Information Sciences 582 (2022) 509–528

Contents lists available at ScienceDirect

Information Sciences

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

Unsupervised semantic-aware adaptive feature fusion network for arrhythmia detection

Panpan Feng [a,d], Jie Fu [a,e], Zhaoyang Ge [a,d], Haiyan Wang [c,d], Yanjie Zhou [b,*], Bing Zhou [a,d], Zongmin Wang [a,d]

[a] School of Information Engineering, Zhengzhou University, Zhengzhou, Henan 450000, China
[b] School of Management Engineering, Zhengzhou University, Zhengzhou, Henan 450000, China
[c] State Key Laboratory of Mathematical Engineering and Advanced Computing, Zhengzhou, Henan 450000, China
[d] Cooperative Innovation Center of Internet Healthcare, Zhengzhou University, Zhengzhou, Henan 450000, China
[e] Pattern Recognition Laboratory of Institute of Automation, Chinese Academy of Sciences, Beijing 100000, China

ARTICLE INFO

Article history:
Received 5 May 2021
Received in revised form 13 September 2021
Accepted 15 September 2021
Available online 17 September 2021

Keywords:
Arrhythmia detection
ECG individual differences
Unsupervised domain adaptation
Multi-view learning

ABSTRACT

An electrocardiogram (ECG) consists of complex P-QRS-T waves. Detecting long-term ECG recordings is time-consuming and error-prone for cardiologists. Deep neural networks (DNNs) can learn deep representations and empower automatic arrhythmia detection. However, when applying DNNs in practice, they usually suffer from domain shift that exits between the training data and testing data. Such shift can be caused by the high variability contained in ECG signals between patients and internal-variability of heartbeats for same patients, leading to degrading performance and impeding generalization of DNNs. To tackle this problem, we propose an unsupervised semantic-aware adaptive feature fusion network (USAFFN) to reduce such shift by alleviating the semantic distribution discrepancy between the feature spaces of two domains. Furthermore, an ECG contains rich information from different angles (beat, rhythm, and frequency levels), which is essential for arrhythmia detection. Therefore, a multi-perspective adaptive feature fusion (MPAFF) module is introduced to extract informative ECG representations. The experimental results show that the detection performance of our approach is highly competitive with the upper bound of alternative methods on the ARDB, and the generalization is confirmed on the INCART and LTDB.

1. Introduction

Arrhythmias refer to a large group of conditions in which there is abnormal activity or behavior in the heart [1]. Electrocardiograms (ECGs) are used routinely by medical personnel for the diagnoses of arrhythmias [1]. Current arrhythmia analysis typically adopts a beat-by-beat analysis strategy. The complex P-QRS-T waves and subtle changes of ECGs make it time-consuming and error-prone for professional cardiologists to analyze long-term recordings. Therefore, it is necessary to exploit automatic arrhythmia detection of beat-by-beat analysis based on machine learning strategy. However, the discrepancies in waveforms and characteristics of ECGs among different subjects and subject groups pose a major challenge for reli-

* Corresponding author.
E-mail address: ieyjzhou@zzu.edu.cn (Y. Zhou).

https://doi.org/10.1016/j.ins.2021.09.046
0020-0255/

able automatic heartbeat detection [2]. Such discrepancies are determined by the specific physiological structure of each individual heart.

A series of literature [2–7] on the inter-subject arrhythmia detection task have been proposed, mainly including patient-independent methods, patient-specific methods and domain-adaptation methods. The patient-independent methods tend to train a model with data from multiple patients and then directly apply this model to detect arrhythmias for new individuals. For example, Niu et al. [7] designed a convolution neural network utilizing data from multiple patients for heartbeat classification. Due to without adapting to the high inter-subject variability in ECG data, the patient-independent models generally perform poor generalization. In reality, therefore, it is difficult to apply these models to detect arrhythmias. From this, a natural solution is to take the specificity of each new individual in ECG data into account for adapting the patient-independent models, and thus, producing the patient-specific methods. Generally, the patient-specific methods tend to train a specific model for each individual. For instance, Hu et al. [2] proposed a mixture-of-experts (MOE) approach for patient-adaptable heartbeat classification. This work combined a global classifier and a subject-specific classifier, in which the global classifier was trained with large labeled ECG signals from multiple subjects and the subject-specific classifier was trained with several samples from the particular subject. However, two natural issues prevent the patient-specific methods from being applied in practice. Firstly, labeling a certain amount of ECG samples with each subject for training is impossible in reliable applications; Secondly, the parameters saved in the model increase linearly with increased numbers of patients. Instead of taking the specificity of each individual into account for adapting, several researchers term the discrepancies among ECG signals from inter-subject groups as domain shifts, and thus, they introduced domain-adaptation methods. Domain adaptation seeks to deal with the problem where training data (i.e. source domain) and test data (i.e. target domain) share a similar task but follow dissimilar distributions. Generally, in this setting, source domain contains labeled data and target domain contains unlabeled data. The domain-adaptation methods make full use of existing sufficient annotated data (source domain) to transfer valuable knowledge to unseen samples (target domain), thus bridging the gap between two domains. For example, Chen et al. [4] adapt the established classifier from source domain to target domain with two domain-regularizer loss functions. Niu et al. [5] proposed an unsupervised domain adaptation network to extract domain invariant representations via adversarial training for arrhythmia detection. Nevertheless, this adversarial domain adaptation method performs semantic-unawareness domain adaptation, thus negative transfer of knowledge may occur. In addition, the above methods do not take into account the diversity of information contained in ECG signals, so the features learned are not robust enough, resulting in poor performance.

Inspired by the above observations and considerations, an unsupervised semantic-aware adaptive feature fusion network (USAFFN) is proposed for inter-subject arrhythmia detection. Its main motivation is shown in Fig. 1. In particular, Fig. 1(a) presents the traditional adversarial domain adaptation framework, and it is dedicated to learn a shared feature space between different domains. In this framework, different categories of samples from different domains are regarded as a global feature space to align, but the semantic information contained in different categories of samples is ignored, resulting in the mixture of the feature spaces of different domains. In order to make up for the shortcoming, a semantic-aware adversarial domain adaptation framework is proposed, in which category information from different domains is integrated into the adversarial domain adaptation process, as shown in Fig. 1(b). This operation is able to align the global feature spaces of different domains and ensure the discriminability between the feature spaces of different categories. The experimental results turn out that compared with the traditional adversarial domain adaptation approach, our scheme mitigates the negative transfer of representations and the representations have better discrimination. In particular, the framework of the USAFFN is shown in Fig. 2. The USAFFN consists of a feature extractor, a domain discriminator, and a heartbeat classifier. The feature extractor is used to extract features of samples from different domains; The domain discriminator is mainly used to guide the feature extractor to realize a shared feature space of different domains alignment in the adversarial training process. At the same time, the semantic information from different domains is integrated into the feature extractor to ensure the discrimination between different feature spaces of categories; The heartbeat classifier is used to predict and output final results of the model. In addition, to extract informative ECG representations, a multi-perspective adaptive feature fusion (MPAFF) module is explored for the feature extractor. Its construction is inspired by the fact that ECGs carry rich information from different angles, and the information also plays variant roles in arrhythmia detection. The framework of the MPAFF module is shown in Fig. 3(b). It includes two sub-branches: the first sub-branch aims to extract multi-perspective features by a multi-branch attention block, and the second is to adaptively learn weights of the extracted features, and then fuse these weighted features together to obtain robust features.

The merits of this paper are summarized as follows:

- An unsupervised semantic-aware domain adaptation framework is developed for arrhythmia detection, and it is the first time to take the idea of semantic-aware adversarial domain adaptation and multi-perspective adaptive feature fusion into account for unsupervised domain adaptation. The experimental results show it achieves excellent performance.
- The USAFFN is proposed to alleviate semantic distribution discrepancies between the feature spaces of two domains, so as to achieve excellent performance on the target domain.
- The MPAFF module is explored to learn robust and informative ECG representations, where ECGs are characterized by complex P-QRS-T waves, subtle changes, and multiple views.
- The effectiveness and generalization of our framework are demonstrated on several benchmark databases. The experimental results are consistent with the theoretical analysis and the visualization analysis.

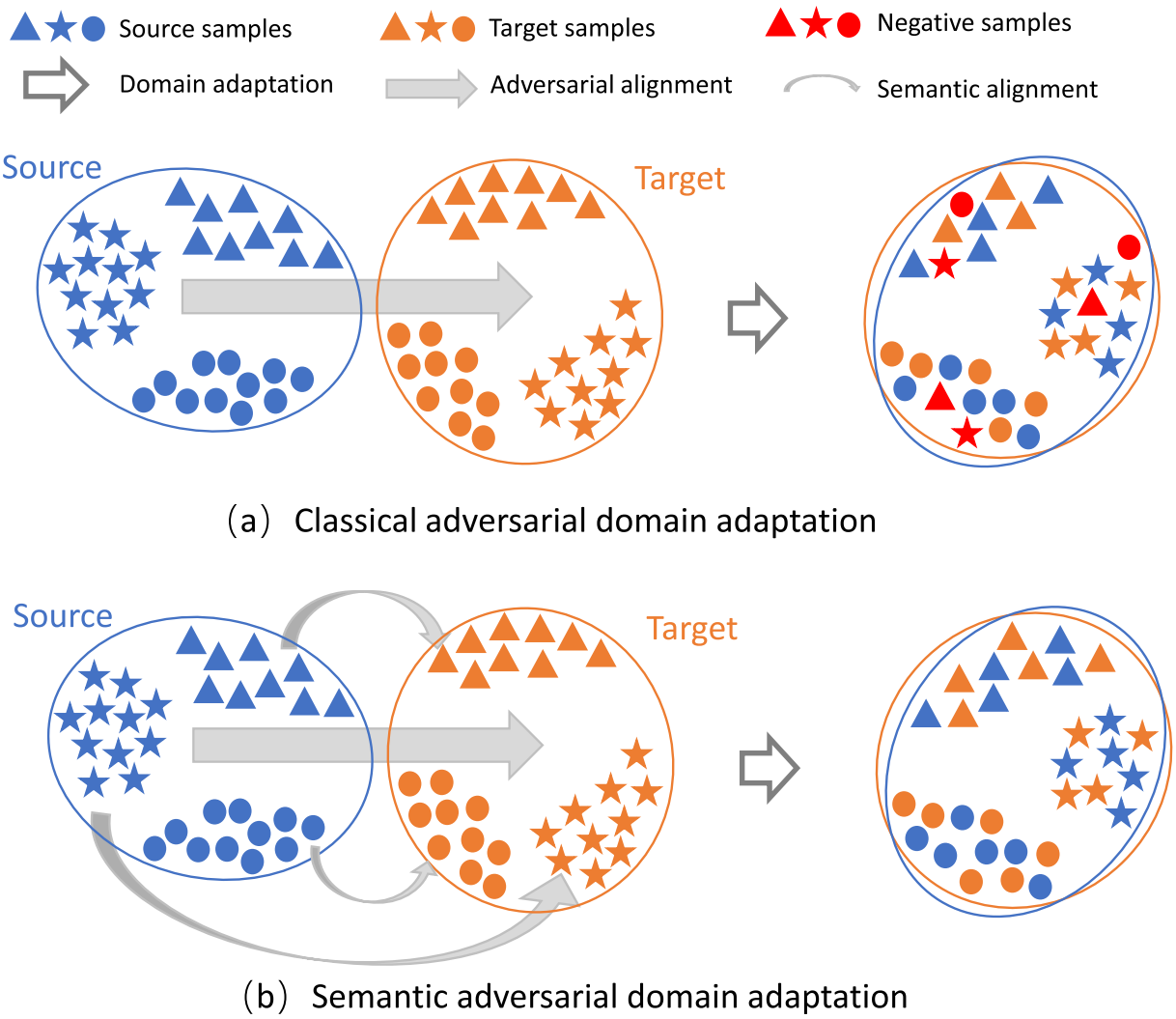


Fig. 1. Classical adversarial domain adaptation vs. semantic adversarial domain adaptation. (a): Traditional adversarial domain adaptation mines the similarity between different domains via the global feature space alignment. It performs category-agnostic domain adaptation. However, this may result in negative transfer of features. (b): Semantic adversarial domain adaptation integrates the category information from different domains into the traditional adversarial domain adaptation process. It mitigates the negative transfer of features via aligning the category distribution statistics of different domains.

The remainder of this paper is structured as follows. Section 2 introduces the related work. Section 3 describes the details of our method. Section 4 presents experiments and results. Finally, conclusions and future research directions are presented in Section 5.

2. Related work

This section includes three related aspects: arrhythmia detection, domain adaptation and multi-view learning, as detailed below.

2.1. Arrhythmia detection

Many attempts have been made to automatic arrhythmia detection in recent years [6–15], which are commonly divided into two types: conventional machine learning algorithms and deep learning algorithms. Traditional machine learning algorithms typically require handcrafted features to characterize ECG signals. These features contain wavelet [12], RR intervals [13], high-order statistics [14], wave amplitude [15], and so on. Lately, given the strong characteristic learning ability and outstanding performance of DNNs, arrhythmia detection models based on DNNs have been widely proposed [6–11]. Several researchers designed patient-independent arrhythmia detection models that do not adapt to the specificity of each new sub-

ject in ECG data [7,8,11]. Niu et al. [7] proposed a multi-perspective convolution neural network for inter-patient heartbeat classification. Nevertheless, this work required careful preprocessing of raw signals and eliminated patient-specific ECG representations that are meaningful for heartbeat classification. Mousavi et al. [8] combined a convolution neural network and a sequence-to-sequence model for arrhythmia detection. However, their constructed classification task is simple by selecting samples. Wang et al. [11] proposed an intelligent ECG-assisted annotation system, that not only supplemented labeled data, but also significantly reduced the workload compared with manual annotation. Nevertheless, this work required manual intervention. While these DNNs-based patient-independent models performed well within a dataset, it is unsatisfactory to apply these models directly to other datasets due to the distribution discrepancies of different test conditions. Some researchers developed patient-specific models for arrhythmia detection to adapt the specificity of each new subject in ECG data [6,9,10]. Zhai et al. [6] combined a dual heartbeat coupling approach and a convolution neural network for inter-patient arrhythmia classification. However, it involves selecting the most representative heartbeats. Ye et al. [9] proposed an automatic subject-adaptable heartbeat classifier based on semi-supervised learning and multi-view learning. It selected several high-confidence samples of each particular subject to train the model. Golany et al. [10] introduced a generative model to synthesize subject-specific ECG data to fine-tune the pre-trained model that trained with data of other subjects. Nevertheless, these algorithms typically require manual intervention in developing patient-specific models or the number of the models is equal to that of subjects, which hinders their application in practice.

2.2. Domain adaptation

Domain adaptation [16], with the ability leveraging the labeled source domain to learn a model for the target domain, is a subproblem of transfer learning. It typically includes two domains: source domain and target domain. It aims to address the problem where training data (i.e. source domain) and testing data (i.e. target domain) share a similar task but follow different probability distributions. Its main goal is to mine similarities between the two domains in order to improve the performance of the model in the target domain. Recently, domain adaptation based on DNNs (i.e. deep domain adaption) has attracted much attention [16]. Researches have done in-depth researches on deep domain adaptation and achieved inspiring results. For example, Chen et al. [17] constructed an image-to-image translation generate network based on CycleGAN [18] for unsupervised domain adaptation in chest X-ray segmentation. Unsupervised domain adaptation has also been applied to ECG classification [4,5,19–21]. Bazi et al. [19] proposed a domain transfer SVM model for arrhythmia detection, but required manual intervention. Jin et al. [20] developed a domain adaptive residual network to detect atrial fibrillation. Ammour [21] proposed a de-noising auto-encoder for feature learning and used hidden layers to reduce the data shift to detect atrial fibrillation. Beyond that, inspired by the success of unsupervised domain adversarial training [22–24], a few representative approaches have been available for ECG classification [4,5]. Specifically, Chen et al. [4] adopted an unsupervised domain adaptation model to classify heartbeat, which aligned category distributions between the two domains by introducing two domain-regularizer loss functions. Niu et al. [5] employed an unsupervised adversarial domain adaptation network for arrhythmia classification. This method aligned the global movement between the two domains, but ignored the semantic information in the samples, which may lead to negative feature transfer. In addition, ECG signals contain multiple patterns (beat, rhythm, and frequency levels) that have been ignored by previous feature extractors. In contrast, in this work, multi-view ECG features and sematic information based on unsupervised domain adaptation are taken into account for arrhythmia detection.

2.3. Multi-view learning

Multi-view learning is interpreted to boost learning performance by integrating the diversity of multi-view data. The views can be attained from diverse sources or constructed from different perspectives. Multi-view mechanisms have demonstrated their success based on two principles: complementary principle and consensus principle [25]. The former principle typically adopts a multi-perspective setting. Each perspective may contain knowledge that the others do not. This multi-perspective setting provides a comprehensive description of the data. In addition, compared with a single perspective, randomness and uncertainty are significantly reduced. The latter principle is commonly used to build different models in order to agree on the predictions for the same sample. Researchers have conducted in-depth studies on multi-view mechanisms and achieved inspiring results [3,26,27]. Ye et al. [3] employed the consensus principle to implement automatic arrhythmia detection by combining a general classifier and a specific classifier. Li et al. [26] utilized the complementary principle to construct three-dimensional inputs to enhance arrhythmia detection performance. Huang et al. [27] developed a multi-view feature fusion model for heartbeat classification. In this work, our model not only considers multi-view inputs, but also introduces a multi-perspective adaptive feature fusion module to obtain robust features.

3. Method

The discrepancies of ECG signals between different subjects pose a great challenge to deep learning algorithms. Therefore, our goal is to effectively mitigate the discrepancies and adapt the source model based on deep learning to the target domain.

3.1. Problem formulation

In this work, we model the inter-subject discrepancies as an unsupervised domain adaptation setting for arrhythmia detection. Specifically, there are two different but related domains, namely, the source domain D_S and the target domain D_T. The labeled samples from the source domain are defined as $D_S = \{X_S, Y_S\} = \left\{\left(x_S^i, y_S^i\right)\right\}_{i=1}^{N_S}$ and the unlabeled data from the target domain are defined as $D_T = \{X_T\} = \left\{x_T^i\right\}_{i=1}^{N_T}$. The goal is to learn a model trained with $\{X_S, Y_S\}$ and X_T so that it can correctly predict the labels of X_T.

To achieve this goal, traditional adversarial domain adaptation methods align the marginal distribution of the feature space between two domains by adversarial learning. These methods involve two processes. Firstly, these methods learn a model by minimizing the classification loss L_C on the source domain. Specifically,

$$L_C(f_F, f_C) = E[l(f_C(f_F(X_S)), Y_S)], \tag{1}$$

where $E[\cdot]$ is the statistical expectation, $l(\cdot,\cdot)$ denotes an appropriate loss function. f_F and f_C represent the feature extractor and the classifier, respectively.

Secondly, we aim to obtain domain-invariant representations to minimize the feature distribution distances between the two domains. These domain-invariant representations are obtained by the f_F, when they are able to confuse the domain discriminator f_D. The f_D aims to distinguish samples source and target domains. The f_F and f_D act as two opponents: the f_F would try to obtain domain-invariant representations to confuse the f_D, while the f_D, on the contrary, would try to detect whether these representations from the source domain or the target domain. The competition between f_F and f_D drives them to improve themselves, until the f_D is indistinguishable from the feature representations obtained by the f_F. Formally, this process can be achieved by iteratively optimizing the domain discrimination loss L_D and the domain confusion loss L_F:

$$\min_{\theta_D} \mathrm{L}_D(\mathrm{X_S}, \mathrm{X_T}, \theta_F; \theta_D), \tag{2}$$

$$\min_{\theta_F} \mathrm{L}_F(\mathrm{X_S}, \mathrm{X_T}, \theta_D; \theta_F). \tag{3}$$

The implementation details of the above will be introduced in Section 3.2.3. However, the main limitation of these traditional adversarial adaptation methods is that they mainly focus on aligning global domain statistics without considering the crucial semantic information for each category.

3.2. Our approach

To deal with the limitation mentioned above, a novel unsupervised domain adaptation framework is proposed for arrhythmia detection. Fig. 2 shows the outline of the framework. Briefly, an unsupervised semantic-aware adaptive feature fusion network (USAFFN) is developed. The USAFFN embeds category information into the adversarial domain adaptation process by introducing a semantic loss function. In addition, we leverage the fact that ECG signals contain rich information from different angles (beat, rhythm and frequency levels). A multi-perspective adaptive feature fusion (MPAFF) module is explored to obtain latent representations. The USAFFN is composed of three components: a shared feature extractor F with the MPAFF module, a domain discriminator D, and a heartbeat classifier C. The USAFFN and the MPAFF module are illustrated in Fig. 3(a) and (b) respectively.

In the following, the whole procedure of our approach, including preprocessing, network and optimization, is introduced in detail.

3.2.1. Preprocessing

Given the raw ECG data, the preprocessing procedure includes the following four steps:

1) Denoising: Noise usually distorts ECG signals and impairs arrhythmia detection. In order to increase the signal-to-noise ratio, we perform the baseline removal which is the same as that in the article [4].

2) Peak detection: After suppressing the baseline wander, the signals need to be separated into segments for classification. Before segmentation, the detection of R-peaks is required. The commonly used R-peaks detection methods include the Pan-Tompkins algorithm [28] and the gqrs package [29]. Recently, several algorithms [30–32] based on deep learning have been available for R-peaks detection. Herein, in order to look at the effect of our proposed module, we apply the algorithm [33] which is the same as that in the article [4] for R-peaks detection.

3) Segmentation: An ECG signal is represented as a sequence of P-QRS-T waves: a P wave followed by a QRS set and finally a T wave. After detecting R-peaks, instead of dividing continual ECG signals into separate units [34], we segment ECG signals into heartbeats by starting at the 0.14s (50 points at a sampling rate of 360) before the R-peak and ending at the 0.28s (100 points at a sampling rate of 360) after the R-peak, guaranteeing each segment contains one heartbeat. On this basis, we resize each segment to 128 sample points according to the selection of segment width in the article [26].

4) Input: Since beat-to-beat correlation can advance the performance of arrhythmia detection [6], we extract RR interval (a time interval between two consecutive R peaks) information including the pre-RR ration [35] (the ratio of the current pre-RR interval to the average of all pre-RR intervals of the corresponding records) and near-pre-RR ration [26] (the ratio of the

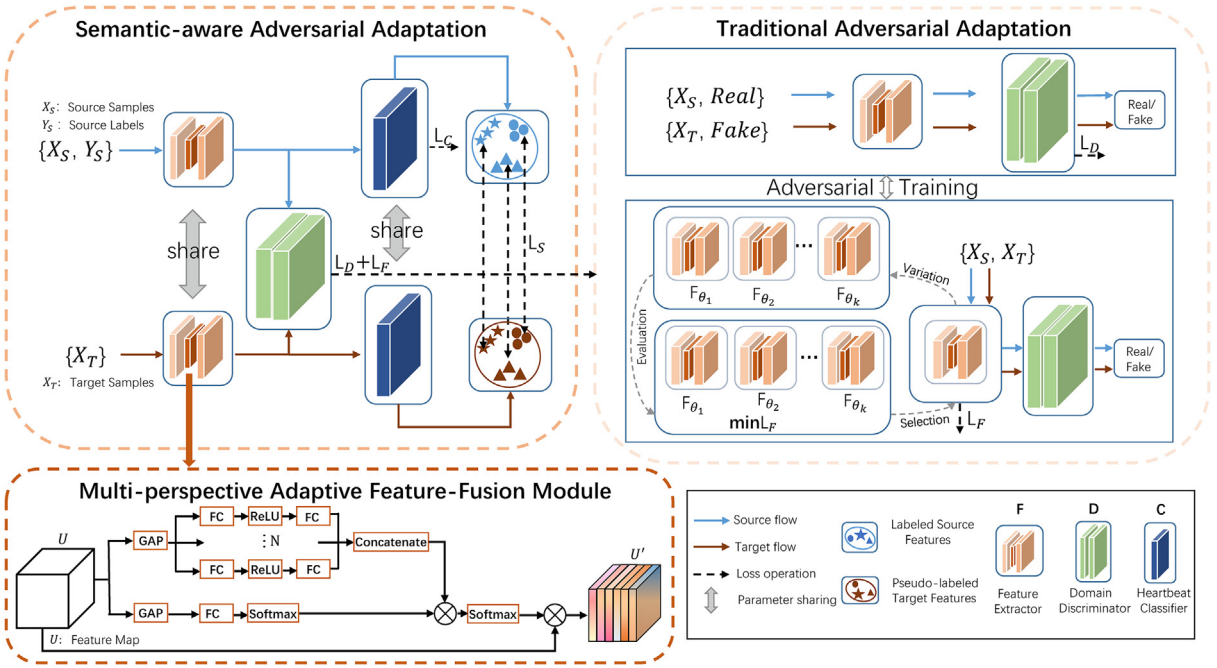


Fig. 2. The overview of our semantic-aware adversarial adaptation framework for arrhythmia detection. Traditional adversarial adaptation approaches align the marginal distribution of the feature space between source domain and target domain without considering the semantic information contained in samples. Our semantic-aware adversarial adaptation incorporates the category information into the traditional adversarial adaptation process by introducing a semantic loss function L_S. Moreover, we leverage the fact that ECG signals contain rich information from different angles (beat, rhythm and frequency levels). The MPAFF module is explored to obtain rich feature representations. The overall training process contains four loss functions: heartbeat classification loss function L_C, domain discrimination loss function L_D, domain confusion loss function L_F and semantic loss function L_S.

current pre-RR interval to the average of ten pre-RR intervals before the heartbeat). In addition, the ECG segment contains intra-heartbeat knowledge, the RR interval information and ECG segment are stacked together to form the input of the network.

3.2.2. Network

Fig. 3(a) shows the structure of the USAFFN. It includes three components: a shared feature extractor with the MPAFF module, a domain discriminator and a shared heartbeat classifier.

1) Feature extractor: We leverage the fact that consists of complex P-QRS-T waves, and these waves contain rich information from different perspectives (beat, rhythm, and frequency levels). A multi-perspective adaptive feature fusion (MPAFF) module is proposed to enhance the power of extracting feature representations. Besides, the skip connection is used to keep the original features not be suppressed and it also avoid the degeneration of the network. The structure of the MPAFF module is shown in Fig. 3(b), which is described below.

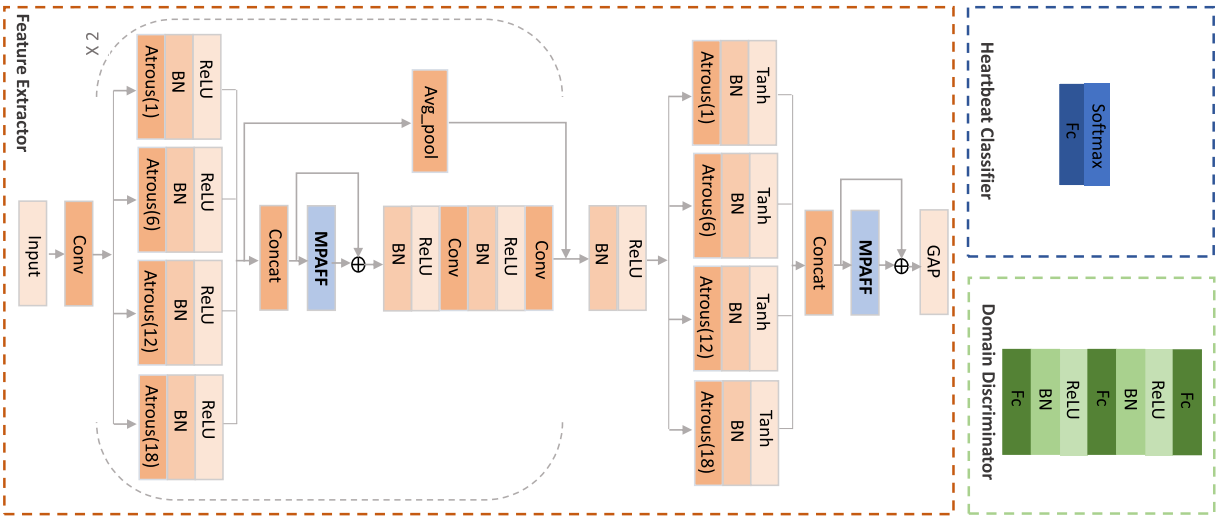
MPAFF module: As shown in Fig. 3(b), the MPAFF module consists of two sub-branches. The first sub-branch aims to obtain multiple perspective representations via a multi-branch attention module. The second sub-branch attempts to adaptively learn the weights of different features obtained by the first sub-branch. The first sub-branch is a tree-shaped structure containing multiple branches of the same size. It should be noted that although the structure of each branch is the same, they do not share parameters. According to Table 7, the number of the branches is selected as 3. Each branch can be interpreted as a single squeeze-and-excitation (SE) layer [36], which accounts for the interdependencies between channels. Specifically, in order to sum out the spatial information of the feature map, the squeeze operation is realized by global average pooling, denoted as F_{sq}. Then, in order to obtain the importance of each feature map, one fully connected (Fc) layer, one Relu layer and another Fc layer are used. This operation is termed as F_{se}. Thus, the output U^i of the i-th branch is obtained:

$$U^i = F_{se}(F_{sq}(U)) \in R^{C\times 1}, \tag{4}$$

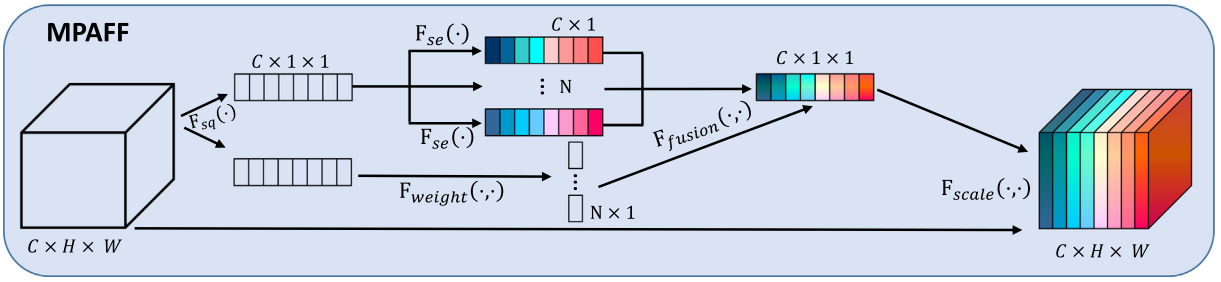
where U is the input feature map $U \in R^{C\times H\times W}$.

Concatenating the outputs of the above N branches together, the fusion space of the feature representation is obtained:

$$U_{all} = \left[U^1, U^2, \ldots, U^N\right] \in R^{C\times N}. \tag{5}$$



(a) The unsupervised semantic-aware adaptive feature fusion network.



(b) The multi-perspective adaptive feature fusion module.

Fig. 3. (a): The structure of the USAFFN. (b): The structure of the MPAFF module.

In the second sub-branch, in order to obtain the weight of the feature fusion output U_{all}, two operations are performed. Firstly, global average pooling is introduced to remove the spatial dimension, and then an adaptive weighting operation including one Fc layer and a softmax function is applied:

$$F_{weight}(U) = \text{softmax}(W_1 F_{sq}(U)) \in R^{N\times 1}, \tag{6}$$

where $W_1 \in R^{N\times C}$ is the weight matrix of softmax layer.

Then, the adaptive weighted result is obtained:

$$U_{all} F_{weight}(U) \in R^{C\times 1}. \tag{7}$$

Finally, the adaptive output $\tilde{U}$ is obtained:

$$\tilde{U} = F_{scale}(U, \text{sigmoid}(U_{all} F_{weight}(U))) \in R^{C\times H\times W}, \tag{8}$$

where the operation F_{scale} refers to the channel-wise multiplication.

In this way, compared with only one branch layer, the MPAFF module utilizes multi-perspective adaptive feature representation fusion to extract richer information.

2) Domain discriminator: The domain discriminator is used to distinguish whether the features obtained by the feature extractor come from the source domain or the target domain. It is usually implemented by performing binary classification. In the optimization process, the domain discriminator and the feature extractor act as two opponents: the goal of feature extractor is to obtain domain invariant features, so as to confuse the domain discriminator, while the domain discriminator attempts to distinguish whether these features come from the source domain or the target domain. The competition encourages them to improve themselves until the domain discriminator cannot distinguish the features obtained by the feature extractor. The structure of the domain discriminator is shown in Fig. 3(a).

3) Heartbeat classifier: To decode the deep representations from the feature extractor, a simple yet efficient classifier is applied, as shown in Fig. 3(a), with one Fc layer, followed by a softmax function.

3.2.3. Optimization

It is noted that traditional adversarial adaptation methods align global domain statistics without considering the semantic information contained in samples. Thus, the USAFFN incorporates category information into the adversarial adaptation process by introducing a semantic loss function L_S. The optimization process based on the traditional adversarial unsupervised domain adaptation method is presented as follows. Firstly, the desired heartbeat classifier is obtained by minimizing the weighted cross-entropy classification loss function L_C on the source domain:

$$L_C(X_S, Y_S; \theta_F, \theta_C) = -\frac{1}{N_s}\sum_{i=1}^{N_s} w_i y_i \log\left(f_C\left(f_F\left(x_i^s\right)\right)\right), \tag{9}$$

where θ_C is the parameter of the heartbeat classifier f_C; θ_F is the parameter of the feature extractor f_F.

Secondly, in order to obtain the domain-invariant features, a domain discriminator f_D with the parameter θ_D is used to distinguish whether the features obtained by the f_F come from the source domain or the target domain. These deep domain-invariant representations are obtained by optimizing two loss functions: the domain discrimination loss function L_D and the domain confusion loss function L_F:

$$L_D(X_S, X_T, \theta_F; \theta_D) = -\sum_d \mathrm{I}[y_D = d] \log q_d, \tag{10}$$

$$L_F(X_S, X_T, \theta_D; \theta_F) = -\sum_d \frac{1}{D} \log q_d, \tag{11}$$

where y_D denotes the domain label, q_d corresponds to the softmax activated by the domain discriminator: $q_d = \mathrm{softmax}(\theta_D^T f_D(x; \theta_F))$. $\frac{1}{D}$ is the probability of domain. The loss function L_D only updates the parameter θ_D, seeking to obtain the best domain discriminator. The loss function L_F only updates θ_F, attempting to obtain domain-invariant features. These features are obtained when the best domain discriminator performs poorly. Obviously, the competition encourages the f_F and the f_D to improve themselves until the f_D cannot distinguish the features obtained by the f_F. According to the parameters obtained from the previous iteration, the θ_D and θ_F are iteratively updated to optimize.

Algorithm 1. The calculation process of the semantic loss function

Input: Labeled samples S, unlabeled samples T, batch size M, iterations N, global and local centers for two domains: C_S, C_T, C_S^k and C_T^k

Output: Semantic alignment loss function: L_S

$L_S = 0$

$C_S = \frac{1}{|S|}\sum_{(x,y)\in S} f_F(x)$, $C_T = \frac{1}{|T|}\sum_{(x,y)\in T} f_F(x)$

for $i = 1$ to N

  S_i = Random Sampling (S,M), T_i = Random Sampling (T,M)

  $\widetilde{T}_i = \mathrm{Annotating}(f_F, f_C, T_i)$

  $C_{S_{(i)}} \leftarrow \frac{1}{|S_i|}\sum_{(x_j,y_j)\in S_i} f_F(x_j)$, $C_{T_{(i)}} \leftarrow \frac{1}{|T_i|}\sum_{(x_j,y_j)\in T_i} f_F(x_j)$

  $C_S \leftarrow (1-\alpha)C_S + \alpha C_{S_{(i)}}$, $C_T \leftarrow (1-\alpha)C_T + \alpha C_{T_{(i)}}$ **do**

  for $k = 1$ to K **do**

    $C_S^k = \frac{1}{|S^k|}\sum_{(x,y)\in S^k} f_F(x)$, $C_T^k \leftarrow \frac{1}{\left|\widetilde{T^k}\right|}\sum_{(x,y)\in \widetilde{T^k}} f_F(x)$

    $C_{S_{(i)}}^k \leftarrow \frac{1}{|S_i^k|}\sum_{(x_j,y_j)\in S_i^k} f_F(x_j)$, $C_{T_{(i)}}^k \leftarrow \frac{1}{|\widetilde{T_i^k}|}\sum_{(x_j,y_j)\in \widetilde{T_i^k}} f_F(x_j)$

    $C_S^k \leftarrow (1-\alpha)C_S^k + \alpha C_{S_{(i)}}^k$, $C_T^k \leftarrow (1-\alpha)C_T^k + \alpha C_{T_{(i)}}^k$

    $L_S \leftarrow L_S + d(C_S, C_T) + d\left(C_S^k, C_T^k\right)$

  end for

end for

However, the global alignment based on traditional adversarial adaptation approaches ignores the semantic information contained in samples. Therefore, the USAFFN is used to incorporate category information into the adversarial adaptation process by introducing a semantic loss function L_S. The L_S implements category-level migrations via aligning the global centers and the local centers on the two domains:

$$L_S(X_S, Y_S, X_T; \theta_F) = d(C_S, C_T) + \sum_{k=1}^{K} d\left(C_S^k, C_T^k\right), \tag{12}$$

where K is the number of categories. C_S and C_S^k are the global center and the k-th local center for the source domain S respectively. C_T and C_T^k are the global center and the k-th local center for the target domain T respectively. $d(.,.)$ is a distance metric function. In this study, we use the cosine similarity distance $d(x_1,x_2) = 1 - \frac{x_1 \bullet x_2}{\|x_1\|\|x_2\|}$, because for softmax classifiers, the angle can well represent the difference in feature vectors of different categories [37]. Due to lack of semantic information in the target domain, C_T^k is not obtained. The pseudo-labels are assigned to the samples from the target domain by the pre-trained model on the source domain. The calculation process of the semantic loss function L_S is shown in Algorithm 1.

To sum up, the object function of jointly updating the f_F and the f_C is formulated as:

$$L = L_C + w_f L_F + w_s L_S, \tag{13}$$

where w_f and w_s are the weights of L_F and L_S, respectively.

3.2.4. Theory analysis

In this section, based on the domain adaptation theory [23], the feasibility of our model is analyzed. The bound $\varepsilon_{\mathcal{T}}(h)$ relating the expected error on the target samples is represented as follows.

Theorem 1. *Given the source domain $\mathcal{S}$ and the target domain $\mathcal{T}$, assume that $\mathcal{H}$ is a hypothesis space. For any hypothesis $h \in \mathcal{H}$,*

$$\varepsilon_{\mathcal{T}}(h) \leqslant \varepsilon_{\mathcal{S}}(h) + d_1(\mathcal{D}_{\mathcal{T}}, \mathcal{D}_{\mathcal{S}}) + \min\{\varepsilon_{\mathcal{S}}(f_{\mathcal{S}}, f_{\mathcal{T}}), \varepsilon_{\mathcal{T}}(f_{\mathcal{S}}, f_{\mathcal{T}})\}, \tag{14}$$

where $\varepsilon_{\mathcal{S}}(h)$ is the expected error on the source samples, $d_1(\mathcal{D}_{\mathcal{S}}, \mathcal{D}_{\mathcal{T}})$ presents a discrepancy distance across the two domains $\mathcal{S}$ and $\mathcal{T}$, $f_{\mathcal{S}}$ and $f_{\mathcal{T}}$ are classification functions for the source and target domain respectively.

Proof. The first term $\varepsilon_{\mathcal{S}}(h)$ can be minimized easily with source label samples. $d_1(\mathcal{D}_{\mathcal{S}}, \mathcal{D}_{\mathcal{T}})$ is used to measure the distribution divergence of the two domains [22]. Recall that $\varepsilon_{\mathcal{T}}(h) = \varepsilon_{\mathcal{T}}(h, f_{\mathcal{T}})$ and $\varepsilon_{\mathcal{S}}(h) = \varepsilon_{\mathcal{S}}(h, f_{\mathcal{S}})$. According to the triangle inequality, for any classification functions f_1, f_2 and f_3, we have $\varepsilon(f_1, f_2) \leqslant \varepsilon(f_1, f_3) + \varepsilon(f_2, f_3)$. Then,

$$\begin{aligned}
\varepsilon_{\mathcal{T}}(h) &= \varepsilon_{\mathcal{T}}(h) + \varepsilon_{\mathcal{S}}(h) - \varepsilon_{\mathcal{S}}(h) + \varepsilon_{\mathcal{S}}(h, f_{\mathcal{T}}) - \varepsilon_{\mathcal{S}}(h, f_{\mathcal{T}}) && (15)\\
&= \varepsilon_{\mathcal{S}}(h) + \varepsilon_{\mathcal{T}}(h, f_{\mathcal{T}}) - \varepsilon_{\mathcal{S}}(h, f_{\mathcal{T}}) + \varepsilon_{\mathcal{S}}(h, f_{\mathcal{T}}) - \varepsilon_{\mathcal{S}}(h, f_{\mathcal{S}}) && (16)\\
&\leqslant \varepsilon_{\mathcal{S}}(h) + d_1(\mathcal{D}_{\mathcal{S}}, \mathcal{D}_{\mathcal{T}}) + \varepsilon_{\mathcal{S}}(h, f_{\mathcal{S}}) + \varepsilon_{\mathcal{S}}(f_{\mathcal{S}}, f_{\mathcal{T}}) - \varepsilon_{\mathcal{S}}(h, f_{\mathcal{S}}). && (17)\\
&\therefore \varepsilon_{\mathcal{T}}(h) \leqslant \varepsilon_{\mathcal{S}}(h) + d_1(\mathcal{D}_{\mathcal{S}}, \mathcal{D}_{\mathcal{T}}) + \varepsilon_{\mathcal{S}}(f_{\mathcal{S}}, f_{\mathcal{T}}). && (18)
\end{aligned}$$

In the same way, then,

$$\begin{aligned}
\varepsilon_{\mathcal{T}}(h) &= \varepsilon_{\mathcal{T}}(h) + \varepsilon_{\mathcal{S}}(h) - \varepsilon_{\mathcal{S}}(h) + \varepsilon_{\mathcal{T}}(h, f_{\mathcal{S}}) - \varepsilon_{\mathcal{T}}(h, f_{\mathcal{S}}) && (19)\\
&= \varepsilon_{\mathcal{S}}(h) + \varepsilon_{\mathcal{T}}(h, f_{\mathcal{T}}) - \varepsilon_{\mathcal{T}}(h, f_{\mathcal{S}}) + \varepsilon_{\mathcal{T}}(h, f_{\mathcal{S}}) - \varepsilon_{\mathcal{S}}(h, f_{\mathcal{S}}) && (20)\\
&\leqslant \varepsilon_{\mathcal{S}}(h) + \varepsilon_{\mathcal{T}}(h, f_{\mathcal{S}}) + \varepsilon_{\mathcal{T}}(f_{\mathcal{S}}, f_{\mathcal{T}}) - \varepsilon_{\mathcal{T}}(h, f_{\mathcal{S}}) + d_1(\mathcal{D}_{\mathcal{T}}, \mathcal{D}_{\mathcal{S}}). && (21)\\
&\therefore \varepsilon_{\mathcal{T}}(h) \leqslant \varepsilon_{\mathcal{S}}(h) + d_1(\mathcal{D}_{\mathcal{S}}, \mathcal{D}_{\mathcal{T}}) + \varepsilon_{\mathcal{T}}(f_{\mathcal{S}}, f_{\mathcal{T}}). && (22)
\end{aligned}$$

To sum up, then,

$$\varepsilon_{\mathcal{T}}(h) \leqslant \varepsilon_{\mathcal{S}}(h) + d_1(\mathcal{D}_{\mathcal{T}}, \mathcal{D}_{\mathcal{S}}) + \min\{\varepsilon_{\mathcal{S}}(f_{\mathcal{S}}, f_{\mathcal{T}}), \varepsilon_{\mathcal{T}}(f_{\mathcal{S}}, f_{\mathcal{T}})\}. \tag{23}$$

Unfortunately, without real target labels, the classification function $f_{\mathcal{T}}$ is not obtained directly. Therefore, the pseudo-labels are utilized to optimize. It is defined as $R = \min\{\varepsilon_{\mathcal{S}}(f_{\mathcal{S}}, f_{\mathcal{T}}), \varepsilon_{\mathcal{T}}(f_{\mathcal{S}}, f_{\mathcal{T}})\}$. It is showed that the USAFFN aims to optimize the upper bound for R. According to the triangle inequality, then

$$\begin{aligned}
R &= \min\{\varepsilon_{\mathcal{S}}(f_{\mathcal{S}}, f_{\mathcal{T}}), \varepsilon_{\mathcal{T}}(f_{\mathcal{S}}, f_{\mathcal{T}})\} && (24)\\
&\leqslant \min\left\{\varepsilon_{\mathcal{S}}(f_{\mathcal{S}}, h) + \varepsilon_{\mathcal{S}}(h, f_{\mathcal{T}}), \varepsilon_{\mathcal{T}}\left(f_{\mathcal{S}}, f_{\widetilde{\mathcal{T}}}\right) + \varepsilon_{\mathcal{T}}\left(f_{\widetilde{\mathcal{T}}}, f_{\mathcal{T}}\right)\right\} && (25)\\
&\leqslant \min\left\{\varepsilon_{\mathcal{S}}(f_{\mathcal{S}}, h) + \varepsilon_{\mathcal{S}}\left(h, f_{\widetilde{\mathcal{T}}}\right) + \varepsilon_{\mathcal{S}}\left(f_{\widetilde{\mathcal{T}}}, f_{\mathcal{T}}\right), \varepsilon_{\mathcal{T}}\left(f_{\mathcal{S}}, f_{\widetilde{\mathcal{T}}}\right) + \varepsilon_{\mathcal{T}}\left(f_{\widetilde{\mathcal{T}}}, f_{\mathcal{T}}\right)\right\}, && (26)
\end{aligned}$$

where $f_{\widetilde{\mathcal{T}}}$ is the pseudo target labeling function. Since source labels are obtained, then h is easily obtained in the hypothesis space to approximate the $f_{\mathcal{S}}$. Our method use the pre-trained model in the source domain to assign pseudo labeled samples in the target domain, then $\varepsilon_{\mathcal{S}}\left(h, f_{\widetilde{\mathcal{T}}}\right)$ and $\varepsilon_{\mathcal{S}}\left(f_{\widetilde{\mathcal{T}}}, f_{\mathcal{T}}\right)$ will be minimized. The $\varepsilon_{\mathcal{T}}\left(f_{\widetilde{\mathcal{T}}}, f_{\mathcal{T}}\right)$ denotes the discrepancy between the real target labeling function $f_{\mathcal{T}}$ and the pseudo target labeling function $f_{\widetilde{\mathcal{T}}}$. The discrepancy can be minimized when the target labeling function is trained on the target domain. Since $\varepsilon_{\mathcal{T}}\left(f_{\mathcal{S}}, f_{\widetilde{\mathcal{T}}}\right) = E_{x\sim\mathcal{T}}\left[l\left(f_{\mathcal{S}}(x), f_{\widetilde{\mathcal{T}}}(x)\right)\right]$, where $l(\cdot,\cdot)$ is typically the 0–1 loss function , the source labeling function $f_{\mathcal{S}}$ should return k for source samples in category k. If the feature representations of target samples in category k are similar to those of source samples in category k, the target samples can be predicted the same as that of the target pseudo labeling function. Consequently, $\varepsilon_{\mathcal{T}}\left(f_{\mathcal{S}}, f_{\widetilde{\mathcal{T}}}\right)$ is expected to be small.

In summary, the USAFFN is able to further minimize the R by aligning the global and the local centroid in the source domain and the pseudo labeled target domain. Therefore, utilizing the obtained pseudo labels of the target samples is effective for unsupervised domain adaptation.

4. Experiments

4.1. ECG databases

In this work, ECG signals are required from three databases of varying difficulty: the MITBIH arrhythmia database (ARDB) [38], the St. Petersburg Institute of Cardiological Technics (INCART) [39], and the MITBIH Long-Term ECG database (LTDB) [39]. They are available on Physionet [39]. These databases are briefly described as followings. For all these databases, they contain annotations for timing information and heartbeat category information verified by independent experts. The heartbeat labeling adopts the ANSI/AAMI EC57:1998 standard [40]. The standard developed by the Association for the Advancement of Medical Instruments (AAMI) recommends that the heartbeat types of ECG waveforms are clustered to five beat classes: N (normal or bundle branch block), S (supraventricular ectopic beat), V (ventricular ectopic beat), F (fusion beat) and Q (unassigned beat). This standard is referred to as the heartbeat type in this study. Table 1 shows that the heartbeat types are mapping to the AAMI heartbeat categories. This paper mainly focuses on four classes: N, V, S, F, while class Q is not considered. The details of the ECG signals from the three databases are shown in Table 2. Example signals for all experimental databases are shown in Fig. 4.

1) ARDB: This database, the most common benchmark, contains 48 records and each includes two-channel ECG signals for a 30-min duration. They are selected from 24-h recordings of 47 individuals. Continuous ECG signals are band-pass filtered at 0.1–100 Hz and then digitized at 360 Hz. According to the AAMI recommendation, 4 paced records (102, 104, 107, 217) are excluded in the ARDB.

2) INCART: This database consists of 75 annotated recordings extracted from 32 Holter records. Each record is 30 min long and contains 12 standard leads, and each sampled at 257 Hz. The reference annotation files contain over 175,000 beat annotations in all. These raw signals are collected from patients suffering from coronary artery disease (15 women and 17 men, aged 18–80; mean age: 58). They have no pacemakers, and most have ventricular ectopic beats.

3) LTDB: This database includes 7 long-term annotated ECG recordings extracted from Holter records. Each recording contains two-lead signals of approximately 14 to 22 h, and each sampled at 128 Hz. The heartbeat type annotations are reviewed manually.

In order to validate the effectiveness of our model, the performance is evaluated on the unlabeled target domain. Besides, several ablation studies are explored to verify the effectiveness of the MPAFF module and the semantic-aware adaptation module. Further, to verify the generalization ability, we evaluate the performance of the model in the INCART and LTDB.

To be specific, several experiments are conducted to better evaluate the performance of our proposed model for unsupervised adaptation classification across three different domain shifts. In all these experiments, the labels of the target domain are withheld. Table 3 briefly shows the three different domain adaptation tasks. For the domain adaptation task: DS1-DS2, the ARDB is divided into two datasets: DS1 and DS2 by convention [41]. Specifically, the dataset DS1 consists of 22 records (101, 106, 108, 109, 112, 114, 115, 116, 118,119, 122, 124, 201, 203, 205, 207, 208, 209, 215, 220, 223, 230) and the dataset DS2 includes the other 22 records (100, 103, 105, 111, 113, 117, 121, 123, 200, 202, 210, 212, 213, 214, 219, 221, 222, 228, 231, 232, 233, 234). This partition ensures that no heartbeats in the DS2 come from the same subject in the DS1, known as the inter-subject paradigm [42]. The labeled dataset DS1 is used as the source domain, and the unlabeled dataset DS2 is used

Table 1
Mapping the heartbeat types to the AAMI heartbeat categories.

AAMI	MIT-BIH	Heartbeat Types
N	N	Normal beat
	L	Left bundle branch block beat
	R	Right bundle branch block beat
	e	Atrial escape beat
	j	Nodal(Junctional) escape beat
S	A	Atrial premature beat
	a	Aberrated atrial premature beat
	J	Nodal(Junctional) premature beat
	S	Supraventricular premature or ectopic beat
V	V	Premature ventricular contraction
	E	Ventricular escape beat
F	F	Fusion of ventricular and normal beat
Q	/	Paced beat
	f	Fusion of paced and normal beat
	Q	Unclassifiable beat

Table 2
The database and its class representation in this work.

Database	The number of heartbeat types				Fs(Hz)	#Record
	N	V	S	F		
ARDB	89705	6982	2766	802	360	44
INCART	152348	19863	1944	219	257	75
LTDB	600110	64081	1499	2906	128	7

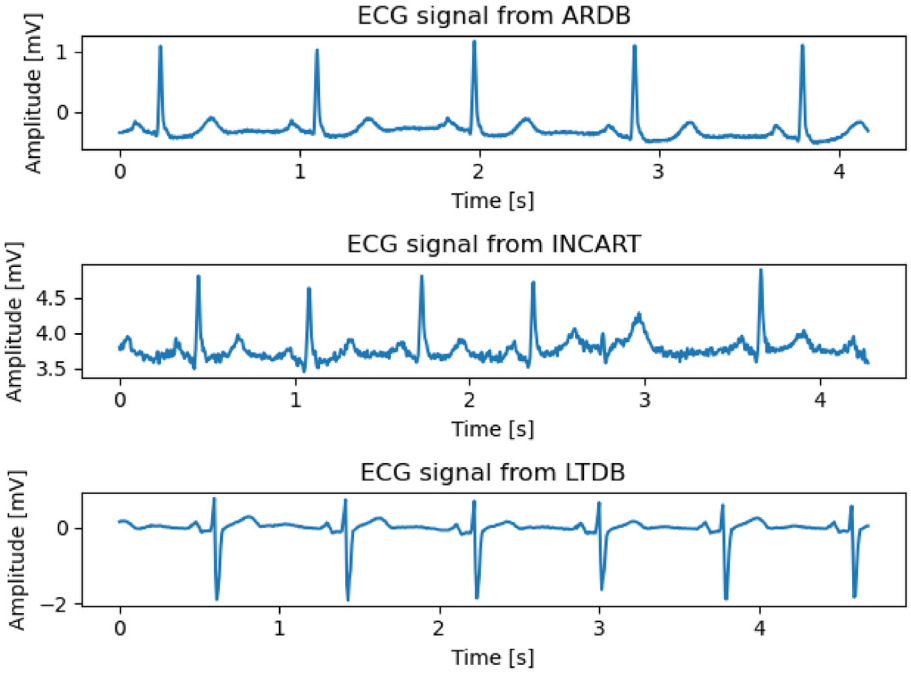


Fig. 4. Example ECG signals from ARDB, INCART, and LTDB.

Table 3
Different domain adaptation tasks in this study.

Domain adaptation tasks	Training data	Testing data
DS1-DS2	labeled dataset: DS1 unlabeled dataset: DS2	labeled dataset: DS2
ARDB-INCART	labeled dataset: ARDB unlabeled dataset: INCART	labeled dataset: INCART
ARDB-LTDB	labeled dataset: ARDB unlabeled dataset: LTDB	labeled dataset: LTDB

as the target domain to verify the effectiveness of the USAFFN. In this paper, the inter-subject paradigm is employed to compare the performance of our model with other algorithms on the ARDB. Note that for the samples in the target domain DS2, only the initial five minute unlabeled heartbeats are used for training. The brief distribution statistics of the two datasets DS1 and DS2 are shown in Table 4. In addition, the performance of our model is evaluated on the INCART and LTDB for verifying the generalization ability. It also presents the positive effect of our model across different domain shifts. Obviously, the discrepancies between ECG signals from different databases are much larger than those from different datasets in the same database. These discrepancies between signals from different databases are from individual characteristics, differences in collecting devices or conditions, and so on. More precisely, for domain adaptation task: ARDB-INCART, we refer to the labeled samples from the ARDB as the source domain and the unlabeled samples from the INCART or LTDB as the target domain. Since these different databases have different channels and sampling rates, the signals of the target domain select the same channel as the source domain data and re-sample to 360 Hz. For the adaptation task: ARDB-LTDB, due to the limited number

Table 4
The distribution statistics of the DS1 and DS2.

Datasets	The number of heartbeat types				Total
	N	V	S	F	
DS1	45820	3786	944	414	50964
DS2	43885	3196	1822	388	49291

of patients on the LTDB, 7000 recordings are randomly sampled from the class N, and the same recordings on the ARDB are extracted from other classes for training.

4.2. Evaluation metrics

In order to evaluate the performance of our model, the following four performance metrics are required: the overall classification accuracy (Acc), sensitivity (Sen), positive productivity (Ppr), the F1 score. TP, TN, FP, and FN represent true positives, true negatives, false positives, and false negatives respectively.

For the category k, define the following quantities as the parameters used in the scoring formulas,

$$\begin{aligned} \mathrm{TP}_k &= \mathrm{T}_{kk}, \\ \mathrm{TN}_k &= \left(\sum_k \mathrm{T}_{kk}\right) - \mathrm{T}_{kk}, \\ \mathrm{FP}_k &= \left(\sum_j \mathrm{T}_{ik}\right) - \mathrm{T}_{kk}, \\ \mathrm{FN}_k &= \left(\sum_i \mathrm{T}_{ik}\right) - \mathrm{T}_{kk}, \end{aligned} \tag{27}$$

where Acc is the ratio of correctly classified patterns to the total classified patterns, Sen is the rate of correctly classified events among all events, and Ppr is the rate of correctly classified events among all detected events. In addition, the F1 score is expressed as the harmonic mean of the Ppr and Sen, and it is used as a comprehensive criterion. They are defined as follows: $\mathrm{Sen} = (\mathrm{TP})/(\mathrm{TP} + \mathrm{FN})$, $\mathrm{Ppr} = (\mathrm{TP})/(\mathrm{TP} + \mathrm{FP})$, $\mathrm{F1} = (2 \times \mathrm{Sen} \times \mathrm{Ppr})/(\mathrm{Sen} + \mathrm{Ppr})$, $\mathrm{Acc} = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN})$.

4.3. Implementation details

Our method is trained with a PyTorch implementation and is performed in an Intel Corel I7 and a GeForce GTX 1080 Ti Graphics Cards. The Adam optimizer trains our proposed network with a mini-batch size of 256. We start the training stage with an initial learning rate of 0.01 and decay it by a factor of 10 for every 100 epochs. The weight hyper-parameters of the domain confusion loss function and the semantic alignment loss function are set to 0.01 and 0.1 respectively. The hyper-parameter α in the calculation of the semantic loss function is 0.1. If no performance improvement is observed in twenty iterations, we terminate the training process early. We randomly choose 90% of the samples in the source domain to optimize the model parameters, and then use the remaining 10% to select the optimal model. The maximum number of training iterations is set to 200 in the pre-trained process, and the maximum number of training iterations in the adaptation process is set to 300.

4.4. Results and discussions

To better evaluate the performance of our model, we conduct four experiments. Section 4.4.1 aims to validate the effectiveness of the USAFFN by comparing the performance of our model with the current methods on the ARDB. Section 4.4.2 is to verify the effectiveness of the MPAFF module and the semantic-aware adaptation module. Section 4.4.3 attempts to demonstrate the generalization ability of our method. Section 4.4.4 is to explore the impact of several hyperparameters on the model performance.

4.4.1. The effectiveness of our method

Comparison with alternative methods: Table 5 shows the performance of our method and the results obtained by several representative alternative algorithms [3–7,12,19,26,43–45] with the same inter-subject scheme on the DS2. These ECG classification algorithms are roughly divided into three categories: patient-independent, patient-specific, and domain-adaptation practice. The patient-independent ECG classification approaches [7,12,26,43–45] typically use population data to train general classifiers and then directly predict new patients. The patient-specific methods [3,6] introduce patient-adjustable models for new patients and usually require manual intervention. The domain-adaptation algorithms[5,4,19] aim to achieve excellent performance on the target domain by alleviating the distribution discrepancy between the feature

Table 5
Comparison of the performance of inter-subject arrhythmia detection methods on the DS2.

Method	N(%)			V(%)			S(%)			F(%)			Overall Acc(%)
	Sen	Ppr	F1	Sen	Ppr	F1	Sen	Ppr	F1	Sen	Ppr	F1	
Li et al. [26][1]	91.8	98.9	95.2	95.1	90.1	92.5	89.0	35.4	50.6	32.2	20.3	24.9	91.4
Shi et al. [12][1]	92.1	99.5	95.6	95.1	88.1	91.4	91.7	46.2	61.4	61.6	15.2	24.3	92.1
Sellami et al. [43][1]	88.5	98.8	93.3	92.0	72.1	80.8	82.0	30.4	44.3	68.3	26.5	38.1	95.3
Niu et al. [7][1]	98.9	97.4	98.1	85.7	94.1	89.7	76.5	76.6	76.6	0.2	0.7	0.3	96.4
Raj et al. [44][1]	91.0	99.3	94.9	89.5	85.4	87.4	80.9	50.9	62.4	93.5	12.6	22.2	89.9
Zhai et al. [45][1]	96.3	99.0	97.6	87.5	90.9	89.2	93.8	59.0	72.5	31.1	19.8	24.1	95.2
Ye et al. [3][2]	99.7	97.3	98.4	91.8	98.0	94.7	61.4	90.7	73.2	9.0	38.2	14.5	97.0
Zhai et al. [6][2]	99.6	98.5	99.0	93.8	92.4	93.0	76.8	74.0	75.3	79.5	62.3	69.8	96.1
Niu et al. [5][3]	93.9	97.4	95.6	85.1	57.8	68.8	76.6	73.2	74.8	38.4	44.9	41.3	92.3
Chen et al. [4][3]	95.3	98.4	96.8	95.7	93.8	94.7	78.5	88.6	83.2	43.8	8.8	14.6	94.3
Bazi et al. [19][3]	–	–	–	–	–	–	–	–	–	–	–	–	93.0
Our method[3]	97.0	98.7	97.9	87.6	91.3	89.4	86.0	67.6	75.7	64.9	32.3	43.1	95.7

[1] Patient-independent practice.
[2] Patient-specific practice.
[3] Domain-adaptation practice.

spaces of the two domains. Since patient-specific classifiers typically require additional expert annotations for each test subject, these methods are impractical in real-world applications. Although semi-supervised learning has been used to label new patient data in these works [3,6], it is still necessary to select data from each new patient's data as input, which undoubtedly increases the workload. In particular, when the number of new subjects is large enough, the workload is undesirable. Besides, these algorithms [5,4,19] based on domain adaptation technology have been widely developed in semi-supervised learning to improve the performance of heartbeat classification. As shown in Table 5, compared with these algorithms [5,4,19], the overall accuracy of our model achieves the state-of-the-art on the domain adaptation task DS1-DS2.

Specifically, Niu et al. [7] proposed a novel deep learning framework for inter-patient heartbeat classification based on a symbolization approach and a multi-perspective convolution neural network. The framework achieved a high classification accuracy (96.4%), but the F1 score of class F is low (0.3%). The symbolization approach jointly represented the morphology and rhythm of the heartbeat, normalized the signal, and alleviated the effect of inter-patient variation based on the baseline correction. However, the symbolic representations eliminate patient-specific representations that are meaningful for classification. Besides, the raw ECG signals need to be carefully preprocessed (beat alignment and baseline correction). Shi et al. [12] constructed two hierarchical weighted XGBoost classifiers, which undoubtedly increased the complexity. Li et al. [26] proposed an automatic heartbeat classifier based on DNN, which required a complicated process to calculate the beat-to-beat correlation. Compared with our method, the performance of category V in the articles [12,26] shows an advantage to some extent, but the overall performance is not satisfactory. Sellami et al. [43] proposed a convolutional neural network based on a batch-weighted loss function. Despite the overall accuracy was relatively outstanding (95.3%) in the work[43], the power of predicting category S was significantly insufficient, that is, the positive productive was 30.4 % and the F1 score was 44.3 %. Raj et al. [44] extracted features using the sparse decomposition technique, which required hand-crafted features and some professional clinical knowledge, achieving a high sensitivity (93.5%), but low positive productivity (12.6%) for class F. Zhai et al. [45] obtained comparable performance based on a semi-supervised learning system, including a normal beat estimation module and an iterative heartbeat label update algorithm. At the same time, the heartbeats need to be selected for new patients. As shown in Table 5, compared with these patient-independent approaches [7,12,26,43–45], our model achieves the competitive overall performance and the highest F1 score (43.1%) for class F.

In addition, since the patient-independent methods do not consider the differences between individuals of ECG signals, these patient-independent classifiers [7,12,26,43–45] generally perform less well than the patient-specific classifiers [3,6]. To be specific, Ye et al. [3] employed the consensus principle to model a general classifier and a specific classifier for automatic arrhythmia detection. Zhai et al. [6] adapted a patient-specific deep learning system for inter-patient heartbeat classification based on a convolution neural network. They transformed heartbeats into a dual beat coupling matrix and introduced a beat selection approach to select the most representative heartbeats for training. In addition, several labeled ECG data of the test patients were used for training. Therefore, several heartbeats of each new patient are required to annotate. Besides, it involves tedious data preprocessing, requiring the selection of representative heartbeats.

Recently, several articles [5,4,19] based on domain-adaptation practice have been introduced. Chen et al. [4] adopted an unsupervised domain adaptation approach to align category distributions between two domains by introducing two domain-regularizer loss functions. Niu et al. [5] employed an unsupervised adversarial domain adaptation network for arrhythmia detection by obtaining domain invariant representations. Bazi et al. [19] proposed a domain transfer SVM model, which matches two distributions measured with the maximum mean distribution, but requires manual intervention. These domain adaptation methods align the global movement between the two domains, but ignore the semantic information in samples, which may lead to negative transfer of features. The performance of class F in the article [4] has significant advantages over that of class F in the article [5]. These results indicate that the performance of the adversarial domain adaptation

model in the article [5] has some advantages for category F. In contrast, the results obtained by the two domain regularizer loss functions in the work [4] are more friendly to the classes N, V, and S. Bazi et al. [19] only described the overall accuracy. We add sensitivity, positive productivity, and F1 score to make a more comprehensive comparison with other recent references. In general, in terms of the performance of our model, the results of classes N, V, and S are competitive with several other alternative methods.

To sum up, our model shows the competitive performance among these methods. It avoids annotating and selecting ECG data for new subjects, and only requires collecting ECG data per se, which is relatively easy and part of the clinical routine. Our unsupervised semantic-aware domain adaptation framework is a valuable tool to deal with the annotation or detection of a large amount of unlabeled ECG data in clinical applications.

Overall classification performance: From Table 5, on the whole, our model is competitive on the DS2: the overall accuracy is 95.7%; the F1 scores for class V and S are 89.4% and 75.7%, respectively; Especially the F1 score for category F is 43.1%. The confusion matrix on the DS2 is shown in Table 6. To be specific, we draw the following conclusions: 1) For category N, most are misclassified as category S. The reason for this classification error may be that the morphological characteristics of heartbeat type N are similar to those of heartbeat type S. In the same way, for category S, the most significant proportion of those misclassified heartbeat types is category N. 2) For category F, most are misclassified as category N. It is worth noting that category F is not wrongly predicted as category S, and vice versa. Therefore, it can be inferred that there are significant discrepancies between the morphological characteristics of category S and F. 3) Inspired by the performance of each heartbeat type, the characteristic differences among these categories can be inferred.

We construct an unsupervised semantic domain adaptation framework based on adversarial learning with learning category-aware information of both domains, its essence still depends on data-driven to extract features to some extent. Given these limited training samples, it is not easy to train a perfect model for all heartbeat types. Overall, our method has obtained competitive performance for categories N, V, S and promising performance for category F on the DS2.

4.4.2. Ablation studies

In this section, in order to further verify the effectiveness of the MPAFF module and the semantic-aware adaptation module, we evaluate the influence of each module.

Effectiveness of branches in the MPAFF module: The number of branches (Num) in the MPAFF module is interpreted as different perspectives. In general, the more perspectives there are, the richer the extracted representations. However, too many perspectives easily lead to over-fitting and increase the complexity. To choose the best Num that can balance performance and efficiency, we set Num to 1, 2, 3, 4, respectively.

The experimental results are shown in Table 7. The results suggest that: 1) The overall accuracy improves as Num increases. While Num is 4, the overall accuracy appears degradation phenomenon. This means that more Num dose not always indicate more superior performance. To be specific, when Num is 3, the performance is relatively optimal. The reason for this may be that Num = 3 is the optimal trade-off between performance and efficiency. 2) As Num increases, the gains of categories N, V, S are not significant and remain in a small range of about 1% to 10%. However, the gains of category F are relatively significant, and the average range of gains varies well over 10%. The gains validate that the multi-perspective setting enhances the power of learning characteristics between the two domains.

Effectiveness of adaptive weights in the MPAFF module: In order to analyze the impact of adaptive in the MPAFF module, we conduct an additional experiment to investigate the performance by averaging the outputs of multiple branches, and compare it with the performance based on our adaptive weights.

The experimental results are presented in Table 8. From Table 8, we observe that the F1 score of categories F, V, and S are increased by 3.1%, 5.5%, and 18.5% respectively, in the adaptive weight setting. For category N, the overall performance is slightly improved.

The reason for these performance variations is that branches obtain rich characteristics from different perspectives. These characteristics are not equivalent and need to be weighted adaptively instead of giving the same weight. Comparing the experimental results in Table 8, it is confirmed that the adaptive weights have good adaptability.

Effectiveness of the semantic adaptation: To verify the effectiveness of the semantic adaptation, we evaluate the discriminant ability of the model. In addition, we investigate the performance of the adversarial adaptation algorithm, which eliminates the semantic alignment loss in the whole joint loss function. Further, we assess the performance of non-adaptation.

Table 6
The confusion matrix of our method on the DS2.

	Algorithm result			
Ground truth	N	V	S	F
N	42597	221	648	419
V	185	2800	102	109
S	220	34	1568	0
F	127	9	0	252

Table 7
Impact of the number of branches (Num) in the MPAFF module on the DS2.

Setting	N(%)			V(%)			S(%)			F(%)			Overall Acc(%)
	Sen	Ppr	F1	Sen	Ppr	F1	Sen	Ppr	F1	Sen	Ppr	F1	
Num = 1	94.9	97.4	96.1	87.3	90.6	88.9	74.4	71.9	73.1	20.2	5.1	8.2	93.0
Num = 2	96.8	97.4	97.1	85.7	83.4	84.5	66.7	55.9	60.8	43.2	77.3	55.4	94.5
Num = 3	97.0	98.7	97.9	87.6	91.3	89.4	86.0	67.6	75.7	64.9	32.3	43.1	95.7
Num = 4	96.3	97.1	96.7	82.3	85.9	84.1	72.9	51.3	60.2	12.0	39.4	18.4	93.8

Table 8
Impact of adaptive weights in the MPAFF module on the DS2.

Setting	N(%)			V(%)			S(%)			F(%)			Overall Acc(%)
	Sen	Ppr	F1	Sen	Ppr	F1	Sen	Ppr	F1	Sen	Ppr	F1	
w/o adaptive	96.0	97.4	96.7	81.8	86.2	83.9	60.7	54.1	57.2	68.8	28.2	40.0	93.5
w adaptive	97.0	98.7	97.9	87.6	91.3	89.4	86.0	67.6	75.7	64.9	32.3	43.1	95.7

The performance of the three settings (non-adaptation, adversarial adaptation, our adaptation) on the DS2 is presented in Fig. 5. Further, The performance suggests that: 1) In the setting of non-adaptation, in order to obtain a relatively high overall accuracy, the model lays emphasis on the discrimination of classes N and V, and relatively ignores the performance of classes S and F. 2) After the adversarial adaptation, the classification performance of each category and the overall accuracy are improved. Namely, the classical adversarial adaptation method mitigates the degradation caused by the discrepancy of data distributions. 3) After our adaptation, our model significantly improved the performance of classes S and F between the two domains. Specifically, the gains of the F1 scores for classes F and S are 8.1% and 5.6% respectively. The experimental results confirm that our method significantly alleviates the performance degradation caused by the discrepancy of data distribution. Further, in order to intuitively illustrate the positive effect of the adaptation, we visualize the representations extracted by the feature extractor by using the t-distributed stochastic neighbor embedding (t-SNE) [46]. The visual diagrams present the discrepancies between ECG signals from the DS1 and DS2, as shown in Fig. 6. From Fig. 6, we draw the following conclusions: 1) From (a), it is evident that there are discrepancies in the feature space of the two datasets DS1 and DS2 due to individual differences. As a result, the performance of the pre-trained model on the DS1 decreases on the DS2. 2) From (b) →(c), obviously, our model alleviates the misalignment of categories between the two domains. 3) From (c), the feature spaces of the two domains are aligned so that the pre-trained heartbeat classifier on the source domain performs well on the target domain. To sum up, we confirm the effectiveness of our semantic adaptation model.

4.4.3. The generalization of our method

To further verify the generalization of our method, we evaluate the performance of our method across different domain shifts. More precisely, two domain adaptation tasks are constructed: ARDB-INCART and ARDB-LTDB, as shown in Table 3. Further, in order to intuitively illustrate the positive effect of the adaptation process, we visualize the deep representations extracted by the feature extractor by using the t-distributed stochastic neighbor embedding (t-SNE) [46].

The overall accuracy of different unsupervised adaptation methods between ARDB, INCART, and LTDB is shown in Table 9. The classification performance of the non-adaptation, adversarial adaptation, and our model on the INCART and the LTDB is presented in Figs. 7 and 8, respectively. Figs. 9 and 10 present the discrepancies in the feature spaces of the ECG signals from the two domain adaptation tasks: ARDB-INCART and ARDB-LTDB, respectively. Compared with the domain adaptation task: DS1-DS2, the distribution shift of other domain adaptation tasks is significantly larger. From Table 9, the following conclusions are drawn: our method greatly improves the classification accuracy of the three domain adaptation tasks. In general, for domain adaptation tasks: ARDB-INCART and ARDB-LTDB, the overall accuracy of all classes improves significantly from 78.5% to 96.7% and from 52.2% to 90.8%, respectively. Specifically, as shown in Figs. 7 and 8, the darker the color, the less misclassification. Clearly, the performance of the non-adaptation model performs poor. The positive effect of the adversarial adaptation model and our adaptation model is illustrated by observing the darker colors. Further, our adaptation model performs better than that of the adversarial adaptation model. However, not all classes improve. The reasons for this may be twofold. Firstly, the deep neural network is difficult to achieve high accuracy with limited samples; Secondly, several classes do no have correctly labeled target signals before adaptation, and the adaptation is unable to recover the performance of these classes.

From Figs. 6, 9 and 10, we draw the following conclusions: 1) Comparing Figs. 9(a), 10(a) with Fig. 6(a), it is confirmed that the discrepancies of the ECG feature spaces in different databases are much more significant than those in different datasets of the same database. There is consistent with the performance of the non-adaptation method, as shown in Table 9. The greater the discrepancy in the feature spaces between the two domains, the lower the performance of the non-adaptation model. 2) From (b) →(c) of Figs. 9 and 10 respectively, our adaptation model obviously alleviates the misalignment of category distributions between the two domains to some extent. 3) From (c) in Figs. 6, 9 and 10, the representation

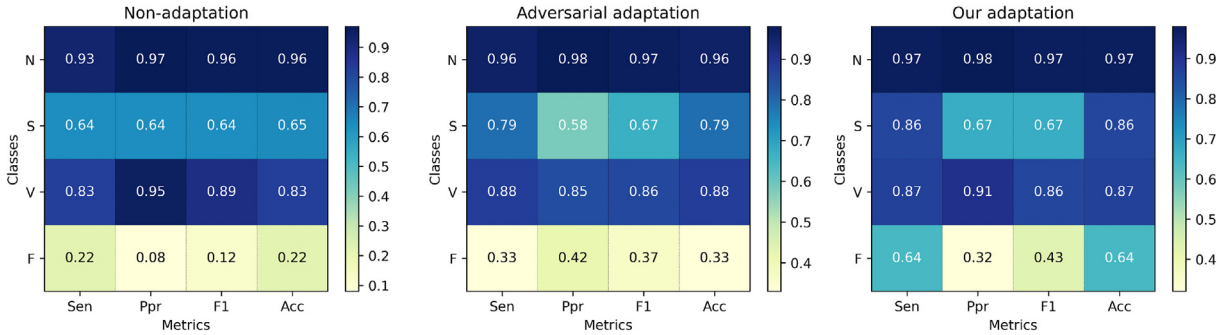


Fig. 5. Domain adaptation task: DS1 → DS2. The classification performance of the non-adaptation, adversarial adaptation, and our adaptation model on the DS2. We observe that our unsupervised adaptation algorithm performs better.

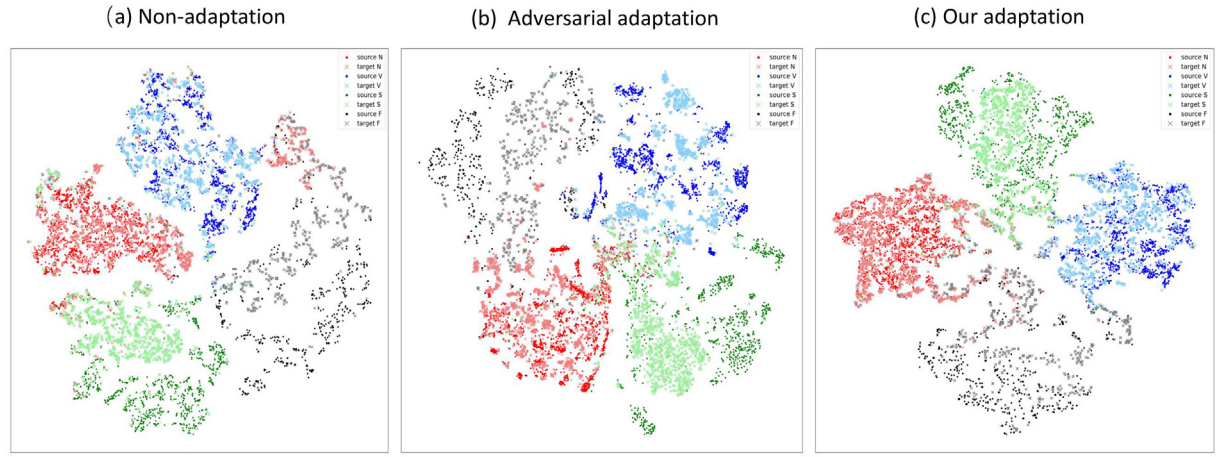


Fig. 6. Domain adaptation task: DS1 → DS2. We confirm the effects of our algorithm on the DS2 via using the t-SNE [46] to visual the feature representations. Red, blue, green and black represent N, V, S and F respectively. The symbol • represents the source domain sample, and the symbol × denotes the target domain sample, as shown in the chart.

Table 9
The overall accuracy of various methods in different domain adaptation tasks.

Method	DS1-DS2	ARDB-INCART	ARDB-LTDB
Non-adaptation	93.8%	78.5%	52.2%
Adversarial adaptation	94.8%	92.9%	79.5%
Our adaptation	95.7%	96.7%	90.8%

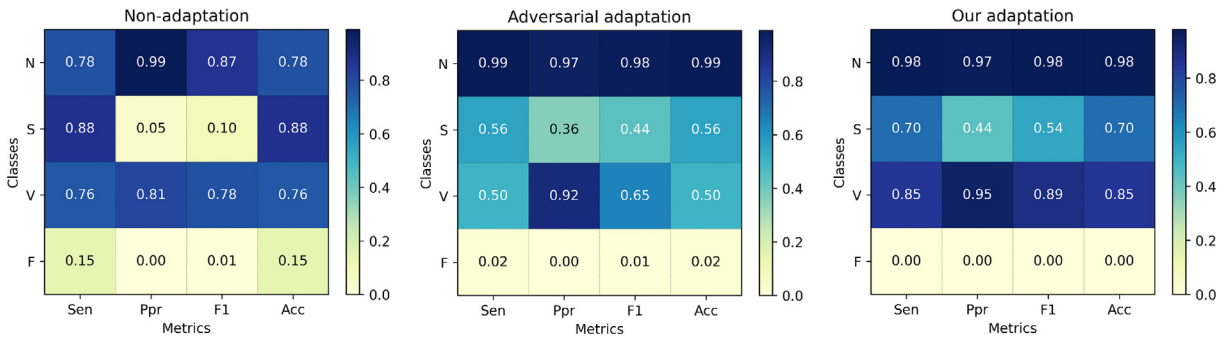


Fig. 7. Domain adaptation task: ARDB → INCART. The classification performance of the non-adaptation, adversarial adaptation, and our adaptation model on the INCART. We observe that our unsupervised adaptation algorithm performs better.

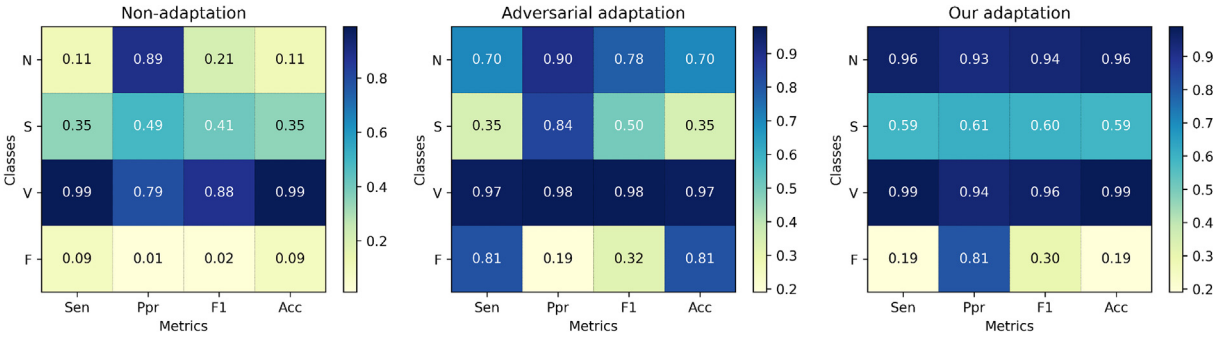


Fig. 8. Domain adaptation task: ARDB → LTDB. The classification performance of the non-adaptation, adversarial adaptation, and our adaptation model on the LTDB. We observe that our unsupervised adaptation algorithm performs better.

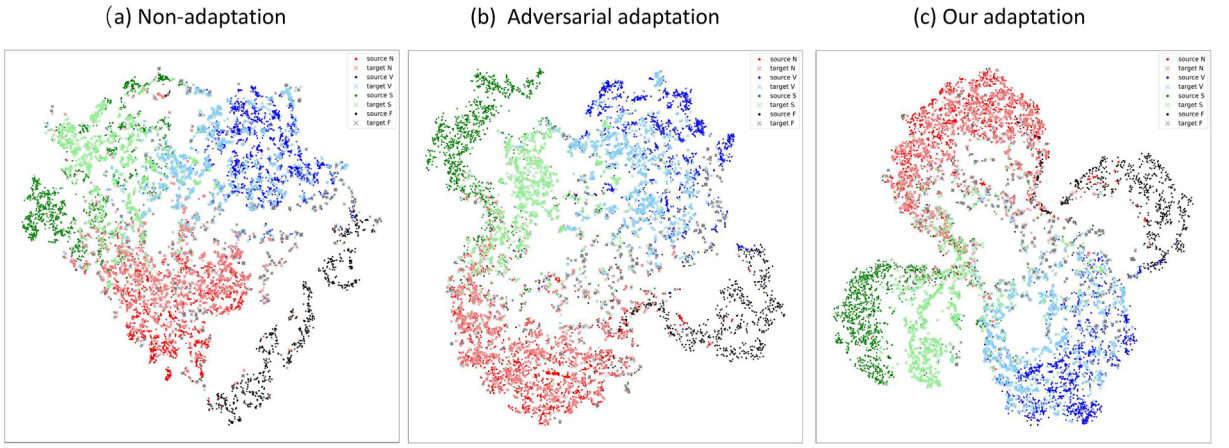


Fig. 9. Domain adaptation task: ARDB → INCART. We confirm the generalization of our model on the INCART via using the t-SNE to visual the feature representations. Red, blue, green and black represent N, V, S and F respectively. The symbol • and × represent samples from source and target domain, respectively.

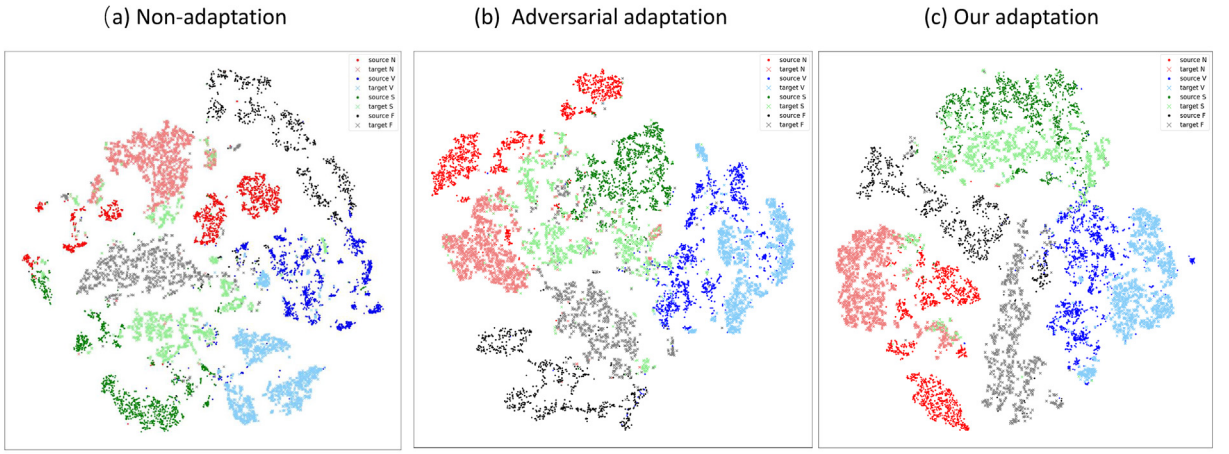


Fig. 10. Domain adaptation task: ARDB → LTDB. We confirm the generalization of our model on the LTDB via using the t-SNE to visual the obtained feature representations. Red, blue, green and black represent N, V, S and F respectively. The symbol • and × represent samples from source and target domain, respectively.

spaces of the two domains in different domain adaptation tasks are aligned, so that the pre-trained heartbeat classifier on the source domain performs well on the target domain. In conclusion, we confirm the generalization of our model.

4.4.4. The hyperparameters of our method

The hyper-parameters of the model are the parameters that control the training process. They determine the properties of the constructed specific domain adaptation task. Therefore, they have an indirect impact on the features extracted by the network. In this study, the hyperparameters, including batch size, the number of iterations, learning rate decay, the weights of the semantic alignment loss function and the domain confusion loss function, are discussed in the domain adaptation task: DS1-DS2.

Batch size: To analyze the impact of batch size on the performance, Fig. 11 shows how the overall accuracy of our model changes with the batch size. With the increasing batch size, the overall accuracy of the model does not change significantly. Only when the batch size is 512, the overall accuracy of the model decreases significantly. Therefore, it is inferred that the effect of batch size on the model is relatively limited. At the same time, the batch size of 256 is a suitable choice for the overall accuracy and avoiding over-fitting.

The number of iterations and learning rate decay: To analyze the impact of whether with the learning rate decay for training the model, Fig. 12 shows the variation trend of train loss with or without learning rate decay as the number of iterations increases in the pre-train stage and the domain adaptation stage, respectively. From Fig. 12, we observe that compared with the model without learning rate decay, the learning rate with decay helps to train the model relatively fast and stable.

Weight factors: Fig. 13 shows the influence of different combinations of weight factors in the loss function on the overall performance. When the weight factor w_f of the model fusion loss function and the weight factor w_s of the model semantic loss function are set to $\{w_f = 0.001, w_s = 0.001\}, \{w_f = 0.001, w_s = 0.01\}, \{w_f = 0.001, w_s = 0.1\}, \{w_f = 0.01, w_s = 0.001\}$, $\{w_f = 0.01, w_s = 0.01\}, \{w_f = 0.01, w_s = 0.1\}, \{w_f = 0.1, w_s = 0.001\}, \{w_f = 0.1, w_s = 0.01\}$, and $\{w_f = 0.1, w_s = 0.1\}$, we evaluate the overall accuracy of our model. As shown in Fig. 13, when the weight factors are set to $\{w_f = 0.01, w_s = 0.1\}$,

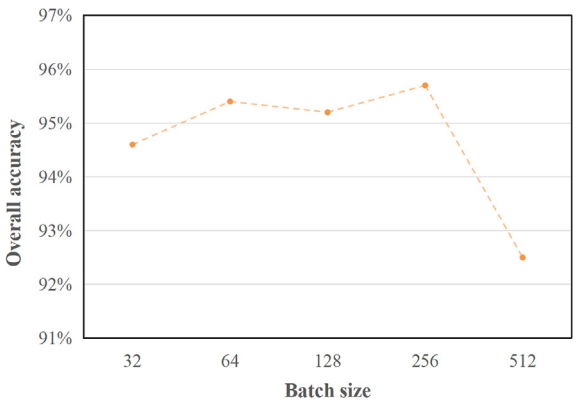


Fig. 11. The effect of batch size on the classification performance.

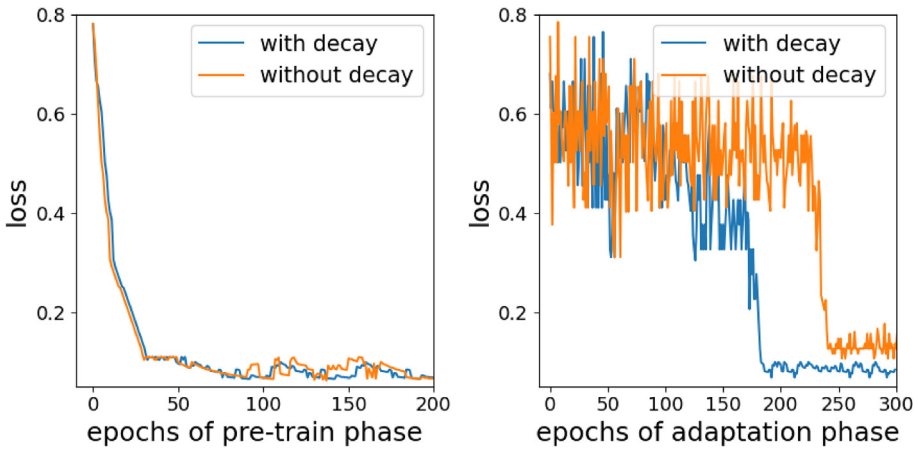


Fig. 12. The effect of learning rate decay on the train loss in the pre-training and adaptation phase.

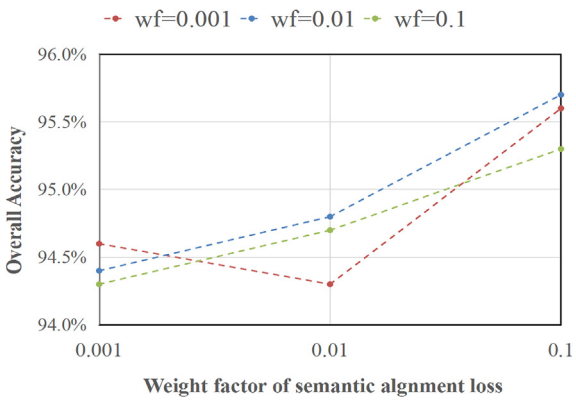


Fig. 13. The effect of weight factors w_f and w_s on the classification performance.

the overall accuracy of our model is the highest. When $w_f = 0.01$, the overall accuracy of our model fluctuates slightly, and the overall accuracy is relatively high compared with other options of 0.1 or 0.001.

5. Conclusion

This paper presents an unsupervised semantic-aware adaptive feature fusion network for inter-subject arrhythmia detection. We view the discrepancies of ECG signals from diverse individuals as domain shifts. Prior unsupervised adversarial adaptation methods in the context of arrhythmia detection ignore the multi-level pattern characteristics and semantic information of ECG signals. Our proposed method considers the multi-perspective representations and semantic-aware information of ECG signals from the two domains. To be specific, we employ a multi-perspective adaptive weight feature fusion module for better extracting deep representations from the two domains and incorporate the category information between two domains in the process of domain adaptation. Our method achieves satisfactory performance and has a great reference advantage in diagnosing ECG signals from diverse sources. Besides, we confirm the generalization capability of our model on the benchmarks ARDB, INCART, and LTDB. Further, we consider some influential factors of the model such as batch size, weight factors, and so on. The experimental results show that the hyperparameters of the model results are suitable.

As for future work, the concept of our unsupervised semantic-aware adaptation can be generalized to other meaningful medical scenarios, such as computer-aided obstructive sleep apnea detection and intelligent annotation. Another potential application is anomaly detection in industry. Since the data distribution may change across different equipment and time, the concept of our approach enables efficient anomaly detection with only unlabeled data that can be collected easily during the working period. Besides, lightweight deep learning models for mobile or other embedded applications have a far-reaching impact on applications. In the future, efforts will be made to explore the application of domain adaptation technology in the lightweight medical field.

CRediT authorship contribution statement

Panpan Feng: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Resources, Writing - original & draft, Visualization. **Jie Fu:** Conceptualization, Methodology, Writing - review & editing, Visualization. **Zhaoyang Ge:** Visualization, Investigation, Resources. **Haiyan Wang:** Investigation, Resources. **Yanjie Zhou:** Conceptualization, Formal analysis, Writing - review & editing, Project administration. **Bing Zhou:** Funding acquisition. **Zongmin Wang:** Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is supported by The Key Research, Development, and Dissemination Program of Henan Province (Science and Technology for the People) under Grant No. 182207310002.

References

[1] Y. Hagiwara, H. Fujita, S.L. Oh, J.H. Tan, R. San Tan, E.J. Ciaccio, U.R. Acharya, Computer-aided diagnosis of atrial fibrillation based on ECG signals: A review, Inf. Sci. 467 (2018) 99–114.
[2] Y.H. Hu, S. Palreddy, W.J. Tompkins, A patient-adaptable ECG beat classifier using a mixture of experts approach, IEEE Trans. Biomed. Eng. 44 (9) (1997) 891–900.
[3] C. Ye, B.V. Kumar, M.T. Coimbra, An automatic subject-adaptable heartbeat classifier based on multi-view learning, IEEE J. Biomed. Health Inform. 20 (6) (2015) 1482–1492.
[4] M. Chen, G. Wang, Z. Ding, J. Li, H. Yang, Unsupervised domain adaptation for ECG arrhythmia classification, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2020, pp. 304–307.
[5] L. Niu, C. Chen, H. Liu, S. Zhou, M. Shu, A deep-learning approach to ECG classification based on adversarial domain adaptation, Healthcare 8 (4) (2020) 437.
[6] X. Zhai, C. Tin, Automated ECG classification using dual heartbeat coupling based on convolutional neural network, IEEE Access 6 (2018) 27465–27472.
[7] J. Niu, Y. Tang, Z. Sun, W. Zhang, Inter-patient ECG classification with symbolic representations and multi-perspective convolutional neural networks, IEEE J. Biomed. Health Inform. 24 (5) (2019) 1321–1332.
[8] S. Mousavi, F. Afghah, Inter-and intra-patient ECG heartbeat classification for arrhythmia detection: a sequence to sequence deep learning approach, in: Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, 2019, pp. 1308–1312.
[9] C. Ye, B.V. Kumar, M.T. Coimbra, Heartbeat classification using morphological and dynamic features of ECG signals, IEEE Trans. Biomed. Eng. 59 (10) (2012) 2930–2941.
[10] T. Golany, K. Radinsky, Pgans: Personalized generative adversarial networks for ECG synthesis to improve patient-specific deep ECG classification, in: Proceedings of the AAAI Conference on Artificial Intelligence, 2019, pp. 557–564..
[11] H. Wang, Y. Zhou, B. Zhou, X. Niu, H. Zhang, Z. Wang, Interactive ECG annotation: An artificial intelligence method for smart ECG manipulation, Inf. Sci. (2021)..
[12] H. Shi, H. Wang, Y. Huang, L. Zhao, C. Qin, C. Liu, A hierarchical method based on weighted extreme gradient boosting in ECG heartbeat classification, Comput. Methods Programs Biomed. 171 (2019) 1–10.
[13] P. Yang, D. Wang, W.-B. Zhao, L.-H. Fu, J.-L. Du, H. Su, Ensemble of kernel extreme learning machine based random forest classifiers for automatic heartbeat classification, Biomed. Signal Process. Control. 63 (2021) 102–138.
[14] H. Wang, H. Dai, Y. Zhou, B. Zhou, P. Lu, H. Zhang, Z. Wang, An effective feature extraction method based on GDS for atrial fibrillation detection, J. Biomed. Inform. 119 (2021) 103819.
[15] V. Mondéjar-Guerra, J. Novo, J. Rouco, M.G. Penedo, M. Ortega, Heartbeat classification fusing temporal and morphological information of ECGs via ensemble of classifiers, Biomed. Signal Process. Control. 47 (2019) 41–48.
[16] C. Zhang, Q. Zhao, Deep Discriminative Domain Adaptation, Inf. Sci. 575 (2021) 599–610.
[17] C. Chen, Q. Dou, H. Chen, P.-A. Heng, Semantic-aware generative adversarial nets for unsupervised domain adaptation in chest x-ray segmentation, in: Proceedings of the International Workshop on Machine Learning in Medical Imaging, 2018, pp. 143–151.
[18] J.-Y. Zhu, T. Park, P. Isola, A.A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: Proceedings of the IEEE International Conference on Computer Vision, 2017, pp. 2223–2232.
[19] Y. Bazi, N. Alajlan, H. AlHichri, S. Malek, Domain adaptation methods for ECG classification, in: Proceedings of the International Conference on Computer Medical Applications, 2013, pp. 1–4.
[20] Y. Jin, C. Qin, J. Liu, K. Lin, H. Shi, Y. Huang, C. Liu, A novel domain adaptive residual network for automatic atrial fibrillation detection, Knowl.-Based Syst. 203 (2020) 106122.
[21] N. Ammour, Atrial fibrillation detection with a domain adaptation neural network approach, in: Proceedings of the International Conference on Computational Science and Computational Intelligence, 2018, pp. 738–743.
[22] Y. Ganin, V. Lempitsky, Unsupervised domain adaptation by backpropagation, in: Proceedings of the International Conference on Machine Learning, 2015, pp. 1180–1189.
[23] S. Ben-David, J. Blitzer, K. Crammer, A. Kulesza, F. Pereira, J.W. Vaughan, A theory of learning from different domains, MLJ 79 (1–2) (2010) 151–175.
[24] C. Zhang, Q. Zhao, Y. Wang, Transferable attention networks for adversarial domain adaptation, Inf. Sci. 539 (2020) 422–433.
[25] B. Liu, X. Chen, Y. Xiao, W. Li, L. Liu, C. Liu, An efficient dictionary-based multi-view learning method, Inf. Sci. 576 (2021) 157–172.
[26] F. Li, Y. Xu, Z. Chen, Z. Liu, Automated heartbeat classification using 3-D inputs based on convolutional neural network with multi-fields of view, IEEE Access 7 (2019) 76295–76304.
[27] Y. Huang, H. Li, X. Yu, A multiview feature fusion model for heartbeat classification, Physiol. Meas..
[28] J. Pan, W.J. Tompkins, A real-time QRS detection algorithm, IEEE Trans. Biomed. Eng. 32 (3) (1985) 230–236.
[29] PhysioNet, Physionet community, 2000..
[30] A. Peimankar, S. Puthusserypady, DENS-ECG: A deep learning approach for ECG signal delineation, Expert Syst. Appl. 165 (2021) 113911.
[31] A. Fotoohinasab, T. Hocking, F. Afghah, A graph-constrained changepoint detection approach for ECG segmentation, in: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2020, pp. 332–336.
[32] A. Fotoohinasab, T. Hocking, F. Afghah, A greedy graph search algorithm based on changepoint analysis for automatic QRS complex detection, Comput. Biol. Med. 130 (2021) 104208.
[33] P.S. Hamilton, W.J. Tompkins, Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database, IEEE Trans. Biomed. Eng. 33 (1986) 1157–1165.
[34] S. Mousavi, F. Afghah, F. Khadem, U.R. Acharya, ECG language processing (ELP): A new technique to analyze ECG signals, Comput. Methods Programs Biomed. 202 (2021) 105959.
[35] H. Huang, J. Liu, Q. Zhu, R. Wang, G. Hu, A new hierarchical method for inter-patient heartbeat classification using random projections and RR intervals, Biomed. Eng. Online 13 (2014) 1–26.
[36] J. Hu, L. Shen, G. Sun, Squeeze-and-excitation networks, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018, pp. 7132–7141.
[37] W. Liu, Y. Wen, Z. Yu, M. Yang, Large-margin softmax loss for convolutional neural networks, in: Proceedings of the International Conference on Machine Learning, 2016, pp. 507–516.
[38] G.B. Moody, R.G. Mark, The impact of the MIT-BIH arrhythmia database, IEEE Eng. Med. Biol. Mag. 20 (3) (2001) 45–50.
[39] A.L. Goldberger, L.A. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, Physiobank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (23) (2000) e215–e220.
[40] ANSI/AAMI EC57, Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms, 1998..
[41] P. De Chazal, M. O'Dwyer, R.B. Reilly, Automatic classification of heartbeats using ECG morphology and heartbeat interval features, IEEE Trans. Biomed. Eng. 53 (12) (2006) 2535–2543.
[42] G. De Lannoy, D. François, J. Delbeke, M. Verleysen, Weighted conditional random fields for supervised interpatient heartbeat classification, IEEE Trans. Biomed. Eng. 59 (1) (2012) 241–247.
[43] A. Sellami, H. Hwang, A robust deep convolutional neural network with batch-weighted loss for heartbeat classification, Expert Syst. Appl. 122 (2019) 75–84.
[44] S. Raj, K.C. Ray, Sparse representation of ECG signals for automated recognition of cardiac arrhythmias, Expert Syst. Appl. 105 (2018) 49–64.
[45] X. Zhai, Z. Zhou, C. Tin, Semi-supervised learning for ECG classification without patient-specific labeled data, Expert Syst. Appl. 158 (2020) 113411.
[46] V.D.M. Laurens, G. Hinton, Visualizing data using t-SNE, J. Mach. Learn. Res. 9 (2008) 2579–2605.