Cross-domain Joint Dictionary Learning for ECG Inference from PPG

Xin Tian, *Member, IEEE,* Qiang Zhu, *Member, IEEE,* Yuenan Li, *Senior Member, IEEE,* Min Wu, *Fellow, IEEE*

***Abstract*—The inverse problem of inferring clinical gold-standard electrocardiogram (ECG) from photoplethysmogram (PPG) that can be measured by affordable wearable internet-of-healthcare-things (IoHT) devices is a research direction receiving growing attention. It combines the easy measurability of PPG and the rich clinical knowledge of ECG for long-term continuous cardiac monitoring. The prior art for reconstruction using a universal basis, such as discrete cosine transform (DCT), has limited fidelity for uncommon ECG shapes due to the lack of representative power. To better utilize the data and improve data representation, we design two dictionary learning frameworks, the cross-domain joint dictionary learning (XDJDL) and the label-consistent XDJDL (LC-XDJDL), to further improve the ECG inference quality and enrich the PPG-based diagnosis knowledge. Building on the K-SVD technique, our proposed joint dictionary learning frameworks largely extend the expressive power by optimizing simultaneously a pair of signal dictionaries for PPG and ECG with the transforms to relate their sparse codes and disease information. The proposed models are evaluated with a variety of PPG and ECG morphologies from two benchmark datasets that cover various age groups and disease types. The results show the proposed frameworks achieve better inference performance than previous methods with average Pearson coefficients being 0.88 using XDJDL and 0.92 using LC-XDJDL, suggesting an encouraging potential for ECG screening using PPG based on the proactively learned PPG-ECG relationship. By enabling the dynamic monitoring and analysis of the health status of an individual, the proposed frameworks contribute to the emerging digital twins paradigm for personalized healthcare.**

***Index Terms*—Joint Dictionary Learning, Sparse Coding, ECG, PPG, Internet-of-Healthcare-Things (IoHT), Digital Twins.**

I. Introduction

CARDIOVASCULAR diseases (CVDs) have become a leading cause of death globally. From alarming reports of the World Health Organization, an estimated 17.9 million people died from CVDs in 2019, representing 32% of all global deaths [2].

Electrocardiogram (ECG) is a widely-used gold-standard for cardiovascular diagnostic procedures. By measuring the electrical activity of the heart and conveying information regarding heart functionality, continuous ECG monitoring is proven to be beneficial for the early detection of CVDs [3]. However, most conventional ECG equipment is restrictive on users' activities. Newer clinical ambulatory ECG monitoring devices, such as the Zio patch [4], have alleviated the above-mentioned issues, although potential skin irritation during long-term adhesive wear remains, especially for people with sensitive skin. In addition, a prescription is needed to obtain the Zio patch, thus not easily accessible to the general public. Apple Watch [5] and wearable devices alike, such as KardiaMobile [6], are moderately affordable and can show real-time ECG without adhesion to the skin, but they generally require active user participation and is usually for short duration measurement of 30-second periods, making it infeasible for long-term continuous ECG monitoring. Table I summarizes the comparison of different ECG sensing techniques discussed above.

TABLE I
COMPARISON OF DIFFERENT ECG SENSING TECHNIQUES.

ECG Sens. Tech.	Cost	Accessibility	Need No Active Participation?	Long-Term & Conts. Monitoring
Standard ECG	Medium*	Low	×	×
Apple Watch [5]	Medium	High	×	×
KardiaMobile [6]	Low	High	×	×
Zio patch [4]	Medium*	Low	✓	✓(skin irritation)
Our Proposed	Low	High	✓	✓(little side effect)

*High cost in the U.S. if without medical insurance.

Given the constraints of the ECG sensors, researchers have made efforts toward long-term continuous ECG monitoring by inferring full ECG waveform from optical sensors, such as the photoplethysmogram (PPG) sensors [1], [7], [8]. PPG sensors are ubiquitously seen in the wearable internet-of-healthcare-things (IoHT) devices and have become a common modality for monitoring heart conditions due to the maturity of the technology and low cost [9]. It measures the optical response of the blood volume changes at the peripheral ends, including fingertips [10], and provides valuable information about the cardiovascular system via daily use of the pulse oximeter. Compared to ECG, PPG is more user-friendly in long-term continuous monitoring without constant user participation.

PPG and ECG are physiologically related as they embody the same cardiac process in two different signal sensing domains. The peripheral blood volume change recorded by PPG is influenced by the contraction and relaxation of the heart muscles, which are controlled by the cardiac electrical signals triggered by the sinoatrial node [11]. The waveform shape (i.e. signal morphology), pulse interval, and amplitude characteristics of PPG provide important information

This work was supported in part by NSF under Grant 2124291.

X. Tian and Q. Zhu were with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, 20742 USA, where the work was carried out, and are now with Meta Inc. (e-mail: xtian17, zhuqiang@terpmail.umd.edu).

Y. Li is with the School of Electrical and Information Engineering, Tianjin University, Tianjin, China (email: ynli@tju.edu.cn). He participated in this research while he was a visiting researcher at the University of Maryland, College Park, MD, USA.

M. Wu is with the Department of Electrical and Computer Engineering, University of Maryland, College Park, MD, 20742 USA (e-mail: minwu@umd.edu).

A preliminary version (4 pages of main text) reporting early-stage results of this work was presented in the 2020 IEEE ICASSP conference and was published in its proceedings (DOI: 10.1109/ICASSP40776.2020.9054242) [1].

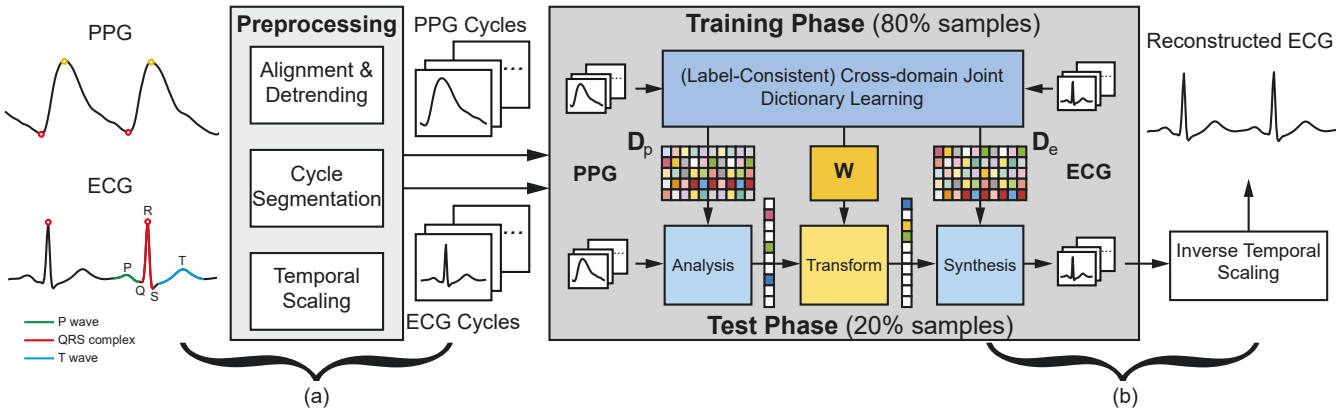


Fig. 1. Illustration of the proposed framework. The ECG and PPG signals are first preprocessed to obtain temporally aligned and normalized pairs of cycles. 80% pairs of ECG and PPG signal cycles from each subject are used for training paired dictionaries $\mathbf{D}_p$, $\mathbf{D}_e$, and a linear transform $\mathbf{W}$ which will be applied in the test phase to infer the ECG signals.

about the cardiovascular system [10], including heart rate, heart rate variability [12], respiration [13], and blood pressure [14]. Therefore, inferring the medical gold-standard ECG signal using the PPG sensor provides a solution to achieve a low-cost, long-term continuous cardiac monitoring, which facilitates further diagnosis and leads to early intervention opportunities, especially for the low-income, disadvantaged populations, who have limited access to affordable preventive care. Our proposed technique embodies the trend of *digital twins* in healthcare [15], which is an emerging technology that plays a pivotal role in advancing personalized healthcare. The aspects of digital twins that our work contributes to are on developing a rich representation of an individual supported by data and models, through which the physiological status of this individual can be dynamically monitored by various sensors through IoHT and analyzed over time.

The previous work of ECG reconstruction from PPG using a universal, data-independent basis of the discrete cosine transform (DCT) [7], [8] has limited fidelity to represent uncommon ECG waveform shapes, especially for the *group-based* case with a broader range of signal morphologies [16]. We focus on such group-based cases in this paper and consider data science and learning techniques with richer representative power to answer the following research question:

- *Group-based model*: Can a single model, trained from a group of subjects with a certain determinant of physiology (e.g., age, weight, disease type, etc.), predict the ECG waveforms from unseen PPG measurements for individuals in the training group?

To overcome the limitation of the DCT method and develop the synergy of model and data, our work aims at improving data representation through a more versatile and adaptive framework based on dictionary learning to demonstrate the feasibility of ECG waveform inference from PPG signal as an inverse filtering problem. In addition to the algorithmic improvement, sparse coding and dictionary learning frameworks are proven to perform efficiently in IoT platforms in terms of cutting down power consumption and computation cost [17], [18]. Thus, by investigating the dictionary learning based approach in this paper, we strike a balance between the model complexity and practical cost in IoT applications.

Our proposed cross-domain joint dictionary learning (XDJDL) method for ECG reconstruction from PPG is summarized in Fig. 1. A further-developed label-consistent XDJDL model (LC-XDJDL) is also proposed when the disease label information is available. The PPG and ECG signals are first preprocessed into normalized signal cycles to facilitate the subsequent training. In the training phase, the ECG-PPG dictionary pair is jointly updated with a stable linear mapping that relates the sparse representations of the two measurements. In LC-XDJDL, an additional linear mapping that enforces the label consistency for the PPG sparse codes will be learned to further improve the ECG reconstruction performance and enrich the knowledge base for diagnosis from PPG.

Extending the preliminary results reported in [1], the contribution of this work is threefold:

1) To the best of our knowledge, together with our earlier dissemination of preliminary results in a conference publication [1], this paper presents the first clinical application of frameworks based on dictionary pair learning for ECG waveform inference from PPG. The XDJDL framework enables the synergistic utilization of the advantages of PPG and ECG signals for better preventive healthcare. Furthermore, a new label-consistent XDJDL (LC-XDJDL) is proposed in this paper to add the label-consistency constraint for PPG sparse codes. The benefits of LC-XDJDL are not limited to improving the quality of ECG reconstruction, but also making the PPG-based heart disease diagnosis interpretable as the transform that relates the sparse codes of PPG to disease labels can help to gain intuitive insights on how a certain disease can impact PPG.

2) On top of the preliminary work reported in our earlier conference publication, extended work is included in this journal version: a) The datasets used for experimental evaluation are expanded to 75 subjects with a variety of age and ECG morphological patterns collected from different cardiovascular pathologies; b) The inferred ECG signals are evaluated in both overall waveform shape and the morphology and timing of the subwaves for a fine-granular evaluation; and c) Our proposed

frameworks are compared with a series of prior art dictionary learning algorithms and outperform them, suggesting that the developed learning methods can accurately infer ECG from PPG.

3) In the discussions, we extend the evaluation of the proposed framework when the cardiac cycle information in the test set can only be obtained from the PPG signals. We analyze the practicality of applying our proposed ECG inference techniques in low-power and low-cost healthcare IoT devices. We also present the limitations of the proposed method in two challenging cases with a leave-one-out experiment and a preliminary experiment on an extra dataset where PPG and ECG are distorted by motion artifacts during running.

II. Related Work

A. ECG Reconstruction From PPG

There are many prior arts extracting physiological parameters [19], [20] or classifying arrhythmia [21]–[24] from the input ECG or PPG signals using machine learning methods. However, direct parameter estimation or automatic diagnosis is insufficient for medical practitioners to interpret. The ECG signal, rather than the derived results via black-box models, is still the gold-standard tool on which cardiologists rely and make further decisions. Our proposed technique in this paper providing the reconstructed ECG waveform offers complementary support and allows the manual check from cardiovascular experts with their medical expertise and clinical experiences.

Very limited prior work has been devoted to PPG-based ECG inference. The pilot study [7], [8] proposed to relate the waveforms of PPG and ECG in the discrete cosine transform (DCT) domain by a linear model. In the participant-specific case where a linear model is trained from and tested for the same individual, this DCT method achieved a mean reconstruction correlation of 0.94. In contrast, for the group-based model, the achieved mean correlation degraded to 0.79. This suggests that there is still substantial room for improvement when extending to the group-based model case where a universal mapping needs to be trained by a wider variety of ECG morphologies from multiple people. To address these above-mentioned issues, we consider dictionary learning based sparse representation for ECG and PPG as it provides a richer and more adaptive representation than the universal dictionary DCT by better leveraging data. And we will use this as a foundation to develop joint dictionary learning models for reconstruction. The theoretical basis of the principled methods, including the DCT and our proposed dictionary learning frameworks, is the underlying physiological and signal processing mechanisms [8]. Neural networks, with strong expressive power and high structural flexibility, are also adopted to solve this problem [25], [26]. However, current neural network based PPG-to-ECG algorithms using black-box end-to-end signal synthesis lack interpretability, and our ongoing efforts are being made to investigate neural networks with low complexity and interpretability.

B. Dictionary Learning

Algorithms that learn a single dictionary for signal representation [27]–[29] have been well-studied. They have been successfully applied to cardiac signal processing, including recent research showing that ECG signals can be well-represented as a sparse linear combination of atoms from an appropriately learned dictionary for such applications as ECG classification and compression [30]–[32].

In the domain of image processing and computer vision, these single dictionary learning strategies [33], [34] have been extended to joint dictionary learning tasks [35], [36]. For image super-resolution [16], [37], [38], coupled dictionary learning frameworks are proposed to learn a dictionary pair for low- and high-resolution image patches while enforcing the similarity of their sparse codes with respect to their dictionaries. One assumption from this model is that the transform matrix between the two sparse codes is an identity matrix. In person re-identification [39] and photo-to-sketch [40] problems, a linear mapping between the codings of input and output images is introduced into the objective function for semi-coupled dictionary learning. In both training schemes, the updates of the mapping and dictionaries are separately done within each iteration, making the dictionary computation less aware of the signal transform.

Our method aims at boosting reconstruction performance from PPG to ECG by using a joint dictionary learning framework. Unlike the super-resolution problem [16], [37], [38] where the input and output reside in the same signal domain, the problem of ECG reconstruction from PPG extends to two different sensing domains and an identity transfer matrix may not be adequate to capture the transform between the domains. To address this, XDJDL introduces a PPG-to-ECG mapping, which spans the two sensing modalities with low waveform correlation, providing more flexibility and generalization for the two learned dictionaries. Different from [39], [40], we update the linear transform and the dictionary in the same step, which can optimize the capability of the obtained dictionaries for both signal representation and transformation. This kind of transform-aware joint dictionary learning formulation is one of the major differences from other coupled dictionary learning frameworks. This framework can also be easily generalized to different constraints. For instance, in the proposed LC-XDJDL model, we add a label-consistency regularization term to the objective function of the XDJDL model, which encourages the transformed sparse codes from the same class to be similar.

III. Proposed Framework

A. Physiological Background and Preprocessing

As briefly discussed in the introduction, PPG and ECG are physiologically related because the corresponding cycles of them represent the same cardiac process during the respective heartbeat measured in two sensing modalities. Different phases in one cardiac cycle progress as follows [10], [41] with fiducial points labeled in part(a) of Fig. 1: The heart's pacemaker at the sinoatrial node first triggers the depolarization and contraction of the atria, resulting in the P-wave of ECG. The wave of depolarization then proceeds to initiate ventricular contraction that generates the QRS wave complex of ECG. Simultaneously, blood is pumped to the body and microvascular blood volume increases, resulting in the ascending slope of PPG. Finally,

ventricular repolarization and relaxation follow to form the T-wave of ECG and blood flows back towards the heart, leading to the descending slope of PPG.

To establish the quantitative relationship between the corresponding cycles of ECG and PPG, we preprocess the two signals during the training phase to obtain temporally aligned and normalized pairs of signals, so that the critical temporal features of both waveforms are synchronized for learning and evaluation. We have followed the procedures used in [7]. The preprocessing method we adopt is rooted in the aforementioned underlying physiological relationships between PPG and ECG signals, which is independent of the dataset selection. First, considering the synchronization issue between separate ECG and PPG devices, we align the whole ECG and PPG sequences according to the moment when the ventricles contract and the blood flows to the vessels, which corresponds to the R peaks of ECG and the onsets of PPG in the same cycle. Both the onset and R peaks are detected by the beat detection functions from the PhysioNet Cardiovascular Signal Toolbox [42]. Then we detrend the aligned signals by a second-order difference operator based algorithm [7] to eliminate the baseline drift related to respiration, motion, vasomotor activity, and change in contact surface [10]. To prepare for the learning of the cycle-wise relation during one heartbeat, the detrended PPG and ECG signals are partitioned into cycles by the *R2R* [7] segmentation scheme, where the partition points are the R peaks of the ECG signal. After the segmentation, each cycle is linearly interpolated to length d to mitigate the influence of the heart rate variation. Finally, we normalize the amplitude of each cycle by subtracting the sample mean and dividing by the sample standard deviation. The preprocessed PPG and ECG signal cycles are stored in data matrices $\mathbf{P}$ and $\mathbf{E}$, respectively.

B. Cross-Domain Joint Dictionary Learning (XDJDL)

1) The K-SVD Model: As one of the most popular dictionary learning methods, the K-SVD model is composed of two main optimization steps: linear sparse coding based on the current overcomplete dictionary with k atoms, and updating the dictionary by the SVD method [27].

Let $\mathbf{X} \in \mathbf{R}^{d\times n}$ be a set of input signals, with each column $\mathbf{x}_i$ being a training sample. K-SVD aims to solve the following ℓ_0-norm constraint problem in Eq. (1):

$$\begin{aligned} &\min_{\mathbf{D},\mathbf{A}} \|\mathbf{X}-\mathbf{D}\mathbf{A}\|_F^2 \\ &\text{s.t.}\ \|\mathbf{a}_j\|_0 \le t_0, \quad j=1,...,n. \end{aligned} \tag{1}$$

where $\mathbf{D} \in \mathbb{R}^{d\times k}$ is the reconstructive dictionary with k atoms; $\mathbf{A} \in \mathbb{R}^{k\times n}$ is the corresponding sparse codes of $\mathbf{X}$, with each column denoted as $\mathbf{a}_j$; and t_0 is a parameter for sparsity constraint. Frobenius norm is used to calculate the element-wise ℓ_2-norm in the given matrix.

In this paper, we aim to tackle the ECG inference from PPG by learning a dictionary pair for ECG and PPG along with a linear transform between the sparse representations of the two signals. The reconstructive dictionary pair characterizes the two structural domains of the two biomedical signals, and the mapping function reveals the intrinsic relationship between ECG and PPG signals in the sparse domain. We impose the linear mapping error as one regularization term in the objective function, and convert it to a problem that can be optimized by the K-SVD dictionary learning method. The details of the model formulation and optimization algorithm are discussed in the following subsections.

2) Proposed XDJDL Model: We denote the PPG and ECG datasets as $\mathbf{P} = [\mathbf{X}_p, \mathbf{T}_p] \in \mathbb{R}^{d\times(n+m)}$ and $\mathbf{E} = [\mathbf{X}_e, \mathbf{T}_e] \in \mathbb{R}^{d\times(n+m)}$, respectively. Each column of $\mathbf{P}$ and $\mathbf{E}$ is denoted as $\mathbf{p}_i \in \mathbb{R}^{d\times 1}$ and $\mathbf{e}_i \in \mathbb{R}^{d\times 1}$, representing one PPG/ECG cycle during the same cardiac cycle. The goal is to learn the patterns (in terms of dictionaries, mappings, etc.) from the training data $\mathbf{X}_p \in \mathbb{R}^{d\times n}$ and $\mathbf{X}_e \in \mathbb{R}^{d\times n}$ to infer the test ECG dataset $\mathbf{T}_e \in \mathbb{R}^{d\times m}$ from PPG $\mathbf{T}_p \in \mathbb{R}^{d\times m}$.

We formulate the XDJDL framework as:

$$\begin{aligned} \min_{\mathbf{D}_e,\mathbf{A}_e,\mathbf{D}_p,\mathbf{A}_p,\mathbf{W}} &\|\mathbf{X}_e-\mathbf{D}_e\mathbf{A}_e\|_F^2 + \alpha\|\mathbf{X}_p-\mathbf{D}_p\mathbf{A}_p\|_F^2 \\ &+ \beta\|\mathbf{A}_e-\mathbf{W}\mathbf{A}_p\|_F^2 \\ &\text{s.t.}\ \|\mathbf{a}_{p,j}\|_0 \le t_p, \text{and}\ \|\mathbf{a}_{e,j}\|_0 \le t_e, \quad j=1,...,n. \end{aligned} \tag{2}$$

where $\mathbf{D}_p \in \mathbb{R}^{d\times k_p}$ and $\mathbf{D}_e \in \mathbb{R}^{d\times k_e}$ are dictionaries learned for $\mathbf{X}_p$ and $\mathbf{X}_e$, respectively; $\mathbf{A}_p \in \mathbb{R}^{k_p\times n}$ and $\mathbf{A}_e \in \mathbb{R}^{k_e\times n}$ are the corresponding sparse coding matrices. Each column of $\mathbf{A}_p$ and $\mathbf{A}_e$ is denoted as $\mathbf{a}_{p,j}$ and $\mathbf{a}_{e,j}$ with the sparsity bounded above by t_p and t_e.

For the objective function in Eq. (2), $\|\mathbf{X}_e-\mathbf{D}_e\mathbf{A}_e\|_F^2$ and $\|\mathbf{X}_p-\mathbf{D}_p\mathbf{A}_p\|_F^2$ are the data fidelity terms for ECG and PPG cycle sets, respectively. The term $\|\mathbf{A}_e-\mathbf{W}\mathbf{A}_p\|_F^2$ represents the mapping error between the sparse codes of ECG and PPG signals, which enforces the transformed sparse codes of PPG to approximate that of ECG. Intuitively, we can enforce the two sparse representations for ECG and PPG from the same cycle to be the same and set the regularization term as $\|\mathbf{A}_e-\mathbf{A}_p\|_F^2$. However, since ECG and PPG are from two different signal sensing modalities and the waveform difference between the two signals is significant, directly pushing their sparse representations to be similar could compromise the generalization of the two learned dictionaries.

From the formulation in Eq. (2), we can jointly learn the dictionaries for ECG and PPG datasets, which produce a good representation for each sample in the training set with strict sparsity constraints. Meanwhile, we learn the linear approximation of the transform that relates the sparse codes of PPG and ECG, and use it to entail the intrinsic relationship between certain PPG atoms and ECG atoms from their corresponding dictionaries.

3) Optimization: Eq. (2) can be rewritten as:

$$\begin{aligned} \min_{\substack{\mathbf{D}_e,\mathbf{A}_e,\mathbf{D}_p,\\ \mathbf{A}_p,\mathbf{W}}} &\left\| \begin{pmatrix} \mathbf{X}_e \\ \sqrt{\alpha}\mathbf{X}_p \\ \mathbf{0} \end{pmatrix} - \begin{pmatrix} \mathbf{D}_e & \mathbf{0} \\ \mathbf{0} & \sqrt{\alpha}\mathbf{D}_p \\ -\sqrt{\beta}\mathbf{I} & \sqrt{\beta}\mathbf{W} \end{pmatrix} \begin{pmatrix} \mathbf{A}_e \\ \mathbf{A}_p \end{pmatrix} \right\|_F^2 \\ &\text{s.t.}\ \|\mathbf{a}_{e,j}\|_0 \le t_e, \text{and}\ \|\mathbf{a}_{p,j}\|_0 \le t_p, \quad j=1,...,n. \end{aligned} \tag{3}$$

where $\mathbf{I}$ is an identity matrix and $\mathbf{0}$ is a zero matrix.

Let $\mathbf{X} \triangleq (\mathbf{X}_e, \sqrt{\alpha}\mathbf{X}_p, \mathbf{0})^{\mathrm{T}} \in \mathbb{R}^{(2d+k_e)\times n}$, $\mathbf{D} \triangleq (\mathbf{D}_e, \mathbf{0}, -\sqrt{\beta}\mathbf{I}; \mathbf{0}, \sqrt{\alpha}\mathbf{D}_p, \sqrt{\beta}\mathbf{W})^{\mathrm{T}} \in \mathbb{R}^{(2d+k_e)\times(k_e+k_p)}$, and

$\mathbf{A} \triangleq (\mathbf{A}_e, \mathbf{A}_p)^{\mathrm{T}} \in \mathbb{R}^{(k_e+k_p)\times n}$. The optimization of (3) can be written as the following problem:

$$\begin{aligned} &\min_{\mathbf{D},\mathbf{A}} \|\mathbf{X}-\mathbf{D}\mathbf{A}\|_F^2, \\ &\text{s.t. } \|\mathbf{a}_{+,j}\|_0 \le t_e, \text{and } \|\mathbf{a}_{-,j}\|_0 \le t_p, \quad j=1,...,n. \end{aligned} \tag{4}$$

where $\mathbf{a}_{*,j}$ denotes the jth column of $\mathbf{A}_*$, and $\mathbf{A}_+$ is defined as the first k_e rows of sparse matrix $\mathbf{A}$ while $\mathbf{A}_-$ is the last k_p rows of sparse matrix $\mathbf{A}$. The formulation in Eq. (4) now resembles Eq. (1), suggesting that K-SVD can be adapted for this optimization. The only difference between Eq. (4) and Eq. (1) is the local sparsity constraint, which will be addressed in the following optimization procedures.

Step 0: Initialization

To initialize $\mathbf{D}$ and $\mathbf{A}$, we need to initialize their components: $\mathbf{D}_e, \mathbf{D}_p, \mathbf{W}, \mathbf{A}_e$, and $\mathbf{A}_p$. First, we randomly select a subset of columns from training data $\mathbf{X}_e$ and $\mathbf{X}_p$ to form $\mathbf{D}_e$ and $\mathbf{D}_p$. Then, we initialize the sparse codes $\mathbf{A}_e$ and $\mathbf{A}_p$ by solving Eq. (7) with respect to $\{\mathbf{D}_e,\ \mathbf{X}_e,\ t_e\}$ and $\{\mathbf{D}_p,\ \mathbf{X}_p,\ t_p\}$, respectively. Finally, we use the ridge regression model to initialize $\mathbf{W}$:

$$\min_{\mathbf{W}} \|\mathbf{A}_e - \mathbf{W}\mathbf{A}_p\|_F^2 + \lambda \|\mathbf{W}\|_F^2. \tag{5}$$

This has a closed-form solution as:

$$\mathbf{W} = \mathbf{A}_e\mathbf{A}_p^{\mathrm{T}}(\mathbf{A}_p\mathbf{A}_p^{\mathrm{T}} + \lambda\mathbf{I})^{-1}. \tag{6}$$

After the initialization, we use a two-step iterative optimization to minimize the energy in (4), whereby step one is sparse coding and step two is dictionary updating by SVD.

Step 1: Sparse Coding

Given $\mathbf{D}$, the step of sparse coding finds the sparse representation $\mathbf{a}_j$ for $\mathbf{x}_j$, for $j = 1, ..., n$, by solving

$$\begin{aligned} &\min_{\mathbf{a}_j} \|\mathbf{x}_j - \mathbf{D}\mathbf{a}_j\|_2^2 \\ &\text{s.t. } \|\mathbf{a}_j\|_0 \le t. \end{aligned} \tag{7}$$

where $\mathbf{a}_j$ is the j^{th} column of the sparse representation matrix $\mathbf{A}$ and $\mathbf{x}_j$ is the j^{th} training sample in matrix $\mathbf{X}$.

Many approaches were proposed to solve Eq. (7) [43]. Here we adopt the orthogonal matching pursuit (OMP) method [44], which is a greedy method that provides a good approximation with convergence in limited iterations. As mentioned earlier, the local sparsity constraints imposed on Eq. (4) will affect the direct application of OMP. One workaround is to solve the following problem in Eq. (8) in place of Eq. (4),

$$\begin{aligned} &\min_{\mathbf{D},\mathbf{A}} \|\mathbf{X}-\mathbf{D}\mathbf{A}\|_F^2 \\ &\text{s.t. } \|\mathbf{a}_j\|_0 \le t_e + t_p, \quad j=1,...,n. \end{aligned} \tag{8}$$

where $\mathbf{a}_j$ is the vertical concatenation of $\mathbf{a}_{+,j}$ and $\mathbf{a}_{-,j}$ in Eq. (4), and t_e and t_p are the sparsity constraints for the upper and bottom parts of $\mathbf{a}_j$, respectively. During the OMP process in each iteration, we will only keep the largest sparse coefficients in $\mathbf{a}_j$ to ensure the local sparsity constraints.

Step 2: Dictionary Update

To update the k^{th} atom, $\mathbf{d}_k$, in dictionary $\mathbf{D}$ and its corresponding coefficients, $\mathbf{a}_R^k$, in the k^{th} row of $\mathbf{A}$, we apply SVD to the residue term $\mathbf{R}_k \triangleq \mathbf{X} - \sum_{j\neq k}\mathbf{d}_j\mathbf{a}_R^j$. In practice, we only select the training samples that use the atom $\mathbf{d}_k$ and avoid filling in the zeros entries of $\mathbf{a}_R^k$ during the update. We do so through denoting the nonzero entries in $\mathbf{a}_R^k$ as $\tilde{\mathbf{a}}_R^k$, and correspondingly, $\mathbf{R}_k$ as $\tilde{\mathbf{R}}_k$. The updated atom $\mathbf{d}_k$ and the related coefficients $\tilde{\mathbf{a}}_R^k$ will then be computed by:

$$\min_{\mathbf{d}_k,\tilde{\mathbf{a}}_R^k} \left\|\tilde{\mathbf{R}}_k - \mathbf{d}_k\tilde{\mathbf{a}}_R^k\right\|_F^2. \tag{9}$$

To solve Eq. (9), we use the SVD method on the residue term [27], i.e., $\tilde{\mathbf{R}}_k = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{\mathbf{T}}$. And then, $\mathbf{d}_k$ and $\tilde{\mathbf{a}}_R^k$ can be updated as follows:

$$\begin{aligned} \mathbf{d}_k &= \mathbf{U}(:,1), \\ \tilde{\mathbf{a}}_R^k &= \mathbf{\Sigma}(1,1)\mathbf{V}^{\mathbf{T}}(1,:). \end{aligned} \tag{10}$$

Note that taking $\mathbf{D} \triangleq (\mathbf{D}_e, \mathbf{0}, -\sqrt{\beta}\mathbf{I}; \mathbf{0}, \sqrt{\alpha}\mathbf{D}_p, \sqrt{\beta}\mathbf{W})^{\mathrm{T}}$ as a whole in the dictionary update phase does not solve this optimization problem because the zero matrices part and the identity matrix part in $\mathbf{D}$ cannot be guaranteed in the update of the dictionary by SVD. A remedy to the above problem is to decompose the dictionary update problem for $\mathbf{D}$ into the following two subproblems by revisiting the matrix form of the optimization problem in Eq. (3).

(i) Update $\mathbf{D}_e, \mathbf{A}_e$:

$$< \mathbf{D}_e^*, \mathbf{A}_e^* > = \underset{\mathbf{D}_e,\mathbf{A}_e}{\operatorname{argmin}} \|\mathbf{X}_e - \mathbf{D}_e\mathbf{A}_e\|_F^2. \tag{11}$$

We use SVD to update all atoms in $\mathbf{D}_e$ and the corresponding nonzero entries in $\mathbf{A}_e$ by solving Eq. (11) with the same procedure as in Eq. (9) and (10). The columns of $\mathbf{D}_e$ are ℓ_2 normalized.

(ii) Update $\mathbf{D}_p, \mathbf{A}_p$, and $\mathbf{W}$:

The updated ECG sparse representation matrix $\mathbf{A}_e^*$ from the subproblem (i) then serves as an input to the second subproblem here to update $\mathbf{W}$, $\mathbf{D}_p$, and $\mathbf{A}_p$ in Eq. (12).

$$< \mathbf{D}_p^*, \mathbf{A}_p^*, \mathbf{W}^* > = \underset{\mathbf{D}_p,\mathbf{A}_p,\mathbf{W}}{\operatorname{argmin}} \left\| \begin{pmatrix} \sqrt{\alpha}\mathbf{X}_p \\ \sqrt{\beta}\mathbf{A}_e^* \end{pmatrix} - \begin{pmatrix} \sqrt{\alpha}\mathbf{D}_p \\ \sqrt{\beta}\mathbf{W} \end{pmatrix} \mathbf{A}_p \right\|_F^2. \tag{12}$$

We treat $(\sqrt{\alpha}\mathbf{D}_p, \sqrt{\beta}\mathbf{W})^{\mathrm{T}}$ as a whole dictionary, and use the SVD method in Eq. (9) and (10) to update it together with the nonzero entries in $\mathbf{A}_p$. The linear transform and the dictionary are updated simultaneously, which addresses the problem of isolated update raised in [39], [40] and is one of the major differences from other coupled dictionary learning models.

After solving the two subproblems, $\mathbf{D}$ and $\mathbf{A}$ can be assembled by filling in the submatrices. The main steps of XDJDL are summarized in Algorithm 1.

Algorithm 1 Cross-domain joint dictionary learning

Input: Training data $\mathbf{X}_e$ and $\mathbf{X}_p$ of ECG and PPG cycles, Testing data $\mathbf{T}_e$ and $\mathbf{T}_p$, and sparsity constraints $t_e,\ t_p$

Training phase:

Initialization:

- Initialize $\{\mathbf{D}_e, \mathbf{D}_p\}$ by randomly selecting atoms from the training data.
- Initialize $\mathbf{A}_e, \mathbf{A}_p$ by solving Eq. (7) with OMP.
- Initialize $\mathbf{W}$ by Eq. (6).

while not converged **do**

- Update $\mathbf{D}, \mathbf{A}$ by combining updated submatrices.
- Sparse coding: compute $\mathbf{A}$ in Eq. (7) with OMP. Zero out the smallest nonzero entries in the columns of $\mathbf{A}$ if any local sparsity constraint does not hold.
- Dictionary update:
  - Update $\mathbf{D}_e, \mathbf{A}_e$ in Eq. (11) by the SVD method illustrated in Eq. (9)(10).
  - Update $\mathbf{D}_p, \mathbf{A}_p, \mathbf{W}$ in Eq. (12) by the SVD method illustrated in Eq. (9)(10).

end while

Testing phase:

for each sample $\mathbf{t}_p^j \in \mathbf{T}_p$ **do**

- Compute sparse code $\mathbf{s}_p^j$ of $\mathbf{t}_p^j$ under $\mathbf{D}_p$ using Eq. (7).
- Calculate $\mathbf{s}_e^j = \mathbf{W}\mathbf{s}_p^j$.
- Compute the reconstructed ECG sample as $\mathbf{r}_e^j = \mathbf{D}_e\mathbf{s}_e^j$, and store it in matrix $\mathbf{R}_e$.

end for

Output: $\mathbf{R}_e$

C. Label Consistent XDJDL (LC-XDJDL)

For cases where the disease type is known or can be predicted, such as from the PPG signals that we have, we can further leverage the disease label. In this subsection, we examine the effect of adding a label consistency regularization term to the objective function in Eq. (2) as follows:

$$\begin{aligned}\min_{\mathbf{D}_e,\mathbf{A}_e,\mathbf{D}_p,\mathbf{A}_p,\mathbf{W}} &\|\mathbf{X}_e - \mathbf{D}_e\mathbf{A}_e\|_F^2 + \alpha\|\mathbf{X}_p - \mathbf{D}_p\mathbf{A}_p\|_F^2 \\ &+ \beta\|\mathbf{A}_e - \mathbf{W}\mathbf{A}_p\|_F^2 + \gamma\|\mathbf{Q} - \mathbf{H}\mathbf{A}_p\|_F^2 \\ &\text{s.t. } \|\mathbf{a}_{p,j}\|_0 \le t_p, \text{and } \|\mathbf{a}_{e,j}\|_0 \le t_e, \quad j = 1, ..., n.\end{aligned} \tag{13}$$

where $\mathbf{Q} \triangleq [q_1, q_2, ..., q_n] \in \mathbb{R}^{r\times n}$ is a discriminative representation matrix [45] in which each column $q_i = [0, 0, ..., 0, 1, 1, 0, .., 0]^{\mathrm{T}} \in \mathbb{R}^{r\times 1}$ corresponds to a discriminative coding for an input signal. The nonzero elements in q_i occur at the corresponding disease label, which is similar to the one-hot encoding with the number of ones as a tunable parameter. The additional regularization term $\|\mathbf{Q} - \mathbf{H}\mathbf{A}_p\|_F^2$ represents the discriminative sparse code error, which enforces the transformed sparse codes of PPG to approximate the discriminative codes in $\mathbf{Q}$. It yields such dictionaries that the signals from the same class have very similar sparse codes, i.e. enforcing the label-consistency in the sparse representations.

We add the label-consistency regularization term for two main purposes: One is to improve the ECG reconstruction quality by using additional class information to constrain the degrees of freedom of the PPG sparse codes. The other is to enrich the knowledge base of PPG for the diagnosis of a certain set of diseases of interest. CVDs weaken the heart functionality, which further impacts the blood circulation in the body, thus PPG manifests certain disease information. By enforcing the consistency between the sparse codes of PPG and disease labels, one can gain insights into how the disease is revealed on PPG by inspecting the specific columns of the PPG sparse coding matrix $\mathbf{A}_p$ and the label matrix $\mathbf{Q}$.

Similarly, Eq. (13) can be written in the matrix form:

$$\begin{aligned}\min_{\substack{\mathbf{D}_e,\mathbf{A}_e,\mathbf{D}_p,\\ \mathbf{A}_p,\mathbf{W},\mathbf{H}}} &\left\| \begin{pmatrix} \mathbf{X}_e \\ \sqrt{\alpha}\mathbf{X}_p \\ \mathbf{0} \\ \sqrt{\gamma}\mathbf{Q} \end{pmatrix} - \begin{pmatrix} \mathbf{D}_e & \mathbf{0} \\ \mathbf{0} & \sqrt{\alpha}\mathbf{D}_p \\ -\sqrt{\beta}\mathbf{I} & \sqrt{\beta}\mathbf{W} \\ 0 & \sqrt{\gamma}\mathbf{H} \end{pmatrix} \begin{pmatrix} \mathbf{A}_e \\ \mathbf{A}_p \end{pmatrix} \right\|_F^2 \\ &\text{s.t. } \|\mathbf{a}_{e,j}\|_0 \le t_e, \text{and } \|\mathbf{a}_{p,j}\|_0 \le t_p, \quad j = 1, ..., n.\end{aligned} \tag{14}$$

The two-step optimization method in Section III-B3 can still be applied to find the optimal solution to both the dictionary pair and the linear mappings $\mathbf{W}$ and $\mathbf{H}$. In the test phase, the PPG sparse representation matrix $\mathbf{A}_p$ is obtained by applying sparse coding with the learned $\mathbf{D}_p$, $\mathbf{H}$, the test sample matrix $\mathbf{T}_p$, and the label matrix $\mathbf{Q}$.

IV. Experimental Evaluation

A. Dataset

The Medical Information Mart for Intensive Care III (MIMIC-III) [46], [47] is a publicly-available database assembled by researchers at MIT. It comprises a large number of ICU patients with de-identified health data from their hospital stays. To evaluate our proposed framework and algorithm, we have extracted a small subset of the MIMIC-III database as follows. First, we select waveforms that contain both lead-II ECG and PPG signals sampled at 125Hz from the MIMIC-III waveform database. Then the selected waveforms are cross-referenced with the corresponding patient profile by subject ID in the MIMIC-III clinical information database. Patients with the four types of CVDs are further selected: congestive heart failure (CHF), myocardial infarction (MI) including ST-segment elevated (STEMI) and non-ST segment elevated (NSTEMI), hypotension (HYPO), and coronary artery disease (CAD). These diseases are all included in the "diseases of the circulatory system" in the ICD-9 international disease classification codes. After that, we analyze the signal pair quality using the PPG SQI function from the PhysioNet cardiovascular signal toolbox [42] and keep the pair segments that are evaluated as "acceptable" or "excellent."

The resulting dataset consists of 33 patients, with each patient having only one of the four diseases in the record. Each patient has three sessions of 5-min ECG and PPG paired recordings collected within several hours, resulting in $34,000+$ ECG/PPG cycle pairs in total. Table II shows the composition of the collected dataset.

TABLE II
COMPOSITION OF THE COLLECTED DATASET

Disease	Number of patients	Number of cycles
CHF	7	7075 (20.6%)
MI	7	7106 (20.8%)
HYPO	7	8281 (24.2%)
CAD	12	11781 (34.4%)
Total	33	34243 (100%)

CHF: congestive heart failure MI: myocardial infarction
HYPO: hypotension CAD: coronary artery disease

B. Metrics for Evaluation

As shown in Fig. 3 (a), a complete ECG cycle contains five major points, including P, Q, R, S, and T, which segment the ECG cycle into P wave, QRS complex, and T wave. The shape information of those waves is useful for further diagnosis. The interval parameters (PR interval, QRS interval, QT interval) defined by those five fiducial points are also important for examining a patient's heart conditions [48]. Thus, to evaluate the quality of the reconstructed ECG, we take into consideration both the morphological metrics and the accuracy of time interval recovery.

1) Evaluation of Waveform Morphology: We apply the Pearson correlation (ρ) and relative root mean squared error (rRMSE) as the metrics for evaluating the ECG morphological reconstruction. They are defined as follows:

$$\begin{aligned} \rho &= \frac{(\mathbf{x}-\bar{x})^{\mathrm{T}}(\hat{\mathbf{x}}-\bar{\hat{x}})}{\|\mathbf{x}-\bar{x}\|_2 \left\|\hat{\mathbf{x}}-\bar{\hat{x}}\right\|_2}, \\ \mathrm{rRMSE} &= \frac{\|\mathbf{x}-\hat{\mathbf{x}}\|_2}{\|\mathbf{x}\|_2}. \end{aligned} \tag{15}$$

where $\mathbf{x}$, $\hat{\mathbf{x}}$, $\bar{x}$, and $\bar{\hat{x}}$ denote the reference ECG cycle, the recovered ECG cycle, and the average of all coordinates of the vectors $\mathbf{x}$ and $\hat{\mathbf{x}}$, respectively.

2) Evaluation of Time Interval Recovery: Three important ECG interval parameters are studied in this work, including the PR interval, the QRS duration, and the QT interval. Normally, the PR interval lasts 0.12-0.20 seconds, which begins from the onset of the P wave and ends at the beginning of the QRS complex. We use the segment from P point to R point of ECG as the approximated PR interval in this paper. A prolonged PR interval can indicate the possibility of first-degree heart blockage [48]. The duration of the QRS complex is normally 0.12 seconds or less, for ventricular depolarization. A prolonged QRS complex indicates impaired conduction within the ventricles. The QT interval is from the onset of the QRS complex to the end of the T wave, which is normally less than 0.48 seconds. A prolonged QT interval may lead to ventricular tachycardia [48].

We apply a combination of several established algorithms [49]–[51] to detect the major fiducial points of both the ground-truth ECG and the reconstructed ECG to obtain the above-mentioned interval parameters. We apply the mean absolute error (MAE) in Eq. (16) to evaluate the time recovery accuracy:

$$\mathrm{MAE} = \frac{1}{N}\sum_{n=1}^{N}|L_{rec}-L_{ref}|. \tag{16}$$

TABLE III
CONFIGURATIONS OF THE MODELS IMPLEMENTED FOR COMPARISON OF ECG RECONSTRUCTION PERFORMANCE. THE CONFIGURATION INCLUDES SPARSITY CONSTRAINTS ON THE REPRESENTATIONS AND THE LEARNABLE LINEAR MAPPING BETWEEN THE REPRESENTATIONS OF PPG AND ECG.

Reconstruction Scheme	Configuration	
	Sparsity Constraint?	Linear Mapping Between Representations?
DCT [8]	n.a.	✓
CPDL [39]	n.a.	n.a.
ScSR [37]	ℓ_1	n.a.
SCDL [40]	ℓ_1	✓
CDL [38]	ℓ_0	n.a.
XDJDL (**proposed**)	ℓ_0	✓

where the L_{rec} and L_{ref} are the interval length (in seconds) of the reconstructed ECG and ground-truth ECG signals, respectively, and N is the total number of cycles for evaluation.

C. Overall Morphological Reconstruction

We compare our proposed XDJDL method with the state-of-the-art in ECG reconstruction from PPG, which used DCT based method [8]. In addition, we apply several representative and state-of-the-art models of coupled or semi-coupled dictionary learning, including CPDL [39], ScSR [37], SCDL [40], and CDL [38], to compare with the proposed XDJDL method on the ECG reconstruction task. The codes for the prior art are downloaded from the respective authors' websites. The configurations of the prior art methods are listed in Table III. The characteristics of these models can be concluded as (1) the way they represent the signals with any sparsity constraints and (2) whether the cross-domain signal representations are assumed to be identical or linearly related by a learnable mapping. We notice that there is newer related work of dictionary learning, such as [52]. We examine and find the problem setup in [52] essentially falls in the category (1) that being an ℓ_1 constrained problem and (2) where the cross-domain mapping is assumed to be identical. This problem setup is similar to the setting in ScSR [37] and [52] was dealing with a different application scenario.

To make a fair comparison, we evaluate the DCT based reconstruction method in the group-model training mode where a single linear transform $\mathbf{W}_{\mathrm{DCT}}$ is learned using training data from all patients, i.e., a generic model is learned to capture the signal relation for a group of subjects. The normalized PPG/ECG cycle length is chosen as $d = 300$. For XDJDL, the dictionary size for ECG cycles is $k_e = 320$, and the dictionary size for PPG cycles is $k_p = 9000$. The sparsity parameters are set to be $t_e = 10$ and $t_p = 10$. The weights for regularization terms are $\alpha = 1$ and $\beta = 1$. For other dictionary learning models, we have also done grid-search for hyperparameter selection. We split the data from each patient into training and test sets, and the training data ratio is 80%.

Table IV shows the quantitative comparison of the ECG morphological reconstruction performance. From the statistics of the sample mean, standard deviation, and median of ρ and rRMSE, we can see that our proposed XDJDL method outperforms both the DCT based algorithm and other representative coupled/semi-coupled dictionary learning models. Specifically,

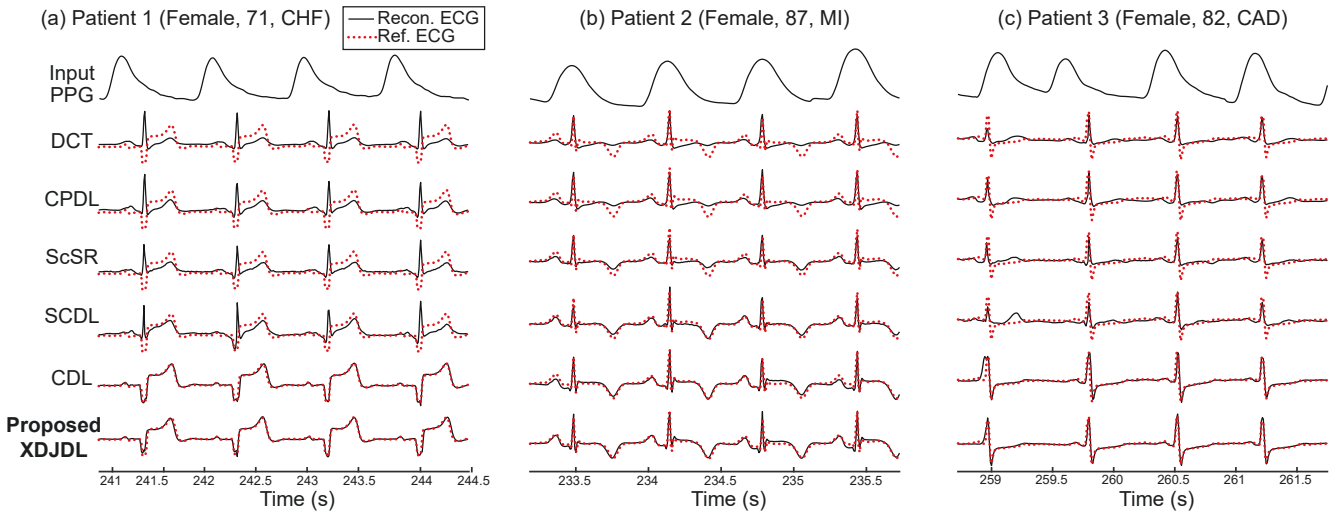


Fig. 2. Qualitative comparison of the ECG signals inferred by different approaches. Examples are (a) a 71-year-old female with congestive heart failure, (b) an 87-year-old female with myocardial infarction, and (c) an 82-year-old female with coronary artery disease. From top to bottom: the input PPG signal from which the ECG is inferred in group-model mode, results by DCT method [8], CPDL [39], ScSR [37], SCDL [40], CDL [38], and our proposed XDJDL.

TABLE IV
QUANTITATIVE PERFORMANCE COMPARISON FOR ECG WAVEFORM INFERENCE

Reconstruction Scheme	ρ			rRMSE		
	$\hat{\mu}$	med	$\hat{\sigma}$	$\hat{\mu}$	med	$\hat{\sigma}$
DCT [8]	0.71	0.83	0.31	0.67	0.60	0.26
CPDL [39]	0.74	0.85	0.31	0.63	0.56	0.35
ScSR [37]	0.82	0.89	0.23	0.54	0.52	0.21
SCDL [40]	0.83	0.89	0.21	0.52	0.49	0.22
CDL [38]	0.85	0.95	0.25	0.49	0.34	0.51
XDJDL (**proposed**)	**0.88**	**0.96**	0.23	**0.39**	**0.29**	0.31

TABLE V
COMPARISON OF ECG SIGNAL INFERENCE AMONG XDJDL, LC1-XDJDL, AND LC2-XDJDL METHODS

Reconstruction Scheme	ρ			rRMSE		
	$\hat{\mu}$	med	$\hat{\sigma}$	$\hat{\mu}$	med	$\hat{\sigma}$
XDJDL	0.88	0.96	0.23	0.39	0.29	0.31
LC1-XDJDL	0.90	0.96	0.20	0.36	0.27	0.28
LC2-XDJDL	**0.92**	**0.97**	0.17	**0.33**	**0.26**	0.25

the average rRMSE is reduced from 0.49 to 0.39, or 20.4% lower than CDL [38], which is the second-best among all competing models.

In Fig. 2, we present three examples of ECG waveform reconstruction to qualitatively compare all the competing models with our proposed XDJDL model. The three patients have different types of disease diagnosis. We observe that even though the waveform variances between the PPGs are relatively smaller than those between the ECGs, our proposed XDJDL method can recover most of the details well in the ECG signal from the PPG signal, suggesting that our method has preserved the intrinsic relation between the atoms from PPG and ECG dictionary pair. In particular, for the second-best CDL [38] method that can reconstruct the overall shape of ECG cycles reasonably well, it has glitches in recovering the details, such as the P wave of the first and last cycle of Patient 2 and the QRS complex of the first cycle of Patient 3.

When the disease information is available, we can apply the proposed label-consistent XDJDL (LC-XDJDL) model from Section III-C to leverage the label information for more accurate monitoring of ECG from the PPG signal. We consider the following scenarios: 1) For cases where the disease information is not directly provided in the test phase, we first predict that from the PPG signals. Here, we have trained an SVM classifier for the PPG multi-class disease classification and chosen the best hyperparameters with a five-fold cross-validation method. The classification accuracy for the PPG test set reaches 92%. We denote the corresponding label-consistent model as LC1-XDJDL. It will take the predicted labels to build the discriminative representation matrix $\mathbf{Q}$. 2) When we have the ground-truth disease labels in the test phase, we can leverage that disease information directly as matrix $\mathbf{Q}$ and the corresponding model is named LC2-XDJDL.

We list the comparison of ECG reconstruction performance using the XDJDL, LC1-XDJDL, and LC2-XDJDL models in Table V. On average, the Pearson coefficient improves from 0.88 to 0.90 with the predicted label information, and to 0.92 with the ground-truth disease type as input. The improvement in terms of the rRMSE is also consistent with the Pearson coefficient. In addition to the reconstruction performance improvement, the label-consistent mapping that relates the PPG sparse codes to disease type in LC-XDJDL helps us understand the role of PPG in diagnosis with a rich ECG knowledge base.

D. Subwave Morphological Reconstruction

In the above subsection, we have shown that our proposed XDJDL outperforms the DCT model and other representative dictionary learning models, and its performance can be better if the disease label (LC-XDJDL) can be utilized for ECG reconstruction and monitoring.

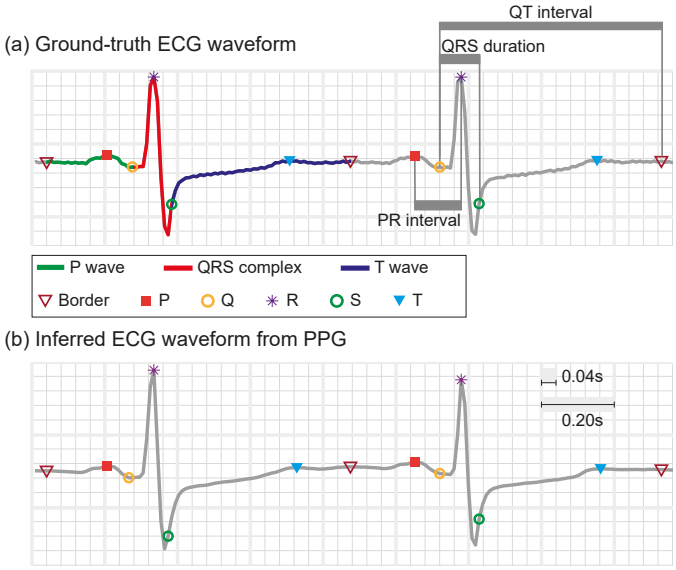


Fig. 3. (a) shows two cycles of the reference ECG signal and (b) shows two cycles of the inferred ECG signal. In the first cycle of (a), the green curve represents the P wave, the red curve is the QRS complex, and the dark blue curve shows the T wave. The PR interval, QRS duration, and the QT interval are all labeled in the second cycle of (a).

In this subsection, we zoom into the reconstruction performance of the subwave of ECG cycles using XDJDL and LC-XDJDL methods. Because each subwave refers to different atrial and ventricular depolarization and re-polarization activities, by zooming in, we can have a better idea of how our methods behave on the inference for different phases of heart activities. A combination of the ECG major point detection algorithms [49]–[51] is used to locate P/Q/R/S/T points of ECG waveform, which helps segment the ECG cycle into subwaves for the evaluation of morphological reconstruction.

Fig. 3 shows an example of the major points detection results on two cycles of the reference ECG (Fig. 3(a)) and the reconstructed ECG (Fig. 3(b)) from a patient with coronary artery disease. In this example, we observe that the locations of the detected major points in both signals are close, indicating a good reconstruction of the ECG waveform. We empirically separate the adjacent ECG cycles at a point that splits the neighboring R-R peaks at the ratio of six to four. After that, a complete ECG cycle is divided into three subwaves, including the P wave that starts from the border point on the left of the ECG cycle and ends at the Q point, the QRS complex from Q to S point, and the T wave from the S point to the right border point. Only a very small portion of reference and reconstructed ECG cycle pairs cannot be detected with a consistent set of fiducial points. The overall number of effective cycles for subwave evaluation is around 92% out of all test cycles, and those effective cycles only have a slightly improved Pearson coefficient (1% on average) compared to the original test dataset.

Table VI lists the reconstruction performance on the three subwaves of the ECG cycle in terms of the mean of Pearson coefficient and rRMSE using XDJDL, LC1-XDJDL, and LC2-XDJDL models. The comparison of results across models is consistent with the results of the overall comparison in Table V. We also observe that the reconstruction for the QRS complex is better than that for the T wave, which is better than that for the P wave. The mean Pearson coefficient of the QRS complex by LC2-XDJDL is 0.94, higher than the overall cycle reconstruction of 0.92, while that of the T wave is slightly lower than the overall performance with the mean Pearson coefficient of 0.89 and that of the P wave is 0.86.

TABLE VI
COMPARISON OF SUBWAVE RECONSTRUCTIONS IN TERMS OF THE MEAN OF ρ AND RRMSE. THE RECONSTRUCTION OF QRS COMPLEX IS THE BEST COMPARED TO THAT OF T WAVE AND P WAVE.

Reconstruction Scheme	$\bar{\rho}$			$\overline{\text{rRMSE}}$		
	P wave	QRS complex	T wave	P wave	QRS complex	T wave
XDJDL	0.81	0.92	0.84	0.53	0.33	0.41
LC1-XDJDL	0.83	0.93	0.86	0.49	0.30	0.37
LC2-XDJDL	0.86	0.94	0.89	0.45	0.28	0.34

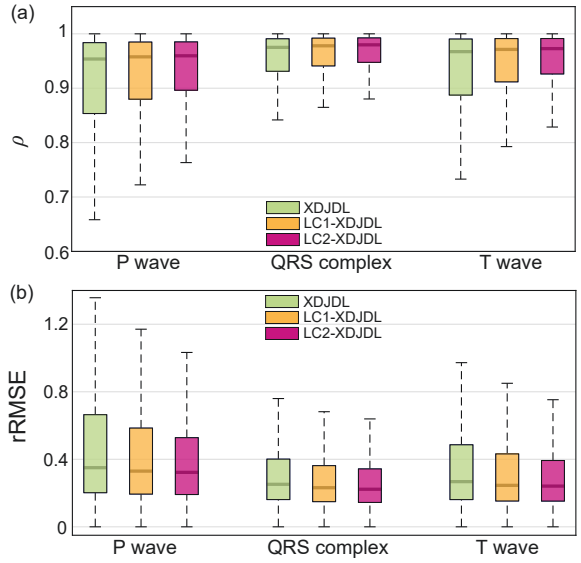


Fig. 4. Comparison of subwave reconstruction performance across XDJDL, LC1-XDJDL, and LC2-XDJDL models. The statistics of (a) Pearson coefficient ρ and (b) rRMSE are summarized using the boxplots.

In addition to the mean of Pearson coefficient and rRMSE, Fig. 4 shows the comparison of the statistics of Pearson coefficient and rRMSE in boxplots for the three subwaves of ECG so that we can see the overall result distribution of the two metrics. We observe that the medians of ρ and rRMSE for each of the three subwaves are very similar across the proposed models. Specifically, the medians of ρ of P wave are 0.95, 0.96, and 0.96, respectively, those of QRS complex are all 0.98, and those of T wave are all 0.97; the medians of rRMSE of P wave are 0.35, 0.33, 0.32, those of QRS complex are 0.25, 0.23, 0.22, and those of the T wave are 0.27, 0.25, and 0.24, respectively. Analysis of these boxplots suggests that our proposed models can preserve the relation between PPG and QRS complex well. The overall reconstruction performance can be improved if the relations between PPG and P and T

TABLE VII
COMPARISON OF TIMING INTERVAL RECOVERY ACCURACY IN MAE.

Reconstruction Scheme	Mean (in seconds)			MAE (in seconds)		
	PR	QRS	QT	PR	QRS	QT
XDJDL	0.164	0.115	0.331	0.030	0.012	0.030
LC1-XDJDL	0.166	0.116	0.331	0.026	0.011	0.027
LC2-XDJDL	0.167	0.115	0.331	0.025	0.010	0.025
Reference	0.172	0.113	0.328	-	-	-

TABLE VIII
QUANTITATIVE COMPARISON OF DIFFERENT SEGMENTATION SCHEMES.

Reconstruction Scheme	ρ			rRMSE		
	$\hat{\mu}$	med	$\hat{\sigma}$	$\hat{\mu}$	med	$\hat{\sigma}$
XDJDL (O2O-1)	0.70	0.84	0.32	0.66	0.57	0.39
XDJDL (O2O-2)	0.80	0.88	0.24	0.55	0.48	0.32
XDJDL (R2R)	0.88	0.96	0.23	0.39	0.29	0.31

waves are better learned.

E. Time Interval Recovery

In addition to the morphological reconstruction evaluation, we evaluate whether the time intervals are well preserved. The labeling of those intervals is shown in Fig. 3.

From columns 2-4 in Table VII, we can compare the average of the reconstructed intervals and the reference intervals. For PR intervals, the difference between the reconstructed and reference is approximately 4%; for QRS durations, such difference is within 3%; and for QT intervals, the difference is less than 1%. This suggests that, on average, the timing information of the intervals is preserved well. From column 5-7 in Table VII, we also notice that the MAEs of the PR interval are 0.030s, 0.026s, and 0.025s using XDJDL, LC1-XDJDL, and LC2-XDJDL models, respectively. The relatively large error in the timing of PR interval recovery is consistent with the result of P wave reconstruction performance shown in Section IV-D. Nevertheless, the MAE of the timing for the QRS complex is around 11ms, which is just a quarter of the smallest grid on the conventional hand copy of ECG recorders (40 ms) and is negligible given the sampling rate (125 Hz) of the ECG signal in the MIMIC III dataset. The MAE of the QT interval is around 27ms, which is less than three-quarters of the smallest grid on ECG graph paper and is around 8% of the QT interval (0.331s).

V. DISCUSSIONS

A. Result Using PPG-based Segmentation Scheme

In Section IV, we have evaluated our proposed models based on the assumption that the cycle information from ECG signals is available to separate the ECG/PPG time-series signals into training and test cycles. But in practice, we may not have the groundtruth of cycle segmentation from ECG. Thus, we consider such realistic scenarios of reconstructing the ECG from the "estimated cycles" of PPG that are segmented by the PPG onsets instead of the R peaks of ECG signals. The PPG onsets are used for segmentation rather than the PPG peaks because of the underlying physiological meaning as we have mentioned in Section III-A. For ease of notation, we denote:

- R2R: segmentation scheme based on R peaks of ECG for both training and test data, which is used in Section IV;
- O2O-1: segmentation scheme based on PPG onsets for both training and test data;
- O2O-2: segmentation scheme based on R peaks of ECG for training data and based on PPG onsets for test data.

Due to the discrepancy between the detected locations of PPG onset and R peak of ECG from the same cycle, the "estimated PPG cycles" using O2O schemes slightly vary from the PPG cycles which are segmented by R2R. To single out the contribution to the ECG reconstruction error due to the discrepancy in the waveform shape rather than the misalignment of the ECG peaks, we evaluate O2O schemes after compensating the time offset between the reconstructed ECG and original ECG signals. This is done by shifting each reconstructed ECG cycle in time so that the original and reconstructed ECG signals are matched according to their R peaks. The comparison result is shown in Table VIII. Compared to R2R, when using the O2O-1 scheme, the average Pearson coefficient drops from 0.88 to 0.70, and the average rRMSE rises from 0.39 to 0.66. And using the O2O-2 scheme can help improve the performance compared to O2O-1, where the mean Pearson coefficient becomes 0.80 and the mean rRMSE becomes 0.55.

B. Evaluation on the Capnobase TBME-RR Dataset

In this subsection, we experimented with the Capnobase TBME-RR database [53] that contains forty-two eight-minute sessions from 29 children and 13 adults during elective surgery and routine anesthesia. Each session corresponds to a unique participant and contains simultaneously recorded PPG and ECG signals. The signals are recorded with a sampling frequency of 300 Hz. The dataset covers a wide range of participant's age, which is from one-year-old to sixty-three-year-old with the median age being fourteen. Thus, this dataset is used for a supplementary evaluation of the proposed method from the angle of age variety in addition to disease variety in the MIMIC-III dataset.

We first pruned the signals according to the artifact labels provided in the dataset and preprocessed the signals using the method in Section III-A to obtain aligned and normalized signal pairs. To be consistent in the evaluation, as in Section IV, we selected the first 80% of the data from each subject as the training set and the rest for testing.

Table IX summarizes the performance comparison using the Capnobase TBME-RR dataset. Our proposed XDJDL method outperforms all the other groups in terms of the mean and median rRMSE by a large margin. Even though the CDL [38] method is 0.1% better than our proposed method in mean and median correlation coefficient ρ, our method achieves 26% smaller $\hat{\sigma}$ of ρ than the CDL method, showing that our proposed method achieves good performance of ECG reconstruction more consistently for all participants.

TABLE IX
QUANTITATIVE PERFORMANCE COMPARISON FOR ECG WAVEFORM INFERENCE USING THE CAPNOBASE TBME-RR DATABASE

Reconstruction Scheme	ρ			rRMSE		
	$\hat{\mu}$	med	$\hat{\sigma}$	$\hat{\mu}$	med	$\hat{\sigma}$
DCT [8]	0.902	0.919	0.066	0.427	0.413	0.128
CPDL [39]	0.956	0.968	0.049	0.282	0.247	0.150
ScSR [37]	0.967	0.976	0.039	0.286	0.247	0.165
SCDL [40]	0.971	0.978	0.038	0.191	0.166	0.101
CDL [38]	**0.980**	**0.991**	0.062	0.219	0.145	0.296
XDJDL (**proposed**)	0.979	0.990	0.048	**0.146**	**0.105**	0.122

C. Feasibility Analysis of The Proposed Method for The Internet-of-Healthcare-Things (IoHT)

In this subsection, we analyze two important practicality issues when applying our proposed ECG reconstruction techniques to healthcare IoT devices. One issue is energy consumption. The sensors used to capture physiological signals, e.g., PPG signals, are mostly wearable devices, which are powered by batteries [9]. Thus, being energy-efficient is necessary to ensure continuous signal acquisition, data transmission, and monitoring. The other issue is computational cost. As mentioned in [9], [54], applications that require lower latency need higher computational capabilities. Thus, the computational load of the algorithms needs to be considered in real-world scenarios.

The first issue about energy consumption in wearable devices can be resolved by the existing mature technologies like the Bluetooth low-energy module commonly applied for low-power wireless communication in wearable healthcare devices [55]. In the test phase of our proposed XDJDL and LC-XDJDL frameworks, PPG signals acquired by wearable devices can be transmitted to the IoT devices, such as smartphones, at low power with the help of those modules. For the second issue about computational cost, with the dictionary pairs constructed locally and stored in the cloud or edge devices, the computational cost is mainly from sparse coding and lightweight matrix multiplication. Since sparse coding via OMP in our proposed methods is proven to be able to be executed on the IoT platform in real-time [18], we envision that our proposed frameworks can satisfy the practical requirements well.

To further evaluate quantitatively the feasibility of applying our proposed method to IoHT platforms, we examine the following metrics to measure the usage of computational resources to reconstruct one ECG cycle:

1) Computational time
2) Energy consumption
3) Memory space

The specifications of the laptop we used for the experiment are as follows: Processor: i7-8650U; Architecture: Intel x86; CPU Frequency: 1.90GHz; Cores: 4; RAM: 24GB. Our test here is designed to resemble an online inference scenario in which new sequences of continuous ECG waveform need to be inferred by the IoHT system with the input PPG waveform. The experiment is repeated 100 times to evaluate the computational time and memory space for each cycle. Note that the actual energy consumption estimation can be complex, as it depends on the operating system, the temperature inside and outside the device, and the efficiency of the power supply. Thus, we use FLOP (Floating-point Operations) here as the measure for energy consumption, as it is independent of hardware configurations given the algorithm. With FLOP, the energy in joule can be estimated as it is proportional to FLOP given FLOPS (FLOP per Second) per watt, i.e., FLOPS/W, specified by the IoT device.

TABLE X
COMPARISON IN TERMS OF COMPUTATIONAL COMPLEXITY, ACCURACY, AND EXPLAINABILITY TO RECONSTRUCT TEST ECG CYCLES.

Reconstruction Scheme	FLOP Consumption	Number of Model Parameters	$\bar{\rho}$	$\overline{r\text{RMSE}}$	Explainability
DCT [8]	0.36M	0.27M	0.71	0.67	✓
XDJDL (proposed)	60.21M	5.67M	0.88	0.39	✓
LC2-XDJDL (proposed)	64.72M	5.90M	0.92	0.33	✓
TANN [25]	518.79M	10.98M	0.94	0.28	×

The computational resources consumed by our proposed methods are as follows: 1) The average computational time for generating one ECG cycle is 15.67 ms and 18.91 ms using XDJDL and LC2-XDJDL methods, respectively. This processing time is one to two orders of magnitude shorter than a heart cycle (around 0.5s to 1s per beat at rest), suggesting that the processing can be done in real-time; 2) The 60 MFLOP/65 MFLOP and 31 MB/43 MB memory space required by the proposed XDJDL/LC2-XDJDL methods are well within the capability of such commonly seen IoT platforms as the Raspberry Pi 3B (RAM: 1 GB, 0.73 GFLOPS/W) [56] for the research prototype that has not been optimized for deployment. Considerable reductions in computing resources are possible with industry-grade implementation.

To evaluate the trade-off between the accuracy and computational cost, we consider the comparison with the DCT method [8] and a deep learning method (TANN) for ECG reconstruction from PPG [25]. TANN [25] was tested using the source code implementation shared by its authors. Since TANN [25] and our proposed approaches are implemented on different platforms (Python versus MATLAB), we mainly focus on the FLOP and the number of model parameters as a surrogate of the memory space for a fair comparison of computational complexity. The FLOP counts the floating-point operations consumed by each algorithm in synthesizing a 300-point ECG cycle from a PPG cycle with the same length. From Table X, we conclude that the DCT method is the most lightweight PPG-to-ECG algorithm and has an advantage in computational consumption while its accuracy is sacrificed by a large margin. As for TANN [25], even though its accuracy performance is 2% better than LC2-XDJDL in terms of mean Pearson coefficient, TANN's FLOP consumption is about an order of magnitude higher than that of our proposed methods and its number of model parameters almost double ours. This comparison suggests that compared to neural network approaches, our algorithm has higher computational efficiency, which facilities the deployment of this new type of ECG monitoring on IoT platforms and supports real-time applications that require lower latency. In light of the advantage of deep learning in representing the multi-scale

and fine-granular details of physiological signals, a forward-looking direction for our ongoing and future effort include to explore the lightweight implementation of neural network based PPG-to-ECG mapping, aiming at achieving a better trade-off between the precision of ECG reconstruction and the time-space costs. We have also indicated the explainability of the proposed methods and the comparison models in the last column of Table X. It is worth noting that TANN is lack of explainability due to its black-box design of the architectures, while our proposed methods are based on the underlying biomedical and statistical relations of PPG and ECG signals and strive for interpretability.

It is not easy to make a fair comparison with the other comparative algorithms based on dictionary learning [37]–[40] in terms of computational resources. This is because 1) Their inference stages require solving optimization problems that have already been implemented in the optimized software, such as .mex files. The .mex files only provide interfaces to execute the external subroutines developed by C/C++, but the source codes for solving the optimization problems are not accessible, making it infeasible to count the FLOP of these algorithms; and 2) They were mainly proposed in the computer vision domain rather than specifically targeted at PPG-to-ECG application.

D. Limitations of The Proposed Method

1) Performance of Leave-One-Out Experiment: As a proof of concept and considering the current moderate amount of available data, we have so far split each patient's data into training and test sets. This corresponds to the trend of "precision medicine" to tailor the healthcare practice to individual patients. Meanwhile, we are curious how the algorithm would behave if the test patient is never seen in the training phase, corresponding to the situation of training models for the whole population or patient groups categorized by gender, age, race, or other ways. We will examine this through leave-one-out experiments.

We apply a pre-clustering process based on the ECG data to select a sub-group of patients with similar ECG features for the leave-one-out experiment. First, we reduce the dimension of the ECG cycles by principal component analysis (PCA), and then we use K-means to cluster the ECG features after PCA. Based on the clustered ECG features, we select the largest cluster of ECGs from 19 patients. The mean Pearson coefficient for the leave-one-out experiment on the 19 patients is 0.74 (std: 0.15, median: 0.77).

From the result, we can see that as expected, the leave-one-out experiment is a more challenging case given the large variability of ECG data morphologies of ICU patients and the limited number of patients in the collected dataset. Based on the results in Section IV-C, we see the encouraging capability of recovering large variations in ECG from relatively small variations in PPG across cycles and patient populations. This suggests a strong potential for predicting ECG from PPG of unseen patients through further research and larger data collection. In our follow-up work, we are considering an improved problem definition and data collection procedure to enhance the generalization capability of learning.

2) Performance Evaluation on A Motion Dataset: So far, we have demonstrated in this paper the feasibility and improved accuracy of ECG waveform inference from PPG using the proposed methods on two benchmark datasets [46], [53] in Section IV-C and Section V-B. Those datasets were collected under a resting condition with relatively small movement artifacts. Noises and artifacts were still present in those datasets but in a controlled manner, which leads to good quality of data acquisition and is beneficial for the feasibility study and accuracy improvement of reconstructing ECG from PPG. In this subsection, we consider a more challenging scenario where IoHT devices are worn during exercise and show the preliminary results with the motion-contaminated signals.

We adopt the 2015 IEEE Signal Processing Cup dataset [57] for evaluation, which consists of paired PPG and ECG signals from 13 participants during physical exercises. This dataset provided by the Samsung Research Lab in the U.S. aimed to facilitate the study of accurate heart rate (HR) monitoring of PPG signals from wrist-type sensors and included ECG signals as a reference. The PPG signals were collected from the wrist while the subjects ran on a treadmill at speeds of 6 km/h, 8 km/h, 12 km/h, or 15 km/h, respectively. Simultaneously, the ECG signals were collected from the chest and the acceleration signal was recorded from the wrist by a three-axis accelerometer. All signals were sampled at 125 Hz. Each subject ran once and the total length of the recording was 5 minutes per subject.

Since the quality of PPG signals is crucial to ECG reconstruction, we first use the absolute error of the PPG estimated HR as a metric to exclude the participants with extremely corrupted PPG signals. Because HR represents the frequency characteristic of PPG that affects the accuracy of determining a PPG cycle. The HR is estimated from the PPG by a state-of-the-art adaptive multi-trace carving (AMTC) [58] algorithm that tracks the HR from the spectrogram of PPG by dynamic programming and adaptive trace compensation. The reference HR values are given in the dataset. Three out of the thirteen participants are excluded as their HR estimation error is quite off likely due to data collection issues and the remaining ten participants' data are used for learning and testing the XDJDL model. In addition, to improve the quality of noise-contaminated PPG, we employ an off-the-shelf denoising algorithm by recursive least square (RLS) adaptive filtering [8], [59]. It treats the contaminated PPG as the sum of the underlying cleaned PPG and motion-induced noise. Since the acceleration signals are correlated with the motion, they are used to estimate the motion artifacts for PPG noise removal. We will compare the ECG reconstruction performance from PPG signals before and after denoising.

Fig. 5(a) shows the comparison of the statistics of Pearson coefficient and rRMSE in boxplots for ECG reconstructed from the PPG signal without denoising (referred to as "raw PPG") and RLS filtered PPG signal (referred to as "cleaned PPG"). The average Pearson coefficient of the reconstructed ECG using raw PPG is 0.49 (median: 0.69, std: 0.51) and using cleaned PPG is improved to 0.61 (median: 0.72, std: 0.37). This improvement can be attributed to that the spurious peaks and waves in the motion-contaminated PPG are removed by

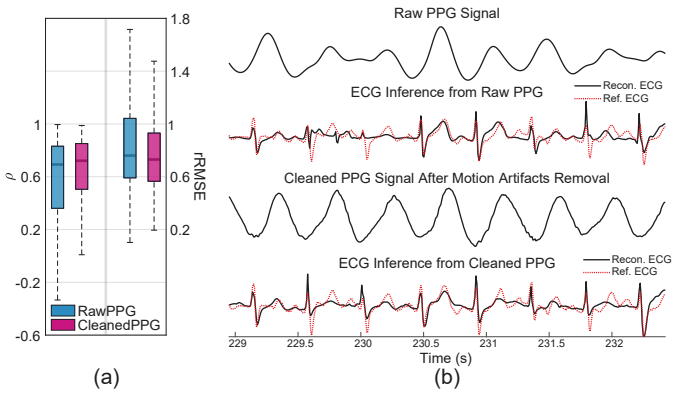


Fig. 5. (a) Statistical distribution of Pearson coefficient (ρ) and rRMSE for reconstructed ECG from PPG signals before denoising ("raw PPG") and PPG signals after denoising ("cleaned PPG"). (b) Qualitative comparison of raw PPG, cleaned PPG, and the ECG signals inferred from them.

the RLS filtering. While the noise due to motion is mitigated, distortions in PPG and even ECG waveforms are still present as shown in Fig. 5(b). Treating potentially corrupted ECG as the reference and distorted PPG as the input might misguide the learning system, and produce unreliable waveform reconstruction. Fig. 5(b) shows an example with close to average performance. We observe that on one hand, the cleaned PPG has clearer cycle shapes than the raw PPG; and on the other hand, some of the physiological characteristics representing the blood flow process are irregular after RLS motion artifacts removal, such as the peak in the third cycle and the ascending and descending slopes in the fifth cycle. Also, the reference ECG signals contain varying ST segment elevations over consecutive cycles during motion. We expect such limitations can be addressed with the development of more advanced PPG and ECG denoising and waveform preserving approaches for preprocessing and the availability of a larger dataset under different types of activities (such as walking, running, driving, climbing stairs, etc).

E. Future Work Towards Explainable AI

Our proposed XDJDL and LC-XDJDL models accomplished to infer the ECG based on PPG by leveraging the biomedical and statistical relationship between the signals. This is an initial effort to demonstrate a potential benefit from our "explainable" AI, rather than black-box data-driven AI, to provide the more user-friendly PPG measurements inferred ECG data for the medical professionals to interpret and offer medical insights. Our framework also helps transfer the rich ECG knowledge base from decades of medical practice to augment the PPG diagnosis for public health.

Given the challenge of making the ECG inference more accurate for an unseen group of subjects, e.g., by age, gender, or other medical and health condition, we are extending our current work with a neural network to further enrich the representation and learn the relation when sufficient data is available. Our ongoing efforts have been focused on both developing a data collection pipeline for more diversity and coverage of training data and exploring an explainable generative model with strong expressive power to improve the generalization performance. With the step-by-step capturing of complex models by utilizing the biomedical, statistical, and physical meanings, as well as harnessing the power of the data, we aim to provide explainable AI with our ongoing efforts.

VI. Conclusions

We have proposed a cross-domain joint dictionary learning (XDJDL) framework and the extended label-consistent XDJDL (LC-XDJDL) model for ECG waveform inference from the PPG signal. Compared to the prior art using the DCT method, our proposed method better leverages the data to improve data representation while extending over a model-based approach. The promising experimental results with the average Pearson coefficient being 0.88 using XDJDL and 0.92 using LC-XDJDL validate that our proposed models can learn the relation between PPG and ECG and reconstruct ECG well. From the analysis for subwave reconstruction and timing of interval recovery, we observe that we can restore the QRS complex and the QT interval in high precision, which is essential for ECG monitoring and to gain more PPG-based diagnosis knowledge. This work reveals the potential of long-term and user-friendly ECG screening from the PPG signals that we can acquire from the daily use of low-cost, low-power wearable devices for IoT and digital twins applications in healthcare.

Our ongoing work and future endeavor will strive to improve the accuracy and generalization capability of the ECG waveform inference from PPG while maintaining sound interpretability. Future research and development may consider improving the robustness to exercise conditions by developing more advanced PPG and ECG denoising and waveform preserving algorithms and collecting datasets under different types of activities.

References

[1] X. Tian, Q. Zhu, Y. Li, and M. Wu, "Cross-Domain Joint Dictionary Learning for ECG Reconstruction from PPG," in *IEEE Int. Conf. Acoustics, Speech and Signal Process. (ICASSP)*, 2020, pp. 936–940.

[2] "Cardiovascular diseases (CVDs)," https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds), Accessed: 2021-09-27.

[3] A. Rosiek and K. Leksowski, "The Risk Factors and Prevention of Cardiovascular Disease: The Importance of Electrocardiogram in the Diagnosis and Treatment of Acute Coronary Syndrome," *Therapeutics and clinical risk management*, vol. 12, p. 1223, 2016.

[4] "How Zio Works," https://www.irhythmtech.com/patients/how-it-works, Accessed: 2020-12-25.

[5] "Take an ECG with the ECG app on Apple Watch," https://support.apple.com/en-us/HT208955, Accessed: 2022-1-25.

[6] "KardiaMobile: Check in on your heart from home," https://store.kardia.com/products/kardiamobile, Accessed: 2022-1-25.

[7] Q. Zhu, X. Tian, C.-W. Wong, and M. Wu, "ECG Reconstruction via PPG: A Pilot Study," in *IEEE EMBS Int. Conf. Biomed. & Health Informatics (BHI)*, Chicago, IL, May 2019.

[8] ——, "Learning Your Heart Actions From Pulse: ECG Waveform Reconstruction From PPG," *IEEE Internet of Things Journal*, vol. 8, no. 23, pp. 16 734–16 748, 2021.

[9] H. Habibzadeh, K. Dinesh, O. Rajabi Shishvan, A. Boggio-Dandry, G. Sharma, and T. Soyata, "A Survey of Healthcare Internet of Things (HIoT): A Clinical Perspective," *IEEE Internet of Things J.*, 2020.

[10] J. Allen, "Photoplethysmography and Its Application in Clinical Physiological Measurement," *Physio. Measurement*, 2007.

[11] A. K. Joshi, A. Tomar, and M. Tomar, "A Review Paper on Analysis of Electrocardiograph (ECG) Signal for the Detection of Arrhythmia Abnormalities," *Int. J. Advanced Research in Electrical, Electronics and Instrumentation Eng.*, 2014.

[12] E. Gil, M. Orini, R. Bailon, J. M. Vergara, L. Mainardi, and P. Laguna, "Photoplethysmography Pulse Rate Variability as A Surrogate Measurement of Heart Rate Variability During Non-stationary Conditions," *Physio. Measurement*, 2010.

[13] A. Johansson, "Neural Network for Photoplethysmographic Respiratory Rate Monitoring," *Medical and Biological Eng. and Comput.*, 2003.

[14] E. C.-P. Chua, S. J. Redmond, G. McDarby, and C. Heneghan, "Towards Using Photo-plethysmogram Amplitude to Measure Blood Pressure During Sleep," *Annals of Biomed. Eng.*, 2010.

[15] K. Bruynseels, F. Santoni de Sio, and J. van den Hoven, "Digital Twins in Health Care: Ethical Implications of An Emerging Engineering Paradigm," *Frontiers in genetics*, vol. 9, p. 31, 2018.

[16] J. Yang, Z. Wang, Z. Lin, S. Cohen, and T. Huang, "Coupled Dictionary Training for Image Super-resolution," *IEEE Trans. Image Process.*, 2012.

[17] Z. Li, H. Huang, and S. Misra, "Compressed Sensing via Dictionary Learning and Approximate Message Passing for Multimedia Internet of Things," *IEEE Internet of Things J.*, 2017.

[18] M. Al Disi, H. Djelouat, C. Kotroni, E. Politis, A. Amira, F. Bensaali, G. Dimitrakopoulos, and G. Alinier, "ECG Signal Reconstruction on the IoT-gateway and Efficacy of Compressive Sensing Under Real-Time Constraints," *IEEE Access*, 2018.

[19] G. Zhang, Z. Mei, Y. Zhang, X. Ma, B. Lo, D. Chen, and Y. Zhang, "A Noninvasive Blood Glucose Monitoring System Based on Smartphone PPG Signal Processing and Machine Learning," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 11, pp. 7209–7218, 2020.

[20] R. Banerjee, A. Sinha, A. D. Choudhury, and A. Visvanathan, "PhotoECG: Photoplethysmography to Estimate ECG Parameters," in *IEEE Int. Conf. Acoustics, Speech and Signal Process.*, 2014.

[21] A. Y. Hannun, P. Rajpurkar, M. Haghpanahi, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level Arrhythmia Detection and Classification in Ambulatory Electrocardiograms Using a Deep Neural Network," *Nature medicine*, vol. 25, no. 1, pp. 65–69, 2019.

[22] U. R. Acharya, H. Fujita, S. L. Oh, Y. Hagiwara, J. H. Tan, and M. Adam, "Application of Deep Convolutional Neural Network for Automated Detection of Myocardial Infarction Using ECG Signals," *Information Sciences*, vol. 415, pp. 190–198, 2017.

[23] S. K. Bashar, D. Han, S. Hajeb-Mohammadalipour, E. Ding, C. Whitcomb, D. D. McManus, and K. H. Chon, "Atrial Fibrillation Detection from Wrist Photoplethysmography Signals Using Smartwatches," *Scientific reports*, vol. 9, no. 1, pp. 1–10, 2019.

[24] N. Paradkar and S. R. Chowdhury, "Cardiac Arrhythmia Detection Using Photoplethysmography," in *2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2017, pp. 113–116.

[25] H.-Y. Chiu, H.-H. Shuai, and P. C.-P. Chao, "Reconstructing QRS Complex From PPG by Transformed Attentional Neural Networks," *IEEE Sensors Journal*, vol. 20, no. 20, pp. 12 374–12 383, 2020.

[26] K. Vo, E. K. Naeini, A. Naderi, D. Jilani, A. M. Rahmani, N. Dutt, and H. Cao, "P2E-WGAN: ECG waveform synthesis from PPG with conditional wasserstein generative adversarial networks," in *Proceedings of the 36th Annual ACM Symposium on Applied Computing*, 2021, pp. 1030–1036.

[27] M. Aharon, M. Elad, A. Bruckstein *et al.*, "K-SVD: An Algorithm for Designing Overcomplete Dictionaries for Sparse Representation," *IEEE Trans. Signal Process.*, 2006.

[28] K. Engan, S. O. Aase, and J. H. Husoy, "Method of Optimal Directions for Frame Design," in *IEEE Int. Conf. Acoustics, Speech, and Signal Process.*, 1999.

[29] J. Mairal, J. Ponce, G. Sapiro, A. Zisserman, and F. R. Bach, "Supervised Dictionary Learning," in *Advances in Neural Inf. Process. Systems*, 2009.

[30] T. Liu, Y. Si, D. Wen, M. Zang, and L. Lang, "Dictionary Learning for VQ Feature Extraction in ECG Beats Classification," *Expert Systems with Applications*, 2016.

[31] A. Majumdar and R. Ward, "Robust Greedy Deep Dictionary Learning for ECG Arrhythmia Classification," in *IEEE Int. Joint Conf. Neural Netw.*, 2017.

[32] D. Craven, B. McGinley, L. Kilmartin, M. Glavin, and E. Jones, "Adaptive Dictionary Reconstruction for Compressed Sensing of ECG Signals," *IEEE J. Biomed. Health Informatics*, 2016.

[33] H. Tang, H. Liu, W. Xiao, and N. Sebe, "When Dictionary Learning Meets Deep Learning: Deep Dictionary Learning and Coding Network for Image Recognition With Limited Data," *IEEE Trans. Neural Netw. Learning Systems*, 2021.

[34] Z. Li, Z. Lai, Y. Xu, J. Yang, and D. Zhang, "A Locality-Constrained and Label Embedding Dictionary Learning Algorithm for Image Classification," *IEEE Trans. Neural Netw. Learning Systems*, 2017.

[35] Y. Sun, Z. Zhang, W. Jiang, Z. Zhang, L. Zhang, S. Yan, and M. Wang, "Discriminative Local Sparse Representation by Robust Adaptive Dictionary Pair Learning," *IEEE Trans. Neural Netw. Learning Systems*, vol. 31, no. 10, pp. 4303–4317, 2020.

[36] Z. Chen, X.-J. Wu, and J. Kittler, "Relaxed Block-Diagonal Dictionary Pair Learning With Locality Constraint for Image Recognition," *IEEE Trans. Neural Netw. Learning Systems*, 2021.

[37] J. Yang, J. Wright, T. S. Huang, and Y. Ma, "Image Super-resolution via Sparse Representation," *IEEE Trans. Image Process.*, 2010.

[38] J. Xu, C. Qi, and Z. Chang, "Coupled K-SVD Dictionary Training for Super-resolution," in *IEEE Int. Conf. Image Process.*, 2014.

[39] K. Li, Z. Ding, S. Li, and Y. Fu, "Discriminative Semi-coupled Projective Dictionary Learning for Low-resolution Person Re-identification," in *Thirty-Second AAAI Conference on Artificial Intelligence*, 2018.

[40] S. Wang, L. Zhang, Y. Liang, and Q. Pan, "Semi-coupled Dictionary Learning with Applications to Image Super-resolution and Photo-sketch Synthesis," in *IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, 2012.

[41] E. A. Ashley and J. Niebauer, *Cardiology explained*. Remedica, 2004.

[42] A. N. Vest, G. Da Poian, Q. Li, C. Liu, S. Nemati, A. J. Shah, and G. D. Clifford, "An Open Source Benchmarked Toolbox for Cardiovascular Waveform and Interval Analysis," *Physio. Measurement*, 2018.

[43] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A Survey of Sparse Representation: Algorithms and Applications," *IEEE Access*, 2015.

[44] J. A. Tropp and A. C. Gilbert, "Signal Recovery from Random Measurements via Orthogonal Matching Pursuit," *IEEE Trans. Inf. Theory*, 2007.

[45] Z. Jiang, Z. Lin, and L. S. Davis, "Label Consistent K-SVD: Learning a Discriminative Dictionary for Recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, 2013.

[46] A. E. Johnson, T. J. Pollard, L. Shen, H. L. Li-wei, M. Feng, M. Ghassemi, B. Moody, P. Szolovits, L. A. Celi, and R. G. Mark, "MIMIC-III, A Freely Accessible Critical Care Database," *Scientific Data*, 2016.

[47] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of A New Research Resource for Complex Physiologic Signals," *Circulation*, 2000.

[48] J. Hampton and J. Hampton, *The ECG Made Easy E-Book*. Elsevier Health Sciences, 2019.

[49] H. Sedghamiz, "BioSigKit: A Matlab Toolbox and Interface for Analysis of BioSignals," *J. Open Source Software*, 2018.

[50] J. Pan and W. J. Tompkins, "A Real-Time QRS Detection Algorithm," *IEEE Trans. Biomed. Eng.*, 1985.

[51] H. Sedghamiz and D. Santonocito, "Unsupervised Detection and Classification of Motor Unit Action Potentials in Intramuscular Electromyography Signals," in *IEEE E-health and Bioengineering Conf.*, 2015.

[52] Z. Gao, H.-Z. Xuan, H. Zhang, S. Wan, and K.-K. R. Choo, "Adaptive fusion and category-level dictionary learning model for multiview human action recognition," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 9280–9293, 2019.

[53] W. Karlen, S. Raman, J. M. Ansermino, and G. A. Dumont, "Multiparameter Respiratory Rate Estimation from the Photoplethysmogram," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 7, pp. 1946–1953, 2013.

[54] M. Asif-Ur-Rahman, F. Afsana, M. Mahmud, M. S. Kaiser, M. R. Ahmed, O. Kaiwartya, and A. James-Taylor, "Toward a Heterogeneous Mist, Fog, and Cloud-Based Framework for the Internet of Healthcare Things," *IEEE Internet of Things J.*, 2019.

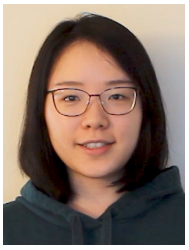
[55] T. Wu, F. Wu, C. Qiu, J.-M. Redouté, and M. R. Yuce, "A Rigid-Flex Wearable Health Monitoring Sensor Patch for IoT-Connected Healthcare Applications," *IEEE Internet of Things Journal*, 2020.

[56] Z. U. Ahmed, M. G. Mortuza, M. J. Uddin, M. H. Kabir, M. Mahiuddin, and M. J. Hoque, "Internet of Things based patient health monitoring system using wearable biomedical device," in *2018 international conference on innovation in engineering and technology (ICIET)*. IEEE, 2018, pp. 1–5.

[57] Z. Zhang, Z. Pi, and B. Liu, "TROIKA: A General Framework for Heart Rate Monitoring Using Wrist-type Photoplethysmographic Signals During Intensive Physical Exercise," *IEEE Trans. Biomed. Eng.*, 2014.

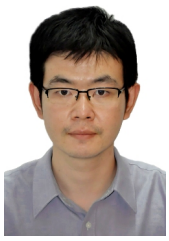
[58] Q. Zhu, M. Chen, C.-W. Wong, and M. Wu, "Adaptive Multi-trace Carving for Robust Frequency Tracking in Forensic Applications," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 1174–1189, May 2020.

[59] E. Khan, F. Al Hossain, S. Z. Uddin, S. K. Alam, and M. K. Hasan, "A Robust Heart Rate Monitoring Scheme Using Photoplethysmographic Signals Corrupted by Intense Motion Artifacts," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 550–562, 2016.



Xin Tian (Member, IEEE) received the B.E. degree in Optoelectronic Information Science and Engineering from Huazhong University of Science and Technology of China in 2017 and her Ph.D. degree at the Department of Electrical and Computer Engineering, University of Maryland, College Park, USA, in 2022. Her research interests are signal processing, data science, and machine learning in smart health. She is a research scientist at Meta since 2022. She was selected as a Future Faculty Fellow and was awarded as an Outstanding Graduate Assistant by the University of Maryland in 2020 and 2021, respectively.

Qiang Zhu (S'17–M'20) received his B.E. degree in control science and engineering from Zhejiang University, Hangzhou, China, in 2010, his M.S. degree in control science and engineering from Shanghai Jiao Tong University, Shanghai, China, in 2014, and his Ph.D. degree in electrical engineering from the University of Maryland, College Park, USA, in 2020. He is a research scientist at Meta since 2020. His research interests are signal processing, machine learning, and information retrieval. He received the Distinguished Teaching Assistance Award in 2016 from the University of Maryland.



Yuenan Li (M'13–SM'21) received his B. Eng. and M. Eng. degrees in measurement technology and instrument, and the Ph.D. degree in information and communication engineering, in 2004, 2006, and 2010, respectively, all from Harbin Institute of Technology, China. He is currently an Associate Professor at the School of Electrical and Information Engineering, Tianjin University, China. From 2019 to 2020, he was a visiting researcher at the University of Maryland, College Park, USA. His research interests include multimedia signal processing, information security and forensics, computer vision, and physiological signal processing.

Min Wu (Fellow, IEEE) received the B.E. degree in automation and the B.A. degree in economics from Tsinghua University, Beijing, China, in 1996 with the highest honors, and the Ph.D. degree in electrical engineering from Princeton University, Princeton, NJ, USA, in 2001. She is a Professor and an Associate Dean of Engineering and a Distinguished Scholar-Teacher with the University of Maryland, College Park, MD, USA. Her research interests include information security and forensics, multimedia signal processing, and applications of data science and machine learning in health and IoT. Prof. Wu received a U.S. NSF CAREER award, a U.S. ONR Young Investigator Award, a TR100 Young Innovator Award from the MIT Technology Review, an IEEE Harriett B. Rigas Education Award, and an IEEE SP Society Meritorious Service Award. She chaired the IEEE Technical Committee on Information Forensics and Security, and has served as the Vice President—Finance of the IEEE Signal Processing Society and the Editor-in-Chief of the IEEE Signal Processing Magazine. She has been elected as 2022-2023 President-Elect of the IEEE Signal Processing Society. She is a Fellow of AAAS and of the U.S. National Academy of Inventors.