claims and contributions?

What are the key results?

What progress have other researchers made on this problem?

How were these results achieved?

Using which techniques evaluated under which methods?



Questions I Consider When Reading a Paper

What is the primary field and subfield of the work?

What problem are the authors trying to address?

What are the primary claims and contributions?

What are the key results?

What progress have other researchers made on this problem?

How were these results achieved?

Using which techniques evaluated under which methods?

What problems, questions, or findings could be expanded on as future work?



Questions I Consider When Reading a Paper

What is the primary field and subfield of the work?

What problem are the authors trying to address?

What are the primary claims and contributions?

What are the key results?

What progress have other researchers made on this problem?

How were these results achieved?

Using which techniques evaluated under which methods?

What problems, questions, or findings could be expanded on as future work?



Discussion: PoseCNN

<https://arxiv.org/abs/1711.00199>



Discussion: NeRF

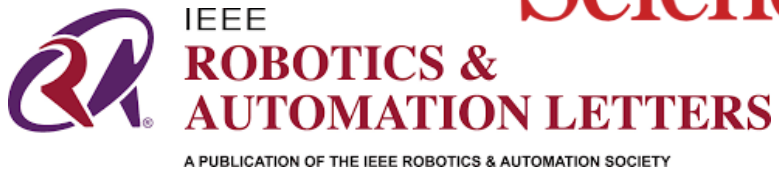
<https://arxiv.org/abs/2003.08934>



DEEPRob

Discussion 5
 How to Read Deep Learning Papers
 University of Michigan | Department of Robotics

 **OCTOBER 1 - 5, 2023**
 IEEE/RSJ International Conference on
 Intelligent Robots and Systems



Science Robotics

