

PhD Viva · Aug 2025

Robust Cardiac Feature Monitoring based on Millimeter-Wave Radar

Yuanyuan Zhang

Electrical Eng. and Electronics, UoL & School of Advanced Technology, XJTLU

Principle Supervisor: Dr. Rui Yang

Co-Supervisor: Prof. Yi Huang

Prof. Steven Guan

Dr. Yutao Yue



Contents

1. Background and Challenges

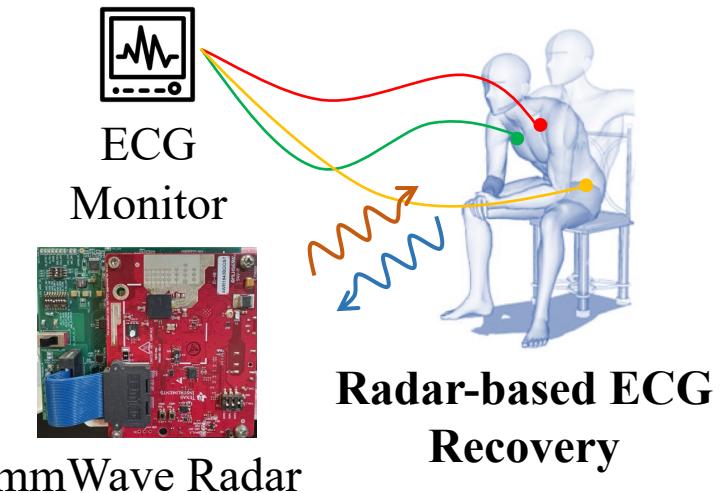
2. Proposed Methods with Results

- Study A: Efficient High-SNR Signal Acquisition
- Study B: Robust Single-cycle ECG Generation
- Study C: Robust Long-term ECG Generation
- Study D: Alleviation of Data Scarcity

3. Conclusions and Future Work

Why we need contactless cardiac features monitoring?

1. Suitable for long-term monitoring or special patients (e.g., burn patients)
2. Suitable for in-cabin monitoring or smart home (e.g., healthcare scenarios)



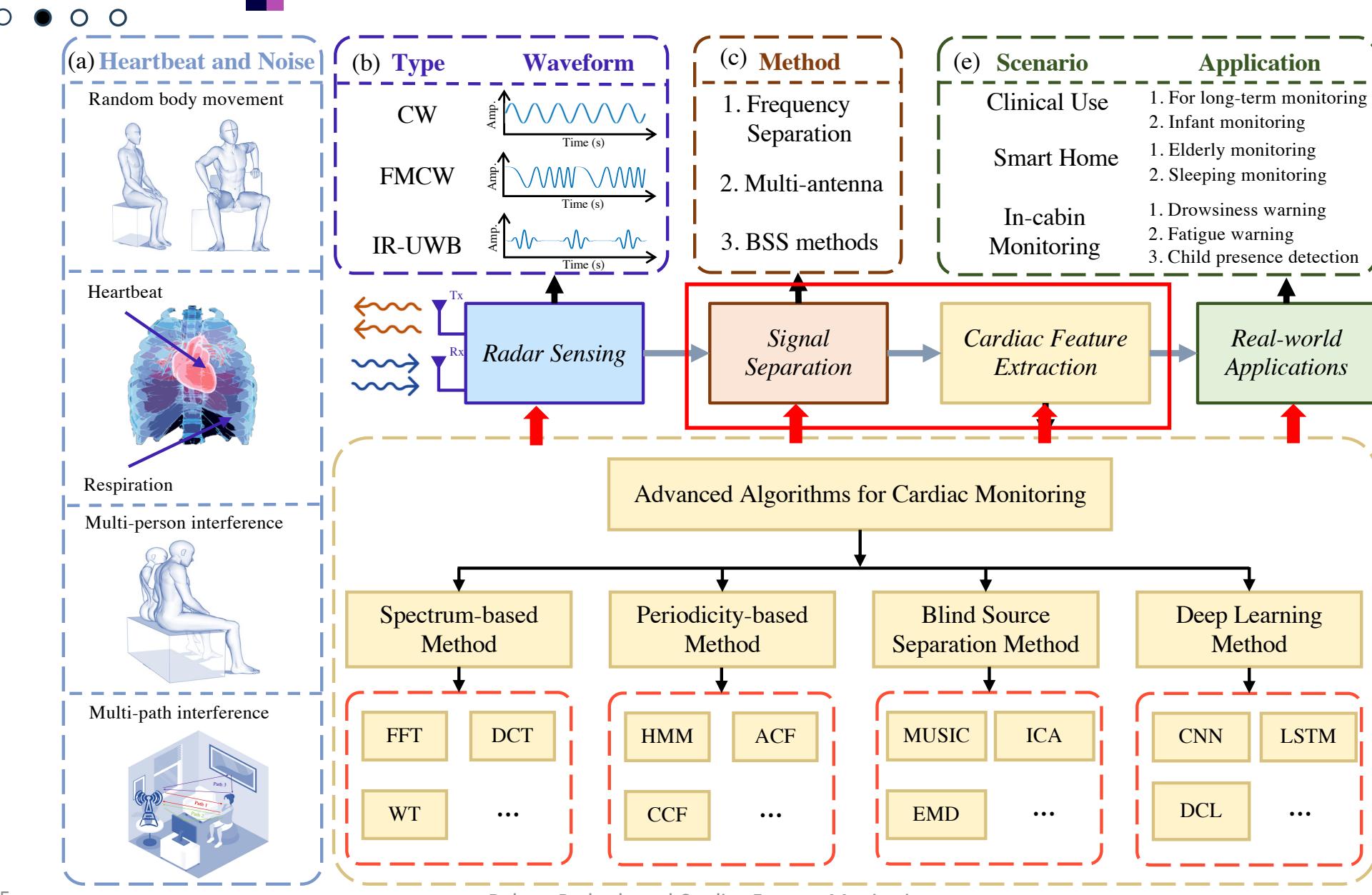
Why we use radar as the contactless sensor?

Sensors	Cameras	Acoustic Sensors	Wi-Fi Routers	Radars
Pros	1. SOTA. CV algorithms or DL frameworks	1. Directly monitor the heart sound	1. No need for extra devices 2. Unobtrusive	1. Unobtrusive 2. Contactless 3. Robustness
Cons	1. Privacy issue 2. Rely on high quality image.	1. Low accuracy 2. Vulnerable to acoustic noise	1. Low accuracy 2. Complex EM environment	1. Require special signal processing design 2. Less investigated



Background

Research Background – General Pipeline



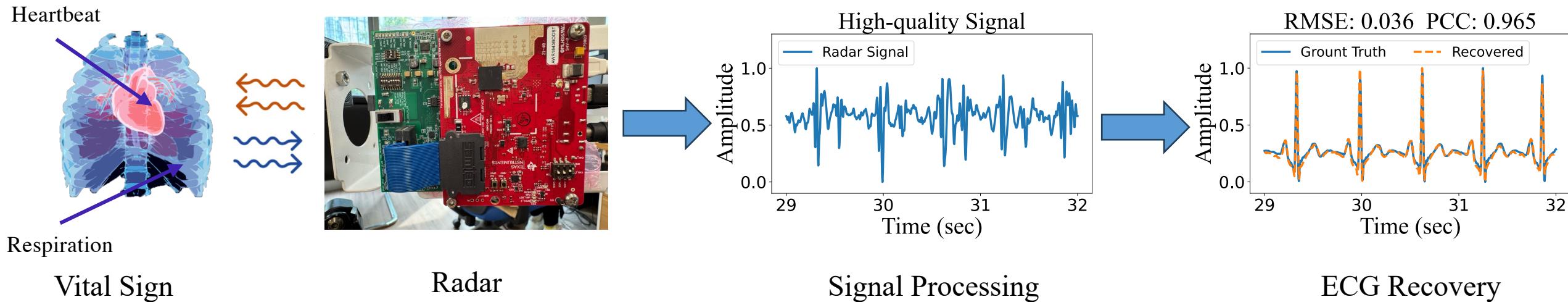


Background



Research Gaps and Contributions – Research Gaps

Illustration for radar-based ECG recovery :



Research gap in radar-based ECG recovery:

1. Design efficient algorithm for **high-SNR** signal collection.
2. Model and realize the **domain transformation** from radar to ECG **for fine-grained cardiac features**.
3. Realize **long-term ECG recovery** while **mitigating the real-world noise** (e.g., body movement, constant noise)
4. Alleviate **data scarcity** for the deployment in new scenarios with **limited data**.



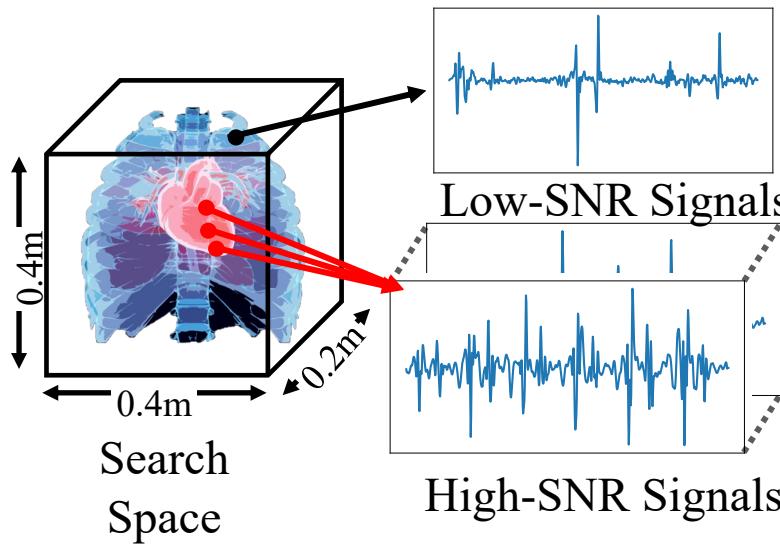
Background

- ○ ○ ●

Research Gaps and Contributions – Contributions

Contributions:

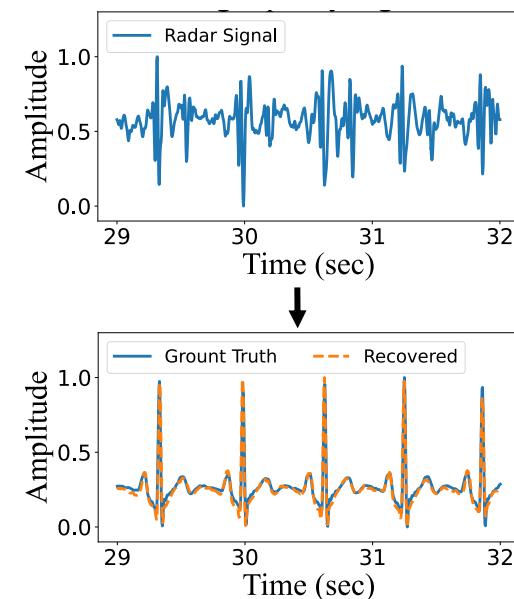
1. Propose a **derivative-free optimization method** to iteratively approach the point with ample cardiac features.
2. Design a **signal model** to describe the relationship between cardiac activities from different domains (radar to ECG).
3. Design **the ODE decoder and multi-task learning framework** to realize ECG recovery and resist constant/abrupt noises.
4. Propose an **augmentation** method and a **transfer learning** framework to alleviate data scarcity.



Chapter 3

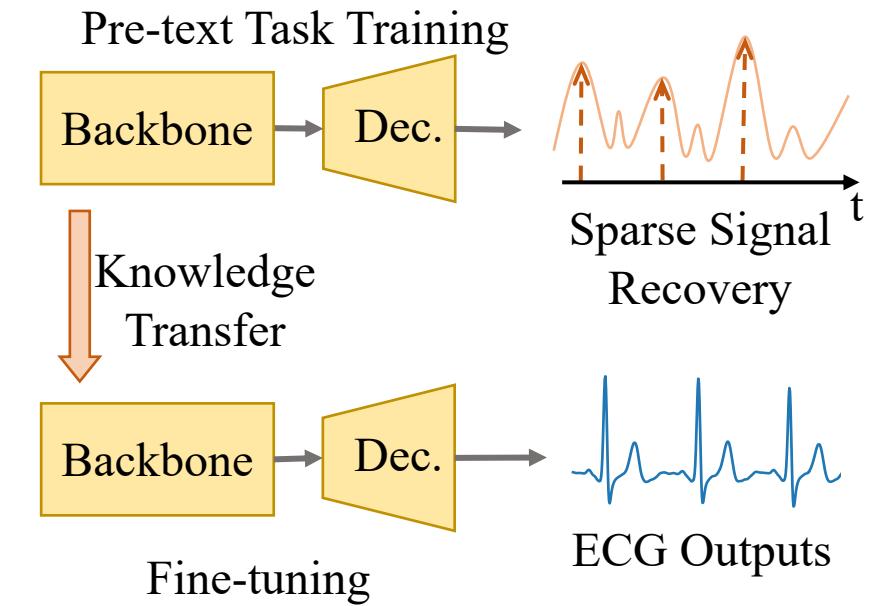
Submitted to TMC

8/11/25



Chapter 4 and 5

Published in TMC and TIM



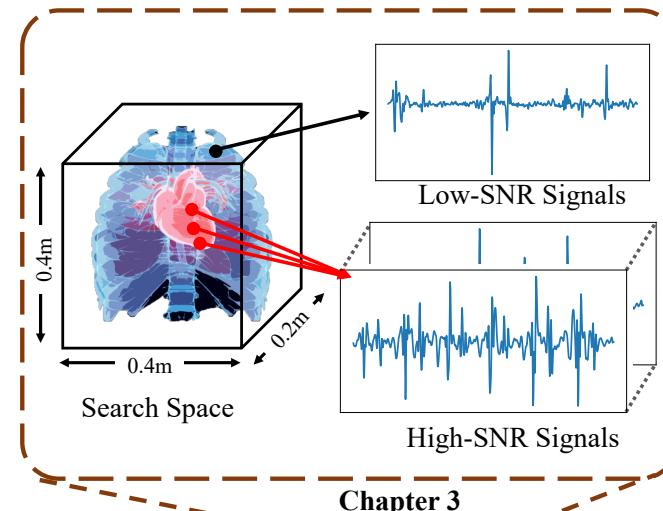


Background



Research Gaps and Contributions – Contributions

Contribution 1



Chapter 3

Chapter 1

Introduction

Efficient High-SNR
Signal Acquisition

Chapter 2

Background and
Literature Review

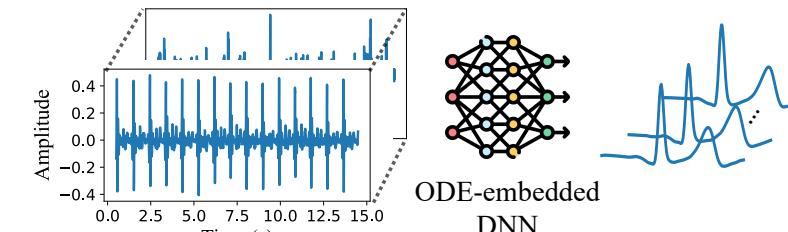
Robust Single-cycle
ECG Generation

Alleviation of Data
Scarcity

Conclusions and
Future Work

Chapter 7

Contribution 2



Pre-text Task Training

Backbone

Knowledge Transfer

Backbone

Fine-tuning

Sparse Signal Recovery

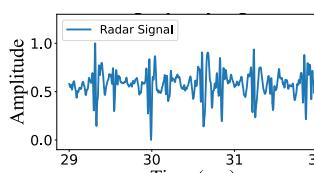
ECG Outputs

Contribution 4

Chapter 6

Robust Long-term
ECG Generation

Chapter 5

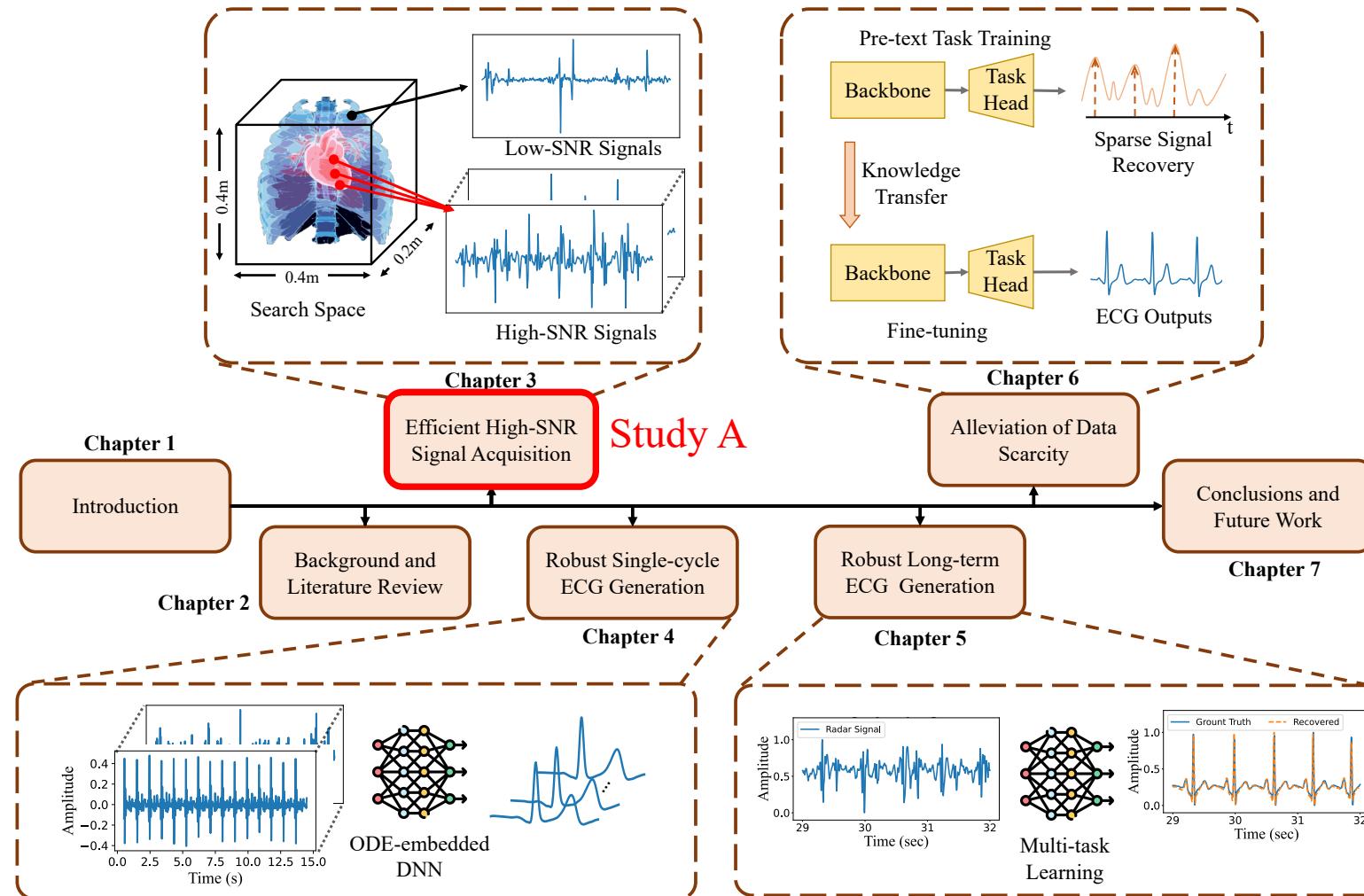


Multi-task
Learning

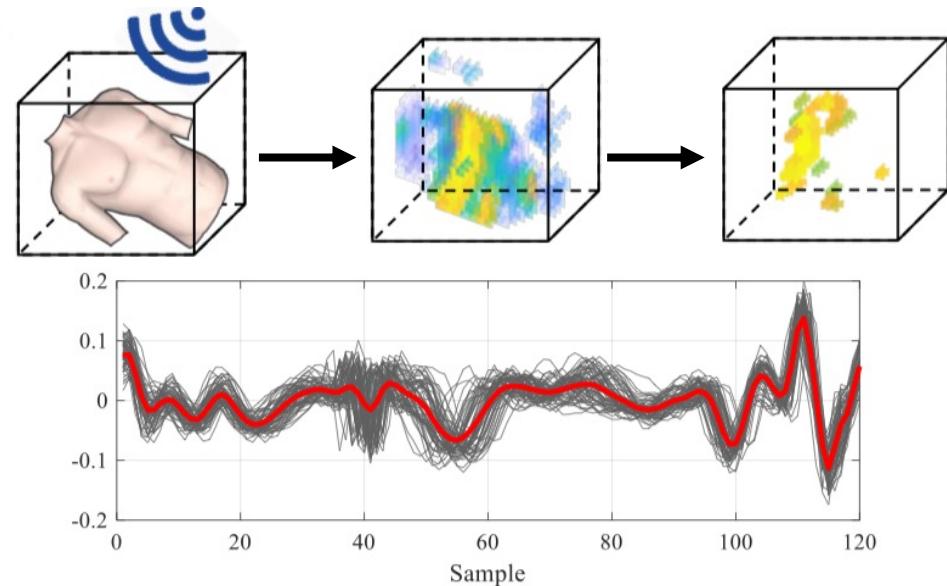
Contribution 3



Part 2 – Proposed Methods

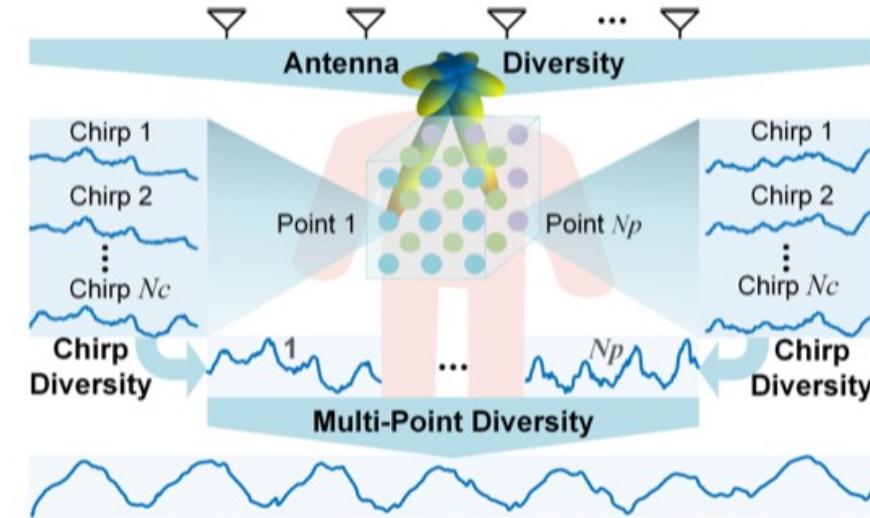


Clustering with exhaustive calculations [1]

**Limitations:**

1. Time consuming
2. Rely on precise body localization
3. Not suitable for long distance

Signal accumulation for noise suppression [2]

**Limitations:**

1. Rely on precise body localization
2. Could amplify non-gaussian noise
3. Only used for coarse vital sign, e.g., heartbeat monitoring

[1] J. Chen, D. Zhang, Z. Wu, F. Zhou, Q. Sun, and Y. Chen, "Contactless electrocardiogram monitoring with millimeter wave radar," IEEE Transactions on Mobile Computing, Dec. 2022.

[2] J. Liu, J. Wang, Q. Gao, X. Li, M. Pan, and Y. Fang, "Diversity-enhanced robust device-free vital signs monitoring using mmWave signals," IEEE Transactions on Mobile Computing, Jun. 2024.

Contributions:

1. Propose **cardio-focusing and -tracking (CFT) algorithm** to find **cardio-focused (CF) points** with clear cardiac features from a **discontinuous objective space**.
2. Provide a faithful approach to **assessing SNR without knowing the ground truth**, avoiding exhaustive calculations
3. The proposed CFT algorithm has been validated on sitting **subjects in various scenarios** and could provide radar measurements with **better SNR** compared with existing methods.



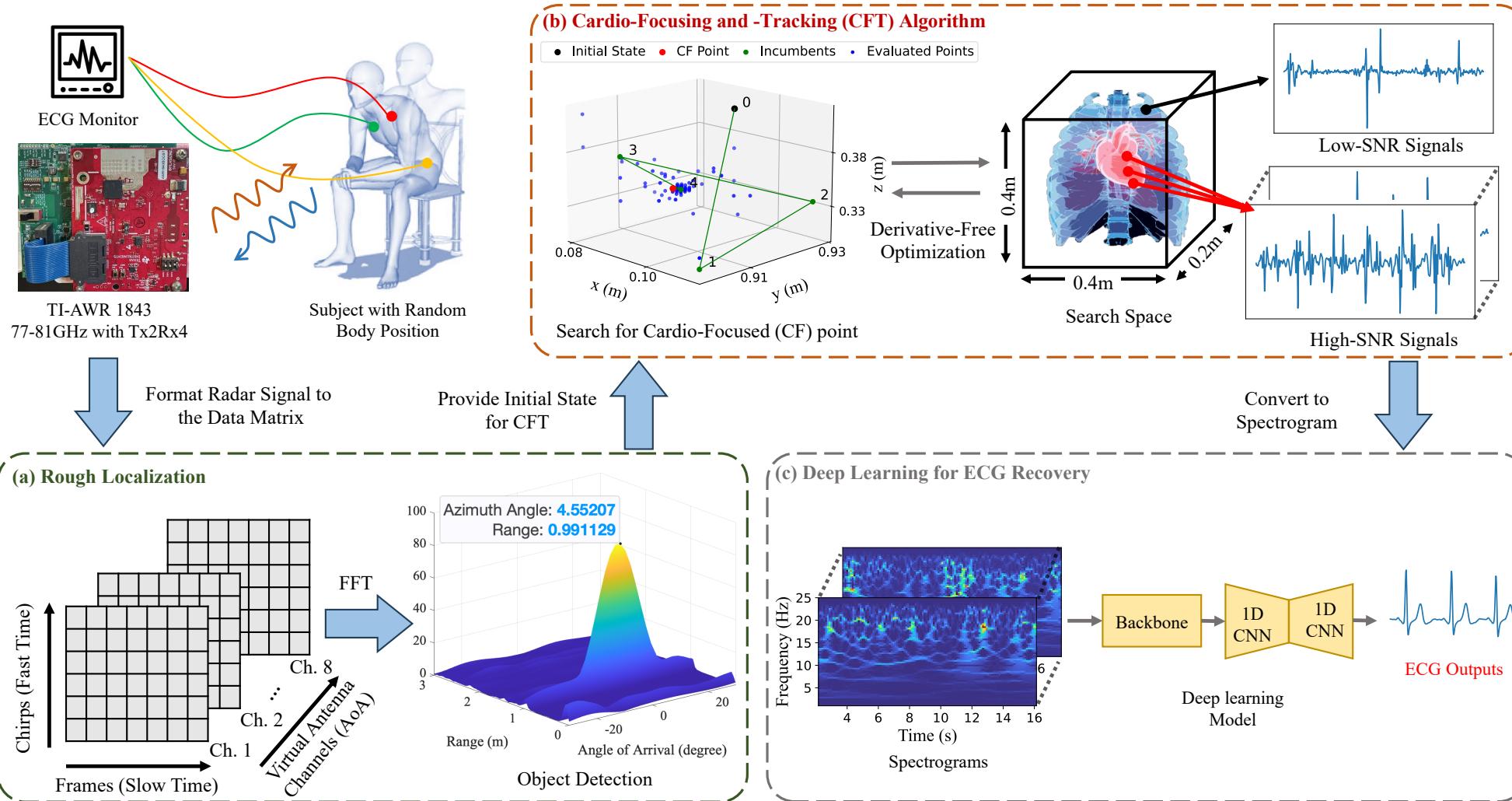
Data Collection scenarios



TI AWR-1843 @77GHz
with 3.8GHz bandwidth



Efficient High-SNR Signal Acquisition – Overview

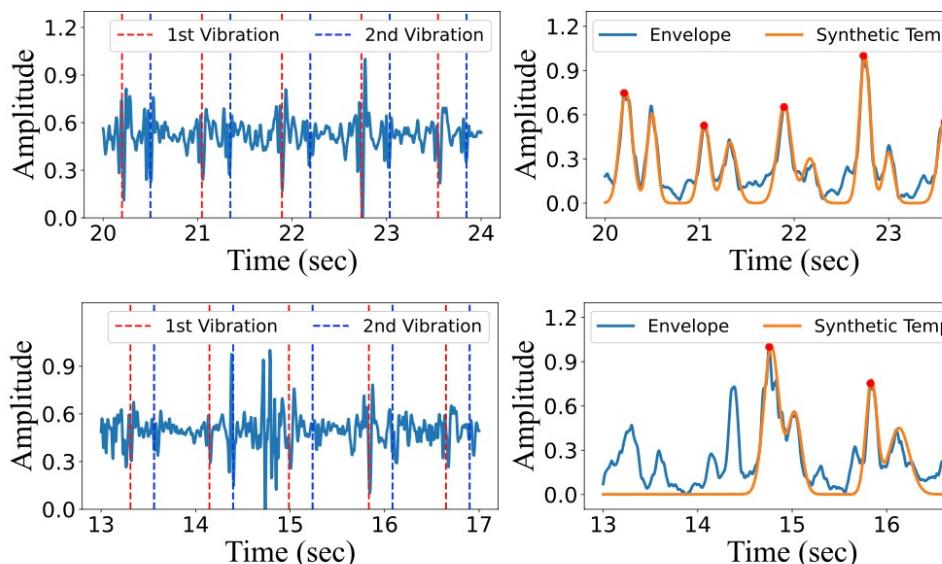




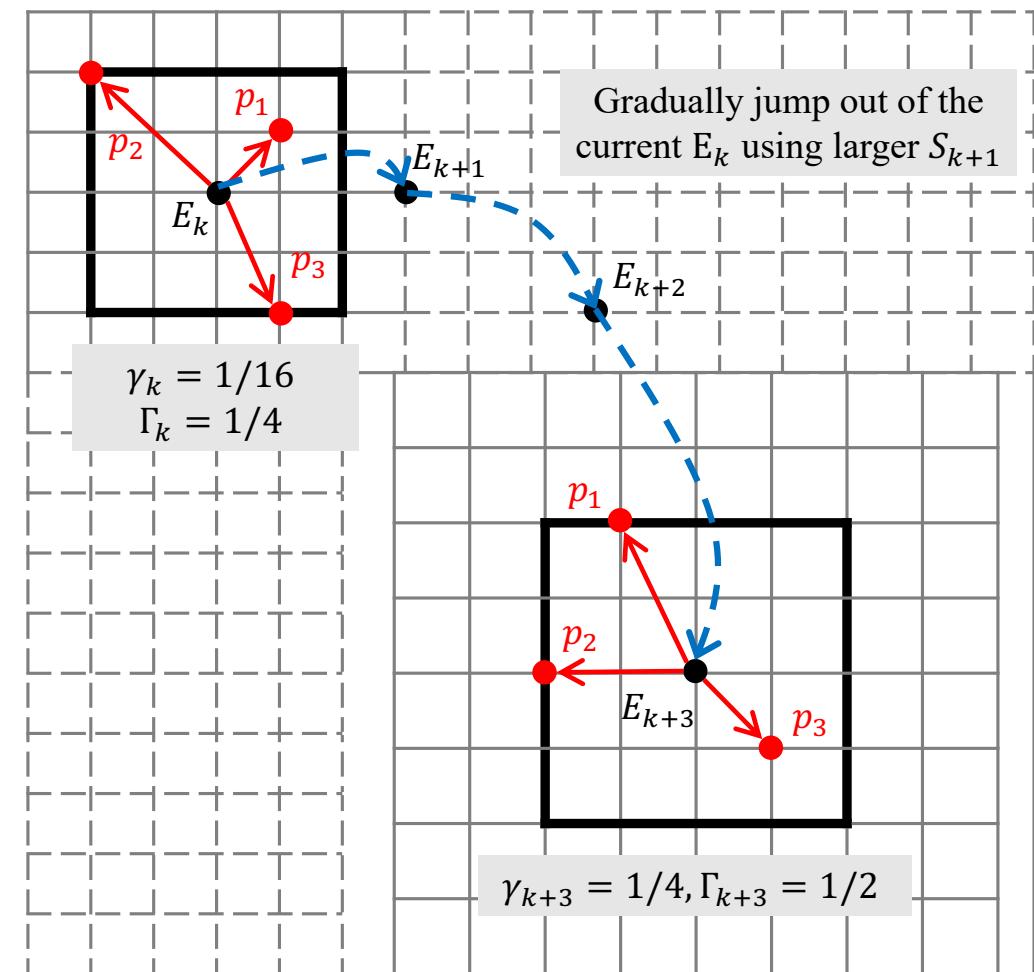
Efficient High-SNR Signal Acquisition – CFT

Advantages of CFT:

1. Provide a faithful approach to **assessing SNR without knowing the ground truth**, avoiding exhaustive calculations.
2. Suitable for the highly discontinuous objective space and could **break local minima by restricting the search space**.



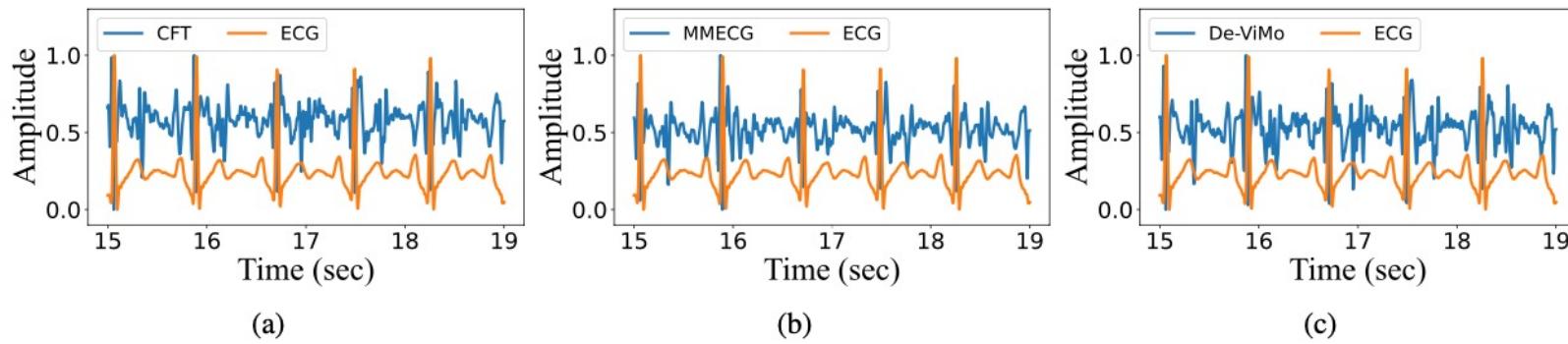
Assess signal SNR with universal template



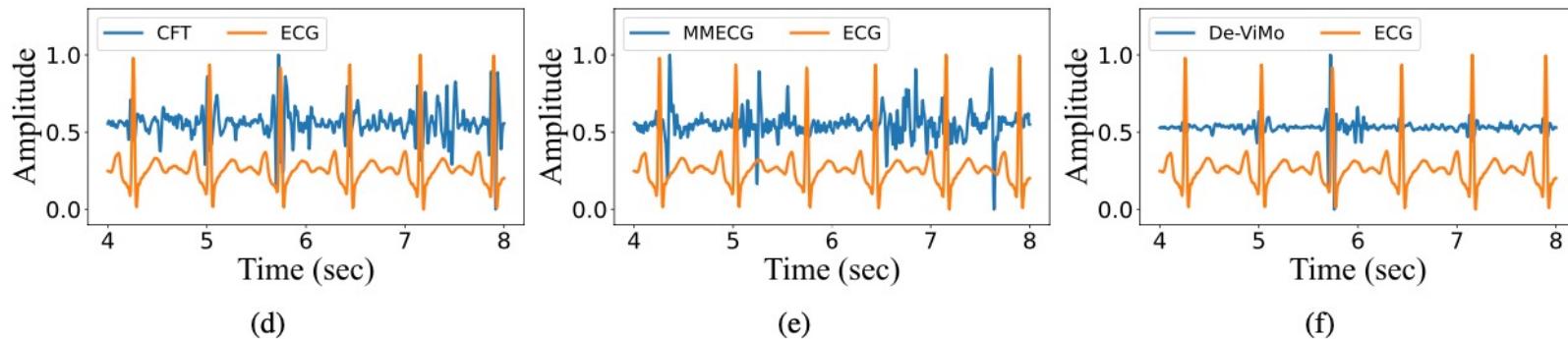
Bold search region based on the spanning grids

Efficient High-SNR Signal Acquisition – Results

**Rough Loc.
≈
CF Point**

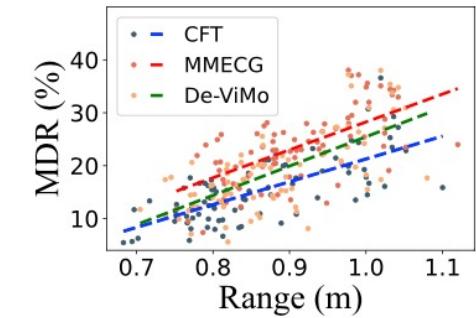
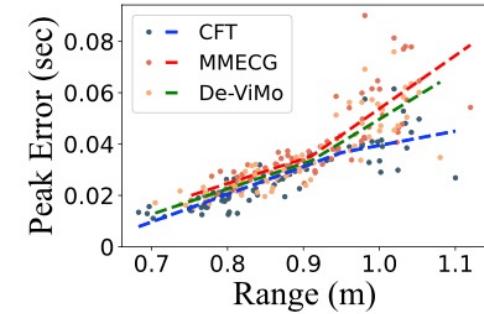


**Rough Loc.
≠
CF Point**



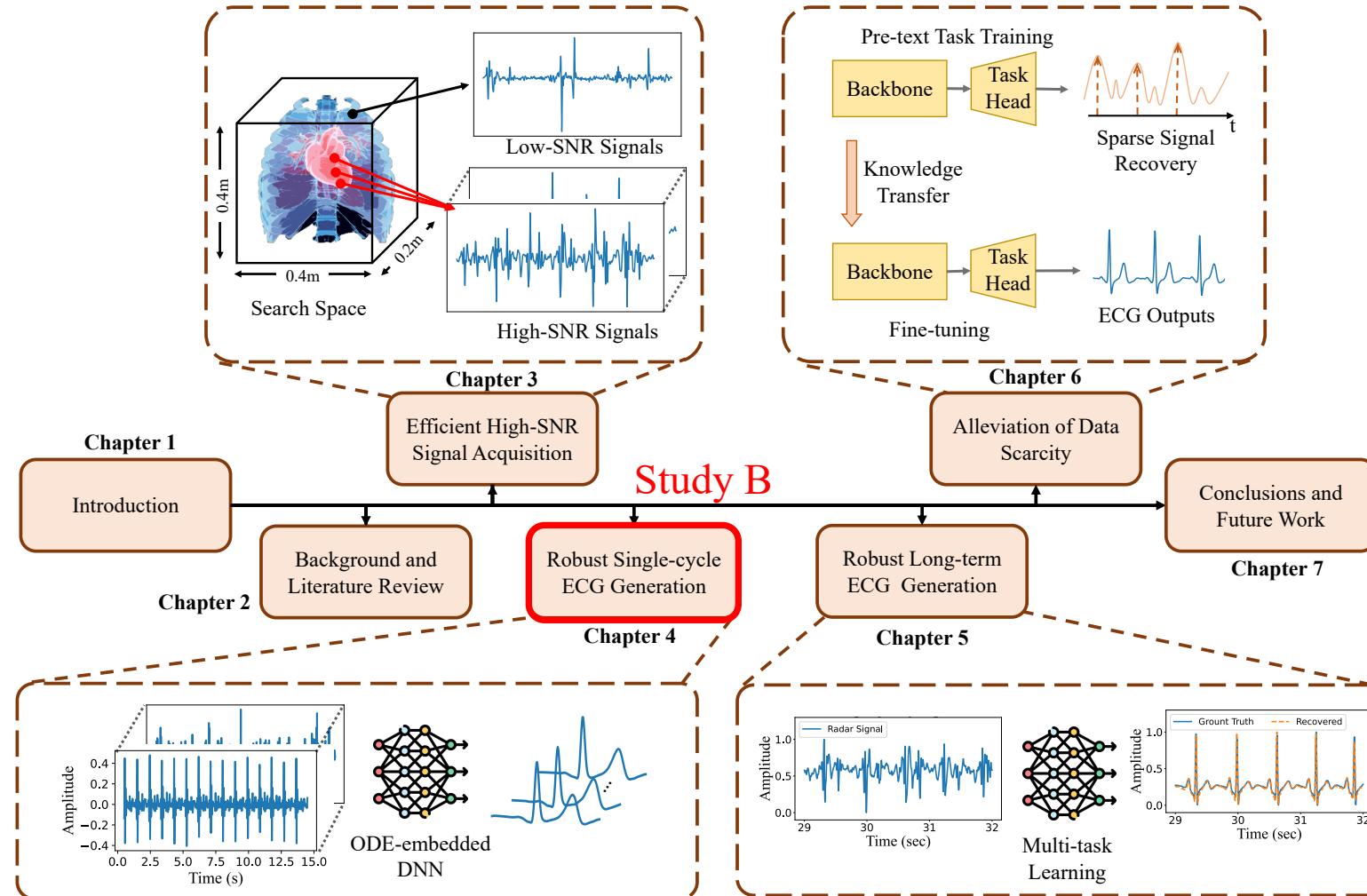
Remarks:

1. Clustering- or accumulation-based method **rely on a good initial estimation**, but the FMCW signal processing could only provide **rough body location**.
2. The **body posture** is subject to **change** during data collection, causing a deviation of CF point by several decimeters. CFT could **track the CF point** by reactivating the optimization if SNR is low.

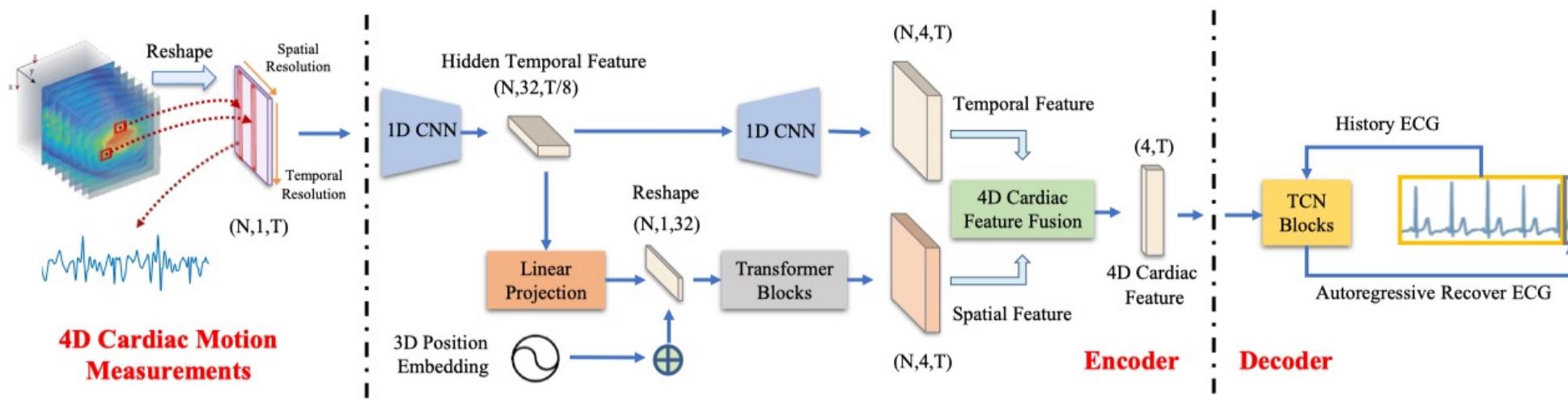




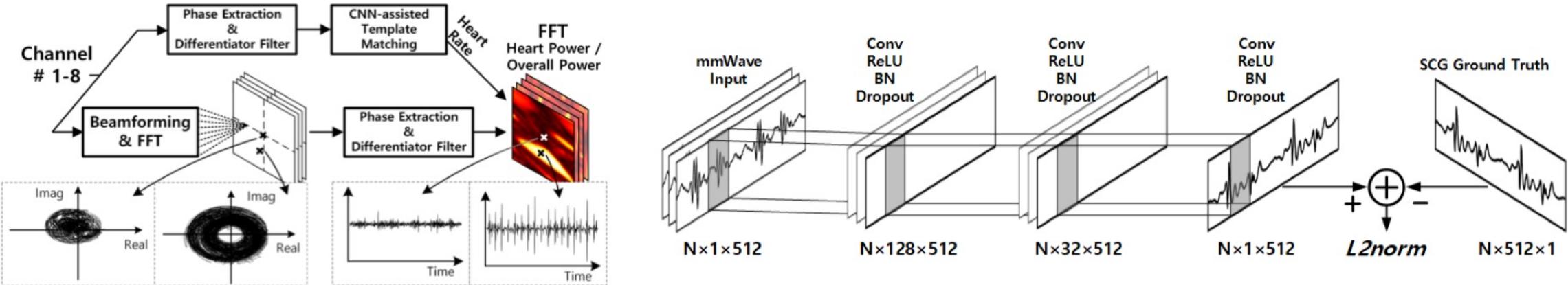
Part 2 – Proposed Methods



[1]



[3]

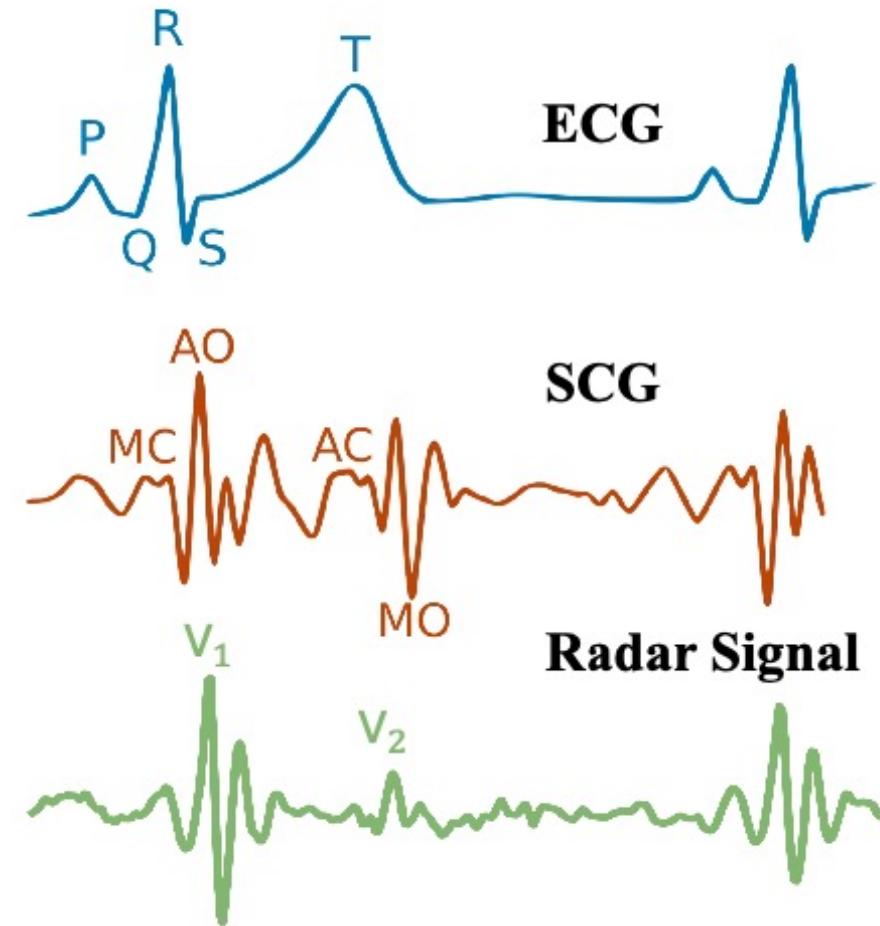


[1] J. Chen, D. Zhang, Z. Wu, F. Zhou, Q. Sun, and Y. Chen, "Contactless electrocardiogram monitoring with millimeter wave radar," IEEE Transactions on Mobile Computing, pp. 1–17, Oct. 2022.

[3] U. Ha, S. Assana, and F. Adib, "Contactless seismocardiography via deep learning radars," in Proceedings of the 26th Annual International Conference on Mobile Computing and Networking (MobiCom), pp. 114, Apr. 2020.

Domain transformation of cardiac features:

1. ECG Feature: **P-wave, QRS-complex, T-wave** for electrical activities of atrial/ventricular depolarization/repolarization.
2. SCG Features: **excitation–contraction coupling** caused valve opening/closure (**AO/AC**) and mitral valve opening/closure (**MO/MC**)
3. Radar Feature: Similar with SCG but only with two prominent vibrations **v1** and **v2** induced by AO and AC.
4. Heart muscle contraction has a **pulsatile nature**, and the bones/tissues in chest area introduce the extra **damping** into the pulse [4], [5].

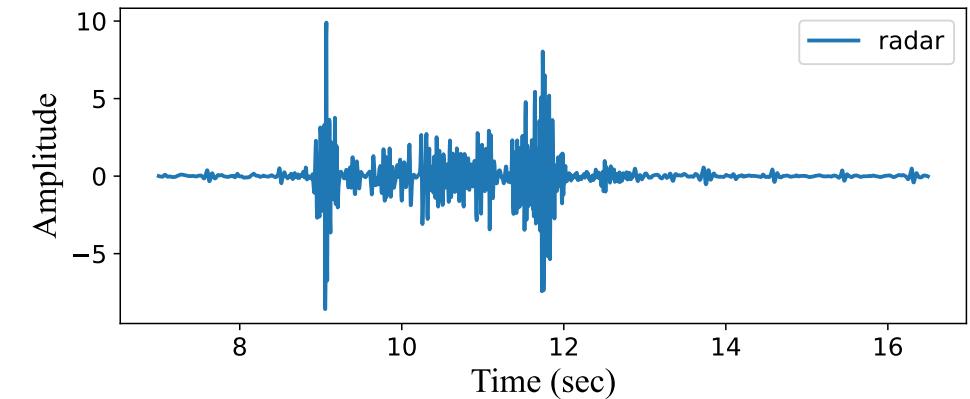


[4] M. Nosrati and N. Tavassolian, "High-accuracy heart rate variability monitoring using Doppler radar based on Gaussian pulse train modeling and FTPR algorithm," IEEE Transactions on Microwave Theory and Techniques, vol. 66, no. 1, pp. 556–567, Jan. 2017.

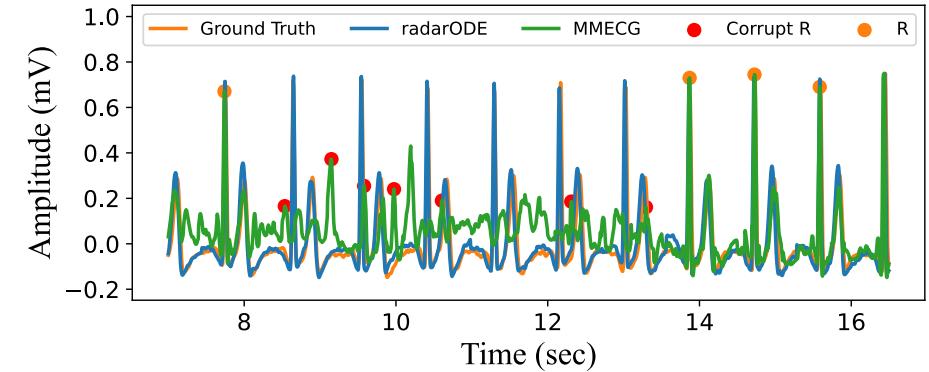
[5] D. R. Morgan and M. G. Zierdt, "Novel signal processing techniques for doppler radar cardiopulmonary sensing," Signal Processing, vol. 89, no. 1, pp. 45–66, Jan. 2009.

Contributions:

1. Propose the **signal model in terms of fine-grained cardiac features**.
2. Design an ODE-embedded module called **single-cycle ECG generator (SCEG)** to **parameterize the radar signal** into sparse representations with morphological meanings.
3. By **fusing the extracted morphological and temporal features**, the proposed radarODE framework is proven to be robust in the presence of body movement noise and could realize accurate reconstruction of the ECG signal.



Distorted radar signal



Recovered ECG signal

Proposed signal model:

Based on the previous intuition, we model the cardiac vibration $\tilde{x}(t)$ as

two **prominent vibrations** and other noises as:

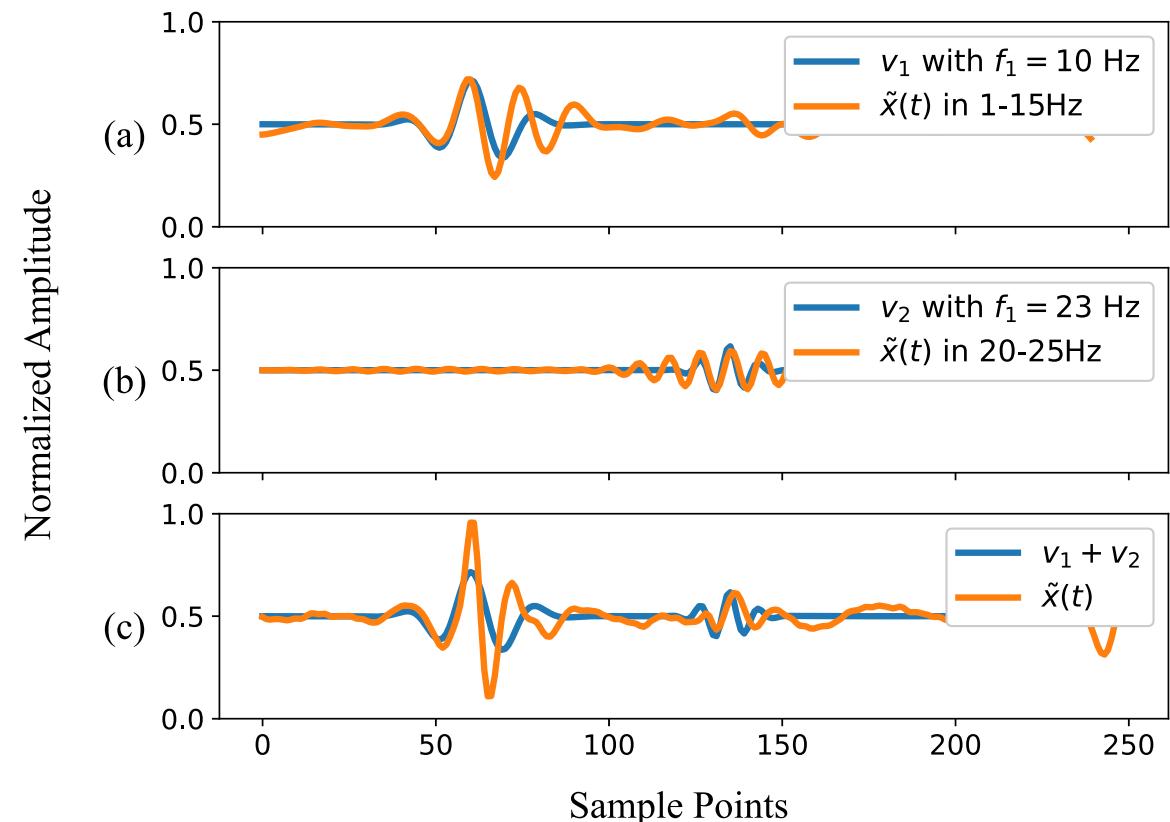
$$\tilde{x}(t) = v_1(t) + v_2(t) + x_n(t)$$

with

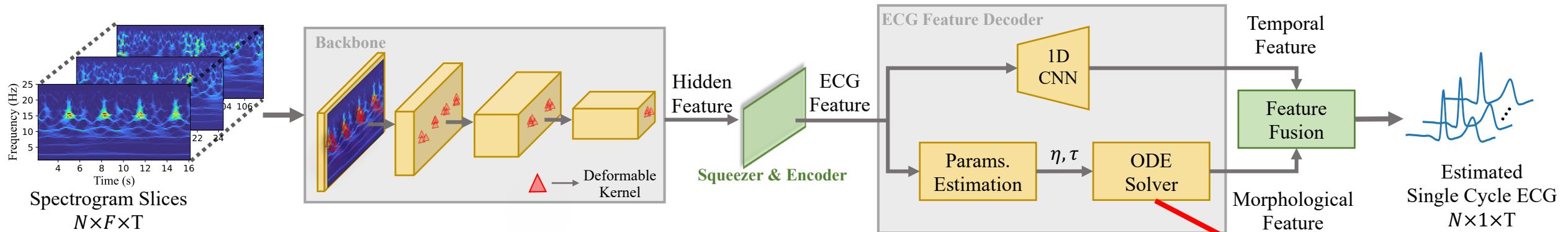
$$v_1 = a_1 \cos(2\pi f_1 t) \exp\left(-\frac{(t - T_1)^2}{b_1^2}\right)$$

$$v_2 = a_2 \cos(2\pi f_2 t) \exp\left(-\frac{(t - T_2)^2}{b_2^2}\right)$$

$a_1, a_2 \rightarrow$ amplitude of the vibrations,
 $b_1, b_2 \rightarrow$ length/width of the vibration,
 $f_1, f_2 \rightarrow$ **central frequency of the vibration**,
 $T_1, T_2 \rightarrow$ **when the vibrations happen**.



Single Cycle ECG Generator (SCEG):



The ODE model describes a **3D trajectory** denoted by (x, y, z) [6]

$$\frac{dx}{dt} = \alpha(x, y)x - \omega y$$

$$\frac{dy}{dt} = \alpha(x, y)y + \omega x$$

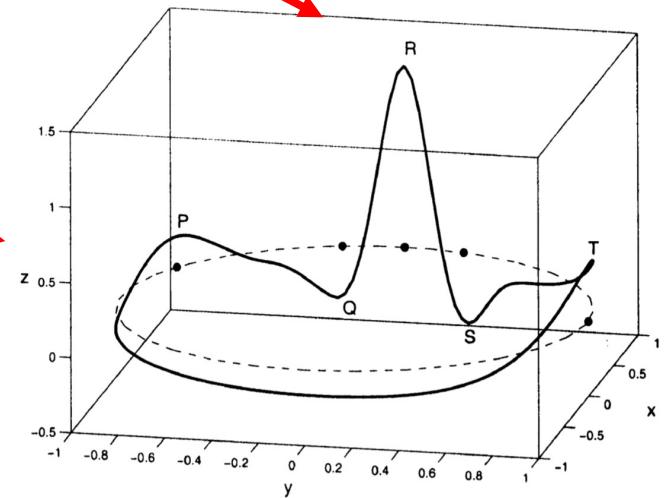
$$\frac{dz}{dt} = - \sum_{e_f \in \mathcal{F}} a_{e_f} \Delta \theta_{e_f}(x, y) e^{-\Delta \theta_{e_f}(x, y)^2 / 2 b_{e_f}^2} - z$$

$$\alpha(x, y) = 1 - \sqrt{x^2 + y^2}$$

$$\Delta \theta_{e_f}(x, y) = (\theta(x, y) - \theta_{e_f}) \mod 2\pi$$

$$\theta(x, y) = \text{atan } 2(y, x) \in [-\pi, \pi]$$

$$e_f \in \mathcal{F} = \{e_P, e_Q, e_R, e_S, e_T\}$$

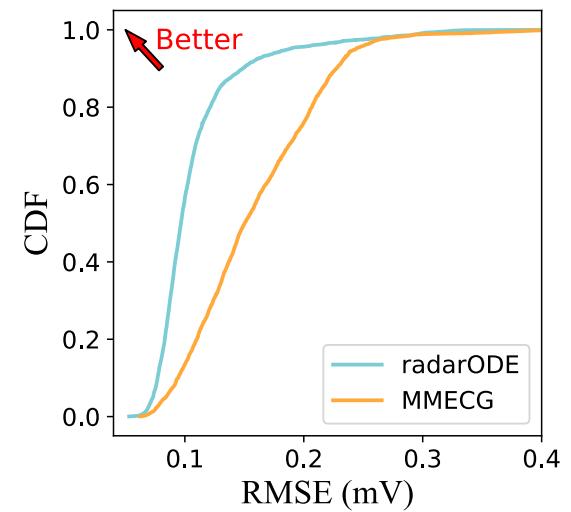
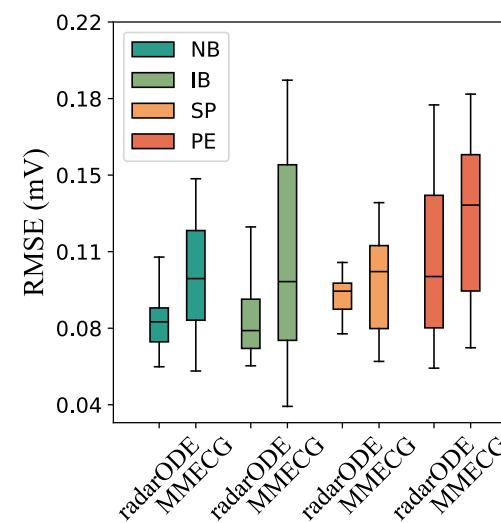
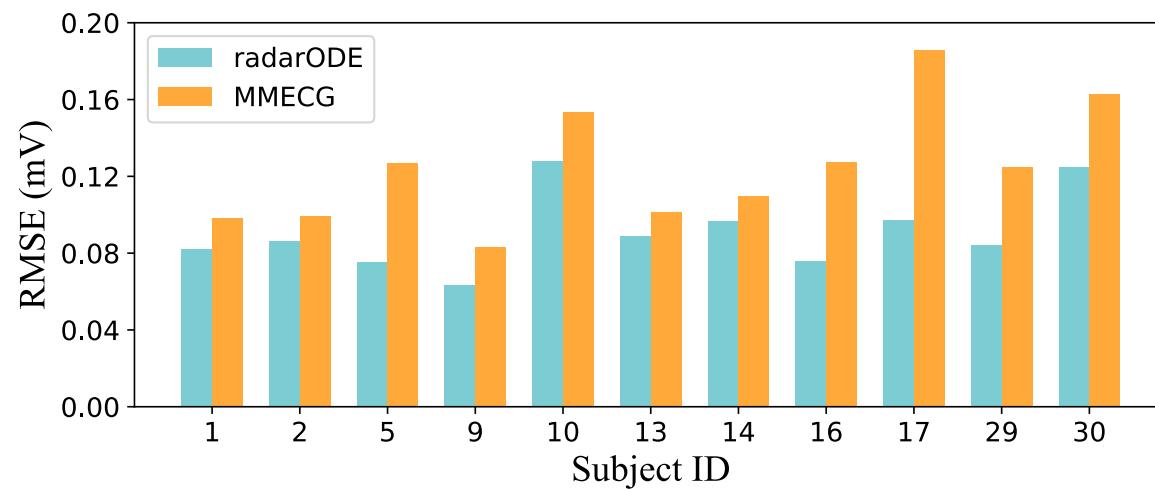
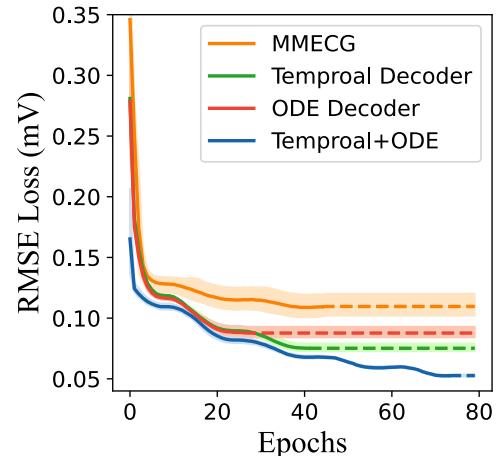


[6] P. E. McSharry, G. D. Clifford, L. Tarassenko, and L. A. Smith, “A dynamical model for generating synthetic electrocardiogram signals,” IEEE transactions on biomedical engineering, vol. 50, no. 3, pp. 289–294, Mar. 2003.

Robust Single-cycle ECG Generation – Results

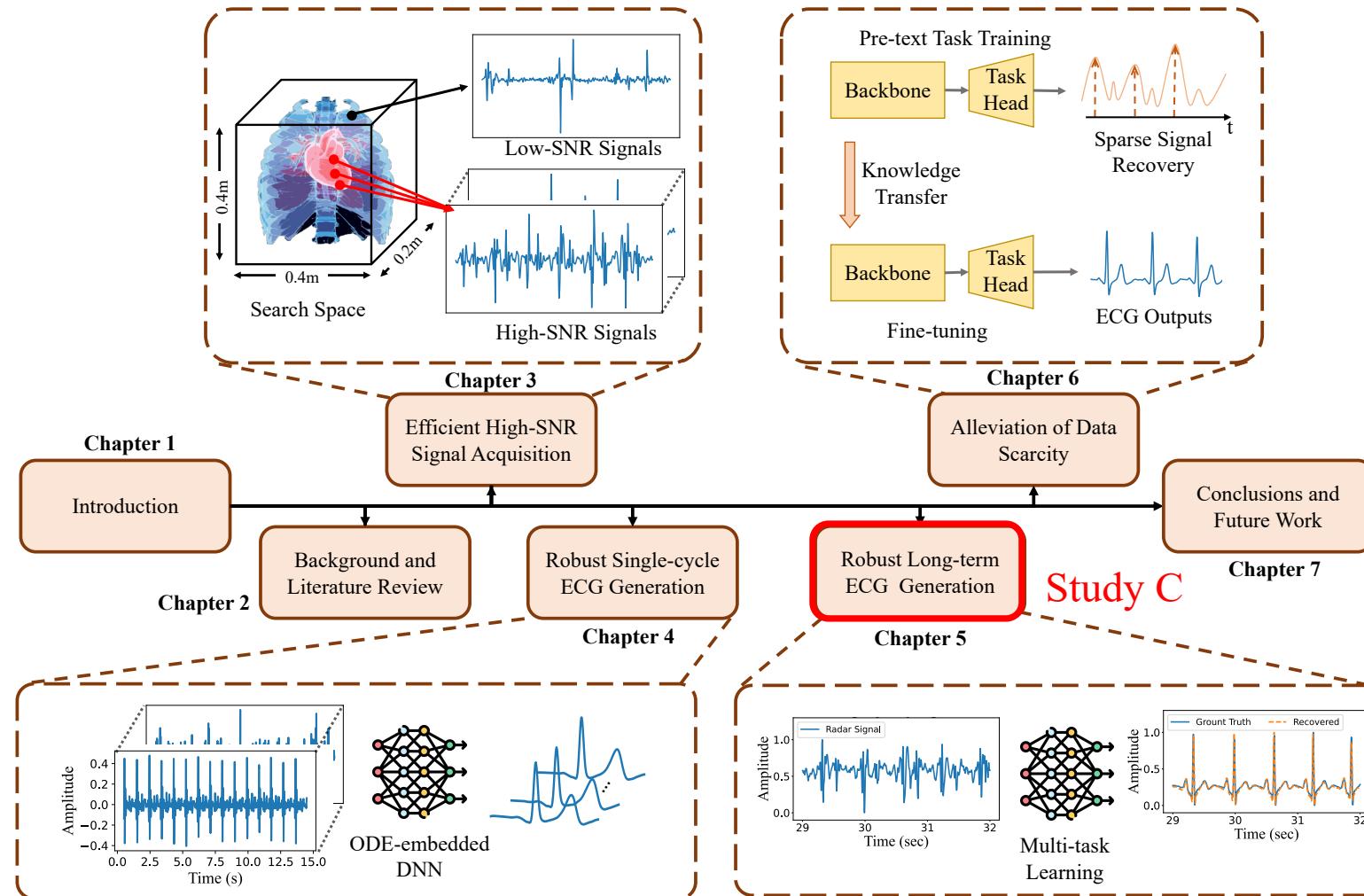
COMPARISON WITH BENCHMARK AND ABLATION STUDY ABOUT TEMPORAL DECODER AND ODE DECODER (WITHOUT RESAMPLING)

Framework	Backbone	Encoder	Decoder	Fusion Method	RMSE (mV)	Correlation
MMECG [33]	Conv1d + Transformer		Transconv1d + TCN	Multiplication	0.091	87.9%
SCEG	DeformConv2d	Conv1d	Initial + Temporal	-	0.086	89.4%
			Initial + ODE	-	0.092	85.5%
			Initial + Temporal + ODE	Multiplication + Stack + Conv2d	0.077	92.6%



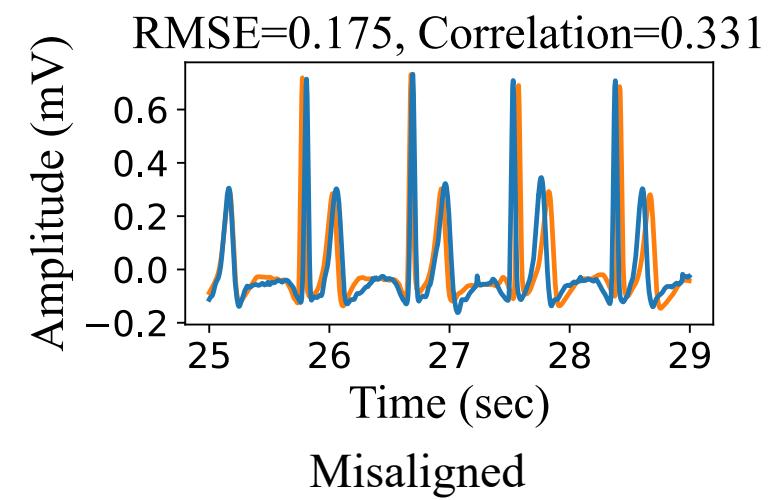
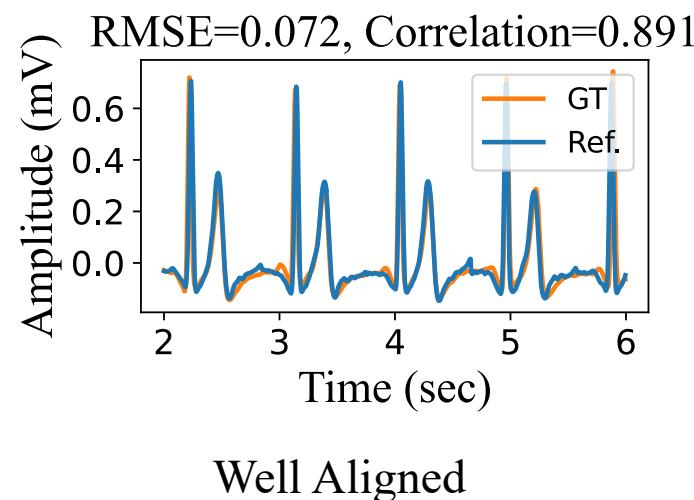
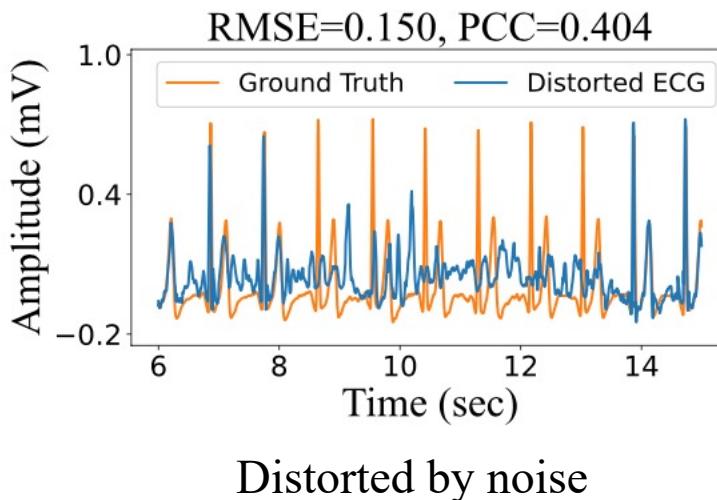


Part 2 – Proposed Methods



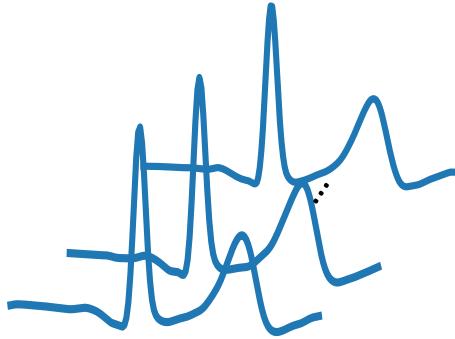
Research gap in radar-based ECG recovery:

1. The **long-term** domain transformation (radar to ECG) lacks theoretical explanation.
2. Noise robustness is **never** considered or validated in ECG recovery, especially for the DL model itself.
3. The **misalignment issue** in the method based on single cardiac cycle.

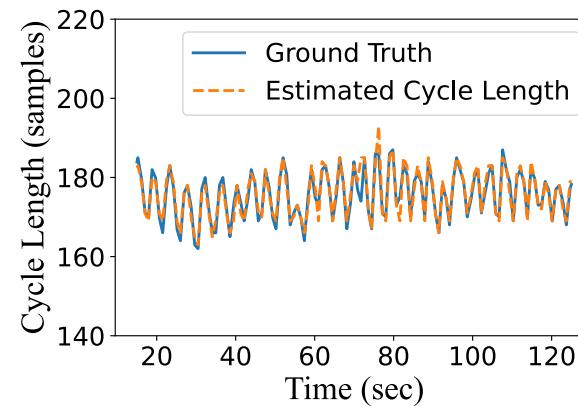


Contributions:

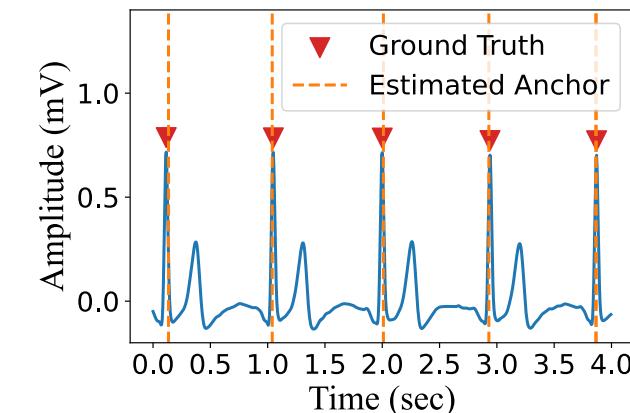
1. Model the radar-based ECG reconstruction **into three tasks** and design an MTL framework named radarODE-MTL to avoid the reintroduction of the time-domain noises and hence improve the noise robustness.
2. A novel optimization strategy **eccentric gradient alignment (EGA)** is proposed for MTL learning by **eccentrically aligning the task gradients**, aiming to **adaptively alter the task priority** to ensure efficient learning progress for the tasks with disparate complexity while **preventing the negative transfer phenomenon** at the same time.

Three Tasks:

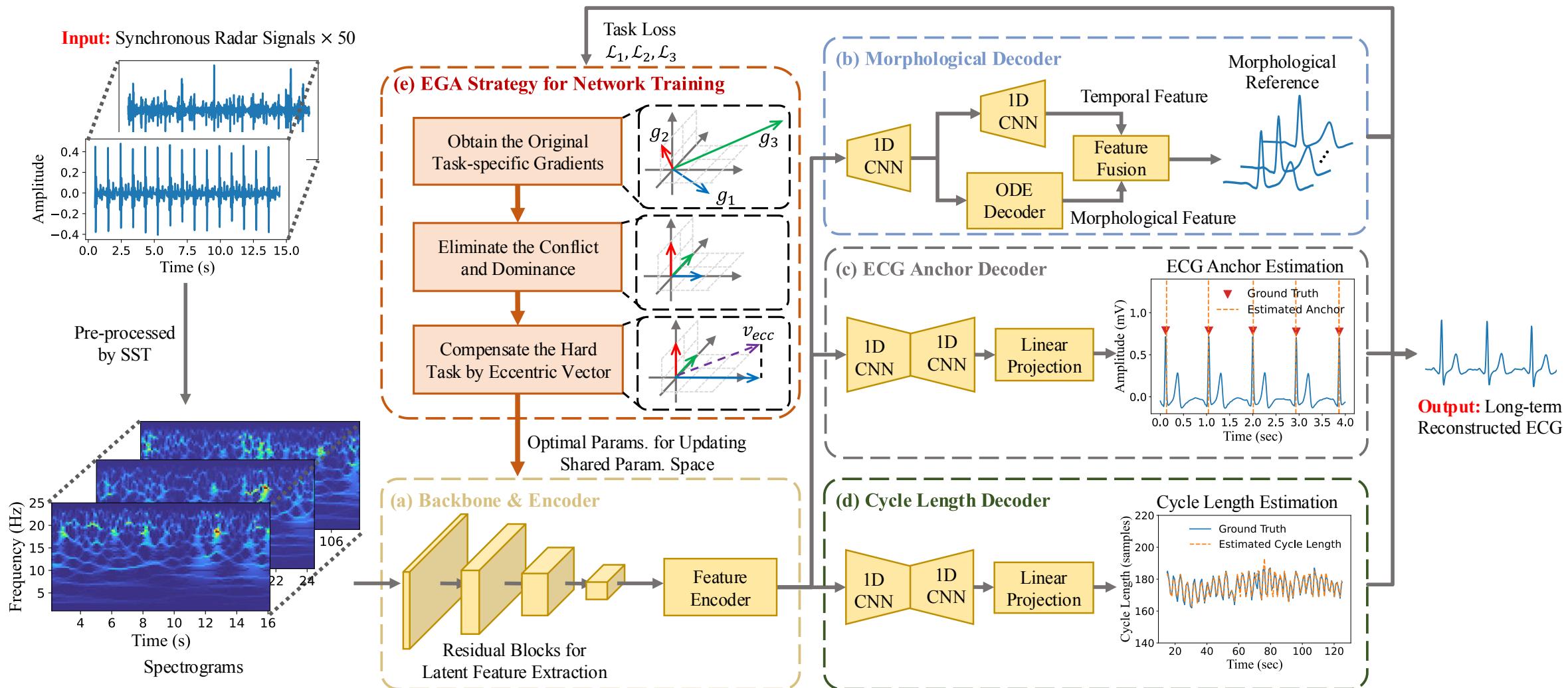
Morphological Recovery

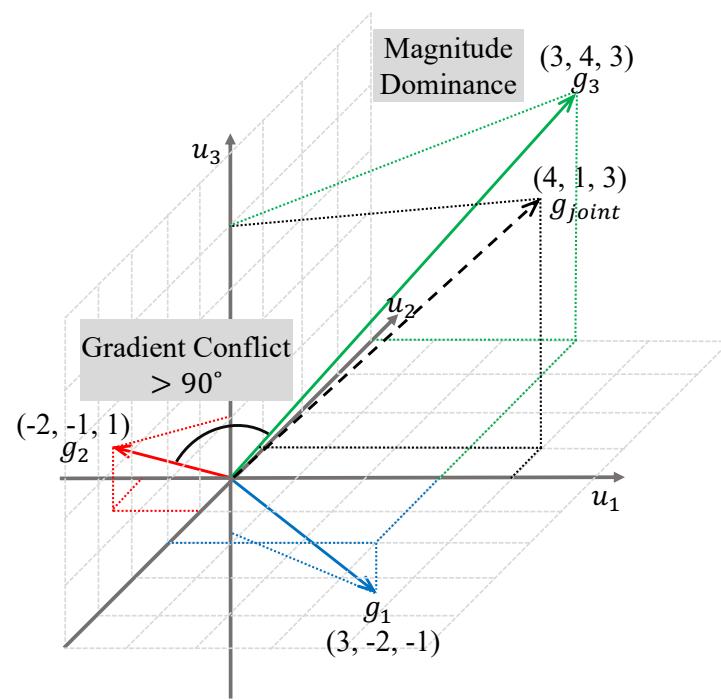


Cycle Length Estimation

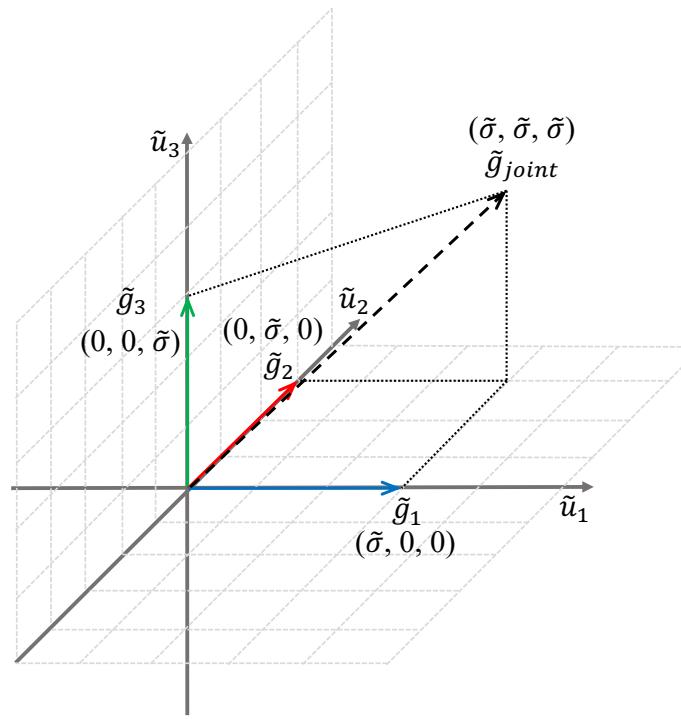


ECG Anchor Estimation

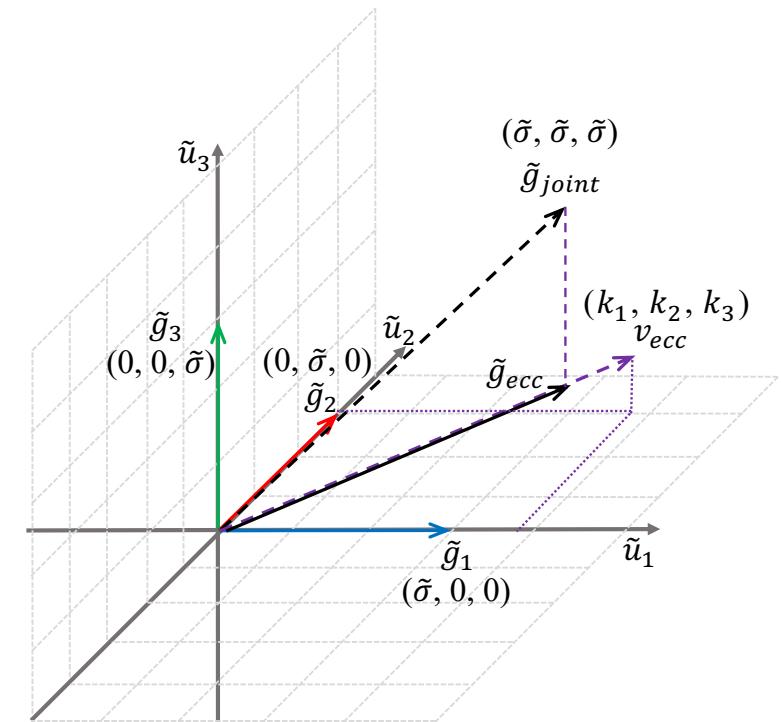




(a) Original gradient space with **gradient conflict** and **magnitude dominance**

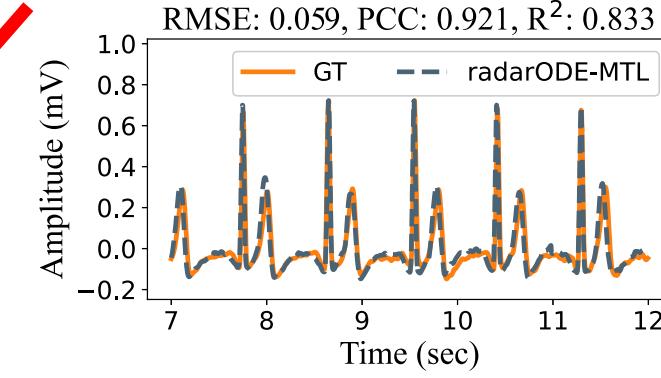
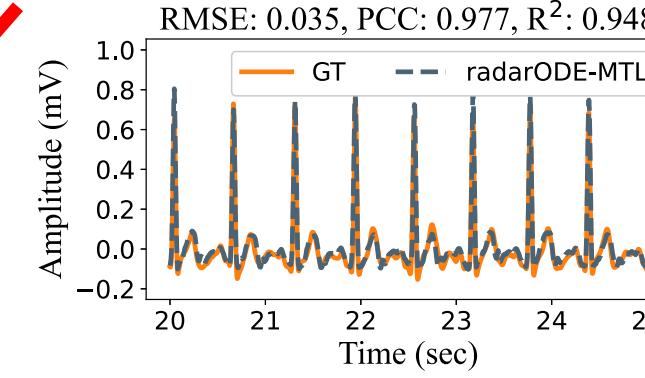
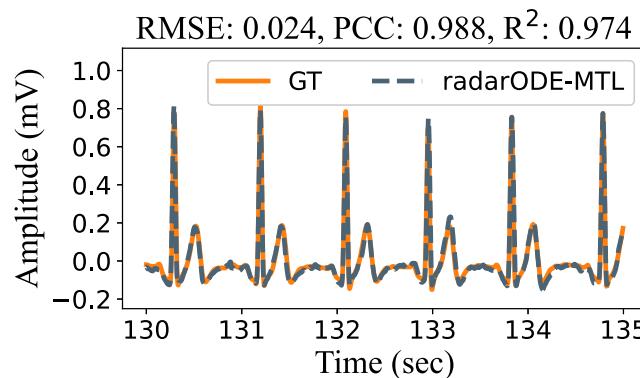
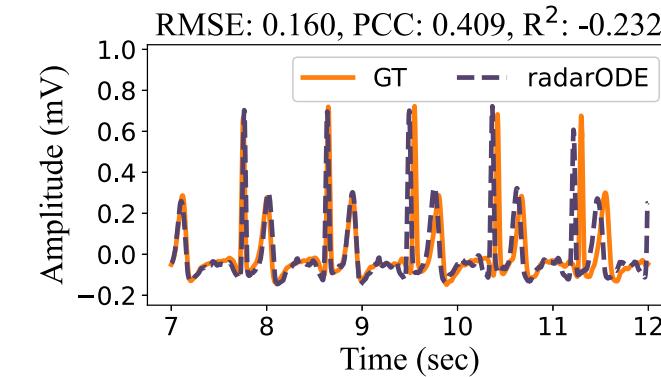
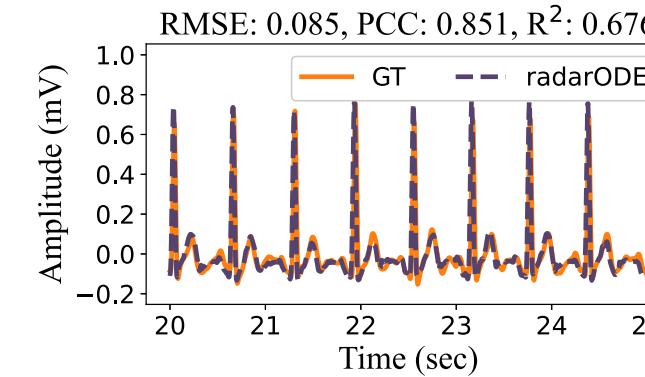
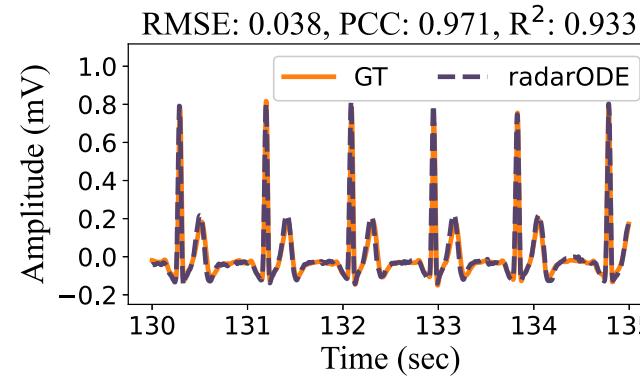
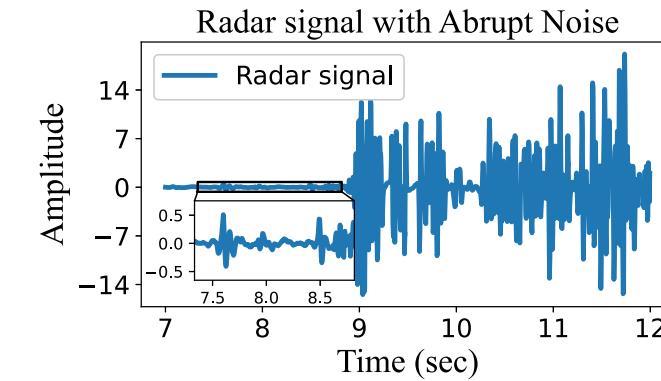
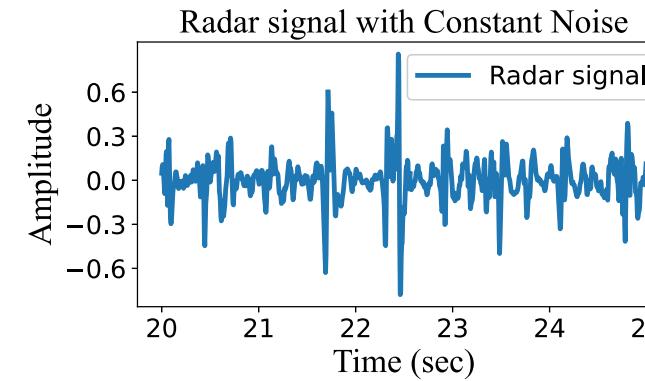
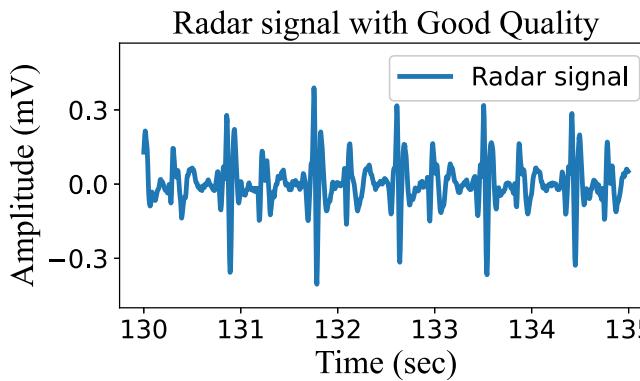


(b) The projection of the original gradient space into the **orthogonal** space with equal “learning rate”

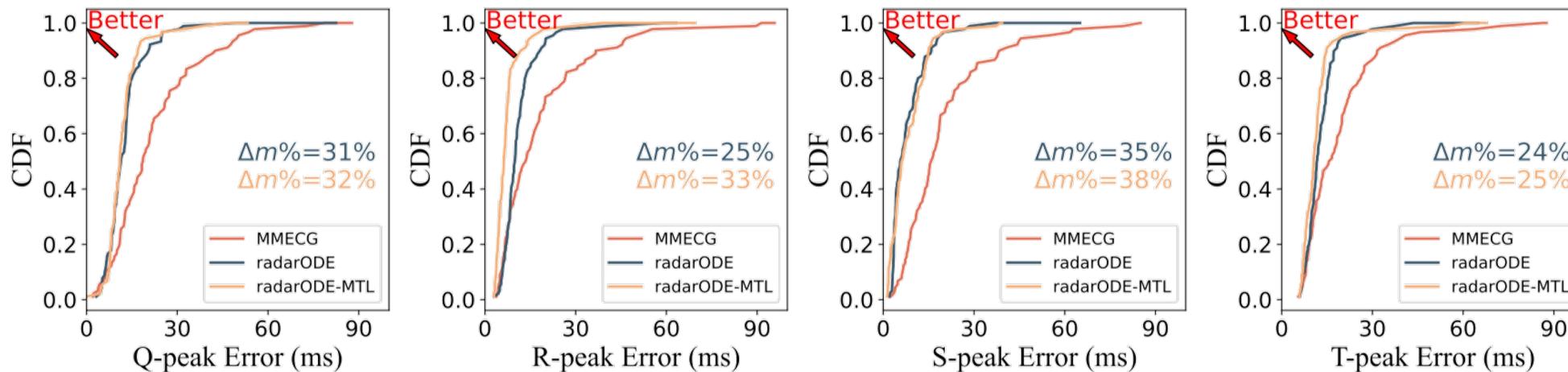


(c) The implementation of eccentric gradient alignment **to slant the joint gradient** to the hard task by introducing the **eccentric vector**

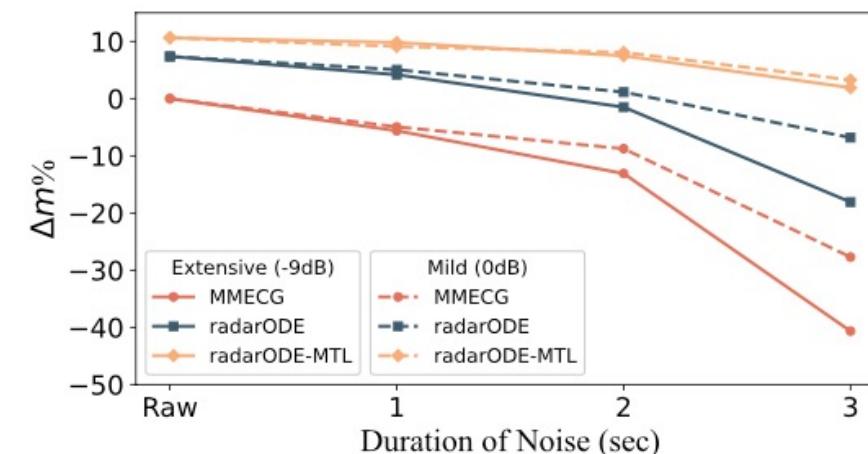
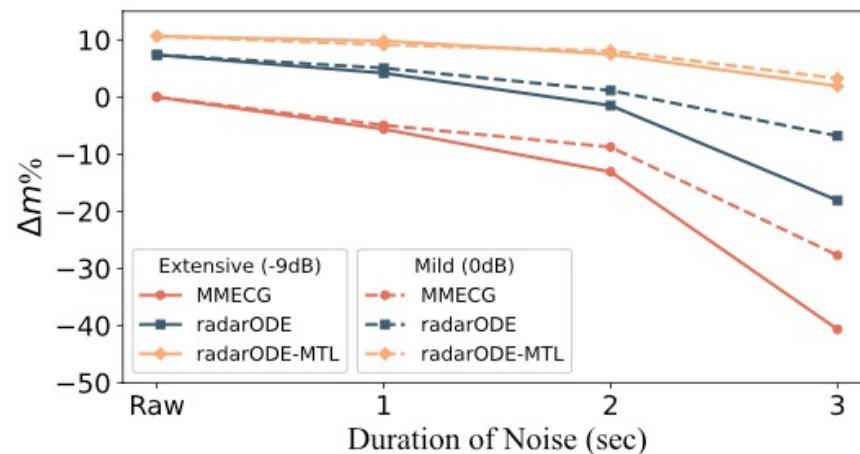
Robust Long-term ECG Generation – Results



Robust Long-term ECG Generation – Results



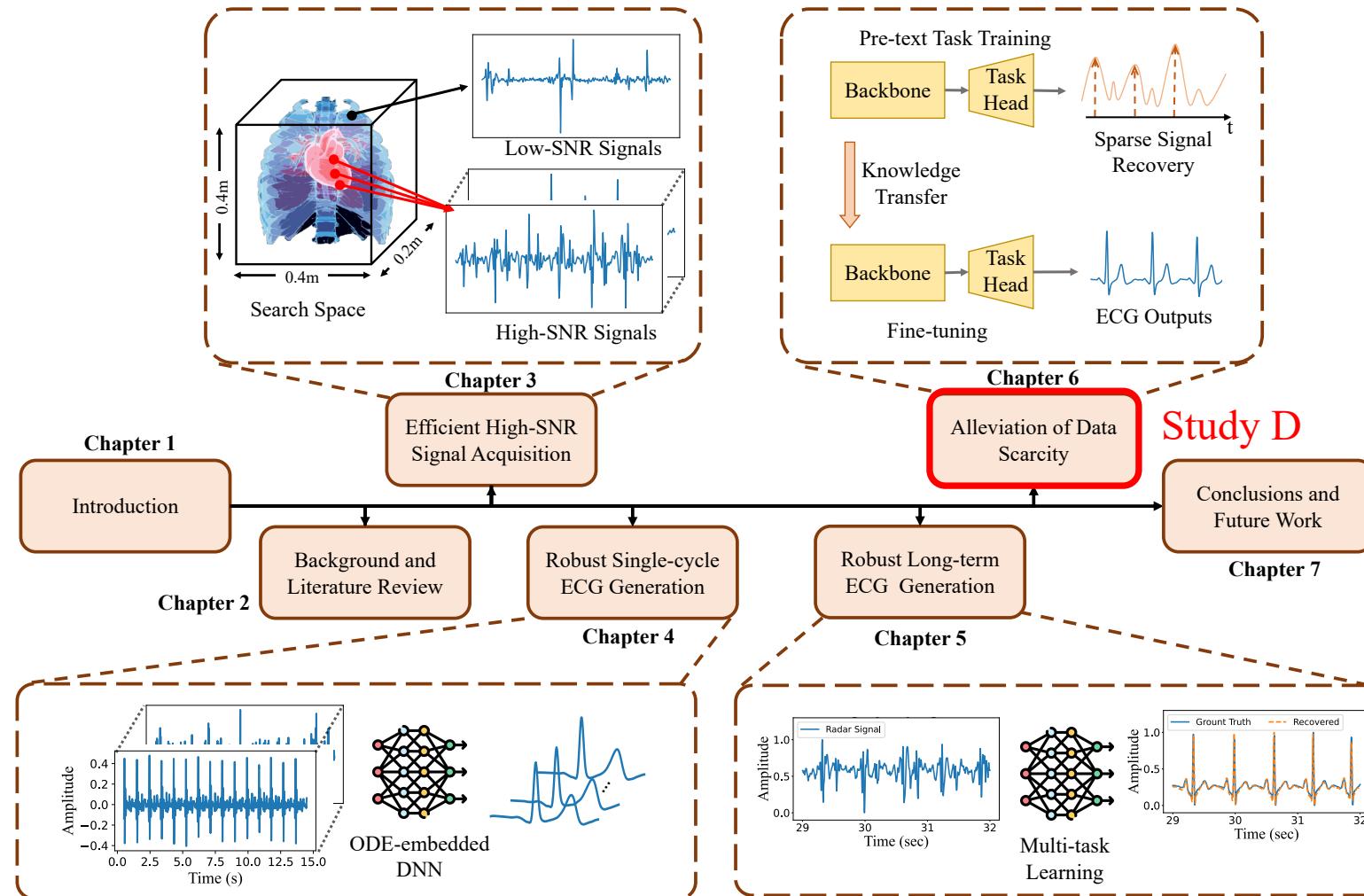
Peak accuracy



Noise robustness test

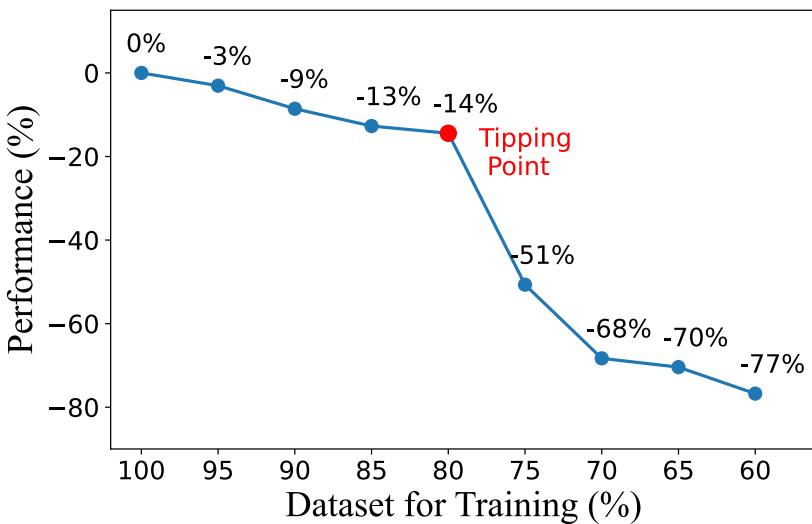


Part 2 – Proposed Methods

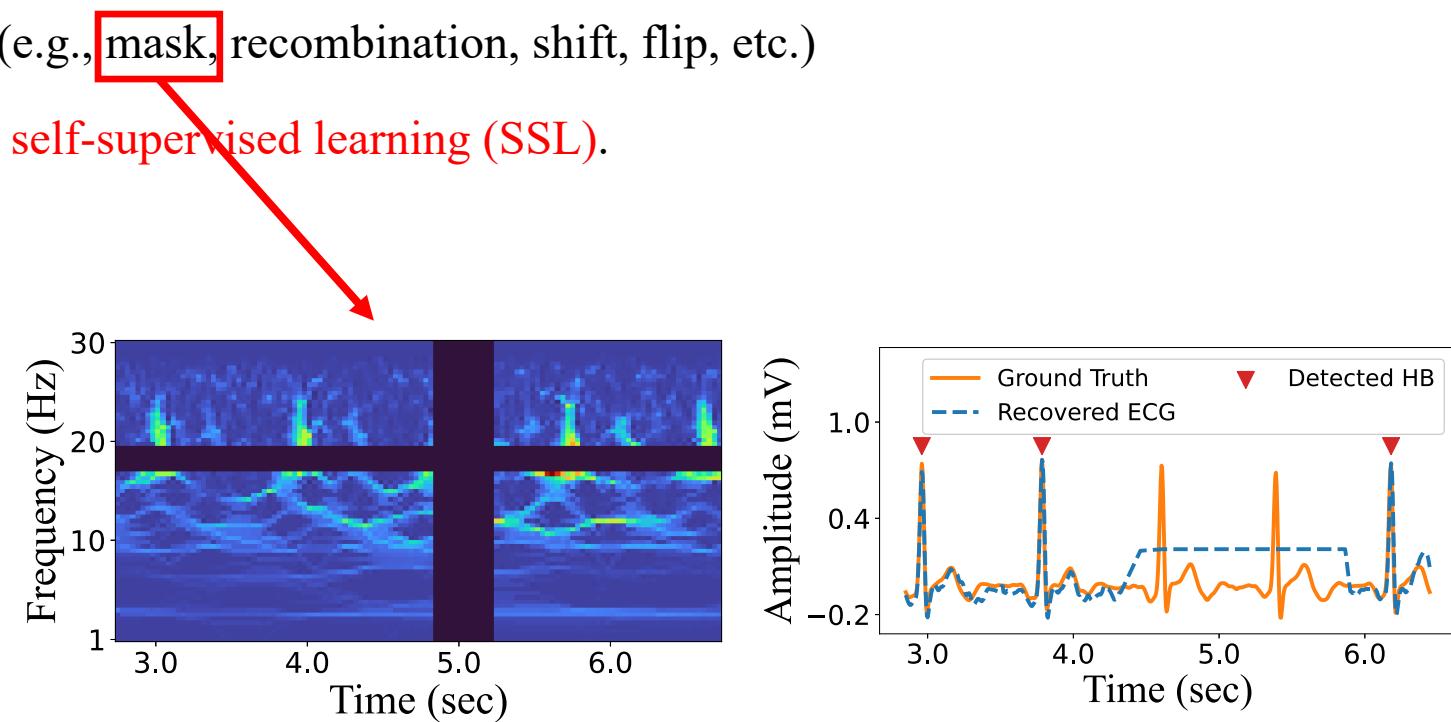


Research gap in radar-based ECG recovery:

1. Limited data for deep learning model training, especially for **new scenarios**.
2. **Unfaithful data augmentation** for regression task. (e.g., **mask**, recombination, shift, flip, etc.)
3. Do not make full use of **unlabeled** radar signal for **self-supervised learning (SSL)**.



Limited data may cause significant degradation



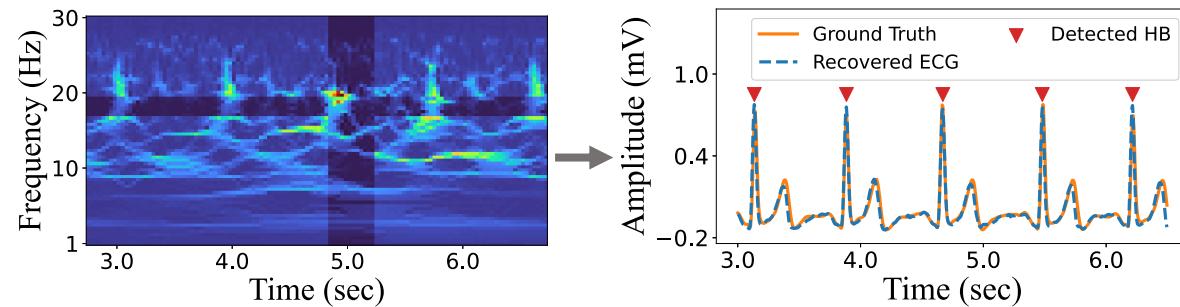
Traditional zero mask may cause missed detection



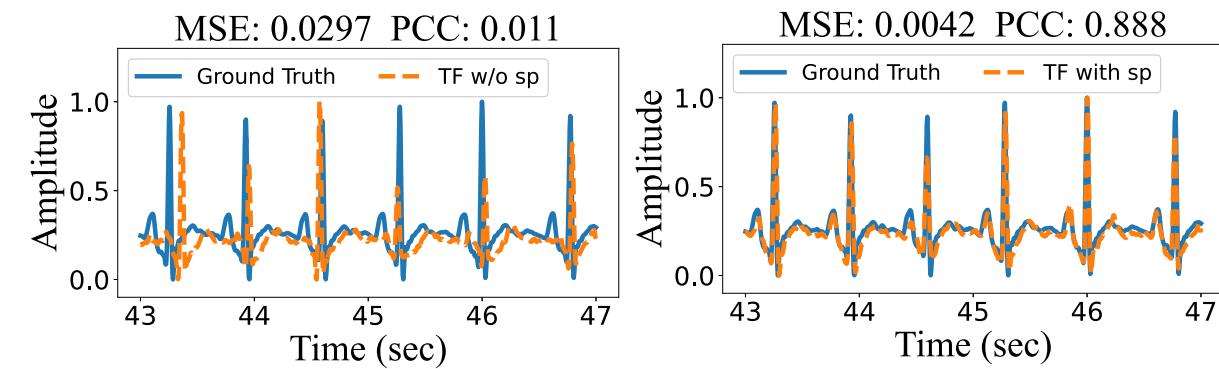
Alleviate Data Scarcity – Contributions

Contributions:

1. Propose **Horcrux** enhances the **diversity** of the limited dataset while preserving the **intrinsic time consistency** in the original signal and **restricting the potential distribution shifts** introduced into the augmented dataset.
2. Propose a transfer learning framework **RFcardi** following an SSL paradigm to effectively learn **the latent representations** from radar signals by leveraging an appropriate **pre-text task**.

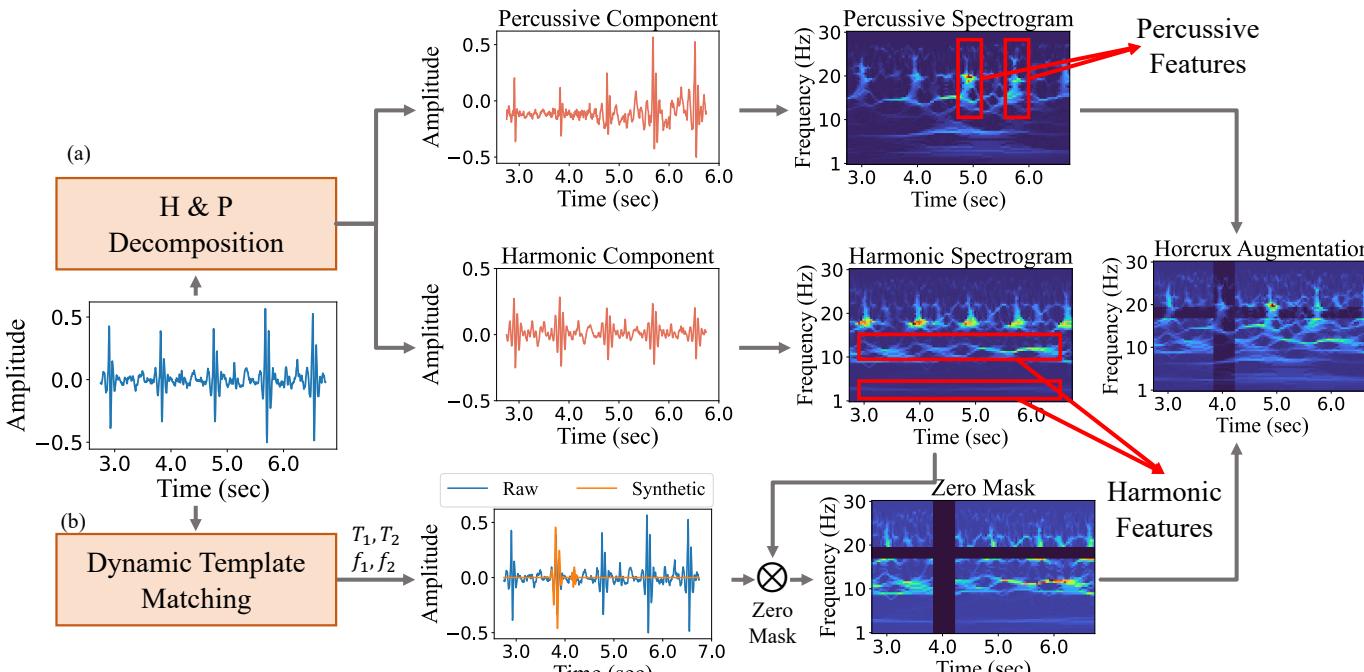


Time-domain consistency preserved using proposed Horcrux

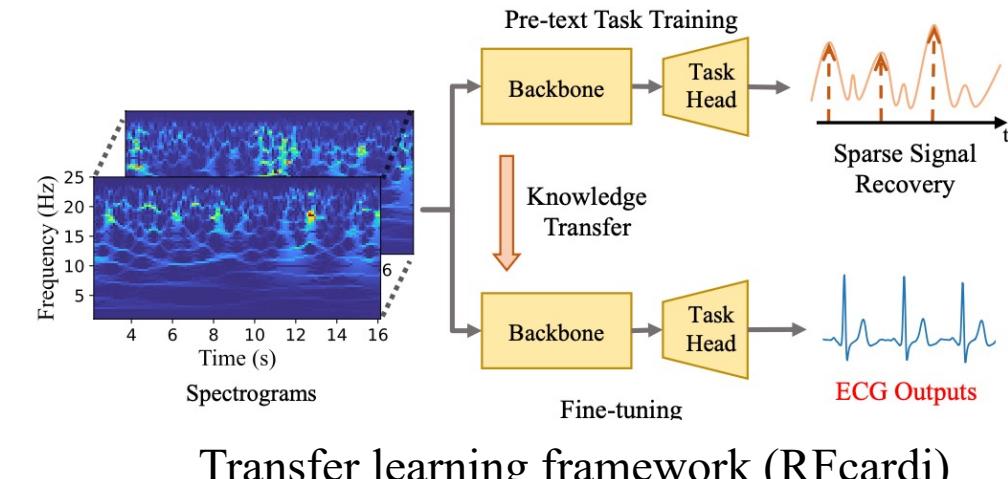


ECG Recovery w/o or with pre-training

Alleviate Data Scarcity – Horcrux and RFcardi



Data augmentation pipeline (Horcrux)



Transfer learning framework (RFcardi)

SSL with pre-text Task

1. Sparse Signal Recovery

$$h = \Phi x + n$$

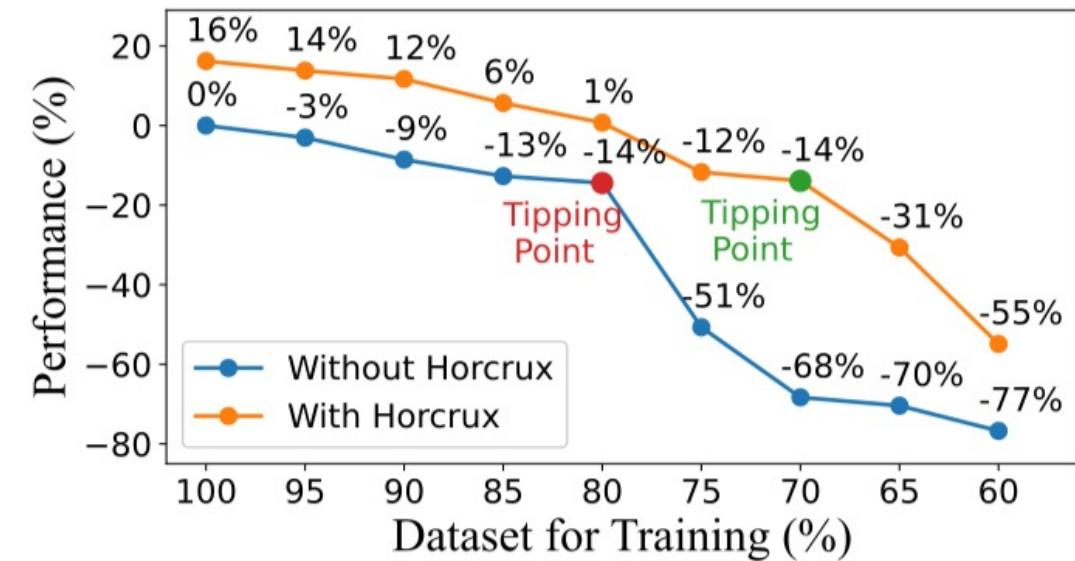
2. Loss function with sparse penalty from $[0, \lambda_s]$

$$\mathcal{L} = \|x - x'\|_2 + \lambda_s \underbrace{\frac{\|x\|_1/\|x\|_2 - 1}{\sqrt{m} - 1}}_{\text{sparse penalty}}$$

TABLE I
COMPARISON OF DIFFERENT AUGMENTATION METHODS

Methods	RMSE (mV) ↓	PCC ↑	H. E. (ms) ↓	MDR ↓	$\Delta m\%$
Baseline	0.096	82.65%	8.82	6.73%	0.0%
C-Mixup [13]	<u>0.088</u>	82.25%	7.55	6.14%	7.80%
ADA [8]	<u>0.097</u>	81.28%	7.20	6.69%	4.12%
RC-Mixup [14]	<u>0.089</u>	<u>83.36%</u>	7.72	5.79%	8.69%
Horcrux (10%)	0.096	82.24%	6.89	6.67%	5.63%
Horcrux (15%)	0.087	<u>84.75%</u>	6.96	5.19%	<u>14.02%</u>
Horcrux (20%)	0.086	85.41%	<u>6.21</u>	<u>5.30%</u>	16.20%
Horcrux (25%)	0.092	81.99%	6.16	6.36%	9.81%
Horcrux (30%)	0.094	80.18%	6.55	6.59%	6.78%

Bold and underline represent the best and the second best results.



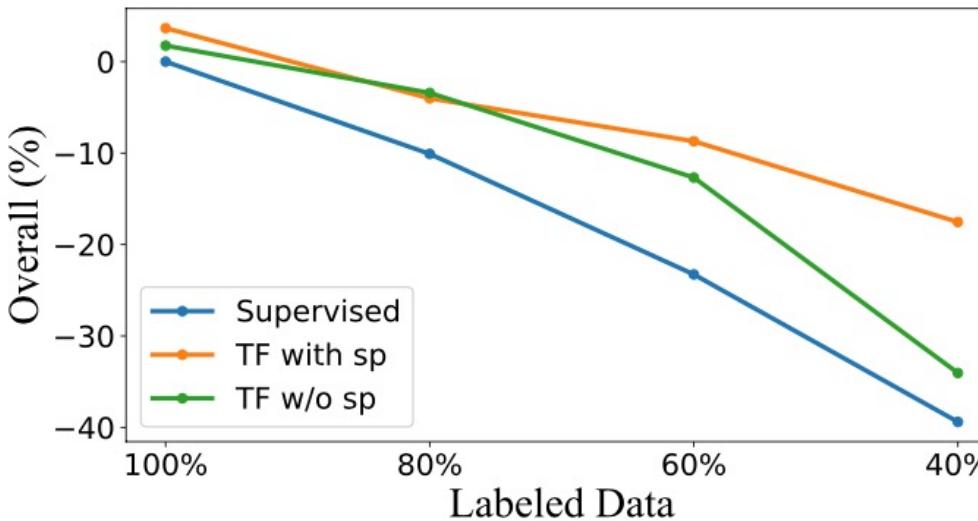


TABLE IV
PERFORMANCE OF ECG RECOVERY USING DIFFERENT PERCENTAGES OF LABELED DATA

Methods	MSE ($\times 10^{-2}$) ↓	PCC ↑	Peak Error (ms) ↓	MDR ↓	Overall ↑	MSE ($\times 10^{-2}$) ↓	PCC ↑	Peak Error (ms) ↓	MDR ↓	Overall ↑
	100% Labeled					80% Labeled				
Supervised	0.80	85.47%	7.61	6.85%	0.00%	0.84	84.60%	8.90	8.04%	-10.09%
TF w/o sp*	0.81	85.35%	8.46	5.51%	1.75%	0.82	86.36%	8.35	7.02%	-3.42%
TF with sp	0.80	85.51%	8.40	5.14%	3.66%	0.81	84.29%	8.31	7.14%	-4.02%
60% Labeled					40% Labeled					Overall ↑
Supervised	0.93	79.91%	10.65	8.93%	-23.27%	0.98	75.89%	11.15	12.15%	-39.4%
TF w/o sp	0.85	83.74%	8.84	8.65%	-12.68%	0.97	76.56%	10.87	10.99%	-34.05%
TF with sp	0.86	84.92%	8.58	7.72%	-8.71%	0.93	78.72%	8.70	9.02%	-17.54%

*TF for transfer learning and sp for sparse penalty



Part 3 – Conclusions and Future Work

- **Part 1 – Summary of contributions**
- **Part 3 – Publications**
- **Part 2 – Future work**



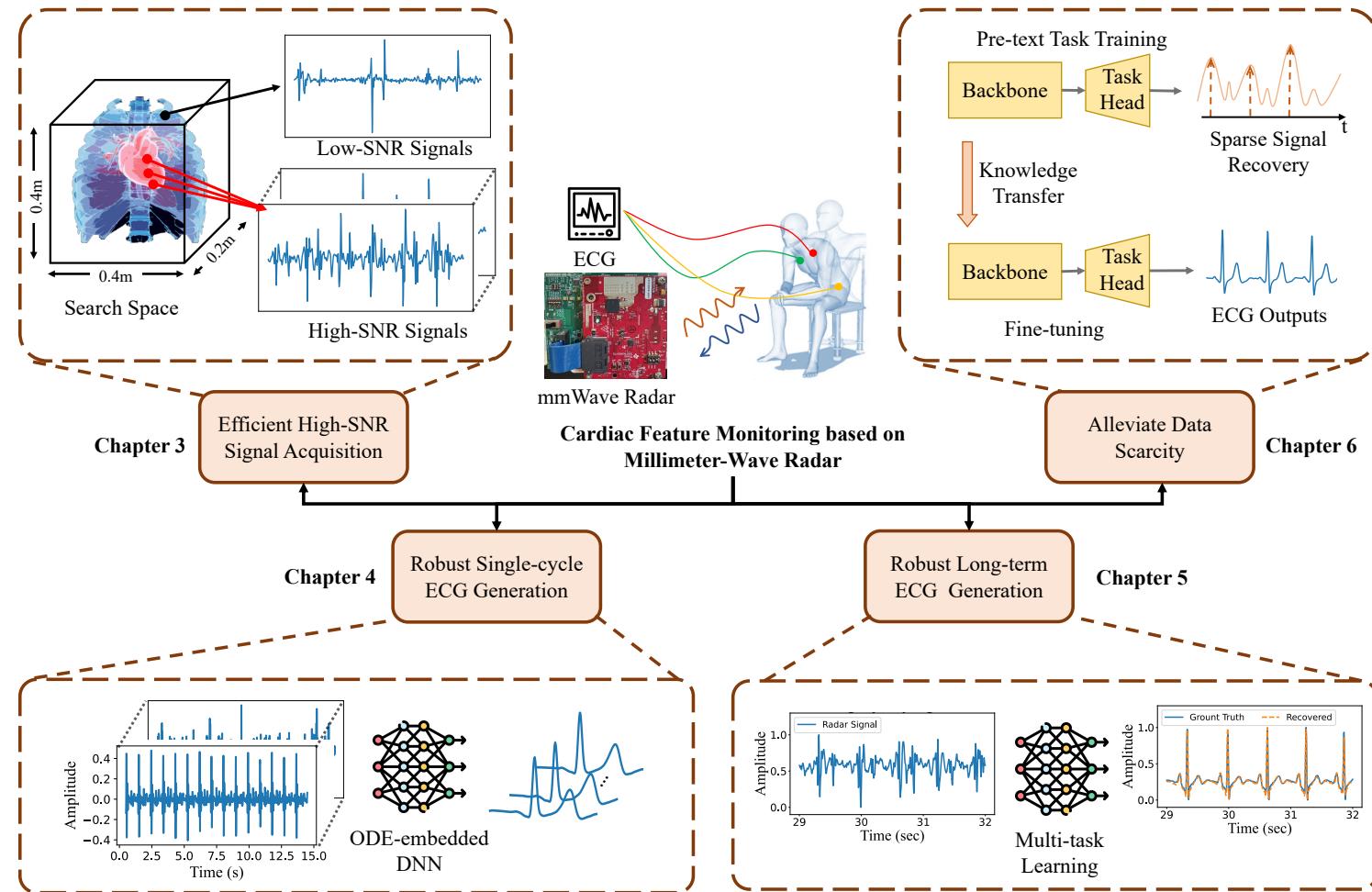
Conclusions



Summary of contributions

Contributions:

1. Efficient data collection with ample cardiac features
2. Model the transformation between fine-grained cardiac features (mechanical domain to electrical domain).
3. Realize the long-term radar-based ECG recovery even under the impact of different noises.
4. Alleviate the data scarcity for the deployment in new scenarios with limited data.





Conclusions

Publications



Publications as first author:

1. **Yuanyuan Zhang**, Rui Yang, Yutao Yue, Eng Gee Lim, Zidong Wang, “An Overview of Algorithms for Contactless Cardiac Feature Extraction From Radar Signals: Advances and Challenges”, *IEEE Transactions on Instrumentation and Measurement*, Jul. 2023. **(Chapter 2)**
2. **Yuanyuan Zhang**, Haocheng Zhao, Sijie Xiong, Rui Yang, Eng Gee Lim, Yutao Yue, “From High-SNR Radar Signal to ECG: A Transfer Learning Model with Cardio-Focusing Algorithm for Scenarios with Limited Data”, *IEEE Transactions on Mobile Computing*. **(Under Major Revision) (Chapter 3)**
3. **Yuanyuan Zhang**, Runwei Guan, Lingxiao Li, Rui Yang, Yutao Yue, Eng Gee Lim, “radarODE: An ODE-Embedded Deep Learning Model for Contactless ECG Reconstruction from Millimeter-Wave Radar”, *IEEE Transactions on Mobile Computing*, Apr. 2025. **(Chapter 4)**
4. **Yuanyuan Zhang**, Rui Yang, Yutao Yue, Eng Gee Lim, “radarODE-MTL: A Multi-Task Learning Framework with Eccentric Gradient Alignment for Robust Radar-Based ECG Reconstruction”, *IEEE Transactions on Instrumentation and Measurement*, Apr. 2025. **(Chapter 5)**
5. **Yuanyuan Zhang**, Sijie Xiong, Rui Yang, Eng Gee Lim, Yutao Yue, “Recover from Horcrux: A Spectrogram Augmentation Method for Cardiac Feature Monitoring from Radar Signal Components”, in *Proceedings of the 47th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC2025)*, Jul. 2025. **(Chapter 6)**



Conclusions



Publications

Publications as co-author:

6. Sijie Xiong, **Yuanyuan Zhang**, Cheng Tang, Haoling Xiong, Yiding Li, Atsushi Shimada, “U-MA: A Unified Framework with Differential Mamba under Parallel U-Net Scheme for Time Series Forecasting”, *Engineering Applications of Artificial Intelligence*. (Under Review)
7. Sijie Xiong, Cheng Tang, **Yuanyuan Zhang**, Haoling Xiong, Youhao Xu, Atsushi Shimada, “CME-Mamba with Enhancing Nonlinear Dependencies for Time Series Forecasting”, *Applied Soft Computing*. (Under Review)
8. Sijie Xiong, **Yuanyuan Zhang**, Cheng Tang, Fumiya Okubo, Yinlong Hu, Atsushi Shimada, “PIDTracker: An Efficient 2D Controllable Holt-Winters Model for Glucose Monitoring”, in *Proceedings of the AAAI Conference on Artificial Intelligence*. (Under Review)
9. Sijie Xiong, Jianing Wang, **Yuanyuan Zhang**, Cheng Tang, Fumiya Okubo, Atsushi Shimada, “GlucoMixer: An Efficient Continuous Glucose Monitoring Model with Mixers”, in *Proceedings of the AAAI Conference on Artificial Intelligence*. (Under Review)



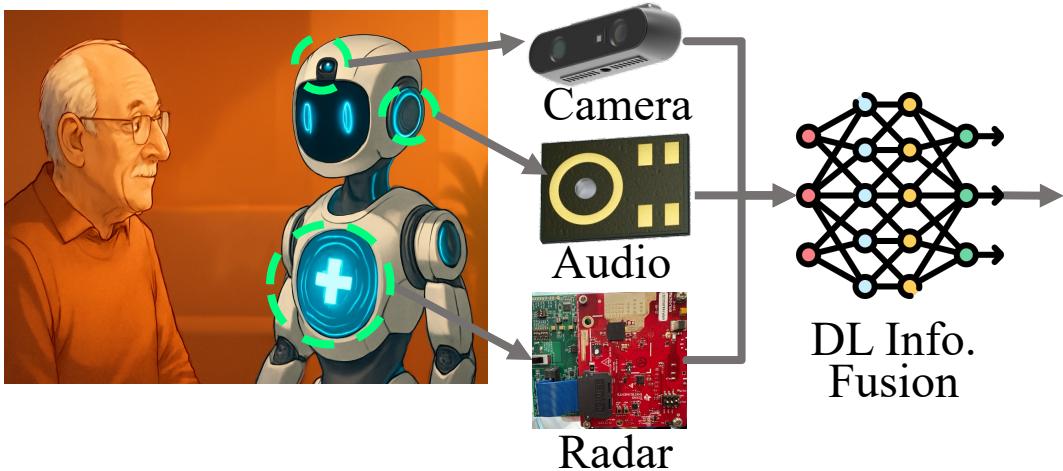
Conclusions



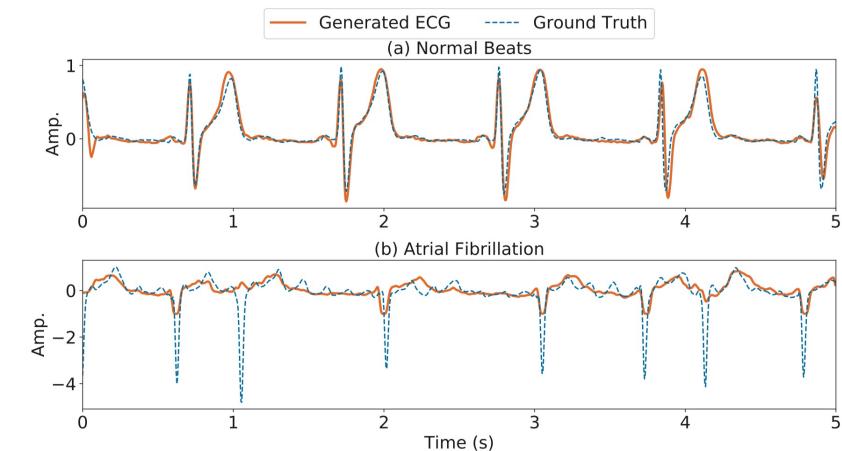
Conclusions – Future work

Future work:

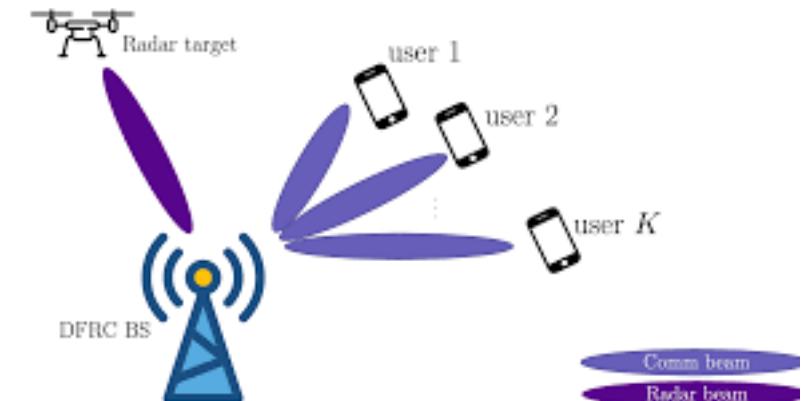
1. Multi-sensor fusion for advanced healthcare monitoring.
2. Evaluation on the dataset for patients.
3. Integrated sensing and communication (for 6G).



Multi-sensor system for healthcare monitoring



ECG signal for patients



ISCA



Thanks for your time!



Supplementary Materials – Chapter 3

Assess signal SNR with universal template

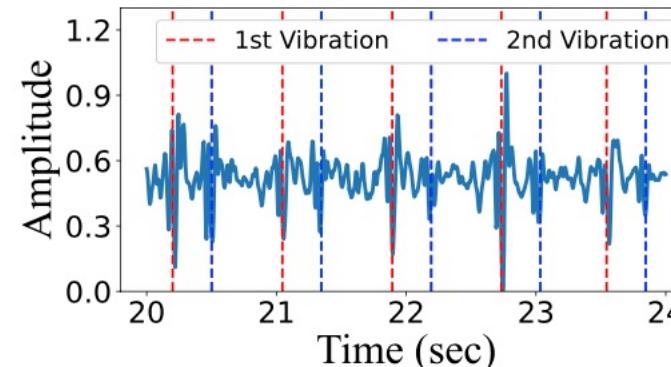
1. Propose the dynamic template

$$h_m(t) = a_1 \exp\left(-\frac{(t - b_1)^2}{2c_1^2}\right) + a_2 \exp\left(-\frac{(t - b_2)^2}{2c_2^2}\right)$$

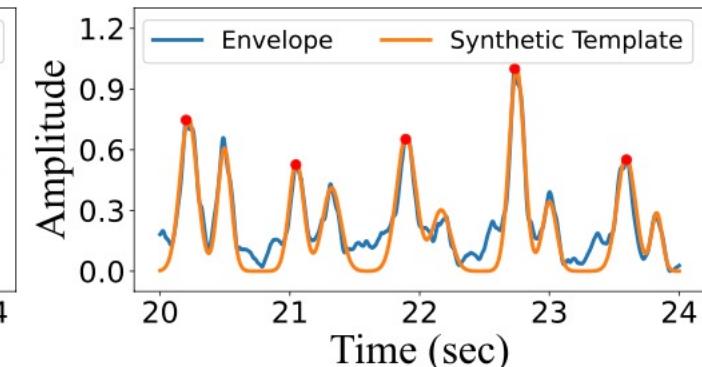
$a_1, a_2 \rightarrow$ amplitude of the vibrations

$b_1, b_2 \rightarrow$ when the vibrations happen

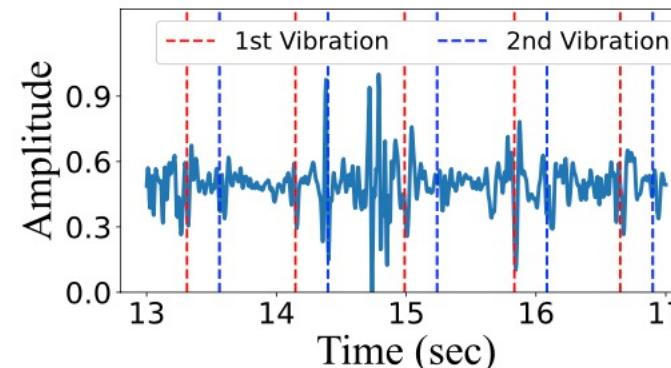
$c_1, c_2 \rightarrow$ length/width of the peak



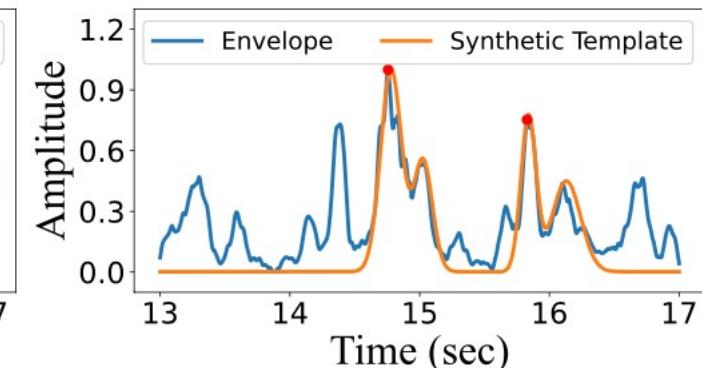
(a)



(b)



(c)



(d)

2. Identify incumbent peaks (red dots)

Peak position and amplitude $\rightarrow a_1, b_1$

3. Run template matching

$$\theta^* = \arg \min_{\theta=\{a_2, b_2, c_1, c_2\}} \|h(E, t) - h_m(t, \theta)\|_2$$



Supplementary Materials – Chapter 3

Range and Angle FFT

1. Received FMCW signal

$$x_R(t) = A_R \exp \left\{ -j \left[2\pi f_c(t - t_d) + \pi \frac{B}{T_c} (t - t_d)^2 \right] \right\}$$

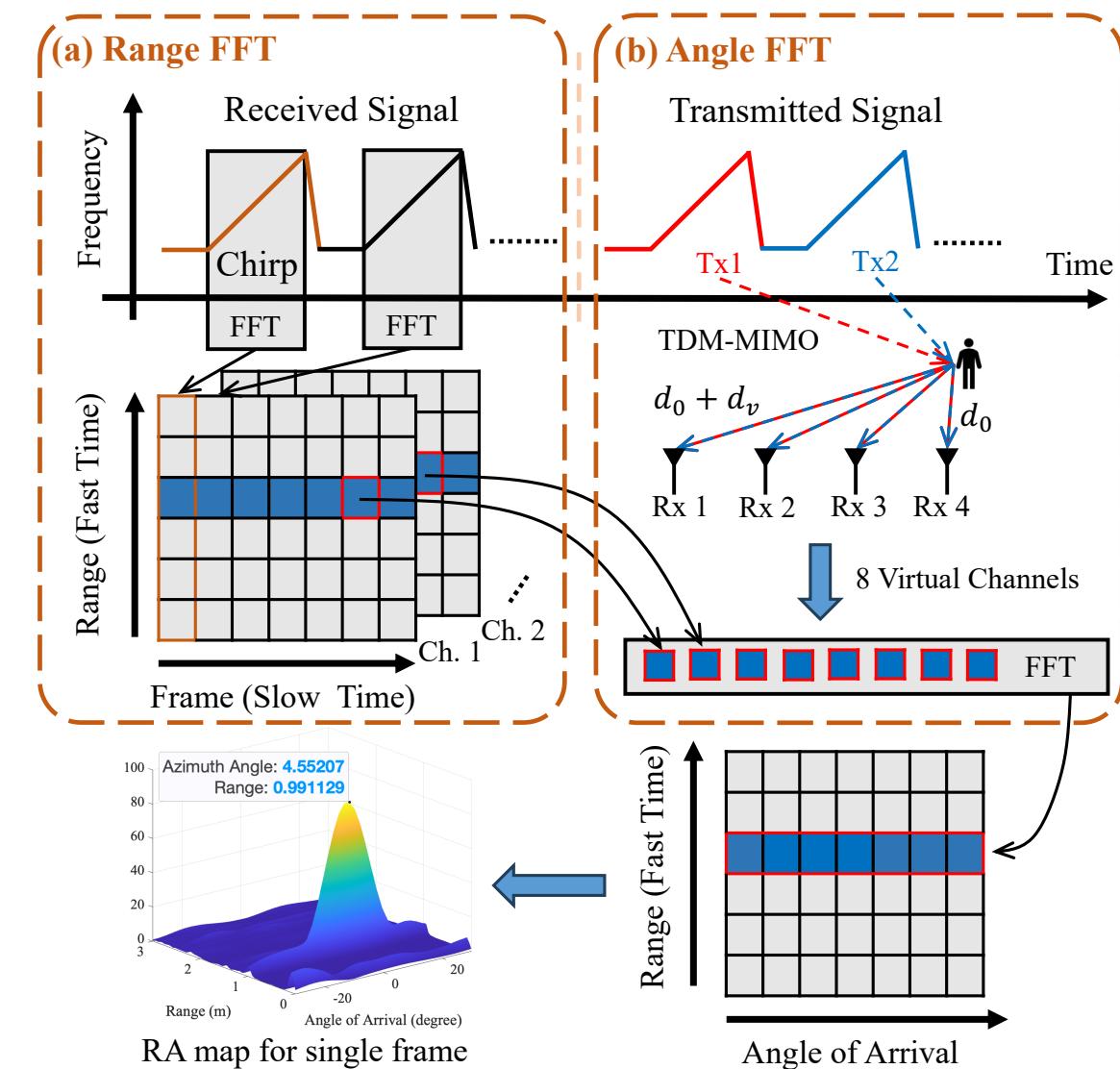
2. Radar-human distance introduce phase shift

$$\phi_s = \frac{4\pi d_0}{\lambda}$$

3. Distance between antennas introduce extra phase shift w.r.t. angle of arrival θ (AoA)

$$\Delta\phi_v = \frac{4\pi d_v}{\lambda}$$

$$d_v = l \sin(\theta)$$



Objective

$$E_b = \arg \min_{E \in \mathbb{R}^n} \{\mathcal{F}(E) : E \in \Omega\} \quad \text{for each iteration } k$$

Initialization

- ### 1. Initialization of grid G_k

$$G_k := \{E_k + \gamma_k D\} \subset \mathbb{R}^n$$

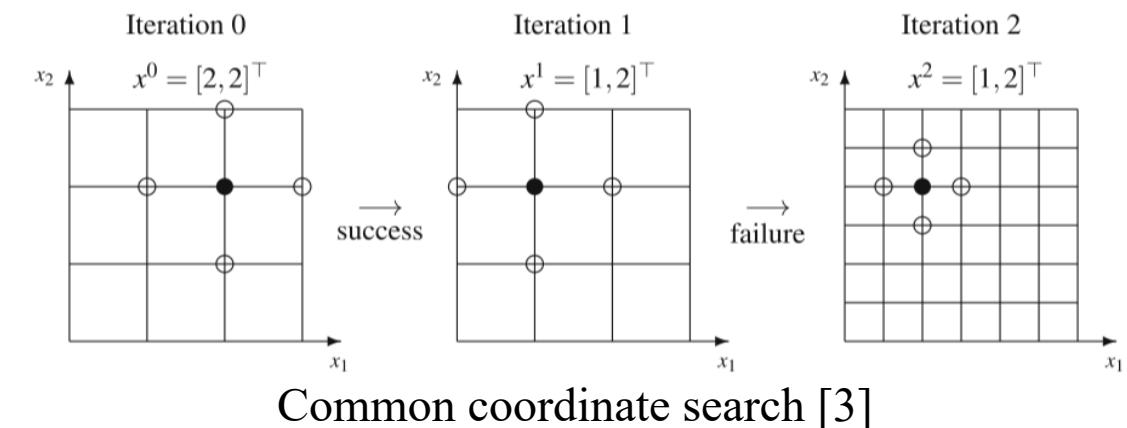
with grid size parameter γ_k and search direction set D

- ## 2. Introduce search region S_k

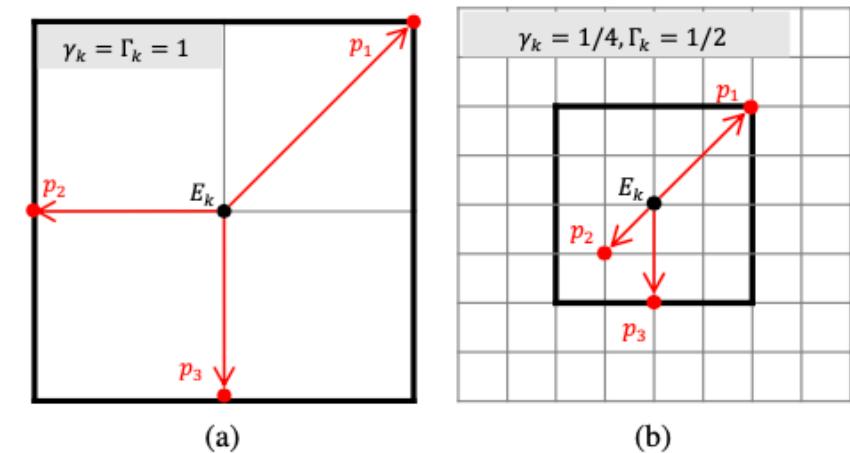
$$S_k := \{E \in G_k : \|E - E_k\|_\infty \leq \Gamma_k a\}$$

$$a = \max \{ \|a'\|_\infty : a' \in D\}$$

with search size parameter Γ_k



Common coordinate search [3]



Proposed CFT Algorithm

[3] J. Larson, M. Menickelly, and S. M. Wild, “Derivative-free optimization methods,” *Acta Numerica*, vol. 28, pp. 287–404, Jun. 2019.



Supplementary Materials – Chapter 3

Main Iteration

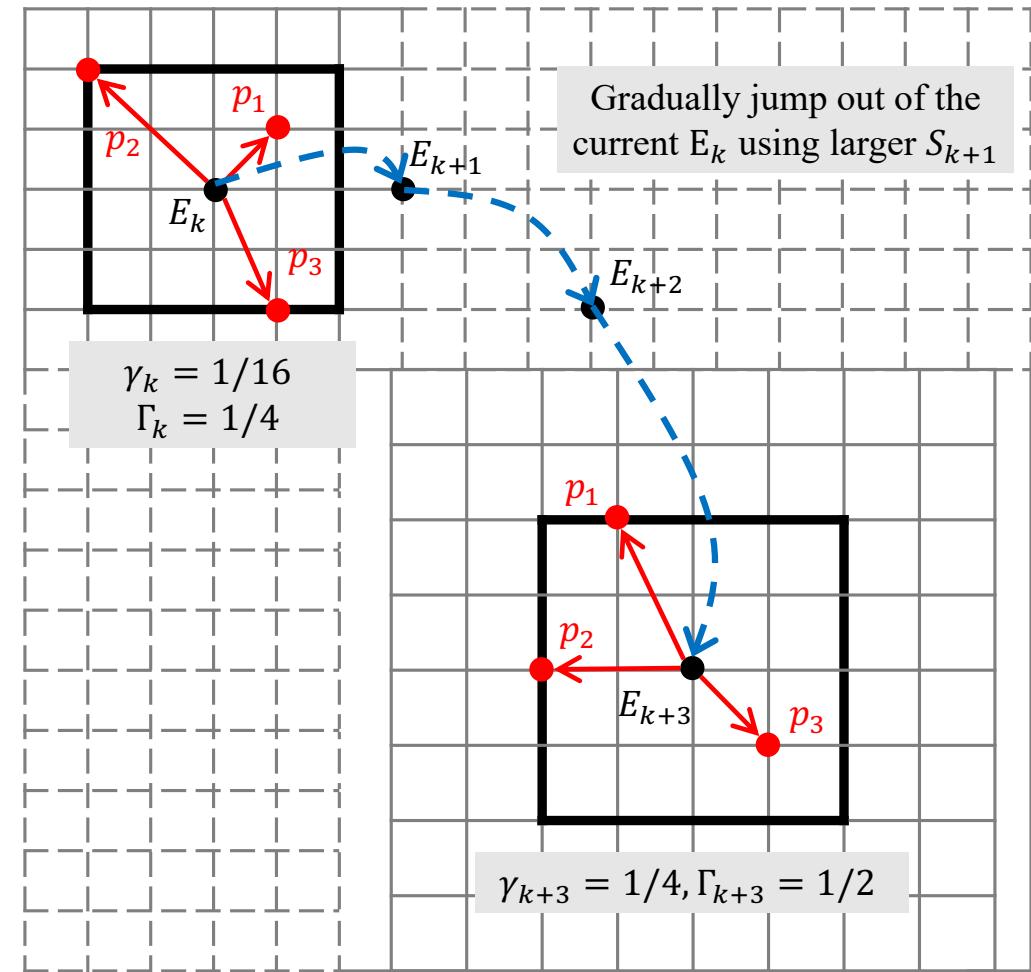
1. Initial Search Stage

Evaluate $\mathcal{F}(E)$ on a subset of grids G_k based on any sampling algorithm (Latin hypercube sampling)

2. Resizing Stage

- If success, move to new point and double Γ_k
- If failed, run the same search inside search region S_k
 - If success, move to new point and double Γ_k
 - If failed, half Γ_k

$$\text{Set } \gamma_{k+1} \leftarrow \min(\Gamma_k, \Gamma_k^2)$$





Supplementary Materials – Chapter 5

Define a MTL problem for n tasks:

$$\theta^* = \arg \min_{\theta \in \mathbb{R}^m} \left\{ \mathcal{F}(\theta) \triangleq \frac{1}{n} \sum_{i=1}^n \mathcal{L}_i(\theta) \right\} : \theta \in \mathbb{R}^m \text{ for shared parameter space}$$

The original task-specific gradient and the gradient matrix :

$$\mathbf{g}_i = \nabla_{\theta} \mathcal{L}_i(\theta), i \in [n] \quad \mathbf{G} = \{\mathbf{g}_1, \dots, \mathbf{g}_i\} \in \mathbb{R}^{n \times m}$$

Project to orth. space

1. Formulate the projection problem:

$$\min \|\mathbf{g}_{joint} - \tilde{\mathbf{g}}_{joint}\|_2^2 \text{ s.t. } \tilde{\mathbf{G}}\tilde{\mathbf{G}}^\top = \mathbf{I} \quad \text{with} \quad \begin{aligned} \mathbf{g}_{joint} &= \mathbf{G}^\top \mathbf{w} \\ \tilde{\mathbf{g}}_{joint} &= \tilde{\mathbf{G}}^\top \mathbf{w} \end{aligned} \quad \mathbf{w} = [1, \dots, 1]^\top$$

Then we have:

$$\|\mathbf{g}_{joint} - \tilde{\mathbf{g}}_{joint}\|_2^2 = \|\mathbf{G}^\top \mathbf{w} - \tilde{\mathbf{G}}^\top \mathbf{w}\|_2^2 \leq \|\mathbf{G}^\top - \tilde{\mathbf{G}}^\top\|_F^2 \|\mathbf{w}\|_2^2$$

2. The final projection problem:

$$\min_{\tilde{\mathbf{G}}} \|\mathbf{G} - \tilde{\mathbf{G}}\|_F^2 \quad \text{s.t. } \tilde{\mathbf{G}}\tilde{\mathbf{G}}^\top = \mathbf{I}$$

Algorithm 1 EGA Optimization Strategy for MTL

- 1: **Input:** Loss values for n tasks $\{\mathcal{L}_1, \dots, \mathcal{L}_i\}, i \in [n]$, T for softmax and t_{warm} for warmup epoch
INITIALIZATION:
 - 2: - Unit eccentric vector $\mathbf{v}_{ecc} = [1, \dots, 1]^\top \in \mathbb{R}^n$
FOR EACH ITERATION:
 - 3: - Get the current epoch as t
 - 4: - Get task-specific gradient $\mathbf{g}_i = \nabla_{\theta} \mathcal{L}_i(\theta), i \in [n]$
 - 5: - Form gradient matrix $\mathbf{G} = \{\mathbf{g}_1, \dots, \mathbf{g}_i\} \in \mathbb{R}^{n \times m}$
 - 6: - Eigen decomposition of Gram matrix as in (13):
 $\mathbf{G}\mathbf{G}^\top = \mathbf{U}(\Sigma\Sigma^\top)\mathbf{U}^\top$ with eigenvalues λ
 - 7: - Get scaling factor for gradient normalization:
 $\tilde{\sigma} = \min(\sqrt{\lambda})$
 - 8: - Calculated the orthogonal and normalized gradient matrix as in (15): $\tilde{\mathbf{G}} = \tilde{\sigma}\mathbf{U}\Sigma^{-1}\mathbf{U}^\top\mathbf{G}$
 - 9: **if** $t = t_{warm}$ **then**
 - 10: - Record the loss values for all the tasks $\mathcal{L}_i(t_{warm})$
 - 11: **else if** $t > t_{warm}$ **then**
 - 12: - Calculate the task difficulty weights as in (17):
 $k_i(t) = \text{softmax}(lr_i(t-1))$
 - 13: - Form eccentric vector $\mathbf{v}_{ecc} = [k_1, \dots, k_i]^\top$
 - 14: **end if**
 - 15: **Output:** $\tilde{\mathbf{g}}_{ecc} = \tilde{\mathbf{G}}^\top \mathbf{v}_{ecc}$ for optimization
-



Supplementary Materials – Chapter 5

3. Find the solution for:

$$\min_{\tilde{G}} \|\mathbf{G} - \tilde{\mathbf{G}}\|_F^2 \quad \text{s.t. } \tilde{\mathbf{G}}\tilde{\mathbf{G}}^\top = \mathbf{I}$$

4. Borrow the solution of orthogonal Procrustes problem using SVD:

$$\begin{aligned} \min & \quad \|\mathbf{A}\Omega - \mathbf{B}\|_F \quad \mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n} \\ \text{s.t.} & \quad \Omega^\top \Omega = \mathbf{I} \quad \mathbf{C} = \mathbf{B}^\top \mathbf{A} \end{aligned} \quad \text{with solution} \quad \mathbf{U}\Sigma\mathbf{V}^\top = \mathbf{C}$$
$$\Omega = \mathbf{V}\mathbf{U}^\top$$

5. The new gradient matrix (using eigen decomposition) is:

$$\tilde{\mathbf{G}} = \mathbf{U}\mathbf{V}^\top \quad \text{with} \quad \mathbf{G}\mathbf{G}^\top = \mathbf{U}(\Sigma\Sigma^\top)\mathbf{U}^\top$$

6. The final solution is:

$$\tilde{\mathbf{G}} = \mathbf{U}\Sigma^{-1}\mathbf{U}^\top \mathbf{G} \quad \text{with unit singular values}$$

The gradient conflict resolved \square

Algorithm 1 EGA Optimization Strategy for MTL

- 1: **Input:** Loss values for n tasks $[\mathcal{L}_1, \dots, \mathcal{L}_i], i \in [n]$, T for softmax and t_{warm} for warmup epoch
INITIALIZATION:
 - 2: - Unit eccentric vector $\mathbf{v}_{ecc} = [1, \dots, 1]^\top \in \mathbb{R}^n$
 - 3: FOR EACH ITERATION:
 - 4: - Get the current epoch as t
 - 5: - Get task-specific gradient $\mathbf{g}_i = \nabla_\theta \mathcal{L}_i(\theta), i \in [n]$
 - 6: - Form gradient matrix $\mathbf{G} = \{\mathbf{g}_1, \dots, \mathbf{g}_i\} \in \mathbb{R}^{n \times m}$
 - 7: - Eigen decomposition of Gram matrix as in (13):
$$\mathbf{G}\mathbf{G}^\top = \mathbf{U}(\Sigma\Sigma^\top)\mathbf{U}^\top$$
 with eigenvalues λ
 - 8: - Get scaling factor for gradient normalization:
$$\tilde{\sigma} = \min(\sqrt{\lambda})$$
 - 9: - Calculated the orthogonal and normalized gradient matrix as in (15):
$$\tilde{\mathbf{G}} = \tilde{\sigma}\mathbf{U}\Sigma^{-1}\mathbf{U}^\top \mathbf{G}$$
 - 10: **if** $t = t_{warm}$ **then**
 - 11: - Record the loss values for all the tasks $\mathcal{L}_i(t_{warm})$
 - 12: **else if** $t > t_{warm}$ **then**
 - 13: - Calculate the task difficulty weights as in (17):
$$k_i(t) = \text{softmax}(lr_i(t-1))$$
 - 14: - Form eccentric vector $\mathbf{v}_{ecc} = [k_1, \dots, k_i]^\top$
 - 15: **end if**
 - 16: **Output:** $\tilde{\mathbf{g}}_{ecc} = \tilde{\mathbf{G}}^\top \mathbf{v}_{ecc}$ for optimization



Supplementary Materials – Chapter 5

Gradient Normalization

1. The magnitude of each gradient \propto eigenvalues λ [4]

$$\tilde{\mathbf{G}} = \mathbf{U}\mathbf{V}^\top \quad \text{with} \quad \mathbf{G}\mathbf{G}^\top = \mathbf{U}(\Sigma\Sigma^\top)\mathbf{U}^\top$$

2. Get scaling factor for gradient normalization:

$$\tilde{\sigma} = \min(\sqrt{\lambda})$$

3. The new gradient matrix is:

$$\tilde{\mathbf{G}} = \tilde{\sigma}\mathbf{U}\Sigma^{-1}\mathbf{U}^\top\mathbf{G}$$

The magnitude dominance resolved \square

Algorithm 1 EGA Optimization Strategy for MTL

```
1: Input: Loss values for  $n$  tasks  $[\mathcal{L}_1, \dots, \mathcal{L}_n], i \in [n]$ ,  
       $T$  for softmax and  $t_{warm}$  for warmup epoch  
INITIALIZATION:  
2: - Unit eccentric vector  $\mathbf{v}_{ecc} = [1, \dots, 1]^\top \in \mathbb{R}^n$   
FOR EACH ITERATION:  
3: - Get the current epoch as  $t$   
4: - Get task-specific gradient  $\mathbf{g}_i = \nabla_\theta \mathcal{L}_i(\theta), i \in [n]$   
5: - Form gradient matrix  $\mathbf{G} = \{\mathbf{g}_1, \dots, \mathbf{g}_n\} \in \mathbb{R}^{n \times m}$   
6: - Eigen decomposition of Gram matrix as in (13):  
       $\mathbf{G}\mathbf{G}^\top = \mathbf{U}(\Sigma\Sigma^\top)\mathbf{U}^\top$  with eigenvalues  $\lambda$   
7: - Get scaling factor for gradient normalization:  
       $\tilde{\sigma} = \min(\sqrt{\lambda})$   
8: - Calculated the orthogonal and normalized gradient  
      matrix as in (15):  $\tilde{\mathbf{G}} = \tilde{\sigma}\mathbf{U}\Sigma^{-1}\mathbf{U}^\top\mathbf{G}$   
9: if  $t = t_{warm}$  then  
10:   - Record the loss values for all the tasks  $\mathcal{L}_i(t_{warm})$   
11: else if  $t > t_{warm}$  then  
12:   - Calculate the task difficulty weights as in (17):  
      $k_i(t) = \text{softmax}(lr_i(t-1))$   
13:   - Form eccentric vector  $\mathbf{v}_{ecc} = [k_1, \dots, k_n]^\top$   
14: end if  
15: Output:  $\tilde{\mathbf{g}}_{ecc} = \tilde{\mathbf{G}}^\top\mathbf{v}_{ecc}$  for optimization
```

[4] D.Senushkin,N.Patakin,A.Kuznetsov, and A.Konushin, "Independent component alignment for multi-task learning," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Sep. 2023



Supplementary Materials – Chapter 5

Introduce eccentric vector

1. Run several warmup epoch until the epoch t_{warm}

2. Record the loss values for all the tasks $\mathcal{L}_i(t_{warm})$

3. Calculate learning rate:

$$lr_i(t-1) = \frac{\mathcal{L}_i(t-1)}{\mathcal{L}_i(t_{warm})} \quad (\text{small } lr_i \text{ for fast learning rate})$$

4. Calculate task difficulty weight and \mathbf{v}_{ecc} :

$$k_i(t) = \text{softmax}(lr_i(t-1)) = \frac{n \exp(lr_i(t-1)/T)}{\sum_{j=1}^n \exp(lr_j(t-1)/T)}$$

$$\mathbf{v}_{ecc} = [k_1, \dots, k_n]^\top \quad (\text{small T enlarge discrepancy})$$

5. Get the final joint gradient:

$$\tilde{\mathbf{g}}_{ecc} = \tilde{\mathbf{G}}^\top \mathbf{v}_{ecc}$$

The difficulty imbalance resolved □

Algorithm 1 EGA Optimization Strategy for MTL

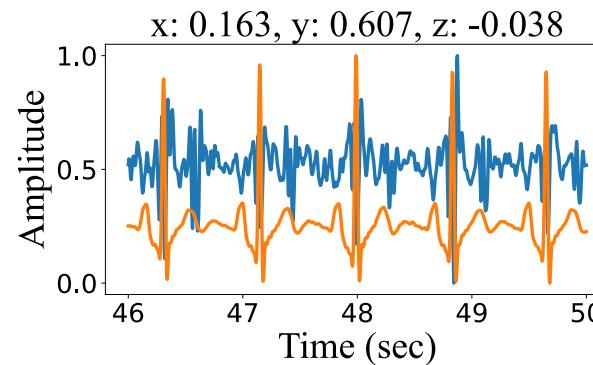
- 1: **Input:** Loss values for n tasks $[\mathcal{L}_1, \dots, \mathcal{L}_n], i \in [n]$,
 T for softmax and t_{warm} for warmup epoch
INITIALIZATION:
 - 2: - Unit eccentric vector $\mathbf{v}_{ecc} = [1, \dots, 1]^\top \in \mathbb{R}^n$
 - FOR EACH ITERATION:
 - 3: - Get the current epoch as t
 - 4: - Get task-specific gradient $\mathbf{g}_i = \nabla_\theta \mathcal{L}_i(\theta), i \in [n]$
 - 5: - Form gradient matrix $\mathbf{G} = \{\mathbf{g}_1, \dots, \mathbf{g}_n\} \in \mathbb{R}^{n \times m}$
 - 6: - Eigen decomposition of Gram matrix as in (13):
 $\mathbf{G}\mathbf{G}^\top = \mathbf{U}(\Sigma\Sigma^\top)\mathbf{U}^\top$ with eigenvalues λ
 - 7: - Get scaling factor for gradient normalization:
 $\tilde{\sigma} = \min(\sqrt{\lambda})$
 - 8: - Calculated the orthogonal and normalized gradient
 matrix as in (15): $\tilde{\mathbf{G}} = \tilde{\sigma}\mathbf{U}\Sigma^{-1}\mathbf{U}^\top\mathbf{G}$
 - 9: **if** $t = t_{warm}$ **then**
 - 10: - Record the loss values for all the tasks $\mathcal{L}_i(t_{warm})$
 - 11: **else if** $t > t_{warm}$ **then**
 - 12: - Calculate the task difficulty weights as in (17):
 $k_i(t) = \text{softmax}(lr_i(t-1))$
 - 13: - Form eccentric vector $\mathbf{v}_{ecc} = [k_1, \dots, k_n]^\top$
 - 14: **end if**
 - 15: **Output:** $\tilde{\mathbf{g}}_{ecc} = \tilde{\mathbf{G}}^\top \mathbf{v}_{ecc}$ for optimization



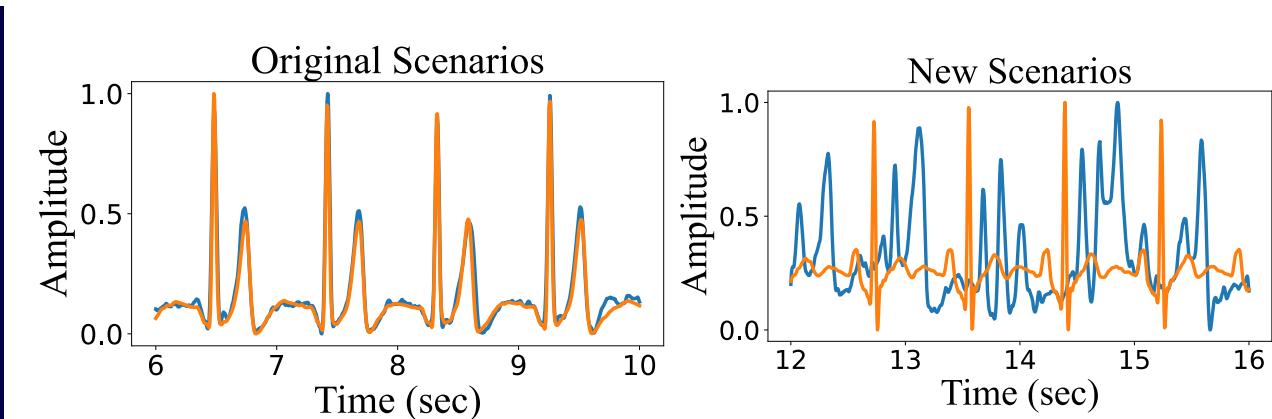
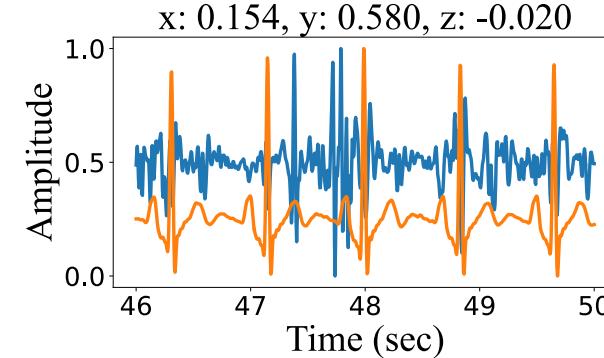
Supplementary Materials – Chapter 6

Challenges:

1. Efficiently extract high-SNR radar signals with ample cardiac features:
 - Assess signal SNR without ground truth.
 - Search for the high-SNR signal within a highly discontinuous space.
2. Learning common representations from unlabeled radar signals to assist supervised training with limited data.



Signal extracted form adjacent points (distance = 3cm)



Inference for the radar signal collected from new scenario