Echo-Robot: Semi-Autonomous Cardiac Ultrasound Image Acquisition Using AI and Robotics

Eliott Laurent, Raska Soemantoro, Kathryn Jenner, Attila Kardos, Gilbert Tang, and Yifan Zhao, *Senior Member, IEEE*

***Abstract*—Echocardiography is a critical tool for diagnosing cardiovascular diseases, offering detailed insights into heart functions. However, its accessibility is currently limited by a shortage of trained sonographers, specific skill requirements, and the physical strain imposed on professionals during repetitive procedures. This article introduces a new robotic system designed to automate the acquisition of transthoracic echocardiography (TTE) images. The system autonomously adjusts the position and orientation of the ultrasound transducer based on analysing real-time ultrasound images, without relying on tomographic data or depth sensors. Initially, the transducer is manually placed on the subject's skin, and the system uses a deep learning approach to grade the quality of ultrasound images captured at each position. The robot then adjusts its position by spiralling outwards from the starting point, moving to the location with the highest image quality score. Next, the system fine-tunes the transducer's orientation in 5-degree increments along all three axes of rotation, informed by another deep learning module to identify the field of view. The robotic system was tested using a cardiac simulator, achieving approximately 80% accuracy in acquiring the A4Ch view when the probe was initially positioned randomly in a 6 by 6 cm area beneath the left nipple. The impact of this work would be rapid diagnostics in the Emergency Departments to reduce the length of stay in hospitals, a reduction of hospital admissions related to heart disease by accessing local healthcare communities, acceleration of clearing the post-Covid backlog, and improved quality of life and longevity of patients.**

***Index Terms*—Cardiac ultrasound, echocardiography, machine learning, autonomous robotics.**

I. Introduction

ULTRASOUND is a safe and cost-effective diagnostic technology used for various medical conditions. The challenge for echocardiography is to provide full anatomical and functional imaging of the heart while mitigating imaging artifacts caused by different types of tissues surrounding the heart, such as lung tissues filled with air, bones that form the ribcage, blood, muscle, and fat tissues. These tissues can introduce attenuation, shadowing, and reflections artifacts when they lie between the heart and the ultrasound probe. Although image acquisition of echocardiography is standardised, ideal locations of acquisition are very narrow and dependent on the position of the heart relative to the ribs and the lungs (known as the echo window). The standard echocardiographic views are shown in Fig. 1, where the two most important echo windows (parasternal and apical areas) and associated typical views with different angulations are highlighted. It usually takes at least 10-15 mins, even for an experienced sonographer, to find the right place and angle for the transducer, using the movement of sliding, tilting, rotating, rocking, and compressing, to acquire artefact-free and non-foreshortened echo images that are appropriate for meaningful clinical use. To date, no transducer can find the right echo window without the sonographer's manual adjustment. According to the Society of Radiographers, the vacancy rate for sonographers in the NHS increased from 6.7% in 2019 to 13.4% in 2023. This was attributed to factors such as increased demand for ultrasound services, difficulty in recruiting and retaining staff, and the impact of the COVID-19 pandemic on the workforce [1]. Moreover, over the years, a high proportion of sonographers suffered from repetitive shoulder injuries and have been out of work for a long time or left the profession. The increasing mismatch between this capacity and demand urgently requests an innovative approach to fill the gap.

Since the late 20th century, automating ultrasonic scanning with robotics for medical applications has attracted significant innovations and research, broadly categorised into teleoperated, semi-autonomous, and autonomous robots.

A. Teleoperated Robots

These systems typically consist of a robot that manipulates the ultrasound probe, and a control station located remotely where a doctor or medical expert manipulates a dummy ultrasound probe. Examples include the MGIUS-R3 [2] from MGI Tech and the MELODY system [3] from AdEchoTech. The MGIUS-R3 uses a 6 degrees of freedom (DOF) manipulator, while the MELODY system is positioned by a local assistant, with a doctor controlling fine adjustments remotely. Wang et al. [4] and Jiang et al. [5] demonstrated the effectiveness of the MGIUS-R3 in remotely diagnosing COVID-19 patients. The MELODY system has been widely used in abdominal [6], [7], vascular [6], obstetric [8], [9], and cardiac [10] ultrasound imaging. Georgescu et al. [9]

Received 7 January 2025; revised 21 March 2025 and 17 May 2025; accepted 15 July 2025. Date of publication 18 July 2025; date of current version 21 August 2025. This article was recommended for publication by Associate Editor V. Iacovacci and Editor P. Dario upon evaluation of the reviewers' comments. *(Corresponding author: Yifan Zhao.)*

Eliott Laurent, Gilbert Tang, and Yifan Zhao are with the Faculty of Engineering and Applied Sciences, Cranfield University, MK43 0AL Bedford, U.K. (e-mail: yifan.zhao@cranfield.ac.uk).

Raska Soemantoro is with the Department of Mechanical and Aerospace Engineering, The University of Manchester, M13 9PL Manchester, U.K.

Kathryn Jenner is with the Simulation Division, Intelligent Ultrasound, CF10 1DY Cardiff, U.K.

Attila Kardos is with the Translational Cardiovascular Research Group, Department of Cardiology, Milton Keynes University Hospital NHS Foundation Trust, MK6 5LD Milton Keynes, U.K.

Digital Object Identifier 10.1109/TMRB.2025.3590471

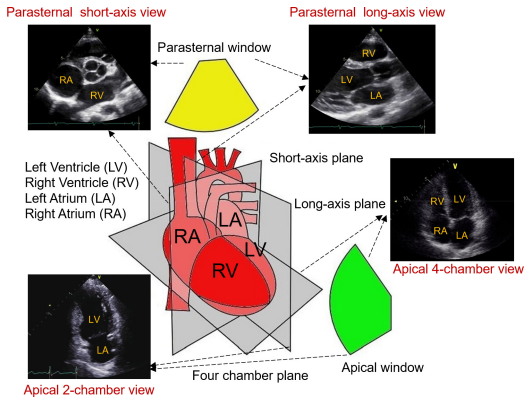


Fig. 1. The standard TTE views.

found that teleoperated robotic systems significantly reduced waiting times while producing images of comparable quality to standard techniques.

Several non-commercial robotic systems have been developed for research purposes, such as the TOURS (Tele-Operated Ultrasound System), which has been tested on over 100 patients for pelvic, obstetric, vascular, and abdominal exams [11], and even used on International Space Station for remote diagnosis [12]. For TTE, Arbeille et al. [13] developed a system with a 3-DOF robotic arm for probe orientation and a motorised plate for translational movements.

B. Semi-Autonomous Robots

Fang et al. [14] developed a robotic ultrasound system using a UR5 robot equipped with two force sensors to assist sonographers in applying the correct amount of force to the probe. This collaborative system significantly reduces the physical strain on sonographers. The study reported a reduction in human-applied force by 32%, along with an improvement in image stability by 8%, mitigating risks such as shoulder injuries caused by the repetitive application of high forces during scanning. However, while this approach addresses ergonomic challenges, it does not address the shortage of skilled sonographers or focus on fully automated scanning procedures, which are critical issues in echocardiography. Carriere et al. [15] used a 7-DOF robot to automatically scan a breast phantom (once placed on it), applying constant force—an essential factor for achieving consistent imaging of the highly deformable breast tissue. Their method, designed to ensure uniform force application across the entire breast, is well-suited to the relatively homogeneous and pliable nature of breast anatomy. However, cardiac ultrasound poses unique challenges, relying on accurate probe positioning within a narrow echo window rather than comprehensive area scanning. This fundamental difference limits the direct applicability of their approach to echocardiography.

Similarly, a system designed for orthopaedic scans [16] adjusts the probe's orientation to maintain perpendicularity to the contact surface, using only force measurements and an ultrasound confidence map, eliminating the need for external sensors. While effective for rigid and predictable anatomical structures, this technique is not directly applicable to cardiac imaging, where probe orientation depends on patient-specific anatomy and is not perpendicular to the skin.

In cardiac imaging, Kim et al. [17] developed a semi-autonomous system for monitoring left ventricular ejection fraction. A notable feature of their system is its repositioning function, which adjusts the probe position when the initial ultrasound image is deemed "partially satisfactory" by a convolutional neural network (CNN). This functionality assumes that the region with the highest intensity corresponds to the left ventricle and re-centres it within the image frame. However, the system's ability to adjust the probe relies on the initial visibility of the left ventricle, limiting its effectiveness in cases where the heart is entirely out of the probe's field of view. Currently, the only commercially available semi-autonomous product is the ABUS system by GE HealthCare [18], designed for breast ultrasound. While a physician manually positions the probe and applies pressure, the system scans the entire breast with a large-footprint transducer in a single sweep. This approach has proven effective for breast cancer diagnosis [19]. However, it is not directly applicable to cardiac ultrasound, where the objective is to locate specific anatomical windows for imaging the heart, rather than scanning an entire region.

C. Autonomous Robots

Recent examples of fully automated robotic ultrasound systems include RobUSt [20], a fully automated system for scanning jugular veins. The system is composed of a 7-DOF robot with integrated torque sensors and an RGB-D camera. The system uses an impedance controller to manage force, applying an average of 6 newtons. The 3D camera captures a point cloud of the phantom's surface, which is then used for path planning. Another example is the system developed by Tan et al. [21], designed for fully autonomous breast scanning. This system features a gantry-like structure, two robotic arms, a multi-structured light system for acquiring a point cloud of the scanning surface, a human-computer interface, and a flexible ultrasound probe clamping device. While innovative, the system's high complexity raises significant concerns about its scalability and cost-effectiveness.

Deng et al. [22] introduced a portable robot-assisted device for autonomous ultrasound acquisitions that incorporates an admittance force control algorithm and a robot-specific pose feedback mechanism to ensure safety and maintain image quality. While the system can autonomously perform follow-up scans, the initial scan still requires the expertise of a skilled sonographer. Nevertheless, leveraging patient data to enhance the accuracy of follow-up diagnostics represents a promising advancement in the field.

In the field of transthoracic echocardiography, Shida et al. [23] designed a robot specifically for acquiring parasternal views. Its innovative design allows the patient to lie directly on the robot, with the key advantage being that if excessive force is applied, the patient can simply straighten up to relieve the pressure. The system demonstrated promising preliminary results during human testing on six participants; however, its scope is limited to the parasternal window. The

design may not be easily adaptable to other imaging windows, such as the apical window.

Tang et al. [24] introduced an autonomous ultrasound scanning robotic system that can locate a region of interest to scan automatically. Then a "snake-like" pattern was used to scan this whole region without stopping at the position where the heart is fully visible. However, the system only adjusts orientation within the in-plane axis, which may be insufficient for real-world applications, as sonographers typically adjust orientation across all axes.

D. Automated Ultrasound Path Planning

In general, path-planning methods for 2D ultrasound can fall under 3 categories: greedy, landmark-based, and feedback-based [25]. Greedy systems rely on comprehensive data collection over a region to obtain a final view of the specific anatomy, which often involves post-processing. In contrast, landmark-based systems require input of a visual element of the body's surface, which is then used to approximate the imaging position. Feedback-based systems utilise sensory data fed in real-time into a rule-based system that moves the transducer.

These three methods have their own benefits and disadvantages. However, the main importance is how they fit according to the needs of the specific imaging routine. For echocardiography, implementation of such automated ultrasound methods has been limited. Despite this, some notable works include the work of Tang et al. [24] which uses a mix of both greedy and feedback systems, and that of Shida et al. [23] which resembles a greedy system, as it involves initial path planning using an exhaustive scan of the anatomy. Considering this limited number of works, the authors believe that this topic is an open area of research that needs to be addressed.

In summary, research on autonomous robotic ultrasound systems for TTE is still very limited. AI has not been fully leveraged to control robots to find the optimal position and orientation of transducers based on real-time image analysis. Moreover, there are no widely available robotic manipulators designed for this purpose. Custom designs are often more challenging to deploy at scale and can be more expensive. This limitation is partly due to the specific probe handling, positioning techniques, and haptic feedback required for echocardiography, which differ from those used in other surface-based ultrasound applications. Another challenge is the use of specific phased-array transducers in TTE and the difficulty of obtaining high-quality images of the heart due to its fast movement compared to other organs.

This paper introduces a first-of-its-kind semi-autonomous robotic ultrasound system for echocardiography, with the potential for full autonomy, capable of acquiring images from both parasternal and apical views. The novelty of this study lies in developing an intelligent path-planning method for automatically guiding an ultrasound transducer to localise the heart at desired echocardiographic views. The main contributions of this study include:

1. An architecture of Echo-Robot that interconnects a robot, ultrasound transducer, and essential software.
2. Two dedicated AI modules for real-time evaluation of ultrasound image quality and echo views.
3. A spiral path planning strategy to effectively localise the appropriate echo windows based on AI outputs with minimum steps.

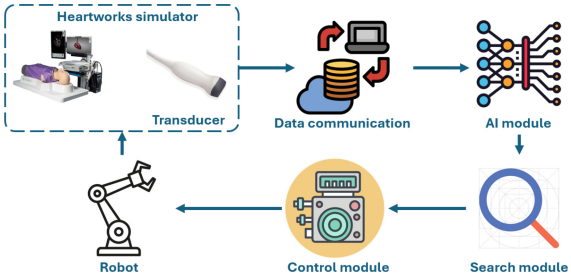


Fig. 2. Key components of Echo-Robot with interlinks.

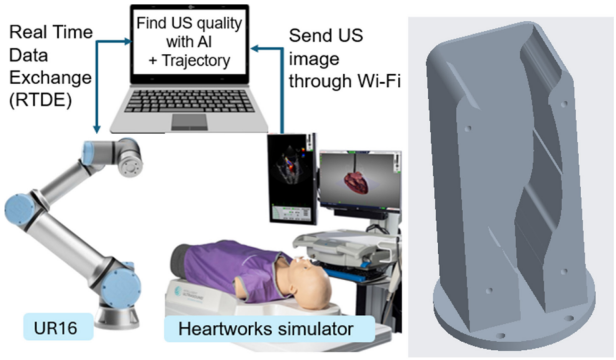


Fig. 3. (a) Echo-Robot hardware setup architecture; (b) the design of the probe end effector.

II. Methodology

As illustrated in Fig. 2, the proposed Echo-Robot consists of six key components: (a) the Heartworks simulator as the testing platform, (b) a data communication module for real-time transfer of ultrasound images to a local laptop, (c) AI modules to assess image quality and identify echo view types, (d) a search module to determine the next movement strategy, (e) a control module to manage the robot's movements, and (f) a robotic arm to position and orient the transducer as needed. The details of each component and the testing method are introduced below.

A. System Hardware and Setup

The hardware system used in this research consists of three main components: the UR16e collaborative robotic manipulator developed by Universal Robots, the Heartworks simulator from Intelligent Ultrasound, and an Acer Swift 5 laptop. The setup of the system with associated data communication is illustrated in Fig. 3(a). The HeartWorks simulator is a state-of-the-art device designed to replicate TTE with high fidelity. The manikin's texture closely mimics that of a real patient, including anatomical structures such as the ribcage. Simulated ultrasound images are generated based on the probe's position, orientation, and the pressure applied.

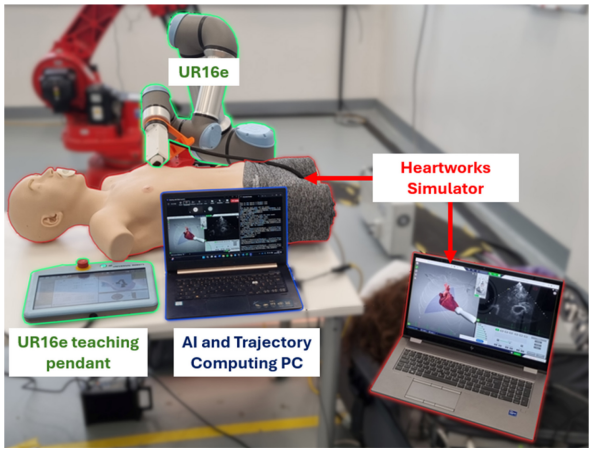


Fig. 4. An implementation snapshot of Echo-Robot.

As in real-world echocardiography, achieving an optimal acoustic window requires adjusting for various obstructions, including lung tissue, the ribcage, and other anatomical and physiological factors that affect ultrasound propagation. To automate and standardise the ultrasound scanning process, a collaborative robotic system is employed to manipulate the probe with precision and consistency. Collaborative robots, such as the UR16e, are frequently used in robotic ultrasound systems due to their embedded safety features, they use force sensors to automatically stop the robot if too much force is applied, which allows them to operate alongside humans. Unlike traditional robotic manipulators that can injure anyone in their path, collaborative robots, or cobots, are designed for safe human interaction and do not need to be enclosed in a protective cage. The UR16e has a payload capacity of 16 kg and a reach of 900 mm, which is sufficient for our use case, as the end effector weighs around 440 g, and the probe is required to move within a range of just a few centimetres. Its 3D force torque sensor has a rated precision of 5 Newtons. It was used by the *force_mode* function to maintain a constant force in the z-direction. Additionally, the UR16e has 6 DOF, enabling it to move in 3D space and achieve any combination of position and orientation within its reach. This capability is essential for manipulating the transducer in TTE. To attach the ultrasound probe to the robot, a custom 3D-printed part was created, as shown in Fig. 3(b). This part is an extrusion of the specific probe's contour, provided by the Heartworks simulator. A plate was secured to this part with screws, ensuring the probe remains fixed in its holder. The end effector itself can be mounted to the UR16e using just four M4 screws. Fig. 4 displays a snapshot of the experimental setup.

B. Communication Module

One of the main challenges in this research was defining an effective communication architecture, as shown in Fig. 3(a). Seamless interaction was required among the robot arm, data processor, and the data acquisition system. For proof of concept, a communication channel was established using Microsoft Teams to transfer simulated ultrasound images to the master computer running AI modules.

The master computer communicated with the robotic manipulator using Real-Time Data Exchange (RTDE) over Ethernet [26]. RTDE messages were predefined using "recipes" specifying input and output data types. Since the robot's tool centre point (TCP) pose—a vector with six elements—could not be directly sent, individual components were transmitted as doubles, representing the robot's desired position and orientation. An advantage of RTDE is compatibility with URScript, the programming language for Universal Robots.

C. AI Modules

This module includes two AI models, as detailed below.

1) Image Quality Assessment Model: This model was designed to assign a quality metric to each real-time image obtained from the ultrasound feed. This quality metric was then used for the window search algorithm. The image quality assessment was performed by a convolutional neural network (CNN), which classifies ultrasound images as either 'good' (positive) or 'bad' (negative) based on a novel echocardiography image quality dataset.

The model was trained using an image quality dataset containing 5,166 images, with 2,852 classified as 'bad' and 2,314 as 'good'. The dataset consists of three parts:

- The first part includes 1,137 anonymised images (22% of the dataset) from patients collected from Milton Keynes University Hospital using a Philips Epique CVX medical cardiac ultrasound system.
- The second part comprises 3,679 synthetic images (71% of the dataset) generated by the HeartWorks ultrasound simulator.
- The third part contains 346 anonymised ultrasound images (7% of the dataset) collected from a healthy participant using a portable GE Vscan Extend ultrasound device.

All images in the dataset were manually labelled based on reference views pre-labelled by healthcare professionals.

The classification model employs the InceptionV3 architecture, consisting of 13 layers and accepting an input size of $299 \times 299 \times 3$. Illustrated in Fig. 5, this model was selected as it provides a balance between accuracy and training/inference time, considering the real-time nature of this application. Other models considered include the VGG16 and VGG19 architectures, but they were found to be too computationally expensive for this application.

The InceptionV3 model was implemented using the TensorFlow machine learning framework, which provides the InceptionV3 model through the *keras.applications* submodule. Prior to training, the images in the dataset were resized to 299×299 pixels and normalised to a range of 0 to 1, improving pixel data distribution and reducing convergence time during training. The model's weights were initialised using the pre-trained ImageNet checkpoint, and training was conducted over 200 epochs with an 80/20 training/validation split of the image quality dataset. Additionally, a final output gate, comprising two sigmoid layers, was added to the InceptionV3 architecture to provide a probabilistic

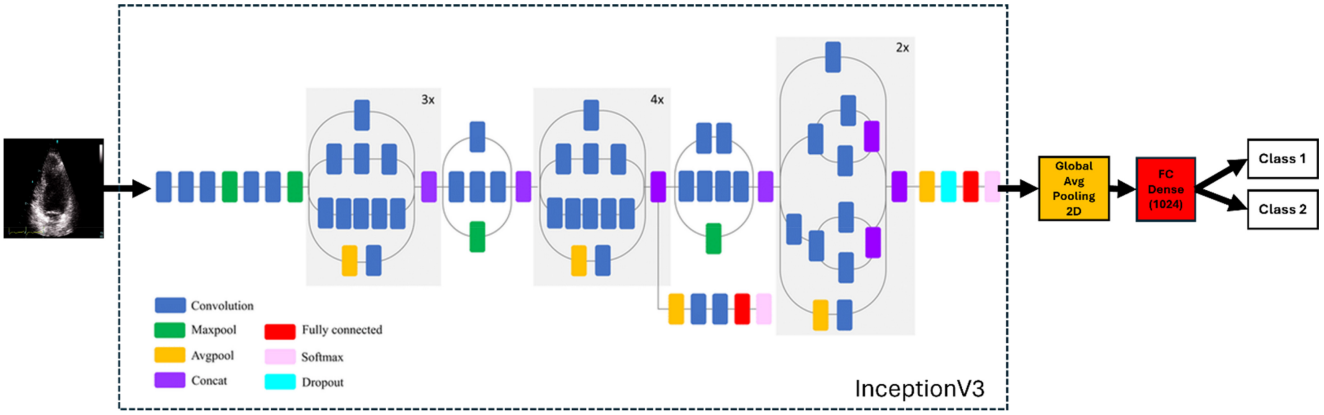


Fig. 5. Diagram of the InceptionV3 models used in this work. Class 1 and Class 2 refer to 'good' and 'bad' as well as 'A4Ch' and 'A2Ch' labels.

representation of the classification results. The model outputs a quality score between 0 and 1, indicating the confidence level that the image is of 'good' quality.

2) Echo-Window Classification Model: This module will automatically classify view types of images: apical 4-chamber (A4Ch) view or apical 2-chamber (A2Ch) view. This step is essential as it will support the system to determine if the scan is completed by checking images from different views. Furthermore, this step may help the search process to find the next view rapidly based on the current view. The approach can be extended to other echocardiographic views. The CNN used in this model closely mirrors the one implemented for image quality assessment but was trained on a different dataset. This dataset contains 12,980 images extracted from the HMC-QU public echocardiographic videos dataset [27]. The dataset includes 6,500 A2Ch images and 6,480 A4Ch images, with all images resized to 299×299 pixels. Approximately 65% of the data were collected from individuals with myocardial infarction, while the remaining 35% were from healthy subjects. All these images were considered as having 'good' quality.

The A4Ch view offers a view of all four heart chambers, while the A2Ch view primarily focuses on the left ventricle and left atrium. Both views are essential for assessing cardiac structure and function, which is why they are the focus of this study [28]. Additionally, both views are acquired from the same position—rotating the transducer from the A4Ch view anticlockwise reveals the A2Ch view. This research simplified the echocardiography image search problem into a 2-dimensional domain, without accounting for the transducer's rotational angle.

After pre-processing, the window classification model was trained for 150 epochs using an 80/20 training/validation split of the HMC-QU dataset. TABLE I provides the configurations used to train the two models [25].

TABLE I
TRAINING PARAMETERS FOR AI MODELS

	Image quality grade module	Video classification module
No. of training epochs (set)	200	150
No. of training epochs (true)	156	118
Batch size	32	32
Optimiser	Stochastic gradient descent	Stochastic gradient descent
Loss	Sparse categorical cross-entropy	Sparse categorical cross-entropy
Learning rate	0.00001	0.00001
Momentum	0.9	0.9
Early stopping condition	No improvement in validation loss during 3 consecutive epochs	No improvement in validation loss during 3 consecutive epochs

D. Search Module

One of the most challenging aspects of this study was developing the search algorithm. In many robotic manipulator applications, the robotic arms follow a predefined trajectory. However, in this study, such an approach is not feasible since the position of the patient's heart is unknown relative to the robot. As a result, the robot must continuously update its understanding of its environment by using the ultrasound probe and adjusting its position based on real-time measurements. This dynamic adjustment of the robot's path during movement is known as feedback-based path planning.

This study focuses on navigation once the ultrasound probe is in contact with the patient's skin, as the robot initially lacks knowledge of the patient's position relative to itself. Once the probe is placed on the skin, the robot can begin navigating using ultrasound data. The initial placement of the probe can be performed by either the patient or a nurse, without requiring precise positioning or specialized skills.

The strategy employed in this project is divided into two phases. In the first phase, the robot adjusts the transducer to a default orientation that is expected to be close to the ideal position. The default orientation in the world reference frame was set to 2.193 radians for the x-axis, 0.610 radians for the y-axis and −0.701 radians for the z-axis. The initial transducer orientation was set according to the recommended guidelines

TABLE II
PARAMETERS OF THE DIFFERENT SPIRAL TRAJECTORIES

	Growth rate (mm)	Angular rate (rad)	Steps
Initial spiral	4.2	0.361	10
'Small' spiral	2	0.628	10
'Big' spiral	5	0.471	10

from the transducer's manufacturer. This orientation is defined with reference to the robot arm's base frame. The robot then fine-tunes its position to identify coordinates that have the desired ultrasound image quality. In the second phase, the system further refines the transducer's orientation to enhance image quality. This adjustment is guided by the image quality assessment model. The search strategy is illustrated in Fig. 6.

Once the probe is placed on the skin, the robot initiates a spiral scanning pattern, as depicted in Fig. 7. This strategy differs from the method used by medical professionals, who typically start at the 5th intercostal space and make fine adjustment to the probe to obtain the A4Ch view. For a robotic system, it is challenging to reliably identify the 5th intercostal space without relying on additional data, such as tomographic data. The advantage of the proposed search strategy is that it enables a simpler and more cost-effective system, given that an untrained medical professional is able to initially place the ultrasound probe relatively close to the ideal position. In the study by Morgan et al. [29], untrained medical students had a mean displacement error of 3.2 cm, which falls within the acceptable error margin for our system. The spiral trajectory is defined by three parameters: growth rate, angular rate, and total number of steps. Increasing the growth rate causes the spiral to expand outward more quickly while increasing the angular speed reduces the total number of steps required to complete a full rotation. During the first spiral, these settings are set to 4.2 mm for the growth rate, 0.3611 radians for the angular rate and 10 steps. If the image quality is slightly below the allowable quality threshold, the robot will start a small spiral trajectory since it is probably close to the ideal position. If the quality score is low, the robot will conduct a bigger spiral because it is probably not near the ideal position. The different parameters are detailed in TABLE II.

At each position along the spiral, the ultrasound quality score was recorded from the first AI module. This score was calculated as the average of four images to enhance reliability. The robot then moves to the position with the highest score. If this score exceeds a set threshold 50%, as scores below this are deemed poor by the AI the robot proceeds to adjust the probe's orientation. If the score is below the threshold, the robot repeats the spiral scan from the point with the best score until a position with a sufficiently high score is identified. Once such a position is found, the robot adjusts the probe's orientation accordingly.

For instance, in Fig. 7, the trajectory begins at the centre of the black spiral, marked with a blue cross. The position indicated by the black cross represents the point with the highest score, but it remains below the threshold. As a result, the algorithm initiates a new trajectory (shown by the red spiral), starting from this high score point. The robot then moves to the position with the overall best score, indicated by the red cross, even if this score is below 50%. This approach is taken because if the quality remains low after two spiral scans, it suggests that the issue may lie with the probe's orientation rather than its position.

For the orientation optimisation, the robot adjusts the probe's tilt in 5-degree increments along the z-axis. After each adjustment, it evaluates the image quality and moves to the position with the best score. The robot then continues to refine the orientation by adjusting along the x-axis and subsequently along the y-axis. This systematic approach ensures that the

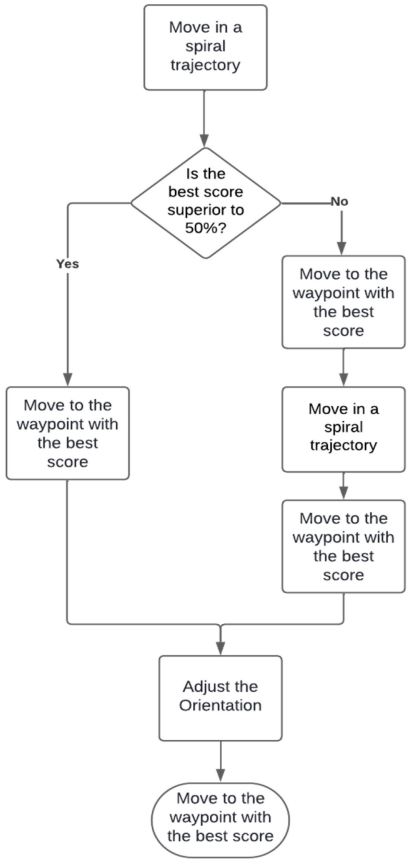


Fig. 6. The adjustment process of position and orientation.

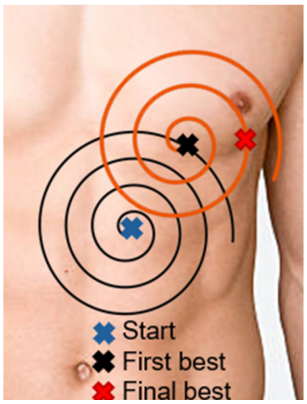


Fig. 7. The proposed spiral search strategy.

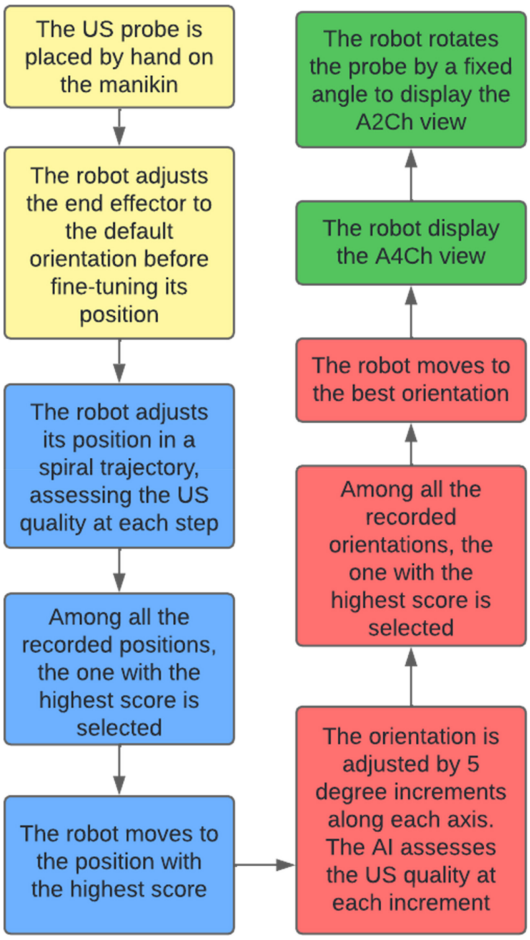


Fig. 8. The complete process of the Echo-Robot.

TABLE III
CONFUSION MATRIX OF THE IMAGE QUALITY ASSESSMENT MODEL

	True 'bad'	True 'good'
Predicted 'bad'	0.998	0.014
Predicted 'good'	0.002	0.986

probe's orientation is optimised to achieve the best possible image quality.

The strategy outlined was initially used to locate the apical A4Ch view, as sonographers typically start by finding this view before attempting to transition to other views, such as the A2Ch view, by rotating the probe. This approach was adapted for this project. The ideal positions for both A4Ch and A2Ch views were determined based on the robot's actual pose. The transition from the A4Ch view to the A2Ch view is then calculated as the difference between these two waypoints. The complete process of the Echo-Robot is summarised in Fig. 8. The process begins with the initialisation phase (yellow), followed by the position adjustment phase (blue), the orientation adjustment phase (red), and concludes with the final acquisition phase (green). At each new step, the AI module is used to assess the quality of the ultrasound image.

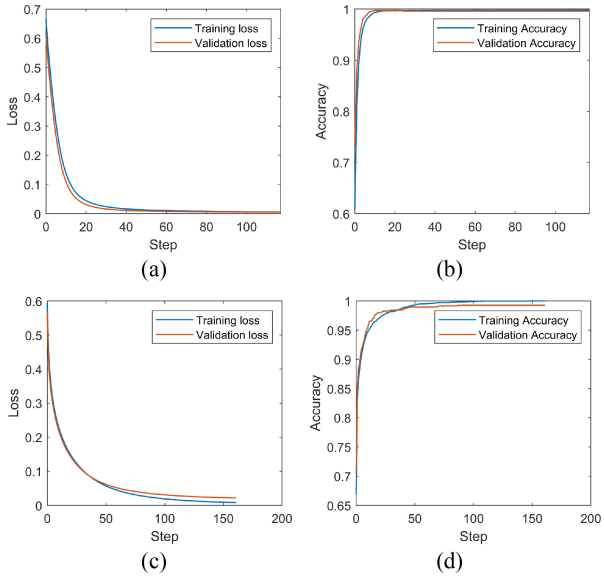


Fig. 9. Training and validation performance: (a) loss and (b) accuracy for the echo window classification model; (c) loss and (d) accuracy for the image quality assessment model.

TABLE IV
CONFUSION MATRIX OF THE WINDOW CLASSIFICATION MODEL

	True A2Ch	True A4Ch
Predicted A2Ch	0.997	0.005
Predicted A4Ch	0.003	0.995

III. RESULTS

TABLE III presents the confusion matrix for the image quality grading module. The number of false negatives is slightly higher than that of false positives, but the overall accuracy is more than 99%. The results of the window classification model are shown in TABLE IV. The model demonstrates a final accuracy of over 99% in recognising both apical views. Detailed loss and accuracy of the training dataset and validation dataset can be found in Fig. 9. While further improvements could be achieved by tuning the hyperparameters, it is noted that the research scope of this work lies mainly in the proof-of-concept of system development.

The reliability of the whole Echo-Robot system was evaluated by testing it in two distinct regions, as depicted in Fig. 10. The first region, shown in blue, is a 6 cm by 6 cm square, with its top-right corner positioned just below the left nipple. The second region, shown in red, forms an L-shape, defined by a 10 cm by 10 cm square from which the blue square has been subtracted.

The ultrasound probe was randomly placed within these two regions, and the resulting image quality was assessed using two metrics. The first metric was an AI-generated quality score, applied exclusively to the A4Ch view. The second metric was a quality rating assigned by an external observer on a scale of 1 to 5, with the criteria for each grade outlined in TABLE V. The probe was manually positioned at random orientations and locations within the defined zones. A total of 32 samples were collected from the blue area and 24 from

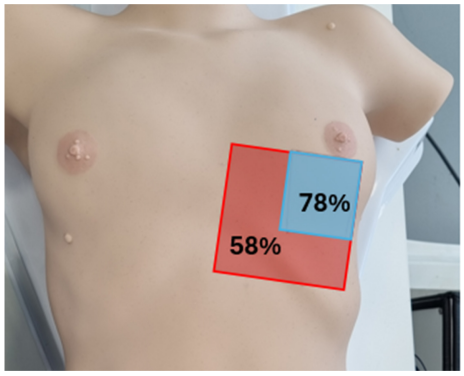
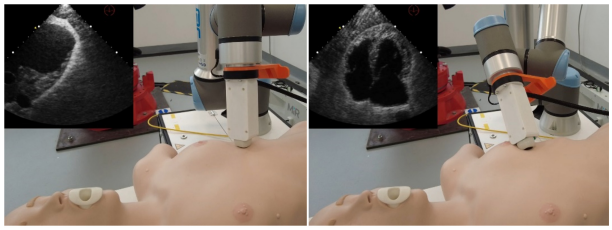


Fig. 10. The first (blue) and second (red) testing zones with corresponding scan performance.



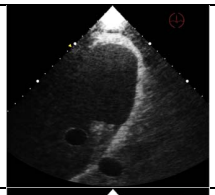
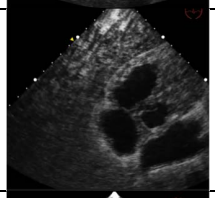
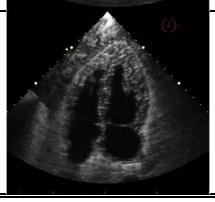
(a) (b)

Fig. 11. Snapshots of the initial (a) and the optimised positions (b) of the transducer.

the red area. Fig. 11 illustrates the initial pose of the probe compared to its position after the robot has located the heart. On average, the robot took between 1 minute 30 seconds and 2 minutes 30 seconds to complete a scan, performing approximately 25 steps with the current configuration. This time could be easily reduced by using more advanced hardware made for AI, which is significantly shorter than the average time for a limited transthoracic echocardiogram conducted manually, approximately 4 minutes, as suggested by Ferrada et al. [30]. In the A4Ch view, within the blue zone, 7 out of 32 scans were graded below 4, indicating insufficient quality for diagnostic purposes, resulting in an accuracy of 78%. Of the 32 scans, 23 were graded 5, and 2 were graded 4, meaning the system achieved perfect results in 72% of cases. The low number of images graded as 4 is largely due to the AI's stringent evaluation of off-centre images, which are often classified as "bad" with high uncertainty. In comparison, images graded as v3 are also classified as "bad," but with significantly lower uncertainty.

For the A2Ch view, diagnostic-quality scans were obtained in 18 out of 32 attempts (56%), with 13 scans graded 5 and 5 graded 4. Performance was significantly lower in the red zone. For the A4Ch view, 10 out of 24 attempts failed to meet diagnostic standards, while 6 attempts were unsuccessful for the A2Ch view. As expected, the closer the probe's initial position is to the heart, the better the system's performance, regardless of the view being captured. More detailed results are presented in TABLE VI. Comparing these findings with the state of the art in autonomous cardiac ultrasound scanning, Shida et al. [23] achieved a 63% success rate in obtaining high-quality parasternal long-axis views when testing on real human

TABLE V
IMAGE QUALITY GRADING FOR THE A4CH VIEW

Grade	Description	Example image
1	Bad image with no heart visible	
2	Bad image, the heart is visible but none of the valves are visible and not all chambers are visible	
3	Image is not good enough for a diagnosis but close to being good enough.	
4	Image is good enough for a diagnosis with all chambers and valves visible, but off-centre	
5	Good image	

subjects. However, their focus was on the parasternal window, not the apical window. Their system scans the region of interest (ROI) using a snake-like trajectory while maintaining the probe's orientation perpendicular to the skin surface. This approach, starting with a perpendicular orientation, is less optimal for scanning in the apical window, where the ideal orientation is not perpendicular to the skin. Moreover, assuming that the probe's initial position follows a 3D normal distribution centred around the ideal location, a spiralling outward from the initial starting point would likely be the most efficient method for locating the optimal position.

Another relevant study is that of Tang et al. [24] whose system is able to recognise the posture of a patient to locate an ROI to scan. However, their solution was tested only on a phantom model that lacks the complexity of simulating lung tissue or rib obstacles that significantly complicate ultrasound scanning. These challenges are accurately modelled using the Heartworks simulator, which is a better representative of real-world conditions.

Finally, Soemantoro et al. [25] proposed an innovative path-planning approach in which the probe is moved in multiple directions (up, down, left, right) to determine in which direction the quality of the image is improving. The

TABLE VI
SCAN PERFORMANCE WHEN RANDOMLY PLACING THE TRANSDUCER IN THE RED OR BLUE ZONE

	A4Ch view		A2Ch View	
	Blue Zone	Red Zone	Blue Zone	Red Zone
Average AI quality score	56%	46%	-	-
Average rating	4.35	3.56	3.64	2.24
Accuracy	78%	58%	56%	25%

movement distance is determined by an algorithm that employs a rolling window to store data about previously visited points. This system was also tested using the Heartworks simulator, allowing for direct comparisons of the starting positions and the relative accuracy of different parameters. In general, the proposed system performs better when starting further from the ideal position. However, when starting closer to the ideal position, the method developed by Soemantoro et al. [25] performs better. It is important to note that their search strategy has not yet been integrated into a robotic system, and real-world results, once integrated, may differ.

IV. DISCUSSION

The results suggest that the proposed navigation strategy is highly promising, with the system achieving an accuracy of approximately 80% when starting in a region close to the heart. They also demonstrate that a more precise initial positioning significantly enhances the accuracy of the final images. If, in future versions of the robotic ultrasound system, the initial starting position can be more precise than the current 6 cm by 6 cm square, further accuracy improvement is expected. One potential solution is the integration of computer vision techniques to assist with the initial probe placement by locating a point below the patient's left nipple, which could accelerate the scan process and improve consistency. Additionally, several factors can easily be optimised to boost accuracy, such as increasing the number of points in each spiral, analysing more images in each waypoint, and fine-tuning the hyperparameters of the AI models. The search algorithm can also be refined by adjusting the parameters defining the spiral shape and thresholds of grading scores.

The force control mechanism is another area for improvement. Currently, the robot is compliant only around the z-axis, a strategy that works due to the relatively flat torso of the simulator and the safety features of Universal Robots, which automatically halt the system if potentially injurious forces are detected at any point, including the end effector or joints. However, this approach may cause discomfort for patients, especially during prolonged scans. It can also lead to imperfect contact between the probe and the skin, reducing image quality, and may introduce vibrations that slightly displace the simulator or patient, further diminishing the system's overall performance. To enhance patient comfort, the force control system should be expanded to offer compliance along all axes, a potential solution of which is to use admittance control. However, admittance control requires more precise force sensors, which would increase the cost of the system.

Furthermore, the current system does not account for patient motion, a factor that becomes critical when scanning real patients. Unlike the simulator, patients are likely move during a scan and making motion compensation essential for maintaining imaging quality. Additionally, the spiral search approach may be less efficient when applied to larger areas or complex anatomical structures. Future works will focus on exploring more intelligent path-planning methods. Machine learning algorithms can be used to predict the heart region based on current and historical imaging data, thereby enhancing the efficiency and precision of the search process.

The system is currently considered semi-automatic as it needs manual placement of the transducer at the initial starting position. This limitation could potentially be tackled by adding a vision-based module that uses AI to automatically recognise key markers (e.g., the nipple or shoulder) and determine the starting position more accurately.

AI plays a critical role in the system's performance. In some cases, the AI assigned low quality scores to images that appeared nearly ideal, primarily because the heart was slightly off-centre, likely a result of such cases being underrepresented in the training data. Conversely, images partially obscured by ribs were occasionally graded more favourably, reflecting uncertainty in scenarios insufficiently captured during training. These observations suggest that a deep understanding of tissue structure could improve model reliability. Incorporating anatomical features, such as through segmentation masks or structural analysis, may enhance the system's ability to distinguish subtle variations in image quality and view. While the current AI models were designed for real-time inference, future iterations may explore whether these techniques can be integrated efficiently. To improve generalisability, training data were sourced from three different transducers; however, this may compromise performance for any one specific device. Additionally, the system was tested on a single body model representing a healthy, average-sized individual. Variations in body size and anatomy, including high body mass index, breast implants, or smaller rib cages (e.g., children) are likely to affect performance. Future work will assess the system's adaptability to improve on their performance.

The system relies solely on ultrasound data, which provides a significant advantage compared to other systems that require depth cameras [20], [31], [32], [33], structured light [21], or tomographic data [34], [35], as these add cost and complexity to the setup.

V. CONCLUSION

This study explores the development of Echo-Robot, the first robotic system designed for autonomous echocardiography from the apical window. Focusing on transthoracic echocardiography, the system uses AI to assess ultrasound image quality and determine the view type. It relies solely on a standard UR16 manipulator, eliminating the need for additional equipment like depth cameras or tomographic data. Tested on the Heartworks simulator, Echo-Robot autonomously adjusts

the position and orientation of the probe to achieve optimal A4Ch and A2Ch views, achieving an accuracy of approximately 80%. By automating ultrasound scanning, Echo-Robot has the potential to address the shortage of trained sonographers, improve access to cardiac care in underserved areas, and reduce physical strain on healthcare professionals. While promising, further improvements in system reliability are needed before clinical trials on human patients can begin.

ACKNOWLEDGMENT

The Heartworks simulator was provided by Intelligent Ultrasound.

REFERENCES

[1] *SoR Evidence to the NHS Pay Review Body*, Soc. Radio, London, U.K., Feb. 2024.

[2] MGI-Tech. "MGIUS-R3." Accessed: Mar. 17, 2025. [Online]. Available: https://en.mgitech.cn/products/instruments_info/11/

[3] AdEchoTech. "MELODY." Accessed: Mar. 17, 2025. [Online]. Available: https://www.adechotech.com/products/

[4] J. Wang et al., "Application of a robotic tele-echography system for COVID19 pneumonia," *J. Ultrasound Med.*, vol. 40, no. 2, pp. 385–390, Feb. 2021, doi: 10.1002/jum.15406.

[5] W. Jiang et al., "Application of a tele-ultrasound robot during COVID19 pandemic," *J. Ultrasound Med.*, vol. 42, no. 3, pp. 595–601, Mar. 2023, doi: 10.1002/jum.16041.

[6] M. Georgescu, A. Sacccomandi, B. Baudron, and P. L. Arbeille, "Remote sonography in routine clinical practice between two isolated medical centers and the university hospital using a robotic arm: A 1-year study," *Telemed. e-Health*, vol. 22, no. 4, pp. 276–281, Apr. 2016, doi: 10.1089/tmj.2015.0100.

[7] S. J. Adams et al., "Initial experience using a telerobotic ultrasound system for adult abdominal sonography," *Can. Assoc. Radiol. J.*, vol. 68, no. 3, pp. 308–314, Aug. 2017, doi: 10.1016/j.carj.2016.08.002.

[8] S. J. Adams et al., "A crossover comparison of standard and telerobotic approaches to prenatal sonography," *J. Ultrasound Med.*, vol. 37, no. 11, pp. 2603–2612, Nov. 2018, doi: 10.1002/jum.14619.

[9] M. Georgescu, A. Sacccomandi, B. Baudron, and P. L. Arbeille, "Remote sonography in routine clinical practice between two isolated medical centers and the university hospital using a robotic arm: A 1-year study," *Telemed. e-Health*, vol. 22, no. 4, pp. 276–281, Apr. 2016, doi: 10.1089/tmj.2015.0100..

[10] S. Avgousti et al., "Cardiac ultrasonography over 4G wireless networks using a tele-operated robot," *Healthc Technol. Lett.*, vol. 3, no. 3, pp. 212–217, Sep. 2016, doi: 10.1049/htl.2016.0043.

[11] P. Arbeille et al., "Teleoperated echograph and probe transducer for remote ultrasound investigation on isolated patients (study of 100 cases)," *Telemed. e-Health*, vol. 22, no. 7, pp. 599–607, Jul. 2016, doi: 10.1089/tmj.2015.0186.

[12] P. Arbeille et al., "Remote echography between a ground control center and the international space station using a tele-operated echograph with motorized probe," *Ultrasound Med. Biol.*, vol. 44, no. 11, pp. 2406–2412, Nov. 2018, doi: 10.1016/j.ultrasmedbio.2018.06.012.

[13] P. Arbeille, R. Provost, K. Zuj, D. Dimouro, and M. Georgescu, "Teles-operated echocardiography using a robotic arm and an Internet connection," *Ultrasound Med. Biol.*, vol. 40, no. 10, pp. 2521–2529, Oct. 2014, doi: 10.1016/j.ultrasmedbio.2014.05.015.

[14] T.-Y. Fang, H. K. Zhang, R. Finocchi, R. H. Taylor, and E. M. Boctor, "Force-assisted ultrasound imaging system through dual force sensing and admittance robot control," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 6, pp. 983–991, Jun. 2017, doi: 10.1007/s11548-017-1566-9.

[15] J. Carriere et al., "An admittance-controlled robotic assistant for semi-autonomous breast ultrasound scanning," in *Proc. Int. Symp. Med. Robot. (ISMR)*, Apr. 2019, pp. 1–7, doi: 10.1109/ISMR.2019.8710206.

[16] Z. Jiang et al., "Automatic normal positioning of robotic ultrasound probe based only on confidence map optimization and force measurement," *IEEE Robot. Autom. Lett.*, vol. 5, no. 2, pp. 1342–1349, Apr. 2020, doi: 10.1109/LRA.2020.2967682.

[17] R. Kim et al., "Robot-assisted semi-autonomous ultrasound imaging with tactile sensing and convolutional neural-networks," *IEEE Trans. Med. Robot. Bionics*, vol. 3, no. 1, pp. 96–105, Feb. 2021, doi: 10.1109/TMRB.2020.3047154.

[18] (GE HealthCare, Chalfont St. Giles, U.K.) *Abus Breast Imaging Ultrasound Products & Solutions*. Accessed: May 31, 2024. [Online]. Available: https://www.gehealthcare.co.uk/products/ultrasound/breast-ultrasound

[19] R. F. Brem et al., "Assessing improvement in detection of breast cancer with three-dimensional automated breast US in women with dense breast tissue: The somoinsight study," *Radiology*, vol. 274, no. 3, pp. 663–673, Mar. 2015, doi: 10.1148/radiol.14132832.

[20] F. Suligoj, C. M. Heunis, J. Sikorski, and S. Misra, "RobUSt-an autonomous robotic ultrasound system for medical imaging," *IEEE Access*, vol. 9, pp. 67456–67465, 2021, doi: 10.1109/ACCESS.2021.3077037.

[21] J. Tan et al., "A flexible and fully autonomous breast ultrasound scanning system," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 3, pp. 1920–1933, Jul. 2023, doi: 10.1109/TASE.2022.3189339.

[22] Z. Deng, X. Hou, C. Chen, X. Gu, Z.-G. Hou, and S. Wang, "A portable robot-assisted device with built-in intelligence for autonomous ultrasound acquisitions in follow-up diagnosis," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–10, May 2024, doi: 10.1109/TIM.2024.3400342.

[23] Y. Shida, S. Kumagai, R. Tsumura, and H. Iwata, "Automated image acquisition of parasternal long-axis view with robotic echocardiography," *IEEE Robot. Autom. Lett.*, vol. 8, no. 8, pp. 5228–5235, Aug. 2023, doi: 10.1109/LRA.2023.3292568.

[24] X. Tang et al., "Autonomous ultrasound scanning robotic system based on human posture recognition and image servo control: An application for cardiac imaging," *Front. Robot. AI*, vol. 11, May 2024, Art. no. 1383732, doi: 10.3389/frobt.2024.1383732.

[25] R. Soemantoro, A. Kardos, G. Tang, and Y. Zhao, "An AI-powered navigation framework to achieve an automated acquisition of cardiac ultrasound images," *Sci. Rep.*, vol. 13, no. 1, Sep. 2023, Art. no. 15008, doi: 10.1038/s41598-023-42263-2.

[26] (Univ. Robots, Odense, Denmark). *Real-Time Data Exchange (RTDE) Guide-22229*. Accessed: Aug. 22, 2024. [Online]. Available: https://www.universal-robots.com/articles/ur/interface-communication/real-time-data-exchange-rtde-guide/

[27] A. Degerli et al., "Early detection of myocardial infarction in low-quality echocardiography," *IEEE Access*, vol. 9, pp. 34442–34453, 2021, doi: 10.1109/ACCESS.2021.3059595.

[28] R. M. Lang et al., "Recommendations for cardiac chamber quantification by echocardiography in adults: An update from the American society of echocardiography and the European association of cardiovascular imaging," *Eur. Heart J. Cardiovasc. Imag.*, vol. 16, no. 3, pp. 233–271, Mar. 2015, doi: 10.1093/ehjci/jev014.

[29] D. R. Morgan, M. B. Sanville, S. Bathula, P. S. Demirel, M. S. Perkins, and M. G. E. Johnson, "Simulator-based training in FoCUS with skill-based metrics for feedback: An efficacy study," *POCUS J.*, vol. 4, no. 2, pp. 33–36, Nov. 2019, doi: 10.24908/pocus.v4i2.13845.

[30] P. Ferrada et al., "Limited transthoracic echocardiogram: So easy any trauma attending can do it," *J. Trauma, Injury, Infect. Crit. Care*, vol. 71, no. 5, pp. 1327–1332, Nov. 2011, doi: 10.1097/TA.0b013e3182318574.

[31] G. Ma, S. R. Oca, Y. Zhu, P. J. Codd, and D. M. Buckland, "A novel robotic system for ultrasound-guided peripheral vascular localization," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2021, pp. 12321–12327, doi: 10.1109/ICRA48506.2021.9561924.

[32] Z. Wang et al., "Full-coverage path planning and stable interaction control for automated robotic breast ultrasound scanning," *IEEE Trans. Ind. Electron.*, vol. 70, no. 7, pp. 7051–7061, Jul. 2023, doi: 10.1109/TIE.2022.3204967.

[33] X. Ma, Z. Zhang, and H. K. Zhang, "Autonomous scanning target localization for robotic lung ultrasound imaging," in *Proc. IEEE Int. Conf. Intell. Robots Syst.Proc.* , 2021, pp. 9467–9474, doi: 10.1109/IROS51168.2021.9635902.

[34] R. Göbl, S. Virga, J. Rackerseder, B. Frisch, N. Navab, and C. Hennersperger, "Acoustic window planning for ultrasound acquisition," *Int. J. CARS*, vol. 12, no. 6, pp. 993–1001, 2017, doi: 10.1007/s11548-017-1551-3.

[35] Y. Bi, C. Qian, Z. Zhang, N. Navab, and Z. Jiang, "Autonomous path planning for intercostal robotic ultrasound imaging using reinforcement learning," 2024, *arXiv:2404.09927*.