

PRECISE-AS: Personalized Reinforcement Learning for Efficient Point-of-Care Echocardiography in Aortic Stenosis Diagnosis

Armin Saadat¹, Nima Hashemi¹, Hooman Vaseli¹, Michael Y. Tsang³, Christina Luong³, Michiel Van de Panne², Teresa S. M. Tsang³, and Purang Abolmaesumi¹

¹ Department of Electrical and Computer Engineering, The University of British Columbia, Vancouver, BC, Canada
{arminsdt, purang}@ece.ubc.ca

² Department of Computer Science, The University of British Columbia, Vancouver, BC, Canada

³ Vancouver General Hospital, Vancouver, BC, Canada **

Abstract. Aortic stenosis (AS) is a life-threatening condition caused by a narrowing of the aortic valve, leading to impaired blood flow. Despite its high prevalence, access to echocardiography (echo)—the gold-standard diagnostic tool—is often limited due to resource constraints, particularly in rural and underserved areas. Point-of-care ultrasound (POCUS) offers a more accessible alternative but is restricted by operator expertise and the challenge of selecting the most relevant imaging views. To address this, we propose a reinforcement learning (RL)-driven active video acquisition framework that dynamically selects each patient’s most informative echo videos. Unlike traditional methods that rely on a fixed set of videos, our approach continuously evaluates whether additional imaging is needed, optimizing both accuracy and efficiency. Tested on data from 2,572 patients, our method achieves 80.6% classification accuracy while using only 47% of the echo videos compared to a full acquisition. These results demonstrate the potential of active feature acquisition to enhance AS diagnosis, making echocardiographic assessments more efficient, scalable, and personalized. Our source code is available at: <https://github.com/Armin-Saadat/PRECISE-AS>.

Keywords: Efficient Echocardiography · Aortic Stenosis · Active Feature Acquisition · Reinforcement Learning · Point-of-Care Ultrasound

1 Introduction

Aortic stenosis (AS) [5] is a life-threatening heart valve disease marked by progressive leaflet thickening and calcification, restricting blood flow from the left ventricle. If untreated, severe AS has a five-year mortality rate of 67% [19]. Early detection is critical but remains inaccessible due to screening limitations.

** T. S.M. Tsang and P. Abolmaesumi are joint senior authors.

Echocardiography (echo) is the gold-standard for AS assessment, but its use is constrained by a shortage of trained personnel, leading to long wait times [16,18]. Point-of-care ultrasound (POCUS) offers a more accessible alternative by enabling non-specialists to perform focused cardiac imaging at the bedside [15]. However, POCUS exams lack comprehensive Doppler imaging and depend on operator expertise. Machine learning (ML) can bridge this gap by automating data acquisition and interpretation, assisting operators in capturing optimal echo views and determining the need for additional imaging-enhancing accuracy while reducing acquisition time.

Active feature acquisition (AFA) has been explored to optimize test-time data collection, with methods such as L_1 regularization enforcing sparsity at the population level and patient-specific techniques utilizing decision trees, mutual information-based selection, or Markov Decision Processes (MDP). However, AFA has been largely limited to low-dimensional data, with high-dimensional medical imaging applications restricted to pixel-level sampling in static images [2,3,4,7], leaving video-based dynamic acquisition unexplored.

We introduce PRECISE-AS, the first active video acquisition framework for echocardiography, leveraging reinforcement learning (RL) to dynamically select the most informative echo videos for each patient. Our method sequentially acquires data until diagnostic confidence is achieved, optimizing the trade-off between accuracy and acquisition time. Unlike static approaches that rely on predefined echo views, PRECISE-AS constructs personalized diagnostic pathways, ensuring patient-specific optimization of video acquisition.

Tested on a dataset of 2,572 patients, PRECISE-AS achieves state-of-the-art accuracy in AS detection while significantly reducing the number of required echo videos. These results demonstrate the potential of AFA in high-dimensional medical imaging, enabling cost-effective, scalable, and patient-tailored echocardiographic assessments at the point of care.

2 Background

Active Feature Acquisition. AFA frames the data acquisition problem as a discrete-time decision-making process [24] and operates under the premise that not all features are always available. The process begins with an empty feature set, sequentially acquiring additional features until it reaches a target prediction accuracy or the acquisition budget is exhausted. AFA seeks to balance prediction error against acquisition cost [12]. Approaches to AFA include greedy algorithms, which iteratively select the feature with the highest marginal information gain [2,8,26], and MDP-based methods that employ reinforcement learning to optimize the entire sequence of feature acquisitions rather than just the next immediate step [4,17,25].

Markov Decision Process. An MDP [14] is defined by (S, S_{end}, A, R, T) , where S is the set of states, $S_{end} \subset S$ the terminal states, and A the set of actions. The transition function $T : S \times A \times S \rightarrow \mathbb{R}$ gives the probability of

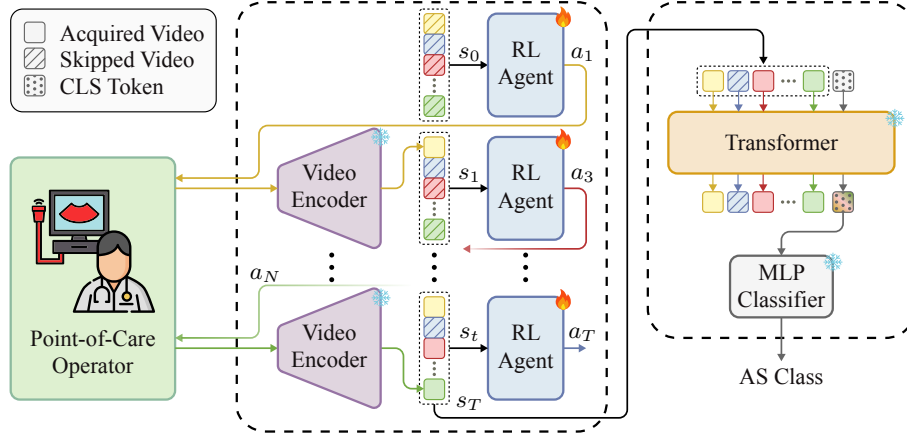


Fig. 1. Overview of the structure. Each action a_i acquires the i -th video and updates the state. At termination (a_T), the final state s_T is passed on for classification.

moving to a new state given the current state and action, while the reward function $R : S \times A \times S \rightarrow \mathbb{R}$ assigns rewards based on these transitions. An episode is a sequence of states, actions, and rewards from an initial state to a terminal state. A policy $\pi : S \rightarrow A$ maps states to actions, and the goal is to find the optimal policy π^* that maximizes the expected cumulative reward.

3 Method

PRECISE-AS (Fig. 1) consists of three key components. 1) A video encoder $f(\cdot)$ that maps an echo video $x_i \in \mathbb{R}^{H \times W \times 3}$ to feature $f(x_i) \in \mathbb{R}^D$, where H , W correspond to the height and width of the frames and D is the embedding size of the features. 2) An RL agent that iteratively selects encoded videos for each patient until it terminates the acquisition process. 3) A classifier $g(\cdot)$ that integrates the selected encoded videos and makes a diagnosis. The encoder and classifier are trained and frozen to be used for training the RL agent. The AS severity label is predicted as follows:

$$y^{\text{pred}} = \arg \max(g(z_1, z_2, \dots, z_N)), \quad \text{where } z_i = [m_i \cdot f(x_i)]. \quad (1)$$

Here, N is the maximum number of videos per study, m_i is a boolean scalar indicating whether x_i is selected, and \cdot denotes a dot product. The following sections describe each component in detail.

3.1 Feature Extraction

PRECISE-AS utilizes the open-source ProtoASNet [21] as its feature extractor model. ProtoASNet, currently the state-of-the-art for AS classification from

B-mode echo videos, is a prototypical network that classifies AS severity while remaining transparent in its decision-making. It learns spatio-temporal prototypes—representative clips highlighting calcification and restricted aortic valve leaflet motion—and compares new inputs against these reference points. By precisely locating clinically relevant features in each video frame, ProtoASNet offers an interpretable approach in which predictions hinge on a handful of evidence-based prototypes rather than opaque, large-scale parameters.

Class-wise similarity scores are computed by gauging how closely an input embedding aligns with each prototype, and a fully connected layer then weights these similarity scores to yield a probability for each AS severity category. This design ensures that the “why” behind each classification remains grounded in clinically significant patterns. ProtoASNet also incorporates aleatoric uncertainty by introducing prototypes that capture ambiguous or poor-visibility regions, flagging uncertain cases when necessary. Rather than using its final classification output, we specifically take the vectors feeding into the final fully connected layer as our feature representation for subsequent steps in our analysis. We train this feature extractor independently before the rest of the framework, then freeze its learned parameters to produce a fixed feature set for training the classifier.

3.2 Feature Selection

PRECISE-AS formulates feature selection as an MDP. Each state $s \in \mathbb{R}^{N \times D}$ represents the current subset of selected videos. In this state space, chosen videos retain their feature embeddings, while masked vectors replace unselected videos. The action space consists of $\{a_T, a_1, \dots, a_N\}$, where a_i selects video x_i and a_T terminates the acquisition process. The start state (s_0) is zero-initialized. As shown in Eq. (2), taking action a_i updates the i -th entry of the state to $f(x_i)$:

$$s_{t+1} = s_t.\text{copy}(), \quad s_{t+1}[i] = f(x_i). \quad (2)$$

Similar to [4], a sparse reward scheme is adopted. Non-terminal states receive a reward of 0, while terminal states receive the following reward:

$$R(s_T) = \mathbf{1}(y_{s_T}^{\text{pred}} = y^{\text{true}}) - \lambda \sum_{i \in s_T} c_i, \quad (3)$$

where s_T denotes a terminal state, and $\mathbf{1}(\cdot)$ is the indicator function that returns 1 if $y_{s_T}^{\text{pred}} = y^{\text{true}}$ (accuracy) or 0 otherwise. The term c_i is the cost of acquiring the i -th video in state s , and λ balances acquisition cost against classification accuracy. To obtain $y_{s_T}^{\text{pred}}$, we feed the terminal state into a frozen pre-trained classifier. The agent can take at most $N + 1$ actions, up to N for acquiring videos and one final termination action a_T . If the agent uses all $N + 1$ actions without invoking a_T , the episode automatically ends with zero reward.

PRECISE-AS uses Q-learning [23] to get the optimal policy and maximize the expected reward. Given the continuous, high-dimensional nature of our state space, we employ Double Deep Q-Networks (DDQN) [20], an enhancement of standard DQN designed to mitigate overestimation bias and stabilize training.

3.3 Feature Integration and Classification

The AS severity classifier has two objectives: (1) providing the reward signal during RL agent training, and (2) making predictions using the subset of videos selected by the agent at inference time. Consequently, the classifier must adapt to varying numbers of available videos and remain robust when some views are missing, while leveraging study-level aggregation to enhance performance beyond individual video-level predictions.

To accomplish this, we frame the problem similarly to natural language processing. Each echo video is treated as a “token,” and we maintain a maximum of N tokens to match the maximum number of echo videos per study. Missing tokens correspond to unselected videos. To ensure positional awareness, we use a consistent ordering of echo views and apply positional encodings, so the model understands which token corresponds to which view. This setup naturally aligns with Transformer [22] encoders, where each sequence of tokens (videos) is passed through the encoder along with a classification token (CLS) that aggregates information at the study level.

During training, we improve the classifier’s robustness to missing data by randomly masking 50% of the video tokens in each epoch. For any masked token (i.e., when $m_i = 0$), its features $f(x_i)$ are replaced with zeros, as shown in Eq. (1). We also apply attention masking to ensure that the CLS token ignores these masked tokens, compelling the model to rely on the remaining unmasked videos for classification. As a result, the classifier becomes flexible, delivering reliable predictions even with incomplete data. After training, the classifier’s parameters are fixed to be used in training the RL agent.

At inference time, the classifier processes a terminal state $s_T \in \mathbb{R}^{N \times D}$, interpreting it as a sequence of N tokens. Any videos the RL agent did not select are masked out (i.e., zeroed), and attention masking ensures that only the acquired videos contribute to the final classification. This design allows the system to produce accurate and efficient AS severity assessments tailored to each patient’s available echo data.

4 Experiments and Results

4.1 Datasets

We conducted our experiments on a private AS dataset sourced from an echo study database at a tertiary care hospital, following institutional review board approval. Echo videos were acquired using Philips iE33, Vivid i, and Vivid E9 systems. AS severity was classified by a Level III echocardiographer using standard Doppler guidelines [8], and only cases with consistent Doppler measurements were retained. A view-detection algorithm [13] automatically identified parasternal long-axis (PLAX) and short-axis (PSAX) cine clips, which were then re-screened by a Level III echocardiographer to remove misclassifications. The final dataset comprises 2,572 patient studies, including 5,055 PLAX and 4,062 PSAX cine clips. The data were split into training, validation, and test

Table 1. Quantitative metrics include balanced accuracy (bACC), weighted F1, and balanced mean absolute error (bMAE). bMAE is computed as the average MAE over classes (labels 0, 1, 2 for no, early, and significant AS). Study-level results are obtained by averaging video prediction probabilities per study, except for PRECISE-AS, which inherently integrates study-level information. Best results are in bold.

Method	Study-level (N=252)			Video-based	Acquired Videos
	bACC \uparrow	F1 \uparrow	bMAE \downarrow		
ProtoPNet [6]	70.9(4.7)	0.69(.07)	0.32(.05)	\times	100%
XProtoNet [11]	73.8(0.8)	0.74(.01)	0.29(.01)	\times	100%
ProtoASNet (Image) [21]	73.9(3.5)	0.74(.04)	0.29(.04)	\times	100%
Huang et al. [10]	74.7(1.6)	0.75(.02)	0.28(.02)	\times	100%
Ahmadi et al. [1]	76.9(1.7)	0.77(.02)	0.25(.02)	\checkmark	100%
XProtoNet (Video) [11]	77.2(1.4)	0.77(.01)	0.25(.02)	\checkmark	100%
Ginsberg et al. [9]	78.3(1.6)	0.78(.01)	0.24(.02)	\checkmark	100%
ProtoASNet [21]	80.0(1.1)	0.80(.01)	0.22(.01)	\checkmark	100%
ProtoASNet+RT4U [21]	80.1(1.4)	0.80(.01)	0.22(.01)	\checkmark	100%
PRECISE-AS (w/o RL)	80.6(0.5)	0.83(.01)	0.20(.01)	\checkmark	100%
PRECISE-AS (with RL)	80.6(0.8)	0.83(.01)	0.20(.01)	\checkmark	47%

sets (80-10-10), ensuring patient-level exclusivity. Each study was labeled with Doppler-derived severity levels, resulting in 1088 normal, 575 early, and 909 significant AS cases. For consistency, we randomly selected two PLAX and two PSAX videos per patient, resulting in four clips per study, ordered as [PLAX₁, PLAX₂, PSAX₁, PSAX₂]. This fixed ordering gives each echo view a consistent meaning. Each video is assigned a unit cost, as all clips are captured using the same imaging modality, making the total cost directly proportional to the number of videos acquired.

4.2 Implementation Details

For feature extraction, we train ProtoASNet according to the implementation details outlined in the original paper. We implement DDQN using two three-layer, fully connected networks for feature selection. The RL agent is trained for 50 epochs using MSE loss. We apply a discount factor of 1.0 to treat immediate and future rewards equally, as each additional echo video inherently carries its own cost. The best model is selected based on its validation accuracy. We utilize a Transformer with six encoder layers, eight attention heads, and a 256-dimensional feed-forward layer for classification. The network is trained using cross-entropy loss for 50 epochs. The framework is implemented in PyTorch, and models are trained on a single 16 GB NVIDIA Tesla V100 GPU.

4.3 Evaluations

Quantitative Assessment. Table 1 compares the performance of PRECISE-AS for classifying AS severity against both image-based and video-based base-

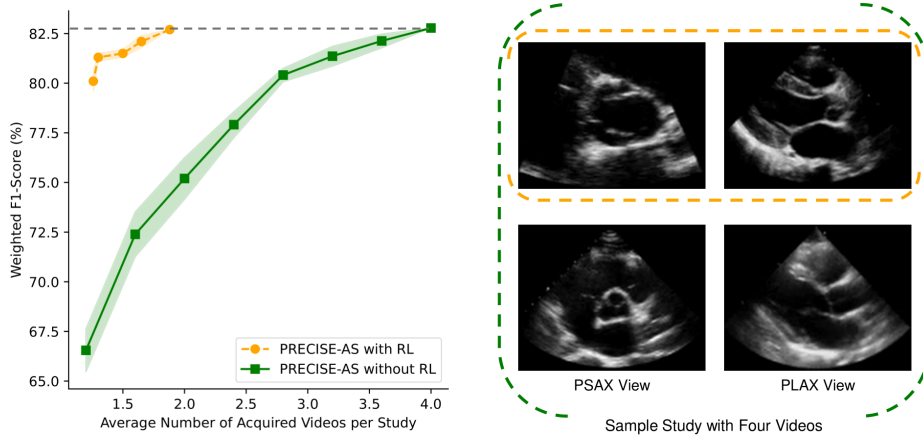


Fig. 2. (Left) F1 score versus the average number of acquired videos per study. Each yellow data point corresponds to a specific cost-coefficient. The solid lines represent the mean performance, while the shaded regions indicate one standard deviation. (Right) A sample study consisting of four echo videos. PRECISE-AS only requires the top two videos to achieve the same patient-level accuracy as the whole study.

lines. Video-based models outperform the image-based ones because they capture the dynamic motion of the aortic valve—a key indicator of AS severity. PRECISE-AS fully exploits this video-level information by employing the video-based ProtoASNet as its encoder. For the baselines, study-level predictions are obtained by averaging the class probabilities predicted at the image or video level. In contrast, PRECISE-AS intrinsically operates at the study level by using an attention mechanism to enrich a CLS token that encapsulates study-level information. This dual advantage explains why PRECISE-AS (without RL) outperforms the baselines when acquiring full studies. Moreover, incorporating RL to actively and efficiently select videos maintains performance while reducing the average number of acquired videos per study to 47% of a full acquisition—as indicated by PRECISE-AS (with RL). Fig. 2 provides a detailed comparison between with and without RL, showing that using RL for video selection as inputs to the classifier, consistently has a better performance-to-number-of-videos ratio, and ultimately obtains the best performance using fewer videos. This means that to get the best performance, PRECISE-AS w/o RL must select 4 videos per study on average, while PRECISE-AS with RL only needs around 2 videos, as shown in a sample study in Fig. 2. Although PRECISE-AS fails to show statistically significant improvement over ProtoASNet ($p > 0.05$), it matches in evaluation metrics while using less than 50% of the echo videos.

Qualitative Assessment. Active data acquisition offers critical insights into the decision pathways leading to a final diagnosis. These insights can inform AS diagnostic guidelines in healthcare by emphasizing the statistical significance of

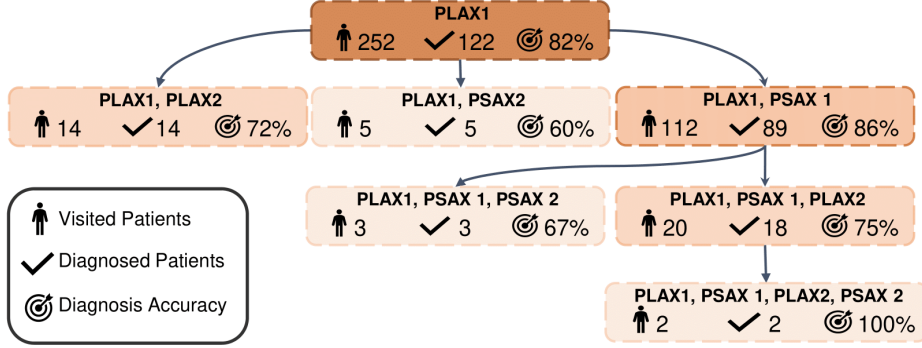


Fig. 3. Personalized diagnostic pathways for AS classification. Each node represents a state defined by the set of acquired videos and shows how many patients reached it. Directed edges illustrate the sequential video acquisitions leading to a final diagnosis.

Table 2. Ablation study of the acquisition cost coefficient (λ) demonstrating the trade-off between acquisition efficiency and classification performance.

Cost Coefficient	Study-level (N=252)			Acquired-Videos	
	bACC \uparrow	F1 \uparrow	bMAE \downarrow	Ratio \downarrow	Count \downarrow
w/o RL	80.6(0.5)	0.83(.01)	0.20(.01)	100%	4.00
0.001	80.6(0.8)	0.83(.01)	0.20(.01)	47%	1.88
0.01	79.9(0.8)	0.82(.01)	0.21(.01)	41%	1.64
0.1	79.0(0.6)	0.81(.01)	0.21(.01)	37%	1.50
0.2	78.5(0.7)	0.81(.01)	0.21(.01)	32%	1.30
0.25	77.1(0.6)	0.80(.01)	0.23(.02)	31%	1.26

key features and modalities. As illustrated in Fig. 3, the optimal starting point is to acquire a PLAX view, followed by a PSAX video, thereby enabling the classifier to examine the aortic valve from two complementary perspectives. We hypothesize that when two videos from the same view are selected, the first is suboptimal. Moreover, the agent adheres to implicit data acquisition rules as it terminates the process and avoids repetitive actions within an episode.

Ablation Study. The cost coefficient (λ) in Eq. 3 plays a crucial role in training the RL agent. Increasing λ causes the model to select fewer videos, which in turn reduces performance (see Table 2). If λ is set excessively high, the RL agent terminates the acquisition process immediately without selecting any videos.

5 Conclusion

PRECISE-AS represents a significant advancement in automated, patient-specific echocardiographic assessment, addressing critical barriers to efficient and accessible AS diagnosis. By leveraging reinforcement learning for active video ac-

quisition, our approach reduces unnecessary imaging while maintaining state-of-the-art diagnostic accuracy. This personalized, adaptive strategy streamlines POCUS workflows, enabling faster, more targeted AS evaluations—a crucial step toward improving early detection and timely intervention, particularly in resource-limited settings. By demonstrating the feasibility of active feature acquisition in high-dimensional medical imaging, PRECISE-AS lays the groundwork for scalable, intelligent echocardiography, ultimately enhancing clinical decision-making, operational efficiency, and patient outcomes.

Acknowledgments. This research was supported in part by the Canadian Institutes of Health Research (CIHR) and the Natural Sciences and Engineering Research Council of Canada (NSERC), and through computational resources and services provided by Advanced Research Computing at the University of British Columbia.

Disclosure of Interests. The authors have no competing interests to declare.

References

1. Ahmadi, N., Tsang, M.Y., Gu, A.N., Tsang, T.S.M., Abolmaesumi, P.: Transformer-based spatio-temporal analysis for classification of aortic stenosis severity from echocardiography cine series. *IEEE Transactions on Medical Imaging* **43**(1), 366–376 (2024). <https://doi.org/10.1109/TMI.2023.3305384>
2. Akrouf, M., Farahmand, A., Jarman, T.: Improving skin condition classification with a question answering model. *CoRR* **abs/1811.06165** (2018). <https://doi.org/10.48550/arXiv.1811.06165>
3. Akrouf, M., Farahmand, A., Jarman, T., Abid, L.: Improving skin condition classification with a visual symptom checker trained using reinforcement learning. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2019*. LNCS, vol. 11767, pp. 549–557. Springer (2019). https://doi.org/10.1007/978-3-030-32251-9_60
4. Bernardino, G., Jonsson, A., Loncaric, F., Martí Castellote, P., Sitges, M., Clarysse, P., Duchateau, N.: Reinforcement learning for active modality selection during diagnosis. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2022*. LNCS, vol. 13431, pp. 592–601. Springer (2022)
5. Carabello, B.A.: Introduction to aortic stenosis. *Circ. Res.* **113**(2), 179–185 (2013). <https://doi.org/10.1161/CIRCRESAHA.113.300156>
6. Chen, C., Li, O., Tao, C., Barnett, A.J., Rudin, C., Su, J.K.: This looks like that: deep learning for interpretable image recognition pp. 8928–8939 (2019). <https://doi.org/10.48550/arXiv.1806.10574>
7. Chen, Y., Hassani, S.H., Karbasi, A., Krause, A.: Sequential information maximization: when is greedy near-optimal? In: *Proceedings of the 28th Conference on Learning Theory (COLT 2015)*. *Proceedings of Machine Learning Research*, vol. 40, pp. 338–363. PMLR (2015), <https://proceedings.mlr.press/v40/Chen15b.html>
8. Covert, I.C., Qiu, W., Lu, M., Kim, N.Y., White, N.J., Lee, S.: Learning to maximize mutual information for dynamic feature selection. In: *Proceedings of the 40th International Conference on Machine Learning (ICML 2023)*. *Proceedings of Machine Learning Research*, vol. 202, pp. 6424–6447. PMLR (2023). <https://doi.org/10.48550/arXiv.2301.00557>

9. Ginsberg, T., Tal, R., Tsang, M.Y., Macdonald, C., Taheri Dezaki, F., van der Kuur, J., Luong, C., Abolmaesumi, P., Tsang, T.S.M.: Deep video networks for automatic assessment of aortic stenosis in echocardiography. In: *Simplifying Medical Ultrasound – ASMUS 2021*. LNCS, vol. 12967, pp. 202–210. Springer (2021)
10. Huang, Z., Long, G., Wessler, B., Hughes, M.C.: TMED 2: a dataset for semi-supervised classification of echocardiograms. In: *DataPerf: Benchmarking Data for Data-Centric AI Workshop at ICML 2022* (2022)
11. Kim, E., Kim, S., Seo, M., Yoon, S.: XProtoNet: diagnosis in chest radiography with global and local explanations pp. 15714–15723 (2021). <https://doi.org/10.1109/CVPR46437.2021.01546>
12. Li, Y., Shan, S., Liu, Q., Oliva, J.B.: Towards robust active feature acquisition. *CoRR* **abs/2107.04163** (2021). <https://doi.org/10.48550/arXiv.2107.04163>
13. Liao, Z., Girgis, H., Abdi, A., Vaseli, H., Hetherington, J., Rohling, R., Gin, K., Tsang, T., Abolmaesumi, P.: On modelling label uncertainty in deep neural networks: automatic estimation of intra-observer variability in 2D echocardiography quality assessment. *IEEE Transactions on Medical Imaging* **39**(6), 1868–1883 (2020). <https://doi.org/10.1109/TMI.2019.2959209>
14. Littman, M.L.: Markov decision processes. In: *International Encyclopedia of the Social & Behavioral Sciences*, pp. 9240–9242. Elsevier, Oxford (2001)
15. Luong, C.L., Ong, K., Kaila, K., Pellikka, P.A., Gin, K., Tsang, T.S.M.: Focused cardiac ultrasonography: current applications and future directions. *J. Ultrasound Med.* **38**(4), 865–876 (2019). <https://doi.org/10.1002/jum.14773>
16. Munt, B., O’Neill, B.J., Koilpillai, C., Gin, K., Jue, J., Honos, G.: Treating the right patient at the right time: access to echocardiography in canada. *Can. J. Cardiol.* **22**(12), 1029–1033 (2006). [https://doi.org/10.1016/S0828-282X\(06\)70318-X](https://doi.org/10.1016/S0828-282X(06)70318-X)
17. Muyama, L., Neuraz, A., Coulet, A.: Deep reinforcement learning for personalized diagnostic decision pathways using electronic health records: a comparative study on anemia and systemic lupus erythematosus. *Artif. Intell. Med.* **157**, 102994 (2024). <https://doi.org/10.1016/j.artmed.2024.102994>
18. Sanfilippo, A.J., Bewick, D., Chan, K.L., Cujec, B., Dumesnil, J.G., Honos, G., et al.: Guidelines for the provision of echocardiography in canada: recommendations of a joint canadian cardiovascular society/canadian society of echocardiography consensus panel. *Can. J. Cardiol.* **21**(9), 763–780 (2005), <https://pubmed.ncbi.nlm.nih.gov/16082436/>
19. Strange, G., Stewart, S., Celermajer, D., Prior, D., Scalia, G.M., Marwick, T., Ilton, M., Joseph, M., Codde, J., Playford, D.: Poor long-term survival in patients with moderate aortic stenosis. *J. Am. Coll. Cardiol.* **74**(15), 1851–1863 (2019). <https://doi.org/10.1016/j.jacc.2019.08.004>
20. Van Hasselt, H., Guez, A., Silver, D.: Deep reinforcement learning with double Q-learning. In: *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence (AAAI-16)*. pp. 2094–2100. AAAI Press (2016). <https://doi.org/10.48550/arXiv.1509.06461>
21. Vaseli, H., Gu, A.N., Ahmadi Amiri, S.N., Tsang, M.Y., Fung, A., Kondori, N., Saadat, A., Abolmaesumi, P., Tsang, T.S.M.: ProtoASNet: dynamic prototypes for inherently interpretable and uncertainty-aware aortic stenosis classification in echocardiography. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*. LNCS, vol. 14225, pp. 368–378. Springer (2023). https://doi.org/10.1007/978-3-031-43987-2_36
22. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need **30**, 5998–6008 (2017). <https://doi.org/10.48550/arXiv.1706.03762>

23. Watkins, C.J.C.H., Dayan, P.: Q-learning. *Mach. Learn.* **8**(3–4), 279–292 (1992). <https://doi.org/10.1007/BF00992698>
24. Yin, H., Li, Y., Pan, S.J., Zhang, C., Tschitschek, S.: Reinforcement learning with efficient active feature acquisition. In: *NeurIPS 2020 Workshop on Machine Learning for Health (ML4H)* (2020). <https://doi.org/10.48550/arXiv.2011.00825>
25. Yu, Z., Li, Y., Kim, J.C., Huang, K., Luo, Y., Wang, M.: Deep reinforcement learning for cost-effective medical diagnosis. In: *Proceedings of the Eleventh International Conference on Learning Representations (ICLR 2023)*. OpenReview (2023). <https://doi.org/10.48550/arXiv.2302.10261>
26. Zannone, S., Hernández-Lobato, J.M., Zhang, C., Palla, K.: Odin: optimal discovery of high-value information using model-based deep reinforcement learning. In: *ICML 2019 Workshop on Real-world Sequential Decision Making*. pp. 1–8 (2019), <https://realworld-sdm.github.io/paper/21.pdf>