



Review

Reinforcement Learning for Ultrasound Image Analysis: A Scoping Review

Maha Ezzelarab ¹, Midhila Madhusoodanan ¹, Shrimanti Ghosh ¹, Geetika Vadali ², Jacob L. Jaremko ¹ and Abhilash Hareendranathan ^{1,*}

¹ Department of Radiology and Diagnostic Imaging, University of Alberta, Edmonton, AB T6G 2T4, Canada; ezzelara@ualberta.ca (M.E.); midhila@ualberta.ca (M.M.); shrimant@ualberta.ca (S.G.); jjaremko@ualberta.ca (J.L.J.)

² College of Computing, Georgia Institute of Technology, Atlanta, GA 30332, USA; geetika.vadali4@gmail.com

* Correspondence: hareendr@ualberta.ca

Abstract

Machine learning using supervised approaches has been widely applied to ultrasound image analysis. In contrast, reinforcement learning (RL), which is well suited for sequential decision-making, is underexplored. Ultrasound workflows involve sequential subtasks like image acquisition, quality assessment, summarization, and interpretation that can be integrated into RL frameworks. This scoping review examined RL applications in ultrasound. A comprehensive search was conducted in Scopus, PubMed, Embase, and MEDLINE for studies published between 2015 and 2026. Eligible studies used RL with clinical ultrasound data, and the data were summarized by application area, methods, and anatomical targets. The review also provides a brief overview of RL concepts as foundational knowledge for understanding various RL formulations used. From 326 records retrieved, 39 studies were included. Most studies used model-free deep RL, with value-based methods being the most common, particularly Deep Q-Network (DQN) and its variants, on retrospective data across diverse anatomical targets, with breast, fetal, and uterine ultrasound being the most frequently represented categories. RL was used to automate tasks including navigation, plane localization, landmark detection, and video summarization. RL in ultrasound imaging is an emerging field of research and has advantages for sequential workflow optimization tasks. Most approaches are at an early stage and have been tested on small datasets, lack consistent evaluation protocols, and report limited clinical translation.

Keywords: reinforcement learning; ultrasound image analysis; sequential decision-making; healthcare applications; medical imaging

Academic Editor: Nikolaos Kourkoumelis

Received: 22 April 2026

Revised: 9 July 2026

Accepted: 21 July 2026

Published: 3 August 2026

Copyright: © 2026 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Medical ultrasound (US) imaging is widely used in clinical practice for diagnostic and triage applications due to its low cost, portability, and ability to provide real-time visualization of anatomical structures [1,2]. It is commonly used both within hospital environments and in emergency care settings. The effectiveness of US imaging is often highly dependent on operator expertise, which affects image acquisition and interpretation [2]. Automated and computer-assisted approaches for US image analysis aim to reduce this dependence on the operator. Over the past decade, machine learning (ML)

techniques have been widely used for various US imaging tasks like segmentation [3], classification [4], object detection [5], and image enhancement [1,2]. Most of these approaches are based on supervised deep learning (DL) and typically learn hierarchical feature representations from labeled data.

Clinical US examinations involve a sequence of interdependent subtasks, like image acquisition, quality assessment, anatomical localization, video summarization, and diagnostic interpretation. For example, sonographers often follow complex and adaptive scanning strategies to identify diagnostically relevant images, with decisions influenced by anatomical visibility, probe positioning, and prior observations [6]. However, conventional supervised learning models typically treat tasks as independent prediction problems and do not explicitly model the sequential dependencies inherent in clinical US workflows.

Rationale for Reinforcement Learning in Ultrasound: Reinforcement learning (RL) provides a natural framework for addressing such sequential decision-making problems. In RL, an agent interacts with an environment and learns a policy to maximize cumulative reward over time [7]. This formulation enables the modeling of long-term dependencies and adaptive decision-making based on prior actions. This makes RL particularly well-suited for tasks like navigation, iterative optimization, and temporal reasoning [7,8]. Deep reinforcement learning (DRL) extends this paradigm by integrating deep neural networks, allowing agents to learn directly from high-dimensional inputs such as medical images and video sequences [7]. By combining representation learning with sequential decision-making, DRL has the potential to address complex ultrasound tasks that go beyond static prediction, including probe navigation, acquisition optimization, and video-based analysis.

Despite its potential, RL in ultrasound imaging remains underexplored and fragmented across diverse tasks, including robotic navigation, standard-plane localization, segmentation, classification, quality assessment, and video summarization. Existing reviews have mainly focused on broader areas, such as artificial intelligence in ultrasound, robotic ultrasound imaging, disease-specific deep learning applications, or reinforcement learning in medical imaging more generally. However, a focused synthesis of RL applications in clinical ultrasound image and video analysis remains lacking. In particular, there is a need to systematically examine how RL has been formulated in ultrasound-specific settings, including state-action-reward design, algorithm selection, anatomical targets, dataset availability, code availability, evaluation strategies, and barriers to clinical translation. Therefore, this scoping review is timely and necessary to map the current evidence, clarify methodological trends, and identify gaps for future ultrasound RL research.

Objectives: The objective of this scoping review is to systematically map and synthesize the existing literature on the application of RL in US image and video analysis. It provides a structured overview of how RL methods have been formulated and applied across different ultrasound imaging tasks and anatomical targets, while also examining dataset characteristics and evaluation strategies. In addition, the review highlights key challenges, emerging trends, and gaps in the current literature, and offers practical guidance to support researchers and clinicians in understanding, applying, and advancing RL methods in ultrasound imaging.

1.1. Background and Related Work

Deep Learning Applications in US Imaging: DL learns useful patterns directly from ultrasound data without relying on handcrafted features. It has been used for image interpretation, measurement, and decision support [1,9,10]. DL mostly addresses computer vision tasks such as classification, segmentation, and detection/localization. For example, DL based approaches have been used to classify thyroid nodules and ovarian cancers from ultrasound images [11,12], segment cardiac structures [13], and detect and localize clinically optimal imaging planes [14]. DL has also been extended to more advanced

applications that support the overall imaging pipeline like image quality assessment [15], US denoising [16,17], video analysis and functional assessment [18], and multimodal approaches to improve diagnostic accuracy [19,20]. However, these supervised approaches treat each task independently and are not designed for sequential decision-making.

1.2. Overview of Reinforcement Learning

RL is a learning paradigm based on interaction, where a system improves its behavior through trial and error using feedback rather than labeled data [7]. This process is analogous to how humans learn from experience.

1.2.1. Reinforcement Learning Framework

The RL framework consists of three main components: an agent, an environment, and a reward signal. The agent represents the decision-making system, the environment defines the setting in which decisions are made, and the reward provides feedback on the effectiveness of the agent's actions, with positive rewards encouraging desirable behavior and penalties discouraging unfavorable actions. As shown in Figure 1, the agent observes a state, selects an action according to its policy, and receives feedback from the environment in the form of a new state and reward [7]. Through repeated interactions, the agent learns a policy that maximizes cumulative reward over time. Unlike conventional supervised one-step prediction, RL can optimize delayed and cumulative rewards across a sequence of decisions or subtasks.

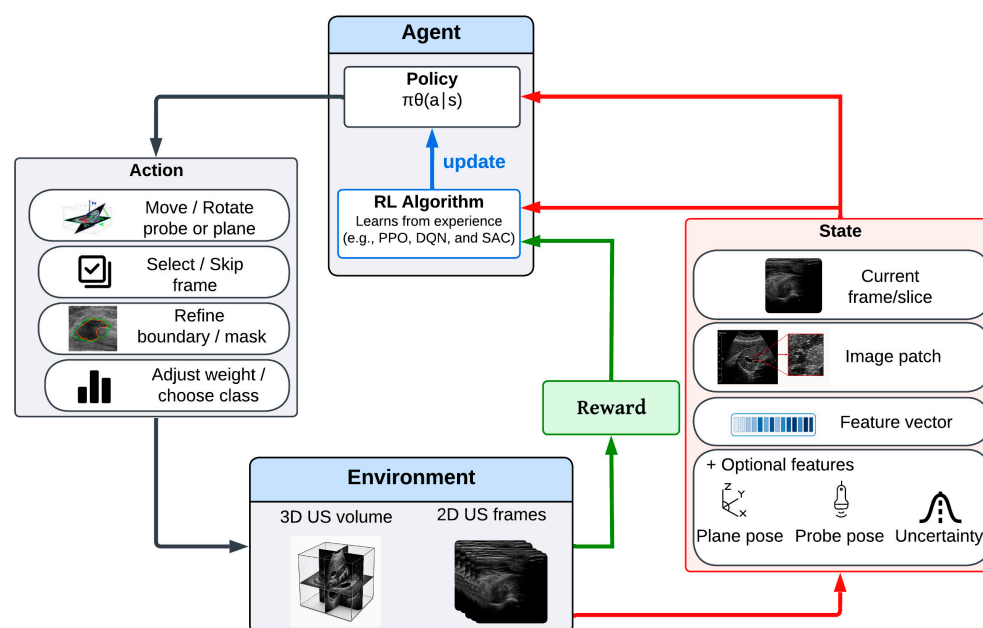


Figure 1. Ultrasound-specific reinforcement learning framework. The agent observes ultrasound-derived states, such as the current frame or slice, image patch, feature vector, or contextual information including plane pose, probe pose, and uncertainty. Based on the learned policy, the agent selects task-specific actions, such as moving or rotating the probe or imaging plane, selecting or skipping frames, refining boundaries or masks, or adjusting weights for classification. The ultrasound environment then returns the next state and a task-dependent reward, which is used by the RL algorithm to update the policy.

RL Framework for Ultrasound: In the context of ultrasound imaging, the agent represents the learning-based decision-making system, often implemented as a neural

network policy, that receives ultrasound-derived states and selects task-specific actions. The environment may correspond to a 3D ultrasound volume, a 2D frame sequence, or a localized image region. The state represents the current observation available to the agent, such as a frame, an imaging plane, an image patch, a learned feature representation, or temporal context from previous steps. The action depends on the task and may involve navigating the probe or imaging plane, selecting informative frames, refining a segmentation boundary, or supporting a classification decision. The reward provides task-specific feedback on the selected action. For example, a keyframe that improves diagnostic accuracy may receive a positive reward, whereas a non-informative or misleading frame may receive a penalty. Through this feedback, the agent learns to select actions that improve ultrasound analysis over time.

1.2.2. Types of Reinforcement Learning Approaches

RL methods are commonly categorized based on whether the agent uses a model of the environment (Figure 2) or not.

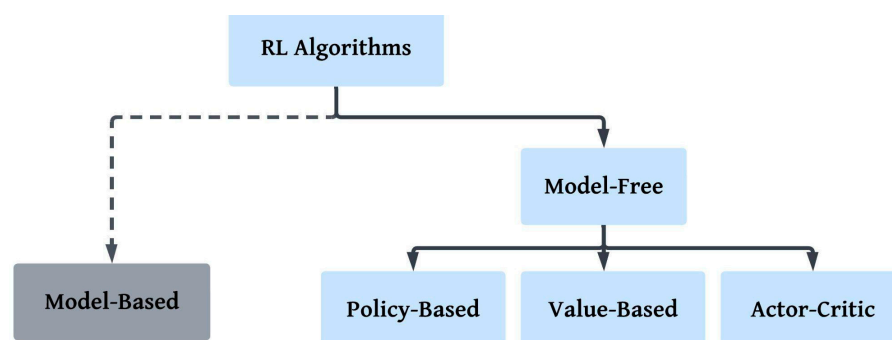


Figure 2. Categorization of RL algorithms. RL algorithms are broadly divided into model-based and model-free approaches. In model-based RL, the agent relies on a provided or learned model of the environment for planning, whereas in model-free RL, the agent learns directly from interactions with the environment. Because model-based RL was rarely represented in the reviewed ultrasound studies, the present review focuses mainly on the three major model-free categories: policy-based, value-based, and actor–critic methods. The dashed arrow indicates that model-based RL is included for completeness, whereas the solid arrows indicate the model-free categories emphasized in this review.

- (a) **Model-Based Reinforcement Learning:** In model-based RL, the agent uses an explicit representation of the environment to predict the outcomes of actions and support planning. Although this can improve decision-making, constructing or learning an accurate environment model is often difficult in real-world medical imaging settings. In the reviewed ultrasound literature, model-based RL was rarely used. Therefore, it is included here only for completeness, and the discussion in this review focuses primarily on model-free methods.
- (b) **Model-Free Reinforcement Learning:** In Model-free RL, instead of explicitly modeling the environment, the agent learns directly from experience by interacting with the data and receiving feedback in the form of rewards. This approach is flexible and has the ability to handle complex visual inputs, as seen in medical imaging. Model-free RL methods can be broadly divided into three categories:
 - i. **Policy-based methods** learn a direct mapping from observations (e.g., ultrasound images or frames) to actions by optimizing a parameterized policy. Rather than explicitly evaluating all possible actions, the model learns to select

actions that maximize expected reward. This makes them well suited for problems with continuous or uncertain decision spaces, such as adjusting probe position or refining anatomical views. REINFORCE is a popular policy-based algorithm that updates the policy by increasing the likelihood of actions that lead to higher rewards.

- ii. **Value-based methods** learn to estimate the expected return (value) of actions in a given state. The agent then selects the action with the highest estimated value. These methods are particularly effective when the action space is discrete, such as selecting frames, adjusting planes in predefined directions, or navigating step-by-step within an image. Deep Q-Network (DQN) is a popular value-based approach that combines Q-learning with deep neural networks to handle high-dimensional visual inputs.
- iii. **Actor-critic methods** combine both approaches by learning a policy (actor) while simultaneously estimating the value of actions (critic). The actor determines which action to take, while the critic evaluates the action and provides feedback to improve the policy. This joint framework often leads to more stable and efficient training and is well suited for complex tasks involving both discrete and continuous decisions, such as coordinated navigation and localization in ultrasound.

Deep RL: This integration of DL with RL, referred to as DRL, allows the agent to make decisions based on rich visual information. In DRL frameworks, neural networks are used to approximate key components of the RL process, such as the policy (which determines the action to take) or the value function (which estimates the quality of actions). For example, convolutional neural networks are often used to extract spatial features from ultrasound images, while recurrent networks or temporal models may capture motion and context across video frames. This enables RL agents to operate effectively in complex environments where the input data is high-dimensional and continuously changing.

In the context of ultrasound imaging, reinforcement learning has been primarily explored for tasks that involve sequential decision-making across the imaging workflow. These include navigation and probe guidance, anatomical plane localization, landmark detection, and video summarization, where the agent iteratively interacts with image data to optimize task-specific objectives. More recently, RL has also been investigated in image understanding and optimization tasks such as segmentation, classification, and quality assessment. These application areas are examined in detail in Section 3.

2. Materials and Methods

This scoping review was conducted in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guidelines. The review protocol was prospectively registered on the Open Science Framework (OSF) and is publicly available at: <https://osf.io/g2ayu> (accessed on 22 April 2026).

2.1. Eligibility Criteria

Studies were selected based on predefined eligibility criteria, as summarized in Table 1. Studies were included if reinforcement learning methods were developed using clinical human ultrasound data. Studies relying exclusively on phantom or simulated data were excluded. However, studies that used clinical data for model development and

incorporated phantom-based experiments only for evaluation or validation were considered eligible.

Table 1. Eligibility criteria for study inclusion and exclusion in this scoping review.

Category	Inclusion Criteria	Exclusion Criteria
Study type	Peer-reviewed articles, conference papers	Review articles, editorials, or commentaries
Methodology	RL or DRL applied	No RL
Data	Clinical ultrasound data	Phantom/simulated
Population	Human studies	Non-human
Language	English	Non-English
Timeframe	2015–2026	Outside range

2.2. Information Sources

A comprehensive literature search was conducted across Scopus, PubMed, Embase and MEDLINE (via Ovid) electronic databases. The search covered studies published between 2015 and 2026, with the most recent search performed in March 2026. All retrieved records were imported into Covidence (Veritas Health Innovation, Melbourne, Australia), an online systematic review management platform, for duplicate removal, screening, and study selection.

2.3. Search Strategy

The search strategy combined keywords related to RL and US imaging. The full search queries used for each database are presented in Table 2.

Table 2. Search queries used across electronic databases for identifying studies on RL in US imaging.

Database	Search Query
PubMed	("reinforcement learning" OR "deep reinforcement learning" OR "q learning" OR "dqn" OR "actor critic" OR "policy gradient") AND ("ultrasound" OR "ultrasonograph" OR "sonograph" OR "point of care ultrasound" OR "pocus")
Scopus	TITLE-ABS-KEY ("reinforcement learning" OR "deep reinforcement learning" OR "q learning" OR "dqn" OR "actor critic" OR "policy gradient") AND TITLE-ABS-KEY ("ultrasound" OR "ultrasonograph" OR "sonograph" OR "pocus" OR "point-of-care ultrasound")
Embase	("reinforcement learning" OR "deep reinforcement learning" OR "q learning" OR "dqn" OR "actor critic" OR "policy gradient") AND ("ultrasound" OR "ultrasonograph" OR "sonograph" OR "point of care ultrasound")
MEDLINE (via Ovid)	("reinforcement learning" OR "deep reinforcement learning" OR "q learning" OR "dqn" OR "actor critic" OR "policy gradient") AND ("ultrasound" OR "ultrasonograph" OR "sonograph" OR "point of care ultrasound")

2.4. Data Charting Process

Title and abstract screening were performed by two reviewers (ME and SG) in Covidence. Full-text screening was then conducted by all authors, and disagreements or uncertainties were resolved through discussion until consensus was reached.

Data from the included studies were charted using a standardized data charting form developed specifically for this review. The form was designed by ME and SG to capture key methodological characteristics and application-specific details related to the use of reinforcement learning in ultrasound imaging. Prior to full data charting, the form was tested on a subset of studies and refined to ensure clarity, consistency, and completeness. Data extraction was conducted by all authors using the standardized form, and extracted data were reviewed by ME and SG for completeness and consistency prior to synthesis. A formal inter-rater agreement statistic was not calculated.

2.5. Data Items

For each included study, key variables were systematically charted to capture both methodological characteristics and application-specific details relevant to the use of reinforcement learning in ultrasound imaging. The selected data items were chosen to enable comprehensive comparison across studies and to support the identification of trends, strengths, and gaps in the literature. Specifically, the extracted variables included the publication year, US task category, RL algorithm used, anatomical region of interest, and dataset characteristics (e.g., sample size, data type, and clinical setting), as well as the availability of datasets and code, evaluation metrics, and key methodological strengths and limitations. To strengthen the methodological synthesis, RL formulation details were also summarized where reported or clearly described, including state representation, action space, reward design, validation maturity, and clinical readiness.

2.6. Critical Appraisal

As this study was conducted as a scoping review, a formal critical appraisal of the methodological quality of individual studies was not performed. The primary aim of this review was to map the extent, range, and nature of existing research, and to identify methodological trends and knowledge gaps.

2.7. Synthesis of Results

The charted data were synthesized using a narrative analysis approach. Studies were grouped according to ultrasound task categories and the RL methodologies employed. The synthesis focused on identifying common RL formulations and algorithmic approaches, dataset characteristics and availability, evaluation metrics and validation strategies, as well as methodological strengths, limitations, and emerging trends within the field. This approach enabled a structured comparison of studies across different application domains and supported the identification of patterns in the use of RL within ultrasound imaging workflows.

In addition to the study-level summary, a compact task-level methodological taxonomy was developed to improve cross-study comparability. This taxonomy summarized recurring RL formulations across ultrasound task categories, including state representation, action space, reward design, evidence maturity, and clinical readiness. Because the included studies were heterogeneous in anatomy, dataset size, validation design, and performance metrics, the taxonomy was intended as a qualitative synthesis rather than a formal quality score or quantitative ranking of methods.

This structured approach enabled a comprehensive mapping of the existing literature on reinforcement learning in ultrasound imaging. By applying predefined eligibility criteria, systematic screening procedures, and a standardized data charting process, this

review ensured consistency and transparency in study selection and analysis. The extracted data provide a broad overview of application domains, methodological approaches, and anatomical targets, forming the basis for identifying key trends, common design patterns, and existing gaps in the field.

3. Results

A total of 326 records were identified and imported into Covidence for screening. Following import, 14 duplicates were identified manually, and an additional 99 duplicates were removed automatically by the platform, resulting in 213 unique records for screening. Title and abstract screening were performed on these 213 records, after which 69 studies were deemed potentially eligible and selected for full-text review.

Following full-text assessment, 30 studies were excluded based on predefined eligibility criteria. The primary reasons for exclusion included the use of only phantom or simulated ultrasound data, absence of reinforcement learning methodologies, use of non-clinical or non-human datasets, non-English publications, and inappropriate study designs. A total of 39 studies met all inclusion criteria and were included in the final synthesis. The study selection process is summarized in the PRISMA flow diagram Figure 3.

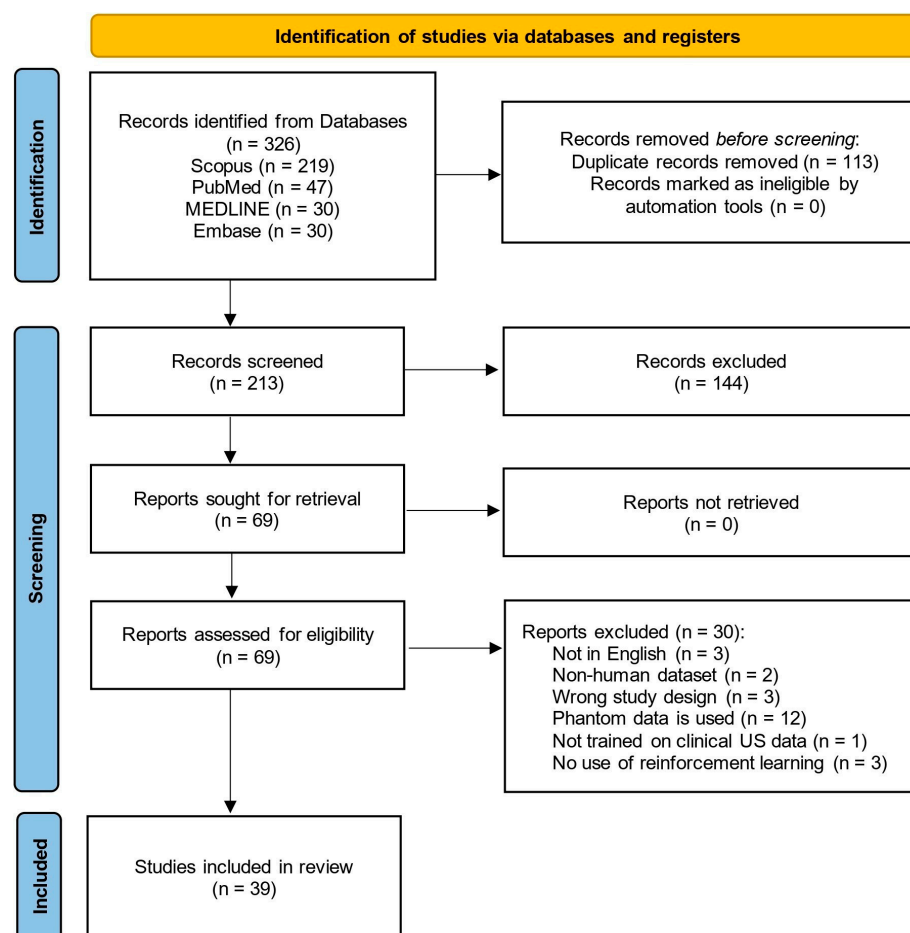


Figure 3. PRISMA flow diagram of study selection.

The characteristics of the included studies are summarized in Table 3. The extracted information includes the study authors, RL algorithms used, US application domain, dataset characteristics, and dataset availability. Table 3 summarizes the characteristics of the 39 included studies across US application domains, reinforcement learning (RL) methods, and dataset properties. The studies span the period from 2019 to 2026, with a noticeable

increase in publications in recent years, reflecting growing interest in RL for US imaging. Although the search window covered studies published from 2015 to 2026, the earliest study meeting the eligibility criteria was published in 2019. The included studies cover a range of task categories, including navigation guidance, plane localization, video summarization and keyframe extraction, segmentation, classification, landmark detection, quality assessment, and training optimization. Among these, navigation and plane localization represent a substantial portion of the literature, while segmentation and classification are less frequently explored.

Across tasks, a variety of RL algorithms were employed. Value-based methods, particularly DQN and its variants, including DDQN and dueling DQN, were the most commonly used, accounting for 26/39 studies (~67%). Policy-based methods, such as REINFORCE, were used in 6/39 studies (~15%), while actor–critic approaches, including PPO and TD3, were reported in 7/39 studies (~18%).

The applications spanned diverse anatomical regions. Because some studies addressed more than one anatomical target, anatomical counts were treated as non-mutually exclusive. Breast and fetal ultrasound were the most represented categories, each appearing in 7/39 studies (~18%), followed by uterine imaging in 6/39 studies (~15%). Spine applications were reported in 4/39 studies (~10%), while cardiac and prostate applications each appeared in 3/39 studies (~8%). Carotid, musculoskeletal, and thyroid applications were represented in 2/39 studies each (~5%). Other anatomical targets, including bladder, brain, kidney, lung, ovarian, and vascular ultrasound, were each explored in only 1/39 study (~3%). Therefore, the total number of anatomical assignments exceeded the number of included studies.

Data availability was limited across the reviewed studies. Most studies used private datasets (28/39 studies, ~72%), while 5/39 studies (~13%) reported that data were available upon request. Fully public datasets were used in only 2/39 studies (~5%), and another 2/39 studies (~5%) used partially available data, such as a publicly available test set. In addition, 2/39 studies (~5%) used mixed data sources. Source code availability was also limited, with code reported as available in only 7/39 studies (~18%), while 32/39 studies (~82%) did not report publicly available code. Dataset sizes varied widely, ranging from small cohorts of fewer than 10 subjects in some navigation studies to larger cohorts of more than 1000 subjects in applications such as plane localization and training optimization.

Table 3. Study-level characteristics of reinforcement learning applications in ultrasound image analysis.

Task	Study (Year)	RL Method	ROI	Data (Avail.)	Size (Subjects)	Code (Avail.)
Navigation Guidance	H. Hase et al., 2020 [21]	DQN	Spine	Request	34	✓
	Y. Chen et al., 2021 [22]	TD3	Kidney	Private	5	✗
	K. Li et al., 2022 [23]	DQN	Spine	Private	17	✗
	C. Shen et al., 2023 [24]	DQN	Heart	Private	13	✗
	G. Ning et al., 2023 [25]	PPO	Vascular	Private	20	✗
	K. Su et al., 2024 [26]	DQN	Thyroid	Request	19	✓
	Y. Shida et al., 2025 [27]	DQN	Heart	Private	24	✗

Plane Localization	K. Su et al., 2025 [28]	PPO	Carotid	Private	10	✗
	X. Wang et al., 2025 [29]	PPO	Spine	Private	5	✗
	H.-L. Hsu et al., 2026 [30]	DQN	Bladder	Private	14	✗
	H. Dou et al., 2019 [31]	DDQN	Fetal brain	Private	-	✗
	X. Yang et al., 2021 [32]	Dueling DQN	Fetal brain/Fetal abdomen/Uterus	Private	-	✓
	X. Yang et al., 2021 [33]	DDQN	Uterus/Fetal brain	Private	476	✗
	Y. Zou et al., 2022 [34]	Dueling DQN	Uterus/Fetal brain	Private	-	✗
	J. Huang et al., 2023 [35]	DDQN	Fetal face	Private	1250	✗
	B. Jiang et al., 2024 [36]	DDQN	Fetal head	Private	6	✗
	T. Liu et al., 2022 [37]	REINFORCE	Fetal US	Private	50	✗
Summarization and Keyframe Extraction	R. Huang et al., 2022 [38]	REINFORCE	Breast	Private	653	✗
	R. P. Mathews et al., 2023 [39]	REINFORCE	Lung US	Private	40	✗
	S. Ghosh et al., 2025 [40]	DQN	Shoulder rotator cuff tendon	Private	114	✗
	S. Ghosh et al., 2025 [41]	DQN	Wrist	Private	100	✗
	Y. Huang et al., 2026 [42]	REINFORCE	Uterus	Private	-	✓
Segmentation	N. Karunayake et al., 2023 [43]	PPO	Breast	Partial (test set available)	-	✗
	W. Yi et al., 2024 [44]	PPO	Prostate	Private	249	✗
	Y. Huang et al., 2025 [45]	DDQN	Breast	Request	BUS: 1892, ABUS: 605	✗
Classification	C. Narmatha et al., 2023 [46]	HHO-DQN	Ovarian	Public	-	✗
	X. Xie et al., 2024 [47]	DQN	Thyroid	Private	203	✗
Landmark Detection	G. Leroy et al., 2020 [48]	DQN	Brain	Partial (test set available)	72	✓
	S. Lian et al., 2021 [49]	DQN	Carotid	Request	95	✗

Quality Assessment	Q. Y. Ran et al., 2023 [50]	DQN	Spine	Request	12	✗
	S. U. Saeed et al., 2021 [51]	DQN	Prostate	Private	-	✗
	S. U. Saeed et al., 2021 [52]	PPO	Prostate	Private	249	✗
	P. Jarosik et al., 2021 [53]	DQN	Breast	Private	-	✗
	S. Liu et al., 2023 [54]	DQN	Fetal brain	Private	-	✗
	A. Saadat et al., 2026 [55]	DDQN	Heart	Private	-	✓
Training Optimization	R. Huang et al., 2021 [56]	REINFORCE	Breast	Mixed	835	✗
	S. Chen et al., 2026 [57]	REINFORCE	Breast	Mixed	2811	✓
	Y. Liu et al., 2025 [58]	DQN	Breast	Private	357	✗
	P. P. Phatale et al., 2025 [59]	DQN	Fetal US	Public	1792	✗

Dataset size is reported as the number of subjects, as described in the original studies. Image-, frame-, video-, and volume-level counts were omitted to improve consistency across studies.

Because detailed study-level characteristics and individual RL algorithms are already presented in Table 3, Table 4 provides a compact task-level taxonomy focused on how RL is formulated across ultrasound applications. The table summarizes the main state, action, and reward components, along with qualitative evidence maturity and clinical readiness.

Table 4. Compact methodological taxonomy of reinforcement learning formulations across ultrasound task categories.

Task/Studies	State	Action	Reward	Evidence Maturity	Clinical Readiness
Navigation guidance [21–30]	Current ultrasound image, probe position, or scanning context	Move the probe, adjust the scanning path, or stop when the target view is reached	Better target-view acquisition, image similarity, scanning success, or measurement agreement	Early to intermediate	Limited clinical workflow validation
Plane localization [31–36]	Current 2D plane sampled from a 3D ultrasound volume	Shift, rotate, or stop the virtual plane search	Smaller distance or orientation error from the target standard plane	Intermediate	Prospective workflow testing needed
Video summarization and keyframe extraction [37–42]	Frame-level, slice-level, or video-level features	Select, skip, or rank frames/slices	More representative, diverse, or diagnostically useful summaries	Early to intermediate	Heterogeneous evaluation metrics

Segmentation [43–45]	Image patch, edge map, contour position, or lesion-region context	Move along a boundary, trace a contour, or remove candidate regions	Better boundary agreement or lesion localization	Early to intermediate	Limited external validation
Classification [46,47]	Image features or diagnostic representation	Choose or refine the diagnostic class	Correct classification or improved diagnostic performance	Early feasibility	Limited clinical evidence
Landmark detection [48–50]	Current landmark estimate and surrounding image context	Move the estimated landmark position	Lower landmark localization error	Intermediate	Clinical robustness remains underexplored
Quality assessment [51–55]	Image/video quality features or downstream-task information	Enhance, select, reject, or choose the most useful view/image	Better image quality, task usefulness, or acquisition efficiency	Early to intermediate	Quality definitions vary
Training optimization [56–59]	Training feedback, class distribution, model output, or sample informativeness	Select samples, adjust weights, or balance training data	Better model performance, robustness, class balance, or annotation efficiency	Intermediate	Indirect clinical relevance

Evidence maturity and clinical readiness were summarized qualitatively from reported study characteristics and are not intended as formal evidence grading.

Table 4 shows that RL is most often used to support sequential decisions in ultrasound workflows, including navigation, search, selection, localization, quality assessment, and training optimization. Because state, action, and reward definitions vary by clinical task and data type, the table is intended to summarize recurring formulation patterns rather than rank individual methods or compare performance directly across tasks.

The included studies were grouped according to the ultrasound tasks addressed. The main categories identified were autonomous navigation, plane localization, video summarization and keyframe extraction, segmentation, classification, landmark detection, quality assessment, and training optimization. Within each category, studies were synthesized based on their RL formulations, algorithms, dataset characteristics, and methodological contributions. A visual summary of the distribution of studies by application category, anatomical target, and RL algorithm category is provided in Figure 4.

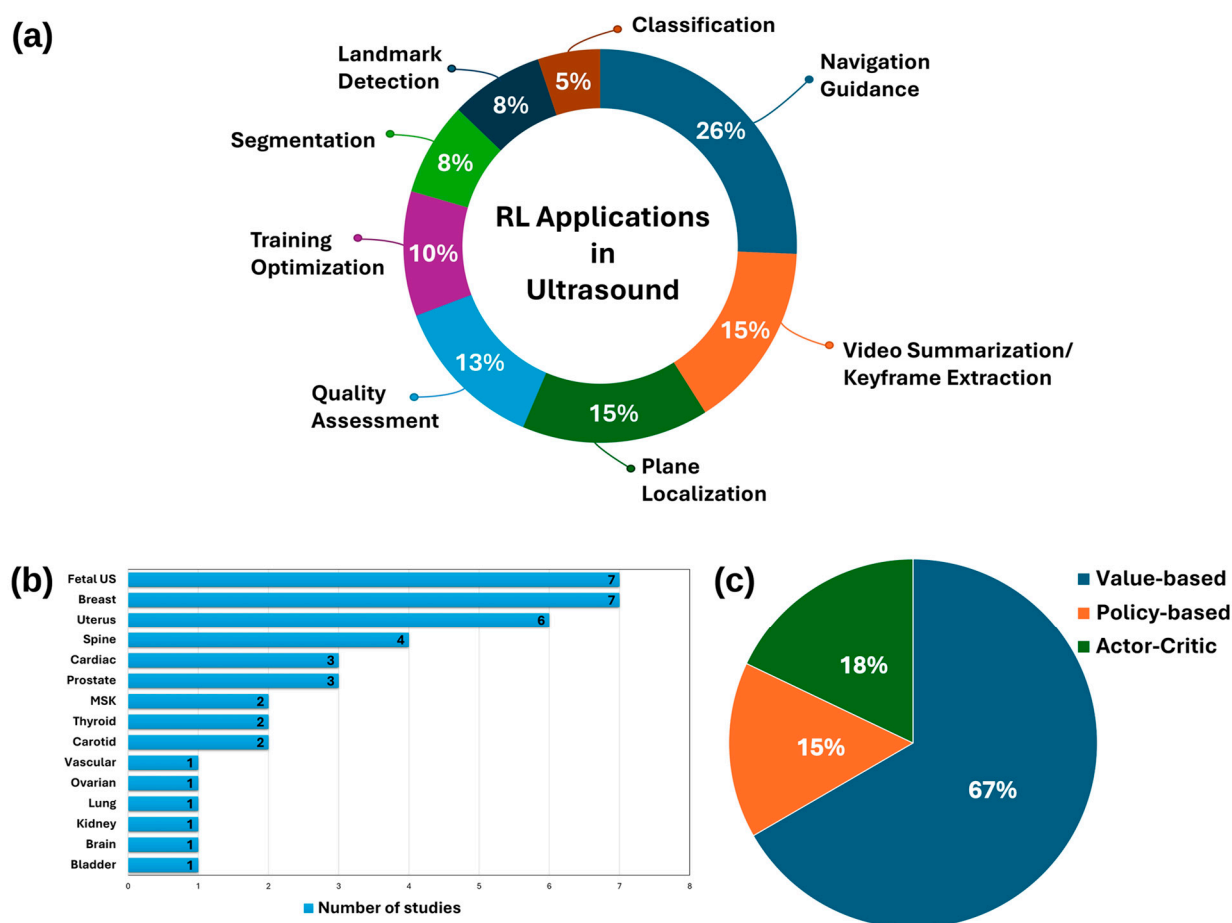


Figure 4. Visual synthesis of included reinforcement learning studies in ultrasound imaging. (a) Distribution of included studies by ultrasound application category. (b) Anatomical targets represented across the reviewed studies. Anatomical counts were treated as non-mutually exclusive because some studies addressed more than one anatomical target. (c) Distribution of reinforcement learning algorithm categories, showing the predominance of value-based methods, particularly DQN and its variants, compared with policy-based and actor–critic approaches.

3.1. Autonomous Navigation Using RL

Ultrasound probe navigation is a key aspect of image acquisition and often requires a high level of operator skill. RL-based automated navigation using robotic systems has been proposed for spinal [21,23,29], cardiac [24,27], vascular [25], thyroid [26], carotid [28] and abdominal [22,30] US. These approaches aim to reduce dependence on operators performing the examination and to lower the risk of repetitive stress injuries for sonographers. Early approaches to autonomous probe navigation focused on incorporating image and force feedback into the kinematics of robotic control systems. However, these models were based on simplified assumptions about the scanning environment and the properties of anatomical tissues, which do not always hold in practical scenarios. RL offers an elegant framework for incorporating environmental uncertainty into the robotic system. In this formulation, the current probe position is considered the state, and a finite set of actions is used to guide positioning. In most cases, the reward is defined in terms of similarity to ultrasound images acquired by an expert sonographer. The papers reviewed in this study primarily used either actor–critic methods or DQN as the underlying reinforcement learning algorithms.

Spine Imaging: H. Hase et al. developed an automatic spine image acquisition approach using DQN [21]. This approach utilized memory buffers and a binary classifier to terminate the scanning task. It was prospectively evaluated based on policy correctness and reachability on 34 subjects. Similarly, K. Li et al. developed a DQN-based approach to guide probe navigation toward different standard views of the spine [23]. This model was trained on 14 subjects and tested on 3 subjects, using SSIM with manually acquired standard images as the evaluation metric. In more recent work, X. Wang et al. developed a dual-agent-based robotic ultrasound path planning approach for spine imaging based on an actor–critic model [29]. Similar to Li et al., this approach was also validated using SSIM on 5 subjects.

Cardiac: C. Shen et al. developed a DQN-based approach for robotic cardiac ultrasound scanning that autonomously acquires high-quality 2D standard planes [24]. The DQN was trained on data from 13 subjects and tested on a cardiac phantom to evaluate the SSIM between automatically and manually acquired images. Similarly, Y. Shida et al. developed a DQN-based approach for autonomous cardiac imaging based on an optimized search behavior generation algorithm [27]. This approach was evaluated on 24 subjects in terms of the probability of reaching the optimal solution and inspection time.

Other Use Cases: K. Su et al. proposed a DQN-based approach for fully autonomous robotic ultrasound for thyroid scanning [26]. This study was conducted on 19 subjects, where TI-RADS scores assigned by experts were compared with those generated by the autonomous scanning system. K. Su et al. extended this work for autonomous carotid scanning using PPO, DQN, and Rainbow DQN algorithms. The study included 10 subjects, with performance evaluated on a smaller test subset of 2 subjects, using SSIM with expert-acquired images as the main evaluation metric [28]. Ning et al. proposed a PPO-based decoupled control strategy for autonomous vascular ultrasound imaging [25]. Y. Chen et al. used an actor–critic model to train a robotic system to acquire kidney images [22]. This approach was evaluated on 5 subjects in terms of measurement error with respect to manually acquired images. Hsu et al. developed an RL-based active-guidance approach for ultrasound bladder scanning [30].

3.2. Plane Localization Using RL

Similar to probe navigation, identifying the standard plane is also a crucial task in ultrasound scanning. For instance, in fetal imaging, clinically meaningful anatomical planes are identified for performing clinical measurements. Early attempts to automate this task used supervised deep learning approaches in which the model estimates the position and orientation of the standard plane from image data [14,60]. While DL approaches have been successful, they analyze each plane independently and often ignore temporal information from adjacent frames. RL uses an alternative formulation for plane localization that models standard plane localization as a sequential navigation task. The agent iteratively adjusts the position and orientation of a virtual plane within a 3D US volume until the target anatomical plane is reached.

Across the reviewed studies, a consistent RL formulation is observed. The state typically corresponds to the currently sampled 2D plane from the 3D volume, sometimes combined with previous observations for temporal context. The actions represent discrete transformations of the plane pose, such as small rotations or translations. The reward is designed to encourage probe movements that bring the predicted plane closer to the annotated ground-truth plane based on geometric distance or orientation differences.

Early RL works such as H. Dou et al. introduced a framework in which a Deep Q-Network agent iteratively searches for the target plane starting from an initialized position, which was extended upon in subsequent works [31]. X. Yang et al. proposed a modified approach to improve search efficiency and robustness through adaptive termination

strategies and evaluation across multiple anatomical planes [32]. Similarly, X. Yang et al. proposed a collaborative multi-agent framework that enables simultaneous localization of multiple anatomical planes through communication between agents [33]. Later studies focused on improving the representation and contextual understanding of the search process. For example, Y. Zou et al. introduced an alternative parameterization of plane pose using tangent-based representations and incorporated anatomical perception modules to improve search stability and robustness [34]. Subsequent work by J. Huang et al. explored methods for incorporating additional anatomical cues, such as segmentation guidance, to facilitate more reliable plane localization [35]. More recently, B. Jiang et al. proposed the use of multiple orthogonal plane observations to provide richer contextual information during the search process, addressing the ambiguity associated with single-slice observations [36].

3.3. Summarization and Key Frame Extraction Using RL

Ultrasound videos are increasingly used alongside conventional 2D imaging in a range of clinical applications. However, most frames within an ultrasound video are redundant or diagnostically less informative, creating a need for automated video summarization and keyframe extraction. RL provides a natural framework for this task by formulating frame selection as a sequential decision-making problem, in which an agent learns to identify a compact subset of informative frames while preserving clinically relevant content.

Across the reviewed studies, the state generally consisted of frame-level image features, temporal embeddings, or spatiotemporal representations extracted from ultrasound video sequences. The action space typically involved selecting, skipping, or ranking frames, while reward functions were designed to balance representativeness, diversity, diagnostic relevance, temporal coverage, or downstream task performance. Early studies primarily used policy-gradient formulations. For example, T. Liu et al. introduced a policy-gradient RL framework with a Bi-LSTM policy network for fetal ultrasound video summarization, optimizing a composite reward based on representativeness, diversity, and diagnostic relevance [37]. R. Huang et al. later proposed a C3D-BiLSTM-based RL framework for breast ultrasound summarization, incorporating spatiotemporal feature learning to improve keyframe localization and support downstream classification under class imbalance [38].

Subsequent studies extended RL-based summarization to additional anatomical regions and methodological settings. R. P. Mathews et al. applied an unsupervised multi-latent space RL formulation to lung ultrasound video summarization, highlighting the trade-off between high compression and preservation of diagnostically useful information [39]. More recently, DQN-based approaches were explored by Ghosh et al. for pediatric wrist ultrasound video summarization and later extended to shoulder ultrasound for identifying keyframes relevant to rotator cuff tear assessment [41,40]. In addition, Y. Huang et al. introduced a hybrid framework combining RL-based slice selection with diffusion models and uncertainty-aware learning for pelvic ultrasound, demonstrating how RL can be integrated with newer generative and uncertainty-aware methods [42].

Overall, the summarization studies show a methodological progression from early policy-gradient models toward DQN-based and hybrid frameworks that combine RL with spatiotemporal feature learning, uncertainty estimation, or downstream diagnostic objectives. However, reported performance metrics varied across studies, including F1-score, precision, recall, compression rate, classification accuracy, and localization performance. Because these studies differed in anatomy, dataset size, video acquisition protocol, validation strategy, and evaluation metric, their performance results should be interpreted descriptively rather than as directly comparable measures of algorithm superiority.

3.4. Segmentation Using RL

US segmentation is well researched using both traditional image analysis and machine learning approaches. Early approaches use handcrafted features and prior assumptions about image structure, such as boundary continuity and intensity variation [61]. With the emergence of deep learning, segmentation has been predominantly formulated as a supervised learning problem, where models learn to generate pixel-wise segmentation masks directly from imaging data [3,62,63].

RL offers an alternative formulation for US segmentation. Instead of predicting the segmentation mask in a single forward pass, RL-based methods model segmentation as a sequential decision-making process, in which agents iteratively explore the image to identify object boundaries or regions of interest. Segmentation is formulated as iterative spatial exploration, with agents performing actions such as patch navigation, contour tracing, or region erasing, guided by reward functions that encourage boundary localization or lesion isolation. The primary objective is to delineate tissues or organ boundaries that are of clinical interest.

The RL formulation defines the state as the local image or contour context; the action performs a discrete local update such as moving to a new patch, following an edge trajectory, or erasing a pixel/voxel. The reward encourages behavior that improves boundary localization. For example, W. Yi et al. formulate segmentation as boundary localization, where a PPO-based agent navigates image patches to identify prostate boundary regions under weak supervision [44]. Similarly, Karunanayake et al. propose a multi-agent RL framework for cooperative contour tracing, in which multiple agents move along edge maps to construct breast tumor boundaries [43]. More recently, Y. Huang et al. introduced an erasing-based formulation, where DDQN agents progressively remove super pixels or super voxels until the lesion classification changes, enabling segmentation from bounding-box supervision in both 2D BUS and 3D ABUS [45]. All these approaches model boundary localization as an iterative process and evaluate performance using standard segmentation metrics such as Dice similarity.

3.5. Classification Using RL

Traditional image classification tasks are typically formulated as supervised learning problems, where a CNN learns a mapping from input images to a particular class label. The motivation for adopting RL in classification lies in its ability to incorporate sequential reasoning, dynamically focus on informative features, and optimize decision policies beyond fixed loss functions. This could be particularly beneficial in medical imaging tasks, where ambiguity, noise, and limited labeled data can challenge conventional supervised approaches.

RL reformulates classification as a sequential decision-making problem in which the classification process is modeled as a Markov Decision Process (MDP), where the state represents the current representation of the input image (e.g., extracted features or intermediate network outputs). The action corresponds to selecting a class label or refining feature selection, and the reward is defined based on the correctness of the prediction or improvement in classification confidence. Through iterative interaction with the environment, the RL agent learns a policy that optimizes long-term rewards.

Some works have explored RL-based classification. C. Narmatha et al. proposed a hybrid HHO-DQN-based RL framework for ovarian cyst classification using US images [46]. The approach integrates CNN-based feature extraction with a DQN, where Harris Hawks Optimization (HHO) is employed to optimize DQN hyperparameters for improved learning. Similarly, X. Xie et al. proposed a reinforced computer-aided diagnosis framework for thyroid cancer using ultrasound imaging by formulating the task as an MDP [47]. Their method integrates collaborative deep learning with RL to optimize

feature selection and diagnostic decisions, achieving improved classification accuracy over standard deep learning-based methods.

3.6. Landmark Detection Using RL

Landmark detection has traditionally been addressed using supervised learning methods, in which models are trained to predict landmark coordinates or heatmaps directly from annotated data. RL reformulates landmark detection as a sequential decision-making problem. An agent iteratively refines landmark locations by interacting with the image environment and receiving reward-based feedback. This allows the model to focus on regions of interest in a step-by-step manner, improving localization accuracy. RL-based approaches also provide flexibility in incorporating spatial constraints and adaptive search strategies. Typically, the state represents the current estimate of the landmark location along with local image features, the action corresponds to spatial movements or adjustments of the landmark position, and the reward is defined based on the reduction in localization error or proximity to the ground-truth landmark.

A communicative RL framework was introduced by G. Leroy et al. for landmark detection in brain images, employing multiple agents that collaboratively localize anatomical structures [48]. By enabling information exchange, the method improves search accuracy compared to single-agent approaches. Extending this direction, S. Lian et al. introduced an RL-based learning-to-learn framework that dynamically optimizes the objective function for landmark detection using heatmap regression [49]. Ran et al. focused on spinal landmark detection in 3D US images for scoliosis assessment, where a multi-scale RL agent is employed to iteratively localize anatomical landmarks [50]. By incorporating structure-aware representations and filtering strategies, the approach achieves improved localization accuracy and supports reliable 3D spine curvature estimation.

These studies model landmark detection as an iterative search process guided by RL, where agents progressively refine predictions rather than performing one-step regression. Despite differences in architectures- ranging from multi-agent communication to meta-learning and multi-scale strategies- they consistently leverage reward-driven optimization to improve localization accuracy, robustness to variability, and adaptability to complex anatomical structures.

3.7. Quality Assessment Using RL

Quality assessment can also be reformulated as a sequential decision-making problem. In this setting, the state typically represents the input image, extracted quality-related features, or downstream task information. The action may involve assigning a quality score, selecting or rejecting an image, choosing an enhancement operation, or deciding whether additional imaging is needed. The reward is usually based on agreement with expert assessment, improvement in image quality, or improvement in downstream task performance. This formulation enables adaptive quality evaluation driven by feedback rather than fixed predefined criteria.

The reviewed studies can be broadly grouped into two directions: image-enhancement-based approaches and task-driven quality assessment approaches. In the first direction, P. Jarosik et al. proposed a pixel-wise RL framework, known as pixelRL, for US image denoising [53]. In this approach, each pixel is treated as an agent that iteratively selects filtering actions, such as Gaussian, median, and bilateral filtering, to enhance image quality. The method was evaluated on breast US images with synthetic noise and showed improvements in PSNR and SSIM compared with noisy inputs. This study illustrates how RL can support local, interpretable, pixel-level enhancement decisions.

A second direction focuses on task-driven image quality assessment, where quality is defined by the usefulness of an image for downstream analysis rather than by visual

appearance alone. S. U. Saeed et al. proposed an RL framework for task-amenable image quality assessment, in which a controller network learns to select or reject images based on their contribution to downstream tasks such as classification and segmentation [51]. The method jointly trains a target-task predictor and a selection policy, using validation-set performance to guide the reward. In subsequent work, the same authors introduced a meta-RL framework for image quality assessment, where an RL-based IQA agent is trained with a task predictor to generalize across different quality definitions with limited supervision [52]. Similarly, S. Liu et al. introduced an RL framework for task-aware image quality evaluation, where the agent selects images according to their contribution to downstream classification or segmentation performance [54].

RL has also been used for active and personalized image acquisition. A. Saadat et al. proposed a personalized RL framework for active echocardiographic video acquisition in aortic stenosis diagnosis, where an RL agent sequentially selects the most informative views [55]. The method dynamically determines whether additional imaging is required, aiming to optimize both diagnostic accuracy and acquisition efficiency.

Across these studies, RL-based quality assessment differs from conventional quality scoring by linking image quality to adaptive decision-making or downstream task performance. However, the definition of “quality” varied across studies, including denoising quality, expert agreement, task usefulness, diagnostic informativeness, and acquisition efficiency. This heterogeneity limits direct comparison across methods and highlights the need for standardized quality definitions, validation strategies, and clinically meaningful evaluation criteria in future RL-based ultrasound quality assessment studies.

3.8. Training Optimization Using RL

RL has also been explored for optimizing training pipelines in ultrasound-based deep learning. Unlike task-level applications such as navigation, localization, or segmentation, these studies use RL to adaptively control aspects of model development, including data selection, modality fusion, loss weighting, and sampling strategy. This is particularly relevant in ultrasound imaging, where datasets are often small, imbalanced, redundant, heterogeneous, or expensive to annotate. Across the reviewed studies, the state generally represented information from the training process, such as image features, modality-specific predictions, class distribution, sample informativeness, model performance, or training dynamics. The action space varied by objective and included selecting informative samples, assigning modality weights, adjusting loss or class weights, and balancing real and synthetic data. Rewards were commonly linked to downstream performance, including accuracy, sensitivity, F1-score, AUC, annotation efficiency, or robustness to missing data.

Several studies used RL to improve data efficiency and training robustness. R. Huang et al. introduced an RL-based adaptive multimodal fusion strategy, where an agent learned optimal weighting across ultrasound imaging modalities to improve classification performance and robustness to missing data [56]. P. P. Phatale et al. proposed a DQN-based active learning framework for fetal ultrasound, in which the agent selected informative samples for annotation to reduce labeling effort while maintaining classification performance [59]. Other studies focused on class imbalance. Y. Liu et al. proposed an RL-based dynamic loss reweighting strategy for imbalanced ultrasound data [58], while S. Chen et al. introduced a REINFORCE-based class-adaptive sampling framework for long-tailed breast ultrasound classification, dynamically balancing real and synthetic data to improve rare-class recognition [57].

Overall, these studies show that RL can support ultrasound AI not only during image interpretation, but also during model training by improving data efficiency, class balance, and robustness. However, reported outcomes varied across studies, including accuracy,

sensitivity, F1-score, AUC, annotation reduction, and robustness measures. Because the studies differed in task objective, dataset size, class distribution, modality setting, and validation protocol, their results should be interpreted descriptively rather than as directly comparable evidence of superiority.

3.9. Cross-Sectional Patterns Across Studies

Across the included studies, several methodological patterns were observed beyond the individual task categories. RL was most frequently applied to ultrasound problems that naturally involve sequential decisions, including probe navigation, plane localization, landmark detection, keyframe selection, quality assessment, and training-data optimization. In these settings, the agent is not simply making a one-step prediction, but is instead learning a policy for a sequence of actions guided by task-specific feedback.

A consistent relationship was observed between task formulation and algorithm choice. Value-based methods, especially DQN and its variants, were common in tasks with discrete action spaces, such as moving a virtual plane, selecting a frame, adjusting a landmark location, or choosing among predefined navigation steps. Actor–critic methods were more frequently used in robotic navigation or path-planning problems, where the control setting may be more complex. Policy-gradient methods were mainly used in summarization and training-optimization studies, where the agent selected informative frames, slices, modalities, or samples to improve downstream performance.

The relationship between datasets and performance metrics was less consistent. Studies varied substantially in anatomical target, data type, cohort size, scanner setting, validation strategy, and reported metrics. For example, navigation and plane-localization studies often used geometric, reachability, or image-similarity metrics, whereas segmentation studies used overlap-based metrics, summarization studies used frame-selection or downstream diagnostic metrics, and training-optimization studies used classification or annotation-efficiency outcomes. Because these metrics evaluate different objectives, they should not be interpreted as directly comparable measures of algorithm superiority.

A qualitative assessment of validation maturity also showed that most studies remain at an early feasibility or intermediate methodological-validation stage. Early feasibility studies generally used small or private datasets, internal validation, controlled environments, or proof-of-concept evaluation. Intermediate studies typically included larger datasets, more detailed ablation experiments, partial data or code availability, or more clinically realistic evaluation settings. Few studies provided evidence of prospective, multi-center, real-time, or workflow-based validation. Therefore, although RL shows strong potential for ultrasound tasks that require adaptive decision-making, the current evidence base remains insufficient to support broad clinical deployment.

4. Discussion

This scoping review highlights several emerging trends in the application of RL to US image analysis. Overall, RL remains a relatively underexplored approach compared with conventional supervised deep learning methods, but the existing studies demonstrate its potential across multiple tasks involving sequential decision-making. The task-level taxonomy further shows that RL has mainly been used where ultrasound analysis can be formulated as a sequence of decisions rather than as a single static prediction.

Across the reviewed literature, RL has been most frequently applied to navigation and localization-related tasks, including probe guidance, anatomical landmark detection, and standard plane localization. These tasks align naturally with the RL framework, as they involve iterative interactions with the imaging environment, where an agent adjusts its actions to reach a target anatomical structure or optimal imaging view. This reflects the

real-world process of ultrasound acquisition, which relies on continuous feedback and adaptive probe manipulation.

In contrast, RL for traditional pixel-wise image analysis tasks such as segmentation is less common. Similarly, applications in classification, diagnostic decision-making, and image quality assessment remain relatively limited. These observations suggest that RL is currently being explored primarily as a complementary approach rather than a replacement for conventional deep learning models, often within hybrid frameworks in which deep neural networks provide feature representations while RL governs decision policies.

From a methodological perspective, most studies rely on value-based RL approaches, particularly variants of DQN. These methods are widely adopted due to their relative simplicity and compatibility with image-based state representations. The predominance of value-based methods likely reflects the nature of many ultrasound tasks reviewed, including probe navigation, plane localization, landmark detection, and frame selection, which are commonly formulated as sequential decision-making problems with discrete action spaces. In such settings, value-based approaches are relatively straightforward to implement and have demonstrated effectiveness in learning image-guided search and navigation policies. In contrast, policy-based and actor–critic methods, which may be advantageous for continuous control problems, have been explored less frequently in the current ultrasound literature. This indicates that current research has primarily focused on discrete decision-making tasks and relatively simple action policies.

A cross-sectional comparison also shows that reported performance metrics are closely tied to task formulation and therefore cannot be compared directly across all studies. Navigation and plane-localization studies commonly used reachability, geometric error, orientation error, image similarity, or measurement agreement. Segmentation studies used overlap-based segmentation metrics, while summarization and keyframe extraction studies reported frame-selection, compression, localization, or downstream diagnostic metrics. Training-optimization studies focused on classification performance, robustness, class imbalance, or annotation efficiency. These differences indicate that the field lacks a unified benchmarking framework and that current performance results should be interpreted within each task context rather than as direct evidence that one RL algorithm is broadly superior to another.

Overall, RL has primarily been explored in ultrasound imaging for navigation, localization, frame selection, and other sequential decision-making tasks, while its application to broader diagnostic workflows remains limited. The field is promising but still methodologically fragmented, with substantial heterogeneity in state-action-reward design, dataset availability, validation protocols, and clinical evaluation.

4.1. Methodological Challenges

Several methodological challenges are consistently observed across the literature. A central issue is the design of reward functions, which varies substantially between studies and strongly influences performance. This variability limits direct comparability across datasets and applications. However, reward design should not be viewed only as a limitation; it is also one of the main methodological contributions of RL. In ultrasound applications, the reward function defines what the agent is encouraged to optimize, such as target localization, anatomical visibility, image quality, diagnostic usefulness, frame diversity, classification performance, or efficient acquisition. Some studies, particularly in video summarization and keyframe extraction, addressed reward design by testing different combinations of reward terms. These reward-specific ablation studies show that individual reward components can substantially affect model behavior and task performance. However, such analyses were not reported consistently across all ultrasound RL

applications, and the clinical or task-specific interpretation of each reward term was not always clearly justified.

Future ultrasound RL studies should therefore make reward design more transparent and clinically grounded. Authors should explicitly report each reward component, explain why it is relevant to the ultrasound task, and evaluate its contribution through reward-specific ablation studies whenever possible. Multi-objective, hierarchical, uncertainty-aware, and human-in-the-loop reward formulations may also help balance competing goals in ultrasound workflows, such as diagnostic usefulness, anatomical consistency, uncertainty reduction, efficient acquisition, and patient comfort. More consistent reporting of state, action, reward, and termination criteria would improve reproducibility and make comparisons across studies more meaningful.

Another methodological challenge is the reliance on strong supervision or constrained initialization strategies. For example, plane localization methods often depend on ground-truth annotations and warm-start initialization, which can restrict exploration and limit the ability of models to perform fully autonomous search. In these studies, learning objectives and evaluation are commonly based on geometric similarity to annotated target planes, using measures such as distance error, angular deviation, or structural similarity. Although useful for controlled evaluation, these metrics may not fully reflect clinical usefulness or robustness in real-world scanning conditions.

RL-based models are also generally more complex than conventional supervised approaches, often requiring multiple engineered components and careful tuning. This increases computational cost and may limit scalability and real-time applicability, as training often requires substantial time and access to GPU-accelerated resources. Furthermore, performance variability across studies suggests limited generalizability, influenced by differences in dataset size, imaging modality, anatomical region, scanner type, and acquisition protocol.

Evaluation protocols also remain inconsistent, particularly for tasks such as video summarization and quality assessment, where standardized metrics are lacking. This makes cross-study comparison difficult and limits the ability to establish best practices. Additionally, many approaches are developed and evaluated in simplified or controlled environments, which may not fully reflect real-world clinical variability.

4.2. Dataset Limitations and Reproducibility

A consistent limitation across the reviewed studies is the reliance on small, single-center, and often private datasets. Most studies used limited numbers of patients, images, videos, or imaging volumes, which restricts the generalizability and robustness of the proposed RL methods. In addition, datasets varied widely in anatomical region, scanner type, acquisition protocol, operator expertise, image quality, and annotation strategy. This heterogeneity makes it difficult to compare methods directly across studies or determine whether reported performance would generalize to different clinical environments.

External validation was also limited across the included literature. Most studies were evaluated retrospectively using internal datasets, while few reported testing across independent institutions, scanners, operators, or patient populations. As a result, the current evidence remains largely feasibility-focused, with limited support for real-world clinical deployment. This is particularly important in ultrasound, where image appearance can vary substantially depending on probe position, operator experience, patient body habitus, scanner settings, and acquisition protocol.

Reproducibility remains another major challenge. Many studies used proprietary clinical datasets, and public access to datasets and source code was limited. This restricts independent validation, benchmarking, and fair comparison between RL approaches. Overall, these limitations highlight the need for larger and more diverse datasets, external

validation, standardized reporting, open data and code sharing where possible, and reproducible benchmarks to support the development and clinical translation of RL-based ultrasound methods.

To improve reproducibility, future studies should report RL design and validation details in a standardized way. At minimum, authors should clearly describe the state representation, action space, reward components and weights, termination criteria, training environment, exploration strategy, hyperparameters, train-validation-test split strategy, and whether splitting was performed at the patient, image, frame, video, or volume level. For clinical ultrasound applications, studies should also report scanner type, acquisition protocol, operator experience where available, annotation procedure, evaluation metrics, and failure cases. Reward-specific ablation studies, external validation across scanners or institutions, prospective testing, and the release of code, trained models, or de-identified datasets where feasible would substantially improve comparability and independent verification.

4.3. Clinical Relevance, Validation Maturity, and Deployment

Although RL shows strong potential for modeling ultrasound workflows, a clear gap remains between methodological development and clinical use. Most studies were evaluated retrospectively on curated datasets, with limited real-time, prospective, or external validation. As a result, it remains unclear how well these methods generalize across different institutions, scanners, operators, patient populations, and acquisition protocols.

To clarify the level of evidence across the field, the included studies can be interpreted qualitatively according to validation maturity. Early feasibility studies generally demonstrate that an RL formulation is technically possible, but often rely on small datasets, controlled environments, internal validation, or private data. Intermediate validation studies include stronger methodological evaluation, such as larger datasets, ablation experiments, more realistic testing environments, or partial availability of data or code. Translational or clinically mature studies would require prospective evaluation, external validation, multi-center testing, real-time implementation, workflow integration, user interaction studies, or evidence that the system improves clinical decision-making or patient outcomes. Using this qualitative interpretation, most of the reviewed literature remains in the early feasibility or intermediate validation stage. Some studies provide promising technical demonstrations, including robotic navigation, standard-plane localization, frame selection, segmentation, and training optimization. However, few studies provide evidence of prospective clinical testing, multi-center generalization, or routine workflow deployment. Therefore, clinical readiness remains limited across the field.

This is especially important in ultrasound, where image quality and anatomical appearance can vary substantially depending on probe position, operator experience, scanner type, and patient-specific factors. Many RL-based methods are developed under controlled conditions, which may not fully reflect routine clinical variability. Therefore, models that perform well in single-center or retrospective settings may not show the same reliability in real-world practice. While RL-based approaches may improve adaptive guidance, efficiency, and workflow integration, their impact on diagnostic accuracy, clinical decision-making, user experience, and patient outcomes remains insufficiently established. Studies should therefore evaluate RL systems in real clinical workflows and assess how they interact with clinicians and sonographers during ultrasound acquisition and interpretation.

4.4. Future Directions

Future studies should focus on making RL-based ultrasound methods more reliable, reproducible, and useful in real clinical settings. One important direction is to design

better reward functions. In ultrasound, goals such as image quality, diagnostic confidence, correct probe movement, and clinical usefulness are difficult to measure with simple rewards. Future work may benefit from using expert knowledge, uncertainty-aware methods, and human feedback during model development. Reward functions should be clinically interpretable, and individual reward components should be evaluated through ablation studies whenever possible.

More research is also needed in anatomical regions that are still less explored. Current studies have focused more on fetal, breast, vascular, cardiac, and abdominal ultrasound. However, areas such as musculoskeletal, pediatric, emergency and point-of-care, thyroid, renal and urinary tract, pelvic, and lung ultrasound have received less attention. These areas may benefit from RL for probe guidance, view selection, image quality assessment, keyframe selection, and temporal analysis.

Future work should also use larger and more diverse datasets, test models across different hospitals and ultrasound machines, and evaluate them in real-time clinical settings. Standard reporting of state, action, reward, training strategy, interaction type, validation method, and performance metrics would make studies easier to compare. In particular, future papers should clearly distinguish whether the RL agent interacts with a physical robotic system, a simulated or reconstructed 3D ultrasound volume, an offline video sequence, an image-analysis model, or a training pipeline. This distinction is important because clinical readiness differs substantially between robotic acquisition, offline image interpretation, and model-training optimization.

Open datasets, shared code, and reproducible benchmarks are also important. Advanced RL methods such as hierarchical RL, multi-agent RL, offline RL, safe RL, and human-in-the-loop RL should be explored further to improve stability, efficiency, interpretability, and clinical translation. However, these methods should be evaluated using clinically meaningful validation protocols rather than only task-specific technical metrics. Future studies should therefore include external validation, scanner- and operator-level robustness testing, failure-case analysis, and assessment of how RL-based tools affect clinical workflow, user confidence, and diagnostic decision-making.

4.5. Limitations of the Review

This review has several limitations. First, it was restricted to studies involving clinical ultrasound data. Studies based exclusively on phantom or simulated datasets were excluded, which may have limited the scope of methodological insights, particularly because many reinforcement learning approaches are initially developed and validated in simulated environments. Second, only English-language publications were included, and the literature search was limited to selected databases. Although major databases were searched, relevant studies indexed elsewhere, unpublished work, or preprints may have been missed. Third, given the rapid evolution of reinforcement learning and medical imaging research, additional studies may have emerged after the search date and are therefore not captured in this review.

Another limitation relates to the heterogeneity of the included studies. The reviewed papers varied substantially in ultrasound task, anatomical region, dataset size, imaging protocol, RL formulation, validation approach, and performance metric. As a result, direct quantitative comparison or pooling of results was not appropriate. Instead, the findings were synthesized narratively to map the current evidence, identify methodological trends, and highlight research gaps. Finally, as this was a scoping review, we did not perform a formal risk-of-bias assessment or quantitative methodological quality scoring of individual studies. Instead, we provided a qualitative discussion of recurring methodological limitations, including small datasets, limited external validation, heterogeneous reward formulations, inconsistent reporting of state-action-reward designs, and restricted

availability of public datasets and source code. Therefore, this review provides a broad picture of current research trends, rather than a formal comparison of study quality or model performance.

The evidence maturity and clinical readiness categories used in this review were qualitative descriptors based on reported study characteristics and were intended to support interpretation, not to provide a formal evidence-grade rating.

5. Conclusions

This scoping review mapped the current landscape of reinforcement learning (RL) in ultrasound image analysis and showed that the field, although still emerging, is expanding across a diverse set of applications. Among the 39 included studies, RL was applied most frequently to tasks that naturally involve sequential decision-making, particularly navigation guidance, plane localization, landmark detection, and video summarization. Across these applications, value-based methods—especially DQN and its variants—were the dominant approach, while policy-based and actor–critic methods were less common. Overall, the literature indicates that RL is particularly well suited to ultrasound workflows in which decisions are interdependent and must be optimized over time.

The review also highlights that RL is not a replacement for conventional deep learning, but rather a complementary paradigm that enables dynamic, reward-driven decision-making on top of learned image representations. This is especially relevant for ultrasound, where image acquisition, localization, quality assessment, and interpretation often occur as iterative and interactive processes. However, the current evidence base remains limited. Most studies rely on small, single-center, private datasets, retrospective or internally focused evaluation, heterogeneous reward formulations, and inconsistent validation protocols. Public datasets, shared source code, prospective validation, and standardized benchmarking frameworks also remain scarce, limiting reproducibility, comparability, and clinical translation.

Anatomically, the reviewed literature is unevenly distributed. RL applications have been explored more frequently in areas such as fetal, breast, vascular, cardiac, and abdominal ultrasound, while several clinically important ultrasound domains remain underexplored. In particular, musculoskeletal ultrasound, pediatric ultrasound, emergency and point-of-care ultrasound, thyroid ultrasound, renal and urinary tract ultrasound, pelvic ultrasound, and lung ultrasound have received limited attention. This imbalance suggests that the potential of RL has not yet been fully examined across the broader range of ultrasound applications where adaptive acquisition guidance, quality assessment, view selection, or temporal reasoning may be valuable.

Taken together, RL holds meaningful promise for next-generation ultrasound systems, particularly in applications requiring adaptive guidance, temporal reasoning, and workflow optimization. However, its clinical utility is not yet established. Further progress will require larger and more diverse clinical datasets, clearer reporting of state-action-reward formulations, clinically meaningful reward designs, standardized evaluation protocols, external validation across scanners and institutions, and prospective real-time testing in clinical workflows. This review therefore serves as both a summary of current developments and a practical entry point for future research in this evolving field.

Author Contributions: Conceptualization, M.E., S.G., G.V., J.L.J. and A.H.; methodology, S.G., M.M. and M.E.; formal analysis, M.E., M.M., S.G. and A.H.; investigation, M.E., M.M., S.G. and A.H.; resources, J.L.J. and A.H.; data curation, M.E., M.M. and S.G.; writing—original draft preparation, M.E., S.G. and A.H.; writing—review and editing, M.E., S.G., G.V., J.L.J. and A.H.; visualization, M.E.; supervision, A.H.; project administration, J.L.J. and A.H.; funding acquisition, J.L.J. and A.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC) Discovery Grant, grant number RES0074575, and the Alberta Innovates AICE Concepts Grant, grant number RES0064064, awarded to A.H.; Alberta Innovates and WCHRI post-doctoral funding awarded to S.G.; and CIHR and CIFAR funding awarded to J.L.J.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new primary data were generated in this study. All charted data supporting the findings of this scoping review are reported within the article and were derived from the cited publications.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ML	Machine learning
DL	Deep learning
RL	Reinforcement learning
DRL	Deep reinforcement learning
US	Ultrasound
DQN	Deep Q-network

References

1. Wang, Y.; Ge, X.; Ma, H.; Qi, S.; Zhang, G.; Yao, Y. Deep Learning in Medical Ultrasound Image Analysis: A Review. *IEEE Access* **2021**, *9*, 54310–54324. <https://doi.org/10.1109/ACCESS.2021.3071301>.
2. Brattain, L.J.; Telfer, B.A.; Dhyani, M.; Grajo, J.R.; Samir, A.E. Machine learning for medical ultrasound: Status, methods, and future opportunities. *Abdom. Radiol.* **2018**, *43*, 786–799.
3. Zhou, Z.; Siddiquee, M.M.R.; Tajbakhsh, N.; Liang, J. UNet++: A nested U-Net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support, LNCS*; Springer: Cham, Switzerland, 2018; Volume 11045, pp. 3–11.
4. Cai, L.; Gao, J.; Zhao, D. A review of the application of deep learning in medical image classification and segmentation. *Ann. Transl. Med.* **2020**, *8*, 713.
5. Zhao, Z.-Q.; Zheng, P.; Xu, S.-T.; Wu, X. Object detection with deep learning: A review. *arXiv* **2018**, arXiv:1807.05511.
6. Sharma, H.; Drukker, L.; Chatelain, P.; Droste, R.; Papageorgiou, A.T.; Noble, J.A. Knowledge representation and learning of operator clinical workflow from full-length routine fetal ultrasound scan videos. *Med. Image Anal.* **2021**, *69*, 101973.
7. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*, 2nd ed.; MIT Press: Cambridge, MA, USA, 2018.
8. Zhou, S.K.; Le, H.N.; Luu, K.; Nguyen, H.V.; Ayache, N. Deep reinforcement learning in medical imaging: A literature review. *Med. Image Anal.* **2021**, *73*, 102193. <https://doi.org/10.1016/j.media.2021.102193>.
9. Akkus, Z.; Cai, J.; Boonrod, A.; Zeinoddini, A.; Weston, A.D.; Philbrick, K.A.; Erickson, B.J. A survey of deep-learning applications in ultrasound: Artificial intelligence-powered ultrasound for improving clinical workflow. *J. Am. Coll. Radiol.* **2019**, *16*, 1318–1328.
10. Yi, J.; Kang, H.K.; Kwon, J.-H.; Kim, K.-S.; Park, M.H.; Seong, Y.K.; Kim, D.W.; Ahn, B.; Ha, K.; Lee, J.; et al. Technology trends and applications of deep learning in ultrasonography: Image quality enhancement, diagnostic support, and improving workflow efficiency. *Ultrasonography* **2021**, *40*, 7–22.
11. Peng, S.; Liu, Y.; Lv, W.; Liu, L.; Zhou, Q.; Yang, H.; Ren, J.; Liu, G.; Wang, X.; Zhang, X.; et al. Deep learning-based artificial intelligence model to assist thyroid nodule diagnosis and management: A multicentre diagnostic study. *Lancet Digit. Health* **2021**, *3*, e250–e259.
12. Gao, Y.; Zeng, S.; Xu, X.; Li, H.; Yao, S.; Song, K.; Li, X.; Chen, L.; Tang, J.; Xing, H.; et al. Deep learning-enabled pelvic ultrasound images for accurate diagnosis of ovarian cancer in China: A retrospective, multicentre, diagnostic study. *Lancet Digit. Health* **2022**, *4*, e179–e187. [https://doi.org/10.1016/S2589-7500\(21\)00278-8](https://doi.org/10.1016/S2589-7500(21)00278-8).

13. Leclerc, S.; Smistad, E.; Pedrosa, J.; Ostvik, A.; Cervenansky, F.; Espinosa, F.; Espeland, T.; Berg, E.A.R.; Jodoin, P.-M.; Grenier, T.; et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography. *IEEE Trans. Med. Imaging* **2019**, *38*, 2198–2210.
14. Baumgartner, C.F.; Kamnitsas, K.; Matthew, J.; Fletcher, T.P.; Smith, S.; Koch, L.M.; Kainz, B.; Rueckert, D. SonoNet: Real-time detection and localisation of fetal standard scan planes in freehand ultrasound. *IEEE Trans. Med. Imaging* **2017**, *36*, 2204–2215.
15. Dong, J.; Liu, S.; Liao, Y.; Wen, H.; Lei, B.; Li, S.; Wang, T. A generic quality control framework for fetal ultrasound cardiac four-chamber planes. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 931–942.
16. Hyun, D.; Brickson, L.L.; Looby, K.T.; Dahl, J.J. Beamforming and speckle reduction using neural networks. *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **2019**, *66*, 898–910.
17. Luijten, B.; Cohen, R.; de Bruijn, F.J.; Schmeitz, H.A.W.; Misch, M.; Eldar, Y.C.; van Sloun, R.J.G. Adaptive ultrasound beamforming using deep learning. *IEEE Trans. Med. Imaging* **2020**, *39*, 3967–3978.
18. Ouyang, D.; He, B.; Ghorbani, A.; Yuan, N.; Ebinger, J.; Langlotz, C.P.; Heidenreich, P.A.; Harrington, R.A.; Liang, D.H.; Ashley, E.A.; et al. Video-based AI for beat-to-beat assessment of cardiac function. *Nature* **2020**, *580*, 252–256.
19. Misra, S.; Yoon, C.; Kim, K.-J.; Managuli, R.; Barr, R.G.; Baek, J.; Kim, C. Deep learning-based multimodal fusion network for segmentation and classification of breast cancers using B-mode and elastography ultrasound images. *Bioeng. Transl. Med.* **2023**, *8*, e10480.
20. Gu, J.; Zhong, X.; Fang, C.; Lou, W.; Fu, P.; Woodruff, H.C.; Wang, B.; Jiang, T.; Lambin, P. Deep learning of multimodal ultrasound: Stratifying the response to neoadjuvant chemotherapy in breast cancer before treatment. *Oncologist* **2024**, *29*, e187–e197.
21. Hase, H.; Azampour, M.F.; Tirindelli, M.; Paschali, M.; Simson, W.; Fatemizadeh, E.; Navab, N. Ultrasound-Guided Robotic Navigation with Deep Reinforcement Learning. In Proceedings of the 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Las Vegas, NV, USA, 24 October 2020–24 January 2021; IEEE: Piscataway, NJ, USA, 2020; pp. 5534–5541. <https://doi.org/10.1109/IROS45743.2020.9340913>.
22. Chen, Y.; Yang, C.; Li, M.; Dai, S.-L. Learning to Predict Action Based on B-ultrasound Image Information. In Proceedings of the 2021 6th IEEE International Conference on Advanced Robotics and Mechatronics (ICARM), Chongqing, China, 3–5 July 2021; IEEE: New York, NY, USA, 2021; pp. 492–497. <https://doi.org/10.1109/ICARM52023.2021.9536054>.
23. Li, K.; Xu, Y.; Wang, J.; Ni, D.; Liu, L.; Meng, M.Q.-H. Image-guided navigation of a robotic ultrasound probe for autonomous spinal sonography using a shadow-aware dual-agent framework. *IEEE Trans. Med. Robot. Bionics* **2022**, *4*, 130–144.
24. Shen, C.; Deng, Z.; Wang, J.; Wang, S.; Chen, C. Towards Autonomous Robotic Ultrasound Scanning Using the Reinforcement Learning-Based Volumetric Data Navigation Method. In Proceedings of the 2023 WRC Symposium on Advanced Robotics and Automation (WRC SARA), Beijing, China, 19 August 2023; IEEE: New York, NY, USA, 2023; pp. 335–340. <https://doi.org/10.1109/WRC SARA60131.2023.10261822>.
25. Ning, G.; Liang, H.; Zhang, X.; Liao, H. Autonomous Robotic Ultrasound Vascular Imaging System with Decoupled Control Strategy for External-Vision-Free Environments. *IEEE Trans. Biomed. Eng.* **2023**, *70*, 3166–3177. <https://doi.org/10.1109/TBME.2023.3279114>.
26. Su, K.; Liu, J.; Ren, X.; Huo, Y.; Du, G.; Zhao, W.; Wang, X.; Liang, B.; Li, D.; Liu, P.X. A fully autonomous robotic ultrasound system for thyroid scanning. *Nat. Commun.* **2024**, *15*, 4004.
27. Shida, Y.; Kumagai, S.; Iwata, H. Robotic navigation with deep reinforcement learning in transthoracic echocardiography. *Int. J. Comput. Assist. Radiol. Surg.* **2025**, *20*, 191–202.
28. Su, K.; Du, G.; Wang, X.; Guan, Q. Tissue-view map for robotic carotid artery ultrasound scanning using reinforcement learning. *IEEE Robot. Autom. Lett.* **2025**, *10*, 5178–5185.
29. Wang, X.; He, Z.; Chen, P.; Wang, Z.; Sun, T. Dual-agent-based robotic ultrasound path planning and interaction control in external-vision-independent environments. *IEEE Trans. Autom. Sci. Eng.* **2025**, *22*, 21674–21685.
30. Hsu, H.-L.; Zahiri, M.; Li, G.Y.; Al Mukaddim, R.; Lee, H.; Wilson, M.G.; Grube, J.; Schmidt, S.; Ghoshal, G.; Raju, B. Active guidance in ultrasound bladder scanning using reinforcement learning. *Sci. Rep.* **2026**, *16*, 5273. <https://doi.org/10.1038/s41598-026-35285-z>.
31. Dou, H.; Yang, X.; Qian, J.; Xue, W.; Qin, H.; Wang, X.; Yu, L.; Wang, S.; Xiong, Y.; Heng, P.-A.; et al. Agent with warm start and active termination for plane localization in 3D ultrasound. In *Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2019; pp. 290–298.
32. Yang, X.; Dou, H.; Huang, R.; Xue, W.; Huang, Y.; Qian, J.; Zhang, Y.; Luo, H.; Guo, H.; Wang, T.; et al. Agent with Warm Start and Adaptive Dynamic Termination for Plane Localization in 3D Ultrasound. *IEEE Trans. Med. Imaging* **2021**, *40*, 1950–1961. <https://doi.org/10.1109/TMI.2021.3069663>.

33. Yang, X.; Huang, Y.; Huang, R.; Dou, H.; Li, R.; Qian, J.; Huang, X.; Shi, W.; Chen, C.; Zhang, Y.; et al. Searching collaborative agents for multi-plane localization in 3D ultrasound. *Med. Image Anal.* **2021**, *72*, 102119. <https://doi.org/10.1016/j.media.2021.102119>.
34. Zou, Y.; Dou, H.; Huang, Y.; Yang, X.; Qian, J.; Zhen, C.; Ji, X.; Ravikumar, N.; Chen, G.; Huang, W.; et al. Agent with tangent-based formulation and anatomical perception for standard plane localization in 3D ultrasound. In *Lecture Notes in Computer Science*; Springer Nature: Cham, Switzerland, 2022; pp. 300–309.
35. Huang, J.; Wang, R.; Jiang, W.; Shao, S.; Chen, T. Agent Based Fetal Face Segmentation for Standard Plane Localization in 3D Ultrasound. In *Proceedings of the 2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Honolulu, Oahu, HI, USA, 1–4 October 2023*; IEEE: New York, NY, USA, 2023; pp. 5317–5322. <https://doi.org/10.1109/SMC53992.2023.10393929>.
36. Jiang, B.; Xu, K.; Graham, E.; Taylor, R.H.; Kang, J.; Unberath, M.; Bector, E.M. Fetal Ultrasound Standard Plane Extraction using Orthogonal Triple-slice Deep Reinforcement Learning Agent. In *Proceedings of the 2024 IEEE Ultrasonics, Ferroelectrics, and Frequency Control Joint Symposium (UFFC-JS), Taipei, Taiwan, 22–26 September 2024*; IEEE: New York, NY, USA, 2024; pp. 1–4. <https://doi.org/10.1109/UFFC-JS60046.2024.10794037>.
37. Liu, T.; Meng, Q.; Huang, J.-J.; Vlontzos, A.; Rueckert, D.; Kainz, B. Video Summarization Through Reinforcement Learning With a 3D Spatio-Temporal U-Net. *IEEE Trans. Image Process.* **2022**, *31*, 1573–1586. <https://doi.org/10.1109/TIP.2022.3143699>.
38. Huang, R.; Ying, Q.; Lin, Z.; Zheng, Z.; Tan, L.; Tang, G.; Zhang, Q.; Luo, M.; Yi, X.; Liu, P.; et al. Extracting keyframes of breast ultrasound video using deep reinforcement learning. *Med. Image Anal.* **2022**, *80*, 102490.
39. Mathews, R.P.; Panicker, M.R.; Hareendranathan, A.R.; Chen, Y.T.; Jaremko, J.L.; Buchanan, B.; Narayan, K.V.; Kesavadas, C.; Mathews, G. Unsupervised multi-latent space RL framework for video summarization in ultrasound imaging. *IEEE J. Biomed. Health Inform.* **2023**, *27*, 227–238. <https://doi.org/10.1109/JBHI.2022.3208779>.
40. Ghosh, S.; Vadali, G.; Singh, A.; Zhou, Y.; Felfeliyan, B.; Wahd, A.; Knight, J.; Panicker, M.R.; Jaremko, J.L.; Hareendranathan, A.R. Shoulder Rotator Cuff Tear Detection from Ultrasound Videos Using Deep Reinforcement Learning. In *Proceedings of the 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI), Houston, TX, USA, 14–17 April 2025*; IEEE: New York, NY, USA, 2025; pp. 1–4. <https://doi.org/10.1109/ISBI60581.2025.10981103>.
41. Ghosh, S.; Vadali, G.; Zhou, Y.; Singh, A.; Knight, J.; Keen, C.; Panicker, M.R.; Hareendranathan, A.R.; Jaremko, J.L. Video Summarization and Fracture Detection in Pediatric Wrist Ultrasound Using Deep Reinforcement Learning. In *Proceedings of the 2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Copenhagen, Denmark, 14–17 July 2025*; IEEE: New York, NY, USA, 2025; pp. 1–4. <https://doi.org/10.1109/EMBC58623.2025.11251547>.
42. Huang, Y.; Xu, Y.; Dou, H.; Deng, J.; Yang, X.; Zheng, H.; Ni, D. Uncertainty-aware diffusion and reinforcement learning for joint plane localization and anomaly diagnosis in 3D ultrasound. In *Medical Image Computing and Computer Assisted Intervention (MICCAI)*; Springer Nature: Cham, Switzerland, 2026; pp. 650–660.
43. Karunanayake, N.; Moodleah, S.; Makhanov, S.S. Edge-driven multi-agent reinforcement learning: A novel approach to ultrasound breast tumor segmentation. *Diagnostics* **2023**, *13*, 3611. <https://doi.org/10.3390/diagnostics13243611>.
44. Yi, W.; Stavrinides, V.; Baum, Z.M.C.; Yang, Q.; Barratt, D.C.; Clarkson, M.J.; Hu, Y.; Saeed, S.U. Boundary-RL: Reinforcement learning for weakly supervised prostate segmentation in TRUS images. In *Lecture Notes in Computer Science*; Springer Nature: Cham, Switzerland, 2024; pp. 277–288.
45. Huang, Y.; Chang, A.; Dou, H.; Tao, X.; Zhou, X.; Cao, Y.; Huang, R.; Frangi, A.F.; Bao, L.; Yang, X.; et al. Flip Learning: Weakly supervised erase to segment nodules in breast ultrasound. *Med. Image Anal.* **2025**, *102*, 103552. <https://doi.org/10.1016/j.media.2025.103552>.
46. Narmatha, C.; Manimegalai, P.; Krishnadass, J.; Valsalan, P.; Manimurugan, S.; Mustafa, M. Ovarian cysts classification using novel deep reinforcement learning with Harris Hawks Optimization method. *J. Supercomput.* **2023**, *79*, 1374–1397.
47. Xie, X.; Tian, Y.; Ota, K.; Dong, M.; Liu, Z.; Jin, H.; Yao, D. Reinforced computer-aided framework for diagnosing thyroid cancer. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2024**, *21*, 737–747.
48. Leroy, G.; Rueckert, D.; Alansary, A. Communicative reinforcement learning agents for landmark detection in brain images. In *Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2020; pp. 177–186.
49. Lian, S.; Luo, Z.; Feng, C.; Li, S.; Li, S. APRIL: Anatomical prior-guided reinforcement learning for accurate carotid lumen diameter and intima-media thickness measurement. *Med. Image Anal.* **2021**, *71*, 102040. <https://doi.org/10.1016/j.media.2021.102040>.
50. Ran, Q.-Y.; Miao, J.; Zhou, S.-P.; Hua, S.-H.; He, S.-Y.; Zhou, P.; Wang, H.-X.; Zheng, Y.-P.; Zhou, G.-Q. Automatic 3-D spine curve measurement in freehand ultrasound via structure-aware reinforcement learning spinous process localization. *Ultrasonics* **2023**, *132*, 107012. <https://doi.org/10.1016/j.ultras.2023.107012>.

51. Saeed, S.U.; Fu, Y.; Baum, Z.M.C.; Yang, Q.; Rusu, M.; Fan, R.E.; Sonn, G.A.; Barratt, D.C.; Hu, Y. Learning image quality assessment by reinforcing task-amenable data selection. In *Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2021; pp. 755–766.
52. Saeed, S.U.; Fu, Y.; Stavrinides, V.; Baum, Z.M.C.; Yang, Q.; Rusu, M.; Fan, R.E.; Sonn, G.A.; Noble, J.A.; Barratt, D.C.; et al. Adaptable image quality assessment using meta-reinforcement learning of task amenability. In *Simplifying Medical Ultrasound*; Springer International Publishing: Cham, Switzerland, 2021; pp. 191–201.
53. Jarosik, P.; Lewandowski, M.; Klimonda, Z.; Byra, M. Pixel-Wise Deep Reinforcement Learning Approach for Ultrasound Image Denoising. In *Proceedings of the 2021 IEEE International Ultrasonics Symposium (IUS), Xi'an, China, 11–16 September 2021*; IEEE: New York, NY, USA, 2021; pp. 1–4. <https://doi.org/10.1109/IUS52206.2021.9593591>.
54. Liu, S.; Ying, Q.; He, S.; Yang, X.; Ni, D.; Huang, R. Hierarchical Agent-Based Reinforcement Learning Framework for Automated Quality Assessment of Fetal Ultrasound Video. In *Proceedings of the 2023 IEEE 20th International Symposium on Biomedical Imaging (ISBI), Cartagena, Colombia, 18–21 April 2023*; IEEE: Piscataway, NJ, USA, 2023; pp. 1–5. <https://doi.org/10.1109/ISBI53787.2023.10230581>.
55. Saadat, A.; Hashemi, N.; Vaseli, H.; Tsang, M.Y.; Luong, C.; Van de Panne, M.; Tsang, T.S.M.; Abolmaesumi, P. PRECISE-AS: Personalized reinforcement learning for efficient point-of-care echocardiography in aortic stenosis diagnosis. In *Lecture Notes in Computer Science*; Springer Nature: Cham, Switzerland, 2026; pp. 474–484.
56. Huang, R.; Lin, Z.; Dou, H.; Wang, J.; Miao, J.; Zhou, G.; Jia, X.; Xu, W.; Mei, Z.; Dong, Y.; et al. AW3M: An auto-weighting and recovery framework for breast cancer diagnosis using multi-modal ultrasound. *Med. Image Anal.* **2021**, *72*, 102137. <https://doi.org/10.1016/j.media.2021.102137>.
57. Chen, S.; Zhou, X.; Wang, Y.; Huang, Y.; Chang, A.; Ni, D.; Huang, R. Subtyping breast lesions via generative augmentation based long-tailed recognition in ultrasound. In *Medical Image Computing and Computer Assisted Intervention (MICCAI)*; Springer Nature: Cham, Switzerland, 2026; pp. 519–529.
58. Liu, Y.; Li, J.; Zhao, C.; Zhang, Y.; Chen, Q.; Qin, J.; Dong, L.; Wang, T.; Jiang, W.; Lei, B.; et al. FAMF-Net: Feature Alignment Mutual Attention Fusion with Region Awareness for Breast Cancer Diagnosis via Imbalanced Data. *IEEE Trans. Med. Imaging* **2025**, *44*, 1153–1167. <https://doi.org/10.1109/TMI.2024.3485612>.
59. Phatale, P.P.; Gargi, S.; Surendran, S. A deep RL-based active learning approach for fetal ultrasound image plane labeling and classification. In *Lecture Notes in Networks and Systems*; Springer Nature: Singapore, 2025; pp. 485–492.
60. Li, Y.; Khanal, B.; Hou, B.; Alansary, A.; Cerrolaza, J.J.; Sinclair, M.; Matthew, J.; Gupta, C.; Knight, C.; Kainz, B.; et al. Standard plane detection in 3D fetal ultrasound using an iterative transformation network. *arXiv* **2018**, arXiv:1806.05550.
61. Noble, J.A.; Boukerroui, D. Ultrasound image segmentation: A survey. *IEEE Trans. Med. Imaging* **2006**, *25*, 987–1010.
62. Ronneberger, O.; Fischer, P.; Brox, T. U-Net: Convolutional networks for biomedical image segmentation. In *Lecture Notes in Computer Science*; Springer: Cham, Switzerland, 2015; pp. 234–241.
63. Isensee, F.; Jaeger, P.F.; Kohl, S.A.A.; Petersen, J.; Maier-Hein, K.H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **2021**, *18*, 203–211.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.