Biomedical Signal Processing and Control 47 (2019) 75–87

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Self-adjustable domain adaptation in personalized ECG monitoring integrated with IR-UWB radar

Wenfeng Yin [a], Xiuzhu Yang [a], Lei Li [a], Lin Zhang [a,*], Nattapong Kitsuwan [b], Ryoichi Shinkuma [c], Eiji Oki [c]

[a] Beijing University of Posts and Telecommunications, Xitucheng Road No. 10, Beijing 100876, China
[b] The University of Electro-Communications, 1-5-1 Chofu-shi, Chofugaoka, Tokyo 182-8585, Japan
[c] Graduate School of Informatics, Kyoto University, Yoshida-honmachi, Sakyo-ku, Kyoto 606-8501, Japan

ARTICLE INFO

Article history:
Received 25 January 2018
Received in revised form 17 July 2018
Accepted 1 August 2018
Available online 23 August 2018

Msc:
92B20
68T05

Keywords:
Transfer learning
Domain adaptation
One-class classification
Self organizing maps
ECG monitoring

ABSTRACT

To enhance electrocardiogram (ECG) monitoring systems in personalized detections, deep neural networks (DNNs) are applied to overcome individual differences by periodical retraining. As introduced previously [4], DNNs relieve individual differences by fusing ECG with impulse radio ultra-wide band (IR-UWB) radar. However, such DNN-based ECG monitoring system tends to overfit into personal small datasets and is difficult to generalize to newly collected unlabeled data. This paper proposes a self-adjustable domain adaptation (SADA) strategy to prevent from overfitting and exploit unlabeled data. Firstly, this paper enlarges the database of ECG and radar data with actual records acquired from 28 testers and expanded by the data augmentation. Secondly, to utilize unlabeled data, SADA combines self organizing maps with the transfer learning in predicting labels. Thirdly, SADA integrates the one-class classification with domain adaptation algorithms to reduce overfitting. Based on our enlarged database and standard databases, a large dataset of 73200 records and a small one of 1849 records are built up to verify our proposal. Results show SADA's effectiveness in predicting labels and increments in the sensitivity of DNNs by 14.4% compared with existing domain adaptation algorithms.

1. Introduction

How to provide a personal diagnosis becomes a trend of the mobile electrocardiogram (ECG) monitoring, after the precise medicine is presented. The reliability of mobile ECG monitoring system is affected by the detection error caused by individual differences. Overcoming the issue of individual differences promotes the clinical usage of mobile ECG monitoring based on the Internet of Things. A mobile cloud hybrid solution for mobile ECG monitoring [1] is put forward to enhance the adaptation to individual differences. This solution applies the deep neural network (DNN) in the arrhythmia detection on the mobile side, and dynamically retrains this neural network on cloud using new personal data. Through the process of periodical retraining, the neural network is adjusted as adaptive to individual differences.

Though such DNN-based mobile ECG monitoring solution achieves the adaptation to individual features due to the virtue of periodical retraining procedures, it is subject to personal data used for fine-tuning. In real mobile health applications, new collected data are almost unlabeled and corresponding categories are imbalanced. Besides, on condition that the size of personal datasets is small, the retraining procedure tends to over-fit. These situations brings about issues on how to exploit unlabeled data, how to prevent over-fitting during retraining, and how to keep the recognition sensitivity for each class on the imbalanced dataset in the ECG monitoring solution.

The domain adaptation is able to relieve over-fitting of neural networks in the target domain, since it controls the learning in target domain with matching to the source domain. As a few labels are available in the target domain in real ECG monitoring applications, the domain adaptation is performable in a semi-supervised way, which exploits unlabeled data by predicting pseudo labels with a basic classifier [14] such as the support vector machine (SVM) [2] or the self-organizing maps (SOM) [19]. Since SOM offers unsupervised morphological clustering, it is adopted in this paper for classifying ECG waveforms. The SOM serves as a rough clustering method and therefore it is employed together with other clustering algorithms, such as the fuzzy c-means (FCM) clustering. However, existing SOM-based clustering is unable to integrate multiple data. Besides, the real ECG monitoring scenarios cannot meet the con-

* Corresponding author.
E-mail address: zhanglin@bupt.edu.cn (L. Zhang).

https://doi.org/10.1016/j.bspc.2018.08.002

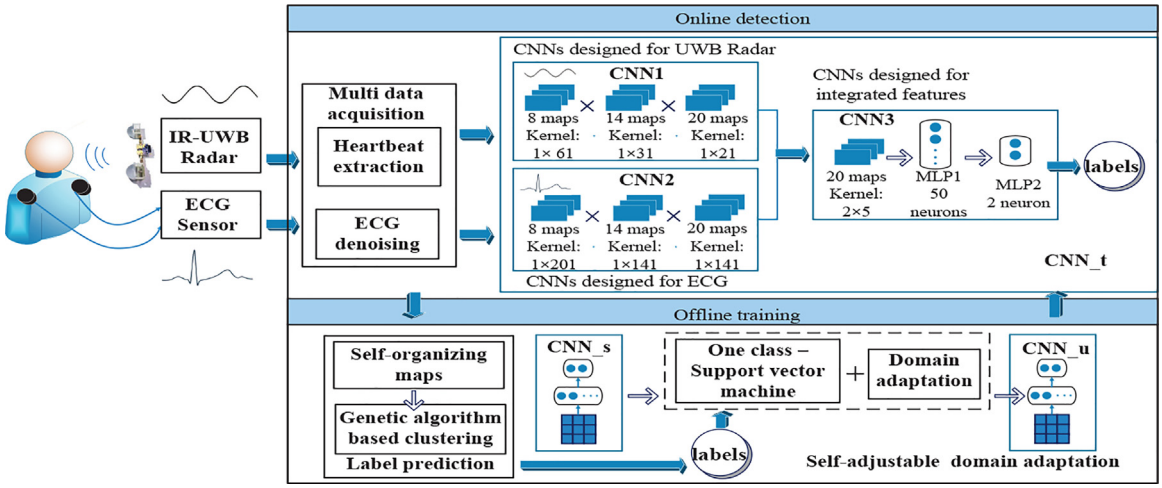

Fig. 1. Proposed strategy for mobile ECG monitoring. CNN_s denotes a standard neural network trained on the MIT-BIH database. CNN_u is adjusted by SADA. CNN_t is updated with weights of CNN_u and used for personalized diagnoses.

straint of existing domain adaptation algorithms that the set of classes is the same across domains. It is common that the target domain only acquires the normal category from a specific person.

This paper proposes a self-adjustable domain adaptation (SADA) strategy in the DNN-based mobile ECG monitoring system to prevent from overfitting to individual differences. Fig. 1 illustrates the architecture of a mobile ECG monitoring solution which is extended by SADA. To fuse ECG and radar data for relieving interferences of motion artifacts and individual differences, SADA adopts a cascade convolutional neural network introduced previously [4]. SADA modifies the SOM-based clustering to fuse multiple data and broadens the application scenarios of domain adaptation algorithms. Firstly, SADA extends the SOM-based clustering by the transfer learning (TL). Secondly, SADA applies the one-class SVM (OC-SVM) with the domain adaptation. By breaking the restriction of domain adaptation with the OC-SVM in our scenario, SADA increases the sensitivity of DNNs. Experimental results verify the effectiveness of SADA by comparisons with existing domain adaptation algorithms.

The rest of this paper is organized in the following. Section 2 gives a summary of related works on the domain adaptation and states details of the personalized ECG monitoring system integrated with IR-UWB radar. Section 3 presents the modification that SADA made to SOM-based clustering, describes issues on the consistency of class sets across domains and explains the extension of existing domain adaptation algorithms by applying the OC-SVM. Section 4 presents experimental results and their analyses. Finally, Section 5 concludes our paper.

2. Related works

2.1. Related works in domain adaptations

The deep structure of DNN is capable of extracting transferable features [13] specific to data and tasks. Since such features are domain-invariant, the deep domain adaptation dependent on DNNs benefits in these domain-invariant features in reducing distribution discrepancies across domains. The deep domain adaptation aligns domains through adding restrictions to the training loss of DNNs with an extra loss of domain fusion or domain classification. The domain fusion loss enables classifiers indistinguishable across domains, while the domain classification loss maintains extracted features distinct on target and source domains. The network architecture in [8] is adjusted to import the domain classification loss into the back-propagation. The network in [9] achieves the domain classification loss and the domain fusion loss simultaneously with a compromise deal. Works in [5–7,15] append the domain fusion loss by calculating the maximum mean discrepancy (MMD) of extracted features on source and target domains.

Specifically, the domain fusion loss is decreased by reducing discrepancies of the marginal distribution or the conditional distribution across domains [3]. The deep transfer network (DTN) [5] poses constraints to both marginal distribution and conditional distribution. The deep adaptation network (DAN) [6] presents a strategy for the optimal multi-kernel selection, to improve the marginal matching performance across multiple layers. The joint adaptation network (JAN) [7] restricts the joint distribution across multiple layers in the neural network. It is no longer necessary to separately adapt the marginal and conditional distribution.

As in [16], the domain adaptation requires that the class space keeps exactly consistent across target and source domains, which hinders its application in real-world scenario such as the mobile health monitoring. Though the network in [9] is applicable in situation that a part of categories lack labeled target data, it still obey the consistency constraint on class space. It is necessary to develop a flexible domain adaptation which is usable in either consistent or inconsistent class space. Then SADA makes a first attempt to broaden the application of domain adaptation to the situation that the class space shrinks in the target domain.

2.2. Personalized ECG monitoring integrated with IR-UWB radar

2.2.1. Multi data acquisition

In addition to the ECG monitoring, it is feasible to apply other contactless heart rate (HR) detections achieved through the capacitively coupled ECG measurement, the radar-based displacement measurement and the audio-based HR measurement [21]. The radar-based HR measurement outperforms other noncontact means due to its robustness in long distances (>10 cm) [22]. Besides, the radar-based HR measurement suppresses the affects of motion artifacts by applying multi antennas [23] or by adopting the generalized warblet transform to cancel random body movements [24]. Therefore, it is effective to apply the radar-based HR measurement to assist in vital sign monitoring scenarios such as the sleep apnea detection based on ECG or electroencephalograms [31–42].

Radar-based HR monitoring is achieved by the continuous-wave (CW) radar and the UWB radar, which includes the stepped frequency CW radar, the frequency-modulated CW radar and the IR-UWB radar [46]. Except the IR-UWB radar, other three kinds

Table 1
Frequency bands of wavelet coefficients in first four levels when ECG sampling rate is 360 Hz.

Wavelet decomposition		ECG features		
Levels	Frequency band (Hz)	Waves	Frequency band (Hz)	Energy ratio (%)
cd_1	90–180	–	–	–
cd_2	45–90	–	–	–
cd_3	22.5–45	QRS	0–38	98
cd_4	11.25–22.5	P	0–20	90
ca_4	0–11.25	T	0–8	92

of radars belong to the doppler radar which extracts vital signs from changing phase information. Because of the null-point problem, doppler radars have limitations in vital sign monitoring [45]. In contrast, IR-UWB radars acquire vital signs from echos' time of arrival and therefore avoid the null-point problem. Besides, IR-UWB radars have advantages in through-the-wall detections. Thus, IR-UWB radar based HR monitoring is selected as a non-invasive measurement in the personalized ECG monitoring system.

The DNN-based ECG monitoring system [4], which SADA is applied to, simultaneously collects ECG recordings and radar data for personalized diagnoses. Since the DNN-based classification is not dependent on predefined features and classifies ECG waveforms more accurately than other classifiers such as the SVM [25], the mobile ECG monitoring deploys neural networks for precise diagnoses. The vital signs carried by radar data vary in diverse individuals [21] and therefore are useful for generalizing DNN-based ECG monitoring against individual differences.

- **ECG acquisition module**

The ECG acquisition module compromises an ECG sensor and a denoising filter, as depicted in Fig. 1. The ECG sensor is based on the BMD101 chip and a bluetooth module BC147413. The denoising filter suppresses the baseline drift by a zero-phase-shift filter, removes power-line interferences at 50 Hz by a notch filter, and reduces the electromyographic interference based on the wavelet decomposition. Since the spectrum of EMG interference is overlapped with that of ECG [28], it is hard to denoise EMG interferences by bandpass filtering. Due to the multiscale frequency resolution, the wavelet decomposition is effective in reducing EMG interferences [28] and extracting features of ECG [26,27] for sleep apnea detections. The wavelet decomposition divides ECG into high frequency bands and low frequency bands in multiple scales. EMG interferences are dominant in high frequency bands while QRS complexes are mainly contained in low frequency bands [48]. Therefore, the wavelet-based denoising is able to reduce EMG interferences by setting thresholds to detailed coefficients of wavelet decompositions in high frequency bands.

In the ECG acquisition module implemented in Fig. 1, each ECG record is resampled from 512 Hz to 360 Hz before denoising. In this case, wavelet coefficients of first four levels respectively occupy frequency bands as shown in Table 1. $cd_1, \ldots, cd_4$ are detailed coefficients of level 1–4, and ca_4 is approximation coefficients of level 4. 98% energy of QRS complexes is distributed in band 0–38 Hz. The peak energy of QRS complexes is concentrated within a frequency band 10–20 Hz [47], which is overlapped with the band of cd_4. In order to reduce EMG interferences with less loss in QRS complexes, the acquisition module decomposes ECG into three levels with bior2.2 wavelet and removes EMG interferences by setting soft thresholds to detailed coefficients of wavelet decompositions in each level.

- **Acquisition module for radar data**

The acquisition module for radar data contains an IR-UWB radar platform and an extractor for heart beat signals. The radar platform is the NVA-R661 radar module, which is implemented on single chip CMOS NVA6201 impulse radar transceiver. The radar signal processing flow contains procedures including the removal of DC component, a band-pass filtering, the removal of clutters based on the principal component analysis (PCA), and the separation of heartbeats and respirations, as illustrated in Fig. 2.

The transmitting frequency band of applied radar module is selected according to the power spectral density (PSD) of transmitting impulses and the Chinese standard on UWB frequencies. Vital signals tend to be buried by clutters and noises when the transmitting PSD is low. Therefore, the frequency band with a higher PSD, provided by the applied radar, is selected as 5.65–7.95 GHz [49]. Cutoff frequencies of bandpass filtering are set as bounds of the transmitting frequency band to reduce noises out of the band. PCA-based methods suppress clutters by removing the component corresponding to the first eigenvalue in the singular value decomposition (SVD) of radar data [30]. Primary components corresponding to first N eigenvalues contain more clutter signals rather than vital signals. The PCA-based method calculates the energy ratio E_n of each components to all components and determines the N components to be removed on condition that $\sum_{i=1}^{N} E_n(i) > 0.85$. Observed in our experiments, the energy ratio of the first component is usually greater than 0.85. Thus N is selected as 1 in this paper.

ECG recordings and radar data are aligned by receiving timestamps before they are dealt with denoising and detections. Continuous data streaming is intercepted into segments by a timewindow without overlapping. After ECG segments are denoised and heartbeat signals are derived from radar segments, ECG segments and heartbeat signals are imported to a convolutional neural network (CNN) for arrythmia detections. Specific processing procedures are elaborated in the previous work [4].

2.2.2. Data fusion and periodical transfer learning

Given that features in ECG and radar data are related but distinct, a cascade convolutional neural network (CCNN) is designed to separatively capture different features from ECG and radar data and then integrates extracted features together. As illustrated in Fig. 1, CCNN comprises the CNN1 for radar data, the CNN2 for ECG recordings and the CNN3 for fusing features. Numbers of neurons, feature maps and layers of CCNN are determined by reference to a multi-lead convolutional neural network (LCNN) designed in [50]. Specifically, convolutional layer numbers of CNN1 and CNN2 are equal to those of LCNN. Map numbers in each layer of CNN1 and CNN2 are set as the same with corresponding layers in LCNN. Neuron numbers in multilayer perceptrons (MLP) of CNN3 are equivalent to MLP of LCNN. The sum of neuron numbers in output maps of CNN1 and CNN2 is equal to neuron numbers in feature maps in the third subsampling layer of LCNN. On the basis of consistent neuron numbers and map numbers with LCNN, advantages of CCNN integrating ECG with radar data are verified in [4]. Input-layer architectures of CNN1 and CNN2 are devised according to respective sampling rates of ECG and radar data. The input layer of CNN1 contains 360 neurons and the input layer of CNN2 contains 1200 neurons.

The cascade neural network is originally trained on the standard MIT-BIH database and a pre-collected dataset of ECG and radar data. To adapt to individual features of a specific person, the cascade neural network is fine-tuned on personal data based on the transfer learning. According to the transfer learning, the weights of standard cascade neural networks are sharable to initialize a new neural network, which is trained with new collected personal data. Through such periodical transfer learning, the standard cascade neural network is generalized to personal data. When the size of personal

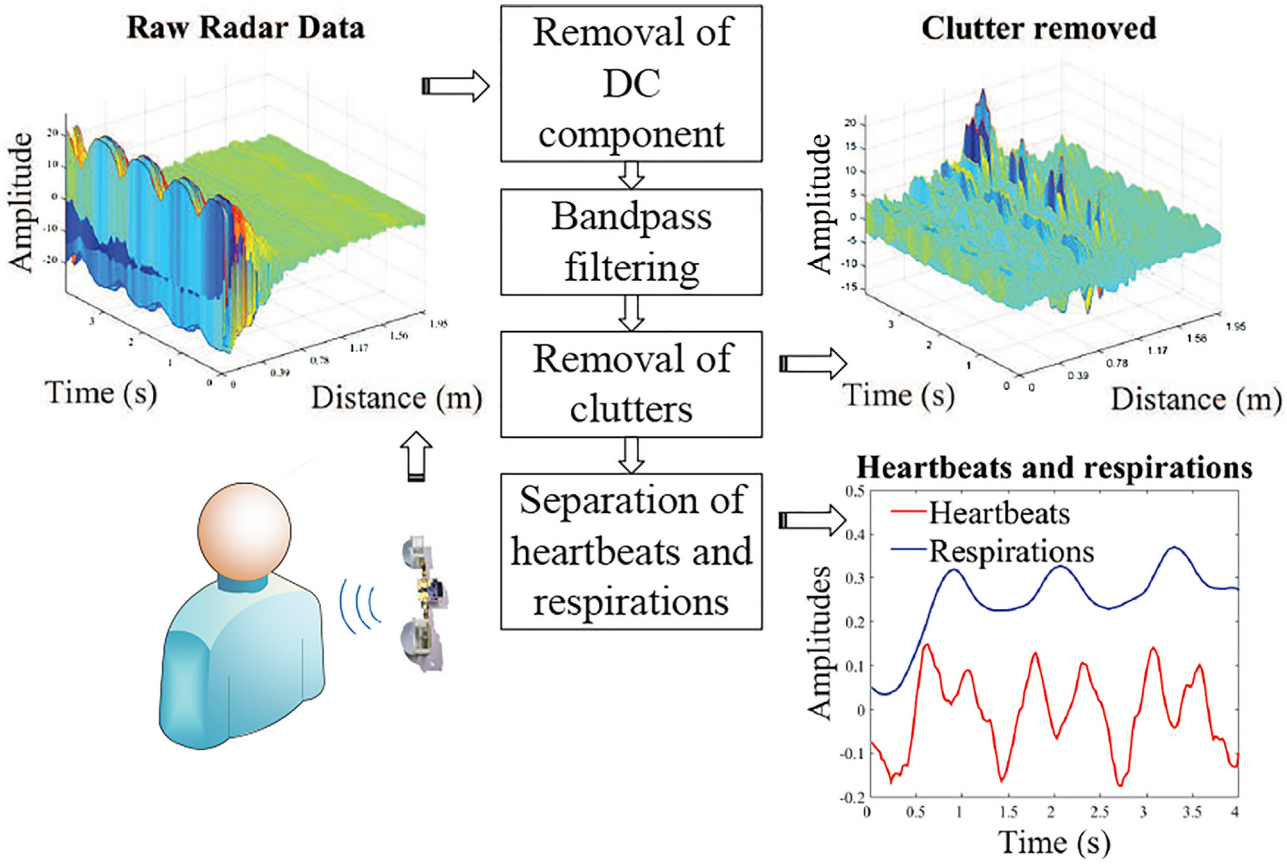

Fig. 2. Processing flow for heartbeat extraction.

dataset is inadequate, the retraining tends to over-fit into the current dataset at each time during periodically updating the neural network.

3. Self-adjustable domain adaptation strategy

3.1. Overview

SADA is implemented based on the ECG monitoring platform mentioned above. As illustrated in Fig. 1, ECG recordings and radar data are collected simultaneously for online detections by a cascade neural network and periodically imported to off-line retaining. SADA is performed in off-line retraining neural networks. Then the updated neural network is applied in online detecting abnormalities in ECG data and radar data. The periodical off-line retraining helps in keeping a robust accuracy in personal data. SADA respectively makes adjustments on the label prediction and the neural network fine-tuning. Since the back propogation of neural networks requires inputting labeled data, SADA predicts labels for unlabeled data by an extended SOM-based clustering. The SOM-based clustering is adjusted by the transfer learning and combined with the genetic algorithm (GA) based FCM clustering. Besides, SADA adopts the one class classification to eliminate limits of the domain adaptation. By deploying the extended domain adaptation, SADA enhances the generalization of neural networks and increases the classification sensitivity.

3.2. Label prediction by SOM

SADA predicts labels for unlabeled data by a morphological clustering based on the SOM and the FCM clustering. Specifically, SOM is modified by reducing errors in similarity measurements with dynamic time wrapping (DTW) [17]. TL is imported into SOM so as to fuse ECG and radar data. GA is combined with FCM clustering as the GA-FCM clustering to prevent FCM clsutering from trapping in local optimums.

3.2.1. Dynamic time wrapping

For clustering ECG recordings, it is more effective to perform the similarity measurement by applying dynamic time wrapping rather than Euclidian distances. DTW offers a robust measurement of similarity for time series. As stated in [11], deploying DTW in clustering ECG waveforms is helpful in increasing the clustering accuracy. As Fig. 3(a) illustrates, the time alignment is not available in a short-time recording, even though it is available for single beat. In this case, the euclidian distance is a measurement so strict that causes interferences in SOM clustering. In Fig. 3(b), DTW matches the point P to the point P_2 while the euclidian distance compares the point P to the point P_1.

3.2.2. Transfer learning based SOM

The IR-UWB radar serves as an effective noncontact meature-ment in the mobile health monitoring. In the situation that ECG data is badly interfered, it is helpful to integrate radar data as an assistance in diagnoses. Since the diagnosis information is consistent in radar data and ECG data, such as beat to beat intervals [18], results of arrhythmia diagnoses are the same for an ECG recording and a synchronized radar signal. Fig. 4 depicts the time relation of an ECG recording and a heartbeat signal extracted from the radar data. The knowledge learned in the radar data is transferable for augmenting the diagnosis accuracy in ECG data. Thus, SADA imports the radar data in clustering ECG data based on the concept of transfer learning.

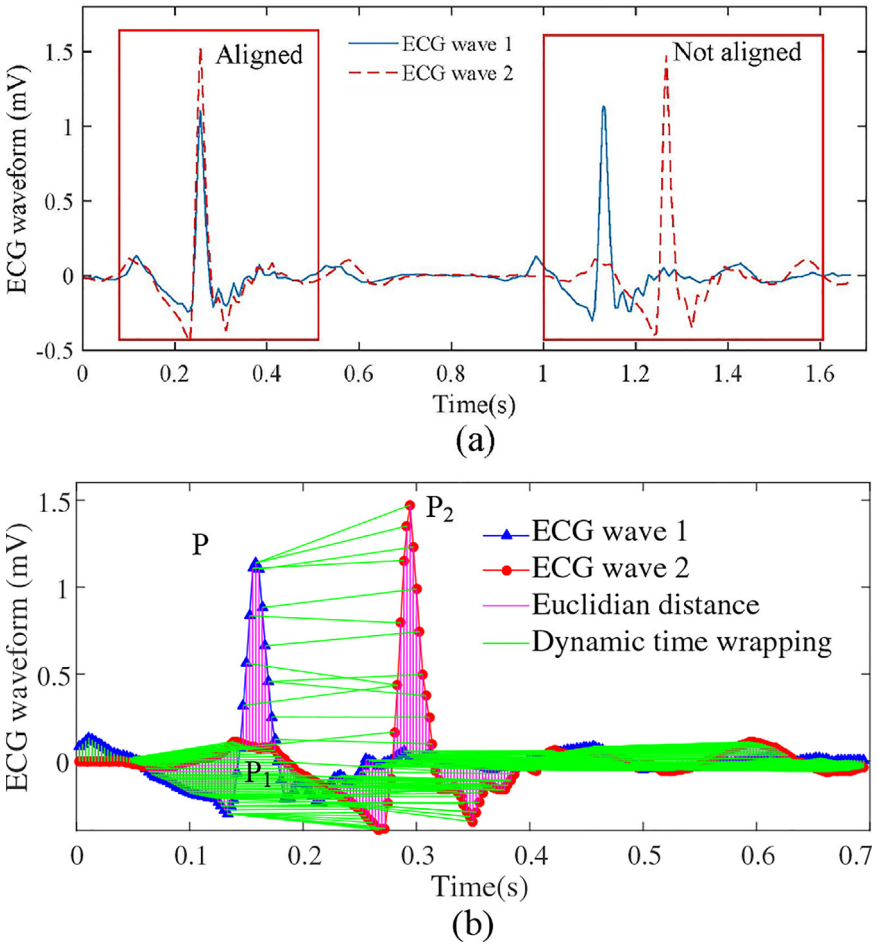

Fig. 3. (a) Time alignment in short-time recordings. (b) Similarity measurement by Euclidian distance and dynamic time wrapping.

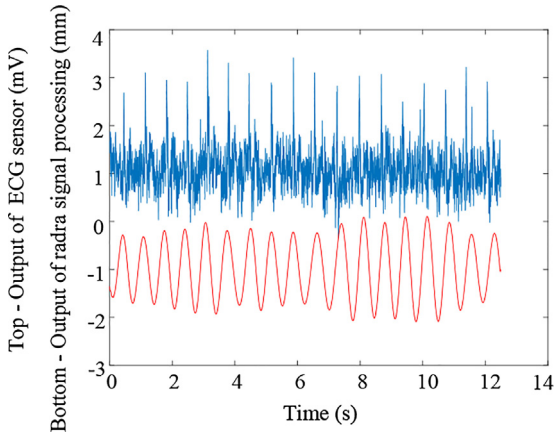

Fig. 4. Heart beat signal extracted from radar data and synchronized ECG signal.

SOM preserves the data distribution and topology during the mapping procedure. In the output map, SOM stimulates neighbor neurons of an activated neuron but suppresses remote neurons. The posterior probabilities of an output neuron are the same in ECG data and radar data, i.e., $P(m|x_{uwb}) = P(m|x_{ecg})$. The posterior probability $P(m|x)$ suggests the possibility that a specific sample x belongs to the mth output neuron. Since the KL divergence is usable in measuring discrepancy of two probability distributions, it is adopted in calculating the distance matrix *Ud* of output neurons. The distance matrix *Ud* carries information about the activated region. The active regions in the output map are distributed at adjacent areas in ECG data and radar data, as marked with red circles in Fig. 5. Such information of activated region is transferable, in the situation that output maps of SOMs for ECG data and radar data have the same topology.

As illustrated in Fig. 6, the transferable distance matrix U_d is obtained from radar data by calculating the KL divergence with the posterior possibility $P(m|x)$ of each output neuron in the first SOM1. Then a second SOM2 initializes its distance matrix by the shared matrix U_d. Specific procedures are described in Algorithm 1.

Algorithm 1. Flow of SOM improved by transfer learning

Input:	Heartbeat data from radar data, and ECG data.
Output:	Index of maps from ECG data to activated output neurons I_x, and centroids of each cluster d_x.
1:	Two self-organizing maps are set up with the same architecture, noted as SOM1 and SOM2.
2:	The first SOM1 obtains the posterior possibility of each output neuron $P(m\|x)$ from the heartbeat data;
3:	A sharing distance matrix U_d is built up through calculating the KL divergence;
4:	The distance matrix of a second SOM2 is initialized by U_d.
5:	The SOM2 generates rough clusters of ECG data, producing I_x and d_x.
6:	**return** I_x and d_x;

3.2.3. *Clustering combined with genetic algorithm*

The SOM is usually implemented in a two-stage clustering, whose second stage employs fine clustering methods such as the fuzzy c-means clustering to merge output neurons of SOM. The

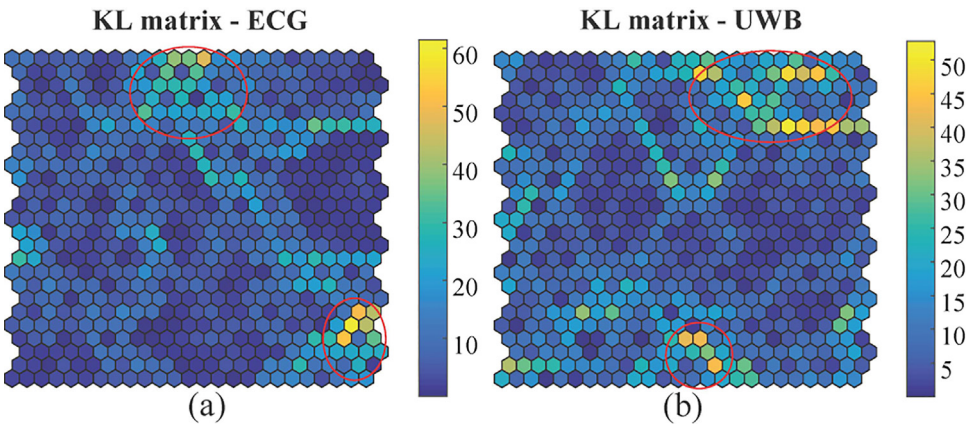

Fig. 5. KL-divergence matrix in (a) ECG data space and (b) UWB-radar data space.

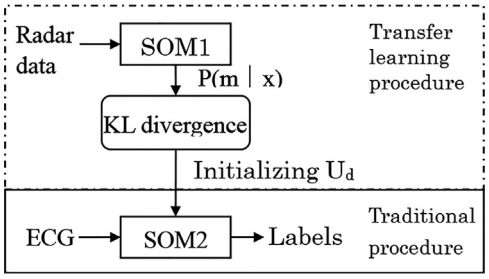

Fig. 6. Flow of SOM improved by transfer learning.

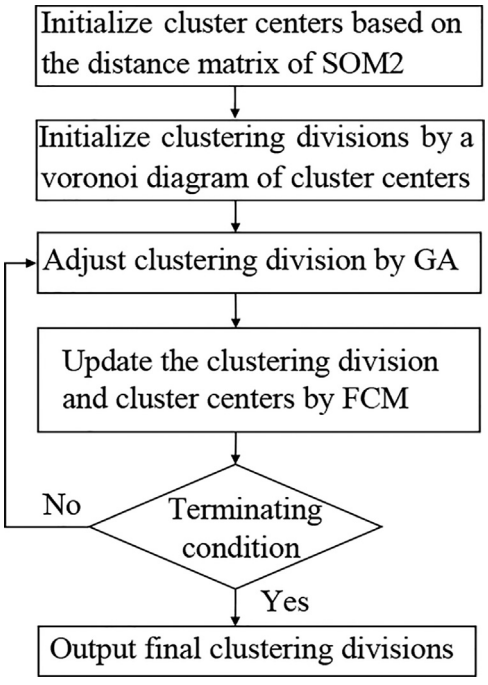

Fig. 7. Flow of clustering combined with genetic algorithm.

FCM clustering is sensitive to initial centers. The genetic algorithm is applicable with the FCM clustering method to optimize the selection of initial centers [20]. Besides, the genetic algorithm is importable into the FCM clustering to prevent trapping in a local optimum.

SADA adopts the genetic algorithm during the tuning iterations of FCM clustering to adjust the clustering division, as shown in Fig. 7. The optimization on clustering division is solved by the genetic algorithm through minimizing the sum of distances to clustering centers within each cluster. Specifically, the genetic algorithm tunes the cluster tags $Y=[y_1, \ldots, y_N]$ by the following optimization function as

$$\min_{t_i \in N+} \sum_{t_i=1}^{L} (D - \delta(Y - t_i)K_i), \tag{1}$$

where $D=[d_1, \ldots, d_N]$, and d_i is an M-dimensional feature vector of the ith node among total N active neurons in the output map of SOM. t_i is a tag for the ith cluster of total L clusters and K_i is the center of the ith cluster.

In the unsupervised clustering, SOM divides input data into rough categories by mapping data onto active neurons, and then the GA-based clustering method assigns specific categories for input data by merging active neurons into corresponding classes.

3.3. Domain adaptation based on OC-SVM

3.3.1. Issues on consistency of class sets

Existing domain adaptation algorithms are restricted to the assumption that the class set is consistent across domains. However, in the application scenario of mobile ECG monitoring, the class set changes at each time of periodical retraining neural networks. Because the appearance of arrhythmia is occasional, it is possible that the class set of new collected data contains only the normal category. Alternatively, the target domain captures one or several kinds of arrhythmias, as depicted in Fig. 8(a).

Specifically, the source domain X^s contains data of the normal category and N kinds of arrythmia, i.e., $X^s = \{x_n^s, x_a^s\}$, $x_a^s = \{x_{a1}^s, x_{a2}^s, \cdots, x_{aN}^s\}$, where x_a^s denotes all arrhythmias exist in the source domain. The target domain X^t comprises data of the normal category and a few abnormal types, i.e., $X^t = \{x_n^t, x_a^t\}$, $x_a^t = \{x_{a1}^t, \cdots, x_{aM}^t\}$, $M \leq N$, where x_a^t denotes all arrhythmias appear in the target domain. X_1^t in Fig. 8(b) illustrates a casual case that the target domain captures M diseases types of N types. Besides, X_2^t in Fig. 8(b) depicts an extreme case that the class set of target domain contains only the normal category, i.e., $x_a^t = \phi$.

When the neural network serves as a two-class classifier for telling normal data from abnormal data in the ECG monitoring scenario, the task is equivalent to a multi-label classification. As shown in Fig. 8(b), for each ECG recording, there are two main labels y_n and y_a indicating whether the current sample is normal. The abnormal label y_a comprises a series of sub tags denoting the corresponding type of arrhythmias, i.e., $y_a=[y_{a1}, \ldots, y_{aN}]$. In order to distinguish normal data from abnormal data, the neural network actually per-

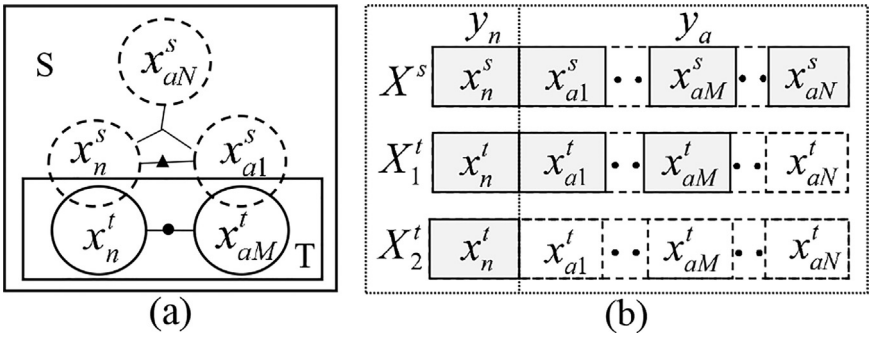

Fig. 8. Distributions in (a) data space and (b) label space.

forms a multi-label classification which recognizes all abnormal tags y_{ai} and then comes down to the abnormal label y_a. In case that the target domain only comprises normal data, the task of neural network is translated from the two-class classification to one-class classification. When the target domain has a few abnormal data, the neural network works as a multi-label classifier, which separates any kind of abnormal data from normal data. As the arrhythmia appears occasionally in the target domain, the class set cannot keep consistent across domains.

Within the source domain, the marginal distribution is $P(x = x_a^s) = \sum_{i=1}^{N} P(x_{ai}^s)$ and $P(x = x_n^s) = 1 - P(x = x_a^s)$. In the target domain, it assumes that a random variable with Binomial distribution B_i decides the possibility that an arrhythmia type exist in a specific ECG sample. If the i th arrhythmia type appeared, $B_i = 1$; otherwise, $B_i = 0$. The vector $B = [B_1, \ldots, B_N]$ indicates all arrhythmia types that appear in an ECG sample. Then, the marginal distribution of target domain is set as $P(x = x_a^t) = \sum_{i=1}^{N} B_i P(x_{ai}^t)$ and $P(x = x_n^t) = 1 - P(x = x_a^t)$. Such random changes in the marginal distributions have influences in the NN-based domain adaptation in classification sensitivity for each arrhythmia type.

Domain adaptation algorithms reduce the distribution discrepancy using the MMD. Taking the deep transfer network [5] for an example, the marginal discrepancy in the feature space is calculated as [10]

$$\mathrm{MMD} = \| \frac{1}{n^s} \sum_{i=1}^{n^s} x_i^s - \frac{1}{n^t} \sum_{j=1}^{n^t} x_j^t \|_2^2 = \mathrm{Tr}(\mathrm{XMX^T}), \tag{2}$$

where $\mathrm{X} = [\mathrm{X}^s, \mathrm{X}^t]$, $\mathrm{X}^s = [x_1^s, \cdots, x_{n^s}^s] \in R^{d \times n^s}$ is a source dataset, and $\mathrm{X}^t = [x_1^t, \cdots, x_{n^t}^t] \in R^{d \times n^t}$ is a target dataset. M is the MMD matrix, whose element M_{ij} is set as

$$M_{ij} = \begin{cases} 1/(n_s)^2, & i \le n^s, j \le n^s \\ 1/(n_t)^2, & i > n^s, j > n^s \\ -1/(n_s n_t), & \text{otherwise.} \end{cases} \tag{3}$$

In our situation, given that class sets are not consistent between domains, SADA attempts to covert class sets to equal sets. SADA adjusts the class set of source domain by adding a sparse diagonal matrix C to the dataset X. Then the MMD is devised as

$$\mathrm{MMD} = \mathrm{Tr}(\mathrm{XCMC^T X^T}), \tag{4}$$

where $C = \mathrm{diag}(c_1, \ldots, c_L)$, and $L = n_s + n_t$. The element c_j $(1 \le j \le L)$ is calculated as

$$c_j = \begin{cases} 1, & j \ge n_s \\ a_j, & 1 \le j \le n_s, \end{cases} \tag{5}$$

where a_j is an indicator whether the jth source data shares the same arrhythmia types with all target data. If the jth source data contains no arrhythmia type included in the target domain, $a_j = 0$; otherwise, $a_j = 1$. However, the indicator vector B is unknown during retraining neural networks, and therefore it is difficult to obtain the sparse matrix C. It is necessary to solve the problem in generating matrix C before the domain adaptation. SADA attempts to acquire the value of c_j by selecting proper source data through one-class classification.

3.3.2. One-class – support vector machine

One-class classification is usable in the situation that only one-class data is available for training. In other cases, when the training dataset is imbalanced in different classes, it is feasible to apply the one-class classification instead of multi-class classification, whose decision boundary is correspondingly deviated. The decision hypersphere of OC-SVM is determined by the quadratic programming [12]:

$$\begin{aligned} &\min_{r,c,\xi} r^2 + \frac{1}{\nu n} \sum_{i=1}^{n} \xi_i \\ &s.t. \quad \left\| c - \phi(x_i) \right\|^2 \le r^2 + \xi_i, \quad \xi_i \ge 0, \end{aligned} \tag{6}$$

where $\phi(x_i)$ is the mapping of ith training sample x_i in the feature space, n is the size of training dataset, c is the center of a hypersphere to enclose selected samples, r is the decision radius of the hypersphere, ξ_i is the error, and ν is a trade-off parameter between r and ξ_i.

The decision function of OC-SVM is set as $f(\mathrm{x}) = \|c - \phi(x)\| - r$. If $f(\mathrm{x}) \le 0$, then x is classified as a target; otherwise, it is recognized as an outlier.

3.3.3. One-class classification based domain adaptation

As shown in Fig. 9, the features of different arrhythmias are distributed in different parts of a feature space projected by kernel functions. Due to individual differences, there are drifts in feature distributions of the same arrythmia. The decision hypersphere of OC-SVM determined by features of target domain is possible to enclose source-domain features irrelative to arrhythmias in target domain. Therefore, the hypersphere needs to be adjusted to select proper source-domain features by adaptively updating the center of a decision hypershpere.

The process of one-class classification based domain adaptation is detailed in the Algorithm 2. During these procedures, the value of a_j in Eq. (5) is generated. Then final selected features of source domain and all features of target domain are used in the domain adaptation by existing algorithms.

Algorithm 2. Flow of one-class classification based domain adaptation

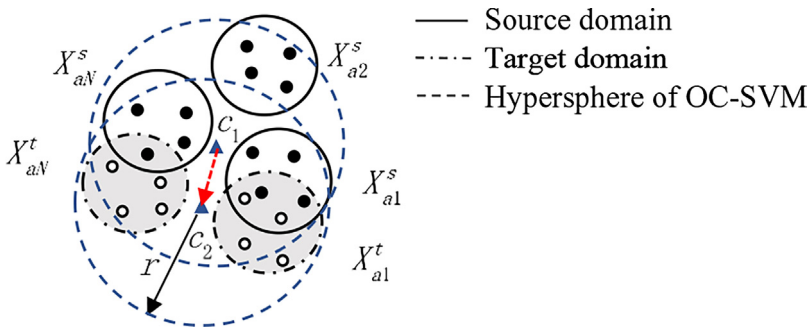

Fig. 9. Distributions of source and target domain in feature space with Gassian kernels.

Input: All features of source domain h^s, all features of target domain h^t, and tags of domains T_d.
Output: Selected features of source domain h^s_x, and an vector $\{a_j\}$ indicating whether a source feature is selected.
The first one-class classification stage:
1: A one class support vector machine is trained by features of target domain in the Hilbert space.
2: The trained OC-SVM is applied to select proper features of source domain, generating values for the vector $\{a_j\}$, $1 \le j \le n_s$.
The iterative one-class classification stage:
3: The selected features of source domain are used to train the next OC-SVM.
4: The next OC-SVM is applied to select proper features of source domain as final selected features h^s_x, updating values in the vector $\{a_j\}$.
5: **return** h^s_x and $\{a_j\}$.

The computational complexity of each one-class classification stage in SADA is $O(L^2)$, where L is the training batch size. Though the one-class classification stage is performed for N times, the extra computational complexity imported by SADA is still $O(L^2)$. The computational complexity of JAN is $O(L^2)$ and DTN has a computational complexity as $O(L)$. The optimization procedure of DAN causes a computational complexity $O(L^3)$. Thus computational complexities of JAN and DAN are not enlarged with adopting SADA. The sensitivity of DTN is increased by applying SADA, while its computational complexity rises as a trade off.

4. Experiments and analysis

In order to validate the performance of SADA, three evaluating parameters are calculated and compared from experiments in the following.

- Accuracy: it is the rate that ECG and radar data are classified into correct categories.
- Specificity: it is the rate that normal ECG data and radar data are categorized as normal.
- Sensitivity: it is the rate that abnormal ECG data and radar data are categorized as abnormal.

4.1. Experiments on SOM

In order to verify the clustering based on SOM and transfer learning, IR-UWB radar data and ECG data are acquired simultaneously from 28 testers in resting states and moving states. After denosing ECG data and extracting heartbeat signals from IR-UWB radar data, synchronized ECG data and heartbeat signals are imported to SOM-based clustering. Experiments are repeated 14 times. 246 records of ECG and heartbeat signals are used at each time and every record lasts for 1.5 seconds. Specific settings of validating schemes are elaborated in Table 2, where the symbol √ denotes the chosen method.

The applied topology of SOM contains 9×9 neurons. Weights of neurons in the SOM are adjusted with Euclidean distances or the DTW barycenter averaging. In settings of GA-based clustering, the population size is 50 and the generation number is 30. The individual size is equal to the number of active neurons in SOM. The crossover fraction is 0.8 and the migration fraction is 0.2.

Table 2
Settings on adopted methods in clustering procedures of four schemes.

Schemes	Rough clustering		Fine clustering		Input data	
	ED	DTW	FCM	GA-FCM	ECG	Radar data
SADA	–	√	–	√	√	√
TL-FCM	–	√	√	–	√	√
ED-FCM	√	–	√	–	√	–
ED-GAFCM	√	–	–	√	√	–

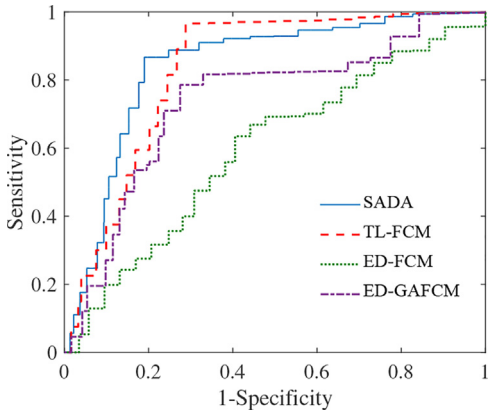

Fig. 10. Comparisons of 4 schemes in receiver operating characteristic space.

Table 3
Evaluating parameters of 4 clustering schemes.

Schemes	Acc (%)	Spec (%)	Sens (%)	F_1	*AUC*
SADA	**82.1**	**80.9**	**86.7**	**0.67**	**0.84**
TL-FCM	79.4	76.1	91.6	0.64	0.83
ED-FCM	72.3	69.3	83.4	0.40	0.59
ED-GAFCM	72.7	70.1	82.2	0.56	0.74

Fig. 10 illustrates the comparison of above four schemes in the receiver operating characteristic (ROC) space. Table 3 shows three parameters mentioned above and two other evaluating parameters including the F_1 score and the area under curves (AUC). As observed in either Table 3 or Fig. 10, SADA achieves a better performance than other three schemes in predicting labels. Evaluating parameters of the scheme TL-FCM are greater than those of the scheme ED-FCM, showing integrating multiple data by TL helps in increasing classification accuracies.

Table 4
Settings of vital sign datasets in two scales.

Datasets	Training dataset		Validation dataset		Test dataset	
	S	T	S	T	S	T
VS	38,700	25,534	V1: 1000	8480	T1: 375	8595
VS_mini	132	1290	V1_mini: 40	258	T1_mini: 40	387

4.2. Experiments on domain adaptation

4.2.1. Experimental settings

(i) Dataset settings in source domain and target domain: The source domain comprises 39075 records derived from the MIT/BIH database and the Chinese Cardiovascular Disease Database (CCDD) [29]. Specifically, 29700 records are extracted from the MIT/BIH database and 9375 records are obtained from CCDD, by dividing a recording into segments which last for 1.5 seconds with a sampling frequency 360 Hz. Radar records in source domain are generated as sinus signals with peaks aligned to R peaks in ECG records.

The target domain contains 34,129 records acquired from our dataset. Each ECG record is a segment for 1.5 s with a sampling rate 512 Hz and each radar record is a segment for 1.5 s with a sampling rate 40 Hz. Since heartbeats usually appear more than twice in 1.5 s, the generated ECG records contain information of RR intervals which are useful for detecting premature beats.

As declared in Table 4, the vital sign dataset is organized in two scales, respectively denoted as VS and VS _*m ini*, where S is the source domain and T is the target domain. The validation dataset of VS is marked as *V*1 and the test dataset of VS is marked as *T*1. The validation dataset of VS _*m ini* is marked as V1 _*m ini* and the test dataset of VS _*m ini* is marked as T1 _*m ini*. The scale of VS _*m ini* is similar to sizes of datasets adopted in [43,44] and adequate for training.

(ii) Validation schemes: SADA is employed in three existing domain adaptation algorithms, including the deep transfer network, the joint adaptation network, and the deep adaptation network. SADA generates a series of extended domain adaptation algorithms, different on the number of iterations of applying the OC-SVM. Specifically, DTN are expanded by implementing the OC-SVM twice, denoted as OC2-DTN. Besides, a standard network is trained in the source domain, named as A_0. A_0 is further fine-tuned in our dataset by transfer learning, denoted as TL-A. Training with each validation scheme is carried out by 14,000 iterations on the caffe platform.

(iii) Neural network settings: The adopted neural network has a cascade architecture presented in [4] to integrate ECG data and radar data. The neural network is retrained by transferring weights of the first three layers and then updating the whole neural network during fine-tuning. As analyzed in [13], features of shallow layers are common and features of deep layers are task specific. According to the strategy of checking transferable layers in [13], first three layers of the cascade convolutional neural network in [4] are sharable across different domains.

4.2.2. Results

Before retraining procedures, the original neural network A_0 is applied to the target domain. A_0 achieves an accuracy 35.7%, a specificity 32.0% and a sensitivity 78.0% in the *V*1 dataset, and obtains an accuracy 32.8%, a specificity 31.4% and a sensitivity 65.0% in the *T*1 dataset. The low values of accuracy and specificity suggest that A_0 performs poorly in the new dataset without fine-tuning. It also indicates that domain adaptation is needed for reducing affects of individual differences. Table 5 depicts results of adjusting A_0 by three existing domain adaptation algorithms and corresponding algorithms extended by SADA on the validation dataset. Results on the test dataset are shown in Table 6.

Table 5
Results of retraining procedure on validation dataset

Schemes	V1 Dataset			V1_mini Dataset		
	Acc (%)	Spec (%)	Sens (%)	Acc (%)	Spec (%)	Sens (%)
TL-A	84.2	87.3	46.1	82.1	89.8	59.8
JAN [7]	83.6	86.0	55.8	80.2	88.2	57.1
OC2-JAN	90.2	92.3	63.5	**91.9**	**95.6**	**82.0**
DAN [6]	71.2	73.0	47.3	80.2	90.9	49.0
OC2-DAN	84.0	86.5	53.5	83.3	90.5	62.0
DTN [5]	73.3	75.9	39.0	72.7	84.5	39.5
OC2-DTN	73.5	76.0	39.5	76.3	87.1	45.7

Table 6
Results of retraining procedure on test dataset.

Schemes	T1 Dataset			T1_mini Dataset		
	Acc (%)	Spec (%)	Sens (%)	Acc (%)	Spec (%)	Sens (%)
TL-A	84.6	86.8	30.0	78.0	85.3	54.8
JAN [7]	83.1	85.7	53.3	79.1	85.2	59.8
OC2-JAN	90.3	92.8	62.1	**87.4**	**91.9**	**72.6**
DAN [6]	72.6	74.7	49.3	75.9	86.4	41.6
OC2-DAN	84.3	87.1	53.3	77.7	84.6	56.0
DTN [5]	74.5	77.4	40.5	71.8	82.4	38.0
OC2-DTN	74.6	77.5	40.6	77.2	88.9	39.5

4.2.3. Analysis

In Tables 5 and 6, TL-A is trained by transfer learning but without domain adaptations. As results of TL-A reveal, the sensitivity cannot be increased as well as the accuracy and the specificity. Since the sensitivity index plays nonnegligible functions in the arrhythmia diagnosis, it is necessary to improve the sensitivity during dynamic retraining.

Focused on evaluating parameters of V1 _*m ini* dataset in Table 5, our proposal improves the sensitivity of DAN by 13% with OC2-DAN, and enhances the sensitivity of DTN by 6.2% with OC2-DTN. The sensitivity value of JAN is lifted up by 24.9% with OC2-JAN. Besides, all accuracies and specificities of three existing algorithms are increased by integrating with our proposal in the V1 _*m ini* dataset. Although the improvements of JAN and DAN in V1 dataset are not as large as those in V1 _*m ini* dataset, the sensitivity shows increments with our proposal in V1 dataset.

As observed in parameters of V1 _*m ini* dataset in Table 6, the sensitivity of DAN is raised by 14.4%, and the sensitivity of DTN is increased by 1.5%. As to JAN, the increment in sensitivity is up to 12.8% in V1 _*m ini* dataset with our proposal. Similarly, evaluating parameters of JAN and DAN take on increasing trends in V1 dataset by integrating our proposal.

Validation datasets V1 and V1 _*m ini* are employed in training as well and therefore evaluating parameters in Table 5 show greater values than those in Table 6. Since the decline of parameters in Tables 5 and 6 is not beyond 2.1% in V1 dataset and 9.4% in V1 _*m ini* dataset, it observes the generalization of neural networks and indicates that SADA does not cause over-fitting.

To verify influences of applying the OC-SVM, the cascade one-class classification stage in Algorithm 2 is repeated for multiple times, and results of JAN are depicted in Fig. 11. Increasing trends are observed in Fig. 11(a), (b) and (c) in both V1 dataset and V1 _*m ini* dataset. Besides, SADA helps in reducing loss values of JAN in

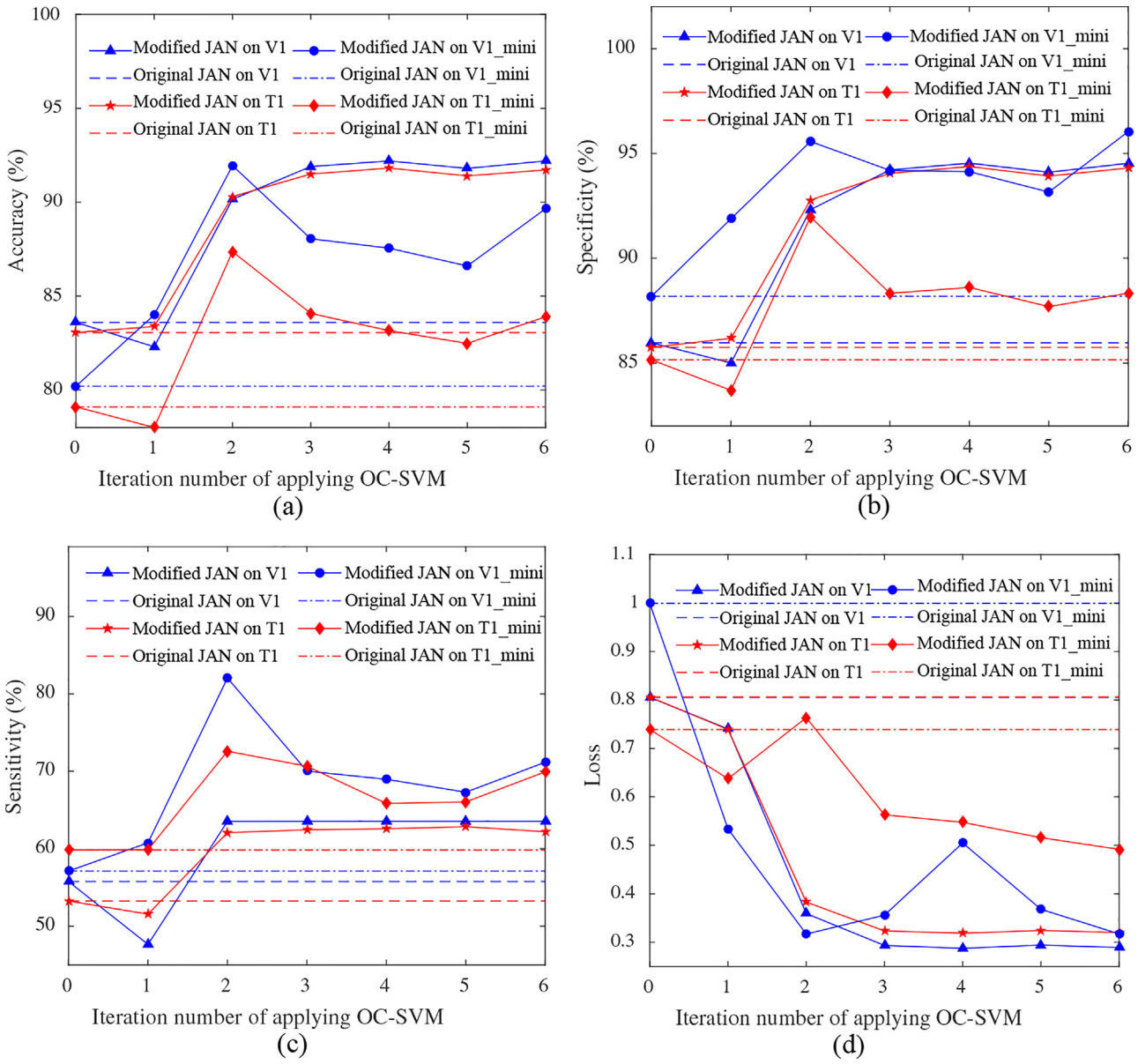

Fig. 11. Iterations of applying OC-SVM to JAN with (a) accuracy, (b) specificity, (c) sensitivity, and (d) loss.

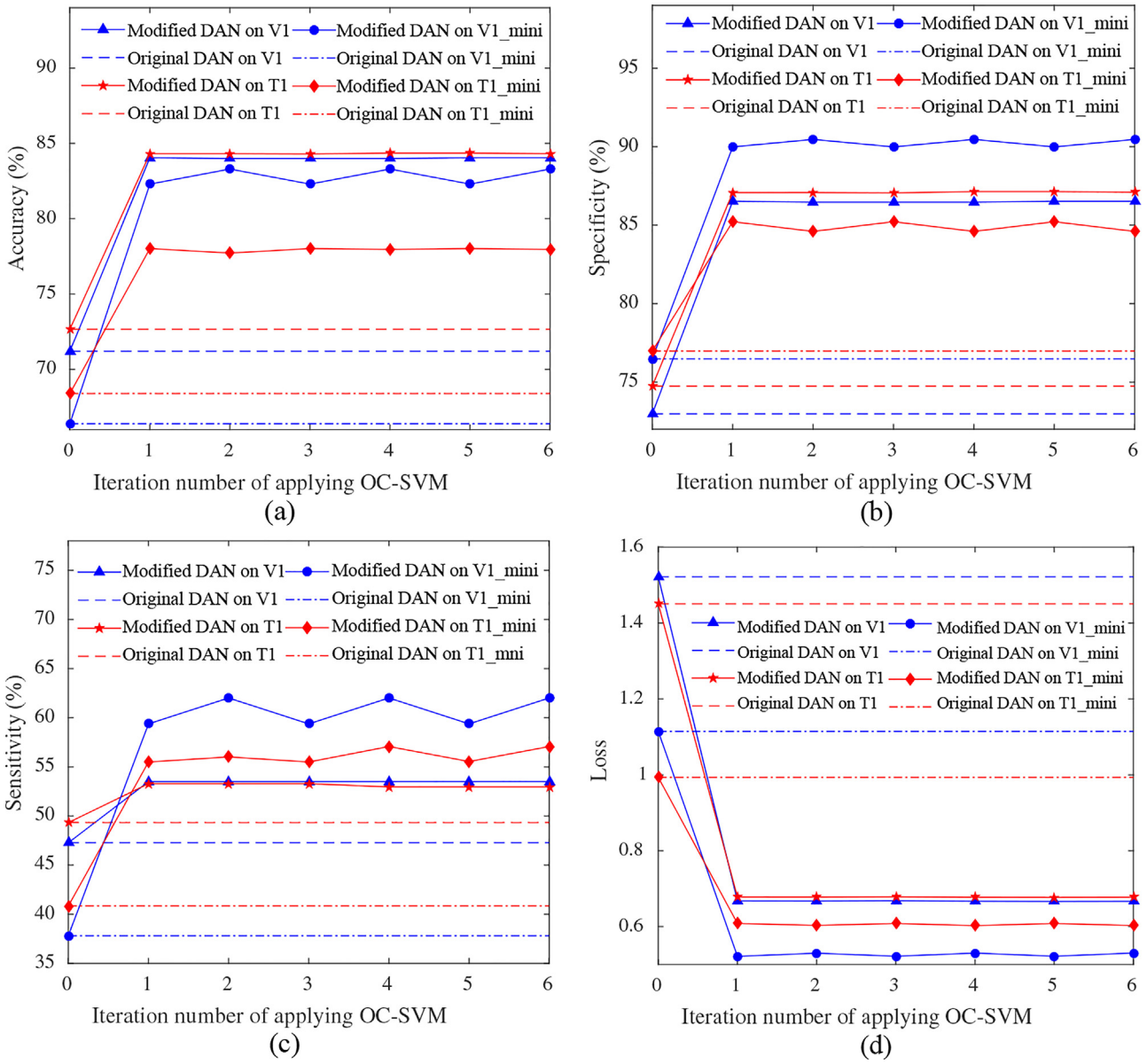

Fig. 12. Iterations of applying OC-SVM to DAN with (a) accuracy, (b) specificity, (c) sensitivity, and (d) loss.

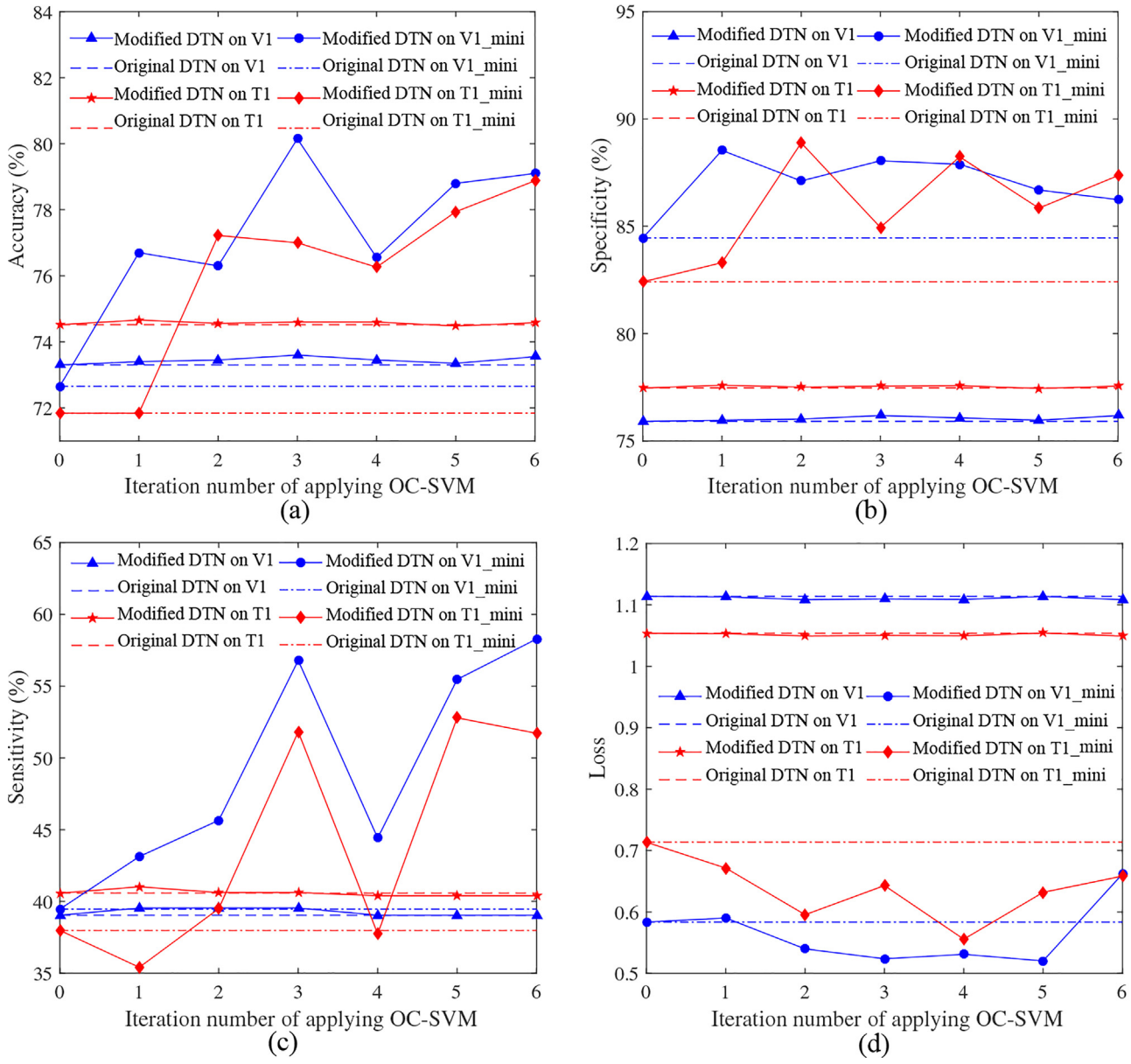

Fig. 13. Iterations of applying OC-SVM to DTN with (a) accuracy, (b) specificity, (c) sensitivity, and (d) loss.

datasets of two scales. Similar results are obtained in experiments that employ OC-SVM with DAN for multiple times, as elaborated in Fig. 12.

Fig. 13 depicts results of applying OC-SVM to DTN for multiple times. As in Fig. 13(a), (b) and (c), SADA lifts up evaluating parameters of DTN in V1 _*m ini* dataset, while evaluating parameters stays stable in V1 dataset. DTN calculates average discrepancies across domains without mapping into a kernel space. SADA applies OC-SVM to select proper sub-domain of source domain for matching. Since average discrepancies change little on the chosen sub-domain and the source domain, DTN shows few increments with SADA in the large-scale dataset.

Table 7 elaborates the time profiling of 100 retraining iterations in CPU and GPU mode on VS dataset with a batch size 500. The CPU version is Intel(R) Pentium(R) G3240 with a running frequency 3.10 GHz. The GPU version is GeForce GTX 1070. The extra time profiling that SADA imports to DTN reaches 283.5 s in the GPU mode. The extra time profiling that SADA imports to DAN reaches 274.9 s in the CPU mode at most. To a specific record, the extra time profiling added to DTN is 5.7 ms in the GPU mode, and the extra time profiling added to DAN is 5.5 ms in the CPU mode. Since the training procedure is periodically performed off-line, such extra time profilling is acceptable.

Table 7
Time profiling of 100 retraining iterations in CPU and GPU mode on VS dataset.

Schemes	CPU mode			GPU mode		
	JAN (s)	DAN (s)	DTN (s)	JAN (s)	DAN (s)	DTN (s)
Original	371.8	424.6	241.3	183.4	260.6	7.2
OC-	387.7	460.8	371.9	213.9	299.9	178.9
OC2-	406.6	509.6	394.8	227.2	321.6	205.9
OC3-	423.1	551.6	415.1	258.9	339.7	227.9
OC4-	447.3	595.4	429.7	275.2	354.8	243.8
OC5-	465.8	644.8	444.4	280.9	384.5	280.6
OC6-	494.0	699.5	461.4	322.8	416.1	290.7

5. Conclusion

This paper proposed the SADA strategy adjusting domain adaptation algorithms in ECG monitoring systems to fit into individual differences. SADA enlarges the dataset of ECG and IR-UWB radar data for training a cascade neural network designed in our previous work. As a semi-supervised domain adaptation algorithm, firstly, SADA predicts labels to exploit unlabeled data by combining SOM and TL. Secondly, SADA selects the matching domain by OC-SVM before deploying domain adaptations. Compared with three existing domain adaptation algorithms, experiments show that SADA increases the sensitivity of DNN by 14.4% in the small-scale dataset V1_*m ini* and 8.8% in large-scale dataset V1.

The sensitivity increments of SADA on small dataset verifies its effectiveness in real personal monitoring scenario whose dataset is not large. Currently, the OC-SVM adopted in SADA is based on single hypersphere induced by Gaussian kernels. In our future work, SADA is to extend the OC-SVM to a classifier built on multi hyperspheres, in order to handle more complicated feature distributions across domains in the kernel space.

References

[1] X. Wang, Q. Gui, B. Liu, Z. Jin, Y. Chen, Enabling smart personalized healthcare: a hybrid mobile-cloud approach for ECG telemonitoring, IEEE J. Biomed. Health Inf. 18 (3) (2014) 739–745.
[2] C. Venkatesan, P. Karthigaikumar, S. Anand Paul, S. Satheeskumaran, R. Kumar, ECG signal preprocessing and SVM classifier-based abnormality detection in remote healthcare applications, IEEE Access. 6 (2018) 9767–9773.
[3] M. Long, J. Wang, G. Ding, J. Sun, P.S. Yu, Transfer feature learning with joint distribution adaptation, Proc. 14th Int. Conf. Comput. Vis. (2013) 2200–2207.
[4] W. Yin, X. Yang, L. Zhang, E. Oki, ECG monitoring system integrated with IR-UWB radar based on CNN, IEEE Access. 4 (2016) 6344–6351.
[5] X. Zhang, F.X. Yu, S.F. Chang, S. Wang, Deep transfer network: unsupervised domain adaptation, Comput. Sci. (2015).
[6] M. Long, Y. Cao, J. Wang, M.I. Jordan, Learning transferable features with deep adaptation networks, Comput. Sci. (2015) 97–105.
[7] M. Long, H. Zhu, J. Wang, M.I. Jordan, Deep transfer learning with joint adaptation networks, Proc. 34th Int. Conf. Machine Learning (2017) 2208–2217.
[8] Y. Ganin, V. Lempitsky, Unsupervised domain adaptation by backpropagation, Proc. 32th Int. Conf. Machine Learning (2015) 1180–1189.
[9] E. Tzeng, J. Hoffman, T. Darrell, K. Saenko, Simultaneous deep transfer across domains and tasks, Proc. IEEE Int. Conf. Comput. Vis. (ICCV) (2015) 4068–4076.
[10] S.J. Pan, I.W. Tsang, J.T. Kwok, Q. Yang, Domain adaptation via transfer component analysis, IEEE Trans. Neural Netw. 22 (2) (2011) 199–210.
[11] A. Rashkovska, D. Kocev, R. Trobec, Clustering of heartbeats from ECG recordings obtained with wireless body sensors, Proc. of the 39th International Convention MIPRO (2016) 461–466.
[12] Z. Noumir, P. Honeine, C. Richard, On simple one-class classification methods, Proc. IEEE Int. Symp. Inf. Theory (2012) 2022–2026.
[13] J. Yosinski, J. Clune, Y. Bengio, H. Lipson, How transferable are features in deep neural networks, NIPS (2014) 3320–3328.
[14] W. Li, L. Duan, D. Xu, I.W. Tsang, Learning with augmented features for supervised and semi-supervised heterogeneous domain adaptation, IEEE Trans. Pattern Anal. Mach. Intell. 36 (6) (2014) 1134–1148.
[15] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, T. Darrell, Deep domain confusion: maximizing for domain invariance, Comput. Sci. (2014).
[16] H. Venkateswara, S. Chakraborty, S. Panchanathan, Deep-learning systems for domain adaptation in computer Vision: learning transferable feature representations, IEEE Signal Process. Mag. 34 (6) (2017) 117–129.
[17] A. Fonseka, D. Alahakoon, S. Bedingfield, GSOM sequence: an unsupervised dynamic approach for knowledge discovery in temporal data, Proc. IEEE Symp. Comput. Intell. Data Mining (2011) 232–238.
[18] W. Hu, Z. Zhao, Y. Wang, H. Zhang, F. Lin, Noncontact accurate measurement of cardiopulmonary activity using a compact quadrature doppler radar sensor, IEEE Trans. Biomed. Eng. 61 (3) (2014) 725–735.
[19] J. Kim, P. Mazumder, Energy-efficient hardware architecture of self-organizing map for ECG clustering in 65-nm CMOS, IEEE Trans. Circuits Syst. II Exp. Briefs 64 (9) (2017) 1097–1101.
[20] Y. Ding, X. Fu, Kernel-based fuzzy c-means clustering algorithm based on genetic algorithm, Neurocomputing 188 (2016) 233–238.
[21] J. Kranjec, S. Begus, G. Gersak, J. Drnovsek, Non-contact heart rate and heart rate variability measurements: a review, Biomed. Signal Process. Control 13 (2014) 102–112.
[22] J. Kranjec, S. Begus, J. Drnovsek, G. Gersak, Novel methods for noncontact heart rate measurement: a feasibility study, IEEE Trans. Instrum. Meas. 63 (4) (2014) 838–847.
[23] M.C. Tang, F.K. Wang, T.S. Horng, Single self-injection-locked radar with two antennas for monitoring vital signs with large body movement cancellation, IEEE Trans. Microw. Theory Tech. 65 (12) (2017) 5324–5333.
[24] S. Kazemi, A. Ghorbani, H. Amindavar, D.R. Morgan, Vital-sign extraction using bootstrap-based generalized Warblet transform in heart and respiration monitoring radar system, IEEE Trans. Instrum. Meas. 65 (2) (2016) 255–263.
[25] R.J. Martis, U.R. Acharya, L.C. Min, ECG beat classification using PCA LDA ICA and discrete wavelet transform, Biomed. Signal Process. Control 8 (5) (2013) 437–448.
[26] A.R. Hassan, M.A. Haque, An expert system for automated identification of obstructive sleep apnea from single-lead ECG using random under sampling boosting, Neurocomputing 235 (2017) 122–130.
[27] A.R. Hassan, Computer-aided obstructive sleep apnea detection using normal inverse Gaussian parameters and adaptive boosting, Biomed. Signal Process. Control 29 (2016) 22–30.
[28] P. Mithun, P.C. Pandey, T. Sebastian, P. Mishra, V.K. Pandey, A wavelet based technique for suppression of EMG noise and motion artifact in ambulatory ECG, Proc. IEEE Conf. Eng. Med. Biol. Soc. (EMBC) (2011) 7087–7090.
[29] J.W. Zhang, X. Liu, J. Dong, CCDD: an enhanced standard ECG database with its management & annotation tools, Int. J. Artif. Intell. Tool 21 (2012) 1C26.
[30] T.M. Grzegorczyk, B. Zhang, M.T. Cornick, Optimized SVD approach for the detection of weak subsurface targets from ground-penetrating radar data, IEEE Trans. Geosci. Rem. Sens. 51 (3) (2013) 1635–1642.
[31] A.R. Hassan, M.I.H. Bhuiyan, Computer-aided sleep staging using complete ensemble empirical mode decomposition with adaptive noise and bootstrap aggregating, Biomed. Signal Process. Control 24 (2016) 1–10.
[32] A.R. Hassan, M.I.H. Bhuiyan, Automatic sleep scoring using statistical features in the EMD domain and ensemble methods, Biocybern. Biomed. Eng. 36 (1) (2016) 248–255.
[33] A.R. Hassan, M.I.H. Bhuiyan, An automated method for sleep staging from EEG signals using normal inverse Gaussian parameters and adaptive boosting, Neurocomputing 219 (2017) 76–87.
[34] A.R. Hassan, M.I.H. Bhuiyan, A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features, J. Neurosci. Methods 271 (2016) 107–118.
[35] A.R. Hassan, A. Subasi, A decision support system for automated identification of sleep stages from single-channel EEG signals, Knowl. Based Syst. 128 (2017) 115–124.
[36] A.R. Hassan, M.I.H. Bhuiyan, Automated identification of sleep states from EEG signals by means of ensemble empirical mode decomposition and random under sampling boosting, Comput. Meth. Prog. Biol. 140 (2017) 201–210.
[37] A.R. Hassan, Automatic screening of obstructive sleep apnea from single-lead electrocardiogram, International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT) (2015) 1–6.
[38] A.R. Hassan, A comparative study of various classifiers for automated sleep apnea screening based on single-lead electrocardiogram, International Conference on Electrical and Electronic Engineering (ICEEE) (2015) 1–4.
[39] A.R. Hassan, M.A. Haque, Computer-aided sleep apnea diagnosis from single-lead electrocardiogram using dual tree complex wavelet transform and spectral features, International Conference on Electrical and Electronic Engineering (ICEEE) (2015) 1–4.
[40] A.R. Hassan, M.A. Haque, Computer-aided obstructive sleep apnea screening from single-lead electrocardiogram using statistical and spectral features and bootstrap aggregating, Biocybern. Biomed. Eng. 36 (1) (2016) 256–266.
[41] A.R. Hassan, M.A. Haque, Computer-aided obstructive sleep apnea identification using statistical features in the EMD domain and extreme learning machine, Biomed. Phys. Eng. Express 2 (3) (2016).
[42] A.R. Hassan, M.A. Haque, Identification of sleep apnea from single-lead electrocardiogram, IEEE International Conference on Computational Science and Engineering (CSE) (2016) 355–360.
[43] H.J. Lee, S.H. Hwang, S.M. Lee, Y.G. Lim, K.S. Park, Estimation of body postures on bed using unconstrained ECG measurements, IEEE J. Biomed. Health Inf. 17 (6) (2013) 985–993.
[44] M. Barni, P. Failla, R. Lazzeretti, A.-R. Sadeghi, T. Schneider, Privacy-preserving ECG classification with branching programs and neural networks, IEEE Trans. Inform. Forensics Secur. 6 (2) (2011) 452–468.
[45] V. Nguyen, A.Q. Javaid, M.A. Weitnauer, Harmonic path (HAPA) algorithm for non-contact vital signs monitoring with IR-UWB radar, Proc. IEEE BioCAS Conf. (2013) 146–149.
[46] X. Hu, T. Jin, Short-range vital signs sensing based on EEMD and CWT using IR-UWB radar, Sensors 16 (12) (2016).
[47] W. Zhang, M. Li, J. Zhao, Research on electrocardiogram signal noise reduction based on wavelet multi-resolution analysis, Wavelet Anal. Pattern Recogn. (2010) 351–354.
[48] L. Smital, M. Vtek, J. Kozumplk, I. Provaznk, Adaptive wavelet wiener filtering of ECG signals, IEEE Trans. Biomed. Eng. 60 (2) (2013) 437–445.
[49] Novelda Internet, NVA620x Preliminary Datasheet Impulse Radar Transceiver, 2013, Available from: www.xethru.com.
[50] L.P. Jin, J. Dong, Deep learning research on clinical electrocardiogram analysis, China Sci. 3 (2015) 398–416.