

Multiple Instance Dictionary Learning for Beat-to-Beat Heart Rate Monitoring from Ballistocardiograms

Changzhe Jiao, *Student Member, IEEE*, Bo-Yu Su, *Student Member, IEEE*, Princess Lyons, *Student Member, IEEE*, Alina Zare, *Senior Member, IEEE*, K. C. Ho, *Fellow, IEEE*, and Marjorie Skubic, *Senior Member, IEEE*

Abstract—A multiple instance dictionary learning approach, Dictionary Learning using Functions of Multiple Instances (DL-FUMI), is used to perform beat-to-beat heart rate estimation and to characterize heartbeat signatures from ballistocardiogram (BCG) signals collected with a hydraulic bed sensor. DL-FUMI estimates a “heartbeat concept” that represents an individual’s personal ballistocardiogram heartbeat pattern. DL-FUMI formulates heartbeat detection and heartbeat characterization as a multiple instance learning problem to address the uncertainty inherent in aligning BCG signals with ground truth during training. Experimental results show that the estimated heartbeat concept found by DL-FUMI is an effective heartbeat prototype and achieves superior performance over comparison algorithms.

Index Terms—Ballistocardiogram, bed sensor, heart rate, Multiple Instance Learning, Dictionary Learning, target detection, target characterization, heartbeat characterization

I. INTRODUCTION

Increasingly more and more devices for realtime heart rate monitoring are becoming available. However, the majority of these devices are intrusive and require continual interaction. For example many heart rate monitoring systems require a user to physically wear the system (*e.g.*, as a watch, chest strap, electrodes, finger sensor, *etc.*) and/or charge batteries frequently. In contrast, devices that use ballistocardiography can provide an unintrusive and, thus, relatively low maintenance, comfortable alternative for heart rate monitoring. These sensing systems record the motion of the human body generated by the sudden ejection of blood into the large vessels at each cardiac cycle [1]. Such motion contains rich information and has gained revived interest due to recent development in measurement technology [2, 3] and a growing interest in managing chronic health conditions through passive sensors in the home [4].

A hydraulic bed sensor system [5–7] has been recently developed to collect data for a person during sleep. This sensor system is the superposition of the ballistocardiogram (BCG) and respiration signals. The purpose of the sensing

C. Jiao, Bo-Yu Su, Dominic Ho, Princess Lyons and Marjorie Skubic are with the Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO, 65211 USA; e-mail: cjiao@mail.missouri.edu. Alina Zare is with the Department of Electrical and Computer Engineering, University of Florida, Gainesville, FL, 32611 USA; e-mail: azare@ufl.edu. Manuscript received June 11, 2017. This material is based upon work supported by the National Science Foundation under Grant No. IIS-1350078 - CAREER: Supervised Learning for Incomplete and Uncertain Data.

system is to support continuous, non-intrusive monitoring of vital signs of older adults in an unstructured natural living environment. The hydraulic bed sensor is placed beneath the mattress and has shown to be flexible, low-cost and non-intrusive for monitoring an individual’s heart rate. Compared with other methods such as electrocardiography (ECG), BCG does not need electrodes or clips to be affixed to the patient’s body and, thus, is ideal for long term in-home monitoring. However, the lack of saliency and large variability in the BCG signal makes it much more difficult to detect individual heartbeats than with an ECG. In this work, we develop a supervised multiple instance learning approach to learn an individual’s personalized heartbeat signature and, then, use this signature to detect individual heartbeats in the BCG signal and derive the heart rate.

Supervised machine learning methods typically require that every training data point is individually coupled with an accurate training label. However, in many applications obtaining accurate training labels can be time consuming, expensive, and/or infeasible. Furthermore, annotators may be inconsistent during labeling providing inherently imprecise labels. Thus, one often needs to face inaccurately or weakly labeled training data. Also, the majority of supervised machine learning methods assume each data point is a *pure* representative of its associated class. In the case of a BCG system, one avenue for collecting heartbeat groundtruth information is the use of a more intrusive sensor during training (*e.g.*, an ECG or a finger pressure sensor). However, with the use of these, there will be misalignment between the BCG signal and groundtruth signal (ECG or finger sensor). The ECG signal is an electrical signal which starts from SA (sinoatrial) node through the heart to trigger the heartbeat. The BCG signal is a mechanical signal obtained from blood flow, thus, delay between the ECG and BCG is expected. The bed sensor is placed under the mattress which will introduce additional delay. Thus, the exact location in a BCG signal that corresponds to the heartbeat is unknown, making it difficult to align with other sensor signals. Furthermore, several BCG sensors may be used in parallel to ensure full spatial coverage of a subject. However, which of these BCG sensors is able to capture a clear heartbeat signal will depend on the subject’s position relative to the sensors. Thus, not only is there uncertainty in alignment between groundtruth signals and the BCG, but there is also uncertainty as to which BCG signal from the collection of sensors is most informative and accurate. Finally, in addition

to the label uncertainty, the bed sensor data has components associated with a subject's respiration or body movement. Thus, the BCG signals are *mixed* samples containing response from not only the heartbeat but also other sources. Training an accurate classifier or learning a representative target concept from this sort of weakly labeled and mixed data is challenging in practice and requires an approach that can learn from uncertain training labels and mixed data.

Multiple-instance learning (MIL) is a variation on supervised learning for problems with incomplete or uncertain training label information [8–10]. In MIL, training data are grouped into positive and negative “bags.” A bag is defined to be a multi-set of data points where a positive bag contains at least one data point from the target class and negative bags are composed entirely of non-target data. Thus, data point-specific training labels are unavailable. Given training data in this form, the majority of MIL methods either: (1) learn target concepts for describing the target class; or (2) train a classifier that can distinguish between individual target and non-target data points and/or bags. Here, *concepts* refer to generalized class prototypes in the feature space; in the case of heartbeat characterization, a concept is the personalized heartbeat pattern of an individual.

Expanding upon the work in [11, 12], this paper formulates the heartbeat characterization for BCG analysis as a MIL problem. However, in contrast to our prior work (which is based upon the earlier *eFUMI* algorithm), this paper presents the DL-FUMI algorithm. DL-FUMI developed in this paper has the following advantages in comparison to our prior work: (1) DL-FUMI adopts the concept of discriminative dictionary learning to estimate multiple personalized heartbeat signatures to account for the variability in an individual's heartbeat pattern; (2) DL-FUMI introduces the use of a linear mixing model which is more appropriate for BCG signals as opposed to the convex mixing model used in [11]; (3) DL-FUMI adopts the Hybrid Detector [13] for heartbeat detection that can take advantage of the multiple estimated heartbeat signatures; (4) DL-FUMI includes a parameter learning step so the detection parameters (*e.g.*, neighborhood, threshold) are automatically determined from test-on-train procedure; and (5) more comprehensive experimental results (40 subjects) and comparisons are presented as opposed to prior work.

II. RELATED WORK

Over the past several years, there has been a renewed interest in the use of the BCG. This revival has been mainly driven by the advancement in piezoelectric sensors, signal processing [2, 3, 14] and the need for a noninvasive long-term vital signs monitoring. Many approaches have been developed for analyzing the BCG signal for heartbeat detection and heart rate monitoring. [5, 15, 16] rely on clustering a collection of BCG signals (or their associated features) to detect heartbeats. Rosales *et al.* [5] extract three features based on the peak to valley amplitude difference of the BCG signal and then combine the k-means, and Fuzzy C means clustering approaches (CA) to detect the heartbeat. Heise *et al.* [17] use window-peak-to-peak deviation (WPPD) to estimate the

heart rate. WPPD first finds the local maximum at every 0.25 second and then applies a low-pass filter with a cutoff frequency equal to 4Hz to smooth the local maxima. From the smoothed local maxima, WPPD identifies the peak location as the heartbeat position to compute the heart rate. [18] first applies a band-pass filter to remove the respiration information and high frequency noise and then computes the short-term energy (EN), from which the peak positions are considered as the heartbeat locations to obtain the heart rate. There are a number of methods that analyze the BCG signal in the frequency domain to extract the heart rate [7, 19–22]. Among the frequency domain based approaches, the method by Su *et al.* [7, 22] uses the Hilbert transform (HT) to process the data which shows better robustness in the heart rate estimate, and it also has better noise resistance when analyzing the data from the real environment compared to process the data in the time domain. However, the time domain beat-to-beat based approaches are more appealing since they reflect real-time heart rate variability for rhythm pattern analysis [23].

Multiple-instance learning: Since the introduction of the MIL framework [8], many methods have been proposed and developed in the literature. The majority of MIL approaches [8, 24, 25] focus on learning a classification decision boundary to distinguish between positive and negative instances/bags from the ambiguously labeled data. Although these approaches are effective at training classifiers given imprecise labels, they generally do not provide an intuitive description or *representative concept* that characterizes the salient and discriminative features of the target class. The few existing approaches that estimate a target concept include Diverse Density (DD) [9] that estimates a concept by minimizing its distance to at least one instance from each positive bag and maximizing its distance from all instances in negative bags. The Expectation-Maximization (EM) version of diverse density (EMDD) [26] iteratively estimates which instance in each positive bag belong to the target class and then only uses those points from the positive bags to estimate a target concept that maximizes the diverse density. *eFUMI* [10, 27] treats each instance as a convex combination of positive and/or negative concepts and estimates the target and non-target concepts using an EM approach. However, these prototype-based methods only find a single target concept and are, thus, unable to account for any large variation in the target class. To address this, DL-FUMI learns a collection of target dictionary atoms (as well as non-target dictionary atoms) to be able to characterize any variation in the target class.

Sparse Coding and Dictionary Learning: Sparse coding refers to the task of decomposing a signal into a sparse linear combination of dictionary atoms [28–31]. Often, in these applications, pre-defined overcomplete dictionaries (*e.g.*, wavelets, curvelets) or a collection of samples from the data are used as the dictionary set. However, it has been repeatedly shown that, rather than using a pre-defined “off the shelf” dictionary, improved signal reconstruction and greater sparsity in the weights can be achieved by estimating a dictionary for a particular application [32–34]. The unsupervised dictionary learning approaches generally estimate the dictionary and a corresponding set of sparse weights by minimizing the recon-

struction error between the original data and its representation using the estimated dictionary and weights. For example, K-SVD [32] alternates between a sparse coding step using orthogonal matching pursuit (OMP) [35, 36] and a dictionary learning step using singular value decomposition until a preset reconstruction error is reached. However, recent approaches [37–39] demonstrate that classification performance can be improved via supervised dictionary learning.

The proposed DL-FUMI shares the goal of aforementioned approaches to learn an effective set of dictionary atoms for the target class but successfully handles problems in which only imprecise multiple instance learning type labels are available. The model assumed by DL-FUMI, in which each instance is a linear combination of target and non-target dictionary atoms, provides the ability to both discriminate between target and non-target characteristics, account for variation in the target class, and learn a dictionary that provides effective representative prototypes of target and non-target instances. Furthermore, DL-FUMI is unique in that each instance (both positive and negative) shares a common set of non-target/background dictionary atoms. In this way, the estimated target dictionary is discriminative and accounts only for the unique variation in positive class since it is estimated from the residual information not accounted for by the shared non-target/background dictionary.

III. HYDRAULIC BED SENSOR

The hydraulic bed sensor (HBS) developed at the Center for Eldercare and Rehabilitation Technology (CERT) at the University of Missouri is a BCG device providing a low-cost, noninvasive and robust solution for capturing physiological parameters during sleep [5–7]. The HBS was designed to maintain an imperceptible flat profile and to be used beneath a bed mattress. The system is comfortable for subjects lying on the mattress (*i.e.*, noninvasive), easy to install, watertight, and durable.

The HBS is composed of a transducer and a pressure sensor as shown in Fig. 1(a). The transducer was designed to be placed under the subject's upper torso. It is 54.5 cm long, 6 cm wide, and is filled with 0.4 liters of water [5–7]. The integrated silicon pressure sensor (Freescale MPX5010GP) attached to the end of the transducer is used for measuring the vibration of human body arising from each heartbeat. It captures the information of heartbeat together with respiration and motion artifact. The signal from each transducer is then amplified, filtered and sampled at 100 Hz.

In order to ensure enough coverage, four transducers are placed in parallel underneath a mattress as shown in Fig. 1(b). The four transducers are identical and independent, but the data quality collected by those four transducers could vary depending on the sleeping position, type of mattress (*e.g.*, material, thickness, *etc.*) and the physical characteristics of the subject (*e.g.*, age, body mass index (BMI), *etc.*).

IV. DATA DESCRIPTION

Given the hydraulic bed sensor system described above, in this study, we use the dataset collected from 40 subjects at the

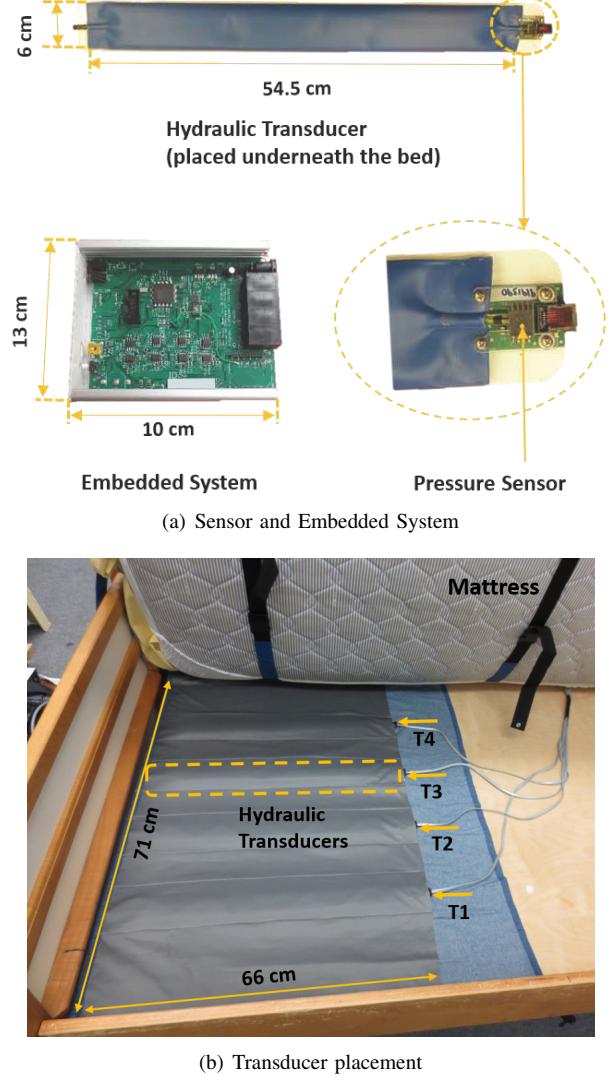


Fig. 1. Hydraulic Bed Sensor System. (a) Hydraulic transducer filled with 0.4 liter of water (top) and embedded system that collects and transmits data from four transducers (bottom). (b) Transducer placement: the transducers are separated 14cm from each other to provide enough coverage for human body.

CERT at the University of Missouri. The data collection from human subjects has been approved by the Institutional Review Board (IRB) at the University of Missouri. To prepare for the data collection, each subject was asked to lie flat on their back for 10 minutes. The gender, age, weight and height of the subjects are listed in Table I. There were 7 females and 33 males. Their ages, weights, heights and BMIs are 18 - 49, 48 - 127 (Kg), 156 - 190 (cm), 18.3 - 37.9, with the average as 29.2, 76.9 (Kg), 175.2 (cm) and 24.9, respectively. The BCG signal was sampled at 100Hz and filtered by a six order Butterworth band-pass filter with 3dB cutoff frequency at 0.4Hz and 10Hz to remove the respiratory component and high frequency noise. For ground-truthing, a piezoelectric pulse sensor (TN1012/ST, ADInstruments) attached to subject's finger was used to record the pulse ejected by a heartbeat.

Feature Extraction from Ballistocardiograms Time Series: In this paper, we used the filtered time domain segments from received BCG signals as training features. Specifically,

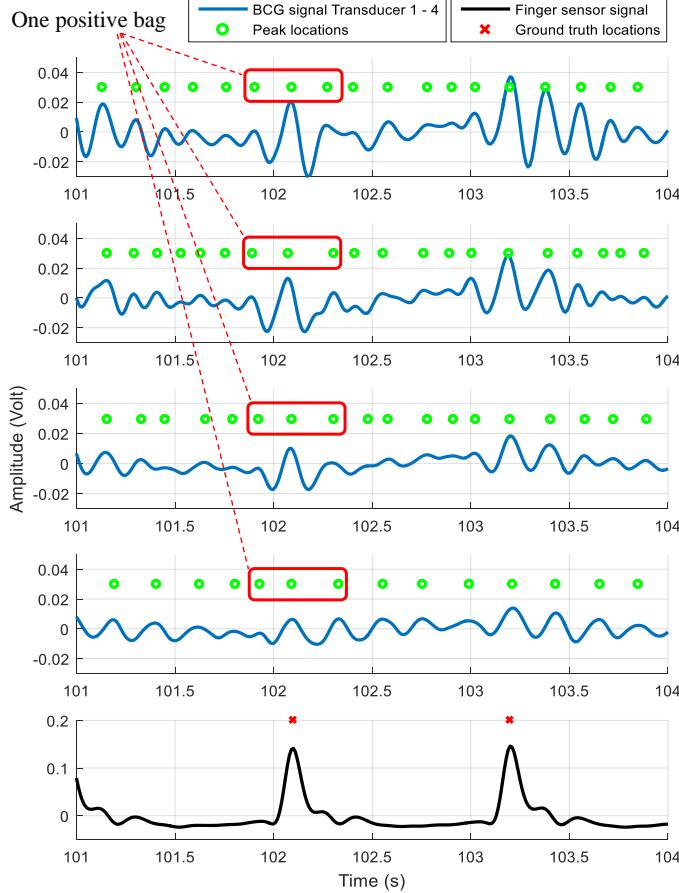


Fig. 2. BCG Signal and Ground Truth Plot. Peaks location were first found as candidate heartbeat (J-peak) locations (green circles), then 3 instances from each transducer closest to the ground-truth location marked by the finger sensor (red cross) were grouped together as a positive bag; instances between two positive bags not placed in any positive bag were grouped together to form one negative bag.

for each subject, we found the signal peaks for the entire BCG signal received by four transducers as candidate J-peak locations (possible heartbeat locations). For each peak, we extract a data segment (simply called instance) that is 91 samples long and centered at the peak (corresponding to 0.91s signal, 45 samples before and after the peak) as the training feature for this peak location. The feature length (0.91s, at 100 Hz sampling rate) was determined empirically and found to be the typical length of a heartbeat pattern. Fig. 2 shows example filtered BCG signals collected by four transducers (blue plots) and the corresponding finger sensor ground truth information (black plots), where the green circles denote every peak location of the filtered BCG signal.

V. MULTIPLE INSTANCE LEARNING PROBLEM IN BALLISTOCARDIOGRAMS

From Fig. 2, it can be seen that near the ground truth locations denoted by the finger sensor, there are prominent peak patterns measured by the four BCG transducers corresponding to heartbeats. However, although all of the sensors are expected to be capturing each corresponding heartbeat

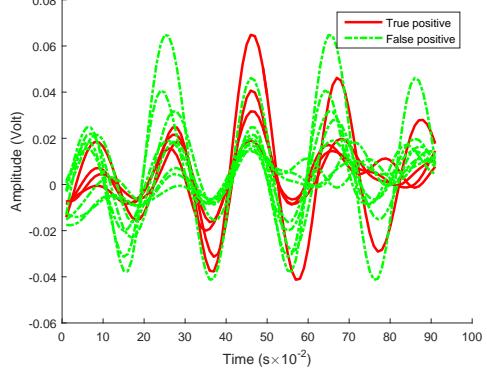


Fig. 3. Plot of one positive bag. A positive bag is also mixed with both true positive instances (heartbeat signal, in red) and false positive instances (non-heartbeat signal, in dotted green), where the true positive heartbeat signals tend to have more prominent J-peaks.

signal simultaneously, there is unavoidable misalignment between the finger sensor and each of the BCG pressure sensors. Furthermore, depending on the location and position of the subject lying on the bed, which of these BCG sensors are able to capture a clear heartbeat signal is difficult to determine. These multiple labeling uncertainties in the training data cast more difficulties to traditional supervised learning methods for heartbeat detection and heart rate estimation from a BCG signal.

In this paper, we continue to investigate the idea of training “bags” to address label uncertainty as well as mis-collection of heartbeat signals in the BCG data. Specifically, each of the extracted sub-signals is treated as individual data points (or “instances”) during training. Each positive labeled training bag (shown as red rectangles in Fig. 2) was formed by grouping the 12 instances across the four transducers (3 instances for each transducer) that are close in time to the ground-truth location marked by the finger sensor (shown as a red cross in Fig. 2). Similarly, one negative bag was formed by grouping instances from four transducers between two positive bags that were not included in any positive bag. Fig. 3 shows an example positive bag. In this figure, we can see that a positively labeled bag contains both true heartbeat patterns (shown in red) and non-heartbeat patterns (shown in dotted green). From Fig. 3 it can also be seen that the assumed heartbeat patterns tend to have more prominent J-peaks. The proposed DL-FUMI algorithm is expected to learn a set of discriminative subject-specific heartbeat concepts from training bags of this type. After learning the heartbeat concept, a signature based detector can then be applied for real-time heartbeat monitoring and heart rate estimation.

VI. DL-FUMI HEARTBEAT CONCEPT LEARNING ALGORITHM

Let $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_N] \in \mathbb{R}^{d \times N}$ be training data where d is the dimensionality of an instance, \mathbf{x}_i , and N is the total number of training instances. The data is grouped into K bags, $\mathbf{B} = \{\mathbf{B}_1, \dots, \mathbf{B}_K\}$, with associated binary bag-level labels, $L = \{L_1, \dots, L_K\}$ where $L_j \in \{0, 1\}$ and $\mathbf{x}_{ji} \in \mathbf{B}_j$ denotes

$$F = \frac{1}{2} \sum_{i=1}^N w_i \left\| \left(\mathbf{x}_i - z_i \sum_{t=1}^T \alpha_{it} \mathbf{d}_t^+ - \sum_{k=1}^M \alpha_{ik} \mathbf{d}_k^- \right) \right\|_2^2 + \lambda \sum_{i=1}^N w_i \left\| \begin{bmatrix} z_i \boldsymbol{\alpha}_i^+ \\ \boldsymbol{\alpha}_i^- \end{bmatrix} \right\|_1 + \sum_{k=1}^M \sum_{t=1}^T \gamma_{kt} \langle \mathbf{d}_k^-, \mathbf{d}_{t \text{old}}^+ \rangle \quad (3)$$

$$E[F] = \sum_{z_i \in \{0, 1\}} P(z_i | \mathbf{x}_i, \boldsymbol{\theta}^{(l-1)}) \left[\frac{1}{2} \sum_{i=1}^N w_i \left\| \mathbf{x}_i - z_i \sum_{t=1}^T \alpha_{it} \mathbf{d}_t^+ - \sum_{k=1}^M \alpha_{ik} \mathbf{d}_k^- \right\|_2^2 + \lambda \sum_{i=1}^N w_i \left\| \begin{bmatrix} z_i \boldsymbol{\alpha}_i^+ \\ \boldsymbol{\alpha}_i^- \end{bmatrix} \right\|_1 \right] + \sum_{k=1}^M \sum_{t=1}^T \gamma_{kt} \langle \mathbf{d}_k^-, \mathbf{d}_{t \text{old}}^+ \rangle \quad (4)$$

the i^{th} instance in bag \mathbf{B}_j . Given training data in this form, DL-FUMI models each instance as a sparse linear combination of target and/or background atoms \mathbf{D} , $\mathbf{x}_i \approx \mathbf{D}\boldsymbol{\alpha}_i$, where $\boldsymbol{\alpha}_i$ is the sparse vector of weights for instance i . Positive bags (*i.e.*, \mathbf{B}_j with $L_j = 1$, denoted as \mathbf{B}_j^+) contain at least one instance composed of some target:

$$\text{if } L_j = 1, \exists \mathbf{x}_i \in \mathbf{B}_j^+ \text{ s.t.} \\ \mathbf{x}_i = \sum_{t=1}^T \alpha_{it} \mathbf{d}_t^+ + \sum_{k=1}^M \alpha_{ik} \mathbf{d}_k^- + \varepsilon_i, \alpha_{it} \neq 0, \quad (1)$$

where ε_i is a noise term. However, the number of instances in a positive bag with a target component is unknown.

If \mathbf{B}_j is a negative bag (*i.e.*, $L_j = 0$, denoted as \mathbf{B}_j^-), then this indicates that \mathbf{B}_j^- does not contain any target:

$$\text{if } L_j = 0, \forall \mathbf{x}_i \in \mathbf{B}_j^-, \mathbf{x}_i = \sum_{k=1}^M \alpha_{ik} \mathbf{d}_k^- + \varepsilon_i \quad (2)$$

Given this problem formulation, the goal of DL-FUMI is to estimate the dictionary¹ $\mathbf{D} = [\mathbf{D}^+ \ \mathbf{D}^-] \in \mathbb{R}^{d \times (T+M)}$, where $\mathbf{D}^+ = [\mathbf{d}_1^+, \dots, \mathbf{d}_T^+]$ are the T target atoms and $\mathbf{D}^- = [\mathbf{d}_1^-, \dots, \mathbf{d}_M^-]$ are the M background atoms. This is accomplished by minimizing (3) which is proportional to the complete negative data log-likelihood, where $\boldsymbol{\alpha}_i^+$ and $\boldsymbol{\alpha}_i^-$ are subsets of $\boldsymbol{\alpha}_i$ corresponding to \mathbf{D}^+ and \mathbf{D}^- , respectively. This model (as in (1) and (2)), differs significantly from most class-specific dictionary learning methods since the background dictionary set is shared among both classes. This provides the ability to estimate a target dictionary \mathbf{D}^+ that is very distinct from non-target data and characterizes the salient and discriminative features of the target class.

The first term in (3) computes the squared residual error between each instance and its estimate using the dictionary. In this term, a set of hidden binary latent variables $\{z_i\}_{i=1}^N$ that indicate whether an instance is or is not a target (*i.e.*, $z_i = 1$ when \mathbf{x}_i contains target) are introduced. For all points in negative bags, $z_i = 0$. For points in positive bags, the value of z_i is unknown. Also, a weight w_i is included where $w_i = 1$ if $\mathbf{x}_i \in \mathbf{B}_j^-$ and $w_i = \psi$ if $\mathbf{x}_i \in \mathbf{B}_j^+$ where ψ is a fixed parameter. This weight helps balance terms when there is a large imbalance between the number of negative and positive instances.

The second term is an l_1 regularization term to promote sparse weights. It also includes the latent variables, z_i , to account for the uncertain presence of target in positive bags.

¹ $[\mathbf{A} \ \mathbf{B}]$ and $\begin{bmatrix} \mathbf{A} \\ \mathbf{B} \end{bmatrix}$ are the concatenation of arrays \mathbf{A} and \mathbf{B} horizontally and vertically, respectively.

The third term is a robust penalty term that promotes discriminative target atoms (and inspired by a term presented in [40]). Instead of using a fixed penalty coefficient, we introduce an adaptive coefficient γ_{kt} defined as:

$$\gamma_{kt} = \Gamma \frac{\langle \mathbf{d}_k^-, \mathbf{d}_t^+ \rangle}{\|\mathbf{d}_k^-\| \|\mathbf{d}_t^+\|} = \Gamma \cos \theta_{kt}, \quad (5)$$

where θ_{kt} is the vector angle between the k^{th} background atom and the t^{th} target atom. Since $\text{sign}(\gamma_{kt}) = \text{sign}(\langle \mathbf{d}_k^-, \mathbf{d}_t^+ \rangle)$, this discriminative term is always positive and will add large penalty when \mathbf{d}_k^- and \mathbf{d}_t^+ have similar shape. Thus, this term encourages a discriminative dictionary by promoting background atoms that are orthogonal to target atoms. In implementation, γ_{kt} is updated once per iteration using $\mathbf{d}_k^{\text{old}}$ and $\mathbf{d}_{t \text{old}}^+$ which are the dictionary values from the previous iteration.

Algorithm 1 DL-FUMI EM algorithm

```

1: Initialize  $\boldsymbol{\theta}^0 = \left\{ \mathbf{D}, \{\boldsymbol{\alpha}_i\}_{i=1}^N \right\}$ ,  $l = 1$ 
2: repeat
3:   E-step: Compute  $P(z_i | \mathbf{x}_i, \boldsymbol{\theta}^{(l-1)})$ 
4:   M-step:
5:   Update  $\mathbf{d}_t^+$  using (11),  $\mathbf{d}_t^+ \leftarrow \frac{1}{\|\mathbf{d}_t^+\|_2} \mathbf{d}_t^+, t = 1, \dots, T$ 
6:   Update  $\mathbf{d}_k^-$  using (10),  $\mathbf{d}_k^- \leftarrow \frac{1}{\|\mathbf{d}_k^-\|_2} \mathbf{d}_k^-, k = 1, \dots, M$ 
7:   for  $q \leftarrow 1$  to  $\text{iter}$  do
8:     Update  $\{\boldsymbol{\alpha}_i\}_{i=1}^{N^+}$  for  $\mathbf{x}_i \in \mathbf{B}_j^+$  using (13), (14)
9:     Update  $\{\boldsymbol{\alpha}_i\}_{i=1}^{N^-}$  for  $\mathbf{x}_i \in \mathbf{B}_j^-$  using (15)
10:  end for
11:   $l \leftarrow l + 1$ 
12: until Stopping criterion met
13: return  $\mathbf{D}, \{\boldsymbol{\alpha}_i\}_{i=1}^N$ 

```

VII. DL-FUMI OPTIMIZATION

Expectation-Maximization is used to optimize (3) and estimate \mathbf{D} . During optimization, the fact that many of the binary latent variables $\{z_i\}_{i=1}^N$ are unknown is addressed by taking the expected value of the log likelihood with respect to z_i as shown in (4). In (4), $\boldsymbol{\theta}^l = \left\{ \mathbf{D}, \{\boldsymbol{\alpha}_i\}_{i=1}^N \right\}$ is the set of parameters estimated at iteration l and $P(z_i | \mathbf{x}_i, \boldsymbol{\theta}^{(l-1)})$ is the probability that each instance is or is not a true target

instance. During the E-step of each iteration, $P(z_i|\mathbf{x}_i, \boldsymbol{\theta}^{(l-1)})$ is computed as:

$$P(z_i|\mathbf{x}_i, \boldsymbol{\theta}^{(l-1)}) = \begin{cases} e^{-\beta \|\mathbf{x}_i - \sum_{k=1}^M \alpha_{ik} \mathbf{d}_k^-\|_2^2} & \text{if } z_i = 0, L_j = 1 \\ 1 - e^{-\beta \|\mathbf{x}_i - \sum_{k=1}^M \alpha_{ik} \mathbf{d}_k^-\|_2^2} & \text{if } z_i = 1, L_j = 1 \\ 0 & \text{if } z_i = 1, L_j = 0 \\ 1 & \text{if } z_i = 0, L_j = 0 \end{cases} \quad (6)$$

where β is a fixed scaling parameter. If \mathbf{x}_i is a non-target instance, then it should be characterized by the background atoms well, thus $P(z_i = 0|\mathbf{x}_i, \boldsymbol{\theta}^{(l-1)}) \approx 1$. Otherwise, if \mathbf{x}_i is a true target instance, it will not be characterized well using only the background atoms and $P(z_i = 1|\mathbf{x}_i, \boldsymbol{\theta}^{(l-1)}) \approx 1$. Please note, z_i is unknown only for the positive bags; for the negative bags, z_i is fixed to 0. This constitutes the *E-step* of the EM algorithm.

The *M-step* is performed by iteratively optimizing (4) for each of the desired parameters. The dictionary \mathbf{D} is updated atom-by-atom using a block coordinate descent scheme [41, 42]. The sparse weights, $\{\alpha_i\}_{i=1}^N$, are updated using an iterative shrinkage-thresholding algorithm [43, 44]. For readability, the derivation of update equations are described in Sec. XI. The method is summarized in Alg. 1.

VIII. HEARTBEAT DETECTION USING ESTIMATED DICTIONARY

Once the heartbeat concept set \mathbf{D} has been estimated using DL-FUMI, heartbeat detection on test data can be performed using the hybrid structured detector (HSD) [13, 45]. The hypotheses used for HSD are:

$$\begin{aligned} \mathbf{H}_0 : \mathbf{x} &\sim \mathcal{N}(\mathbf{D}^- \boldsymbol{\alpha}^-, \sigma_0^2 \Sigma) \\ \mathbf{H}_1 : \mathbf{x} &\sim \mathcal{N}(\mathbf{D}\boldsymbol{\alpha}, \sigma_1^2 \Sigma) \end{aligned} \quad (7)$$

where Σ is the training background covariance and $\boldsymbol{\alpha}$ and $\boldsymbol{\alpha}^-$ are the corresponding sparse codes of test data \mathbf{x} given full concept set \mathbf{D} and background concept set \mathbf{D}^- , respectively.

The confidence that the i^{th} signal is target can be computed using the generalized likelihood ratio test (GLRT) of hypotheses (7),

$$\Lambda_{HSD}(\mathbf{x}_i, \mathbf{D}) = \frac{(\mathbf{x}_i - \mathbf{D}^- \boldsymbol{\alpha}_i^-)^T \Sigma^{-1} (\mathbf{x}_i - \mathbf{D}^- \boldsymbol{\alpha}_i^-)}{(\mathbf{x}_i - \mathbf{D}\boldsymbol{\alpha}_i)^T \Sigma^{-1} (\mathbf{x}_i - \mathbf{D}\boldsymbol{\alpha}_i)}, \quad (8)$$

Eq. (8) indicates that if the reconstruction error of signal \mathbf{x}_i using only the background concept (the numerator) is much larger than that using the entire concept, then the target concept is needed in the reconstruction of signal \mathbf{x}_i and \mathbf{x}_i is likely to be a target. The HSD explicitly models the mixture in BCG signals, utilizes the multiple target concepts and provides a sub-signal detection alternative.

IX. EXPERIMENTAL RESULTS

For each subject, the 10 minutes BCG signal was split into 5 minutes for training and 5 minutes for testing. For DL-FUMI, the parameters used for all the subjects are $T = 3$, $M = 3$, $\lambda = 5 \times 10^{-3}$, $\Gamma = 5 \times 10^{-3}$ and $\beta = 90$. Fig. 4 shows estimated heartbeat concepts for subject No. 10 as an

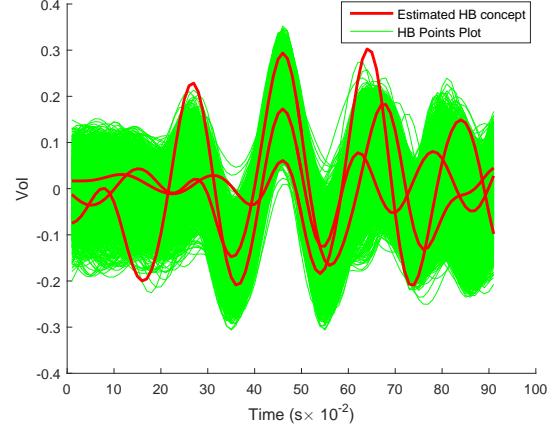


Fig. 4. Estimated Heartbeat Concept by DL-FUMI. The multiple heartbeat concepts estimated by DL-FUMI better account for the variability in heartbeat signals.

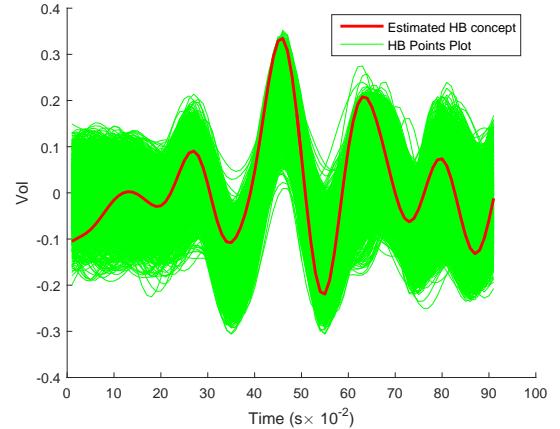


Fig. 5. Estimated Heartbeat Concept by EMDD. A single heartbeat concept estimated by EMDD fits the plots of the heartbeat signal well, but fails to maintain the variability in the heartbeat prototype.

example, where we can see the heartbeat concept estimated by DL-FUMI maintains prominent J-peaks. For comparison, we applied EMDD [26] which is a widely used multiple instance concept learning algorithm to the same data. From Fig. 4 and 5 we can clearly see that although both the heartbeat concepts estimated by DL-FUMI and EMDD have prominent J-peaks, DL-FUMI is able to learn multiple concepts to account for the variability in heartbeat prototype during sleeping, which helps improve performance in heartbeat detection and rate estimation.

After learning heartbeat concepts, heartbeat detection on test data can be carried out. In the results shown in this paper, the HSD was applied to the test data to get a confidence value for each data point to be a true heartbeat signal. Since EMDD only learns a single target concept, we applied K-SVD [32] which is a widely used unsupervised dictionary learning algorithm to the negative labeled training bags to get a set of non-target concepts for EMDD heartbeat detection using HSD. Fig. 6 shows the excerpt of HSD 4 - channel confidence value

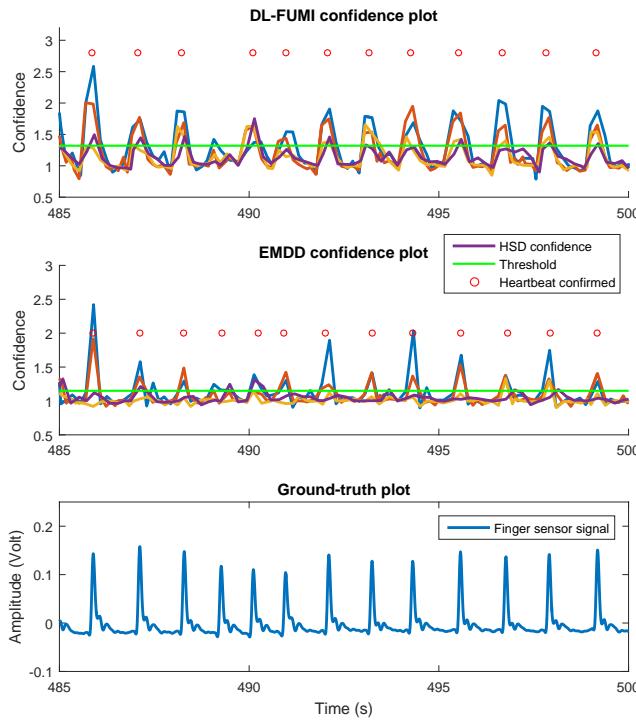


Fig. 6. Confidence Value and Confirmed Heartbeat. The confidence estimated by DL-FUMI (upper) is more consistent than that estimated by EMDD (middle), thus provides more robust detection performance.

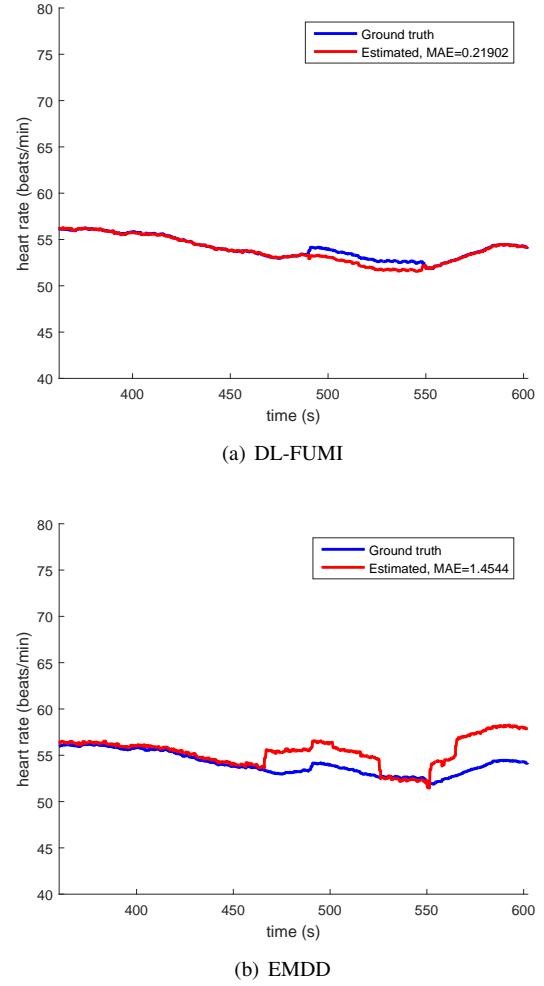


Fig. 8. Heart Rate Estimation. (a) Heart rate estimated by DL-FUMI, where the drop around 500s resulted from one missed detection at 489s. (b) Heart rate estimated by EMDD, several false alarms account for the raise in heart rate estimation.

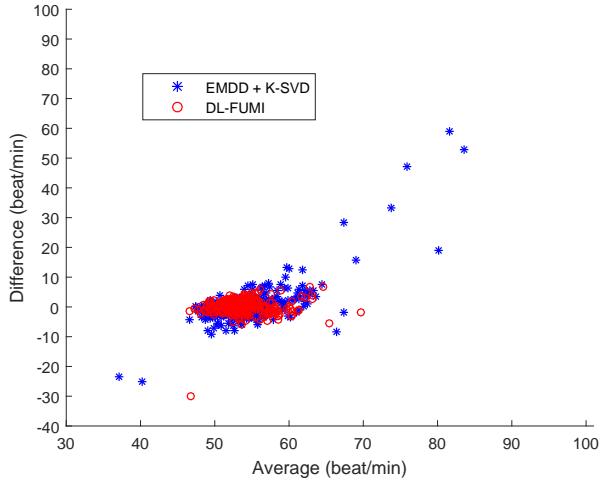


Fig. 7. The Bland Altman plot comparison of DL-FUMI and EMDD for Subject No. 10. The outlier by DL-FUMI (red circle in the bottom left) corresponding to the missed detection appeared around 489s shown in Fig. 6, but the plot of DL-FUMI is overall more compact and has less false alarms as well as missed detections than EMDD.

(485s - 500s) for subject No. 10 estimated by the heartbeat concepts of DL-FUMI (shown in Fig. 4) and EMDD (shown in Fig. 5), respectively. In this procedure, a heartbeat (J-peak) is confirmed through a voting procedure requiring at least two confidence values within a neighborhood (25 samples) that are greater than a threshold (1.32) across all four transducers. The neighborhood and threshold values are determined via cross-validation on training data. From Fig. 6 we can see that the confidence peaks estimated by DL-FUMI match the peaks of the finger sensor signal very well. A missed detection was found at about 489s. For comparison, the confidence estimated by EMDD is not as prominent as DL-FUMI and there are several false alarms found. Fig. 7 shows the beat to beat Bland Altman plot of DL-FUMI and EMDD, where we can see the plot of DL-FUMI is more compact than that of EMDD. For DL-FUMI, there is one outlier in the bottom left corresponding to the missed detection appeared around 489s in Fig. 6.

For heart rate estimation, the average of beat-to-beat heart rates over 1 minute is computed using a sliding window. Fig. 8(a) and Fig. 8(b) show the estimated heart rate on testing

TABLE I
PERFORMANCE OF DL-FUMI AND COMPARISONS ACROSS THE 40 SUBJECTS, BOLD FOR THE BEST, UNDERLINE FOR THE SECOND BEST

Subject	Age, sex (F/M), BMI weight (kg), height (cm)	DL-FUMI	Mean Error (beat/min)	EMDD+K-SVD	EN	WPPD	CA	HT
1	43, M, 23.3, 74.8, 179	0.19 $\pm 7.95 \times 10^{-3}$	0.96 ± 0.82	<u>0.22</u>	1.26	1.76 ± 0.09	0.79	
2	21, M, 23.1, 74, 179	<u>0.85</u> ± 0.05	4.99 ± 1.48	0.58	1.67	1.68 ± 0.14	0.94	
3	33, M, 27.2, 75, 166	<u>0.22</u> $\pm 0.20 \times 10^{-3}$	0.71 ± 0.94	0.17	1.02	1.05 ± 0.03	0.77	
4	23, F, 21.1, 52, 157	<u>0.96</u> ± 0.57	3.55 ± 1.76	14.69	1.28	1.24 ± 0.03	0.85	
5	28, M, 37.9, 127, 183	0.83 ± 0.33	4.83 ± 0.89	2.84	1.25	1.10 ± 0.07	<u>1.00</u>	
6	21, M, 22.2, 75, 184	0.45 ± 0.07	2.61 ± 1.61	20.20	1.26	1.10 ± 0.03	<u>0.66</u>	
7	32, M, 30.1, 82, 165	<u>1.28</u> ± 0.19	4.45 ± 2.23	1.83	1.47	1.98 ± 0.09	0.79	
8	32, M, 27.2, 92, 180	<u>1.46</u> ± 0.89	3.47 ± 2.72	1.65	<u>1.04</u>	1.06 ± 0.04	0.48	
9	29, M, 19.7, 68, 186	2.47 ± 0.60	1.34 ± 0.46	0.77	2.12	2.03 ± 0.04	<u>1.26</u>	
10	24, F, 22.8, 62, 165	0.24 $\pm 1.42 \times 10^{-3}$	3.00 ± 1.72	<u>0.53</u>	1.07	1.39 ± 0.03	0.64	
11	31, M, 22.5, 68, 174	<u>1.09</u> $\pm 9.54 \times 10^{-3}$	2.16 ± 1.20	0.61	1.66	1.42 ± 0.04	<u>1.05</u>	
12	27, M, 24.2, 70, 170	<u>0.30</u> ± 0.10	0.66 ± 0.55	0.18	1.28	1.10 ± 0.04	0.94	
13	29, M, 23.6, 79, 183	<u>0.27</u> ± 0.37	1.14 ± 0.28	0.09	1.03	1.12 ± 0.04	0.53	
14	18, M, 25.6, 83, 180	1.64 ± 0.37	2.88 ± 1.52	6.52	<u>1.63</u>	2.27 ± 0.04	1.62	
15	39, M, 25.0, 74, 172	0.14 $\pm 9.91 \times 10^{-3}$	0.45 ± 0.33	<u>0.27</u>	1.45	1.29 ± 0.04	0.89	
16	26, M, 21.2, 65, 175	<u>0.22</u> ± 0.02	2.09 ± 1.14	6.12	1.25	1.65 ± 0.05	<u>0.50</u>	
17	31, M, 28.9, 100, 186	<u>0.10</u> $\pm 3.21 \times 10^{-3}$	1.66 ± 1.12	0.49	1.04	1.87 ± 0.19	<u>0.46</u>	
18	27, M, 22.3, 70, 177	<u>0.45</u> ± 0.01	1.64 ± 1.72	0.20	1.21	1.07 ± 0.04	0.63	
19	30, M, 25.1, 76, 174	0.15 $\pm 4.31 \times 10^{-3}$	1.18 ± 0.95	<u>0.62</u>	1.45	1.48 ± 0.05	0.71	
20	23, M, 24.7, 73, 172	0.56 ± 0.03	4.43 ± 3.93	<u>0.85</u>	2.35	2.12 ± 0.09	1.42	
21	30, F, 19.7, 57, 170	1.18 ± 0.37	4.08 ± 1.55	4.29	<u>0.90</u>	1.07 ± 0.14	0.23	
22	27, M, 27.5, 86, 177	<u>0.10</u> $\pm 6.17 \times 10^{-3}$	2.12 ± 0.34	<u>0.38</u>	1.65	1.51 ± 0.07	0.44	
23	24, F, 20.8, 60, 170	0.81 $\pm 4.95 \times 10^{-3}$	1.76 ± 0.72	7.65	1.44	1.61 ± 0.13	<u>1.12</u>	
24	25, M, 27.1, 86, 178	<u>0.32</u> ± 0.03	2.21 ± 0.27	7.21	2.74	1.86 ± 0.08	<u>1.14</u>	
25	49, M, 23.0, 83, 190	<u>0.20</u> ± 0.04	0.73 ± 0.75	0.07	1.35	1.41 ± 0.04	0.80	
26	22, M, 26.0, 92, 188	<u>0.25</u> ± 0.02	2.08 ± 1.82	<u>0.49</u>	2.13	2.30 ± 0.26	1.81	
27	33, M, 29.2, 82, 167.6	<u>0.56</u> ± 0.27	4.57 ± 3.26	<u>0.89</u>	1.45	1.49 ± 0.04	1.00	
28	28, M, 21.2, 73, 185.4	0.62 ± 0.22	1.64 ± 1.10	3.21	1.72	1.96 ± 0.08	<u>1.38</u>	
29	34, M, 25.6, 84, 181	0.47 ± 0.08	1.15 ± 0.54	0.09	0.89	0.87 ± 0.06	<u>0.34</u>	
30	34, M, 24.4, 64, 162	<u>0.15</u> ± 0.03	3.67 ± 3.14	<u>0.17</u>	1.06	1.28 ± 0.16	0.34	
31	32, F, 28.2, 77, 165.1	<u>0.99</u> ± 0.12	3.75 ± 1.31	1.47	<u>1.30</u>	2.36 ± 0.13	1.71	
32	22, F, 21.0, 51, 156	<u>1.26</u> ± 0.57	2.53 ± 1.38	10.23	1.59	2.36 ± 0.04	1.03	
33	27, M, 23.8, 77, 180	<u>0.71</u> ± 0.07	0.73 ± 0.23	8.43	1.82	0.98 ± 0.01	0.54	
34	26, M, 23.7, 83, 187	<u>0.61</u> ± 0.13	2.72 ± 2.27	4.65	1.49	1.55 ± 0.08	<u>0.64</u>	
35	28, F, 18.3, 48, 162	<u>0.15</u> ± 0.05	3.98 ± 1.63	<u>0.43</u>	1.34	1.15 ± 0.03	0.76	
36	38, M, 36.9, 120, 180.3	1.44 ± 0.31	<u>1.00</u> ± 0.52	1.68	1.15	1.19 ± 0.16	0.47	
37	37, M, 22.1, 68, 175.3	<u>0.23</u> $\pm 9.86 \times 10^{-3}$	2.42 ± 1.54	0.85	0.74	0.89 ± 0.05	0.40	
38	29, M, 25.8, 79, 175	0.46 ± 0.16	1.37 ± 1.24	6.53	1.13	0.87 ± 0.07	<u>0.47</u>	
39	23, M, 23.0, 68, 172	<u>0.71</u> ± 0.38	4.11 ± 3.43	0.52	1.80	1.84 ± 0.08	1.00	
40	32, M, 30.9, 99, 179	<u>0.37</u> ± 0.04	2.57 ± 2.10	1.94	1.24	1.44 ± 0.08	<u>0.55</u>	
Total average		0.64	2.43	3.02	1.42	1.49	<u>0.83</u>	

TABLE II
THE CORRELATION COEFFICIENTS BETWEEN PERFORMANCE AND AGE,
WEIGHT, HEIGHT, BMI AND GROUNDTRUTH.

	DL-FUMI	EMDD	WPPD	CA	EN	HT
Age	-0.12	-0.37	-0.28	-0.31	-0.46	-0.35
Weight	0.04	-0.06	-0.12	-0.04	-0.21	-0.01
Height	0.00	-0.35	-0.07	-0.01	-0.11	0.01
BMI	0.05	0.11	-0.08	-0.04	-0.21	-0.01
average GT	0.24	-0.01	0.03	-0.05	0.46	-0.11

data for subject No. 10 by DL-FUMI and EMDD, respectively. From Fig. 8(a) we can see that the drop in estimated heart rate around 489s comes from the missed detection of heartbeat in Fig 6. As a comparison, the rise in estimated heart rate in Fig. 8(b) comes from the several false alarms estimated by EMDD.

A comprehensive study and comparison with more state-of-the-art BCG heart rate monitoring algorithms were conducted on the 40 subjects. The four algorithms used for the comparison are window-peak-to-peak (WPPD) [17], clustering approach (CA) [5], energy (EN) [18], and Hilbert transform (HT) [7, 22]. The detail of results are listed in Table I.

The mean error shown in Table I is the mean absolute error (MAE) between the estimation and groundtruthing at every 15s interval, where the bold numbers highlight the best and underlined numbers highlight the second best performance for each subject. Over 40 subjects, the mean error of proposed algorithm is 0.64 (beat/min), which is the best over the comparison algorithms. The mean error of EMDD+K-SVD, EN, WPPD, CA and HT are 2.43, 3.02, 1.42, 1.49 and 0.83 (beat/min), respectively. For EN, there are 3 subjects that have aberrant estimation (with error greater than 10) and make the average result worse. If we ignore these three outliers, the mean error of EN is 2.04 (beat/min).

To examine if the performance of proposed algorithm is related to subject's age, weight, height, BMI, and magnitude of subject's heart rate, we compute the correlation coefficient between the results in Table I and subjects' age, weight, height, BMI and the average of groundtruth heart rate over testing data (denoted as average GT). The sample Pearson correlation coefficient r defined in Eq. (9) measures the linear correlation between two variables x and y , where i is the index of the vari-

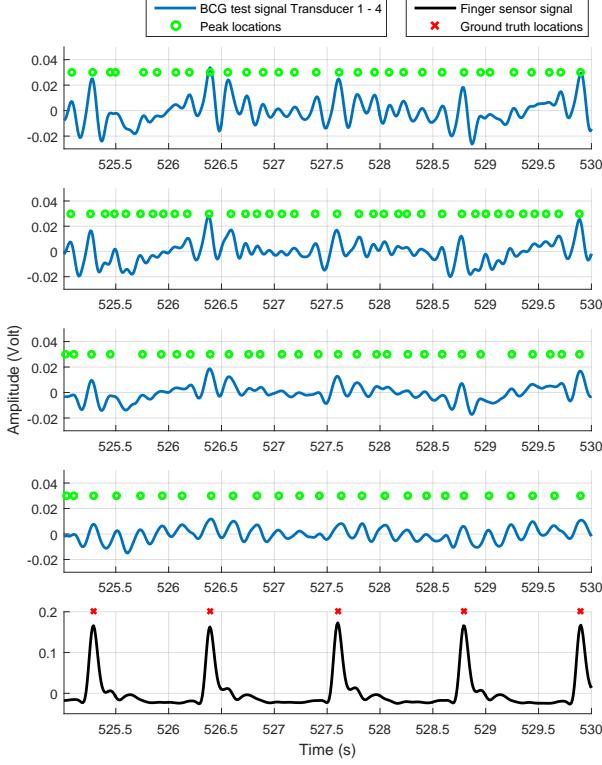


Fig. 9. The 5s plot of subject No. 10.

ables and \bar{x} and \bar{y} are the mean of two variables, respectively. Table II shows the corresponding correlation coefficient of the proposed algorithm and the other five comparison algorithms. For weight, height, and BMI, all six algorithms show no relation. For age, all algorithms consistently show negative relationship with error, this is mainly due to the increase in heartbeat variations with the decrease in age. However, compared with the other five algorithms, the proposed DL-FUMI is more robust to age since DL-FUMI is able to learn multiple concepts to account for heartbeat variability. Based on the results, we can tell that the performance is not related to age, weight, height, and BMI, which validates the robustness of the proposed algorithm.

$$r = \frac{\sum_i (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_i (x_i - \bar{x})^2} \sqrt{\sum_i (y_i - \bar{y})^2}}, \quad (9)$$

The HT algorithm obtains the second best performance and the average estimation error is close to proposed algorithm. However, HT computes the heart rate from the frequency domain and only estimates an overall heart rate from a segment of signal. In contrast, the proposed algorithm does not have this problem, and it can also obtain the beat-to-beat heart rate information. To further evaluate the two algorithms, we pick just 5s signal segment from subject No. 10 shown in Fig. 9. The beat-to-beat detection by DL-FUMI is shown in Fig. 10(a) and the heart rates estimated by both DL-FUMI and HT are shown in Fig 10(b). From Fig. 10(b) we can see that if the subject has high heart rate variation, the HT cannot track the variation well.

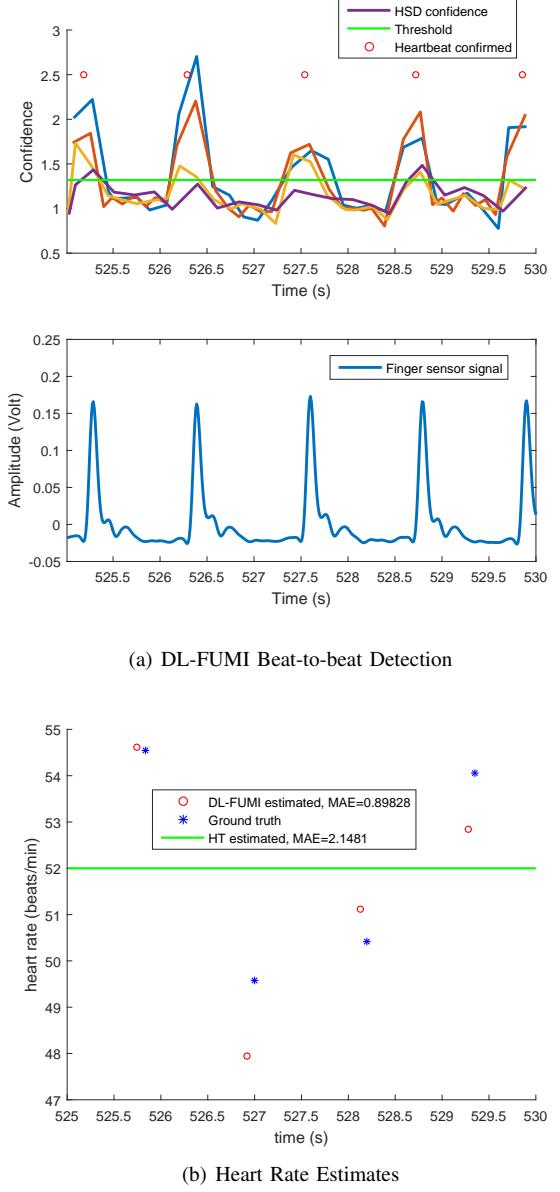


Fig. 10. Heart Rate Estimation from 5s signal. (a) Beat-to-beat Detection by DL-FUMI (upper) and Finer Sensor Plot (ground truth, middle). (b) Beat-to-beat Heart Rate estimated by DL-FUMI (red circles) and finger sensor (blue stars), where the comparison HT algorithm only estimates a constant heart rate (green line) from a period of signal.

X. CONCLUSION

In this paper, heartbeat characterization was formulated as a MIL problem and addressed using the DL-FUMI algorithm. Experimental results show that with enough training data, the algorithm is able to perform well in heartbeat characterization and heart rate estimation with average error less than 1 beat/min. We plan to apply DL-FUMI to the long-term monitored data acquired at TigerPlace, an active aging-in-place retirement community developed by the MU Sinclair School of Nursing and CERT at the University of Missouri [46], and study about how the position, posture, and body movement affect the performance of the proposed algorithm.

$$\mathbf{d}_k^- = \left\{ \sum_{i=1}^{N^+} \left[P(z_i = 1) \psi \alpha_{ik} (\mathbf{x}_i - \sum_{t=1}^T \alpha_{it} \mathbf{d}_t^+ - \sum_{l=1, l \neq k}^M \alpha_{il} \mathbf{d}_l^-) + P(z_i = 0) \psi \alpha_{ik} (\mathbf{x}_i - \sum_{l=1, l \neq k}^M \alpha_{il} \mathbf{d}_l^-) \right] + \sum_{i=1}^{N^-} \left[\alpha_{ik} (\mathbf{x}_i - \sum_{l=1, l \neq k}^M \alpha_{il} \mathbf{d}_l^-) \right] \right. \\ \left. - \Gamma \sum_{t=1}^T \cos \theta_{kt} \mathbf{d}_{t,old}^+ \right\} \left\{ \sum_{i=1}^{N^+} \psi \alpha_{ik}^2 + \sum_{i=1}^{N^-} \alpha_{ik}^2 \right\}^{-1} \quad (10)$$

XI. DERIVATION OF DL-FUMI UPDATE EQUATIONS

This section provides a derivation of DL-FUMI update equations. When updating the dictionary \mathbf{D} , the sparse weights $\{\alpha_i\}_{i=1}^N$ are held fixed. To update one of the atoms in \mathbf{D} , (4) is minimized with respect to the corresponding atom while keeping all other atoms constant. The resulting update equations for \mathbf{d}_t^+ and \mathbf{d}_k^- are shown in (11) and (10).

$$\mathbf{d}_t^+ = \frac{\sum_{i=1}^{N^+} [P(z_i = 1) \alpha_{it} (\mathbf{x}_i - \sum_{l=1, l \neq t}^T \alpha_{il} \mathbf{d}_l^+ - \sum_{k=1}^M \alpha_{ik} \mathbf{d}_k^-)]}{\sum_{i=1}^{N^+} [P(z_i = 1) \alpha_{it}^2]} \quad (11)$$

Note, $P(z_i | \mathbf{x}_i, \boldsymbol{\theta}^{(t-1)})$ is denoted as $P(z_i)$ for simplicity.

When updating the sparse weights, $\{\alpha_i\}_{i=1}^N$, it should be noted that the sparse weight vector α_i for instance \mathbf{x}_i is not dependent on any other instances.

The gradient with respect to α_i without considering the l_1 penalty term is:

$$\frac{\partial F^+}{\partial \alpha_i} = - [P(z_i = 1) \mathbf{D}^+ \mathbf{D}^-]^T \mathbf{x}_i + (P(z_i = 1) \mathbf{D}^T \mathbf{D}^- + P(z_i = 0) [\mathbf{0}_{d \times T} \mathbf{D}^-]^T [\mathbf{0}_{d \times T} \mathbf{D}^-]) \alpha_i. \quad (12)$$

Then α_i at l^{th} iteration can be updated using gradient descent,

$$\alpha_i^l = \alpha_i^{l-1} - \eta_i \frac{\partial F^+}{\partial \alpha_i}, \quad (13)$$

followed by a soft-thresholding:

$$\begin{cases} \alpha_i^{l+} = S_{\lambda P(z_i=1)}(\alpha_i^{l+}), \\ \alpha_i^{l-} = S_\lambda(\alpha_i^{l-}) \end{cases}, \quad (14)$$

s.t. $S_\lambda(\mathbf{x}[i]) = sign(\mathbf{x}[i]) \max(|\mathbf{x}[i]| - \lambda, 0)$, $i = 1, \dots, d$.

Following a similar proof to that in [47], when $\eta_i \in (0, (\lambda_{\max}(P(z_i = 0) [\mathbf{0}_{d \times T} \mathbf{D}^-]^T [\mathbf{0}_{d \times T} \mathbf{D}^-] + P(z_i = 1) \mathbf{D}^T \mathbf{D}^-))^{-1})$, the update of α_i using a gradient descent method with step length η_i monotonically decreases the value of the objective function, where $\lambda_{\max}(\mathbf{A})$ denotes the maximum eigenvalue of \mathbf{A} . For simplicity, η was set as $\eta = \frac{1}{\lambda_{\max}(\mathbf{D}^T \mathbf{D})}$ for all α_i , $\mathbf{x}_i \in \mathbf{B}_j^+$.

A similar update can be used for points from negative bags. The resulting update equation for negative points is:

$$\alpha_i^l = S_\lambda \left(\alpha_i^{l-1} + \frac{1}{\lambda_{\max}(\mathbf{D}^T \mathbf{D}^-)} (\mathbf{D}^{-T} (\mathbf{x}_i - \mathbf{D}^- \alpha_i^{l-1})) \right) \quad (15)$$

The sparse weights corresponding to target dictionary atoms are set to 0 for all points in negative bags.

REFERENCES

- [1] I. Starr, A. Rawson, H. Schroeder, and N. Joseph, "Studies on the estimation of cardiac output in man, and of abnormalities in cardiac function, from the heart's recoil and the blood's impacts; the ballistocardiogram," *Amer. J. of Physiology—Legacy Content*, vol. 127, no. 1, pp. 1–28, 1939.
- [2] E. Pinheiro, O. Postolache, and P. Girão, "Theory and developments in an unobtrusive cardiovascular system representation: ballistocardiography," *The Open Biomed. Eng. J.*, vol. 4, p. 201, 2010.
- [3] O. T. Inan, P.-F. Migeotte, K.-S. Park *et al.*, "Ballistocardiography and seismocardiography: A review of recent advances," *IEEE J. Biomed. and Health Inform.*, vol. 19, no. 4, pp. 1414–1427, 2015.
- [4] M. Skubic, R. D. Guevara, and M. Rantz, "Automated health alerts using in-home sensor data for embedded health assessment," *IEEE journal of translational engineering in health and medicine*, vol. 3, pp. 1–11, 2015.
- [5] L. Rosales, M. Skubic, D. Heise, M. J. Devaney, and M. Schaumburg, "Heartbeat detection from a hydraulic bed sensor using a clustering approach," in *Proc. 34th Annu. Int. IEEE Eng. Med. Biol. Soc. Conf.*, 2012, pp. 2383–2387.
- [6] D. Heise, L. Rosales, M. Sheahen, B.-Y. Su, and M. Skubic, "Non-invasive measurement of heartbeat with a hydraulic bed sensor progress, challenges, and opportunities," in *2013 IEEE Int. Instrum. and Meas. Technol. Conf.*, 2013, pp. 397–402.
- [7] L. Rosales, B. Y. Su, M. Skubic, and K. C. Ho, "Heart rate monitoring using hydraulic bedsensor ballistocardiogram," *J. Ambient Intell. Smart Environ.*, vol. 9, no. 2, pp. 193–207, Feb. 2017.
- [8] T. G. Dietterich, R. H. Lathrop, and T. Lozano-Perez, "Solving the multiple-instance problem with axis-parallel rectangles," *Artificial Intell.*, vol. 89, no. 1-2, pp. 31–17, 1997.
- [9] O. Maron and T. Lozano-Pérez, "A framework for multiple-instance learning," in *Advances in Neural Inf. Process. Syst. (NIPS)*, 1998, pp. 570–576.
- [10] C. Jiao and A. Zare, "Functions of multiple instances for learning target signatures," *IEEE Trans. Geosci. and Remote Sens.*, vol. 53, no. 8, pp. 4670 – 4686, 2015.
- [11] C. Jiao, P. Lyons, A. Zare, L. Rosales, and M. Skubic, "Heart beat characterization from ballistocardiogram signals using extended functions of multiple instances," in *Proc. 38th Annu. Int. IEEE Eng. Med. Biol. Soc. Conf.*, 2016, pp. 756–760.
- [12] C. Jiao and A. Zare, "Multiple instance dictionary learn-

- ing using functions of multiple instances,” in *Int. Conf. Pattern Recognition (ICPR)*, 2016.
- [13] J. Broadwater and R. Chellappa, “Hybrid detectors for subpixel targets,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 11, pp. 1891–1903, Nov. 2007.
- [14] O. A. Postolache, P. M. S. Girao, J. Mendes, E. C. Pinheiro, and G. Postolache, “Physiological parameters measurement based on wheelchair embedded sensors and advanced signal processing,” *IEEE Trans. Instrum. Meas.*, vol. 59, no. 10, pp. 2564–2574, 2010.
- [15] C. Bruser, K. Stadlthanner, S. de Waele, and S. Leonhardt, “Adaptive beat-to-beat heart rate estimation in ballistocardiograms,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 15, no. 5, pp. 778–786, Sept. 2011.
- [16] J. Paalasmaa, H. Toivonen, and M. Partinen, “Adaptive heartbeat modeling for beat-to-beat heart rate measurement in ballistocardiograms,” *IEEE J. Biomed. Health Inform.*, vol. 19, no. 6, pp. 1945–1952, Nov. 2015.
- [17] D. Heise and M. Skubic, “Monitoring pulse and respiration with a non-invasive hydraulic bed sensor,” in *Proc. 32nd Annu. Int. IEEE Eng. Med. Biol. Soc. Conf.*, 2010, pp. 2119–2123.
- [18] K. Lydon, B. Y. Su, L. Rosales, M. Enayati, K. C. Ho, M. Rantz, and M. Skubic, “Robust heartbeat detection from in-home ballistocardiogram signals of older adults using a bed sensor,” in *Proc. 37th Annu. Int. IEEE Eng. Med. Biol. Soc. Conf.*, 2015, pp. 7175–7179.
- [19] J. H. Shin and K. S. Park, “HRV analysis and blood pressure monitoring on weighing scale using bkg,” in *Proc. 34th Annu. Int. IEEE Eng. Med. Biol. Soc. Conf. IEEE*, 2012, pp. 3789–3792.
- [20] J. Alametsa, A. Palomaki, , and J. Viik, “Local ballistocardiographic spectrum studies from signals obtained from limbs and carotid artery with an emfi sensor induced with a tilt table,” in *Proc. 35th Annu. Int. IEEE Eng. Med. Biol. Soc. Conf.*, 2013, pp. 7008–7011.
- [21] Z. Chen, D. Lau, J. T. Teo, S. H. Ng, X. Yang, and P. L. Kei, “Simultaneous measurement of breathing rate and heart rate using a microbend multimode fiber optic sensor,” *J. Biomed. Opt.*, vol. 19, no. 5, pp. 057 001–1–057 001–12, May 2014.
- [22] B. Y. Su, K. C. Ho, M. Skubic, and L. Rosales, “Pulse rate estimation using hydraulic bed sensor,” in *Proc. 34th Annu. Int. IEEE Eng. Med. Biol. Soc. Conf. IEEE*, 2012, pp. 2587–2590.
- [23] Task Force of the European Society of Cardiology, “Heart rate variability standards of measurement, physiological interpretation, and clinical use,” *Eur. Heart J.*, vol. 17, pp. 354–381, 1996.
- [24] S. Andrews, I. Tsochantaridis, and T. Hofmann, “Support vector machines for multiple-instance learning,” in *Advances in Neural Inf. Process. Syst. (NIPS)*, 2003, pp. 577–584.
- [25] Y. Chen, J. Bi, and J. Z. Wang, “MILES: Multiple-instance learning via embedded instance selection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 1931–1947, 2006.
- [26] Q. Zhang and S. Goldman, “EM-DD: An improved multiple-instance learning technique,” in *Advances in Neural Inf. Process. Syst. (NIPS)*, 2001, pp. 1073–1080.
- [27] A. Zare and C. Jiao, “Extended functions of multiple instances for target characterization,” in *6th IEEE 2014 Workshop on Hyperspectral Image and Signal Process.: Evolution in Remote Sens. (WHISPERS)*, pp. 1–4.
- [28] S. G. Mallat and Z. Zhang, “Matching pursuits with time-frequency dictionaries,” *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, 1993.
- [29] S. S. Chen, D. L. Donoho, and M. A. Saunders, “Atomic decomposition by basis pursuit,” *SIAM J. Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.
- [30] E. J. Candès, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, 2006.
- [31] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, 2006.
- [32] M. Aharon, M. Elad, and A. Bruckstein, “K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Trans. on Signal Process.*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [33] B. A. Olshausen *et al.*, “Emergence of simple-cell receptive field properties by learning a sparse code for natural images,” *Nature*, vol. 381, no. 6583, pp. 607–609, 1996.
- [34] B. A. Olshausen and D. J. Field, “Sparse coding with an overcomplete basis set: A strategy employed by v1?” *Vision Research*, vol. 37, no. 23, pp. 3311–3325, 1997.
- [35] Y. C. Pati, R. Rezaifar, and P. S. Krishnaprasad, “Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition,” in *Proc. of the 27th Annu. Asilomar Conf. Signals Syst. Comput. IEEE*, 1993, pp. 40–44.
- [36] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
- [37] Q. Zhang and B. Li, “Discriminative K-SVD for dictionary learning in face recognition,” in *IEEE Conf. Comput. Vision Pattern Recognition (CVPR)*, 2010, pp. 2691–2698.
- [38] Z. Jiang, Z. Lin, and L. S. Davis, “Label consistent k-svd: Learning a discriminative dictionary for recognition,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2651–2664, 2013.
- [39] J. Mairal, F. Bach, and J. Ponce, “Task-driven dictionary learning,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 4, pp. 791–804, 2012.
- [40] I. Ramirez, P. Sprechmann, and G. Sapiro, “Classification and clustering via dictionary learning with structured incoherence and shared features,” in *IEEE Comput. Vision Pattern Recognition (CVPR)*, 2010, pp. 3501–3508.
- [41] D. P. Bertsekas, *Nonlinear Programming*. Athena Scientific, 1999.
- [42] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, “Online learning for matrix factorization and sparse coding,” *J. of Mach. Learning Research*, vol. 11, pp. 19–60, 2010.
- [43] M. A. Figueiredo and R. D. Nowak, “An EM algorithm

- for wavelet-based image restoration,” *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 906–916, 2003.
- [44] I. Daubechies, M. Defrise, and C. De Mol, “An iterative thresholding algorithm for linear inverse problems with a sparsity constraint,” *Commun. on Pure and Appl. Math.*, vol. 57, pp. 1413–1457, 2004.
- [45] J. Broadwater, R. Meth, and R. Chellappa, “A hybrid algorithm for subpixel detection in hyperspectral imagery,” in *IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, vol. 3, 2004, pp. 1601–1604.
- [46] M. J. Rantz, M. Skubic, M. Popescu, C. Galambos, R. J. Koopman, G. L. Alexander, L. J. Phillips, K. Musterman, J. Back, and S. J. Miller, “A new paradigm of technology-enabled vital signs’ for early detection of health change for older adults,” *Gerontology*, vol. 61, no. 3, pp. 281–290, 2015.
- [47] F. Facchinei and J.-S. Pang, *Finite-dimensional variational inequalities and complementarity problems*. Springer Science & Business Media, 2007.