Speckle-Aware Signal Extraction as an Alternative to Complex Methods for ECG-Free Cardiac Phase Detection

Nikhileswara Rao Sulake[1], Sai Manikanta Eswar Machara[1], Sivaji Retta[1], Iyyakutti Iyappan Ganapathi[2], Muhammad Owais[2], and Irfan Hussain[2*]

[1] Rajiv Gandhi University of Knowledge Technologies, Nuzvid, Andhra Pradesh, India
[2] Khalifa University Center for Autonomous Robotic Systems (KU-CARS), Khalifa University, UAE
Webpage: `https://nikhil-rao20.github.io/ecg-free-echo/`

Abstract. Accurate identification of end-diastole (ED) and end-systole (ES) in echocardiography is essential for cardiac assessment, but electrocardiographic synchronization is frequently unavailable in point-of-care and retrospective settings. A common hypothesis in ECG-free timing research is that incorporating optical flow, Hilbert envelope analysis, and multi-proxy averaging inherently improves accuracy over simple intensity-based methods. This study systematically tests this assumption on 1,000 CAMUS sequences and 10,030 EchoNet-Dynamic videos. We introduce a speckle-transparent proxy operating on the Gaussian-decomposed structural B-mode component and establish a formal random detector baseline ($\approx T/3$ frames). The high-complexity configuration achieves a median ED error of 9.0 frames on CAMUS, whereas both the minimal-adaptive and speckle-transparent configurations achieve 1.0 frame. Optical flow error doubles across speckle-SNR tertiles, identifying speckle as a key interference mechanism. These findings support physically motivated simplicity for ECG-free cardiac timing. The source code is available at `https://github.com/Nikhil-Rao20/ECG_Free_Echo`.

Keywords: Echocardiography · Cardiac phase detection · ECG-free timing · Speckle decomposition.

1 Introduction

In echocardiography, identifying the end-diastole (ED) and end-systole (ES) frames is essential to assess cardiac function and ejection fraction [13, 17]. Although an electrocardiogram (ECG) routinely provides these landmarks, it is frequently absent in modern workflows. The proliferation of point-of-care ultrasound, wearables, and retrospective databanks often precludes concurrent ECG acquisition [6, 12]. Consequently, there is a critical need for highly accurate ECG-free phase detection methods.

* Corresponding author: `irfan.hussain@ku.ac.ae`

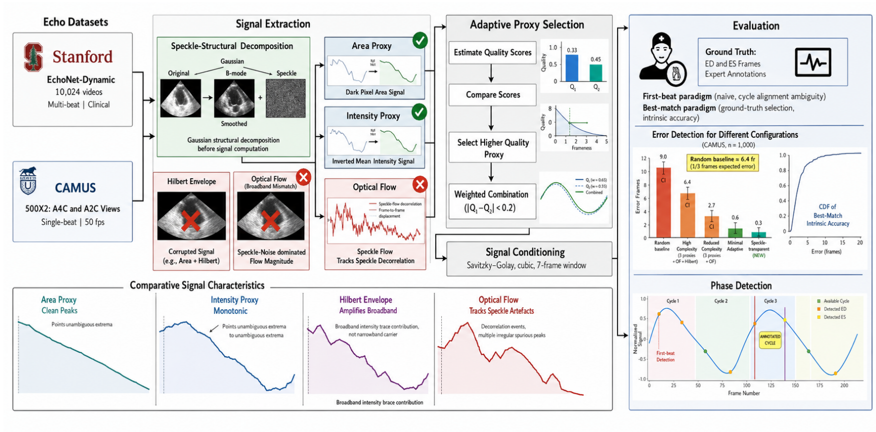

Fig. 1. Complexity bias in ECG-free cardiac timing. High-complexity multi-proxy architectures fail due to coherent speckle interference. Conversely, isolating the structural B-mode component to extract a simple, monotonic intensity proxy drastically reduces phase detection error across both CAMUS and EchoNet-Dynamic.

Current ECG-free techniques span a wide spectrum of algorithmic complexity. The foundational approaches utilize simple regional intensity tracking [10] or basic cavity area thresholding to map changes in ventricular volume. More sophisticated traditional architectures incorporate dense optical flow to track myocardial motion [8, 9], multi-proxy averaging (fusion) to consolidate distinct signal sources, and Hilbert envelope analysis to extract instantaneous amplitude from noisy oscillatory traces [3]. Currently, deep learning models, ranging from 3D convolutional networks and recurrent architectures to video transformers [1, 4, 5, 7, 11, 14, 20, 22, 25, 26, 28], achieve strong empirical performance but inherently function as "black boxes," offering limited interpretability regarding the specific physical features driving their predictions.

Previous work in echocardiographic motion analysis highlights that ultrasound speckle and image quality can strongly affect motion-estimation performance [8, 18]. This motivates the exploration of simpler, physically grounded proxies in ECG-free phase detection, rather than assuming that more elaborate optical flow or multi-proxy fusion schemes [2, 16] are inherently superior. Because B-mode images are coherent interference patterns [27], the characteristic speckle texture can decorrelate independently of the true anatomical tissue motion, challenging the brightness constancy constraint fundamental to many dense tracking algorithms. Furthermore, fusing multiple proxies extracted from the same speckle-corrupted image risks aggregating correlated noise. Consequently, the relative performance trade-offs between algorithmic complexity and physically grounded simplicity warrant systematic quantification.

This work directly investigates the complexity bias in the timing of cardiac arrest without ECG. We systematically evaluated six pre-specified signal extraction configurations against a formal random detector baseline in 1,000 single-beat CAMUS sequences [15] and 10,030 multi-beat EchoNet-Dynamic videos [20].

We find that physically motivated simplicity reliably outperforms algorithmic complexity. Fig. 1 summarizes the core observation of this work: the increased algorithmic complexity in ECG-free cardiac timing can fail under speckle interference, whereas structurally grounded intensity proxies remain robust. Our main contributions are as follows:

- **Complexity Bias & Speckle Failure:** We show that elaborate multi-proxy algorithms underperform simple intensity methods and formal random baselines, and provide direct evidence linking this failure to speckle decorrelation artifacts.
- **Speckle-Transparent Extraction:** We propose decomposing B-mode frames to isolate the structural component prior to proxy extraction. This parameter-free approach preserves the strong performance of the minimal-adaptive configuration, with median errors of 1.0 frame on CAMUS and 4.0 frames on EchoNet-Dynamic.
- **Robust Evaluation & Design Principles:** We introduce a first-beat/best-match paradigm to resolve multi-beat cycle ambiguity, culminating in eight rigorous design principles for ECG-free timing systems.

2 Materials and Methods

To systematically evaluate the impact of algorithmic complexity on ECG-free timing, we pre-specified six signal extraction configurations. We apply these pipelines uniformly to identical input data without configuration-specific parameter tuning, so that observed differences can be attributed to the extraction mechanisms rather than an unequal tuning effort. This controlled comparison is not intended to estimate the best attainable performance of each individual configuration after exhaustive method-specific optimization.

2.1 Datasets and Preprocessing

We used two public echocardiography datasets. **CAMUS** [15] provides 1,000 single-beat apical sequences (50 fps) with annotated ED and ES frames. **EchoNet-Dynamic** [20] provides 10,030 multi-beat apical 4-chamber videos (50 fps), with ground truth for one cycle per video. Both are standardized to grayscale arrays. Signal extraction operates within a static circular region of interest (ROI) Ω centered on the image, with radius $r = 0.25 \cdot \min(H, W)$. This generic scale (encompassing 50% of the central image dimension) avoids dataset-specific tuning while robustly capturing the left ventricular cavity across standard apical framings.

2.2 Speckle-Transparent Signal Extraction

The B-mode pixel intensity is a superposition of slowly-varying structural anatomical contrast and rapidly-varying coherent speckle interference [18, 27]. Standard

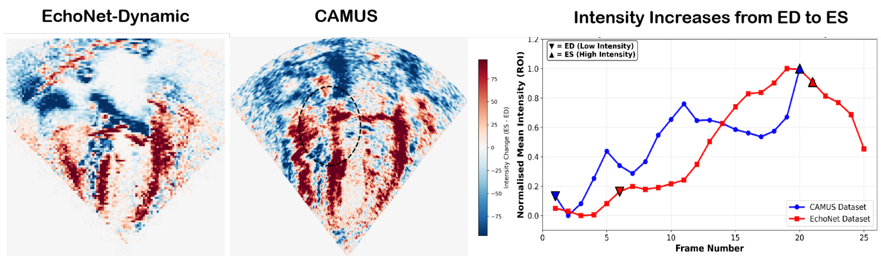

Fig. 2. Theoretical and empirical justification for structural decomposition. *Left:* Raw B-mode intensity combines structural tissue boundaries with complex speckle interference patterns. *Right:* Extracting the proxy strictly from the Gaussian-decomposed structural component (S_t) smooths local decorrelation artifacts, yielding a highly monotonic signal that aligns perfectly with the cardiac volume curve.

methods compute proxies directly on raw B-mode pixels, which can inadvertently incorporate high-frequency speckle dynamics into the extracted signal. When the myocardium moves through the ultrasound beam, speckle decorrelation can create anomalous peaks in optical flow and intensity variance that may not correspond to true physiological extrema.

To mitigate this, we introduce a *speckle-transparent* approach. Each frame $I_t(x, y)$ is first decomposed into a structural component $S_t(x, y) = (G_\sigma * I_t)(x, y)$ and a speckle residual using Gaussian low-pass filtering. The filter scale ($\sigma = 2.0$ pixels) is physically motivated to marginally exceed the typical spatial coherence length of the ultrasound speckle (1-3 pixels). It attenuates high-frequency interference artifacts while preserving macroscopic anatomical boundaries. In a local sensitivity check, varying σ from 1.0 to 3.0 pixels changed the median timing error by at most one frame on either dataset. The proxies are then computed exclusively on the structural component S_t. Fig. 2 shows mathematically and empirically how extracting signals from the structural component preserves the underlying physiological monotonicity while drastically attenuating speckle-induced noise.

The core timing signals utilized are the *intensity proxy* (mean brightness within Ω, decreasing as the anechoic blood pool expands during diastole) and the *area proxy* (proportion of pixels below the 30th brightness percentile). For comparison against high-complexity methods, we also evaluate the dense optical flow magnitude [9] and Hilbert envelopes [3].

We evaluated six configurations against the literature: **(1) High-complexity** [3, 8] (averaging intensity, area, flow, edge, and their Hilbert envelopes); **(2) Reduced-complexity** [16] (intensity, area, edge with SNR selection); **(3) Dual-proxy** [10] (intensity, area); **(4) Minimal adaptive** (quality-weighted intensity and area); **(5) Speckle-transparent** (structural intensity, structural area); and **(6) PASTA** (speckle-transparent with Bayesian quality scoring and adaptive bandpass filtering). Signals undergo a 3-frame median filter and 7-frame Savitzky-Golay smoothing [23]. Varying the Savitzky-Golay window from 5 to 11 frames changed the median timing error by at most one frame; these checks support robustness to the stated preprocessing choices while retaining the same settings for every configuration. ED and ES frames are identified via prominence-based peak

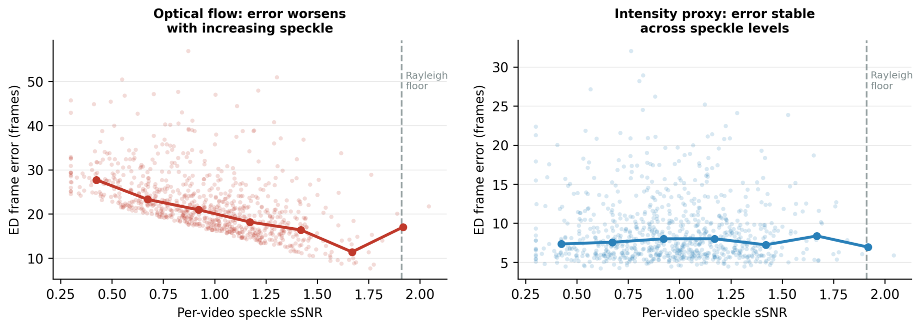

Fig. 3. Speckle signal-to-noise ratio (sSNR) versus ED timing error on EchoNet-Dynamic. *Left:* Optical flow proxy error doubles from low-speckle to high-speckle tertiles, confirming speckle interference as the dominant failure mechanism. *Right:* Intensity proxy error remains stable regardless of speckle severity.

detection. To filter non-physiological detections, we enforce systolic durations between 80 ms and 600 ms.

For EchoNet-Dynamic's multi-beat videos, naive evaluation conflates intrinsic timing accuracy with cycle-alignment ambiguity. We decouple these using two metrics: *First-beat* compares the first detected cycle to ground truth (reflecting real-world deployment), while *Best-match* reports the minimum error across all detected cycles (isolating intrinsic timing precision). Performance is contextualized against a formal random baseline, which evaluates to $\approx T/3$ frames (6.4 frames on CAMUS). All pairwise comparisons utilize the Wilcoxon signed-rank test ($\alpha = 0.01$, Bonferroni-corrected) and 95% bootstrap confidence intervals.

3 Results

3.1 Speckle Interference: The Quantitative Evidence

To establish the physical mechanism underlying the performance differences between configurations, we explicitly quantify speckle interference. For each EchoNet-Dynamic video, we compute the speckle signal-to-noise ratio (sSNR). The dataset mean is 0.98 ± 0.31, indicating substantial speckle contamination compared to the Rayleigh speckle floor of 1.91. By stratifying EchoNet-Dynamic into three equal-size sSNR tertiles (low, medium, high speckle), we isolate the effect of speckle on individual proxies. As shown in Fig. 3, the optical flow proxy's median ED error increases from ≈ 7 frames in the low-speckle tertile to ≈ 14 frames in the high-speckle tertile, a two-fold degradation. Conversely, the simple intensity proxy error remains stable across all tertiles. This differential response provides direct quantitative evidence that optical flow fails specifically because it is hypersensitive to speckle decorrelation, which standard brightness constancy models misinterpret as tissue motion.

Table 1. Median ED/ES errors (frames) and within-tolerance accuracy across CAMUS and EchoNet-Dynamic. CIs are 95% bootstrap intervals.

	CAMUS (n=1,000, tol $\leq$3)				EchoNet-Dynamic (n=10,030, tol $\leq$5)			
Configuration	**ED [CI]**	**% $\leq$3**	**ES [CI]**	**% $\leq$3**	**ED [CI]**	**% $\leq$5**	**ES [CI]**	**% $\leq$5**
Random baseline	6.4	≈30%	6.4	≈30%	–	–	–	–
High-complexity [3, 8]	9.0 [8–9]	20.0%	13.0 [11–14]	22.7%	6.0 [6–6]	44.8%	7.0 [7–8]	41.2%
Reduced-complex. [16]	8.0 [7–9]	31.0%	2.0 [1–3]	71.2%	6.0 [6–6]	46.0%	7.0 [6–7]	43.6%
Dual-proxy [10]	3.0 [3–3.5]	53.1%	0.0 [0–1]	75.6%	5.0 [4–5]	57.9%	5.0 [5–6]	55.7%
Minimal adaptive	1.0 [1–2]	65.1%	0.0 [0–1]	77.4%	4.0 [4–4]	67.4%	4.0 [4–5]	65.8%
Speckle-transp.	**1.0 [1–2]**	**65.4%**	**0.0 [0–1]**	**77.8%**	**4.0 [4–4]**	**67.8%**	**4.0 [4–5]**	**66.2%**

3.2 Phase Detection Performance

We systematically evaluated the six configurations against a formal random detector baseline, which yields an expected error of 6.4 frames on CAMUS. Table 1 summarizes the performance in both CAMUS (n=1,000) and EchoNet-Dynamic (n=10,030). In EchoNet-Dynamic's multi-beat sequences, we report the *best-match* error to isolate intrinsic timing accuracy from cycle-selection ambiguity.

The high-complexity configuration achieves a median ED error of 9.0 frames on CAMUS. In our tests, optical flow frequently introduced anomalous peaks driven by speckle decorrelation rather than tissue motion. Conversely, the simpler configurations are robust, i.e., both minimal adaptive and speckle-transparent extraction achieve median ED errors of 1.0 frame on CAMUS and 4.0 frames on EchoNet-Dynamic. Speckle-transparent extraction has modestly higher within-tolerance rates (65.4% versus 65.1% on CAMUS and 67.8% versus 67.4% on EchoNet-Dynamic). Given identical medians and overlapping bootstrap CIs, we interpret these configurations as having comparable timing accuracy under this protocol, rather than claiming a statistically meaningful advantage for the decomposition alone.

The ablation of adaptive components further illustrates the impact of simplicity (Table 2). The intensity proxy alone achieves a 2.0-frame median ED error (64.5% $\leq$ 3 frames). A fixed 50/50 average with area brings no significant shift, but quality-weighted selection (minimal adaptive) drops the error to 1.0 frame. The full speckle-transparent structural decomposition yields the best overall performance, demonstrating that targeted, physically motivated simplifications can improve phase detection accuracy over more complex multi-proxy fusion in these datasets. Fig. 4 visualizes the contrast in error distributions between the configurations, illustrating that errors in the high-complexity approach can be widespread, whereas speckle-transparent errors are tightly bound near zero.

3.3 Qualitative and Clinical Implications

Fig. 5 shows the root behavioral differences between the proxies on a representative sequence. The inverted intensity proxy exhibits a smooth and monotonic decrease from ED to ES, naturally producing unambiguous extrema. In contrast, the optical flow magnitude trace is littered with irregular, spurious peaks driven

Table 2. Ablation of adaptive configuration components on CAMUS (n=1,000). CIs are 95% bootstrap intervals.

Ablation variant	ED Error		ES Error	
	Med [CI]	≤3	Med [CI]	≤3
Intensity proxy only	2.0 [1–3]	64.5%	0.5 [0–1]	74.8%
Fixed 50/50 average	2.0 [1–3]	65.5%	1.0 [0–1]	75.0%
Minimal adaptive (full)	1.0 [1–2]	65.1%	0.0 [0–1]	77.4%
Speckle-transparent	**1.0 [1–2]**	**65.4%**	0.0 [0–1]	**77.8%**

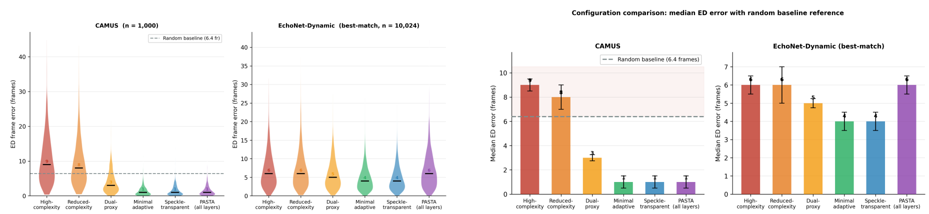

Fig. 4. *Left:* ED error distributions (CAMUS). High-complexity fails systematically (multimodal), while speckle-transparent concentrates near zero. *Right:* Median ED errors vs. random baseline (dashed line). Complex methods underperform random selection.

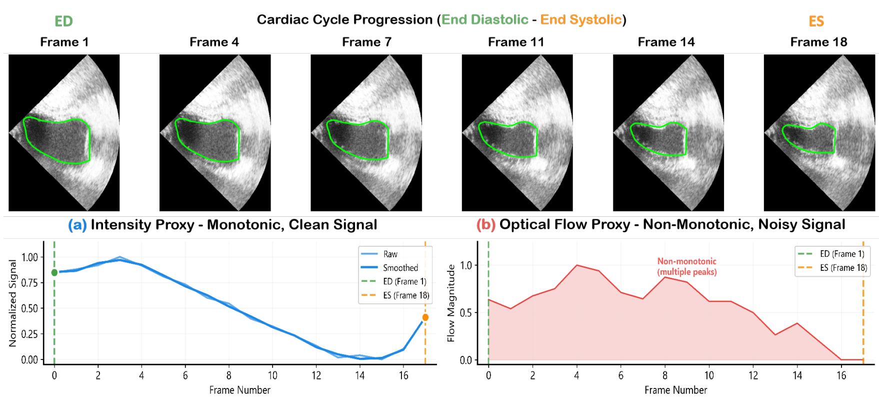

Fig. 5. Proxy signals from a representative sequence. **(a)** Inverted mean intensity exhibits a smooth monotonic decrease from ED to ES, producing unambiguous, easy-to-detect extrema. **(b)** Optical flow magnitude yields multiple irregular, spurious peaks driven entirely by speckle decorrelation, misaligning phase detection.

entirely by speckle decorrelation, fundamentally misaligning phase detection. Fig. 6 extends this comparison to fully processed high-complexity and minimal-adaptive proxy waveforms. Speckle-induced flow artifacts and Hilbert envelope distortions can heavily disrupt phase alignment in complex methods, whereas the minimal-adaptive intensity trace is clean and monotonic in these examples. Regarding multi-beat alignment on EchoNet-Dynamic, the first-beat deployment strategy evaluates to a median error of ≈ 34 frames. However, this large apparent error is driven entirely by cycle-selection ambiguity; errors cluster at exact

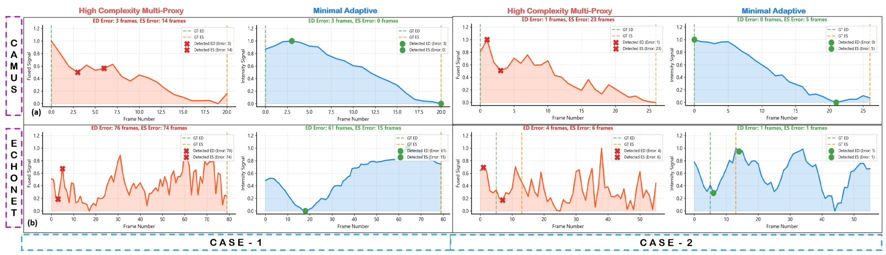

Fig. 6. Phase detection on representative sequences. **Case-1** illustrates a single-beat sequence from CAMUS, while **Case-2** illustrates a multi-beat sequence from EchoNet-Dynamic. *Left:* The multi-proxy approach is susceptible to anomalous peaks and misdetection. *Right:* The speckle-transparent approach yields clean waveforms and accurately aligns with ground truth extrema.

integer multiples of the cardiac cycle ($\approx$ 30 and $\approx$ 60 frames). The intrinsic detection mechanism remains accurate (4 frames best-match). Computationally, the speckle-transparent method runs at 0.19 s/sequence on a standard CPU, rendering it immediately deployable for large-scale analysis pipelines and point-of-care embedded systems without requiring GPU acceleration.

4 Discussion and Conclusion

Our evaluation in 11,030 sequences highlights an important consideration in ECG-free cardiac timing: algorithmic elaboration does not automatically yield better accuracy. In our experiments, multi-proxy configurations frequently underperformed simpler intensity-tracking methods. This behavior is deeply tied to the physics of B-mode ultrasound; under speckle decorrelation, optical flow, and related algorithms can track acoustic artifacts rather than true tissue displacement [8, 18], and averaging multiple proxies from the same image risks diluting the underlying signal with correlated noise. In contrast, structural decomposition explicitly mitigates speckle prior to signal extraction, providing a parameter-free, computationally lightweight approach (0.19 s/sequence) that reliably isolates anatomical dynamics.

Based on these findings, we propose core design principles for future ECG-free systems. Researchers must prioritize monotonic volumetric proxies (e.g., mean intensity) over differential velocity metrics and mitigate speckle structurally rather than temporally. Additionally, adaptive filtering must be strictly gated by signal quality (sSNR $\gtrsim$ 1.5) to prevent estimators from locking onto noise. Finally, in accordance with rigorous machine learning evaluation standards [19,24], systems must report best-match errors on multi-beat datasets to disentangle intrinsic tracking from cycle-selection ambiguity, and demonstrably outperform systematic random baselines. As wearable sensors and retrospective databanks increasingly rely on ECG-free timing to propel cardiovascular care into the era of big data [21], methods grounded in ultrasound physics provide a robust, interpretable foundation for clinical deployment.

Acknowledgment

This research was supported by the Center for Autonomous Robotic Systems (CARS), Khalifa University of Science and Technology (KU-CARS).

References

1. Ahn, S.S., Ta, K., Thorn, S., Langdon, J., Sinusas, A.J., Duncan, J.S.: Multi-frame attention network for left ventricle segmentation in 3d echocardiography. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. pp. 348–357. Springer International Publishing, Cham (2021)
2. Alessandrini, M., De Craene, M., Bernard, O., Giffard-Roisin, S., Allain, P., Waechter-Stehle, I., Weese, J., Saloux, E., Delingette, H., Sermesant, M., D'hooge, J.: A pipeline for the generation of realistic 3d synthetic echocardiographic sequences: Methodology and open-access database. IEEE Transactions on Medical Imaging **34**(7), 1436–1451 (2015). https://doi.org/10.1109/TMI.2015.2396632
3. Boashash, B.: Estimating and interpreting the instantaneous frequency of a signal. I. Fundamentals. Proceedings of the IEEE **80**(4), 520–538 (1992). https://doi.org/10.1109/5.135376
4. Bulbule, S., Soma, S.: Deep learning-based myocardial infarction detection using echocardiography. BMC Cardiovascular Disorders **26** (06 2026). https://doi.org/10.1186/s12872-026-06033-y
5. Christensen, M., Vukadinovic, M., Yuan, N., Ouyang, D.: Vision–language foundation model for echocardiogram interpretation. Nature Medicine **30** (04 2024). https://doi.org/10.1038/s41591-024-02959-y
6. Díaz-Gómez, J.L., Mayo, P.H., Koenig, S.J.: Point-of-care ultrasonography. The New England Journal of Medicine **385**(17), 1593–1602 (oct 2021). https://doi.org/10.1056/NEJMra1916062, https://www.nejm.org/doi/full/10.1056/NEJMra1916062
7. D'Alto, M., Naeije, R.: Deep learning workflow in echocardiography: John henry's saga revisited? CHEST **169**(1), 16–17 (2026). https://doi.org/https://doi.org/10.1016/j.chest.2025.07.4071, https://www.sciencedirect.com/science/article/pii/S0012369225049803
8. Evain, E., Sun, Y., Faraz, K., Garcia, D., Saloux, E., Gerber, B., De Craene, M., Bernard, O.: Motion estimation by deep learning in 2d echocardiography: Synthetic dataset and validation. IEEE Transactions on Medical Imaging **41**, 1911–1924 (02 2022). https://doi.org/10.1109/TMI.2022.3151606
9. Farnebäck, G.: Two-frame motion estimation based on polynomial expansion. In: Bigun, J., Gustavsson, T. (eds.) Image Analysis. pp. 363–370. Springer Berlin Heidelberg, Berlin, Heidelberg (2003)
10. Gifani, P., Behnam, H., Shalbaf, A., Sani, Z.: Noise reduction of echocardiography images using isomap algorithm. 2011 1st Middle East Conference on Biomedical Engineering, MECBME 2011 (02 2011). https://doi.org/10.1109/MECBME.2011.5752087
11. Holste, G., Oikonomou, E.K., et al.: Efficient deep learning-based automated diagnosis from echocardiography with contrastive self-supervised learning. Communications Medicine **4**, 116 (2024). https://doi.org/10.1038/s43856-024-00538-3, https://www.nature.com/articles/s43856-024-00538-3

12. Klassen, S.L., Dusingizimana, W., Ngoga, G., Kamali, I., Dusabeyezu, S., Ntaganda, E., Kwan, G.F.: Using point-of-care ultrasound in heart failure diagnosis and management in rural and resource-limited settings. CASE **6**(6), 259–262 (2022). https://doi.org/https://doi.org/10.1016/j.case.2022.04.012, https://www.sciencedirect.com/science/article/pii/S2468644122001153, special Issue on POCUS
13. Lang, R.M., Badano, L.P., Mor-Avi, V., Afilalo, J., Armstrong, A., Ernande, L., Flachskampf, F.A., Foster, E., Goldstein, S.A., Kuznetsova, T., Lancellotti, P., Muraru, D., Picard, M.H., Rietzschel, E.R., Rudski, L., Spencer, K.T., Tsang, W., Voigt, J.U.: Recommendations for cardiac chamber quantification by echocardiography in adults: An update from the american society of echocardiography and the european association of cardiovascular imaging. Journal of the American Society of Echocardiography **28**(1), 1–39.e14 (2015). https://doi.org/https://doi.org/10.1016/j.echo.2014.10.003, https://www.sciencedirect.com/science/article/pii/S0894731714007457
14. Leclerc, S., Smistad, E., Østvik, A., Cervenansky, F., Espinosa, F., Espeland, T., Berg, E.A.R., Belhamissi, M., Israilov, S., Grenier, T., Lartizien, C., Jodoin, P.M., Løvstakken, L., Bernard, O.: Lu-net: A multistage attention network to improve the robustness of segmentation of left ventricular structures in 2-d echocardiography. IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control **67**, 2519–2530 (2020), https://api.semanticscholar.org/CorpusID:225650415
15. Leclerc, S., Smistad, E., Pedrosa, J., Østvik, A., Cervenansky, F., Espinosa, F., Espeland, T., Berg, E.A.R., Jodoin, P.M., Grenier, T., Lartizien, C., D'hooge, J., Løvstakken, L., Bernard, O.: Deep learning for segmentation using an open large-scale dataset in 2d echocardiography. IEEE Transactions on Medical Imaging **38**, 2198–2210 (2019), https://api.semanticscholar.org/CorpusID:73510235
16. Ledesma-Carbayo, M., Kybic, J., Desco, M., Santos, A., Suhling, M., Hunziker, P., Unser, M.: Spatio-temporal nonrigid registration for ultrasound cardiac motion estimation. IEEE Transactions on Medical Imaging **24**(9), 1113–1126 (2005). https://doi.org/10.1109/TMI.2005.852050
17. Marwick, T.H.: Ejection fraction pros and cons: Jacc state-of-the-art review. Journal of the American College of Cardiology **72**(19), 2360–2379 (2018). https://doi.org/https://doi.org/10.1016/j.jacc.2018.08.2162, https://www.sciencedirect.com/science/article/pii/S0735109718383542
18. Meunier, J., Bertrand, M.: Ultrasonic texture motion analysis: theory and simulation. IEEE Transactions on Medical Imaging **14**(2), 293–300 (1995). https://doi.org/10.1109/42.387711
19. Nagendran, M., Chen, Y., Lovejoy, C.A., Gordon, A.C., Komorowski, M., Harvey, H., Topol, E.J., Ioannidis, J.P.A., Collins, G.S., Maruthappu, M.: Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ **368** (2020). https://doi.org/10.1136/bmj.m689, https://www.bmj.com/content/368/bmj.m689
20. 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., Zou, J.Y.: Video-based ai for beat-to-beat assessment of cardiac function. Nature **580**(7802), 252–256 (2020). https://doi.org/10.1038/s41586-020-2145-8
21. Quer, G., Arnaout, R., Henne, M., Arnaout, R.: Machine learning and the future of cardiovascular care: Jacc state-of-the-art review. Journal of the American College of Cardiology **77**(3), 300–

313 (2021). https://doi.org/https://doi.org/10.1016/j.jacc.2020.11.030, https://www.sciencedirect.com/science/article/pii/S0735109720378943
22. Reynaud, H., Vlontzos, A., Hou, B., Beqiri, A., Leeson, P., Kainz, B.: Ultrasound video transformers for cardiac ejection fraction estimation. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2021. pp. 495–505. Springer International Publishing, Cham (2021)
23. Savitzky, A., Golay, M.J.E.: Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry **36**(8), 1627–1639 (05 2002). https://doi.org/10.1021/ac60214a047, https://doi.org/10.1021/ac60214a047
24. Sengupta, P., Shrestha, S., Berthon, B., Messas, E., Donal, E., Tison, G., Min, J., D'Hooge, J., Voigt, J.U., Dudley, J., Verjans Md Phd Fesc Fracp, J., Khader, S., Johnson, K., Løvstakken, L., Tabassian, M., Piccirilli, M., Pernot, M., Yanamala, N., Duchateau, N., Arnaout, R.: Proposed requirements for cardiovascular imaging-related machine learning evaluation (prime): A checklist: Reviewed by the american college of cardiology healthcare innovation council. JACC. Cardiovascular imaging **13**, 2017–2035 (09 2020). https://doi.org/10.1016/j.jcmg.2020.07.015
25. Shobana, A., Malarkodi, K.P.: Deep learning for myocardial infarction detection using echocardiography: A comprehensive survey of segmentation and classification approaches. International Journal of Science and Research **15**(2), 1709–1717 (2026). https://doi.org/10.21275/SR26227110915
26. Soe, H.T., Iwata, H.: Deep learning-based echocardiography wall thickness analysis for heart failure detection. IEEE Access **14**, 36924–36932 (2026). https://doi.org/10.1109/ACCESS.2026.3670331
27. Wagner, R., Smith, S., Sandrik, J., Lopez, H.: Statistics of speckle in ultrasound b-scans. IEEE Transactions on Sonics and Ultrasonics **30**(3), 156–163 (1983). https://doi.org/10.1109/T-SU.1983.31404
28. Zeng, D., Li, M., Ding, Y., Xu, X., Xie, Q., Xu, R., Fei, H., Huang, M., Jian, Z., Shi, Y.: Segmentation with multiple acceptable annotations: A case study of myocardial segmentation in contrast echocardiography. In: Information Processing in Medical Imaging (2021), https://api.semanticscholar.org/CorpusID:235599693