東海大学紀要情報理工学部
Vol. 25, 2025, pp. 8–13
DOI: 10.18995/24352152.25.8

Conceptual Prototype of Mixed Reality-Based Ultrasound Training System

by

KAWATTIKUL Khanabhorn[*1], HAMAMOTO Kazuhiko[*2]

(received & accepted)

Abstract

Training a novice ultrasonographer is rather complicated because the image appearance depends on ultrasound probe movement and scanning technique. Conventional methods of training often rely on real patients or physical phantoms, making them less flexible and available. Mixed reality (MR) has the potential to provide an alternative to such training contexts in the form of interactive, reproducible, and relatable environments. This article describes a conceptual prototype framework and a MR ultrasound training system which combines AI-generated ultrasound images with sensor-based interaction. Using a conditional generative adversarial network, ultrasound images are generated and integrated in the system from a pipeline that is deployment oriented. The gyro-sensor captures ultrasound-probe-like movements and relates user actions to variations in ultrasound images shown on the screen. The system is characterized by a user experience, interface clarity, and interaction design for beginners. Training is planned with a stepwise sequence of observation, guided practice, and autonomous practice phases. Practical observations about the operation of the image configuration and generation are also given under a single implementation schema for training-oriented system design.

Keywords: Mixed Reality, Novice Ultrasonographer, Ultrasound Training, User Experience Design, Interaction Design, Mannequin, AI-Generated Medical Images

1. Introduction

Ultrasound imaging is used in medicine diagnosis due to its non-invasive nature, gives real-time feedback, and is low cost. However, for a novice ultrasonographer, learning how to use ultrasound is still hard. The quality of the image depends on the position, orientation, and pressure of the ultrasound probe (US-probe), which requires a lot of practice under expert supervision. Traditional training relies on real patients or physical phantoms. Patient availability is often limited, phantoms are costly to maintain, and ethical concerns arise when inexperienced users practice on real individuals. As a result, many medical schools struggle to provide sufficient supervised practice opportunities for beginners [1-2].

Mixed reality (MR) has been explored as an alternative training approach by integrating virtual content into real-world environments [2]. In ultrasound education, MR-based systems allow learners to practice US-probe handling without continuous dependence on real patients. Many existing systems use sensor-based tracking, such as gyro-sensors, to capture US-probe movement and orientation, which is critical because ultrasound images (US-images) change dynamically with US-probe position and angle [3-6]. While these systems demonstrate technical feasibility, most focus on realism or task completion and give limited attention to interaction design and user experience tailored to a novice ultrasonographer.

At the same time, advances in artificial intelligence have enabled the generation of realistic medical images. Conditional generative adversarial networks (CGANs) can synthesize US-images conditioned on anatomical context or US-probe-related parameters, offering a potential solution to data scarcity and reducing the need for additional patient scans [5-8]. Prior work mainly evaluates image quality or model performance [8-11], and few studies investigate how AI-generated images can be integrated into interactive training systems, particularly within MR environments [12].

From a learning perspective, ultrasound training needs to include educational principles that facilitate the pedagogical approach of learning through experience: active practice and timely feedback. First, clear way-finding and guidance can help reduce cognitive load for novices. MR learning research highlights the value of natural gestures, spatial feedback, and gradual changes in task difficulty; however the majority of ultrasound training systems are technology-centric with the lack of experiential design focus.

Furthermore, practical limitations such as image generation speed are often not considered. The feeling of naturalness in the ultrasound training system requires that image generation follows US-probe motion. Excessive latency can disrupt the learning experience, and reliance on high-end GPUs limits accessibility in educational settings with standard computing resources.

In this paper, we propose a conceptual design of MR-based ultrasound training system designed for a novice ultrasonographer. Generated US-images of the AI-model input of CGAN combined with

＊1 Ph.D. Student, Graduation School of Science and Technology, Tokai University (Shinagawa Campus)
＊2 Professor, School of Information Science and Technology, Tokai University (Shonan Campus)

gyro-sensor based tracking were integrated to simulate realistic probe interaction. In contrast with full system realization or user evaluation work, this research primarily addresses experience-driven design, where the focus is on user experience and interaction design from an amateur perspective. We also present empirical findings on ultrasound image generation behavior within a training-specific image configuration to determine whether it is feasible for interactive use on moderately factored computers.

The rest of this paper is structured as follows: Section 2 describes the conceptual framework underlying the proposed MR ultrasound training system. Section 3 details the system and prototype design, including AI image integration, user experience and interaction design, and computational feasibility observations. Section 4 provides a discussion of the design contributions, implications, and limitations of this work. Section 5 concludes the paper and presents future research directions.

2. Conceptual Framework

This section describes the conceptual framework of our MR ultrasound training system, covering learning principles, user experience objectives, interaction design, and system components. We explain how educational considerations guided the design within practical technical constraints.

2.1 Design Rationale

The system design is motivated by challenges commonly faced by a novice ultrasonographer, particularly the difficulty of understanding how US-probe movement influences the displayed ultrasound image. This relationship is often difficult to grasp through textbooks or static images alone. To address this challenge, the proposed system supports learning through direct interaction and clear visual feedback, with MR serving as the primary platform for training.

MR enables virtual ultrasound images to be presented spatially while users move physically, supporting embodied learning through action and perception. AI-generated images provide flexible training data without requiring patients, while gyro-sensor simulation links probe movement to image changes. The goal is to replicate real scanning practice for a beginner or a novice ultrasonographer.

2.2 Learning and Interaction Design

The framework is based on experiential learning theory, which emphasizes learning through practice, feedback, and reflection. For a novice ultrasonographer, repeated practice with clear feedback supports both technical handling and image interpretation.

According to this perspective, the training is divided into three stages. In the first stage, it focuses on observation, where learners explore US-images and basic anatomy, the second stage, guided practice occurs, and system feedback is utilized to guide US-probe movement, And the final stage allows more independence so learners can work with reduced guidance. This framework represents an effort to alleviate cognitive load by progressing the tasks and pairing visual feedback with actions taken by the user.

2.3 Linking Learning to System Design

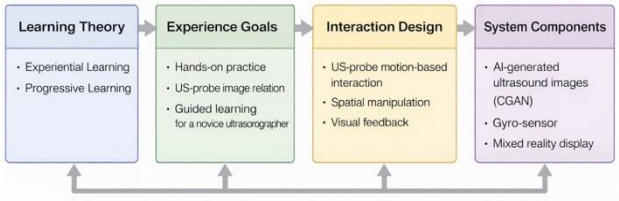


Fig.1 Conceptual mapping from learning principles to interaction design and system components in the proposed mixed reality ultrasound training system.

As shown in Fig. 1, there is an explicit mapping of educational goals to interaction mechanisms and technical components, allowing learning principles to be translated into system design. Principles of experiential and progressive learning encourage engagement — using sensor-based US-probe tracking in our work, users can investigate the relationship between probe motion and ultrasound image appearance. Gyro-sensor data are mapped to controlled changes in AI-generated ultrasound images, providing immediate visual feedback that supports perceptual learning.

The interface and interaction designs are simple and easy for a novice ultrasonographer to navigate through. Every training stage presents limited information with clear visual cues to reduce cognitive load while reinforcing US-probe–image relationships. Design considerations are also constrained by interactive performance requirements, so the tradeoff between AI image generation parameters and system responsiveness is balanced to achieve stable real-time operation on computers with moderate specifications. This design-oriented mapping helps to show how learning objectives inform interaction design and system implementation.

3. System and Prototype Design

3.1 System Overview

The proposed system is a MR–based training prototype designed for a novice ultrasonographer. Its primary goal is to support early-stage learning by allowing learners to explore the relationship between US-probe movement and US-image appearance through interactive experience in a controlled physical training environment.

The system employs a US-probe-like object equipped with a gyro-sensor to capture orientation and motion data during user interaction. In the prototype implementation, interaction is performed on a physical breast phantom rather than a full-body mannequin. The breast phantom provides a stable and anatomically representative surface for probe placement and movement, ensuring a consistent relationship between the learner's physical actions and the visual feedback presented by the system. As users manipulate the US-probe on the phantom surface, corresponding changes are reflected in the displayed ultrasound images. This setup enables simulation of basic US-probe handling without the need for a real ultrasound device or a human subject.

To support intuitive understanding of the training scenario, a conceptual illustration of the proposed MR ultrasound training system

is shown in Fig. 2. This illustration provides an overview of the interac-tion context, including the learner, US-probe, and MR environment, and is intended to help readers grasp the overall training concept before detailed system components are introduced.

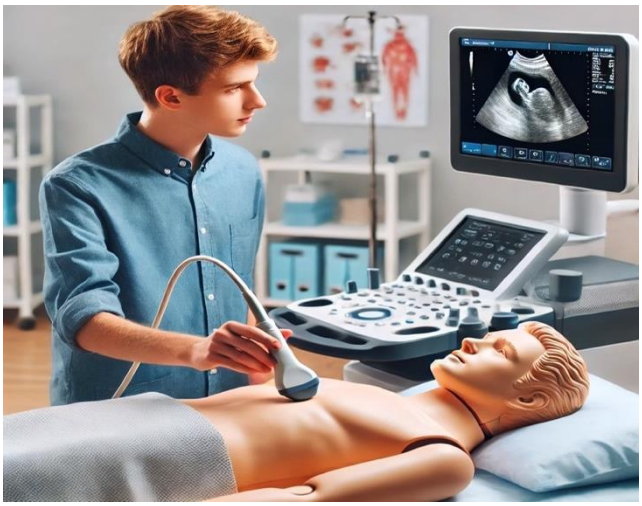


Fig. 2 Conceptual illustration of a mannequin-based mixed reality ultrasound training scenario for a novice ultrasonographer.*

Fig. 3 presents a photograph of the implemented prototype setup, showing the US-probe-like device and the breast phantom used during interaction. This physical configuration serves as the source of the ultrasound images visualized in the training interface, as discussed later in Section 3.3 and shown in Fig. 5.

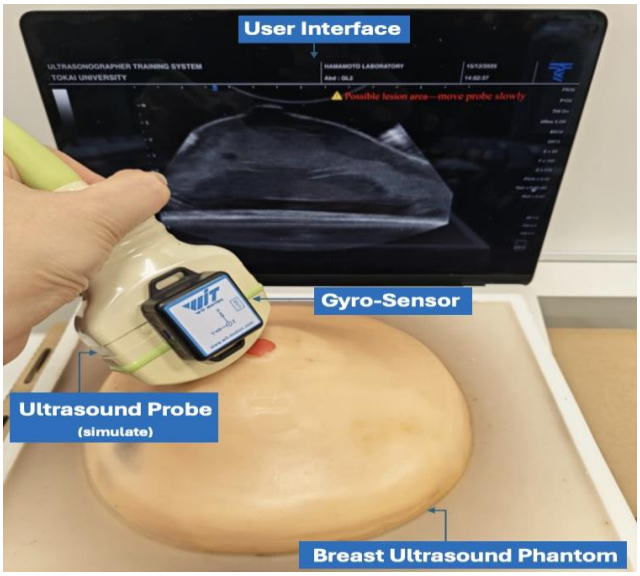


Fig.3 Prototype interaction setup showing the US-probe-like device with a gyro-sensor operated on a breast ultrasound phantom. The ultrasound image generated during interaction is displayed on the user interface.

＊This image is generated for illustrative purposes only and does not represent the implemented prototype or system architecture described in this study.

As shown in Fig. 4, the system architecture of the proposed training prototype integrates sensor-based US-probe interaction, AI-based ultrasound image generation, and MR visualization within a unified pipeline.

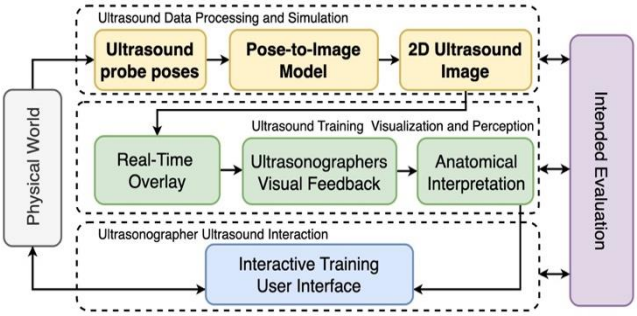


Fig. 4 System architecture of the proposed mixed reality ultrasound training prototype, illustrating sensor-based US-probe interaction on a breast phantom, AI-generated ultrasound images, and mixed reality visual feedback. Evaluation-related components represent future work.

To generate the ultrasound images, the system utilizes a conditional generative adversarial network. The generated images provide flexible and repeatable visual content suitable for training purposes, rather than serving as a replacement for clinical ultrasound data [4,9,13]. Design considerations related to image generation speed and interaction responsiveness are addressed as part of the overall system design and are discussed later in this paper.

3.2 Integration of AI-Generated Ultrasound Images

Within the system pipeline [13], the AI image generation module functions as a dynamic image source within the training system. Input parameters derived from US-probe orientation and scanning position on a physical mannequin are mapped to corresponding ultrasound image outputs, with the mannequin serving as a stable spatial reference for consistent US-probe–image alignment.

The generated US-images are passed to the MR visualization module and presented to the user in real time or near real time, allowing learners to observe image changes that directly correspond to their physical probe movements.

Image generation speed is treated as a system-level design constraint to support interactive training. The AI model and image configuration are selected to balance visual representativeness and computational efficiency. Details of the image configuration and observed generation behavior are reported later to provide reference information for training-oriented system design.

3.3 User Experience and Interface Design

The user interface of the proposed system is designed to support a novice ultrasonographer during early-stage training by helping them understand the relationship between US-probe movement and US-image appearance through direct interaction and visual feedback.

The training is designed to follow a three-stage structure. During the first stage, they will observe US-images and basic anatomical features

without directly engaging with the input. Then guided practice starts, with visual cues and interface as support. Last, they complete practice independently while providing less information and encouragement to facilitate individualized practice of skills. By using clear, simple design, the interface design is optimized to reduce cognitive load [14].



Fig. 5 Example user interface of the mixed reality ultrasound training prototype, illustrating minimal visual elements and stage-specific guidance.

As shown in Fig. 5, the displayed ultrasound image is generated during interaction with the breast phantom described in Section 3.1, allowing learners to directly associate probe movement on the physical training object with changes in image appearance. The presentation of visual cues such as markers and text that help attract attention (decision support) should be limited to brief instructions or feedback. This design is based on well-established usability concepts such as visibility of system status, consistency, and immediate feedback, which have been modeled on Nielsen's usability heuristics [15]. These principles help encourage intuitive interaction and predictable behavior of the system as users, which are key tenets for a novice ultrasonographer.

Real-time feedback is provided by mirroring gyro-sensor–captured changes in US-probe orientation directly in the shown US-images so that users can link physical behaviors with matched changes in images. The experiential design throughout the training series provides balanced guided and independent learning by the learner; however, there is more robust interface provision in the starting stages, while gradually decreasing the amount of guidance to encourage interaction and practice with the learning. The interface is designed for training purposes and does not aim to replicate the full complexity of clinical ultrasound systems.

3.4 Interaction Design with Gyro-Sensor

Interaction design in the proposed system focuses on simulating basic US-probe manipulation using sensor-based input. A gyro-sensor embedded in a US-probe-like object captures orientation and rotational movement during interaction with a physical mannequin. Although simple, this design mimics basic scanning behavior as it remains within easy reach of a novice ultrasonographer.

In the prototype, a BWT61CL sensor (WITMOTION) [16] is used to capture US-probe orientation. The sensor is an AHRS-based IMU capable of measuring three-axis orientation angles, which are transmitted to the system and used as input for interaction control in the training system . In this study, orientation information is utilized to represent US-probe tilt and rotation during scanning interaction, rather than to achieve precise physical tracking.

Gyro-sensor orientation information is linked with changes in US-images in the MR environment. The movement of the US-probe-like object on the mannequin (tilt, rotate) gives corresponding change in image appearance [4,9,13]. The mapping is designed to be consistent and predictable, allowing users to associate physical movement with visual output through repeated interaction. The interaction model emphasizes immediate visual feedback and avoids complex gesture sets to reduce cognitive load in early training. Interaction actions correspond with the training steps specified in the UX design.

In guided practice, visual cues might indicate adequate US-probe orientation or movement. To encourage independent interaction based on learned US-probe–image relationships, later stages slowly reduce these indicators. Additionally, the presence of a gyro-sensor provides interaction that can be achieved without specialized medical equipment and is accessible in a setting, like a classroom, or simulation lab. Instead, the design highlights functional correspondence between US-probe movement and the response to an image, more appropriate for novice ultrasonographer than exact physical accuracy.

3.5 Computational Feasibility Observation

This subsection assesses the computational feasibility of integrating AI-generated US-images into the proposed MR training prototype. Rather than conducting a comparative performance evaluation, this work focuses on reporting the image configurations used in the prototype and the observed image generation behavior under a single implementation setup, as summarized in Table 1.

The prototype was implemented and tested on a MacBook Air with Apple M2 processor (8-core CPU, integrated M2 GPU, 16 GB unified memory), a common device in academic environments. US-images were generated using a CGAN model developed by the authors. The trained model [13] was converted to ONNX format for system deployment [17]. A backend module handled image generation and inference, while a frontend module managed interaction and MR visualization.

Table 1. Image configuration and observed generation speed used in prototype

Image Resolution (pixels)	Intended Use in Training	Observed Generation Time (ms)	Interaction Suitability
523 x 732	High-detail reference view	19.56	Moderate
512 x 720	Standard training view	18.87	Moderate
384 x 576	Interactive training view	13.68	Highly Suitable
256 x 384	Lightweight Interaction	11.42	Suitable

Multiple image resolutions were chosen as shown in Table 1 for different training applications. Reference viewing was conducted using higher resolutions, while lower resolutions provided better performance for more interactive training tasks. Of the tested configurations, a resolution of 384 x 576 pixels was chosen as the primary setting for interactive training, as it preserves the visual characteristics of the model while retaining the ground truth images used during model training (523 x 732 pixels) and offers more responsive interaction compared to higher-resolution configurations. Decreased resolution images delivered relatively stable image updates in response to user interactions, facilitating the training flow without severe interruption.

Our findings imply that adaptation of image resolution with training context can be considered to avoid the clash between visual clarity and interaction responsiveness in MR training contexts. As evidenced by the observed image generation behavior, we can apply AI-generated US-images in a MR training prototype using a training-oriented image configuration. Although the system does not attempt clinical real-time operation, its responsiveness can be interpreted as adequate for a novice ultrasonographer that require immediate visual feedback. Our results serve as a baseline reference for system designers; however, further optimization and evaluation are necessary for these applications.

4. Discussion

In this paper, we propose an abstract prototype of a MR ultrasound training system which combines AI-generated ultrasound images with sensor-based interaction. This study is based on contributions to design, implications for operation, and limitations rather than the study's effectiveness or observed outcomes.

4.1 Design Contribution and Implications

The use of AI-generated US-images [13] demonstrates that synthetic medical images can serve as flexible training artifacts. Using a CGAN via an ONNX-based pipeline [17], the architecture separates model development from training system implementation. Such a separation reinforces adaptability and maintainability, which are important in educational environments. We use a gyro-sensor with a physical breast phantom to perform basic US-probe interactions without needing elaborate medical apparatus. The design can be integrated into classroom, and simulation lab environments, where it provides a well-defined interaction model to help a novice ultrasonographer see the connection between the motion of the US-probe and changes in the US-image with integrated decision support text.

4.2 Computational Feasibility in a Training Context

The observed findings give a first indication about the computational feasibility of deploying artificial intelligence generated US-images in MR training setup. Instead of benchmarking performance, this work focuses on how different image resolutions affect interaction responsiveness within the same system. It would be reasonable to assume that low-resolution images are preferable for interaction-intensive training tasks, while higher resolution images might be used for reference viewing. This results in a common trade-off in the design of training systems; in this context, interaction quality takes priority over visual detail. The results suggest that AI-generated images can be applied in training-oriented MR systems using commonly available computing devices.

4.3 Limitations

A number of limitations need to be recognized. First, since the system proposed is only a conceptual prototype, a formal user study was not carried out. As a result, learning efficiency and usability were not quantitatively assessed. Second, the computational observations were made by using a single hardware configuration and do not represent general system performance. Furthermore, interaction design is limited to fundamental US-probe orientation associated with interaction with a physical mannequin, failing to account for such factors as force feedback or tissue deformation. The mannequin allows for a relatively stable and controlled training environment but does not represent the variability in human anatomy. These features would be relevant to advanced ultrasound training but were beyond the scope of this study, which is aimed at a novice ultrasonographer.

4.4 Future Work

Future work will focus on implement a functional prototype and have user studies with a novice ultrasonographer. It seems that design of interaction should be improved for further interaction, where there are additional feedback mechanisms for better training experience. We will also investigate system optimization and evaluation in different hardware environments to facilitate wider deployment.

5. Conclusion and Future Work

In our paper, we described a conceptual prototype design of an immersive MR ultrasound training system for a novice ultrasonographer. The proposed method leverages AI-generated ultrasound images with sensor-based interaction and MR visualization for experience-based training. Instead of discussing full system implementation or performance comparisons, the emphasis of the paper is on design rationale, user experience, interaction design, and practical feasibility concerns. In the paper, we demonstrate that US-images generated by conditional GANs can be incorporated into a MR training environment using a training-oriented image configuration.

From the reported observations it is clear that with appropriately fine-grained images generated by AI, interactive training will still occur with minimal computational load. This provides practical guidance towards designing MR training systems that balance the quality of interactions with constraints imposed by systems.

Moving further will be practical development of a research-focused functional prototype suited for user studies testing using a beginner ultrasonographer. More refinement of interaction design will be added to this, including more mechanisms for feedback to improve effectiveness. It is also intended to conduct broader testing across system configurations to facilitate scalable deployment in education.

6. References

[1] K. Y. Leung, K. Bala, J. Y. Cho, S. Gokhale, A. Kikuchi, P. Liang, C. L. Ong, Q. B. Nguyen-Phuoc, T. Wataganara and Y. L. Wan, "Utility and challenges of ultrasound education for medical and allied health students in Asia," *Hong Kong Med J,* vol. 30, no. 1, pp. 75-79, Feb 2024.

[2] C. F. Dietrich, C. Lucius, M. B. Nielsen, E. Burmester, S. C. Westerway, C. Y. Chu, G. Condous, X.-W. Cui, Y. Dong, G. Harrison, J. Koch, B. Kraus, C. P. Nolsøe and L. Nayahangan, "The ultrasound use of simulators, current view, and perspectives: Requirements and technical aspects (WFUMB state of the art paper)," *Endosc Ultrasound,* vol. 12, no. 1, pp. 38-49, 2023.

[3] M. E. Nylund, S. Jain, E. Tegnander, . E. J. L. Jensen, . E. Prasolova-Førland, F. Linsdeth and G. Kiss , "Mixed reality training application to perform obstetric pulsed-wave Doppler ultrasound," *Education and Information Technologies,* vol. 29, p. 7519–7551, 2024.

[4] A. Sawaguchi, Y. Sugamiya, A. Yamada, T. Kaizoji and A. Takanishi, "Development of Probe-shape Controller and Training Display for the Ultrasound Home-Training Simulator for Beginners," *Journal of the Robotics Society of Japan,* vol. 42, no. 5, pp. 447-480, 2024.

[5] B. KARTHIK, "A Review of cGANs for Text-to-ImageGeneration," *International Journal of Research Publication and Reviews,* vol. 5, no. 11, pp. 7810-7814, November 2024.

[6] Y. Hu, E. Gibson, L.-L. Lee, W. Xie, . D. C. Barratt, T. Vercauteren and J. A. Noble, "Freehand Ultrasound Image Simulation with Spatially-Conditioned Generative Adversarial Networks," in *Molecular Imaging, Reconstruction and Analysis of Moving Body Organs, and Stroke Imaging and Treatment. RAMBO CMMI SWITCH 2017 2017 2017. Lecture Notes in Computer Science()*, Quebec City, Canada, 2017.

[7] S. Liu, Y. Wang, X. Yang, B. Lei, L. Liu, S. X. Li, D. Ni and T. Wang, "Deep Learning in Medical Ultrasound Analysis: A Review," *Engineering,* vol. 5, no. 2, pp. 261-275, 2019.

[8] I. Goodfellow, Y. Yere, D. Warde-Farley, J. Pouget-Abadie, M. Mirza, B. Xu, Y. E. Rachmad, S. Ozair and A. Courville, "Generative Adversarial Networks," 2014.

[9] R. J. Chen, Y. M. Lu, Y. T. Chen, D. F. K. Williamson and F. Mahmood, "Synthetic data in machine learning for medicine and healthcare," *Nat Biomed Eng,* vol. 5, no. 6, pp. 493-497, 2021.

[10] A. S. Coyner, J. S. Chen, K. Chang, P. Singh, S. Ostmo, R. V. P. Chan, . M. F. Chiang, . J. Kalpathy-Cramer and . J. P. Campbell , "Synthetic Medical Images for Robust, Privacy-Preserving Training of Artificial Intelligence: Application to Retinopathy of Prematurity Diagnosis," *Ophthalmology Science,* vol. 2, no. 2, Elsevier.

[11] X. Wang, J. Yang, B. Zhou, L. Tang and Y. Liang, "Integrating mixed reality, augmented reality, and artificial intelligence in complex liver surgeries: Enhancing precision, safety, and outcomes," *iLIVER,* vol. 4, no. 2, June 2025.

[12] J. Potočnik, S. Foley and E. Thomas, "Current and potential applications of artificial intelligence in medical imaging practice: A narrative review," *Journal of Medical Imaging and Radiation Sciences,* vol. 54, no. 2, pp. 376-385, June 2023.

[13] K. Kawattikul and K. Hamamoto, "CGAN-Based Cross-Sectional Ultrasound Image Generation from Probe-Pose Vectors," in *2025 IEEE International Conference on Imaging Systems and Techniques (IST)*, Strasbourg, France, 2025.

[14] J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cognitive Science,* vol. 12, no. 2, pp. 257-285, 1988.

[15] J. Nielsen and R. Molich, "Heuristic evaluation of user interfaces," in *ACM CHI'90*, Seattle, 1990.

[16] WitMotion Shenzhen Co.Ltd, "User manual-BWT61CL Bluetooth 2.0 AHRS IMU sensor," 2020. [Online]. Available: http://wiki.wit-motion.com/english. [Accessed 10 January 2026].

[17] "Open Neural Network Exchange (ONNX)," [Online]. Available: https://github.com/onnx/onnx.