2018 IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)
4-7 March 2018
Las Vegas, Nevada, USA

A Deep Multi-task Learning Approach for ECG Data Analysis

Jinlong Ji, *Student Member, IEEE,* Xuhui Chen, *Student Member, IEEE,*
Changqing Luo, *Student Member, IEEE,* and Pan Li, *Member, IEEE*

***Abstract*— Deep learning is an advanced representation learning method and can automatically discover hidden features from raw data. Researchers have attempted to adopt it for ECG data analysis in the past few years. However, traditional deep learning algorithms usually require great efforts and experience to fine-tune the neural networks during their training processes. Moreover, these algorithms may suffer from a sharply declined accuracy when a well-trained model is directly applied to analyze the data from another group of patients. To address these issues, we propose a deep multi-task learning scheme for ECG data analysis which only requires limited efforts to fine-tune the network and can enable the trained model to be well applied to other datasets. Specifically, we first convert the ECG data analysis problem into a multi-task learning problem by dividing the ECG data analysis into multiple tasks. We then construct the multiple datasets for each task. Finally, we design a deep parameter-sharing network which inserts parameter-sharing neural layers in traditional neural networks. We conduct experiments by using the MIT-BIH database to validate the performance of our proposed scheme. Results illustrate that our proposed scheme can improve the accuracy of ECG data analysis by up to about 5.1%.**

I. Introduction

In the past decades, academia and industries have proposed to analyze electrocardiography (ECG) signals by various machine learning methods, such as decision tree [1], support vector machine [2], and hidden Markov models [3]. One very important process of such methods is feature extraction which is used to extract hidden informative attributes from raw ECG data. To perform feature extraction, researchers have proposed many schemes that can be classified into two types: the manually based and the automatically based. For the former ones, feature extraction is based on cardiologists' medical experience and professional knowledge. For example, physicians extract the features, such as RR-interval, the R peak amplitude and the QRS duration from raw ECG data, and diagnose the arrhythmia types by using the decision tree method. For the latter ones, feature extraction is conducted automatically by using digital signal processing techniques. For example, researchers can use discrete wavelet transform (DWT) and principal component analysis (PCA) to extract features from ECG signals automatically [4]. However, such automatic methods may ignore some features in the time domain. Moreover, they still require manual efforts, e.g., choosing DWT level for the DWT-based methods.

Deep learning, an advanced representation learning method, can automatically discover hidden features from raw data. Since deep learning based methods have strong capabilities in nonlinear functional learning, they have been widely used in multi-dimensional signal processing problems like computer vision, speech recognition, and natural language processing [5]. Due to such powerful capabilities, many researchers consider employing deep learning based methods for ECG data analysis. By doing so, ECG data analysis does not require manually extracting features anymore. More importantly, some previous works [4], [6], [7] show that these methods are able to achieve very high accuracy.

However, the neural network based learning methods still suffer from some issues for ECG data analysis. First, deep learning based methods need to fine-tune the neural networks during their training process, which requires lots of efforts and experience. Second, ECG signals are personalized. The same disease may have different signs for different patients. Moreover, it is possible for two distinct diseases to have approximately identical effects on normal ECG signals [8]. Thus, a model, which is well trained by using some patients' data, may suffer from a sharply declined accuracy when it is directly applied to analyze the data from another group of patients. Therefore, to enable deep learning to be used for ECG data analysis, we need to carefully design the scheme.

In this paper, we propose a deep multi-task learning scheme for ECG data analysis. Specifically, we first divide the ECG data analysis into multiple tasks, and thus the ECG data analysis problem is converted into a multi-task learning problem. Based on the tasks, we then construct the corresponding datasets. Finally, to conduct multi-task learning over these datasets, we design a deep parameter-sharing network which inserts parameter-sharing neural layers to traditional neural networks. In particular, our proposed scheme can require very few efforts to fine tune the network. Moreover, the trained model can also be easily transferred to the data from other patients. We summarize the major contributions as follows.

- To the best of our knowledge, we are the first to propose a deep multi-task learning scheme for ECG data analysis.
- We convert the ECG data analysis into a multi-task problem for the accuracy improvement of the ECG data analysis.
- To conduct deep multi-task learning, we develop a parameter-sharing neural network which inserts parameter-sharing neural layers to traditional neural networks.
- Our experiment results show that our proposed scheme can improve the accuracy of ECG data analysis by up

J. Ji, X. Chen, C. Luo and P. Li are with the Department of Electrical Engineering and Computer Science, Case Western Reserve University, Cleveland, OH 44106. Email: { jxj405, xxc296, cxl881, lipan}@case.edu

to about 5.1%.

II. Related Works

A. Deep Neural Network Approach for ECG Analysis

Since deep neural networks achieved a great triumph in computer vision, speech recognition, and natural language processing et al. in last decade, researchers attempt to apply these cutting-edge technologies to ECG analysis field in recent years. In 2015, Zachary C. Lipton [6] trained an LSTM recurrent neural network to automatically diagnose heart diseases. Rajpurkar et at. [7] proposed a 34-layer convolutional neural network to classify the heartbeat arrhythmias, and exceed the average cardiologist performance in both sensitivity and precision.

B. Multi-task Learning

Multi-task learning (MTL) solves multiple similar tasks simultaneously to exploit commonalities and differences across these different tasks and improve the performance of each task. Typically, machine learning researchers only care about optimizing metrics for a particular desired task, which is single task learning (STL). Nevertheless, the MTL systems are given a collection of related tasks that all need to be solved. Theoretical and experimental studies have shown that the overall prediction quality can be improved by processing the tasks jointly and sharing information between them. Moreover, if the tasks are similar but the data is in different domains, in such MTL system, the system can acknowledge domain invariant features. Thus, the MTL systems are more transferable than the STL systems. With these advantages, MTL systems has led to success in many applications of machine learning, from natural language processing and speech recognition to computer vision and drug discovery [9].

III. Deep MTL ECG Analysis Scheme

As shown in Figure 1, we propose a deep MTL scheme for ECG data analysis. The proposed scheme consists of three main processes, expanding single specified ECG analysis task to multiple tasks, constructing corresponding datasets for each task, and designing deep parameter-sharing networks for these multiple tasks. Specifically, for an ECG analysis task, like abnormal heartbeats detection or arrhythmia classification, we expand this single task to multiple related tasks. Then we group the labeled data into several datasets constructed for each task. Finally, we design a deep parameter-sharing neural network to conduct the MTL for ECG analysis. In the following, we describe the details of these three processes. Since the datasets construction heavily depends on how we expand the single task to multiple tasks, we introduce them together in one section.

A. Multiple Tasks and Related Datasets

How to expand the single ECG analysis task to multiple related tasks is the most crucial issue in our deep MTL ECG analysis scheme. The purpose of our scheme is to improve the performance of ECG analysis problem, like abnormal heartbeats detection or arrhythmia classification. Based on the source patients where the data is collected from, and the target patients where the learning system will be applied to, we propose two different ways to expand single ECG analysis task to multiple tasks, and construct corresponding datasets. Particularly, if the ECG data is collected from some patients, and the learning system will be used to analyze the same patients' ECG data, we consider this as the consistent source-target scenario. Otherwise, if the ECG data is collected from some patients, and the learning systems will be transferred to analysis the other patients' ECG data, we name this as the inconsistent source-target scenario.

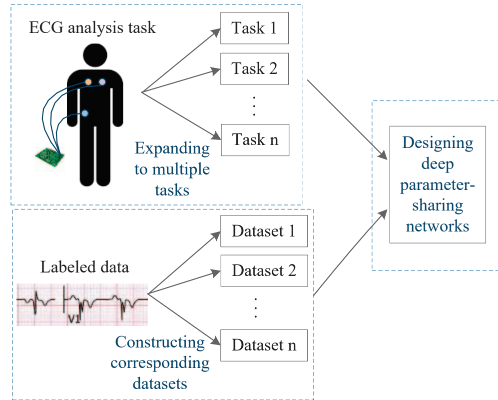

Fig. 1. Deep MTL ECG Analysis Scheme.

1) Consistent Source-Target Scenario: If the source and target are the same group of patients, the metrics we care about is to increase the analysis precision. To improve the single analysis task's performance, we expand single analysis task to multiple tasks. For a specified ECG analysis task, the STL system requires to learn the details of every aspect from the data, but each task in an MTL system only learn its own aspect. By the jointly learning, these tasks can achieve better overall predictions than the STL system.

We take abnormal heartbeats detection and arrhythmia classification as the examples for ECG analysis. For abnormal heartbeats, STL systems treat it as a single anomaly detection problem, which the systems attempt to capture the common pattern of most data and find out the ones different from this pattern. To make abnormal heartbeats detection as multiple tasks, we reconstruct the labeled dataset as two new datasets. In one dataset, most data are normal heartbeats and only a few are abnormal ones, just like the original dataset. On the contrary, in the other dataset, most data are abnormal heartbeats and only a few are normal ones. Corresponding to these two datasets, we have two anomaly detection tasks. Each task learns to capture the common pattern of most data in its own dataset, and as one system, the two learned common patterns of normal and abnormal heartbeats can help each other to know better of its own objective, which will improve the performance of the entire system. Similarly, we expend arrhythmias classification, which typically is a single multi-class classification task, to several bi-classification tasks, each of which is to determine whether the abnormal heartbeat belongs to the type of arrhythmias or not. By learning these tasks jointly, the overall arrhythmia classification precision will increase.

2) Inconsistent Source-Target Scenario: If the source and target are not the same group of patients, due to the slight difference among different patients' ECG signal, the metrics we care about is to learn a transferable system for ECG analysis. The key issue in the inconsistent source and target scenario is to utilize MTL to capture the patients' invariant patterns for ECG analysis. We treat the analysis task on ECG records of different patients as similar learning tasks in different domains. Thus, we build an MTL system to learn the domain invariant features over these tasks.

We still take abnormal heartbeats detection and arrhythmias classification as the examples. For abnormal heartbeats, we take different patients' ECG records as different tasks' dataset and treat each task as an anomaly detection task. When we use a parameter-sharing neural network to solve these task together, the neural network will capture the invariant features among all the tasks. Similarly, we design an MTL system for arrhythmia classifications among different patients' ECG records.

B. Deep Parameter-sharing Neural Networks

The deep parameter-sharing neural network in our proposed deep MTL analysis scheme is not a specified neural network. The deep parameter-sharing neural network is a type of neural networks, which are modified from existing neural networks by the method we propose. As shown in Figure 2, each task can use any existing neural network (the blue block in Figure 2) to extract its private features, then we insert the shared neural layers (the red block in Figure 2) connected to all these tasks. The output of this shared neural layers is connected to each task's own output layer to generate their own results.

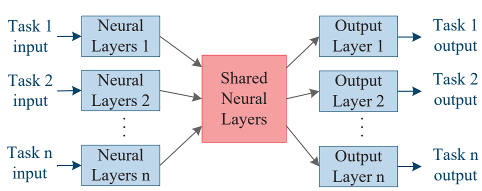

Fig. 2. Deep parameter-sharing neural network.

In particulary, the output of task i's own neural layers is:

$$\mathbf{h}_i = NN\left(x_i\right) \tag{1}$$

The output of shared neural network layers is:

$$\mathbf{h}_s = NN\left(\mathbf{h}_1, \mathbf{h}_2, ..., \mathbf{h}_n\right) \tag{2}$$

The output of task i is:

$$y = softmax\left(w\mathbf{h}_s + b\right) \tag{3}$$

Since both abnormal heartbeats detection and arrhythmias classifications are classification problems, we formulate the loss function for task i as:

$$L_i = -\sum_{i=1}^{N}\sum_{j=1}^{C} y_i^j \log\left(\hat{y}_i^j\right) \tag{4}$$

Combining all the loss functions together, we obtain the loss function for the entire network as:

$$L = \sum_{i=1}^{n} L_i \tag{5}$$

Raw ECG data is time series data of electrical measurement. In deep learning, researchers always utilize one-dimension convolutional neural networks (1D-CNN) or recurrent neural networks (RNN) to capture the features of time sequence data, and both of them can be used to analyze ECG data. In the following, we take these two basic types of neural networks as examples to explain how to build a deep parameter-sharing neural network.

1) Parameter-sharing One-dimension Convolutional Neural Network: As shown in Figure 3, for task i, its training dataset goes through a 1D convolutional neural layer and a 1D max-pooling layer. Then a flatten layer converts the extracted features tensor to a feature vector v_i. Concatenating with feature vectors from other tasks, the common feature vector for all tasks can be represented as $v = \{v_1, v_2, ..., v_n\}$. After several fully-connected layers, whose parameters are shared by all related tasks, the network connects to each task's own desired output layer.

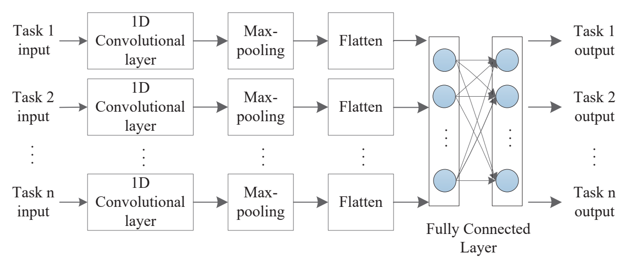

Fig. 3. Deep parameter-sharing convolutional neural network.

2) Parameter-sharing Recurrent Neural Network: As shown in Figure 4, for task i, its training dataset input goes through a recurrent neural layer to extract the hidden state vector v_i. Concatenating with hidden state vectors from other tasks, the common feature vector for all tasks can be represented as $v = \{v_1, v_2, ..., v_n\}$. After several fully-connected layers, whose parameters are shared by all related tasks, the network connects to each task's own desired output layer.

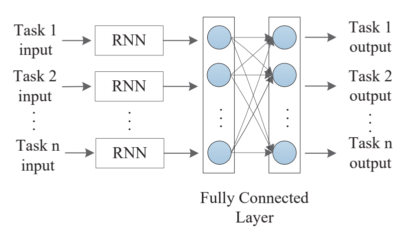

Fig. 4. Deep parameter-sharing recurrent neural network.

IV. Scheme Evaluation

We evaluate our framework by conducting simulations with data from MIT-BIH Arrhythmia Database [10]. Specif-

TABLE I

EXPERIMENTS RESULTS IN CONSISTENT SOURCE-TARGET SCENARIO

	Detection		Classification	
	STL	MTL	STL	MTL
1D-CNN	86.27%	91.47%	79.43%	83.47%
RNN	84.39%	89.23%	80.25%	84.76%
LSTM	89.51%	94.76%	83.84%	87.92%
Avg.	86.72%	91.82%	81.17%	85.38%

TABLE II

EXPERIMENTS RESULTS IN INCONSISTENT SOURCE-TARGET SCENARIO

	Detection		Classification	
	STL	MTL	STL	MTL
1D-CNN	83.16%	87.02%	79.87%	82.43%
RNN	81.35%	85.23%	76.24%	81.29%
LSTM	85.26%	89.11%	81.39%	85.01%
Avg.	83.26%	87.12%	79.17%	82.91%

ically, we use the records 106, 114, 119, 201, 203 and 205 to construct appropriate dataset for our experiments. As discussed in Section III, we take abnormal heartbeats detection and arrhythmia classification as ECG analysis problems, and conduct experiments of both problems in consistent source-target scenario and inconsistent source-target scenario, respectively. For abnormal heartbeats detection, we focus on the identify arrhythmias from normal heartbeats. For arrhythmias classification, we adopt the same criteria in hospitals. Specifically, in out of all types of arrhythmias, physicians are most interested in 4 types, namely Non-sustained ventricular tachycardia (NSVT), Supraventricular tachycardia (SVT), Atrial fibrillation (AF), and Atrial flutter (AFL). Continuous occurrences of these four types of arrhythmias may lead to life-threatening situations like sudden death or stroke. As discussed in Section III, for each problems, we use 1D-CNN, RNN, and LSTM (Long-Short Term Memory network, a variant of RNN) as benchmark STL deep learning methods. Then we compare their performance with their corresponding MTL parameter-sharing neural networks' performance.

A. Experiments on Consistent Source-target Scenario

In consistent source-target scenario, we take records 106, 114, 119, 201, 203 and 205 from MIT-BHI Arrhythmia Database and prepare the data as described in Section III for STL neural networks and MTL parameter-sharing neural networks, respectively. Since most ECG data are normal, we balance the dataset during the training and testing process. In our experiments, as shown in Table I, compared to the performance of STL networks, the average accuracy of MTL increase 5.10% for detection, and 4.21% for classification. The accuracy is the ratio of corrected predict results to all predict results. The experiments results illustrate that deep MTL parameter-sharing systems improve accuracy of each specified problem, and they makes the ECG analysis more accurate.

B. Experiments on Inconsistent Source-target Scenario

In inconsistent source-target scenario, we take records 106, 114, 119, 201, 203 and 205 from MIT-BHI Arrhythmia Database, but we consider records 106, 114, 119, 201, and 203 as source, and record 205 as target. Still, we prepare the dataset as described in Section III for STL networks and MTL parameter-sharing networks, respectively. Similarly to previous scenario, the dataset is balanced and the accuracy is the ratio of corrected predict results to all predict results. As shown in Table II, when transferring the trained networks to a new target, the average accuracy of MTL networks improve the performance of STL ones by up to 3.86% in detection, and 3.74% in classification. The experiments results illustrate that deep MTL parameter-sharing neural networks are more transferable than the STL networks, and they makes the ECG analysis more stable.

V. CONCLUSIONS

In this paper, we have investigated ECG data by employing deep learning based methods. To improve the analysis performance, we have proposed a deep multi-task learning scheme. Specifically, we first convert the ECG data analysis problem into a multi-task learning problem by dividing the ECG data analysis into multiple tasks. We then construct multiple datasets for the tasks. Finally, we design a deep parameter-sharing network which inserts parameter-sharing neural layers in traditional neural networks. We have conducted experiments by using the MIT-BIH database and the results show that the ECG analysis accuracy can be improved by up to 5.1% with our scheme.

REFERENCES

[1] Omar Behadada and Mohammed Amine Chikh. An interpretable classifier for detection of cardiac arrhythmias by using the fuzzy decision tree. *Artificial Intelligence Research*, 2(3):45, 2013.

[2] Inan Guler and Elif Derya Ubeyli. Multiclass support vector machines for eeg-signals classification. *IEEE Transactions on Information Technology in Biomedicine*, 11(2):117–126, 2007.

[3] Benoît Frénay, Gaël De Lannoy, and Michel Verleysen. Improving the transition modelling in hidden markov models for ecg segmentation. In *ESANN*, 2009.

[4] Xuhui Chen, Jinlong Ji, Kenneth Loparo, and Pan Li. Real-time personalized cardiac arrhythmia detection and diagnosis: A cloud computing architecture. In *Biomedical & Health Informatics (BHI), 2017 IEEE EMBS International Conference on*, pages 201–204. IEEE, 2017.

[5] Vivienne Sze, Yu-Hsin Chen, Tien-Ju Yang, and Joel Emer. Efficient processing of deep neural networks: A tutorial and survey. *arXiv preprint arXiv:1703.09039*, 2017.

[6] Zachary C Lipton, David C Kale, Charles Elkan, and Randall Wetzell. Learning to diagnose with lstm recurrent neural networks. *arXiv preprint arXiv:1511.03677*, 2015.

[7] Pranav Rajpurkar, Awni Y Hannun, Masoumeh Haghpanahi, Codie Bourn, and Andrew Y Ng. Cardiologist-level arrhythmia detection with convolutional neural networks. *arXiv preprint arXiv:1707.01836*, 2017.

[8] Shweta H Jambukia, Vipul K Dabhi, and Harshadkumar B Prajapati. Classification of ecg signals using machine learning techniques: A survey. In *Computer Engineering and Applications (ICACEA), 2015 International Conference on Advances in*, pages 714–721. IEEE, 2015.

[9] Sebastian Ruder. An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*, 2017.

[10] George B Moody and Roger G Mark. The impact of the mit-bih arrhythmia database. *IEEE Engineering in Medicine and Biology Magazine*, 20(3):45–50, 2001.