2022 IEEE International Conference on Bioinformatics and Biomedicine (BIBM) | 978-1-6654-6819-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/BIBM55620.2022.9995362

AdapSQA: Adaptive ECG Signal Quality Assessment Model for Inter-Patient Paradigm using Unsupervised Domain Adaptation

Hui Li†, Yu Zhang*‡, Jiyang Han† Yu Yan‡, Yu Liu§,and Hui Yang*†
*Key laboratory of Flexible Electronics of Zhejiang Provience,
Ningbo Institute of Northwestern Polytechnical University, Ningbo, China
†School of Life Science, Northwestern Polytechnical University, Xi'an, China
‡School of Computer Science, Northwestern Polytechnical University, Xi'an, China
§Department of Cardiology, Nanjing University Medical School Affiliated
Nanjing Drum Tower Hospital, Nanjing, China
Corresponding Author: Yu Zhang Email: zhangyu@nwpu.edu.cn

***Abstract*—Signal quality assessment (SQA) is an important topic in the field of wearable electrocardiogram (ECG) monitoring. Existing ECG SQA models focus on the intra-patient paradigm where the training and testing data are from same individuals. However, due to the individual differences in ECG morphology, features extracted from the the training patients may not be applicable to the new patient. Therefore, these models may suffer severe performance degradation in the inter-patient paradigm which is closer to the reality. In this paper, we propose a novel adaptive ECG SQA model called AdapSQA for the inter-patient paradigm using unsupervised domain adaptation in order to enhance its feature extraction adaptability to the new patient. To realize our AdapSQA, a lightweight baseline model for ECG SQA is first built for better feature extraction in wearable systems. Then, a domain adaptation layer is introduced to align the feature distribution of the training patients and the the new patient by minimizing the distance between the two domains. In this way, a baseline model can be adaptive to a new patient without extra annotation. To evaluate the proposed model, a patient-specific ECG Noise Dataset was generated based on the public datasets since there is no public open source of interest. Experimental results demonstrate that our proposed AdapSQA outperforms state-of-the-art approaches in term of the average inter-patient accuracy to 93.67% with a smaller standard deviation of 4.41%, and is able to achieve lightweight deployment for wearable systems.**

***Index Terms*—ECG, signal quality assessment, inter-patient, unsupervised domain adaptation**

I. Introduction

Fast-developing wearable electrocardiogram (ECG) brings the problem of poor signal quality due to the complex living environment, leading to frequent errors in signal analysis. Therefore, it is essential to perform signal quality assessment (SQA) before final analysis in wearable ECG monitoring systems [1]. By discarding unacceptable ECG data, SQA can help improve the performance of the final analysis as well as reduce transmission power in wearable monitoring systems [2, 3].

Various signal quality index [4, 5] have been proposed for ECG signal, meanwhile, machine learning methods [6, 7] and deep learning methods [8, 9] have been used to improve the assessment performance. Existing methods have achieved good performance in the intra-patient paradigm (i.e., training data and testing data are from the same individuals). However, these models may suffer severe performance degradation in the inter-patient paradigm where training data and testing data are from different individuals. In real application scenario, a well-trained ECG SQA model is often used to assess new different patients. Due to huge individual differences in ECG morphology, features extracted from the training patients may not be applicable to the new patient. Our motivation experiments in Section III show that a well-trained model may suffer up to 50% reduction in term of accuracy as for a new patient. Therefore, it is essential to extract effective features for new patient to enhance performance in inter-patient paradigm of SQA.

A natural way is to apply supervised training for each new patient [10]. However, annotating ECG records requires strong expertise and is labor-intensive. In contrast, unsupervised domain adaptation is preferred since it can help extract domain-invariant features by reducing the distance between the labeled source domain and the unlabeled target domain in feature space without extra annotation [11, 12].

In this paper, a novel adaptive ECG SQA model (AdapSQA) is proposed using unsupervised domain adaptation to enhance its feature extraction adaptability to the new patient without extra annotation. The AdapSQA can be obtained through the pre-training stage and patient adaptation stage. A lightweight model for ECG SQA is first built as the baseline in the pre-training stage for better feature extraction with less computational costs. In the patient adaptation stage, a domain adaptation layer is introduced into the network to align the feature distribution of the training patients and the new patient

This work is supported by Ningbo Natural Science Foundation (No. 202003N4057), the Key Research and Development Program of Jiangsu, China (BE2020665), the National Natural Science Foundation of China (Nos. 62172336 and 62032018), and Nanjing Medical Science and Technology Development Fund (ZKX21014)

by minimizing the distances between the two domains in the feature space. In this way, the baseline model can be adaptive to the new patient to realize AdapSQA. In addition, to evaluate the proposed model, a patient-specific ECG Noise Dataset was generated based on the two famous datasets since there is no public open source of interest. Experimental results show that our proposed AdapSQA outperforms state-of-the-art approaches in term of the average inter-patient accuracy to 93.67% with a smaller standard deviation of 4.41%, and can be easily deployed on wearable devices.

The remainder of this paper is organized as follows. Section II describes the materials; Section III presents the motivation experiments. Section IV illustrates the details of our model. Section V shows the experiment results and analysis. Section VI concludes the paper.

II. MATERIALS

A. Dataset Generation

To study the inter-patient paradigm of ECG SQA, long-term noise ECG collected from different patients is required. Since there is no open source of interest, we generated a patient-specific ECG Noise Dataset by adding real noise data to the clean ECG signal according to the Association for the Advancement of Medical Instrumentation (AAMI) convention.

The clean ECG signal are taken from the MIT-BIH Arrhythmia Database [13]. It contains 48 30-minute ambulatory ECG recordings from 47 patients. 28 records labeled clean or noisy in very limited periods ($\leq$ 10 seconds) during the entire recording of each record in channel modified limb lead II from different patients are chosen in this paper. The noise data were taken from the MIT-BIH Noise Stress Test Database (NSTDB) [14]. It provides recordings of 3 representative noise signals which are typically found in ambulatory ECG recordings and include electrode motion artefact (em), baseline wander (bw), and muscle artefact (ma).

The detailed implementation of dataset generation can be described as follows. Clean ECG signal is first divided into 10s segments with 8s overlapping for data augmentation. To contaminate a clean segment, a noise segment of equal length is randomly selected from the noise signal with the given type. Followed by the guidelines of NSTDB, the noise segment is scaled according to the given SNR and then added to the clean segment according to (1) and (2). By contaminating half segments of each record, a patient-specific ECG Noise Dataset is generated. In order to make the dataset closer to reality, the noise type is randomly selected from em, ma, and bw. According to [15], the SNR is randomly selected from the range (-6,6) for em and ma and the range (-12,0) for bw.

$$y = x + a * v \tag{1}$$

$$a = \sqrt{exp(\frac{-ln(10) * S}{10})\frac{P_x}{P_v}} \tag{2}$$

B. Dataset Splitting

28 records are split into two subsets without intersection: DS1 (101, 106, 109, 114, 115, 119, 122, 124, 205, 209, 215, 220, 230) and DS2 (100, 103, 113, 117, 202, 213, 214, 219, 221, 231, 233, 234). DS1 is considered as the source data (training patient data), and each record in DS2 represents the target data (new patient data).

III. MOTIVATION EXPERIMENTS

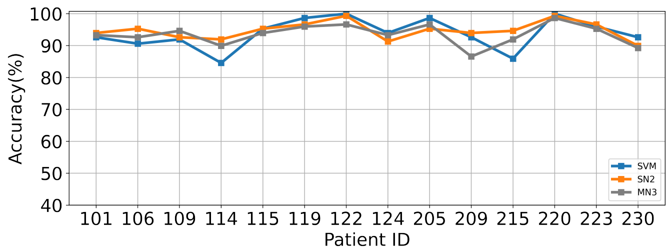
In this section, we present motivation experiments on the assessment performance of ECG SQA in the inter-patient paradigm. In physiological signal analysis, such as arrhythmia classification and emotion recognition, it has long been a problem that a well-trained model may suffer performance degradation when applied to a new patient due to huge changes in signal morphology between individuals [16], [17]. However, this problem has not been studied in physiological SQA fields.

An ECG SQA model is trained in a supervised way and tested on data from each patient in the intra-patient paradigm (without intersection) and inter-patient paradigm to see the comparison performance. Three existing models are chosen to explore this problem. As a representative of machine learning methods, a classic SVM model proposed in [4] is selected. 6 classic SQIs are calculated from the raw ECG data as the input features of the SVM model. Two of the most popular lightweight CNN models: ShufftleNetV2 (SFN2) [18] and MobileNetV3 (MN3) [19] are also selected, as most of the existing deep learning SQA methods are computationally intensive and not suitable for wearable systems. Since the two models are designed for image data, minor changes are performed on the code to adjust the input shape. Accuracy is used as the metric to measure the assessment performance. For statistical analysis, maximum (MAX), minimum (MIN), average (AVG), and standard deviation (SD) are introduced to notice the variability in accuracy across patients.

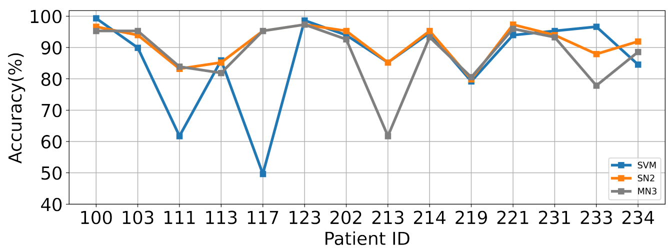
Fig.1 shows the test accuracy of three models on each patient in the intra-patient paradigm and inter-paradigm. From the figure, we can observe that the test performance of both three models varies slightly in the range of 0.8 to 1 across patients in the intra-patient paradigm. While, in the inter-patient paradigm, the variation in test performance across patients is dramatic and accompanied by severe performance degradation. For the SVM model, testing accuracy on Patient 117 was below 50%, which is meaningless for balance binary classification.

Furthermore, statistical analysis of the test accuracy on patients in the two paradigms is shown in Table I. Compared to the intra-patient paradigm, the average test accuracy of patients in the inter-patient paradigm was reduced by 9.37%, 4.05%, and 6.83% for the three models. Meanwhile, the standard deviation more than doubled in the inter-patient paradigm indicating significant impact of individual differences.

From the above motivation experiments, we can conclude that it is necessary and more challenging to study the inter-patient paradigm of ECG SQA due to nonnegligible performance degradation caused by individual differences. To



(a)



(b)

Fig. 1: Test performance of well-trained ECG SQA models on each patient in both two paradigms. (a) Test accuracy on each patient in the intra-patient paradigm. (b) Test accuracy on each patient in the inter-patient paradigm.

TABLE I: Statistical analysis of test accuracy on patients in intra-patient paradigm and inter-patient paradigm

Model	Intra-patient paradigm				Inter-patient paradigm			
	MAX	MIN	AVG	SD	MAX	MIN	AVG	SD
SVM	100.00	84.56	94.23	4.60	99.33	49.66	84.86	13.91
SN2	99.33	89.93	95.03	2.66	97.32	79.87	90.98	5.63
MN3	98.66	86.58	93.83	3.15	97.32	61.74	87.00	9.60

improve the feature extraction adaptability to a new patient in the inter-patient paradigm, building an adaptive ECG SQA model is essential.

IV. Methods

A. Framework

The framework of the AdapSQA can be divided into two stages: pre-training and patient adaptation as shown in Fig.2(a).

In the pre-training stage, a lightweight baseline model for ECG SQA is first designed and trained through labeled data from limited patients in a supervised way. The input data is mapped to the feature space through the feature extractor and achieve properties for each candidate class: acceptable and unacceptable through the classifier. Cross entropy as training label loss is used to regularize the model.

Then, in the patient-adaptation stage, the training data and the new patient data are input into a parallel structure of the baseline model with shared weight separately. A domain adaptation layer is embedded into the structure between the feature extractor and classifier to make the model adaptive to a new patient through aligning the feature distribution of the two domains. The objective of this stage is to minimize the training label loss and domain confusion loss in a balanced way to help the model learn a common representation for both training patients and the new patient.

B. Lightweight Baseline Model

In this subsection, a lightweight baseline model for ECG SQA is designed for better feature extraction with less computational effort. Since SQA is employed before signal analysis in wearable systems, lightweight structure is preferred in the consideration of computational resource limitation. The structure of the lightweight ECG SQA model is shown in Fig.2(b).

1-dimension(1D) convolution neural network (CNN) shows its ability in feature extraction of time sequence physiological signal. The standard convolution operation has the effect of filtering features based on the convolutional kernels and combining features to produce a new representation. To meet the lightweight requirement, 1D depthwise separable convolution is presented to replace standard convolution by factorizing it into a depthwise convolution and a pointwise convolution. Depthwise convolution is used to apply a single filter for each input channel separately. Pointwise convolution, a simple 1×1 convolution, is then used to create a linear combination of the output of the depthwise layer. [19] The computation cost of a 1D standard convolution and a 1D depthwise separable convolution can be calculated as follows with padding:

$$W_k \times M \times N \times W_F \tag{3}$$

$$W_k \times M \times W_F + M \times N \times W_F \tag{4}$$

where the computational cost depends multiplicatively on the number of input channels M, the number of output channels N the kernel width W_k, and the feature map size W_F. In this way, we can get we reduction in the computation of:

$$\frac{W_k \times M \times W_F + M \times N \times W_F}{W_k \times M \times N \times W_F} = \frac{1}{N} + \frac{1}{W_K} \tag{5}$$

Since the kernel width of the baseline model is set to be 3, about $1/3$ computational cost is required compared with the standard convolution. This factorization brings the advantage of significantly reducing the amount of computation requirement. The principle of 1D depthwise separable convolution is shown in Fig.2 (c).

After multiple convolutional layers and a global average pooling layer, the feature extractor maps an input tensor into a 256-D feature vector. The classifier consists of a fully connected layer and a softmax layer, which produces a probability for each category. Rectified linear unit 6 (ReLu6) is also deployed in each layer to increase the nonlinearity.

C. Domain Adaptation Layer

A domain adaptation layer is used to guide the model to learn common feautre representations for both source domain (training patients data) and target domain (the new patient data) by aligning the distribution of the two domains. In

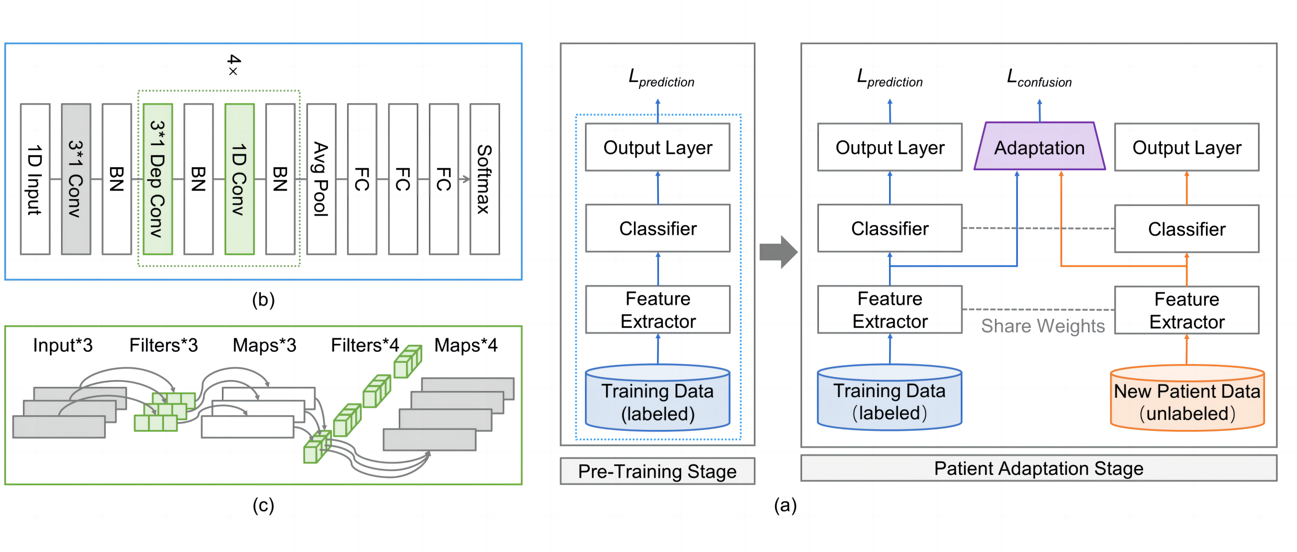


Fig. 2: The overview of the methods. (a) The framework of our proposed AdapSQA. (b) The structure of the lightweight baseline model. (c) The principle of 1D depthwise separable convolution.

TABLE II: Performance variance between before adaptation and after adaptation on patient-specific ECG Noise Dataset

Patient	Before adaptation				After adaptation(CORAL)				After adaptation(MMD)			
	Acc	Pre	Recall	F1	Acc	Pre	Recall	F1	Acc	Pre	Recall	F1
100	95.07	100.00	92.00	0.96	95.97	100.00	92.00	0.96	95.97	100.00	92.00	0.96
103	97.99	100.00	96.00	0.98	98.66	100.00	97.33	0.97	98.66	100.00	97.33	0.99
111	87.92	91.30	84.00	0.88	87.92	91.03	84.00	0.88	86.58	89.86	82.67	0.86
113	87.25	81.11	97.33	0.88	86.58	80.90	96.00	0.88	87.92	86.08	90.67	0.88
117	97.32	97.33	97.33	0.97	97.32	97.33	97.33	0.97	97.99	97.37	98.67	0.98
123	96.64	98.61	94.67	0.97	96.64	98.61	94.67	0.97	97.32	98.63	96.00	0.97
202	95.97	98.59	93.33	0.96	95.97	98.59	93.33	0.96	95.97	100.00	92.00	0.96
213	**57.72**	**54.35**	**100.00**	**0.70**	**66.44**	**60.00**	**100.00**	**0.75**	**89.26**	**86.42**	**93.33**	**0.90**
214	98.66	98.67	98.67	0.99	98.66	98.67	98.67	0.99	98.66	98.67	98.67	0.99
219	85.91	78.12	100.00	0.88	85.91	78.12	100.00	0.88	88.59	82.22	98.67	0.90
221	97.32	100.00	94.67	0.97	97.32	100.00	94.67	0.97	97.32	100.00	94.67	0.97
231	95.97	100.00	92.00	0.96	96.64	100.00	93.33	0.97	96.64	100.00	93.33	0.97
233	**75.84**	**67.57**	**100.00**	**0.81**	**79.19**	**70.75**	**100.00**	**0.83**	**88.59**	85.37	**93.33**	**0.89**
234	91.95	89.87	94.67	0.92	92.62	91.03	94.67	0.93	91.95	89.87	94.67	0.92
Avg	90.17	89.68	95.33	0.92	91.13	90.47	95.24	0.92	93.67	93.89	94.00	0.94

this way, the model learned from limited patients can be well adaptive to the new patient. Maximum mean discrepancy (MMD) loss [20] and correlation alignment (CORAL) loss [21] are often used to regularize the network to minimize the distribution distance between the two domains.

MMD loss was put forward to measure the distance of two different but related domains. Denote by $D_s = \{x_i^s\}_{i=1}^M$ and $D_t = \{x_j^t\}_{j=1}^N$ two sets of samples drawn i.i.d. from the distributions s and t, respectively. By mapping the data to a reproducing kernel Hilbert space (RKHS) using function $\phi(\cdot)$, The empirical approximation to MMD is obtained as follows,

$$MMD^2(D_s, D_t) = \|\frac{1}{M}\sum_{i=1}^{M}\phi(x_i^s) - \frac{1}{N}\sum_{j=1}^{N}\phi(x_j^s)\|_H^2 \quad (6)$$

CORAL loss was defined to align the second-order statistics of the source distribution and the target distribution,

$$CORAL(D_s, D_t) = \frac{1}{4d^2}\|C_s - C_t\|_F^2 \quad (7)$$

where d is the dimension of the x, $C_s(C_t)$ denote the feature covariance matrices and $\|\cdot\|_F^2$ denote the squared matrix Frobenius norm.

MMD loss and CORAL loss are used respectively to achieve domain adaptation for the ECG SQA model to select the best AdapSQA.

V. Experiments

A. Experiment Setting

Several experiments were performed on the patient-specific ECG Noise Dataset to evaluate the proposed model. Since the objective of the proposed model is to improve the inter-patient performance of ECG SQA, we mainly focus on the variation of test performance before and after patient adaptation in the inter-patient paradigm. Then, t-SNE [22] is used to intuitively demonstrate the effect of the proposed model on domain-invariant feature extraction. Comparison experiments with related unsupervised domain adaptation methods are also included. Finally, the computational cost of the proposed model was evaluated to demonstrate its applicability to wearable systems.

Four standard metrics are used to evaluate the assessment performance: overall accuracy (Acc), Precision (Pre), recall (Recall), and F1 score (F1). Model parameters and floating-point operations (FLOPs) are also used to evaluate the computational complexity of the model theoretically.

B. Results Analysis

1) Assessment Performance: Assessment performance of the model before and after adaptation with two different domain confusion losses are reported in TableII. Generally, MMD loss performs better than the CORAL loss in the ECG SQA problem. We can observe that patients who suffered severe performance degradation, i.e., Patient 213 and Patient 233, achieved significant performance improvements after adaptation indicating the effeteness of the domain adaptation. For Patient 213, the improvement is more pronounced, with a 31.54% increase on Acc. Meanwhile, other patients can maintain a relatively acceptable performance after adaptation.

Confusion matrixes on Patient 213 and Patient 233 are also shown in Fig.3 to analyze the performance improvement concretely. The number of false positive samples reduces obliviously after adaptation for both patients, which is the key to increasing the accuracy. Results show that our proposed patient-specific model can significantly improve the discriminative quality of acceptable samples.

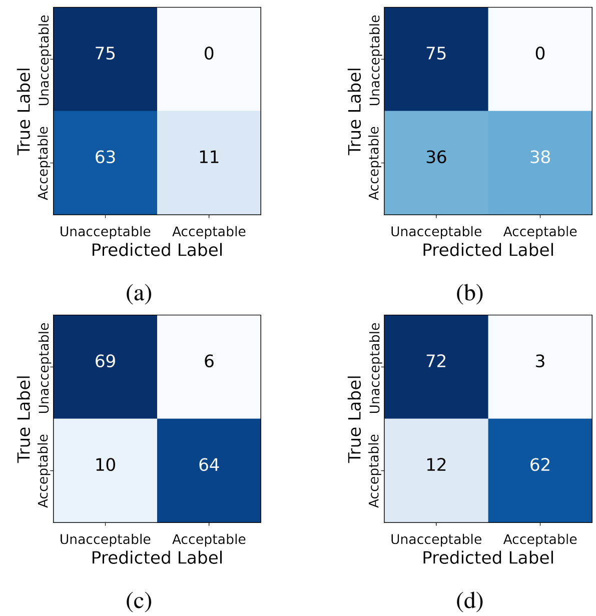


Fig. 3: Confusion Matrix on Patient 213 and Patient 233. (a) Patient 213, before adaptation. (b) Patient 233, before adaptation. (c) Patient 213, after adaptation. (d) Patient 233, after adaptation.

2) Feature Extraction Visualization: To more intuitively demonstrate the effect of the proposed method, we use the t-SNE algorithm to visualize the output features of the feature extractor. The t-SNE of features extracted before and after adaptation by using the proposed model are shown in Fig.4.

The domain distribution shift between the source data and the target data is decreased obliviously while the discriminability of acceptable data and unacceptable data is improved. The proposed model can successfully reduce the discrepancy between the two domains in the feature space, making the learned features in the same category more consistent.

3) Comparison Performance with Related Methods: Comparison performance of the proposed model with some similar unsupervised domain adaptation methods is shown in Table III. The SVM and the SQA are the baseline models trained on the training data without adaptation. TCA and CORAL are two classic domain adaptation methods based on the SVM model, while AdapSQA(MMD) and AdapSQA(CORAL) are the proposed model regularizing with MMD loss and CORAL loss separately.

TABLE III: Comparison results with related unsupervised domain adaptation methods.

Type	Methods	AVG $\pm$ SD (Acc%)
SVM_based	SVM [4]	84.86 ± 13.91
	CORAL [23]	89.02 ± 5.94
	TCA [24]	65.54 ± 21.2
CNN_based	SQA	90.17 ± 10.93
	AdapSQA (CORAL)	91.13 ± 8.90
	AdapSQA (MMD)	93.67 ± 4.41

Compared with these methods, our proposed model has significant advantages where the accuracies reach 91.13% and 93.67% for CORAL loss and MMD loss. The model with MMD loss performs particularly well, the standard deviation is the smallest among all methods, indicating that our method can weaken the impact of individual differences on the assessment performance of ECG SQA in inter-paradigm.

4) Lightweight Evaluation: Since the proposed model will be finally deployed on wearable devices, the size and computational complexity of the model also need to be considered. The total number of parameters for the proposed model is 0.54 M and the FLOPs of the model are 1.36 MMac. Therefore, the AdapSQA is lightweight and can be easily deployed on smartphones or other wearable devices.

VI. Conclusion

In this paper, a novel AdapSQA is proposed to tackle the performance degradation problem in the inter-patient paradigm for ECG SQA through extracting domain-invariant features. To realize AdapSQA, a lightweight baseline model for ECG SQA is first designed and pre-trained to achieve effeteness feature extraction with less computation cost. Then, to alleviate the influence of individual differences that leads to performance degradation in the new patient, a domain adaptation layer is

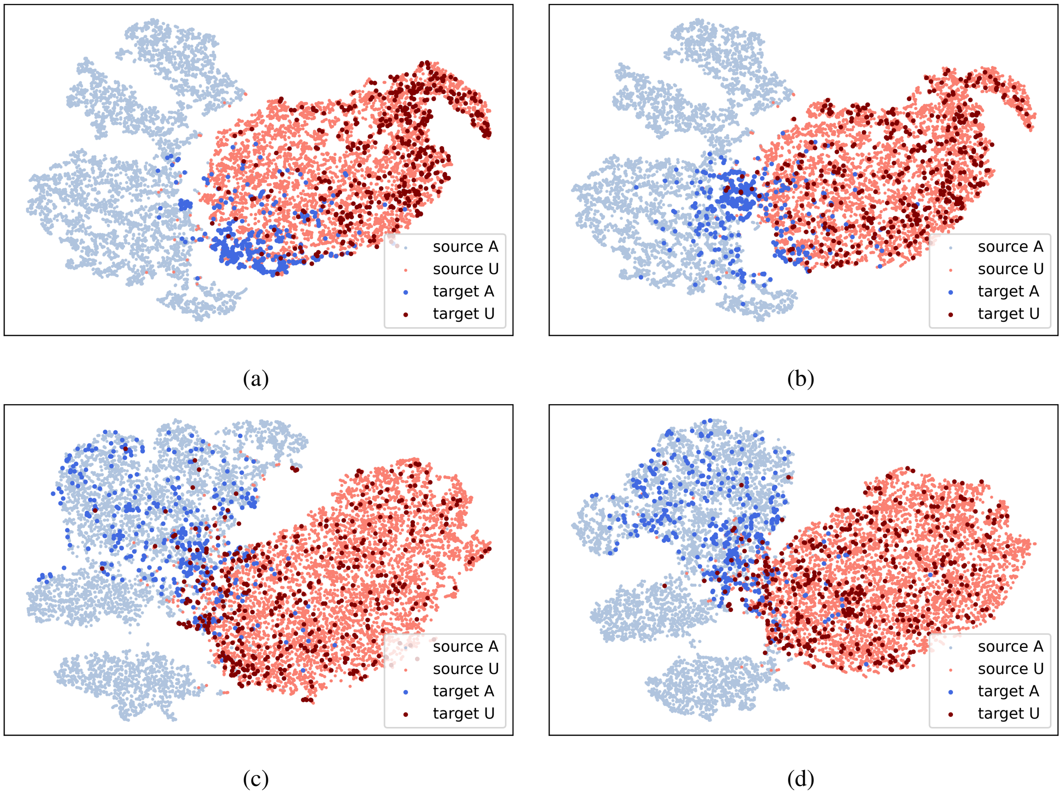


Fig. 4: The t-SNE visualization results of the feature vectors of the proposed model. (a) Patient 213, before adaptation. (b) Patient 233, before adaptation. (c) Patient 213, after adaptation. (d) Patient 233, after adaptation. Acceptable data and unacceptable data are abbreviated with A and U.

employed in the parallel structure of the baseline model to extract common features for both training patients and the new patient. By minimizing the distance between the two domains in the feature space, a baseline model can be adapted to a patient. The proposed model can significantly improve the inter-patient performance on accuracy with smaller variations across patients, outperforming the related domain adaptation methods. Moreover, the proposed model is lightweight and can be easily deployed on wearable devices.

References

[1] C. Orphanidou, "Signal quality assessment in physiological monitoring: state of the art and practical considerations," 2017.

[2] U. Satija, B. Ramkumar, and M. S. Manikandan, "Real-time signal quality-aware ecg telemetry system for iot-based health care monitoring," *IEEE Internet of Things Journal*, vol. 4, no. 3, pp. 815–823, 2017.

[3] J. Behar, J. Oster, Q. Li, and G. D. Clifford, "Ecg signal quality during arrhythmia and its application to false alarm reduction," *IEEE transactions on biomedical engineering*, vol. 60, no. 6, pp. 1660–1666, 2013.

[4] G. Clifford, J. Behar, Q. Li, and I. Rezek, "Signal quality indices and data fusion for determining clinical acceptability of electrocardiograms," *Physiological measurement*, vol. 33, no. 9, p. 1419, 2012.

[5] C. Liu, X. Zhang, L. Zhao, F. Liu, X. Chen, Y. Yao, and J. Li, "Signal quality assessment and lightweight qrs detection for wearable ecg smartvest system," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 1363–1374, 2018.

[6] C. Orphanidou and I. Drobnjak, "Quality assessment of ambulatory ecg using wavelet entropy of the hrv signal," *IEEE journal of biomedical and health informatics*, vol. 21, no. 5, pp. 1216–1223, 2016.

[7] Y. Shahriari, R. Fidler, M. Pelter, Y. Bai, A. Villaroman, and X. Hu, "Electrocardiogram signal quality assessment based on structural image similarity metric," *IEEE Transactions on Biomedical Engineering*, pp. 1–1, 2017.

[8] B. Taji, A. D. Chan, and S. Shirmohammadi, "Classifying measured electrocardiogram signal quality using deep belief networks," in *2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC)*. IEEE, 2017, pp. 1–6.

[9] N. Seeuws, M. De Vos, and A. Bertrand, "Electrocardiogram quality assessment using unsupervised deep learning," *IEEE Transactions on Biomedical Engineering*, vol. 69, no. 2, pp. 882–893, 2021.

[10] S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-

specific ecg classification by 1-d convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2015.

[11] G. Wilson and D. J. Cook, "A survey of unsupervised deep domain adaptation," *ACM Transactions on Intelligent Systems and Technology (TIST)*, vol. 11, no. 5, pp. 1–46, 2020.

[12] R. Ning, C. P. Chen, and T. Zhang, "Cross-subject eeg emotion recognition using domain adaptive few-shot learning networks," in *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. IEEE, 2021, pp. 1468–1472.

[13] G. B. Moody and R. G. Mark, "The impact of the mit-bih arrhythmia database." *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2002.

[14] G. B. Moody, W. K. Muldrow, and R. G. Mark, "Noise stress test for arrhythmia detectors," *Computers in Cardiology*, vol. 11, pp. 381–384, 1984.

[15] Q. Li, C. Rajagopalan, and G. D. Clifford, "A machine learning approach to multi-level ecg signal quality classification," *Computer methods and programs in biomedicine*, vol. 117, no. 3, pp. 435–447, 2014.

[16] Z. Ahmad and N. Khan, "A survey on physiological signal based emotion recognition," *arXiv preprint arXiv:2205.10466*, 2022.

[17] G. Wang, M. Chen, Z. Ding, J. Li, H. Yang, and P. Zhang, "Inter-patient ecg arrhythmia heartbeat classification based on unsupervised domain adaptation," *Neurocomputing*, vol. 454, pp. 339–349, 2021.

[18] N. Ma, X. Zhang, H.-T. Zheng, and J. Sun, "Shufflenet v2: Practical guidelines for efficient cnn architecture design," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 116–131.

[19] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan *et al.*, "Searching for mobilenetv3," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 1314–1324.

[20] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, "Deep domain confusion: Maximizing for domain invariance," *arXiv preprint arXiv:1412.3474*, 2014.

[21] B. Sun and K. Saenko, "Deep coral: Correlation alignment for deep domain adaptation," in *European conference on computer vision*. Springer, 2016, pp. 443–450.

[22] L. Van Der Maaten, "Accelerating t-sne using tree-based algorithms," *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 3221–3245, 2014.

[23] B. Sun, J. Feng, and K. Saenko, "Return of frustratingly easy domain adaptation," *AAAI Press*, 2015.

[24] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang, "Domain adaptation via transfer component analysis," *IEEE transactions on neural networks*, vol. 22, no. 2, pp. 199–210, 2010.