CardiacNet: Learning to Reconstruct Abnormalities for Cardiac Disease Assessment from Echocardiogram Videos

Jiewen Yang[1], Yiqun Lin[1], Bin Pu[1], Jiarong Guo[1],
Xiaowei Xu[3✉], and Xiaomeng Li[1,2✉]

[1] The Hong Kong University of Science and Technology
{jyangcu, ylindw, jguoaz}@connect.ust.hk, {eebinpu, eexmli✉}@ust.hk
[2] HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, China
[3] Guangdong Provincial People's Hospital, Guangzhou, China
xiao.wei.xu@foxmail.com✉

Appendix A: Dataset Example

We provide four different echocardiogram videos with annotations that cover the normal case, the case with Atrial Septal Defect (ASD), and the case with Pulmonary Arterial Hypertension (PAH). This dataset has been stripped of all private information about the patients. Hospitals authorize this dataset and have ethical approval. The Fig. 1 shows four different examples of patients with ASD. The abnormality areas have been marked for easier understanding. Fig. 2 and Fig. 3 also present four different examples of PAH patients and normal cases, respectively. For all three figures, the vertical alphabet denotes different cases, while the horizontal axis is the frames that sample from every 10 frames in order. For complete video visualization, please refer to our attachment in the *supplementary.zip* for more dataset examples.

Appendix B: Algorithm Pipeline

Based on the **CardiacNet** that we have proposed in Section 3 of the manuscript, the overall pipeline of CardiacNet with the **C**onsistency **D**eformation **C**odebook (CDC) module and **C**onsistency **D**eformation **D**iscriminator (CDD) can be formulated as Algorithm 1.

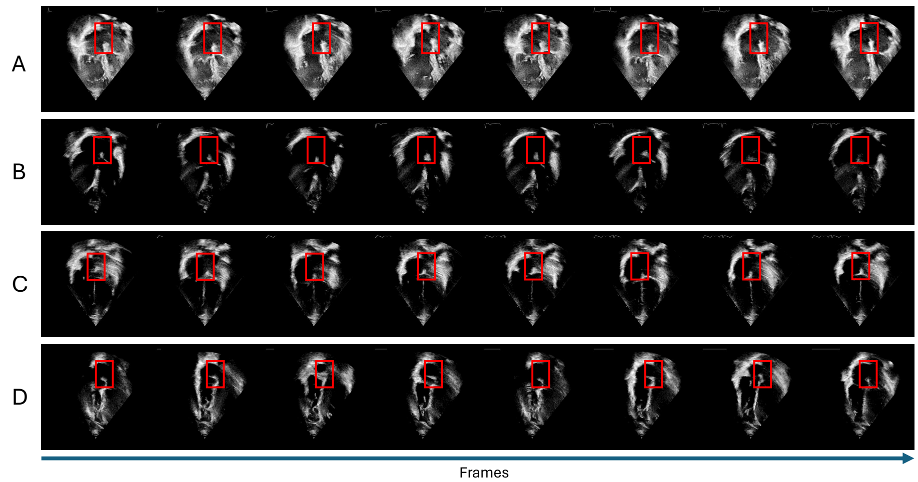

Fig. 1: Four examples of the ASD case.

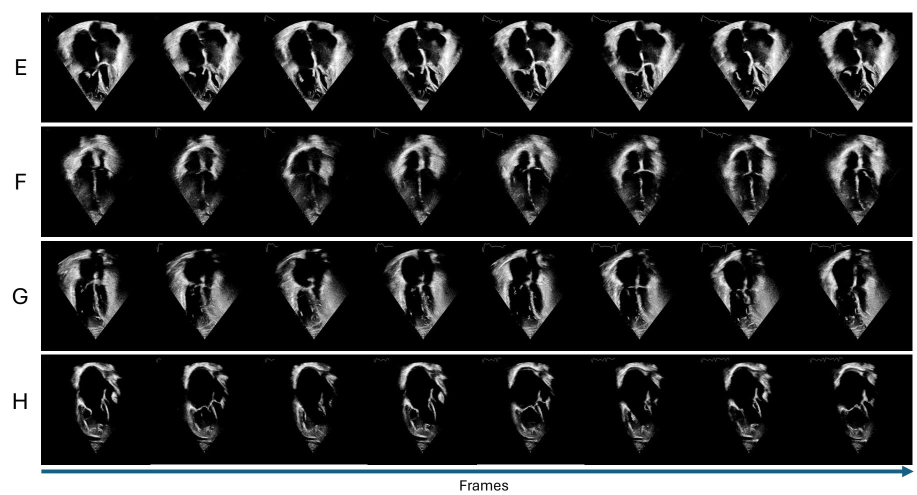

Fig. 2: Four examples of the PAH case.

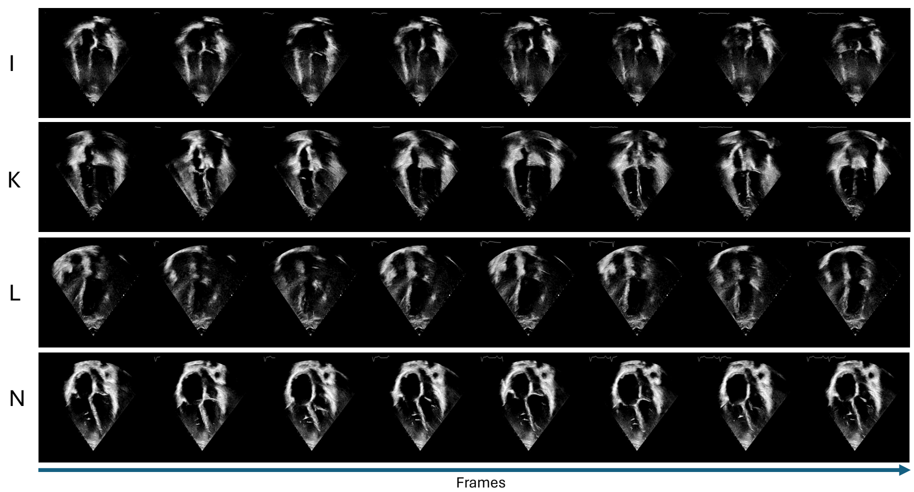

Fig. 3: Four examples of the Normal case.

Algorithm 1: The overall pipeline of CardiacNet

Output:

$\mathcal{L}_{\text{overall}}$: The overall loss of the CardiacNet;
$\mathcal{L}_{\text{CDC}}$: The loss from the consistency deformation codebook;
$\mathcal{L}_{\text{CDD}}$: The loss from the consistency deformation discriminator;

Input:

I: Input echocardiogram videos sample from normal and abnormal cases and $I \in \{X, Y\}$;
θ: To represent the reconstruction network for reconstructing abnormal results from normal cases and normal results from abnormal cases and $\theta \in \{A, B\}$;
$\phi(\cdot)$: The feature extractor and decoder for reconstructing cases between normal and abnormal;
$\xi(\cdot)$: The feature extractor of the $\phi(\cdot)$;
F: Feature map extracted by encoder $I(\cdot)$;
P: The positional encoding;
$\mathcal{M}$: The memory bank of data samples, random initialization;
$\mathcal{Z}$: The codebook of feature extractor $\phi(\cdot)$ for feature quantization.
$\sigma(F, \mathcal{Z}, P)$: The quantization operation, refer to Equation 2;
$\tilde{F}$: The quantized feature of F;
$\mathcal{L}_q(\xi(I), \tilde{F})$: The quantization loss of $\phi(\cdot)$, refer to Equation 3;

Convert input I to non-overlapping patches and perform random masking with a random ratio of 0.5 to 0.7,
$F_I \leftarrow \xi^\theta(I)$; Entering the I to feature extractor $\xi^\theta(\cdot)$ to get the feature F,
$\tilde{F}_I \leftarrow \sigma(F_I, \mathcal{Z}^\theta, P^\theta)$; the quantization of feature F_I,
$\mathcal{L}_q(\xi^\theta(I), \tilde{F}_I) \leftarrow$ Equation 3; Compute the quantization loss,
$\mathcal{Z}'_{\text{new}} \leftarrow (1 - \omega_1) \cdot \mathcal{Z} + \omega \cdot \mathcal{Z}_{\text{new}}$; Update the codebook via Equation 4, where ω is the weight for updating the current codebook that is set as 0.01,
Compute the reconstructed result $\phi^\theta(I)$ with the quantized feature $\tilde{F}$,
$L_{\text{CDD}} \leftarrow \mathcal{L}_{\text{adv}}(\phi^A(X), Y) + \mathcal{L}_{\text{adv}}(\phi^B(Y), X)$; Discriminate the reconstruction result and the real sample with Equation 9,
$I^R \leftarrow \phi^{A/B}(\phi^{B/A}(I))$; Reconstruct the $\phi^\theta(I)$ back to the original image as I^R,
$\mathcal{L}_{\text{recon}} \leftarrow ||I - I^R||_1$; Compute the reconstruction loss with L1 norm.

Optimization Between Network $\phi^A(\cdot)$ and $\phi^B(\cdot)$:
Update the codebook with the quantized feature $\tilde{F}$ for $\mathcal{M}$ of $\phi(\cdot)$,
$\mathcal{L}_{\text{OT}} \leftarrow$ Equation 5, compute the transport distance between the distribution of $\mathcal{M}^A$ and $\mathcal{M}^B$,
$\overline{\mathcal{M}} \leftarrow$ Compute the centroid representation $\overline{\mathcal{M}}$ from $\mathcal{M}$,
$\mathcal{L}_{\text{dis}} \leftarrow$ Compute the distance between the quantized feature and centroid $\overline{\mathcal{M}}$ via Equation 7,
$\mathcal{L}_{\text{CDC}} \leftarrow$ compute the overall loss of module CDC via the Equation 8.

Overall Loss:
$\mathcal{L}_{\text{overall}} = \mathcal{L}_{\text{CDC}} + \mathcal{L}_{\text{CDD}} + \mathcal{L}_{\text{recon}}$