AI-Assisted Ultrasound Instruction for Novices: Can We Avoid Overreliance?

Jack Fitzgerald
Colorado State University
jack.fitzgerald@colostate.edu

Nathaniel Blanchard
Colorado State University
nathaniel.blanchard@colostate.edu

ABSTRACT

Being trained to competently use an ultrasound device typically takes a few years of training, and access to training programs is limited. With the rapid advancements being made in the field of machine learning, researchers are exploring ways to bridge this gap by developing AI-powered systems that can quickly upskill ultrasound novices. These systems have been shown to improve novice performance in the short term, but it is unclear if they are truly upskilling their users. It's possible that users of such systems are overrelying on the guidance that they provide, placing a hurdle in their path towards mastering the necessary skills to become competent sonographers. I believe that it's important to keep overreliance in mind when developing a system for AI-guided ultrasound, since it could potentially have the opposite effect and ultimately decrease the pool of competent sonographers in the long run. I plan to explore the cutting edge of the field to see what tools we have at our disposal, in particular, I have found world modeling and learning latent action spaces to be an exciting direction that could be applied to this problem. The generalizability of world modeling and the ability to learn a latent action space can be utilized to create an agent that imitates and aligns with the actions that professional ultrasound instructors take during traditional training programs. Ideally, such an agent could reduce the overreliance problem by performing actions that are conducive to learning and is not constantly giving away "correct answers" to every problem a sonogorapher trainee encounters during hands-on exercises.

Keywords

Ultrasound, AI-guidance, Overreliance, Latent Actions, World Modeling

1. PROBLEM INTRODUCTION

Point-of-care ultrasound (POCUS) has become a powerful tool in the arsenal of medical practitioners thanks to decades of innovation and miniaturization, donning the title of the "new stethoscope" [6, 25]. Despite it's demonstrated value and importance in modern medicine, becoming trained in the use of an ultrasound device is costly and difficult. Introducing ultrasound training early in a medical professionals career, such as in their undergraduate studies, has shown to pay off [24]. However, early career ultrasound training is limited by equipment costs, lack of faculty funding, and decisions over how to fit ultrasound training into an already packed undergraduate curriculum. As the demand for competent sonographers increases, utilizing AI to guide and teach early-career medical practitioners and ultrasound novices appears to be a promising direction towards addressing these shortcomings

1.1 Alternative Approaches to Teaching Ultrasound

Less conventional approaches to teaching the usage of an ultrasound device have been proposed to increase access to hands-on practice. [4] developed a system that tracked the position of a probe on a manikin, and a monitor would display ultrasound images collected from real patients based on the probe's position. [28] developed the Perk Tutor system, which provided visualizations and assistance for learning ultrasound-guided needle insertions. A follow-up study with Perk Tutor showed that undergraduate students that practiced with it had improved performance compared to a control group afterwards in a test without the Perk Tutor assistance [18]. [15] developed a similar system to Perk Tutor for ultrasound-guided needle insertions and found improved novice performance in certain cases, however they did not investigate if performance persisted when not using the guidance system. While these studies show promising results, one study has shown that fully self-directed ultrasound training with simulation software is insufficient [19]. One work developed a serious game, where the player scans boxes in a toy factory, to help with spatially understanding ultrasound [8]. Mixed reality has also been used to develop a sophisticated ultrasound teaching tool [29].

1.2 AI Ultrasound Guidance

Much work has been done in the last decade to leverage AI to aid in ultrasound scans, including assistance in scan plane acquisition and detection of anatomy within the ultrasound image. [5] developed a CNN model for obstetric ultrasounds to detect standard planes and localize anatomy. [9] Learns from experts performing echocardiograms by training a model to predict the next rotational movement of the ultrasound probe that aligns with a desirable scan plane of the heart. Further work expanded on this and trained a model

Jack Fitzgerald, and Nathaniel Blanchard. AI-Assisted Ultrasound Instruction for Novices: Can We Avoid Overreliance?. In Anthony Botelho, Maria Mercedes T. Rodrigo, Adish Singla, Hiroaki Ogata, Hyojeong So, and Young Hoan Cho (eds.) Proceedings of the 19th International Conference on Educational Data Mining, Seoul, Republic of Korea, June, 2026, pp. 823–825. International Educational Data Mining Society (2026).

https://doi.org/10.5281/zenodo.21040028

on expert eye-gaze, indicating to a novice sonographer where they should be looking at in an ultrasound image during a scan [20]. Reinforcement learning has been used to predict movements toward the desired scan plane by collecting ultrasound scans in such a way that allows for offline training [21].

World modeling techniques [3, 13] have been used to perform probe movement guidance [31, 14] by making good use of unlabeled data. However, this task necessitates often expensive equipment such as robot arms or optical tracking systems to obtain accurate 6D poses of the ultrasound probe during scanning. [9] makes use of an inertial measurement unit (IMU) device, which tends to be cheaper, but noise in the accelerometer makes translational data unusable and thus is only able to predict rotational adjustments to the probe.

1.3 Novice Performance with AI Guidance

AI-guided ultrasound systems that provide probe position guidance [17, 22] and segmented anatomy overlays [7] have been shown to improve novice performance. However, these works do not investigate if these methods of AI guidance facilitate long-term learning or induce overreliance. Although these systems have demonstrated performance improvement in novices, is it enough to fully address the shortage of trained sonographers? If overreliance proves to be an issue, then that would be concerning towards the long term development of skills [16]. I propose that to avoid overreliance on AI guidance, the system should try to closely align to traditional ultrasound training that happens during hands-on practice with volunteers or phantoms.

2. PROPOSED SOLUTION

Developing an AI guidance system that teaches sonography to novices on par with traditional instruction requires the system to understand what types of actions and behaviors (verbal, gestures, etc.) that traditional ultrasound instructors display during hands-on exercises. Existing work has introduced and analyzed datasets containing teacher actions and behavior in classroom settings [1, 33, 26, 11, 23]. However, similar datasets that label teacher behavior and actions for ultrasound instructors do not exist. Work has been done to wholistically represent the actions that sonographers take with the aim of improving sonography education [10], which can be useful in informing dataset curation and possible actions, but these could be different than the actions instructors take. I propose that curating a dataset of instructors guiding students through an ultrasound scanning procedure and annotating relevant actions is the penultimate step towards a guidance system that aligns with traditional ultrasound training.

A promising line of research that has recently picked up steam is in latent action modeling [12, 32, 30, 27], which seeks to learn possible actions from a dataset and represent them in latent space. Difficulties with privacy and costs immediately arise when collecting data of ultrasound instruction (e.g. one-on-one scanning sessions with a patient, sonography student, and an instructor), and labeling such a dataset with appropriate actions adds another layer of costs. These problems in data collection motivate a self-supervised approach towards learning a useful action space for the task of modeling actions that sonogrpahy make. Moreover, pretraining on web-scale data via world modeling to learn a rich action space from unrelated tasks [12] could be frutiful, since certain generic actions may transfer (such as pointing, or verbally indicating mistakes). Models trained in such a way, such as VJEPA2 [2], have demonstrated excellent sample efficiency on downstream tasks. Therefore, perhaps only a relatively small amount of ultrasound instructor action-labeled data would be enough to fine tune a powerful model. This would be the ultimate step of the proposed solution.

3. IMPACT

I anticipate that this research will provide insights to the benefits and shortcomings in applying world modeling and latent action space modeling to education, but more specifically to the domain of hands-on ultrasound training. Developing an action-labeled dataset for ultrasound instruction is also critical, as it can directly impact the ability of the model to derive appropriate representations of the instructor actions that are actually taking place during such training.

4. REFERENCES

[1] B. Albreiki, N. Zaki, and H. Alashwal. A systematic literature review of student'performance prediction using machine learning techniques. *Education Sciences*, 11(9):552, 2021.

[2] M. Assran, A. Bardes, D. Fan, Q. Garrido, R. Howes, M. Muckley, A. Rizvi, C. Roberts, K. Sinha, A. Zholus, et al. V-jepa 2: Self-supervised video models enable understanding, prediction and planning. *arXiv preprint arXiv:2506.09985*, 2025.

[3] M. Assran, Q. Duval, I. Misra, P. Bojanowski, P. Vincent, M. Rabbat, Y. LeCun, and N. Ballas. Self-supervised learning from images with a joint-embedding predictive architecture. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 15619–15629, 2023.

[4] C. J. Banker, P. C. Pedersen, and T. L. Szabo. Interactive training system for medical ultrasound. In *2008 IEEE Ultrasonics Symposium*, pages 1350–1354. IEEE, 2008.

[5] C. F. Baumgartner, K. Kamnitsas, J. Matthew, T. P. Fletcher, S. Smith, L. M. Koch, B. Kainz, and D. Rueckert. Sononet: real-time detection and localisation of fetal standard scan planes in freehand ultrasound. *IEEE transactions on medical imaging*, 36(11):2204–2215, 2017.

[6] A. Bledsoe and J. Zimmerman. Ultrasound: The new stethoscope (point-of-care ultrasound). *Anesthesiology clinics*, 39(3):537–553, 2021.

[7] J. S. Bowness, A. J. Macfarlane, D. Burckett-St Laurent, C. Harris, S. Margetts, M. Morecroft, D. Phillips, T. Rees, N. Sleep, A. Vasalauskaite, et al. Evaluation of the impact of assistive artificial intelligence on ultrasound scanning for regional anaesthesia. *British journal of anaesthesia*, 130(2):226–233, 2023.

[8] B. Byl, M. Süncksen, and M. Teistler. A serious virtual reality game to train spatial cognition for medical ultrasound imaging. In *2018 IEEE 6th International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–4. IEEE, 2018.

[9] R. Droste, L. Drukker, A. T. Papageorghiou, and J. A. Noble. Automatic probe movement guidance for freehand obstetric ultrasound. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 583–592. Springer, 2020.

[10] L. Drukker, H. Sharma, R. Droste, M. Alsharid, P. Chatelain, J. A. Noble, and A. T. Papageorghiou. Transforming obstetric ultrasound into data science using eye tracking, voice recording, transducer motion and ultrasound video. *Scientific Reports*, 11(1):14109, 2021.

[11] Z. Gang, Z. Wenjuan, H. Biling, C. Jie, H. Hui, and X. Qing. A simple teacher behavior recognition method for massive teaching videos based on teacher set. *Applied Intelligence*, 51(12):8828–8849, 2021.

[12] Q. Garrido, T. Nagarajan, B. Terver, N. Ballas, Y. LeCun, and M. Rabbat. Learning latent action world models in the wild. *arXiv preprint arXiv:2601.05230*, 2026.

[13] D. Ha and J. Schmidhuber. World models. *arXiv preprint arXiv:1803.10122*, 2(3):440, 2018.

[14] H. Jiang, Z. Sun, N. Jia, M. Li, Y. Sun, S. Luo, S. Song, and G. Huang. Cardiac copilot: Automatic probe guidance for echocardiography with world model. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 190–199. Springer, 2024.

[15] E. J. Kim, J. Min, J. Song, K. Song, J.-H. Song, and H. J. Byon. The effect of electromagnetic guidance system on early learning curve of ultrasound for novices. *Korean journal of anesthesiology*, 69(1):15–20, 2016.

[16] N. Kosmyna, E. Hauptmann, Y. T. Yuan, J. Situ, X.-H. Liao, A. V. Beresnitzky, I. Braunstein, and P. Maes. Your brain on chatgpt: Accumulation of cognitive debt when using an ai assistant for essay writing task. *arXiv preprint arXiv:2506.08872*, 4, 2025.

[17] A. Kumar, E. Baum, C. Parmer-Chow, and J. Kugler. Limited echocardiogram acquisition by novice clinicians aided with deep learning: A randomized controlled trial. *Biology Methods and Protocols*, 10(1):bpaf083, 2025.

[18] H. Lia, Z. Keri, M. S. Holden, V. Harish, C. H. Mitchell, T. Ungi, and G. Fichtinger. Training with perk tutor improves ultrasound-guided in-plane needle insertion skill. In *Medical Imaging 2017: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 10135, pages 232–239. SPIE, 2017.

[19] F. D. Mackay, F. Zhou, D. Lewis, J. Fraser, and P. R. Atkinson. Can you teach yourself point-of-care ultrasound to a level of clinical competency? evaluation of a self-directed simulation-based training program. *Cureus*, 10(9):e3320, 2018.

[20] Q. Men, C. Teng, L. Drukker, A. T. Papageorghiou, and J. A. Noble. Multimodal-guidenet: Gaze-probe bidirectional guidance in obstetric ultrasound scanning. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 94–103. Springer, 2022.

[21] F. Milletari, V. Birodkar, and M. Sofka. Straight to the point: Reinforcement learning for user guidance in ultrasound. In *International Workshop on Preterm, Perinatal and Paediatric Image Analysis*, pages 3–10. Springer, 2019.

[22] V. Mor-Avi, B. Khandheria, R. Klempfner, J. I. Cotella, M. Moreno, D. Ignatowski, B. Guile, H. J. Hayes, K. Hipke, A. Kaminski, et al. Real-time artificial intelligence–based guidance of echocardiographic imaging by novices: image quality and suitability for diagnostic interpretation and quantitative analysis. *Circulation: Cardiovascular Imaging*, 16(11):e015569, 2023.

[23] J. Naim, J. Cao, F. Tasneem, J. Jacobs, B. Milne, J. Martin, and T. Sumner. Towards actionable pedagogical feedback: A multi-perspective analysis of mathematics teaching and tutoring dialogue. In *Proceedings of the 18th International Conference on Educational Data Mining*, pages 328–341, 2025.

[24] S. M. Provenzano, I. Mageto, A. González Frontera, C. Uter, S. Casey, A. Badillo, and A. O. Fakoya. Ultrasound as a foundational tool in medical education: A literature review. *Advances in Medical Education and Practice*, pages 2167–2179, 2025.

[25] E. Sabath. Point of care ultrasonography as the new "laennec sthetoscope". *World Journal of Nephrology*, 13(1):90542, 2024.

[26] V. Sharma, M. Gupta, A. Kumar, and D. Mishra. Star-3d: A holistic approach for human activity recognition in the classroom environment. *Information*, 15(4):179, 2024.

[27] B. Tharwat, Y. Nasser, A. Abouzeid, and I. Reid. Latent action pretraining through world modeling. *arXiv preprint arXiv:2509.18428*, 2025.

[28] T. Ungi, D. Sargent, E. Moult, A. Lasso, C. Pinter, R. C. McGraw, and G. Fichtinger. Perk tutor: an open-source training platform for ultrasound-guided needle insertions. *IEEE Transactions on Biomedical Engineering*, 59(12):3475–3481, 2012.

[29] J. Wang, J. Zhang, J. N. Capizzano, M. Sigakis, X. Wang, and V. Popov. explainmr: Generating real-time textual and visual explanations to facilitate ultrasonography learning in mr. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*, pages 1–18, 2025.

[30] S. Ye, J. Jang, B. Jeon, S. Joo, J. Yang, B. Peng, A. Mandlekar, R. Tan, Y.-W. Chao, B. Y. Lin, et al. Latent action pretraining from videos. *arXiv preprint arXiv:2410.11758*, 2024.

[31] Y. Yue, Y. Wang, H. Jiang, P. Liu, S. Song, and G. Huang. Echoworld: Learning motion-aware world models for echocardiography probe guidance. In *Proceedings of the Computer Vision and Pattern Recognition Conference*, pages 25993–26003, 2025.

[32] C. Zhang, T. Pearce, P. Zhang, K. Wang, X. Chen, W. Shen, L. Zhao, and J. Bian. What do latent action models actually learn? *arXiv preprint arXiv:2506.15691*, 2025.

[33] J. Zhao, H. Zhu, and L. Niu. Bitnet: A lightweight object detection network for real-time classroom behavior recognition with transformer and bi-directional pyramid network. *Journal of King Saud University-Computer and Information Sciences*, 35(8):101670, 2023.