VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

A micro neural network for healthcare sensor data stream classification in sustainable and smart cities

Research Article

A Micro Neural Network for Healthcare Sensor Data Stream Classification in Sustainable and Smart Cities

Jin Wu,[1,2] **Le Sun** [ID][1,2] **Dandan Peng** [ID][3] **and Siuly Siuly**[4]

[1]*Engineering Research Center of Digital Forensics, Ministry of Education,*
Nanjing University of Information Science and Technology, Nanjing, China
[2]*Department of Jiangsu Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET),*
Nanjing University of Information Science and Technology, Nanjing 210044, China
[3]*School of Computer Science and Network Engineering, Guangzhou University, Guangzhou, Guangdong, China*
[4]*Centre for Applied Informatics, College of Engineering and Science, Victoria University, Melbourne, Australia*

Correspondence should be addressed to Le Sun; 002813@nuist.edu.cn

A smart city is an intelligent space, in which large amounts of data are collected and analyzed using low-cost sensors and automatic algorithms. The application of artificial intelligence and Internet of Things (IoT) technologies in electronic health (E-health) can efficiently promote the development of sustainable and smart cities. The IoT sensors and intelligent algorithms enable the remote monitoring and analyzing of the healthcare data of patients, which reduces the medical and travel expenses in cities. Existing deep learning-based methods for healthcare sensor data classification have made great achievements. However, these methods take much time and storage space for model training and inference. They are difficult to be deployed in small devices to classify the physiological signal of patients in real time. To solve the above problems, this paper proposes a micro time series classification model called the micro neural network (MicroNN). The proposed model is micro enough to be deployed on tiny edge devices. MicroNN can be applied to long-term physiological signal monitoring based on edge computing devices. We conduct comprehensive experiments to evaluate the classification accuracy and computation complexity of MicroNN. Experiment results show that MicroNN performs better than the state-of-the-art methods. The accuracies on the two datasets (MIT-BIH-AR and INCART) are 98.4% and 98.1%, respectively. Finally, we present an application to show how MicroNN can improve the development of sustainable and smart cities.

1. Introduction

International Telecommunication Union (ITU) and the United Nations Economic Commission for Europe (UNECE) jointly put forward the construction scheme of a sustainable smart city [1, 2]. The scheme aims to use information technology to improve the level of people's living standards and increase the efficiency of urban services [3]. Problems, such as uneven distribution of medical resources and low efficiency of disease treatment, have gradually become prominent in urban construction [4, 5]. Many research works [6, 7] explore advanced Internet of Things (IoT) and artificial intelligence technologies to solve these problems to promote the development of urban intelligence and sustainability.

The rapid development of deep learning technology and the Internet of Medical Things (IoMT) has brought new opportunities and challenges to medical development in the construction of smart cities [3]. In recent years, some algorithms [6, 8] based on deep learning have been proposed to classify healthcare sensor data streams to solve the problem of medical problems in the process of urban development. Deep convolution neural network (CNN) [9] and deep recurrent neural network (RNN) [10] are two popular methods for classifying healthcare sensor data streams. The former is mainly represented by the one-dimensional convolutional neural network, which can extract the features of one-dimensional time series data [11]. The latter mainly serializes the neurons to process the serialized

data, so that the neurons among the hidden layers can be related to each other [10]. Most of the existing healthcare sensor data classification methods are improved based on the above two methods. However, these methods are difficult to deploy in edge devices because of their large time and space complexity [12].

To reduce the reasoning time and spatial complexity of the model, different lightweight neural network models are proposed in the literature [13, 14]. These methods can be divided into three scenarios: artificially designed lightweight neural network, neural network model compression algorithm, and automatic design of neural network structures [15]. In the first scenario, the model is made lightweight by reducing the number of parameters, for example, limiting the number of channels of features [16, 17], using decomposition convolution operation or $1 * 1$ convolution kernel [18], etc. However, the design process of this scenario needs a lot of time [19]. The second scenario mainly uses knowledge distillation [20] and network slimming [21] to compress the network model. Unfortunately, these methods often realize the lightweight of the model at the cost of sacrificing the performance of the model. The third scenario is to automatically design a neural network architecture to solve a specific task according to a certain search strategy [15, 22, 23]. When using the methods based on the above scenarios to classify healthcare sensor data streams, the accuracy of the models is not very high. It is mainly because these models do not consider how to distinguish classes with similar features [24, 25].

In contrast to the above methods, this paper proposes a novel model that ensures the classification accuracy of each class while ensuring the lightweight of the model, called MicroNN. Since RNN has the advantage of memory preservation for time series data, the architecture based on multilayered RNN [26] is used as the feature extractor of MicroNN. In addition, to improve the identification ability of MicroNN between classes with similar features [27], Kullback Leibler divergence (KL divergence) is introduced in this paper. Experiments show that the overall accuracy and the classification accuracy of each class using MicroNN exceed other work. Our main contributions are as follows:

(i) MicroNN model is composed of a microfeature extractor and some miniclassifiers.

(ii) MicroNN uses a method based on KL divergence to eliminate shared knowledge among classes.

(iii) We conduct comprehensive experiments based on time complexity and space complexity.

The rest of this paper is organized as follows: section 2 presents the related work, section 3 introduces the proposed model, section 4 shows the experiment, section 5 describes an application scenario of MicroNN, and section 6 summarizes this work.

2. Related Work

E-health has become a part of the development of sustainable and smart cities [2, 32]. With the mature development of

Table 1: The space complexity comparison between MicroNN and state-of-the-art methods.

Work	Methods	Model size (MB)
Liu et al. [26]	CNN	39.5
Chen et al. [28]	CNN	32.9
Jun et al. [29]	LSTM	16.2
Saadatnejad et al. [30]	LSTM	15.4
Faust et al. [31]	Bi-LSTM	27.6
Ours	MicroNN	13.7

deep learning and IoMT, healthcare sensor data stream classification based on edge computing has become possible [1, 33, 34]. It will effectively alleviate the uneven distribution of urban medical resources and further accelerate the intelligence development of cities.

According to the survey [6], different diseases are bothering mankind, which seriously threaten human life and quality of life. Nowadays, how to detect and avoid related diseases as soon as possible has become a major issue in urban development [1, 35]. Therefore, disease diagnosis based on healthcare sensor data stream classification has become a hot research topic. Many pieces of research use traditional machine learning methods to classify healthcare sensor data streams, which rely heavily on the characteristics of manual design. Behadada and Chikh [36] proposed a method based on the fuzzy decision tree to improve the detection of arrhythmias. Nasiri et al. [37] designed a model based on the support vector machine and genetic algorithms to diagnose cardiac arrhythmia with relatively high accuracy. Bensujin and Hubert [38] raised a method by combining the K-means clustering algorithm and bacterial foraging optimization algorithm to examine the heart situation of a person. Sharipov [39] used principal component analysis to improve the cardiac diagnosis via ECG. Jadhav et al. [40] proposed static backpropagation algorithms and the momentum learning rule for diagnosing heart diseases.

At present, because of the excellent performance of deep learning technology in the fields of image classification and text recognition, more research works are trying to apply the deep learning model in the field of disease diagnosis. Liu et al. [26] developed a model based on a multiple-feature-branch convolutional neural network for checking the patient's abnormal heartbeat. Chen et al. [28] proposed a new end-to-end scheme using a convolutional neural network (CNN) for automated ECG analysis. Saadatnejad et al. [30] proposed multiple long-short term memory (LSTM) models to monitor the status of heart activity. Faust et al. [31] proposed a bidirectional LSTM for beat detection. Jun et al. [29] used a CNN model with more layers by transforming the healthcare sensor data into a two-dimensional gray image.

Our work is different from the above work. In Table 1, we compare MicroNN with the discussed methods in terms of space complexity. It can be found that the space complexity of the models discussed is relatively larger than MicroNN. It makes some models not widely used in portable devices or edge devices. Therefore, this paper not only considers the accuracy of the model but also further considers the space complexity of the model (Table 2).

TABLE 2: Meaning of the main notations.

| Notation	Meaning
$X = \{x_1, x_2, \ldots, x_n\}$	The raw physiological information record, n is the length of the record.
$S_i = \{v_1, v_2, \ldots, v_r\}$	S_i refers to the ith slice after being segmented.
$S_{i,j}$	$S_{i,j}$ represents the jth point of i$^{\text{th}}$ slice.
$F_i = \{e_1, e_2, \ldots, e_m\}$	The features after feature extractor.
$\text{RNN}^{[1]}$, $\text{RNN}^{[2]}$	$\text{RNN}^{[1]}$ represents the collection of the RNN model at the first level, $\text{RNN}^{[2]}$ represents the collection of the RNN model at the second level.
$X \sim P_X^{\text{class}_i}$	It refers to the data distribution of each class.
$W_1, W_2,$ and W_3	They are the weights of a miniclassifier.
$f_1, f_2, \ldots, f_n$	They are the output of the miniclassifier.
η, c, and π	They are all hyperparameters in the paper.
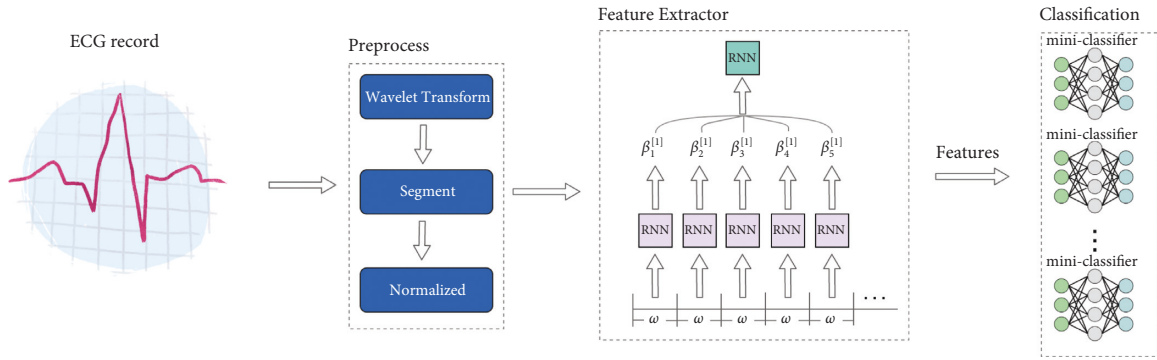
FIGURE 1: The workflow of MicroNN.

3. Our Proposed Model

3.1. System Overview. MicroNN mainly includes three parts: preprocessing model, microfeature extractor, and miniclassifiers. Figure 1 shows the overall architecture of MicroNN. Table 2 is an explanation of the notations used in the paper. The workflow of MicroNN is as follows: a physiological information record $X = \{x_1, x_2, \ldots, x_n\}$. The preprocessing model splits the record into slices with equal length n, and each slice refers to $S_i = \{v_1, v_2, \ldots, v_r\}$.

Then, the microfeature extractor is used to extract the features of S_i, $F_i = \{e_1, e_2, \ldots, e_m\}$. Finally, the feature F_i of S_i is input into each miniclassifier f_i to obtain the corresponding score. Hence, the label of heartbeat S_i is y, as shown in (1).

$$y = \text{argmax}_i (f_1, f_2, \ldots, f_n). \tag{1}$$

3.2. Preprocessing Model. Physiological signals are mainly measured by some mobile edge devices. However, as physiological signals have the characteristics of low amplitude and low frequency, it is easy to be disturbed by noise in the acquisition process [39]. These noises mainly come from internal or external interference [36]. Therefore, the wavelet transform [41] is used to denoise the original signal in this paper. Firstly, the original data is decomposed into nine scales. Then, the wavelet coefficients of nine scales will be processed by threshold operation [41]. Finally, we reconstruct the original data by inverse wavelet transform. Figure 2 shows the changes in physiological signal records (such

as ECG) before and after denoising. Secondly, each physiological signal record is segmented into slices based on the annotations provided by the standard file [42]. Each slice S_i was normalized, $S_{i,j} = S_{i,j}/\|S_i\|_2$, where $S_{i,j}$ represents the j th point of S_i and $\|S_i\|_2$ refers to the 2-norm of a heartbeat slice S_i.

3.3. Microfeature Extractor and Miniclassifiers. In the past, many research works used the convolutional neural network (CNN) as a feature extraction model. However, as CNN needs more computing and storage resources [26], it is difficult to deploy it in edge devices. Consider that the recurrent neural network (RNN) has a memory function in the processing of medical time series data and that its volume is smaller than that of the convolutional neural network [28]. Inspired by ShaRNN [43], this paper mainly adopts the collection of multilevel RNNs as the feature extractor (see Figure 1).

Firstly, it should be noted that we set the RNN collection with two levels. We set the slice data after preprocessing as $S_i = \{v_1, v_2, \ldots, v_r\}$, and we will divide it into some slices whose size is ω. S_i will generate n/ω slices, and we use A_k to represent each slice. Then, we set up an RNN model for each slice:

$$\beta_k^{[1]} = \text{RNN}^{[1]}(A_k), \quad k \in \left[1, \frac{n}{\omega}\right]. \tag{2}$$

Here, $\text{RNN}^{[1]}$ represents the RNN model of the first level, and $\beta_k^{[1]}$ refers to the output of kth slice by $\text{RNN}^{[1]}$. Therefore, we can get the result $[\beta_1^{[1]}, \beta_2^{[1]}, \ldots, \beta_{n/\omega}^{[1]}]$ after the training of RNNs collection of the first level.
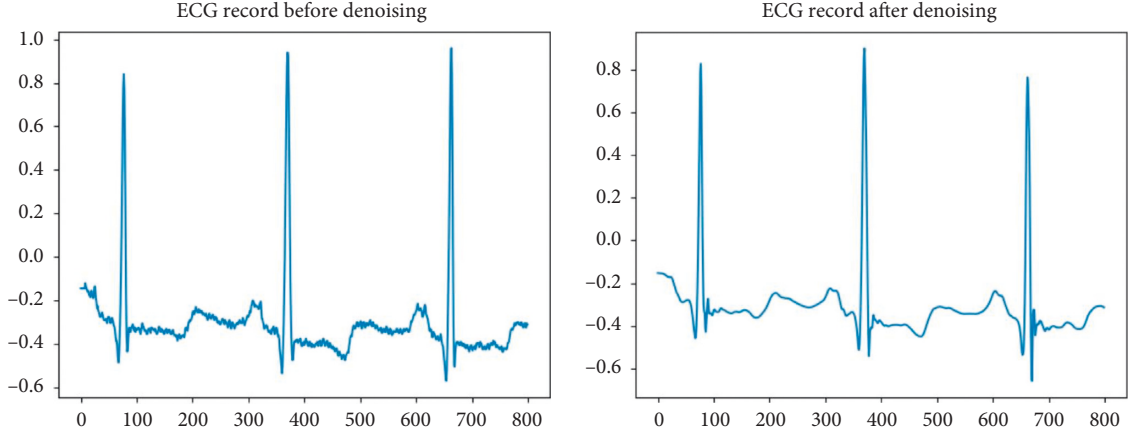
FIGURE 2: The changes in a physiological signal record represented by ECG before and after denoising. (a) ECG record before denoising. (b) ECG record after denoising.

In the next step, we feed the result into the RNN of the second level, and the output is

$$\beta^{[2]} = \text{RNN}^{[2]}\big(\beta_1^{[1]}, \beta_2^{[1]}, \ldots, \beta_{n/\omega}^{[1]}\big), \quad y = \mathscr{F}\big(\beta^{[2]}\big), \quad (3)$$

where $\text{RNN}^{[2]}$ represents the RNN model of the second level, $\mathscr{F}$ refers to the activation function, and y is the extracted feature. It should be noted that $\text{RNN}^{[1]}$ or $\text{RNN}^{[2]}$ can be any RNN model, such as RNN, LSTM, Bi-LSTM, GRU, and so on.

In the selection of a classifier for MicroNN, we adopt a per-class classification model. The model will establish a separate miniclassifier for each class of the task (see the part of classification in Figure 1). All miniclassifiers are connected with the feature extractor. In addition, to improve the performance of the classifier, we employ a loss function called one-class [24] in the training process:

$$\text{loss} = E_{X\sim P_X^{\text{class}_i}}\big[-\log\big(\sigma\big(f_i(X)\big)\big)\big] + \eta \cdot E_{X\sim P_X^{\text{class}_i}}\left\|\frac{\partial f_i(X)}{\partial X}\right\|_2^c$$
$$+ \pi \cdot \big\|\theta_i - \mu_{1:i-1}^*\big\|_2^2, \quad (4)$$

where $X \sim P_X^{\text{class}_i}$ refers to the data distribution of each class, σ is the activation function, and η, c, and π are all hyperparameters.

The first term in the loss function is negative log likelihood. Its purpose is to maximize the score of class_i during training. However, if there is no constraint to the negative log likelihood, it will lead to an unlimited increase in the score. Therefore, the second term, which is called H-reg, is applied in the loss function. It can reach a balance with the negative log likelihood. The structure of per-class classification is a multilayer perceptron with three layers, as shown in (5).

$$E_{X\sim P_X^{\text{class}_i}}\left\|\frac{\partial f_i(X)}{\partial X}\right\|_2^c = E_{X\sim P_X^{\text{class}_i}}\left\|\frac{\partial W_3 \cdot [\sigma(W_2 \cdot \sigma(W_1 \cdot X))]}{\partial X}\right\|_2^c. \quad (5)$$

We can see that the derivation result of H-reg in the training process is related to the weight (W_1, W_2 and W_3).

Therefore, H-reg can restrict the phenomenon of the unlimited growth of weight, which the negative log likelihood brings.

To make the parameters of classifiers between different classes in the same parameter space, the method uses the parameters from 1 to $i-1$ miniclassifiers to initialize the parameters of the ith miniclassifier. Considering the existence of similar features between different classes, deep learning models have difficulty distinguishing classes in the process of training. During the testing stage, a method based on KL-divergence [44] is used to reduce the shared knowledge between classes, as described in the third term of the loss function. Assuming that there are T miniclassifiers in MicroNN, the calculation of shared knowledge among T miniclassifiers is as shown in (6).

$$\rho_{1:T}^* = \arg\min \sum_{i=1}^{T} \varphi_i \text{KL}\big(P_i \| P_{1:T}\big), \quad (6)$$

where φ_i is the mixing ratio with $\sum_{i=1}^{T} \varphi_i = 1$, and P_i refers to the posterior parameter distribution of the ith miniclassifier. The parameters of the ith miniclassifier are updated by (7).

$$\vartheta_i^* = \vartheta_i - \tau \cdot \rho_{1:T}^*, \quad (7)$$

where τ is a hyperparameter.

4. Performance Analysis

The experiments are conducted on a computer with a GPU of Intel (R) Core (TM) i9-11900K and 64.00 GB memory. Experiments are done on two different ECG datasets to evaluate the performance of MicroNN. In the experiment, we divide each dataset into training sets, validation sets, and test sets, and their proportions are 6:2:2, respectively. To better evaluate the performance of the model, we mainly use precision (Pre), recall (Rec), and $F1$-score ($F1$) in the paper. Their relationship is as follows:

$$F1 = \frac{2 \cdot \text{Pre} \cdot \text{Rec}}{\text{Pre} + \text{Rec}}. \quad (8)$$

TABLE 3: The performance comparison between MicroNN and state-of-the-art methods based on MIT-BIH-AR.

Work	Overall ACC(%)	N (%)			S (%)			V (%)		
		PRE	REC	F1	PRE	REC	F1	PRE	REC	F1
MicroNN	98.4	99.0	99.2	99.1	95.1	93.3	94.2	96.5	97.3	96.8
Llamedo and Martinez [45]	78.0	99.1	78.0	87.3	41.0	76.0	53.3	88.0	83.0	85.4
De Chazal et al. [46]	81.9	99.2	86.9	92.6	38.5	75.9	51.1	81.9	77.7	80.0
He et al. [47]	95.1	97.6	97.5	97.6	59.4	83.8	69.5	90.2	80.4	85.0
Zhai and Tin [48]	97.6	98.5	97.6	98.0	74.0	76.8	75.4	92.4	93.8	93.1
Lee et al. [49]	98.1	99.6	97.4	98.5	77.6	91.5	84.0	86.0	89.2	87.6
Li et al. [50]	98.1	98.0	99.8	98.9	94.7	68.7	79.6	91.1	95.5	93.2
Niu et al. [51]	97.5	97.4	98.9	98.1	76.6	76.5	76.5	94.1	85.7	89.7

TABLE 4: The performance comparison between MicroNN and state-of-the-art methods based on INCART.

Work	Overall ACC(%)	N (%)			S (%)			V (%)		
		PRE	REC	F1	PRE	REC	F1	PRE	REC	F1
MicroNN	98.1	99.0	99.0	99.0	88.3	91.1	85.6	95.5	95.0	95.2
Merdjanovska and Rashkovska [52]	94.3	97.7	93.8	95.7	69.3	75.0	72.0	95.7	86.1	90.6
Bidias àMougoufan et al. [53]	81.9	97.7	95.9	96.8	61.8	80.8	70.0	60.9	69.1	64.7
Sun et al. [42]	99.7	99.7	100	99.8	60.8	90.2	72.7	99.0	94.2	96.5

4.1. Datasets Description. The details of the two datasets used in the experiment are as follows:

(1) MIT-BIH arrhythmia database (MIT-BIH-AR) includes the ECG record of 47 subjects studied by the BIH arrhythmia laboratory, and the sampling rate is 360 Hz. It contains 48 half-hour excerpts of two-channel ambulatory ECG recordings. In the experiment, we use the ECG record based on the MLII lead of MIT-BIH-AR. The full name of MIT-BIH is Massachusetts Institute of Technology, Beth Israel Hospital [42].

(2) St Petersburg INCART 12-lead arrhythmia database (INCART) consists of 75 annotated records from 32 humans, and the sampling rate is 257 Hz. Each record lasts for a half-hour and has the data of 12 standard leads. In the experiment, we use the ECG record based on the II lead of INCART.

4.2. Performance of MicroNN. At first, we compared the performance of MicroNN with existing methods at MIT-BIH-AR and INCART (see Tables 3 and 4). Micro has achieved good performance in ACC and F1. As can be seen from Table 3, the low accuracy of other methods is mainly because of the low F1 of class S. It is because class N and class S have many similar characteristics. The model is prone to recognition errors. However, MicroNN's F1 in class S is much higher than other methods, which shows that MicroNN effectively reduces the shared knowledge among classes during training. Similarly, we can see from Table 4 that although the performance of MicroNN in classes N and V is not as good as partial work, MicroNN far exceeds other work in the classification of class S. It is mainly because that MicroNN can effectively solve the problem of the fuzzy boundary.

4.3. Measuring Time and Space Complexity of MicroNN. Table 2 compares the space complexity of MicroNN with other work, which shows that MicroNN is lightweight in terms of space complexity. In addition, we also measure the trend of training time and accuracy of MicroNN based on the change in the number of sample numbers in MIT-BIH-AR and INCART.

It can be seen from Figures 3 and 4 that the accuracy and training time of MicroNN increase with the increase of the number of instances of different datasets on the whole. In MIT-BIH-AR, when the number of instances reaches about 4000, the accuracy reaches 98.4% and tends to be stable. The training time is 23 seconds. For INCART, the number of instances reaches up to 4300 approximately, corresponding to the highest accuracy (98.1%), and the time of training is 27 seconds.

4.4. Threats to Validity. In the paper, threats to the validity of our proposed method are discussed from two perspectives: external validity and internal validity [14].

(1) Threats to internal validity: To prevent the occurrence of overfitting, we divide each dataset into a training set, validation set, and test set. We observed the change in classification accuracy based on different validation sets to check whether the classification model has overfitting.

(2) Threats to external validity: To verify the generalization of the model, we compared MicroNN on two different datasets. The experimental results show that the performance of MicroNN is better than other models.
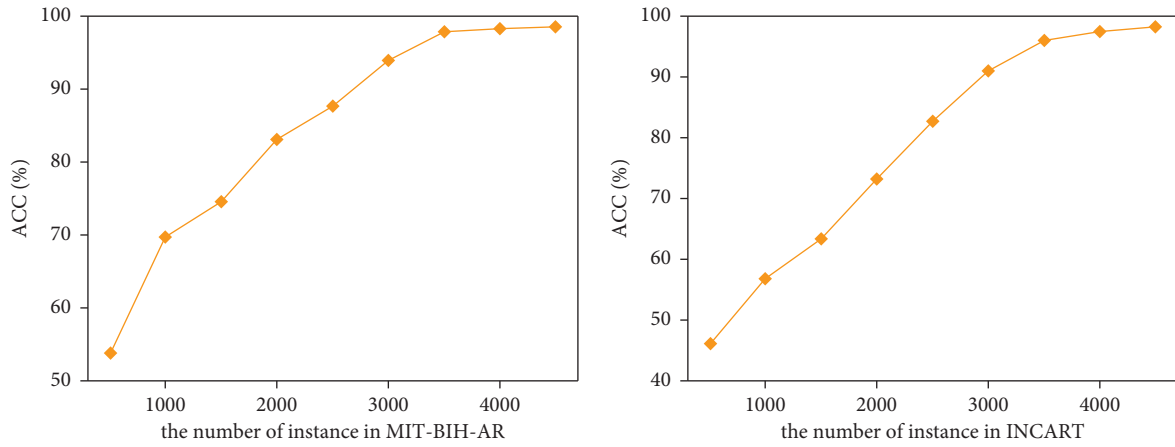
FIGURE 3: Accuracy with respect to the instance numbers in MIT-BIH-AR and INCART.
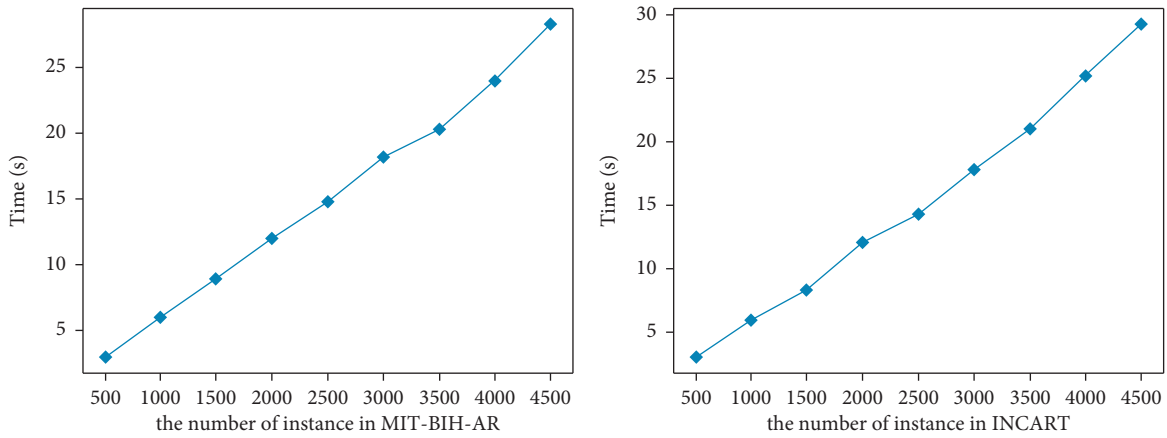

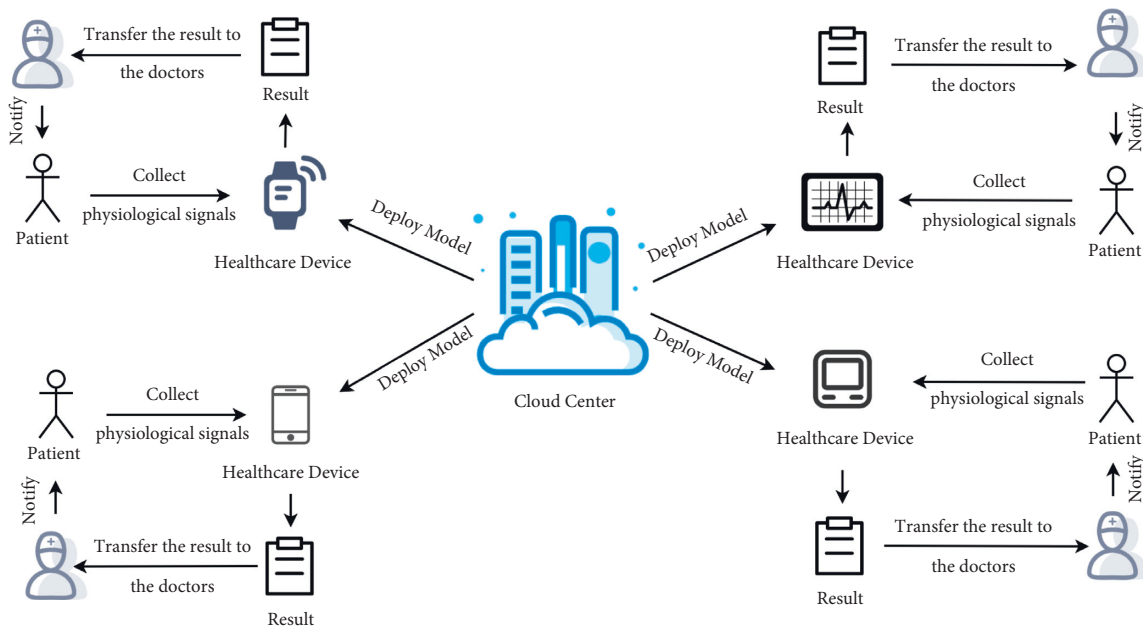FIGURE 4: Time with respect to the instance numbers in MIT-BIH-AR and INCART.

FIGURE 5: An application of MicroNN on the Internet of Medical Things based on edge computing.

5. An Engineering Application of MicroNN

Deep learning research on healthcare sensor data stream classification has attracted extensive attention [33, 54, 55]. However, we still face many challenges in the process of development. For example, the current urban medical resources are insufficient compared with the soaring urban population. The treatment efficiency cannot meet the needs of patients in time [4].

In this paper, we deploy MicroNN in edge devices to effectively improve the efficiency of medical treatment. Figure 5 shows an application example of MicroNN based on edge computing. Different healthcare devices have the function of classifying healthcare sensor data streams. The healthcare devices will classify the collected physiological signals of patients. Then, the results will be used to assist doctors in judging the condition of patients. Finally, the doctor will inform the patient of the specific situation. Therefore, MicroNN plays a certain role in promoting the development of sustainable and smart cities.

6. Conclusion and Future Work

In this paper, we propose a lightweight neural network model called MicroNN for classifying healthcare sensor data streams. It is composed of a microfeature extractor based on multiple recurrent neural networks (RNNs) and multiple miniclassifiers based on a full connection layer with three layers. At the same time, the method based on KL divergence is used to remove the shared knowledge among different classes to improve the performance of the model. In the experiment, we compared the accuracy, time complexity, and space complexity of the model with other models based on two different ECG datasets. MicroNN shows better performance than other works. In a word, MicroNN is a lightweight and efficient model. We will further improve the accuracy of MicroNN while ensuring the lightweight of the model and extend experiments on other healthcare sensor datasets.

Data Availability

The labeled datasets used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

No potential conflicts of interest were reported by the authors.

Acknowledgments

References

[1] Z. Sang and K. Li, "ITU-T standardisation activities on smart sustainable cities," *IET smart cities*, vol. 1, no. 1, pp. 3–9, 2019.

[2] L. Sun, R. Zhou, and D. Peng, "Automatically building service-based systems with function relaxation," *IEEE Transactions on Cybernetics*, pp. 1–14, 2022.

[3] S. Heitlinger, N. Bryan-Kinns, and R. Comber, "The right to the sustainable smart city," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pp. 1–13, ACM, New York, USA, May 2019.

[4] K. Okano, "Regional uneven distribution of healthcare resources related to medical imaging," *Journal of JART-English edition-*, vol. 7, pp. 18–27, 2021.

[5] Z. Qu, Z. Cheng, and W. Liu, "A novel quantum image steganography algorithm based on exploiting modification direction," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 7981–8001, 2019.

[6] Z. Ebrahimi, M. Loni, M. Daneshtalab, and A. Gharehbaghi, "A review on deep learning methods for ecg arrhythmia classification," *Expert Systems with Applications X*, vol. 7, Article ID 100033, 2020.

[7] L. Sun, Q. Yu, D. Peng, S. Subramani, and X. Wang, "Fogmed: a fog-based framework for disease prognosis based medical sensor data streams," *Computers, Materials & Continua*, vol. 66, no. 1, pp. 603–619, 2020.

[8] K. Wang, Y. Zhao, Q. Xiong et al., "Research on healthy anomaly detection model based on deep learning from multiple time-series physiological signals," *Scientific Programming*, vol. 2016, Article ID 5642856, 9 pages, 2016.

[9] T. Wen and Z. Zhang, "Deep convolution neural network and autoencoders-based unsupervised feature learning of eeg signals," *IEEE Access*, vol. 6, pp. 25399–25410, 2018.

[10] C. Zhang, G. Wang, J. Zhao, P. Gao, and J. Lin, "Patient-specific ECG classification based on recurrent neural networks and clustering technique," in *Proceedings of the 2017 13th IASTED International Conference on Biomedical Engineering (BioMed)*, pp. 63–67, IEEE, Innsbruck, Austria, February 2017.

[11] S. Harbola and V. Coors, "One dimensional convolutional neural network architectures for wind prediction," *Energy Conversion and Management*, vol. 195, pp. 70–75, 2019.

[12] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: the communication perspective," *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017.

[13] S. Roy, J. Li, B.-J. Choi, and Y. Bai, "A lightweight supervised intrusion detection mechanism for iot networks," *Future Generation Computer Systems*, vol. 127, pp. 276–285, 2022.

[14] L. Sun and J. Wu, "A scalable and transferable federated learning system for classifying healthcare sensor data," *IEEE Journal of Biomedical and Health Informatics*, 2022.

[15] Y. Zhou, S. Chen, Y. Wang, and W. Huan, "Review of research on lightweight convolutional neural networks," in *Proceedings of the 2020 IEEE 5th Information Technology and Mechatronics Engineering Conference (ITOEC)*, pp. 1713–1720, IEEE, Chongqing, China, June 2020.

[16] H.-Y. Chen and C.-Y. Su, "An enhanced hybrid mobilenet," in *Proceedings of the 2018 9th International Conference on Awareness Science and Technology (iCAST)*, pp. 308–312, IEEE, Taipei, Taiwan, September 2018.

[17] S. Huang, A. Liu, S. Zhang, T. Wang, and N. N Xiong, "BD-VTE: a novel baseline data based verifiable trust evaluation scheme for smart network systems," *IEEE transactions on network science and engineering*, vol. 8, no. 3, pp. 2087–2105, 2021.

[18] H. Liu, X. Mi, and Y. Li, "Smart deep learning based wind speed prediction model using wavelet packet decomposition, convolutional neural network and convolutional long short term memory network," *Energy Conversion and Management*, vol. 166, pp. 120–131, 2018.

[19] H. Park and Y. Kim, "Prediction of strength of reinforced lightweight soil using an artificial neural network," *Engineering Computations*, vol. 28, no. 5, pp. 600–615, 2011.

[20] J. Wang, W. Bao, L. Sun, X. Zhu, B. Cao, and P. S. Yu, "Private model compression via knowledge distillation," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 1190–1197, 2019.

[21] Z. Liu, J. Li, Z. Shen, G. Huang, S. Yan, and C. Zhang, "Learning efficient convolutional networks through network slimming," in *Proceedings of the IEEE international conference on computer vision*, pp. 2736–2744, IEEE, Venice, Italy, August 2017.

[22] J. Yin, W. Lo, S. Deng, Y. Li, Z. Wu, and N Xiong, "Colbar: a collaborative location-based regularization framework for QoS prediction," *Information Sciences*, vol. 265, pp. 68–84, 2014.

[23] Q. Yu and L. Sun, "LPClass: lightweight personalized sensor data classification in computational social systems," *IEEE Transactions on Computational Social Systems*, pp. 1–11, 2022.

[24] W. Hu, Q. Qin, M. Wang, J. Ma, and B. Liu, "Continual learning by using information of each class holistically," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 35, no. 9, pp. 7797–7805, 2021.

[25] Z. Qu, H. Sun, and M. Zheng, "An efficient quantum image steganography protocol based on improved EMD algorithm [J]," *Quantum Information Processing*, vol. 20, no. 2, pp. 1–29, 2021.

[26] W. Liu, Q. Huang, S. Chang, H. Wang, and J. He, "Multiple-featurebranch convolutional neural network for myocardial infarction diagnosis using electrocardiogram," *Biomedical Signal Processing and Control*, vol. 45, pp. 22–32, 2018.

[27] M. A. Rodriguez and M. J. Egenhofer, "Determining semantic similarity among entity classes from different ontologies," *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 2, pp. 442–456, 2003.

[28] M. Chen, G. Wang, P. Xie et al., "Region aggregation network: improving convolutional neural network for ecg characteristic detection," in *Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 2559–2562, IEEE, Honolulu, HI, USA, July 2018.

[29] T. J. Jun, H. M. Nguyen, D. Kang, D. Kim, D. Kim, and Y.-H. Kim, "Ecg arrhythmia classification using a 2-d convolutional neural network," 2018, http://arxiv.org/abs/1804.06812.

[30] S. Saadatnejad, M. Oveisi, and M. Hashemi, "Lstm-based ecg classification for continuous monitoring on personal wearable devices," *IEEE journal of biomedical and health informatics*, vol. 24, no. 2, pp. 515–523, 2020.

[31] O. Faust, A. Shenfield, M. Kareem, T. R. San, H. Fujita, and U. R. Acharya, "Automated detection of atrial fibrillation using long short term memory network with rr interval signals," *Computers in Biology and Medicine*, vol. 102, pp. 327–335, 2018.

[32] Y. Wang, L. Sun, and S. Subramani, "Cab: classifying arrhythmias based on imbalanced sensor data," *KSII Transactions on Internet and Information Systems (TIIS)*, vol. 15, no. 7, pp. 2304–2320, 2021.

[33] W. J. Tan, X. Y. H. Lim, T. Lee, S. C. Wong, H. J. Koh, and D Yeo, "Enhancing the caregiving experience of family care partners in Singapore through an arts programme for persons with dementia: an exploratory study," *Aging & Mental Health*, pp. 1–7, 2021.

[34] Z. Qu, S. Chen, and X. Wang, "A secure controlled quantum image steganography algorithm[J]," *Quantum Information Processing*, vol. 19, no. 10, pp. 1–25, 2020.

[35] P. Lu, "A position self-adaptive method to detect fake access points," *Journal of Quantum Computing*, vol. 2, no. 2, p. 119, 2020.

[36] O. Behadada and M. A. Chikh, "An interpretable classifier for detection of cardiac arrhythmias by using the fuzzy decision tree," *Artificial Intelligence Research*, vol. 2, no. 3, pp. 45–58, 2013.

[37] J. A. Nasiri, M. Naghibzadeh, H. S. Yazdi, and B. Naghibzadeh, "Ecg arrhythmia classification with support vector machines and genetic algorithm," in *Proceedings of the 2009 Third UKSim European Symposium on Computer Modeling and Simulation*, pp. 187–192, IEEE, Athens, Greece, November 2009.

[38] C. Bensujin and C. Hubert, "Detection of st segment elevation myocardial infarction (stemi) using bacterial foraging optimization technique," *Int J Eng Technol*, vol. 6, no. 2, pp. 1212–1223, 2014.

[39] K. Sharipov, "International journal of advanced research in science," *engineering and technology*, vol. 27, p. 2979, 2020.

[40] S. M. Jadhav, S. L. Nalbalwar, and A. A. Ghatol, "Artificial neural network models based cardiac arrhythmia disease diagnosis from ecg signal data," *International Journal of Computer Applications*, vol. 44, no. 15, pp. 8–13, 2012.

[41] A. Arneodo, G. Grasseau, and M. Holschneider, "Wavelet transform of multifractals," *Physical Review Letters*, vol. 61, no. 20, p. 2281, 1988.

[42] L. Sun, Y. Wang, Z. Qu, and N. N. Xiong, "BeatClass: a sustainable ECG classification system in IoT-based eHealth," *IEEE Internet of Things Journal*, vol. 9, no. 10, pp. 7178–7195, 2022.

[43] D. K. Dennis, D. Acar, V. Mandikal et al., "Shallow rnns: a method for accurate time series classification on tiny devices," 2019.

[44] S. W. Lee, J. H. Kim, J. Jun, and J. W. Ha, "Overcoming catastrophic forgetting by incremental moment matching," *Advances in Neural Information Processing Systems*, p. 30, 2017.

[45] M. Llamedo and J. P. Martínez, "Heartbeat classification using feature selection driven by database generalization criteria," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 3, pp. 616–625, 2011.

[46] P. De Chazal, M. O'Dwyer, and R. B. Reilly, "Automatic classification of heartbeats using ecg morphology and heartbeat interval features," *IEEE Transactions on Biomedical Engineering*, vol. 51, no. 7, pp. 1196–1206, 2004.

[47] J. He, J. Rong, L. Sun, H. Wang, and Y. Zhang, "An advanced two-step dnn-based framework for arrhythmia detection," *Advances in Knowledge Discovery and Data Mining*, vol. 12085, p. 422, 2020.

[48] X. Zhai and C. Tin, "Automated ecg classification using dual heartbeat coupling based on convolutional neural network," *IEEE Access*, vol. 6, pp. 27 465–527 472, 2018.

[49] M. Lee, T.-G. Song, and J.-H. Lee, "Heartbeat classification using local transform pattern feature and hybrid neural fuzzy-logic system based on self-organizing map," *Biomedical Signal Processing and Control*, vol. 57, Article ID 101690, 2020.

[50] Y. Li, Y. Pang, J. Wang, and X Li, "Patient-specific ECG classification by deeper CNN from generic to dedicated," *Neurocomputing*, vol. 314, pp. 336–346, 2018.

[51] J. Niu, Y. Tang, Z. Sun, and W Zhang, "Inter-patient ECG classification with symbolic representations and multi-perspective convolutional neural networks," *IEEE journal of biomedical and health informatics*, vol. 24, no. 5, pp. 1321–1332, 2020.

[52] E. Merdjanovska and A. Rashkovska, "Cross-database generalization of deep learning models for arrhythmia classification," in *Proceedings of the 2021 44th International Convention on Information, Communication and Electronic Technology (MIPRO)*, pp. 346–351, IEEE, Opatija, Croatia, September 2021.

[53] J. B. Bidias à Mougoufan, J. S. A. Eyebe Fouda, M. Tchuente, and W. Koepf, "Three-class ecg beat classification by ordinal entropies," *Biomedical Signal Processing and Control*, vol. 67, Article ID 102506, 2021.

[54] J. Guo and B. Li, "The application of medical artificial intelligence technology in rural areas of developing countries," *Health equity*, vol. 2, no. 1, pp. 174–181, 2018.

[55] M. Wu, L. Tan, and N. Xiong, "A structure fidelity approach for big data collection in wireless sensor networks," *Sensors*, vol. 15, no. 1, pp. 248–273, 2014.