Building Continually Improving and Reliable Autonomy through Human-Robot Collaboration

Huihan Liu
The University of Texas at Austin

I. Introduction

In recent years, learning-based robots have shown remarkable capabilities in controlled research settings. Yet, they often fail to generalize reliably in diverse, unstructured environments, undermining their real-world deployability. To address the challenge, one line of prior work focuses on improving robot generalization by scaling up data and policy models [1, 2, 3, 17, 41, 49]. These methods collect large and diverse datasets from human demonstrators [11, 20, 38] and train advanced policy architectures [6, 30] before deploying the robots to real-world environments. While such scaling improves robot performance, these methods remain brittle when faced with out-of-distribution inputs and unpredictable corner cases common in the real world. When failures occur, they lack mechanisms for detection and recovery, undermining human trust. To ensure safe deployment, research in human-robot teaming [37, 44] addresses reliability through shared autonomy between robots and humans [18, 40]. This framework enables humans to prevent or correct robot errors, ensuring safety; however, it does not inherently improve robot performance over time, and as a result, human operators have a high workload from frequent interventions.

Given these limitations, my research focuses on a central question: **How can we ensure reliable deployment of learning-based robots, while continuously improving their performance during deployment?** To address this, I propose to **build continually improving and reliable autonomy through human-robot collaboration**. In this paradigm, robots are deployed with human teaming and *learn on the job*, *i.e.,* learn during deployment from humans. This approach not only ensures safe and reliable deployments via human-robot teaming but also continuously enhances robot autonomy to reduce human workload over time.

I identify three key challenges in realizing this paradigm: (I) improving robot policy from human feedback constantly from deployment; (II) enabling robots to self-monitor and proactively detect errors, hence minimizing the need for constant human oversight; and (III) developing intention-aware robots that adapt to human actions for effective teaming. To address these challenges, I contribute in three corresponding areas: (i) a *human-in-the-loop framework* for lifelong learning from deployment [25]; (ii) an algorithm for *robot self-monitoring* that detects errors during deployment [26, 27]; and (iii) an *intention-aware* system for human-robot teaming that enhances collaboration by interpreting human intents [28].

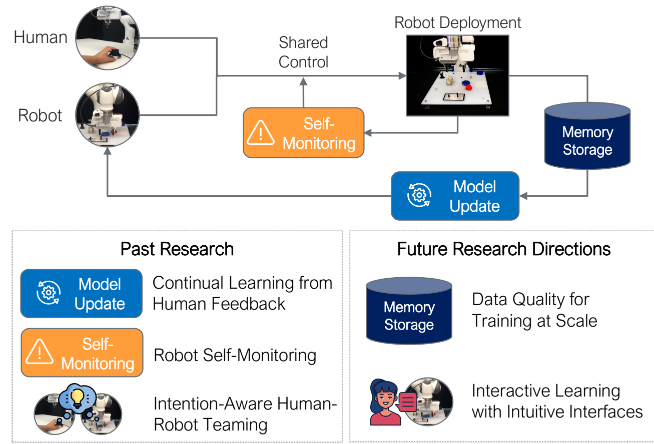

Fig. 1. **Research Overview.** My research improves robot autonomy over long-term deployments through human-robot teaming. To realize this paradigm, I have contributed in three main areas: (i) continual learning with human feedback, (ii) robot self-monitoring, and (iii) intention-aware human-robot teaming. In the future, I will explore data quality for robot learning at scale and interactive learning with intuitive interfaces.

Together, they advance my research vision of reliable learning-based robots deployable in human-centered environments.

II. Past & Current Research

A. Continual Learning from Human Feedback

My work [25] established the framework of improving robot autonomy over long-term deployments through human-robot teaming. This framework is characterized by a deployment-training loop (see Fig. 1). During deployment, a human operator can monitor and intervene in the robot's policy execution. The human can take over control when necessary and handle challenging situations to ensure safe and reliable task execution. The collected human feedback data is used to continually improve the robot policy. As the training and deployment loop repeats, the robot makes fewer mistakes and becomes more autonomous. This framework can allow the robot systems to continuously learn and improve while being reliable.

A key learning challenge of this framework originates from the ever-growing dataset of robot rollouts and human interventions, which includes suboptimal robot behaviors. Learning from such deployment data requires us to use it for policy updates selectively. Our key insight is that we can assess the importance of varying training data based on human interventions for policy learning. To this end, we re-weight training samples based on the presence and timing of interventions: transitions preceding an intervention are labeled as "low-quality", as the human believes the robot is about to fail; while human demonstrations or interventions are deemed "high-quality" because they happen at critical states where help is needed. We then apply weighted behavioral cloning [22, 34, 43, 53, 57, 60] to reinforce corrective behaviors and avoid replicating errors, which leads to consistent autonomous policy improvement over multiple rounds of training and deployment. As the pioneering work to establish a framework for learning from humans during deployment, our research has influenced subsequent works in continual policy improvement [5, 19, 31, 48], scalable data collection [32, 51, 55, 59] and robot foundation model training [9, 21, 49].

B. Robot Self-Monitoring

While our framework [25] ensures reliable deployment, constant human supervision incurs human workloads. Rather than relying on continuous human monitoring, can humans intervene only when necessary? To tackle the problem, my work [26, 27] incorporates a robot self-monitoring mechanism into our human-in-the-loop framework in Section II-A, allowing the robots to monitor and predict errors autonomously.

Our design follows two conceptual ideas: First, it adopts a *model-based* approach, which trains a predictive model of the environment dynamics for failure prediction. We train a visual world model [4, 14], which simulates future policy rollouts and predicts upcoming failures. The visual world model is pretrained in diverse robot environments, enabling the sharing of learned representations of downstream anomaly predictors across various tasks. Second, it uses an *intervention-informed* approach to train its error predictor using human interventions over long-term deployments. The algorithm harnesses the inherent structure of human interventions to continually learn an error predictor without encountering explicit failures, thus ensuring reliable task execution. As the robot improves its autonomy during deployment, the human-robot interaction evolves. Consequently, the failure labels are continuously updated to refine the error predictor, aligning it with the changing human risk assessments. We instantiate these ideas in a multi-task interactive robot fleet learning setting, where a multi-task policy is deployed across large robot fleets, and a runtime monitoring mechanism efficiently supervises multiple tasks. Our work contributes to scalable and reliable robot autonomy by influencing a line of research in failure prediction [35, 36, 58] and deployment-time refinement [8, 10, 39, 56].

C. Intention-Aware Human-Robot Teaming

In our above framework [25, 26, 27], humans mainly only intervene upon errors. As robots become more integrated into daily life, robots and humans will collaborate on tasks as teammates. In these scenarios, humans must maintain a *sense of agency* [7, 15, 29, 54]. Effective human-robot teaming, therefore, requires that robots *anticipate* human intent to enhance task performance and user satisfaction.

To this end, we propose [28], a human-robot teaming system that enables robots to infer diverse human intents and collaborate on long-horizon, open-world tasks. First, to infer diverse intents, it uses a visual language model (VLM) that interprets user teleoperation signals with commonsense reasoning, using self-consistency [52] for confidence estimation. Second, to assist users with open-world tasks, it uses a diverse library of parameterized skills [45] that execute diverse contact-rich behaviors once the human's intent is inferred. Finally, by running human control and VLM inference in parallel, the system minimizes delays and enhances reactivity. This human-robot teaming paradigm complements our prior work [25, 26, 27] in Section II-A, II-B, expanding the scope of reliable robot deployment to human-centric environments.

III. Future Research

While my previous work has established a framework for deploying learning-enabled robots, achieving general-purpose, reliable robotic systems in the human world requires addressing additional challenges. To further advance my mission, I identify two future research directions:

A. Data Quality for Robot Learning at Scale

In the future, as robot policies continue to scale, deploying large-scale robot foundation models will generate massive datasets of uneven quality [16]. Automatically curating such datasets will be crucial both for reducing the memory storage burden and for improving the model performance and efficiency of training, as has been observed by the large vision and language model training [33, 42, 47]. Deployment datasets demonstrate two properties: 1) they have a large amount of redundant and repetitive data; 2) they also contain suboptimal data that hurts learning. To address the challenges, I plan to investigate data curation for large robot foundation models with 1) deduplication [23] of redundant datasets, and 2) automatic removal [12] of suboptimal datapoints by assessing the data quality. With data curation, I aim to improve computational efficiency for training and to optimize policy performance.

B. Interactive Learning with Intuitive Interfaces

In my past research, humans provide feedback by direct intervention. As robots are being deployed into households, the structured intervention will shift toward more intuitive forms of interaction. End users often prefer natural feedback modalities like language corrections or hand gestures rather than direct control [46, 50]. Therefore, developing algorithms to improve robot policies using these alternative inputs is essential. As a first step, I will explore policy updates driven by human language feedback, leveraging pretrained VLMs to bridge the gap between language and low-level motor actions. Furthermore, I aim to investigate learning from human hand gesture corrections by leveraging the similar kinematic form factor between humanoids and humans [13, 24] for better skill transfer. My long-term direction is to build a unified visual-language-action interface capable of integrating diverse forms of human feedback. Robots that learn from humans intuitively can enable collaboration that enriches and empowers human lives, ultimately augmenting our capacities for better living.

References

[1] Konstantinos Bousmalis, Giulia Vezzani, Dushyant Rao, Coline Devin, Alex X. Lee, Maria Bauza, Todor Davchev, Yuxiang Zhou, Agrim Gupta, Akhil Raju, Antoine Laurens, Claudio Fantacci, Valentin Dalibard, Martina Zambelli, Murilo Martins, Rugile Pevceviciute, Michiel Blokzijl, Misha Denil, Nathan Batchelor, Thomas Lampe, Emilio Parisotto, Konrad Żołna, Scott Reed, Sergio Gómez Colmenarejo, Jon Scholz, Abbas Abdolmaleki, Oliver Groth, Jean-Baptiste Regli, Oleg Sushkov, Tom Rothörl, José Enrique Chen, Yusuf Aytar, Dave Barker, Joy Ortiz, Martin Riedmiller, Jost Tobias Springenberg, Raia Hadsell, Francesco Nori, and Nicolas Heess. Robocat: A self-improving generalist agent for robotic manipulation, 2023.

[2] Anthony Brohan, Noah Brown, Justice Carbajal, Yevgen Chebotar, Joseph Dabis, Chelsea Finn, Keerthana Gopalakrishnan, Karol Hausman, Alex Herzog, Jasmine Hsu, Julian Ibarz, Brian Ichter, Alex Irpan, Tomas Jackson, Sally Jesmonth, Nikhil J Joshi, Ryan Julian, Dmitry Kalashnikov, Yuheng Kuang, Isabel Leal, Kuang-Huei Lee, Sergey Levine, Yao Lu, Utsav Malla, Deeksha Manjunath, Igor Mordatch, Ofir Nachum, Carolina Parada, Jodilyn Peralta, Emily Perez, Karl Pertsch, Jornell Quiambao, Kanishka Rao, Michael Ryoo, Grecia Salazar, Pannag Sanketi, Kevin Sayed, Jaspiar Singh, Sumedh Sontakke, Austin Stone, Clayton Tan, Huong Tran, Vincent Vanhoucke, Steve Vega, Quan Vuong, Fei Xia, Ted Xiao, Peng Xu, Sichun Xu, Tianhe Yu, and Brianna Zitkovich. Rt-1: Robotics transformer for real-world control at scale, 2022.

[3] Anthony Brohan, Noah Brown, Justice Carbajal, Yevgen Chebotar, Xi Chen, Krzysztof Choromanski, Tianli Ding, Danny Driess, Avinava Dubey, Chelsea Finn, Pete Florence, Chuyuan Fu, Montse Gonzalez Arenas, Keerthana Gopalakrishnan, Kehang Han, Karol Hausman, Alexander Herzog, Jasmine Hsu, Brian Ichter, Alex Irpan, Nikhil Joshi, Ryan Julian, Dmitry Kalashnikov, Yuheng Kuang, Isabel Leal, Lisa Lee, Tsang-Wei Edward Lee, Sergey Levine, Yao Lu, Henryk Michalewski, Igor Mordatch, Karl Pertsch, Kanishka Rao, Krista Reymann, Michael Ryoo, Grecia Salazar, Pannag Sanketi, Pierre Sermanet, Jaspiar Singh, Anikait Singh, Radu Soricut, Huong Tran, Vincent Vanhoucke, Quan Vuong, Ayzaan Wahid, Stefan Welker, Paul Wohlhart, Jialin Wu, Fei Xia, Ted Xiao, Peng Xu, Sichun Xu, Tianhe Yu, and Brianna Zitkovich. Rt-2: Vision-language-action models transfer web knowledge to robotic control, 2023.

[4] Tim Brooks, Bill Peebles, Connor Holmes, Will DePue, Yufei Guo, Li Jing, David Schnurr, Joe Taylor, Troy Luhman, Eric Luhman, Clarence Ng, Ricky Wang, and Aditya Ramesh. Video generation models as world simulators. 2024.

[5] Yuhui Chen, Shuai Tian, Shugao Liu, Yingting Zhou, Haoran Li, and Dongbin Zhao. Conrft: A reinforced fine-tuning method for vla models via consistency policy. 2025. URL https://api.semanticscholar.org/CorpusID:276250518.

[6] Cheng Chi, Zhenjia Xu, Siyuan Feng, Eric Cousineau, Yilun Du, Benjamin Burchfiel, Russ Tedrake, and Shuran Song. Diffusion policy: Visuomotor policy learning via action diffusion, 2024.

[7] Maggie A. Collier, Rithika Narayan, and Henny Admoni. The sense of agency in assistive robotics using shared autonomy, 2025. URL https://arxiv.org/abs/2501.07462.

[8] Yinpei Dai, Jayjun Lee, Nima Fazeli, and Joyce Chai. Racer: Rich language-guided failure recovery policies for imitation learning. *ArXiv*, abs/2409.14674, 2024. URL https://api.semanticscholar.org/CorpusID:272826633.

[9] Ria Doshi, Homer Walke, Oier Mees, Sudeep Dasari, and Sergey Levine. Scaling cross-embodied learning: One policy for manipulation, navigation, locomotion and aviation. *ArXiv*, abs/2408.11812, 2024. URL https://api.semanticscholar.org/CorpusID:271915701.

[10] Maximilian Du, Alexander Khazatsky, Tobias Gerstenberg, and Chelsea Finn. To err is robotic: Rapid value-based trial-and-error during deployment. *ArXiv*, abs/2406.15917, 2024. URL https://api.semanticscholar.org/CorpusID:270702674.

[11] Frederik Ebert, Yanlai Yang, Karl Schmeckpeper, Bernadette Bucher, Georgios Georgakis, Kostas Daniilidis, Chelsea Finn, and Sergey Levine. Bridge data: Boosting generalization of robotic skills with cross-domain datasets, 2021.

[12] Alex Fang, Albin Madappally Jose, Amit Jain, Ludwig Schmidt, Alexander Toshev, and Vaishaal Shankar. Data filtering networks, 2023. URL https://arxiv.org/abs/2309.17425.

[13] Zipeng Fu, Qingqing Zhao, Qi Wu, Gordon Wetzstein, and Chelsea Finn. Humanplus: Humanoid shadowing and imitation from humans, 2024. URL https://arxiv.org/abs/2406.10454.

[14] David Ha and Jürgen Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2018.

[15] Patrick Haggard. Sense of agency in the human brain. *Nature Reviews Neuroscience*, 18(4):196–207, 2017.

[16] Joey Hejna, Chethan Bhateja, Yichen Jian, Karl Pertsch, and Dorsa Sadigh. Re-mix: Optimizing data mixtures for large scale imitation learning, 2024. URL https://arxiv.org/abs/2408.14037.

[17] Eric Jang, Alex Irpan, Mohi Khansari, Daniel Kappler, Frederik Ebert, Corey Lynch, Sergey Levine, and Chelsea Finn. Bc-z: Zero-shot task generalization with robotic imitation learning, 2022.

[18] Shervin Javdani, Siddhartha S. Srinivasa, and J. Andrew Bagnell. Shared autonomy via hindsight optimization. In *RSS*, volume 2015, 2015.

[19] Yunfan Jiang, Chen Wang, Ruohan Zhang, Jiajun Wu, and Fei-Fei Li. Transic: Sim-to-real policy transfer by learning from online correction. *ArXiv*, abs/2405.10315, 2024. URL https://api.semanticscholar.org/CorpusID:

269791461.

[20] Alexander Khazatsky, Karl Pertsch, Suraj Nair, Ashwin Balakrishna, Sudeep Dasari, Siddharth Karamcheti, Soroush Nasiriany, Mohan Kumar Srirama, Lawrence Yunliang Chen, et al. Droid: A large-scale in-the-wild robot manipulation dataset, 2024.

[21] Moo Jin Kim, Karl Pertsch, Siddharth Karamcheti, Ted Xiao, Ashwin Balakrishna, Suraj Nair, Rafael Rafailov, Ethan Foster, Grace Lam, Pannag R. Sanketi, Quan Vuong, Thomas Kollar, Benjamin Burchfiel, Russ Tedrake, Dorsa Sadigh, Sergey Levine, Percy Liang, and Chelsea Finn. Openvla: An open-source vision-language-action model. *ArXiv*, abs/2406.09246, 2024. URL https://api.semanticscholar.org/CorpusID:270440391.

[22] Ilya Kostrikov, Ashvin Nair, and Sergey Levine. Offline reinforcement learning with implicit q-learning. In *ICLR*, 2021.

[23] Katherine Lee, Daphne Ippolito, Andrew Nystrom, Chiyuan Zhang, Douglas Eck, Chris Callison-Burch, and Nicholas Carlini. Deduplicating training data makes language models better. In *Annual Meeting of the Association for Computational Linguistics*, 2021. URL https://api.semanticscholar.org/CorpusID:235829052.

[24] Jinhan Li, Yifeng Zhu, Yuqi Xie, Zhenyu Jiang, Mingyo Seo, Georgios Pavlakos, and Yuke Zhu. Okami: Teaching humanoid robots manipulation skills through single video imitation, 2024. URL https://arxiv.org/abs/2410.11792.

[25] Huihan Liu, Soroush Nasiriany, Lance Zhang, Zhiyao Bao, and Yuke Zhu. Robot learning on the job: Human-in-the-loop autonomy and learning during deployment. In *Robotics: Science and Systems (RSS)*, 2023.

[26] Huihan Liu, Shivin Dass, Roberto Martín-Martín, and Yuke Zhu. Model-based runtime monitoring with interactive imitation learning. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2024.

[27] Huihan Liu, Yu Zhang, Vaarij Betala, Evan Zhang, James Liu, Crystal Ding, and Yuke Zhu. Multi-task interactive robot fleet learning with visual world models. In *Conference on Robot Learning (CoRL)*, 2024.

[28] Huihan Liu, Rutav Shah, Shuijing Liu, Jack Pittenger, Mingyo Seo, Yuchen Cui, Yonatan Bisk, Roberto Martín-Martín, and Yuke Zhu. Casper: Inferring diverse intents for assistive teleoperation with vision language models, 2025.

[29] Janeen D Loehr. The sense of agency in joint action: An integrative review. *Psychonomic Bulletin & Review*, 29 (4):1089–1117, 2022.

[30] Ajay Mandlekar, Danfei Xu, Josiah Wong, Soroush Nasiriany, Chen Wang, Rohun Kulkarni, Li Fei-Fei, Silvio Savarese, Yuke Zhu, and Roberto Martín-Martín. What matters in learning from offline human demonstrations for robot manipulation, 2021.

[31] Russell Mendonca, Emmanuel Panov, Bernadette Bucher, Jiuguang Wang, and Deepak Pathak. Continuously improving mobile manipulation with autonomous real-world rl. *ArXiv*, abs/2409.20568, 2024. URL https://api.semanticscholar.org/CorpusID:272986643.

[32] Suvir Mirchandani, Suneel Belkhale, Joey Hejna, Md Evelyn Choi, Sazzad Islam, and Dorsa Sadigh. So you think you can scale up autonomous robot data collection? *ArXiv*, abs/2411.01813, 2024. URL https://api.semanticscholar.org/CorpusID:273708589.

[33] Niklas Muennighoff, Alexander M. Rush, Boaz Barak, Teven Le Scao, Aleksandra Piktus, Nouamane Tazi, Sampo Pyysalo, Thomas Wolf, and Colin Raffel. Scaling data-constrained language models, 2023. URL https://arxiv.org/abs/2305.16264.

[34] Ashvin Nair, Abhishek Gupta, Murtaza Dalal, and Sergey Levine. Awac: Accelerating online reinforcement learning with offline datasets. In *arXiv preprint arXiv:2006.09359*, 2021.

[35] Kensuke Nakamura, Ran Tian, and Andrea V. Bajcsy. Not all errors are made equal: A regret metric for detecting system-level trajectory prediction failures. 2024. URL https://api.semanticscholar.org/CorpusID:268264177.

[36] Kensuke Nakamura, Lasse Peters, and Andrea Bajcsy. Generalizing safety beyond collision-avoidance via latent-space reachability analysis, 2025. URL https://arxiv.org/abs/2502.00935.

[37] Manisha Natarajan, Esmaeil Seraj, Batuhan Altundas, Rohan Paleja, Sean Ye, Letian Chen, Reed Jensen, Kimberlee Chang, and Matthew Gombolay. Human-robot teaming: Grand challenges. *Current Robotics Reports*, 4:1–20, 08 2023. doi: 10.1007/s43154-023-00103-1.

[38] Open X-Embodiment Collaboration et al. Open X-Embodiment: Robotic learning datasets and RT-X models. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, Yokohama, Japan, 2024.

[39] Abhishek Paudel, Xuesu Xiao, and Gregory J. Stein. Multi-strategy deployment-time learning and adaptation for navigation under uncertainty. URL https://api.semanticscholar.org/CorpusID:273146838.

[40] Siddharth Reddy, Sergey Levine, and Anca D. Dragan. Shared autonomy via deep reinforcement learning. 2018.

[41] Scott Reed, Konrad Zolna, Emilio Parisotto, Sergio Gomez Colmenarejo, Alexander Novikov, Gabriel Barth-Maron, Mai Gimenez, Yury Sulsky, Jackie Kay, Jost Tobias Springenberg, Tom Eccles, Jake Bruce, Ali Razavi, Ashley Edwards, Nicolas Heess, Yutian Chen, Raia Hadsell, Oriol Vinyals, Mahyar Bordbar, and Nando de Freitas. A generalist agent, 2022.

[42] Noveen Sachdeva, Benjamin Coleman, Wang-Cheng Kang, Jianmo Ni, Lichan Hong, Ed H. Chi, James Caverlee, Julian McAuley, and Derek Zhiyuan Cheng. How to train data-efficient llms, 2024. URL https://arxiv.org/abs/2402.09668.

[43] Fumihiro Sasaki and Ryota Yamashina. Behavioral cloning from noisy demonstrations. In *ICLR*, 2021. URL https://openreview.net/forum?id=zrT3HcsWSAt.

[44] Julie A. Shah. Human-robot teaming using shared mental

models. 2012. URL https://api.semanticscholar.org/CorpusID:8106353.

[45] Rutav Shah, Albert Yu, Yifeng Zhu, Yuke Zhu, and Roberto Martín-Martín. Bumble: Unifying reasoning and acting with vision-language models for building-wide mobile manipulation, 2024. URL https://arxiv.org/abs/2410.06237.

[46] Pratyusha Sharma, Balakumar Sundaralingam, Valts Blukis, Chris Paxton, Tucker Hermans, Antonio Torralba, Jacob Andreas, and Dieter Fox. Correcting robot plans with natural language feedback, 2022. URL https://arxiv.org/abs/2204.05186.

[47] Ben Sorscher, Robert Geirhos, Shashank Shekhar, Surya Ganguli, and Ari S. Morcos. Beyond neural scaling laws: beating power law scaling via data pruning, 2023. URL https://arxiv.org/abs/2206.14486.

[48] Yiran Tao, Guixiu Qiao, Dan Ding, and Zackory Erickson. Incremental learning for robot shared autonomy. *ArXiv*, abs/2410.06315, 2024. URL https://api.semanticscholar.org/CorpusID:273229223.

[49] Octo Model Team, Dibya Ghosh, Homer Walke, Karl Pertsch, Kevin Black, Oier Mees, Sudeep Dasari, Joey Hejna, Tobias Kreiman, Charles Xu, Jianlan Luo, You Liang Tan, Lawrence Yunliang Chen, Pannag Sanketi, Quan Vuong, Ted Xiao, Dorsa Sadigh, Chelsea Finn, and Sergey Levine. Octo: An open-source generalist robot policy, 2024.

[50] Stefanie Tellex, Nakul Gopalan, Hadas Kress-Gazit, and Cynthia Matuszek. Robots that use language. *Annu. Rev. Control. Robotics Auton. Syst.*, 3:25–55, 2020. URL https://api.semanticscholar.org/CorpusID:213739887.

[51] Chen Wang, Haochen Shi, Weizhuo Wang, Ruohan Zhang, Fei-Fei Li, and Karen Liu. Dexcap: Scalable and portable mocap data collection system for dexterous manipulation. *ArXiv*, abs/2403.07788, 2024. URL https://api.semanticscholar.org/CorpusID:268363547.

[52] Xuezhi Wang, Jason Wei, Dale Schuurmans, Quoc Le, Ed Chi, Sharan Narang, Aakanksha Chowdhery, and Denny Zhou. Self-consistency improves chain of thought reasoning in language models, 2023. URL https://arxiv.org/abs/2203.11171.

[53] Ziyu Wang, Alexander Novikov, Konrad Zolna, Jost Tobias Springenberg, Scott Reed, Bobak Shahriari, Noah Siegel, Josh Merel, Caglar Gulcehre, Nicolas Heess, and Nando de Freitas. Critic regularized regression. volume 33, pages 7768–7778, 2020.

[54] Wen Wen, Yoshihiro Kuroki, and Hajime Asama. The sense of agency in driving automation. *Frontiers in Psychology*, 10:2691, 2019.

[55] Philipp Wu, Yide Shentu, Zhongke Yi, Xingyu Lin, and P. Abbeel. Gello: A general, low-cost, and intuitive teleoperation framework for robot manipulators. *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 12156–12163, 2023. URL https://api.semanticscholar.org/CorpusID:262217072.

[56] Yilin Wu, Ran Tian, Gokul Swamy, and Andrea Bajcsy. From foresight to forethought: Vlm-in-the-loop policy steering via latent alignment, 2025. URL https://arxiv.org/abs/2502.01828.

[57] Haoran Xu, Xianyuan Zhan, Honglei Yin, and Huiling Qin. Discriminator-weighted offline imitation learning from suboptimal demonstrations. In *ICML*, pages 24725–24742. PMLR, 2022.

[58] Jia-Fong Yeh, Kuo-Han Hung, Pang-Chi Lo, Chi-Ming Chung, Tsung-Han Wu, Hung-Ting Su, Yi-Ting Chen, and Winston H. Hsu. Aed: Adaptable error detection for few-shot imitation policy, 2024. URL https://arxiv.org/abs/2402.03860.

[59] Zhaxizhuoma, Kehui Liu, Chuyue Guan, Zhongjie Jia, Ziniu Wu, Xin Liu, Tianyu Wang, Shuai Liang, Pengan Chen, Pingrui Zhang, Haoming Song, Delin Qu, Dong Wang, Zhigang Wang, Nieqing Cao, Yan Ding, Bin Zhao, and Xuelong Li. Fastumi: A scalable and hardware-independent universal manipulation interface with dataset. 2024. URL https://api.semanticscholar.org/CorpusID:276095231.

[60] Konrad Zolna, Alexander Novikov, Ksenia Konyushkova, Çaglar Gülçehre, Ziyu Wang, Yusuf Aytar, Misha Denil, Nando de Freitas, and Scott E. Reed. Offline learning from demonstrations and unlabeled experience. In *CoRR*, 2020.