It's Time to Talk Human-Centered Research at MICCAI

Sue Min Cho[1⋆], Catalina Gomez[1⋆], Katharina Breininger[2], Francis Creighton[3], Xiaoqing Guo[4], Dean Ho[5], Masaru Ishii[3], Pierre Jannin[6], Marta Kersten[7], Seong Tae Kim[8], Nassir Navab[9], Cheng Ouyang[10], Shandong Wu[11], Paul Yi[12], Maria A. Zuluaga[13], and Mathias Unberath[1]

[1] Johns Hopkins University, USA
[2] Universität Würzburg, Germany
[3] Johns Hopkins Medicine, USA
[4] Hong Kong Baptist University, Hong Kong, China
[5] National University of Singapore, Singapore
[6] Inserm Université de Rennes, France
[7] Concordia University, Canada
[8] Kyung Hee University, South Korea
[9] Technical University of Munich, Germany
[10] University of Oxford, UK
[11] University of Pittsburgh, USA
[12] St. Jude Children's Research Hospital, USA
[13] EURECOM, France
{scho72, unberath}@jhu.edu

Abstract. Integrating human-centered research practices in medical image computing (MIC) and computer-assisted intervention (CAI) is increasingly recognized as important for realizing the potential benefits of emerging technologies in clinical workflows. Human-centered research offers methods for aligning engineering solutions with human needs and expectations. These considerations have gained momentum in the field of explainable artificial intelligence (XAI) and mixed reality (MR) but are of much broader relevance as all user-facing applications – from diagnostic algorithms to surgical intervention systems – must consider human factors. The challenges arising from the incorporation of user research are multifaceted and, in the context of MICCAI, require tools and methods not currently well represented in the community. Indeed, current human-centered research efforts span diverse topics and approaches, reflecting an early stage of development within the community. This paper provides an overview of human-centered research, gathers insights into the current state of human-centered practices at MICCAI across both MIC and CAI, and points to resources from established fields. Our aim is for this work to serve as an educational foundation and to foster dialogue around human-centered research in the MICCAI community.

Keywords: Human-centered design · Human-subjects research · human-in-the-loop · Medical image computing · Computer-assisted intervention

⋆ Equal contribution

1 Introduction

Realizing the full potential of technologies across medical image computing (MIC) and computer-assisted intervention (CAI) requires a profound evolution in both our methodology and perspective. These technologies – from diagnostic algorithms to surgical interventions – ultimately involve human users. Bridging the gap between technological potential and practical application demands a holistic strategy that not only drives technical advancement but also prioritizes essential aspects of real-world applicability.

In terms of methodology, the emphasis on safe and effective solutions across both MIC and CAI requires rigorous technical validation. In MIC, technical performance metrics ensure algorithms perform accurately and reliably within clinical settings [12], with attention to robustness under varied imaging conditions [7] and approaches addressing data scarcity and bias [14]. In CAI, technical validation typically involves assessing precision, registration accuracy, and real-time performance of surgical robots, navigation systems, and guidance tools [13]. Across both domains, technical rigor is foundational.

However, the path toward truly transformative integration of these technologies extends beyond technical considerations. The success of diagnostic algorithms depends not only on their accuracy but on how clinicians interpret and act on their outputs [19]. The success of surgical technologies depends not only on their precision but also on how surgeons operate with them and how they integrate into workflows [2]. Across MIC and CAI, human-centered research offers methods for understanding these dimensions by complementing technical development with attention to real-world user needs and expectations.

The MICCAI community, renowned for its multidisciplinary nature – spanning medical imaging, computer vision, data science, AI, surgical systems, and more in the complex domain of medicine – has historically made meaningful impact through cross-disciplinary discourse and innovation. Initiatives like CLINICCAI have strengthened connections between computing researchers and clinicians; integrating human-centered methodologies will similarly require sustained collaboration with human-computer interaction (HCI), human factors, and user experience (UX) experts. Given the high-stakes nature of healthcare and the diversity of its stakeholders, sharing learnings from human-centered evaluation – both successes and failures – will be essential for collective progress.

Recognizing the growing relevance of human-centered perspectives, we have established the MICCAI Special Interest Group on Human Factors and Human-Computer Interaction (SIG-HCI). Our goal is to establish a platform committed to openness, transparency, and knowledge-sharing within the MICCAI community. We aim to provide advice and resources to MICCAI Conference organizers, reviewers, and authors, as well as engage the community in MICCAI events to raise awareness of human-centered design, human factors, and best practices of HCI.

As an inaugural effort of the SIG, this stepping-stone paper provides an introduction to human-centered research for the MICCAI community and examines its current state through a narrative review of 118 papers from 2020–2025. Sec-

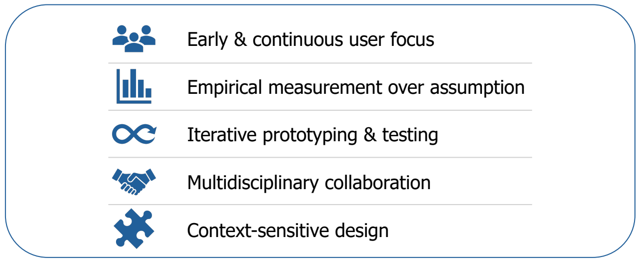

Fig. 1. Core principles of HCR, adapted from ISO 9241-210 [9].

tion 2 presents principles and methods, Section 3 characterizes current practices, and Section 4 discusses pathways forward.

2 What is Human-Centered Research?

Human-centered research (HCR) prioritizes human needs, behaviors, and capabilities in both the design and evaluation of technologies. Related terms include human-centered design and user-centered design; we use HCR broadly to encompass both design processes and empirical evaluation of how humans interact with technologies. The field draws on human factors and ergonomics, HCI, and UX research, providing theories, methods, and validated instruments. In specialized domains like healthcare, researchers often develop technologies for users whose skills and experiences they do not share. HCR addresses this gap by engaging stakeholders throughout development, ensuring the "right" problem is solved rather than assumed [16].

Several core principles guide HCR [9], summarized in Figure 1. First, there should be an early and continuous focus on users and their tasks – not just at the end of development, but from the earliest stages of problem definition. Second, claims about user experience should be grounded in empirical measurement of behavior and perception, not assumptions. Third, design should be iterative, with repeated cycles of prototyping, testing, and refinement based on user feedback. Fourth, HCR is inherently multidisciplinary, requiring collaboration across technical, clinical, and human factors expertise. Finally, context matters: the same technology may succeed or fail depending on the application setting, user expertise, time pressures, and existing workflows. This variability underscores why human-centered evaluation must complement technical benchmarks.

HCR employs diverse methodologies, including qualitative approaches (e.g., interviews, observations) and quantitative approaches (e.g., performance metrics, validated questionnaires). Data collection may be observation-based, score-based (e.g., standardized scales), or sensor-based (e.g., eye tracking, physiological measures), each offering different insights into user experience and behavior.

Depending on the stage of the design process, HCR provides validation methodologies that can be broadly organized into *formative* approaches, which shape and refine designs during development, and *summative* approaches, which evaluate how well mature designs perform [18]. Formative research uses qualitative methods such as contextual inquiry and shadowing [5] to understand user needs, leading to prototypes of varying fidelity that enable rapid, low-cost iteration before designs mature. Summative research assesses how well a mature design performs against predefined benchmarks and user needs, often involving larger samples and more controlled conditions to support broader conclusions.

Usability evaluation collects evidence from observed data through studies that involve participants representative of real users performing realistic tasks [10]. While usability testing can occur during formative stages with prototypes, summative usability studies typically employ more refined systems and standardized protocols to assess whether the design meets acceptance criteria. Usability studies are an important first step in human-centered evaluation, establishing whether users find systems workable and satisfactory. However, comprehensive HCR must go beyond usability to examine broader impacts: effects on user skills and expertise over time, changes to work practices and job roles, and risks such as skill degradation or overtrust in systems.

Performance and behavior evaluations analyze logs and outcomes such as completion and error rates, task time, and learnability. Standardized questionnaires such as the NASA Task Load Index and System Usability Scale provide validated measurements, often paired with interviews.

Throughout both formative and summative stages, researchers employ *empirical* methods (deriving conclusions from observed user data) and *non-empirical* methods (applying established heuristics or cognitive models). Non-empirical approaches such as heuristic evaluation and cognitive walkthroughs offer efficient ways to identify usability issues [17], particularly when access to representative users is limited. These methods do not replace empirical evaluation but complement it, enabling rapid iteration and helping prioritize issues for further investigation.

3 Current Landscape of Human-Centered Research at MICCAI

Having introduced what HCR entails, we now survey its current presence at MICCAI. We conducted a narrative review of proceedings from 2020 to 2025[14], searching for terms including "participant," "user," and "human", and for related phrases such as "interaction design," "human-centered," and "user interface" to identify potentially relevant papers. We defined these terms to retrieve papers that involved human participants in evaluation (e.g., user studies, expert ratings), incorporated human-derived behavioral data (e.g., gaze, annotations), or explicitly addressed user-facing design considerations. We then manually screened the papers from this search, and we identified 118 papers that engage with human-centered themes across medical image computing (86 papers) and computer-assisted intervention (32 papers). This survey is illustrative

[14] https://link.springer.com/conference/miccai

Table 1. Landscape of HCR at MICCAI (2020–2025 narrative review).

Category	**Count**	**Sample References**
Type of Human Involvement		
User studies	35	[15, 20, 6]
Expert evaluation/comparison	25	[26, 27]
Human-derived data	12	[25, 4]
Simulated interaction	35	[28, 8]
Discussion without user involvement	11	
Select Topics		
Interactive segmentation/annotation	25	[11, 24]
Gaze-based methods	9	[3]
Explainability/interpretability	5	[23, 22]
AR/Navigation/Guidance	8	[1, 21]

rather than exhaustive, intended to characterize the breadth of topics explored, types of human involvement, methods employed, and outcomes measured, rather than to provide definitive counts or comprehensive coverage.

3.1 Topics Explored

Human-centered themes appear across diverse research areas. Interactive segmentation and annotation is the most common, with 25 papers proposing systems that accept clicks, scribbles, bounding boxes, or text prompts to guide segmentation. Gaze-based methods represent a growing area, using clinician's eye movements for model supervision, error analysis, or attention guidance. Explainability and interpretability research examines how model outputs can be made understandable to clinical users. Data synthesis and augmentation work often involves human evaluation of generated image realism or clinical utility.

In diagnostic applications, screening and diagnosis systems frequently include comparison to radiologist performance. These studies involve humans as a reference benchmark, measuring whether system accuracy matches or exceeds expert performance. This differs from evaluation where humans are studied as end users, examining how clinicians interpret, trust, or act on system outputs in practice. Both forms of human involvement are relevant, but they address different research questions. Report generation systems are similarly evaluated against human-written reports or by expert review, which establishes output quality; human-centered evaluation would further examine how clinicians use generated reports in their workflow.

In interventional applications, augmented reality (AR), navigation, and guidance systems help surgeons with spatial orientation and planning. Perception and feedback systems enhance what operators can see or sense during procedures, and surgical video and reconstruction work addresses visualization needs. Robotics and tracking research examines human-robot interaction dynamics, while ultrasound and probe guidance systems support real-time imaging acquisition.

3.2 Types of Human Involvement

Papers vary substantially in how humans participate in the research. In 35 papers, user studies were conducted in which participants, typically clinicians, surgeons, or trainees, performed tasks with the system and their behavior or performance was measured. These span interactive segmentation studies measuring interaction time, AR navigation studies with surgeons, perception studies comparing visualization techniques, and usability evaluations of guidance systems. Another common pattern, appearing in 25 papers, involves expert evaluation or comparison, where system outputs are rated by experts or compared against expert performance without experts interacting with the system as end users. Reader studies comparing AI to radiologists, and expert ratings of image quality or realism, fall into this category.

12 papers incorporate human-derived data, such as radiologist gaze recordings or clinician annotations, without involving new human participants, leveraging previously collected behavioral data for model training or supervision. A notable pattern involves interactive capability paired with simulated evaluation: 35 papers propose systems supporting user input but evaluate using simulated clicks, prompts, or scribbles rather than actual users. This pattern is particularly common in interactive segmentation research. Finally, 11 papers discuss potential benefits related to human factors such as usability, trustworthiness, or workflow integration, without conducting empirical human evaluation to support these claims.

3.3 Methods and Outcomes

Among papers involving humans, several methodological approaches appear. Reader studies present cases to clinicians who provide diagnoses or ratings, enabling comparison between algorithms and human performance or measurement of computer-assisted accuracy. Expert rating or review involves specialists evaluating outputs for quality, realism, or clinical acceptability without performing a full diagnostic task. Usability studies observe users interacting with systems, measuring task completion, errors, interaction time, or subjective assessments. Controlled experiments systematically vary conditions, such as comparing visualization techniques or feedback modalities, to measure effects on user performance or perception. Perceptual evaluations assess whether generated or processed images appear realistic or clinically appropriate to human observers, while clinical validation evaluates systems with representative users performing realistic tasks in settings approximating actual use.

When users are involved, they are most commonly expert clinicians (radiologists, pathologists, cardiologists) or surgeons, with trainees, residents, or novices appearing less frequently. The most common outcomes involve task performance: accuracy, completion time, and error rates. Comparison to human baseline appears frequently, establishing whether computer(-aided) systems match or exceed unaided human performance. Perceptual quality or realism is assessed particularly in work involving image synthesis or reconstruction. Usability and interaction efficiency, measured as time or number of interactions required, appear in evaluations of interactive systems. Some constructs relevant to clinical adoption are measured less frequently, including cognitive load, trust, and factors related to how users develop appropriate reliance over time.

3.4 Observations

A substantial portion of papers propose interactive capabilities, including systems that accept clicks, prompts, scribbles, or other user inputs, but evaluate these capabilities using simulated interactions rather than studies with actual users. 35 papers follow this pattern, leaving open questions about whether the proposed interaction models match real user behavior, whether users find these systems efficient and usable, and how performance observed with simulated inputs translates to practice. This gap between interactive framing and empirical user evaluation represents an opportunity for future research to examine how these systems perform with real users in realistic contexts.

The two domains show different patterns of user involvement. Among CAI papers, ∼47% explicitly conduct user studies, compared to ∼26% of MIC papers. This difference likely reflects the procedural nature of CAI, where technologies directly support human operators during time-critical interventions. Human factors considerations have long been more inherent to interventional and surgical systems, and the human role is difficult to abstract away. In MIC, where technologies often support asynchronous interpretation tasks, such perspectives are comparatively newer. The human user may be easier to omit from evaluation, though no less important to eventual deployment success.

Many papers compare AI outputs to expert performance, addressing whether AI can match human accuracy on diagnostic or analytic tasks. Fewer papers study how humans actually work with AI systems once deployed, examining whether users trust outputs appropriately and integrate recommendations into their decision-making effectively. These are distinct research questions with different methodological requirements and different implications for clinical translation. The former establishes technical capability; the latter examines whether that capability translates into benefit when humans are in the loop.

Some aspects of human-centered evaluation remain underexplored across both domains. Studies predominantly involve experienced clinicians or surgeons, with less attention to how trainees learn to use new systems or how expertise level affects interaction patterns. Constructs central to the human factors literature, including trust calibration, cognitive workload, and situation awareness, are infrequently measured despite their relevance to whether technologies will be adopted and used effectively in practice. Longitudinal questions about how use evolves over time, how users develop appropriate reliance, or how systems affect clinical outcomes remain largely unaddressed within MICCAI proceedings.

4 Looking Forward

The observations in Section 3 reveal both emerging engagement with human-centered research and opportunities for growth. We discuss three pathways forward: embracing methodological diversity, building community infrastructure, and learning from established fields.

4.1 Embracing Methodological Diversity

Our survey reveals that empirical user studies and expert evaluations are the most common forms of HCR at MICCAI. These summative approaches are valuable but, as discussed in Section 2, represent only part of the HCR toolkit. Formative methods such as contextual inquiry and prototyping help shape system design before formal evaluation, while non-empirical approaches like heuristic evaluation offer efficient alternatives when user access is limited. Contributions also come in diverse forms: qualitative studies, design iterations, field deployments, and early-stage explorations that surface unexpected challenges. As MICCAI engages more deeply with HCR, reviewers and organizers may benefit from exposure to these traditions to recognize valuable contributions that fall outside purely technical evaluation frameworks.

4.2 Building Community Infrastructure

Shared infrastructure can lower barriers for researchers new to HCR and promote consistency across the MICCAI community. One form of infrastructure involves shared instruments and study protocols. Validated questionnaires for measuring constructs adapted for MIC and CAI contexts could enable researchers to build on established measurement approaches rather than developing ad hoc instruments. Reporting guidelines analogous to Metrics Reloaded [12] could promote transparency, and coordinated international consensus efforts could further establish shared best practices for HCR. Infrastructure can also address sample size challenges. Recruiting clinicians is resource-intensive, and individual groups may struggle to achieve adequate statistical power. Cross-institutional collaborations enabling multi-site studies, and shared datasets containing human behavioral data (e.g., annotations, response times, gaze patterns), could support research difficult for any single group to conduct alone.

4.3 Learning from Established Fields

HCR has a long history in fields such as HCI, human factors and ergonomics, and human-robot interaction (HRI). These disciplines offer established methods, validated instruments, and accumulated knowledge directly relevant to MIC and CAI. The HCI community provides extensive methods for user-centered design, usability evaluation, and interaction studies. Human factors bring traditions from aviation and safety-critical systems, with approaches for assessing workload, situational awareness, and error. HRI addresses trust in autonomous systems, shared control, and human-machine teaming. Engaging with these fields through their literature, conferences, or direct collaborations can enrich HCR at MICCAI. Rather than developing methods in isolation, researchers can build on established foundations and adapt them to the specific contexts of medical image computing and computer-assisted intervention.

5 Conclusion

This paper has introduced human-centered research to the MICCAI community through an overview of HCR principles and methods, a narrative review of MICCAI proceedings (2020–2025), and pathways forward. Our survey reveals emerging engagement with HCR, while also pointing to opportunities in methodological diversity, community infrastructure, and connections to established fields. We hope this paper serves as an accessible foundation for researchers interested in human-centered perspectives and contributes to ongoing dialogue about developing technologies that work not only in technical benchmarks but in the hands of real users.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

Data Availability Statement. The data and source materials supporting the findings of the narrative review are available from the corresponding author upon reasonable request.

References

1. Acar, A., Atoum, J., Connor, P.S., Pierre, C., Lynch, C.N., Kavoussi, N.L., Wu, J.Y.: Navius: navigated augmented reality visualization for ureteroscopic surgery. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 433–443. Springer (2025)
2. Amini Rarani, S.: Smart technologies and digital innovations for improving perioperative patient safety: a review. Patient Safety in Surgery **19**(1), 31 (2025)
3. Anikina, A., Ibragimova, D., Mustafaev, T., Mello-Thoms, C., Ibragimov, B.: Longitudinal anatomical attention maps for recognizing diagnostic errors from radiologists' eye movements. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 315–325. Springer (2025)
4. Awasthi, A., Chung, B.V., Vu, A.M., Le, N., Agrawal, R., Deng, Z., Wu, C., Nguyen, H.V.: Maarta: Multi-agentic adaptive radiology teaching assistant. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 358–368. Springer (2025)
5. Van der Bijl-Brouwer, M., Dorst, K.: Advancing the strategic impact of human-centred design. Design studies **53**, 1–23 (2017)
6. Cho, S.M., Wu, W., Kilmer, E., Taylor, R.H., Unberath, M.: Feeling the stakes: Realism and ecological validity in user research for computer-assisted interventions. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 189–197. Springer (2025)
7. Drenkow, N., Sani, N., Shpitser, I., Unberath, M.: A systematic review of robustness in deep learning for computer vision: Mind the gap? arXiv preprint arXiv:2112.00639 (2021)
8. Fu, J., Li, H., Lu, T., Zhang, S., Wang, G.: Um-sam: Unsupervised medical image segmentation using knowledge distillation from segment anything model. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 616–626. Springer (2025)

9. International Organization for Standardization: ISO 9241-210 Ergonomics of Human-System Interaction - Part 210: Human-Centred Design for Interactive Systems. https://www.iso.org (2010)
10. Lazar, J., Feng, J.H., Hochheiser, H.: Research methods in human-computer interaction. Morgan Kaufmann (2017)
11. Liu, W., Ma, C., Yang, Y., Xie, W., Zhang, Y.: Transforming the interactive segmentation for medical imaging. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 704–713. Springer (2022)
12. Maier-Hein, L., Reinke, A., Godau, P., Tizabi, M.D., Buettner, F., Christodoulou, E., Glocker, B., Isensee, F., Kleesiek, J., Kozubek, M., et al.: Metrics reloaded: recommendations for image analysis validation. Nature methods pp. 1–18 (2024)
13. Marcus, H.J., Ramirez, P.T., Khan, D.Z., Layard Horsfall, H., Hanrahan, J.G., Williams, S.C., Beard, D.J., Bhat, R., Catchpole, K., Cook, A., et al.: The ideal framework for surgical robotics: development, comparative evaluation and long-term monitoring. Nature medicine **30**(1), 61–75 (2024)
14. Pavlak, M., Drenkow, N., Petrick, N., Farhangi, M.M., Unberath, M.: Data audit: Identifying attribute utility- and detectability-induced bias in task models. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 442–452. Springer (2023)
15. Podvin, B., Collins, T., Saibro, G., Innocenzi, C., Yang, Y., Milana, F., Keeza, Y., Ufitinema, G., Ujemurwego, F., Torzilli, G., et al.: Samusa: Segment anything model 2 for ultrasound annotation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 509–518. Springer (2025)
16. Reyes, J., El-Mufti, N., Gorman, S., Xie, D., Kersten-Oertel, M.: User-centered design for surgical innovations: a ventriculostomy case study. In: Workshop on the Ethical and Philosophical Issues in Medical Imaging. pp. 51–62. Springer (2022)
17. Rogers, Y., et al.: New theoretical approaches for hci. Annual review of information science and technology **38**(1), 87–143 (2004)
18. Sankowski, O., Krause, D.: The human-centredness metric: early assessment of the quality of human-centred design activities. Applied Sciences **13**(21), 12090 (2023)
19. Shneiderman, B.: Human-centered AI. Oxford University Press (2022)
20. Sommersperger, M., Dehghani, S., Matten, P., Mach, K., Nasseri, M.A., Roodaki, H., Eck, U., Navab, N.: Semantic virtual shadows (svs) for improved perception in 4d oct guided surgery. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 408–417. Springer (2023)
21. Sommersperger, M., Dehghani, S., Matten, P., Mach, K., Roodaki, H., Eck, U., Navab, N.: Intelligent virtual b-scan mirror (ivbm). In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 418–428. Springer (2023)
22. Tang, D., Donnelly, J., Barnett, A.J., Semenova, L., Jing, J., Hadar, P., Karakis, I., Selioutski, O., Zhao, K., Westover, M.B., et al.: This eeg looks like these eegs: Interpretable interictal epileptiform discharge detection with protoeeg-knn. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 615–625. Springer (2025)
23. Wåhlstrand Skärström, V., Johansson, L., Alvén, J., Lorentzon, M., Häggström, I.: Explainable vertebral fracture analysis with uncertainty estimation using differentiable rule-based classification. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 318–328. Springer (2024)
24. Williams, H., Pedrosa, J., Cattani, L., Housmans, S., Vercauteren, T., Deprest, J., D'hooge, J.: Interactive segmentation via deep learning and b-spline explicit

active surfaces. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 315–325. Springer (2021)
25. Wu, S., Chen, J., Jin, Z., Zhang, P., Gao, Z., Feng, J., Zhang, X., Shen, D.: Graph-based neighbor-aware network for gaze-supervised medical image segmentation. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 229–239. Springer (2025)
26. Xie, C., Yoshii, Y., Kitahara, I.: Sv-drr: High-fidelity novel view x-ray synthesis using diffusion model. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 572–582. Springer (2025)
27. Yan, K., Yin, X., Xia, Y., Wang, F., Wang, S., Gao, Y., Yao, J., Li, C., Bai, X., Zhou, J., et al.: Liver tumor screening and diagnosis in ct with pixel-lesion-patient network. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 72–82. Springer (2023)
28. Zhou, T., Li, L., Bredell, G., Li, J., Konukoglu, E.: Quality-aware memory network for interactive volumetric image segmentation. In: International conference on medical image computing and computer-assisted intervention. pp. 560–570. Springer (2021)