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*]

The Hong Kong University of Science and Technology [1]
HKUST Shenzhen-Hong Kong Collaborative Innovation Research Institute, Futian, Shenzhen, China[2]
Guangdong Provincial People's Hospital, Guangzhou, China[3]

Code: https://github.com/xmed-lab/CardiacNet

Echocardiography for Cardiac Function Assessment

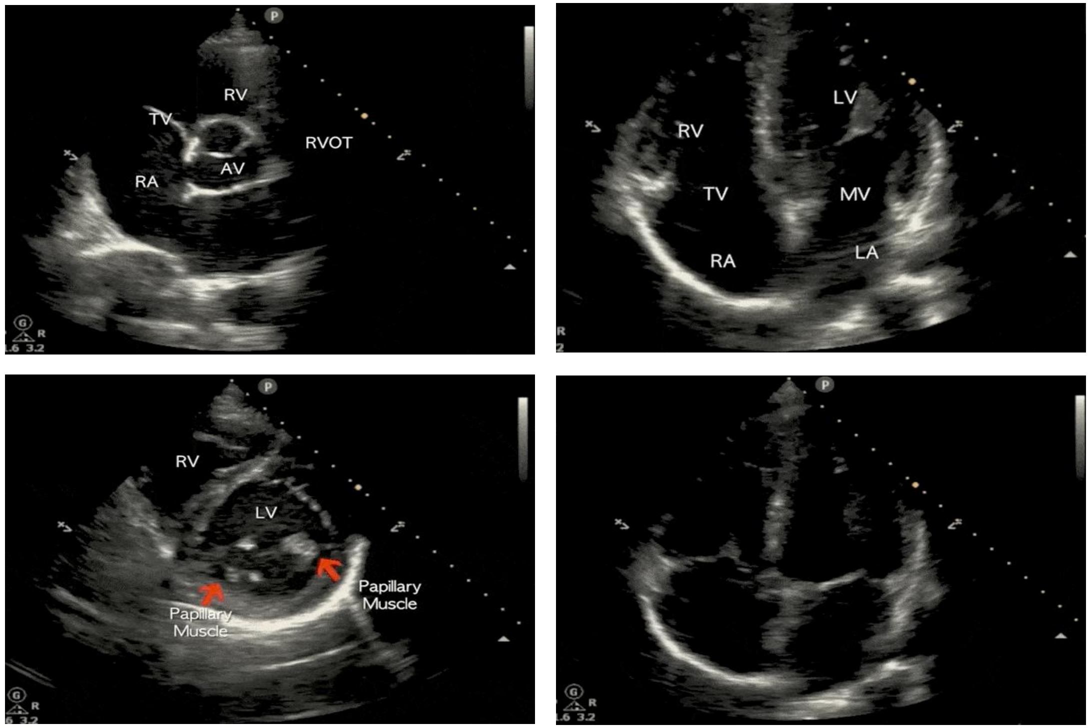


Figure 1 : Examples of Echocardiogram videos [1].

Echocardiography uses ultrasound waves to produce videos of the heart, which can scan cardiac structures and their motion such as valves, vessels, ventricle and atrium.

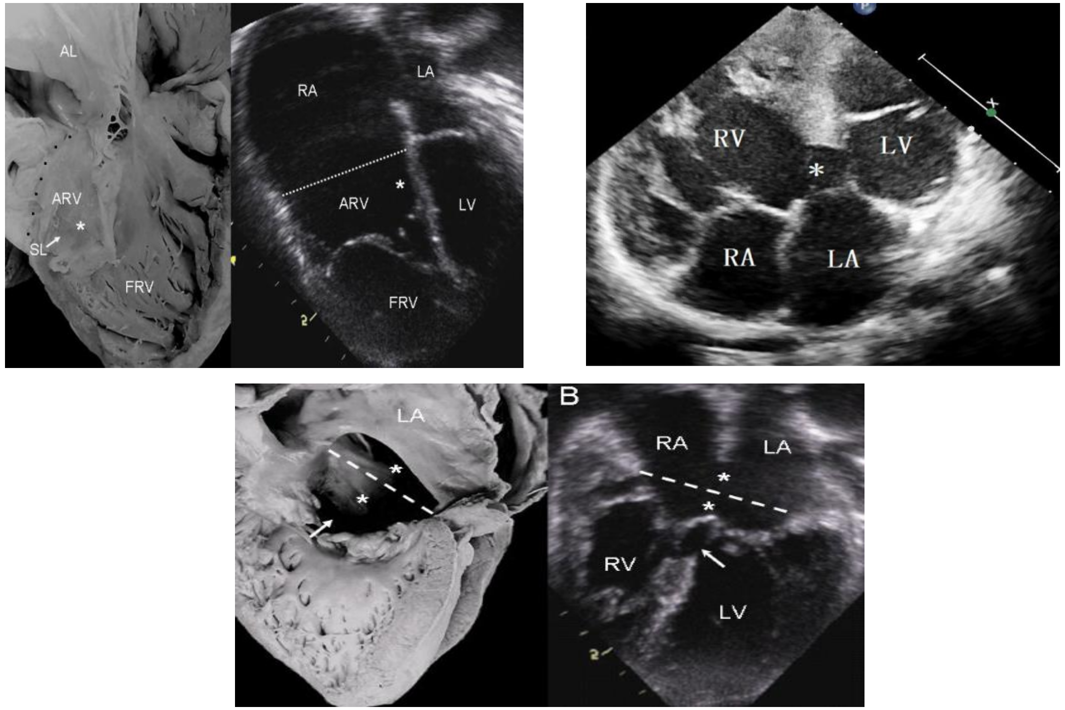


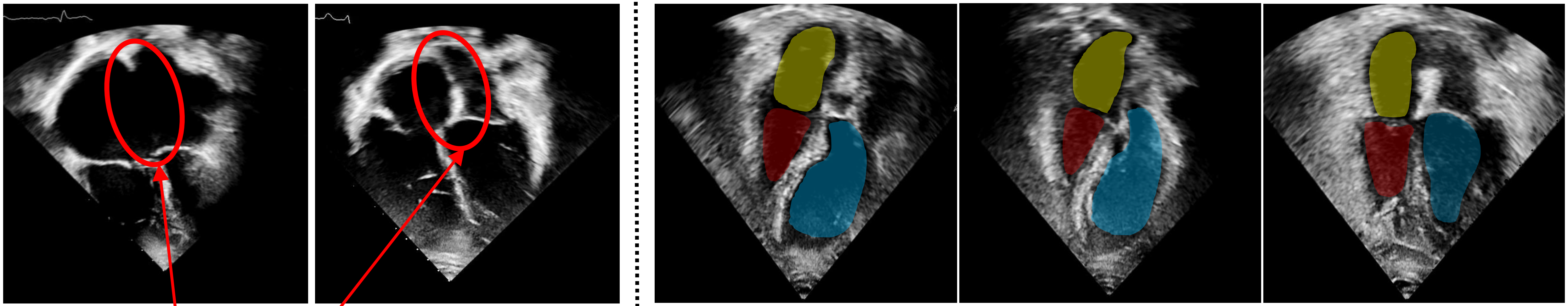
Figure 2 : Example of Ebstein's Anomaly [2], Double-outlet Right Ventricle [3], and complete atrioventricular septal defect [4] from left to right.

Echocardiography can help clinicians identify many cardiovascular diseases, such as congenital heart defects (CHDs).

[1] Cardiac Ultrasound (Echocardiography) Made Easy: Step-By-Step Guide, Vi Dinh, https://www.pocus101.com/cardiac-ultrasound-echocardiography-made-easy-step-by-step-guide/
[2] https://en.wikipedia.org/wiki/Ebstein%27s_anomaly [3] https://journals.sagepub.com/doi/10.1177/2150135117692973
[4] https://cardiovascularultrasound.biomedcentral.com/articles/10.1186/1476-7120-6-33

Challenges for Echocardiography

Structures are similar across normal and abnormal cases.



(a). Abnormal (ASD) (b). Normal – Case 1 (c). Abnormal (PAH) (d). Normal – case 2 (e). Normal – case 3

Figure 3 : Examples of Echocardiogram from two different cardiovascular diseases.

Structure Abnormality refers to cardiac diseases that exhibit clear and distinctive abnormalities in a localized region

Motion Abnormality refers to diseases may not have clear distinctive abnormalities in structures but can be detected through motion abnormalities of local cardiac structure.

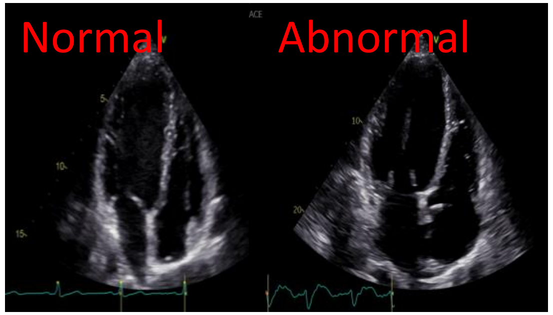


Figure 4 : Examples of Normal (Left) and Heart Failure (Right).

A Benchmark Dataset for Cardiac Disease Assessment

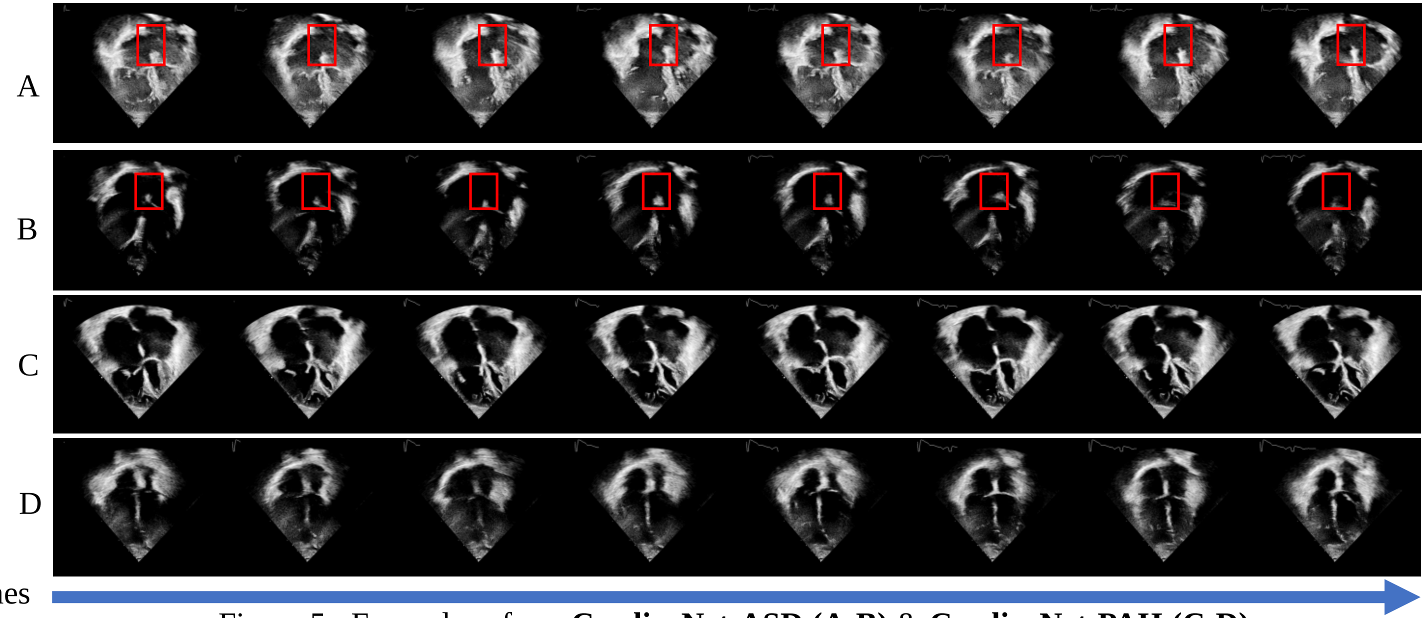


Figure 5 : Examples of our **CardiacNet-ASD (A-B)** & **CardiacNet-PAH (C-D)**.

Current Problems of public datasets

- Videos with **Low Resolution**;
- Without video with **Motion Abnormality**;
- No able to analysis **Structure Abnormality**;
- Mainly **focus on Ejection Fraction Evaluation**;

Advances of our **CardiacNet-ASD/PAH**

- Videos with **High Resolution and Quality**;
- Include **diverse arrays** of cardiac diseases;
- Specifically **for Cardiac Disease assessment**;
- Enable research in both **Structure and Motion Abnormality**;

Table 1: Summary statistics of datasets CardiacNet-PAH and CardiacNet-ASD and two public datasets CAMUS [5] and EchoNet [6].

Dataset	CardiacNet-PAH (Ours)						CardiacNet-ASD (Ours)					
Attri-butes	Total Videos	Total Images	PAH Cases	Normal Cases	Other Cases	Resol-ution	Total Videos	Total Images	ASD Cases	Normal Cases	Other Cases	Resol-ution
	496	44,363	342	154	0	720p	231	13,471	100	131	0	720p

Dataset	CAMUS [5]						EchoNet-Dynamic [6]					
Attri-butes	Total Videos	Total Images	EF≥55% Cases	EF≤50% Cases	50%<EF<55% Cases	Resol-ution	Total Videos	Total Images	EF≥55% Cases	EF≤50% Cases	50%<EF<55% Cases	Resol-ution
	500	10,000	201	178	121	480p	10,300	1,755,250	6961	2246	1093	120p

[5] Leclerc, S., et al.: Deep learning for seg-mentation using an open large-scale dataset in 2d echocardiography. IEEE TMI 38(9), 2198–2210 (2019)
[6] Ouyang, D., et al.: Video-based ai forbeat-to-beat assessment of cardiac function. Nature 580(7802), 252–256 (2020)

New Method for Cardiac Disease Assessment

Motivations

- Pervious studies focus on global information and **show difficulty in capturing local representations**.
- Existing approaches mainly design for CT, MRI and X-ray, which **rarely consider temporal information**.
- Most research **study on abnormalities with only structural details** such as tumors, bone fractures, and anomalous cardiac structures.

Our Solutions

- Formulate **global and local** cardiac structure information both **temporally and globally**.
- Incorporate the **prior cardiac knowledge** to gain a better understanding of the diseases in terms of their **local structural details** and **motion changes**.
- Enabling the **visualization of abnormalities** in the absence of abnormalities annotation.

Overall Pipeline of Our CardiacNet

Bidirectional Reconstruction Design

Simulates the deformation process from "normal" to "abnormal" cases and the reverse process.

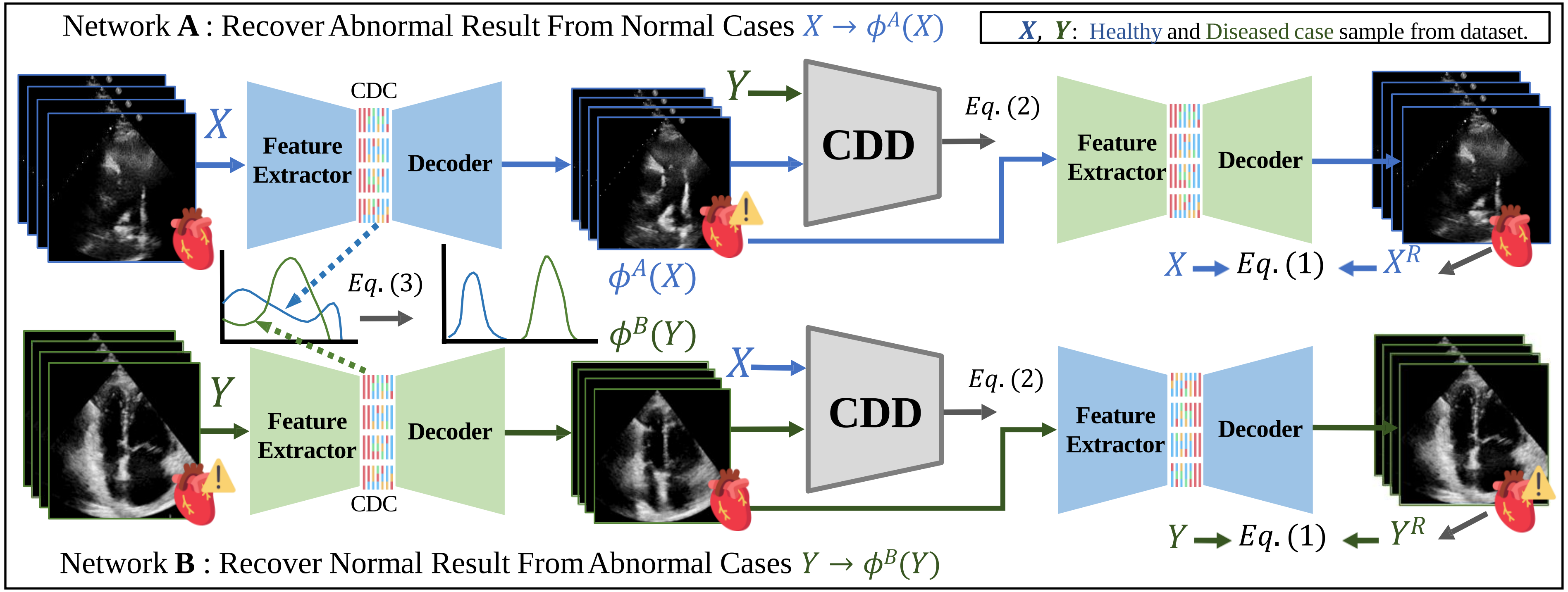


Figure 6 : The overview of our CardiacNet sample normal case X and abnormal case Y, reconstruct the corresponding abnormal and normal results.

Module 1 : Consistency Deformation Codebook (CDC)

Learn the local structural abnormalities and motion changes associated with the diseases.

Module 2 : Consistency Deformation Discriminator (CDD)

Improve the quality of reconstructed videos and maintaining spatiotemporal consistency with the real videos.

$Eq.(1)$: $L1$ Reconstruction Loss;
$Eq.(2)$: Global and Local Discrimination;
$Eq.(3)$: Distribution Optimization;
X, Y : Normal / Abnormal inputs;
X^R, Y^R : Reconstruction Results;
$\phi^{A/B}$ (X/Y) : Network with same structure;

One-way Process of Bidirectional Network (CDC)

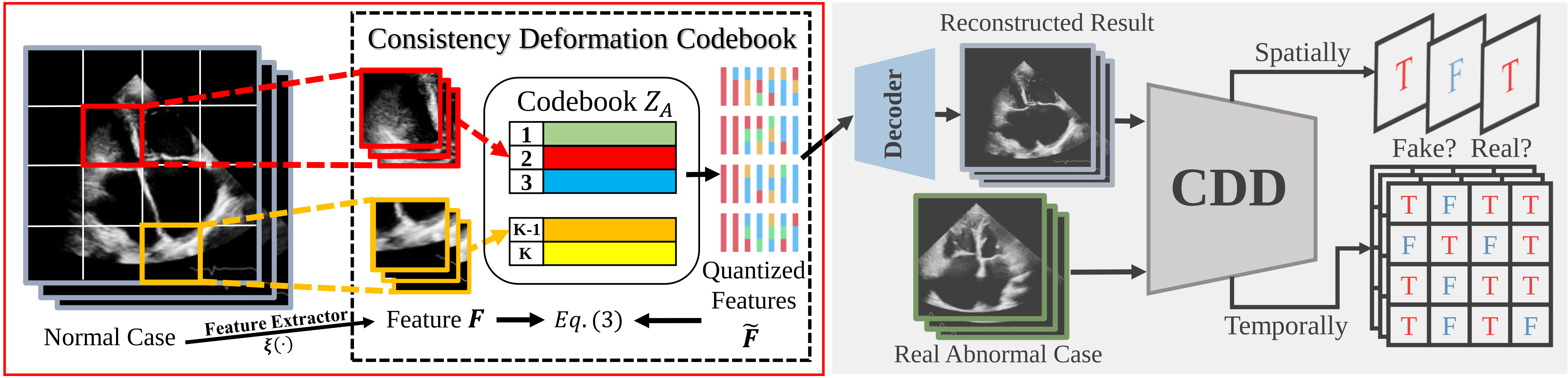


Figure 7 : The overview of our CardiacNet sample normal case X and abnormal case Y, reconstruct the corresponding abnormal and normal results.

Motivation : The network understands the representation of a specific disease that can also reconstruct normal from abnormal or its reverse process → CDC constructs the regional representation via Codebook ↓ Update Codebook with $Eq.(3)$ During Training

Reconstructed Feature | Extracted Feature | Codebook Entries | Positional Encoding

Step 1 : $\boxed{\tilde{F}} = \sigma(\boxed{F}, \boxed{\mathcal{Z}}, \boxed{\mathcal{P}}) := \left(\arg\min_{Z_k \in \mathcal{Z}} \left\| (F_{n,i,j} + P_n) - Z_k \right\|_2^2 \right)_{n,i,j}$

Step 2 : $\mathcal{L}_q(\boxed{\xi}(I), \tilde{F}) = \left\| \boxed{sg}[\xi(\boxed{I})] - \tilde{F} \right\|_2^2 + \lambda \cdot \left\| sg[\tilde{F}] - \xi(I) \right\|_2^2,$

Feature Extractor | Stop-Gradient | Input

One-way Process of Bidirectional Network (CDD)

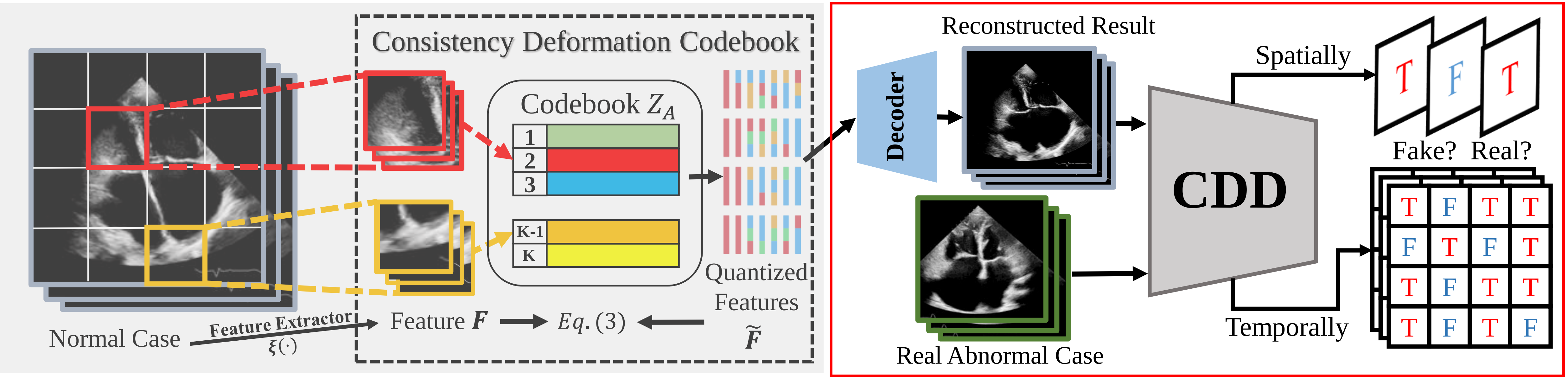


Figure 7 : The overview of our CardiacNet sample normal case X and abnormal case Y, reconstruct the corresponding abnormal and normal results.

Motivation : The discriminator acts as an adversary that forces the reconstructed results to **conform with the real data in semantic properties**. → Both Locally (each region) and Globally (whole video), we guarantee reconstructed with high-quality and remain consistent with real case → Using one-way process as an example

Spatial Consistency Discriminates

Spatial Consistency Discriminates

$$\mathcal{L}_{\text{adv}}(\phi^A(X), Y) = \left(\log(\boxed{\eta^T}(\phi^A(X))) + \log(1 - \eta^T(Y))\right) + \sum_{n=1}^{t}\left[\log(1 - \eta^S(\hat{X}_n)) + \log(\boxed{\eta^S}(\hat{Y}_n))\right] + \sum_{i=1,j=1}^{h,w}\left[\log(1 - \eta^T(\boxed{\hat{X}_{i,j}})) + \log(\eta^T(\boxed{\hat{Y}_{i,j}}))\right].$$

Convert X to Non-overlap Patches $\hat{X}$

Convert Y to Non-overlap Patches $\hat{Y}$

Consistency Deformation Codebook

Motivation : Expect to **maximize the distance of deformations** between the normal and abnormal sets.

Problem : Entries of codebooks being irrelevant and redundant, an entry in the same position of different codebooks is **non-matching and non-equivalent**.

Solution : Utilize memory banks to store features and **approximates distribution** of sets iteratively. Introduce the **transport distance optimization** to distinguish the distribution of normal and abnormal sets.

$Eq.(4)$: $\mathcal{L}_{\text{dis}}\left(\tilde{F}, \overline{\mathcal{M}}\right) = \left\| \tilde{F} - \boxed{\overline{\mathcal{M}}} \right\|_2^2$ Centroid of Memory Bank

Memory Bank

Number of Feature dimension

Number of samples in memory bank

$Eq.(5)$: $\mathcal{L}_{\text{OT}}\left(\boxed{\mathcal{M}}^A, \mathcal{M}^B\right) = \sum_{i=1}^{\boxed{d}} \sum_{j=1}^{\boxed{J}} \left\| \mathcal{M}^A_{j,i} - \mathcal{M}^B_{\pi^i(j),i} \right\|_2^2$

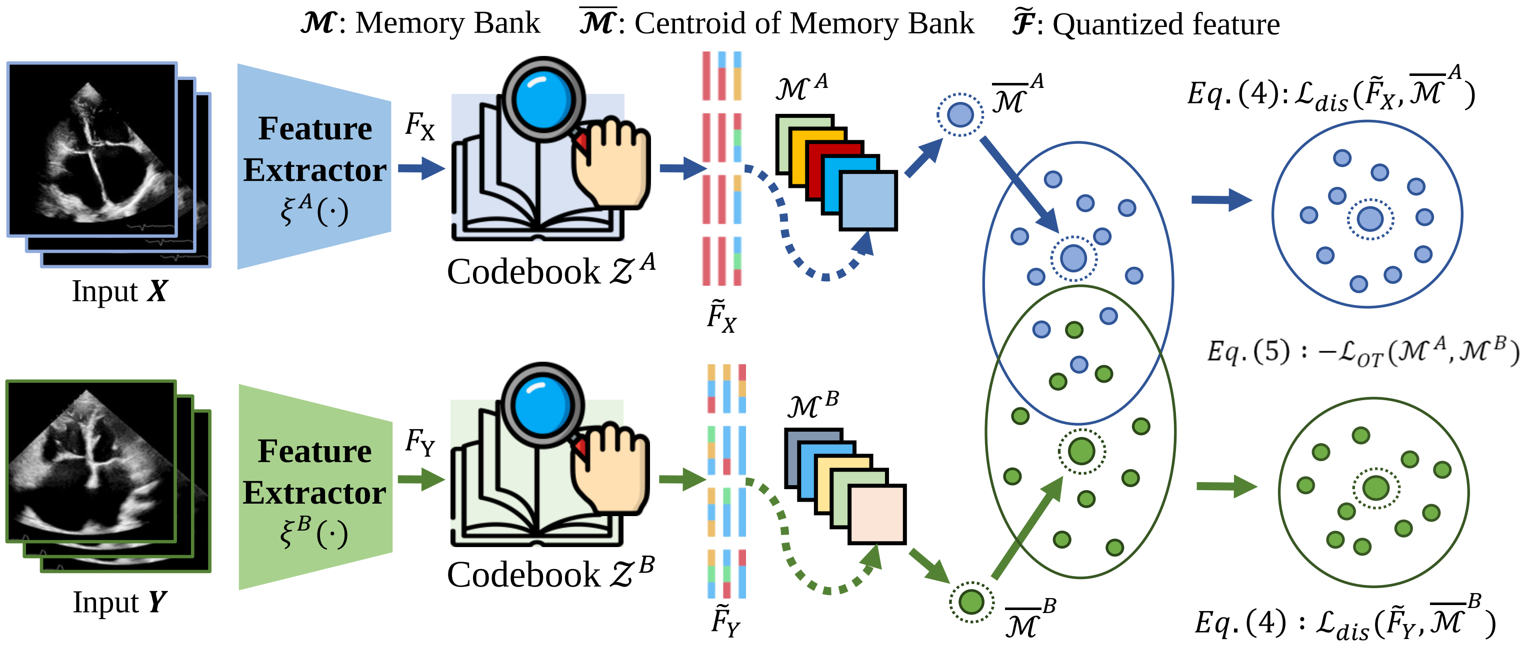


Figure 8 : The optimal transport distance optimization between two networks.

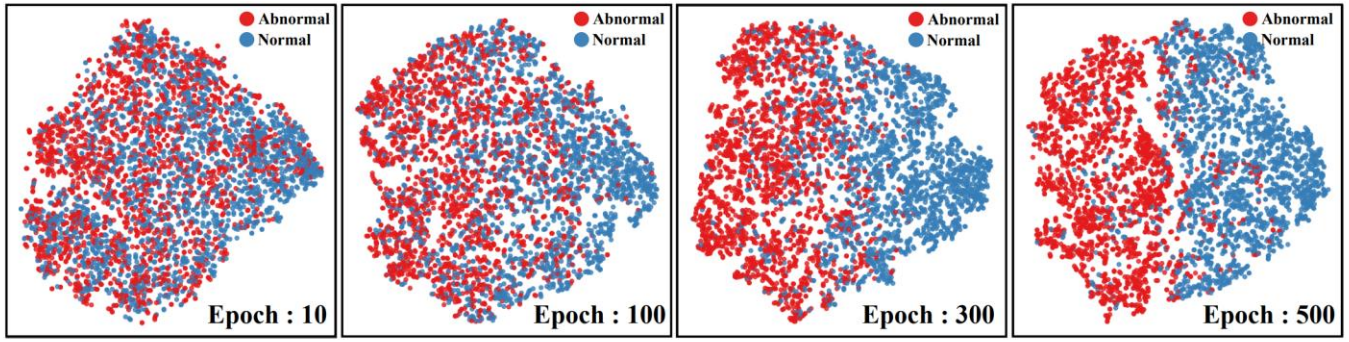


Figure 9 : The visualization of t-SNE results between learned embedding of normal and abnormal cases by our CardiacNet.

Experiments and Results in CardiacNet-PAH/ASD

We use the **Area Under the ROC Curve (AUC)** and **Classification Accuracy (ACC)** to evaluate the performance of trained networks in classifying anomalies. Also, we use **Fréchet Inception Distance (FID)** to evaluate the quality of recovery video.

Table 2: Results in CardiacNet-PAH and CardiacNet-ASD. For ASD, we also introduce the DICE score to evaluate whether recovered images are consistent with the original image in the ventricles and atrium of cardiac structures. **Bold** and Underline denotes the best and the second-best result, respectively.

Methods	Datasets							Efficiency		
	CardiacNet-PAH			CardiacNet-ASD						
	FID↓	AUC-ROC↑	ACC↑	FID↓	DICE↑	AUC-ROC↑	ACC↑	Time↓	MParams↓	TFlops↓
Classification Network										
ResNet3D	-	77.32	71.43	-	-	72.25	75.86	2.479	47.02	0.202
AGXNet	-	76.09	72.41	-	-	76.52	72.41	2.873	12.31	0.210
EchoNet	-	81.63	80.95	-	-	83.62	82.75	2.653	33.19	0.848
DeepGuide	-	82.45	81.63	-	-	<u>85.02</u>	<u>84.79</u>	3.780	15.60	0.748
DiffMIC	-	81.73	79.59	-	-	82.81	81.48	1182	88.56	38.58†
HiFuse	-	<u>84.11</u>	<u>83.67</u>	-	-	81.08	79.31	3.183	135.7	5.106
Reconstruction-Based Methods										
Vanilla GAN	18.90	52.37	46.15	19.07	63.55	60.54	58.62	2.221	12.95	0.842
DAE	16.39	58.91	57.69	15.38	65.80	54.09	53.77	1534	159.4	78.08†
VTGAN	17.66	58.32	51.72	18.10	65.13	70.92	68.97	38.50	243.3	1.423
Att. UNet	18.42	57.29	55.17	18.95	64.30	69.81	62.06	2.621	34.88	4.081
Wolleb et al.	<u>16.12</u>	70.42	67.35	<u>15.78</u>	68.61	67.88	65.51	1488	89.87	45.13†
DeScarGAN	16.59	64.21	71.42	17.04	68.52	71.33	68.97	2.756	8.528	2.756
Diff-SCM	15.57	64.23	61.22	16.37	63.26	69.23	70.83	1295	53.41	40.37†
CyTran	16.40	72.69	69.38	16.93	70.21	<u>74.35</u>	72.41	2.769	1.191	0.125
CardiacNet (Ours)	**14.73**	**89.32**	**85.71**	**15.22**	**73.52**	**91.24**	**89.63**	4.523	28.34	7.949

Experiments and Results in Public Datasets

We use the **Area Under the ROC Curve (AUC)** and **Classification Accuracy (ACC)** to evaluate the performance of trained networks in classifying anomalies. Also, we use **Fréchet Inception Distance (FID)** to evaluate the quality of recovery video.

Table 3: Results in CAMUS [5] and EchoNet [6]. **Bold** and Underline denotes the best and the second-best result, respectively.

Methods	Datasets							
	CAMUS				EchoNet			
	FID↓	MAE↓	AUC↑	ACC↑	FID↓	MAE↓	AUC↑	ACC↑
Classification / Regression Network								
ResNet3D	-	7.59	70.34	68.00	-	5.44	78.80	75.44
AGXNet	-	6.91	76.58	72.00	-	5.17	78.46	80.02
DeepGuide	-	6.72	79.66	74.00	-	4.70	84.33	79.59
EchoNet	-	<u>6.30</u>	<u>80.75</u>	<u>76.00</u>	-	4.22	83.19	81.52
HiFuse	-	6.34	80.26	76.00	-	<u>4.08</u>	<u>85.73</u>	<u>82.41</u>
Reconstruction-Based Methods								
Vanilla GAN	17.24	12.59	65.11	66.00	17.36	20.23	50.18	50.60
VTGAN	16.95	13.72	61.62	56.00	15.83	12.87	61.56	61.05
Att. UNet	17.72	9.48	65.60	62.00	16.44	8.25	65.09	61.92
CyTran	15.82	8.52	66.42	66.00	15.07	7.59	68.45	66.53
DeScarGAN	15.56	6.80	73.24	68.00	14.19	7.23	73.24	71.08
Wolleb et al.	15.17	8.06	75.96	74.00	**13.18**	8.50	72.38	69.57
CardiacNet (Ours)	**14.64**	**5.97**	**83.09**	**80.00**	<u>13.25</u>	**3.83**	**86.52**	**84.70**

Ablation Studies and Visualization

- Our CDC can **help** distinguish cardiac structural and motion **abnormalities**;
- Both **global and local** discriminators can **contribute to the CDD module**;
- The combination of CDD, CDC in CardiacNet achieves the **best performance** in both reconstruction and classification tasks;
- Our CardiacNet also show it ability in **visualizing the abnormalities**.

Table 4: Effectiveness of CDC and CDD. Results report in CardiacNet-PAH.

CDC	CDD	Results FID	AUC	ACC
✗	✗	18.90	52.37	46.15
✓	✗	16.82	80.27	79.59
✗	✓	17.09	52.46	53.84
✓	✓	**14.73**	**89.23**	**85.71**

Table 5: Ablation study of Positional Encoding and Optimal Transport in **only CDC module.**

Pos. Encode	Opt. Trans	Results FID	AUC	ACC
✗	✗	18.90	52.37	46.15
✓	✗	17.41	62.44	65.38
✗	✓	18.06	78.39	75.51
✓	✓	**16.82**	**80.27**	**79.59**

Table 6: Ablation study of Global and Local discriminator in CDD module (Enabling CDC).

Global. CDD	Local. CDD	Results FID	AUC	ACC
✗	✗	16.82	80.27	79.59
✓	✗	15.62	82.41	83.67
✗	✓	15.41	84.57	81.63
✓	✓	**14.73**	**89.23**	**85.71**

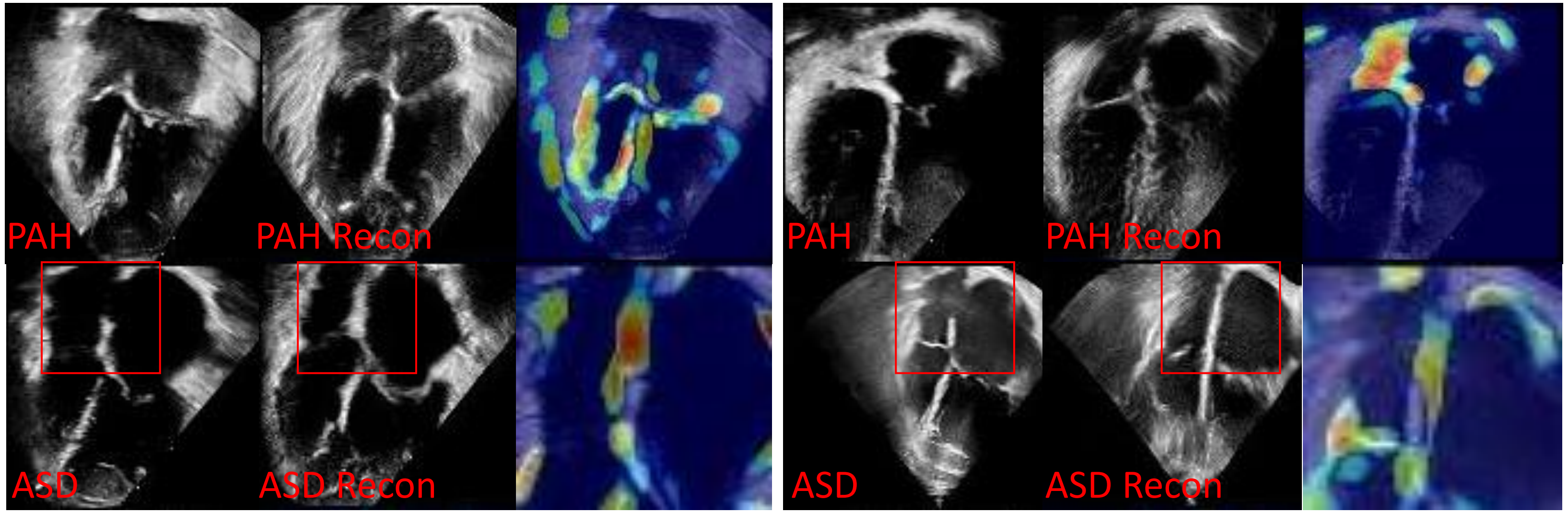


Figure 10 : Some visualization of our proposed method in CardiacNet-PAH and CardiacNet-ASD datasets.

Visualization of Different Methods in CardiacNet-ASD

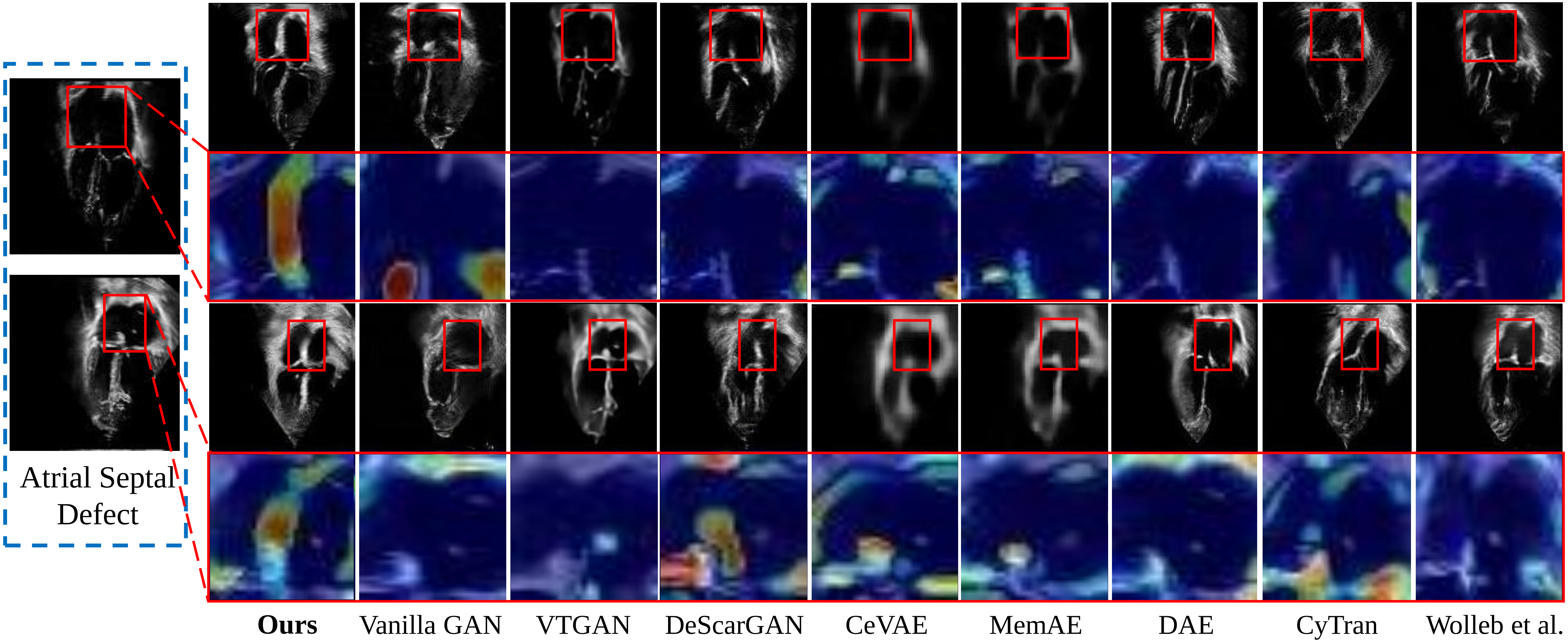


Figure 11 : Visualization of ours and other reconstruction-based methods in CardiacNet-ASD dataset.

Our Contribution

- We introduce **CardiacNet-PAH** and the **CardiacNet-ASD** that specifically designed for cardiac disease assessment using Echocardiogram videos;

- Our proposed **CardiacNet method is a novel approach** that can capture local structural details and cardiac motion changes, enabling accurate assessment of cardiac diseases via Echocardiogram videos.

- Our **CardiacNet surpasses prior work** in classifying PAH and ASD with an improvement of 2.1% and 5.0% in accuracy. The CardiacNet also achieves a relative reduction of 5.2% compared to prior arts in the EF prediction task.

Thank You!

Code & data: ***https://github.com/xmed-lab/CardiacNet***

Contact :
Jiewen Yang - jyangcu@connect.ust.hk
Xiaomeng Li - eexmli@ust.hk