Latent Motion Spectral Regularity for Congenital Heart Defect Screening in Fetal Echocardiography

Yingyu Yang[1], Qianye Yang[1], Can Peng[1], Bojana Salovic[2], Elena D'Alberti[2], Julene Carvalho[3,4], Olga Patey[2,3,4], Aris T. Papageorghiou[2], and J. Alison Noble[1]

[1] Institute of Biomedical Engineering, Department of Engineering Science, University of Oxford, Oxford, UK
yingyu.yang@eng.ox.ac.uk
[2] Nuffield Department of Women's and Reproductive Health, University of Oxford, Oxford, UK
[3] Brompton Centre for Fetal Cardiology, Royal Brompton and Harefield Hospitals, Guy's and St Thomas' NHS Foundation Trust, London, UK
[4] Fetal Medicine Unit, St George's University Hospitals NHS Foundation Trust, London, UK

Abstract. Congenital heart defects (CHD) are a leading cause of fetal morbidity and mortality. Early prenatal detection using ultrasound can inform delivery management and postnatal intervention planning. Automated CHD screening from fetal ultrasound can support clinicians, particularly in settings with limited specialist resources. In this work, we propose an anomaly detection framework for CHD screening from four-chamber view (4CV) fetal echocardiography videos, one of the most widely acquired fetal cardiac views. Our method learns a motion representation from normal videos, yielding a one-dimensional latent motion trajectory over time. In normal hearts, this trajectory evolves smoothly and quasi-periodically across the cardiac cycle. In CHD, deviations from normal dynamics introduce increased high-frequency fluctuations. We capture this phenomenon using a cycle-adaptive spectral regularity score, defined as the ratio of high-frequency power to total power within sliding windows spanning a fixed number of cardiac cycles and robustly aggregated over videos of arbitrary length. Evaluated on one in-centre test set and two external-centre cohorts, the proposed framework outperforms reconstruction-based and latent-distance anomaly baselines, demonstrating improved effectiveness and generalizability across sites and across a range of critical CHD conditions. On the in-centre test set, our method achieves performance comparable to clinical expert readers, highlighting its potential to support fetal CHD screening.

Keywords: Anomaly detection · Fetal echocardiography · Motion representation · Spectral analysis.

1 Introduction

Fetal congenital heart defect (CHD) comprises structural or functional abnormalities arising during pregnancy, affecting approximately 0.8% of live births worldwide [6] and remaining a leading cause of infant morbidity and mortality [15]. Ultrasound is the standard prenatal screening tool, with structured assessment of recommended cardiac planes during the second-trimester scan [4]. However, CHD detection in routine care remains variable, particularly outside specialist centres [3]. Automated CHD screening tools could be useful in settings with limited specialist resources [6, 1]. Among standard views, the four-chamber view (4CV) is widely acquired and clinically informative. Reliable automated analysis of 4CV echocardiography is a key component of scalable fetal CHD screening.

Most existing 4CV CHD detection pipelines operate on static images due to computational efficiency and simpler supervision. Image-based one-class approaches learn normality and flag deviations, including adversarial one-class models [7] and GAN variants using reconstruction- or attention-guided scores [5]. Supervised classification on curated (e.g., Doppler-enhanced) 4CV images can also achieve strong performance when labels are available [9]. However, frame-based decisions typically rely on selecting or sampling representative frames and do not explicitly model the periodic dynamics that encode functional regularity. Moreover, structural defects and functional irregularities are often interleaved: altered anatomy can induce abnormal contraction patterns, suggesting that motion provides complementary cues for detecting abnormality.

Video-based methods address this limitation by leveraging temporal information. SONO produces timeline-style anomaly indications from supervised fetal substructure detector, which is time-consuming and labor-intensive [8]. TVAE models cardiac periodicity by imposing a fixed periodic latent trajectory structure shared across videos and detects anomalies using reconstruction-based criteria and maximum a posteriori (MAP) scoring [12]. However, it was developed on infant 4CV data rather than fetal echocardiography, and its reconstruction-driven scoring primarily reflects appearance/anatomy deviations rather than irregular motion patterns, which may reduce robustness under fetal-specific variability and domain shifts. DISCOVR learns broadly transferable spatiotemporal representations from normal echocardiography videos via self-supervision and demonstrates strong downstream performance, including fetal CHD detection [10]. However, that method requires labeled CHD cases for operating-point selection, which is less practical when CHD cases are rare and heterogeneous. Overall, there remains a gap for normal-only, motion-aware fetal 4CV anomaly detection methods that generalize reliably in heterogeneous clinical settings.

In this work, we propose a motion-centric framework for fetal CHD screening from 4CV echocardiography videos. Since normal fetal cardiac videos are abundant whereas CHD cases are relatively rare and heterogeneous, we formulate screening as an anomaly detection problem that learns a model of normality from normal videos and flags deviations at test time. Our approach (i) adapts a latent motion trajectory model for unsupervised cardiac phase feature extrac-

tion [17, 18] to produce a 1D latent trajectory that tracks phase progression, and (ii) introduces a **cycle-adaptive spectral regularity score** that quantifies abnormality as elevated high-frequency content in this latent trajectory, capturing irregular/off-manifold dynamics. We evaluate on three independent datasets (one in-centre test set and two external test sets) spanning multiple hospitals and ultrasound vendors, and demonstrate improved effectiveness and cross-site generalizability compared with video anomaly detection baselines TVAE [12] and DISCOVR [10]. Our proposed method also shows performance approaching clinical experts on the in-centre test set.

2 Methods

The proposed framework consists of two stages. First, we adopt a latent motion trajectory (LMT) model trained with normal fetal 4CV videos to extract a 1D latent motion signal that summarizes cardiac phase dynamics . In contrast to prior LMT formulations using a 2D motion subspace [17, 18], we enforce a 1D trajectory to obtain a compact phase-related signal suitable for spectral scoring. (Fig. 1 (a-b)) Second, at inference phase, we score each test video using a spectral regularity measure computed from this 1D signal (Fig. 1 (c)).

2.1 Latent motion trajectory model (1D)

Given a video $\{X_t \in \mathbb{R}^{H\times W}\}_{t=1}^{T}$, an encoder $\mathcal{E}$ produces frame features $h_t = \mathcal{E}(X_t) \in \mathbb{R}^D$, with D the hidden dimension. Each frame is embedded as an anchor-plus-motion decomposition constrained to a shared 1D axis (Fig. 1(b)):

$$z_t = z^s + \alpha_t \mathbf{v}, z_t \in \mathbb{R}^D \tag{1}$$

where $z^s \in \mathbb{R}^D$ is a video-specific anchor, $\mathbf{v} \in \mathbb{R}^D, \|\mathbf{v}\|_2 = 1$ is a learnable global direction shared across videos, and $\alpha_t \in \mathbb{R}$ is the scalar motion coordinate forming the latent trajectory $\alpha_{1:T}$.

We parameterize z^s and α_t via lightweight MLPs (Fig. 1(a)):

$$z^s = g_s\left(\frac{1}{T}\sum_{t=1}^{T} h_t\right), \qquad \alpha_t = g_m(h_t). \tag{2}$$

2.2 Self-supervised decoder heads

With previously defined encoder and latent motion trajectory design, we train the autoencoder using one of two self-supervised decoder heads (both evaluated in our experiments, (Fig. 1(a))).

Reconstruction head (LMT-Rec). A decoder $\mathcal{D}_{\text{rec}}$ predicts reconstructed frames $\hat{X}_t = \mathcal{D}_{\text{rec}}(z_t) \in \mathbb{R}^{H\times W}$. We optimise frame-wise reconstruction and also decode z^s to reconstruct the mean appearance [17]:

$$\mathcal{L}_{\text{rec}}(X) = \frac{1}{T}\sum_{t=1}^{T} ||\mathcal{D}_{\text{rec}}(z^s) - X_t||_2^2 + \sum_{t=1}^{T} ||\mathcal{D}_{\text{rec}}(z_t) - X_t||_2^2 \tag{3}$$

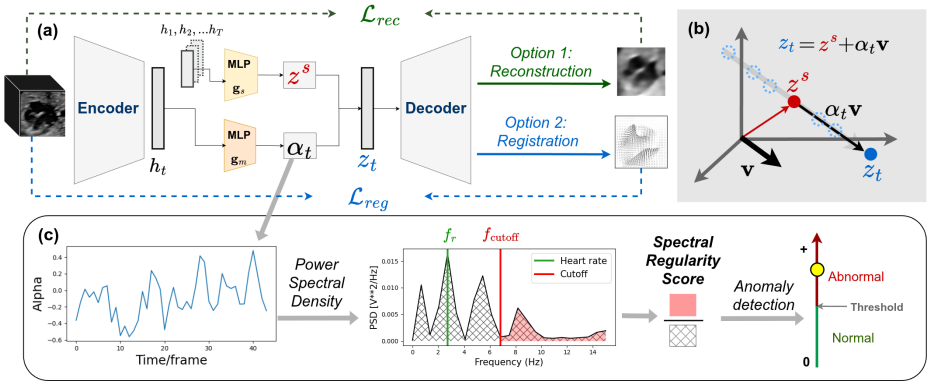

Fig. 1. Method overview. (a) We train an autoencoder that encodes a latent motion trajectory (LMT) on normal fetal 4CV videos using one of two self-supervised instantiations: a pixel-reconstruction model or a registration model that predict deformation fields. (b) At train time, the per-frame latent code is constrained to $z_t = z^s + \alpha_t \mathbf{v}$ (blue circles), where z^s is a video-specific anchor (red circle) and α_t is a scalar motion coordinate along a shared direction $\mathbf{v}$, forming a 1D latent trajectory $\alpha_{1:T}$. (c) At inference, we extract $\alpha_{1:T}$ from a test video and compute the spectral regularity score. Videos with scores above a threshold set on a normal validation set are flagged as anomalous.

Registration head (LMT-Reg). A decoder $\mathcal{D}_{\text{reg}}$ predicts a stationary velocity field (SVF) $\mathbf{u}_t \in \mathbb{R}^{2\times H\times W}$ per frame relative to an implicit reference [2, 18]. The deformation from frame i to j is approximated as $\Phi_{j\leftarrow i} = \exp(\mathbf{u}_j - \mathbf{u}_i)$ [13, 18]. We train by maximizing similarity between short-range temporal neighbors:

$$\mathcal{L}_{\text{reg}} = -\sum_{f=1}^{5}\sum_{t=1}^{T-f}\Big[\text{NCC}\big(X_t \circ \Phi_{t\leftarrow t+f}, X_{t+f}\big) + \text{NCC}\big(X_{t+f} \circ \Phi_{t+f\leftarrow t}, X_t\big)\Big], \quad (4)$$

where $\circ$ denotes warping and NCC is normalized cross-correlation.

2.3 Spectral regularity score

After training on normal videos, the LMT model outputs $\alpha_{1:T}$. α_t of normal cases is expected to be smooth and quasi-periodic over the cardiac cycle. CHD cases may deviate from normal dynamics, introducing additional high-frequency (HF) content. We therefore quantify spectral irregularity by the proportion of HF power in α_t (Fig. 1(c)).

We first estimate the dominant heart-rate frequency f_r from the global power spectral density (PSD), $f_r = \arg\max_{f>0} \text{PSD}(\alpha_{1:T}, f)$, where $\text{PSD}(\cdot)$ is computed using Welch's method [14]. Using f_r and the frame rate f_{FPS}, we define a cycle-adaptive window length L that spans two cardiac cycles $L = \text{round}(2\, f_{\text{FPS}}/f_r)$ to reduce spectral leakage. For each sliding segment $\alpha_{t:t+L-1}$

with stride 1 frame, we compute

$$s[t] = \frac{\sum_{f > f_{\text{cutoff}}} \text{PSD}(\alpha_{t:t+L-1}, f)}{\sum_f \text{PSD}(\alpha_{t:t+L-1}, f)},$$

We set the cutoff frequency f_{cutoff} to be $2.5f_r$ to treat the fundamental and first two harmonics as regular periodic content and penalize higher-frequency irregularities. The video-level spectral regularity score is $S = \text{median}\, s[t], t = 1, 2, ...T - L + 1$. Figure 2 shows detailed examples of spectral regularity score computation.

3 Experiments

Dataset We collected retrospective fetal 4CV videos acquired during the second trimester of gestation from three hospitals, John Radcliffe Hospital, Oxford University Hospitals NHS Foundation Trust, Oxford, UK (Hospital 1), St George's University Hospitals NHS Foundation Trust, London, UK (Hospital 2), and Royal Brompton and Harefield Hospitals, Guy's and St Thomas' NHS Foundation Trust, London, UK (Hospital 3). The dataset forms part of a retrospective cohort collected under the CAIFE study, which aims to develop artificial intelligence models for the early detection of congenital heart disease from ultrasound videos [11]. **Training/validation** (85%/15%) include 1662 normal 4CV videos from 994 participants at Hospital 1. The test set contains 534 videos from 366 participants across three independent cohorts: (1) **Set 1 (Hospital 1)**: 170 videos from 138 normal participants and 60 videos from 24 CHD participants. There is no participant overlap between training/validation and test. As with the training/validation data, all videos were acquired on GE Voluson machines. This forms the in-centre test set. (2) **Set 2 (Hospital 2)**: 109 videos from 69 normal participants and 95 videos from 67 CHD participants. Vendors include Fujifilm ARIETTA850 (70%), Canon TUS-AI800 (19%), GE Voluson (7%), and Samsung HERAW10 (4%). (3) **Set 3 (Hospital 3)**: 48 videos from 32 normal participants and 52 videos from 36 CHD participants. All videos from Set 3 were acquired on Fujifilm ARIETTA machines. We include in this study the CHD conditions that usually demonstrate lesions in 4CV, including coarctation of the aorta (COA), hypoplastic left heart syndrome (HLHS), dysplastic tricuspid valve or ebstein anomaly (DTV/EA), complete atrioventricular septal defect with unbalanced ventricles (CAVSD/U), pulmonary atresia with the intact interventricular septum (PA/IVS), tricuspid atresia with ventricular septal defects and normally related great arteries (TAVSD/NRA), mitral atresia with ventricular septal defect (MA/VSD), aortic stenosis (AS) and univentricular heart (UVH). The diagnosis of all studies were confirmed after birth.

Implementation For the training/validation set (Hospital 1 normal), the heart ROI was cropped using a public automatic cropping tool [16] and augmented with in-plane rotations (0°, 90°, 180°, 270°) during training to increase orientation variability. For the test sets, to reduce confounding probe/fetal motion and

standardize the evaluation across vendors, heart ROIs were manually cropped by an experienced engineer and verified by an experienced obstetrician (without using diagnostic labels). All test videos were resampled to 30 Hz using anti-aliasing resampling. This maintains sufficient temporal resolution for the fetal heart rate range and ensures consistent frame rate across vendors.

For LMT models (LMT-Rec and LMT-Reg), we temporally downsample by a factor of 2 for training and randomly sample 25 consecutive frames (spatially 128×128) from each cropped video. We follow the same LMT architecture as in [17], with the dimension of latent motion subspace set to 1. LMTs are optimized with Adam (learning rate 1×10^{-4}, batch size 16) for 500 epochs, and we select the checkpoint with the lowest validation loss. We compute video-level score according to Sec. 2.3.

Baselines TVAE [12] and DISCOVR [10] are trained from scratch on the same training set (resampled to 30 Hz to be consistent with test data) using their official implementations . For DISCOVR, we use a kNN-based (20 neighbours) anomaly score computed in the learned embedding space. For TVAE, we use the default MAP-based anomaly score with 100 iterations. We report the mean score of 5 randomly sampled clips as the video-level score.

Evaluation To evaluate anomaly detection performance, we aggregate video-level scores to a participant-level score by taking the maximum across each participant's videos. Two clinical experts (one junior obstetrician and one senior specialist) independently reviewed each test video (blinded to diagnosis) and provided a binary normal/abnormal decision per video. We then converted video-level expert decisions to participant-level labels using an OR rule: a participant is marked abnormal if any of their videos is judged abnormal. This aggregation reflects a screening setting where any suspicious clip for a participant should trigger referral. For all methods, the operating threshold is set to the 80th percentile of scores on the normal validation set. We use this percentile to tolerate occasional residual non-ideal clips (e.g., view mismatch or unexpected motion) present in the validation data.

4 Results

Overall performance across sites. Table 1 reports participant-level performance. Metrics are computed using 2,000 stratified bootstrap resamples over test participants within each site (mean $\pm$ standard deviation across bootstrap runs). On the in-centre test set (Set 1, Hospital 1), LMT-Reg achieves the best performance among all anomaly detection methods (AUROC 92.72 ± 3.45, F1 72.49 ± 5.16), substantially improving over TVAE (AUROC 78.36 ± 5.92) and DISCOVR (AUROC 74.21 ± 6.37). At the operating point, LMT-Reg matches the expert's sensitivity (91.67 vs 91.56/91.73) while markedly improving specificity (89.28 vs 73.24/67.43). Figure 3 (a) further shows the high agreement between merged expert opinion (OR rule, see Sec. 3) and LMT-Reg prediction with respect to different CHD subtypes.

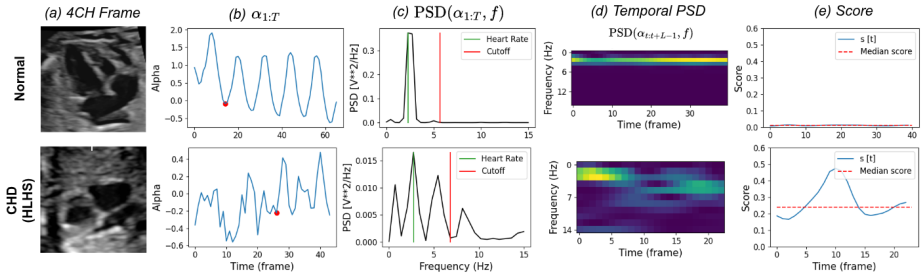

Fig. 2. Qualitative examples from Set 1. Top row: normal case. Bottom row: CHD (HLHS) case. (a) A representative frame from the input 4CV video. (b) Extracted 1D latent motion trajectory $\alpha_{1:T}$. The red marker indicates the time point shown in (a). (c) Global PSD of $\alpha_{1:T}$ (Welch) used to estimate the dominant cardiac frequency f_r and define the cutoff frequency f_{cutoff}. (d) Sliding-window PSD of α using a cycle-adaptive window length (stride 1 frame), visualized as a time–frequency map. (e) Windowed spectral regularity scores $s[t]$. The red dashed line denotes the final video-level score $S = \text{median}_t\, s[t]$. In this example, the CHD case yields a higher score than the normal case.

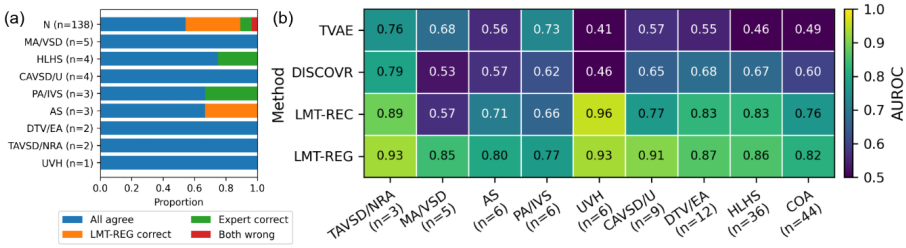

Fig. 3. Subtype analysis. (a) Subtype-wise composition of agreement between LMT-Reg predictions and the merged expert assessment (senior/junior) on site 1, showing the proportions of cases where both agree, only LMT-REG is correct, only the expert is correct, or both are wrong relative to the participant-level CHD ground truth. (b) AUROC for detecting each CHD subtype vs Normal (all test sets pooled). Numbers in brackets indicate participants per subtype (Normal: 239).

On the external test sets (Sets 2–3), all methods degrade under domain shift. Nevertheless, LMT-Reg remains the strongest deep learning method on both sites (Set 2 AUROC 75.56 ± 4.21; Set 3 AUROC 67.25 ± 6.78) and maintains high sensitivity across all test sets (Set 1 91.67 ± 5.55, Set 2 83.63 ± 4.47, Set 3 77.92 ± 6.93), which is desirable in a screening setting. Specificity decreases on the external test sets, likely due to vendor-dependent domain shifts. Because the operating threshold was fixed using the in-centre normal validation set acquired on GE Voluson machines, external normal videos from different vendors may receive higher anomaly scores, leading to more false positives. This suggests that future deployment may benefit from additional vendor-specific normal training data

Table 1. Participant-level evaluation on test sets. AUROC: Area Under the Receiver Operating Characteristic curve. BACC: Balanced Accuracy. Sen.: Sensitivity. Spe.: Specificity. Bold: best metric (excluding expert metrics). Yellow color: expert metrics.

Test	Method	AUROC	F1	Sen.	Spe.
1	*Expert (Senior)*	—	52.58 ± 6.54	91.56 ± 5.86	73.24 ± 3.74
	Expert (Junior)	—	47.99 ± 6.34	91.73 ± 5.79	67.43 ± 4.06
	TVAE	78.36 ± 5.92	43.06 ± 5.27	70.75 ± 9.19	72.46 ± 3.80
	DISCOVR	74.21 ± 6.37	36.68 ± 4.55	70.79 ± 9.26	62.49 ± 4.16
	Ours (LMT-Rec)	86.37 ± 4.61	65.47 ± 7.66	67.06 ± 9.76	$\mathbf{93.47} \pm 2.09$
	Ours (LMT-Reg)	$\mathbf{92.72} \pm 3.45$	$\mathbf{72.49} \pm 5.16$	$\mathbf{91.67} \pm 5.55$	89.23 ± 2.62
2	*Expert (Senior)*	—	88.24 ± 3.05	87.58 ± 4.22	89.90 ± 3.59
	Expert (Junior)	—	81.39 ± 3.78	79.21 ± 5.09	85.33 ± 4.23
	TVAE	60.38 ± 4.91	28.04 ± 6.03	19.38 ± 4.75	$\mathbf{82.53} \pm 4.62$
	DISCOVR	46.83 ± 4.89	58.38 ± 3.17	74.54 ± 5.33	21.67 ± 4.97
	Ours (LMT-Rec)	69.45 ± 4.56	67.99 ± 3.73	74.71 ± 5.20	56.30 ± 5.94
	Ours (LMT-Reg)	$\mathbf{75.56} \pm 4.21$	$\mathbf{72.77} \pm 3.14$	$\mathbf{83.63} \pm 4.47$	55.10 ± 5.89
3	*Expert (Senior)*	—	87.47 ± 4.68	78.03 ± 7.29	100.00 ± 0.00
	Expert (Junior)	—	67.21 ± 7.39	52.56 ± 8.48	96.78 ± 3.12
	TVAE	52.82 ± 7.22	19.03 ± 8.50	11.06 ± 5.43	$\mathbf{96.82} \pm 3.11$
	DISCOVR	50.45 ± 7.11	56.25 ± 5.61	61.06 ± 7.83	37.48 ± 8.65
	Ours (LMT-Rec)	45.81 ± 7.07	51.33 ± 5.77	55.77 ± 7.99	31.46 ± 8.15
	Ours (LMT-Reg)	$\mathbf{67.25} \pm 6.78$	$\mathbf{68.31} \pm 4.66$	$\mathbf{77.92} \pm 6.93$	43.65 ± 8.97

or target-site/vendor calibration. Finally, compared with LMT-Rec, LMT-Reg consistently provides better cross-site performance, suggesting that registration-based self-supervision promotes a representation that transfers more robustly across vendors than pixel reconstruction.

Figure 2 shows representative cases. Normal videos typically yield smooth, quasi-periodic latent trajectories $\alpha_{1:T}$, whereas CHD cases more often produce trajectories with increased temporal irregularity and higher-frequency fluctuations, resulting in elevated spectral regularity scores. Since the LMT model extracts frame-wise features learned from normal data (without imposing an explicit temporal model), increased high-frequency content in $\alpha_{1:T}$ can arise from genuinely abnormal motion and/or temporal ambiguity introduced by out-of-distribution appearance in CHD cases. In both cases, the loss of periodic regularity is captured by the proposed spectral score.

Subtype analysis. To examine detection variability across diagnoses, we perform subtype-versus-normal detection for each CHD subtype and summarize AUROC values in Fig. 3 (b). All test cases are pooled together. Across most subtypes, LMT-based methods achieve higher AUROC than TVAE and DISCOVR, and LMT-Reg is typically the best-performing variant. We note that several subtypes have small sample sizes (as low as $n = 3$ participants), so these

subtype-level results should be interpreted as exploratory. For CHD types with larger participant number, such as COA (n=44), HLHS (n=36) and DTV/EA (n=12), the proposed LMT-Reg demonstrates higher AUROC than other methods, demonstrating its potential for prenatal screening of critical CHD.

5 Conclusion

We presented an anomaly detection framework trained on normal fetal 4CV echocardiography videos for fetal CHD screening. Our approach builds on a latent motion trajectory model to extract a 1D phase-related latent trajectory of cardiac dynamics, and we introduce a spectral regularity score that measures abnormality as the proportion of high-frequency power in this latent motion signal. Across three independent test sets spanning multiple hospitals and ultrasound vendors, the proposed method, particularly the registration-based variant (LMT-Reg) consistently outperformed strong video baselines (TVAE and DISCOVR) and showed improved robustness under cross-site domain shift. Subtype-wise analysis further indicates broad screening potential, with AUROC exceeding 0.80 for 8 of 9 CHD conditions in our cohort. Notably, on the in-centre test set, LMT-Reg achieved performance comparable to clinical experts (a junior obstetrician and a senior fetal cardiologist). Overall, these results suggest that modeling motion regularity in a compact latent space provides an effective and generalizable cue for fetal CHD screening.

Acknowledgments. This work was supported by the InnoHK-funded Hong Kong Centre for Cerebro-cardiovascular Health Engineering (COCHE) Project 2.1 (Cardiovascular risks in early life and fetal echocardiography) and partly supported by the UKRI grant EP/X040186/1 (Turing AI Fellowship). Ethics approval was granted by the East Midlands, Leicester Central Research Ethics Committee (23/EM/0023). Co-authors J. Alison Noble and Aris Papageorghiou were supported by the Oxford Partnership Comprehensive Biomedical Research Centre with funding from the NIHR Biomedical Research Centre (BRC) funding scheme.

Disclosure of Interests. The authors have no competing interests.

References

1. Arnaout, R., Curran, L., Zhao, Y., Levine, J.C., Chinn, E., Moon-Grady, A.J.: An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease. Nature medicine **27**(5), 882–891 (2021)
2. Arsigny, V., Commowick, O., Pennec, X., Ayache, N.: A log-euclidean framework for statistics on diffeomorphisms. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 924–931. Springer (2006)
3. Bakker, M.K., Bergman, J.E., Krikov, S., Amar, E., Cocchi, G., Cragan, J., de Walle, H.E., Gatt, M., Groisman, B., Liu, S., et al.: Prenatal diagnosis and prevalence of critical congenital heart defects: an international retrospective cohort study. BMJ open **9**(7), e028139 (2019)

4. Carvalho, J., Axt-Fliedner, R., Chaoui, R., Copel, J., Cuneo, B., Goff, D., Gordin Kopylov, L., Hecher, K., Lee, W., Moon-Grady, A., et al.: Isuog practice guidelines (updated): fetal cardiac screening. Ultrasound Obstet Gynecol **61**(6), 788–803 (2023)
5. Chotzoglou, E., Day, T., Tan, J., Matthew, J., Lloyd, D., Razavi, R., Simpson, J., Kainz, B.: Learning normal appearance for fetal anomaly screening: Application to the unsupervised detection of hypoplastic left heart syndrome. Machine Learning for Biomedical Imaging **1**, 1–25 (2021). https://doi.org/https://doi.org/10.59275/j.melba.2021-g4dg, https://melba-journal.org/2021:012
6. D'Alberti, E., Patey, O., Smith, C., Šalović, B., Hernandez-Cruz, N., Noble, J.A., Papageorghiou, A.T.: Artificial intelligence-enabled prenatal ultrasound for the detection of fetal cardiac abnormalities: a systematic review and meta-analysis. eClinicalMedicine (2025)
7. Gong, Y., Zhang, Y., Zhu, H., Lv, J., Cheng, Q., Zhang, H., He, Y., Wang, S.: Fetal congenital heart disease echocardiogram screening based on dgacnn: adversarial one-class classification combined with video transfer learning. IEEE transactions on medical imaging **39**(4), 1206–1222 (2019)
8. Komatsu, M., Sakai, A., Komatsu, R., Matsuoka, R., Yasutomi, S., Shozu, K., Dozen, A., Machino, H., Hidaka, H., Arakaki, T., et al.: Detection of cardiac structural abnormalities in fetal ultrasound videos using deep learning. Applied Sciences **11**(1), 371 (2021)
9. Lei, W., Wen, C., Li, H., Yang, S., Shen, K., Yuan, H., Li, H., Xu, H., Gao, X., Zhang, S., et al.: An interpretable deep learning model for first-trimester fetal cardiac screening. npj Digital Medicine (2025)
10. Mishra, D., Salehi, M., Saha, P., Patey, O., Papageorghiou, A.T., Asano, Y.M., Noble, A.: Self-supervised learning of echocardiographic video representations via online cluster distillation. In: The Thirty-ninth Annual Conference on Neural Information Processing Systems (2025)
11. Patey, O., Hernandez-Cruz, N., D'Alberti, E., Salovic, B., Noble, J.A., Papageorghiou, A.T.: Prenatal detection of congenital heart defects using the deep learning-based image and video analysis: protocol for clinical artificial intelligence in fetal echocardiography (caife), an international multicentre multidisciplinary study. BMJ open **15**(6), e101263 (2025)
12. Ryser, A., Manduchi, L., Laumer, F., Michel, H., Wellmann, S., Vogt, J.E.: Anomaly detection in echocardiograms with dynamic variational trajectory models. In: Machine Learning for Healthcare Conference. pp. 425–458. PMLR (2022)
13. Vercauteren, T., Pennec, X., Perchant, A., Ayache, N.: Symmetric log-domain diffeomorphic registration: A demons-based approach. In: International conference on medical image computing and computer-assisted intervention. pp. 754–761. Springer (2008)
14. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S.J., Brett, M., Wilson, J., Millman, K.J., Mayorov, N., Nelson, A.R.J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, İ., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., van Mulbregt, P., SciPy 1.0 Contributors: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. Nature Methods **17**, 261–272 (2020). https://doi.org/10.1038/s41592-019-0686-2

15. Xu, J., Li, Q., Deng, L., Xiong, J., Cheng, Z., Ye, C.: Global, regional, and national epidemiology of congenital heart disease in children from 1990 to 2021. Frontiers in Cardiovascular Medicine **12**, 1522644 (2025)
16. Yang, Q., Cui, K., Hu, Y., Peng, C., Hernandez-Cruz, N., Ahuja, R., D'Alberti, E., Raymond, C., Patey, O., Papageorghiou, A., et al.: A deep learning framework for fetal heart tracking in ultrasound videos: Toward enhanced congenital heart defects detection. In: International Conference on Functional Imaging and Modeling of the Heart. pp. 219–230. Springer (2025)
17. Yang, Y., Yang, Q., Cui, K., Peng, C., D'Alberti, E., Hernandez-Cruz, N., Patey, O., Papageorghiou, A.T., Noble, J.A.: Latent motion profiling for annotation-free cardiac phase detection in adult and fetal echocardiography videos. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 316–325. Springer (2025)
18. Yang, Y., Yang, Q., Peng, C., D'Alberti, E., Patey, O., Papageorghiou, A.T., Noble, J.A.: Orientation-robust latent motion trajectory learning for annotation-free cardiac phase detection in fetal echocardiography (2026), https://arxiv.org/abs/2602.06761