Received February 23, 2019, accepted March 18, 2019, date of publication March 22, 2019, date of current version April 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2907076

Real-Time Detection of Acute Cognitive Stress Using a Convolutional Neural Network From Electrocardiographic Signal

JIAYUAN HE[1], (Member, IEEE), KE LI[2], XIAOLI LIAO[2], PING ZHANG[2], AND NING JIANG[1], (Senior Member, IEEE)

[1]Waterloo Engineering Bionics Lab, Department of Systems Design Engineering, Faculty of Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada

[2]Key Laboratory for Neuroinformation, Research Center for Digital Health Engineering Technology, School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China

Corresponding authors: Ke Li (colinlike@163.com) and Ning Jiang (ning.jiang@uwaterloo.ca)

This work was supported in part by the Collaborative Research and Development Grant from NSERC (CRD#: 503545), in part by the TalentEdge Fellowship Program from the Ontario Center of Excellence under Grant 28273, and in part by the Sichuan Science and Technology Program under Grant 18GJHZ0248.

ABSTRACT As stress is related to many mental and physical health problems, monitoring stress and its management is getting increasingly important in modern societies. Because of the advantage of convolutional neural network (CNN) in automatic feature learning, this study is proposed to use CNN to achieve accurate and fast detection of acute cognitive stress from heart rate variability (HRV). The traditional mental arithmetic calculation was adopted as the stressor for a total of twenty participants, during which one-lead electrocardiogram (ECG) was acquired. Six conventional HRV methods for inferring cognitive stress were extracted from the ECG signals, and their performance in identifying acute cognitive stress was compared with the proposed CNN-based method. The experimental results showed that with a super-short (10 s) time window, the detection error rate of CNN was 17.3%, which is significantly better than the performance of all six conventional HRV methods ($> 7.2\%$, $p < 0.01$). Further analysis showed that the improvement achieved by the proposed CNN methods mainly came from the decrease in false stress sample detection. This study demonstrated the possibility of super-short windows and the advantage of CNN on acute cognitive stress detection. Its outcome would benefit practical applications of real-time stress detection via HRV.

INDEX TERMS Cognitive stress, electrocardiogram (ECG), heart rate variability (HRV), convolutional neural network (CNN).

I. INTRODUCTION

Stress is normally presented when the mental and physiological resources available could not meet the corresponding demands of an individual [1], such as a person is questioned over the area he/she is not familiar with, or a student still has several problems unsolved with a few minutes left in a final exam. In modern societies, stress is prevalent among ordinary people and its consequences can severely affect their daily living, as well as their health. According to the statistics from American Physiological Association and American Institute of Stress, in 2014, there were 77% of US people regularly experiencing physical symptoms, and 73% of US people regularly experiencing psychological symptoms, both of which were caused by stress [2]. Further, 33% of people felt that they lived with extreme stress, and 48% of people confirmed that stress had a negative impact on both their personal and professional life. The scenario in the EU was similar, more than 22% of employees thought that their health was at risk because of their intense work [3]. The subsequent cost for stress caused health care and missed work was countless. In the US, it reached up to 300 billion dollars per year [2]. In the UK, 13 million working days were lost every year with a cost of 12 billion pounds [4]. Consequently, to get rid of stress and keep a healthy lifestyle, stress detection and management was necessary and essential for ordinary people.

The physiological response to stress would be initially reflected on the activities of the autonomic nervous system (ANS) [5]. ANS consists of two branches, the sympathetic and parasympathetic. The exposure to stress events

The associate editor coordinating the review of this manuscript and approving it for publication was Qingxue Zhang.

would break the balance between the two branches. The parasympathetic branch would be suppressed, while the sympathetic branch would be hyper-activated. This information would be transmitted to the cardiac activity, which could be measured by electrocardiogram (ECG) signals. As such, ECG signal is usually adopted as a source signal for stress measurement.

Heart rate variability (HRV) is the variation between the interval of two consecutive R peaks (heartbeats) of ECG signals [6]. As only R peaks are required for calculation, HRV is robust to noise and disturbances because R peaks have the highest amplitude in ECG. As such, HRV is commonly used as an investigation tool for stress detection. To obtain HRV measurement, normally three ECG electrodes are needed, attached on the right leg (RL), right arm (RA) and left leg (LL), to collect the standard Lead II ECG signal [7]. Once all the R peaks are detected, HRV features can be extracted for stress detection [8]. Normally a selected time window is chosen, and the features are calculated within the window. In the HRV based stress detection literatures, the window length was usually adopted in minutes [9], [10]. As such, inferences of cognitive stress could only be made with at least a minute delay, not possible for real-time stress detection. For feature extraction, many methods were proposed based on the physiological symptoms of stress, which induced the increase of heart rate beats [9], balance change between the sympathetic and parasympathetic branch of ANS [6], chaotics of heart rate rhythm [11], etc. The common features of stress detection included heart rate (HR), the power ratio of the low-frequency band (0.04 - 0.15 Hz) to the high-frequency band (0.15 – 0.4 Hz), etc. However, there were studies found that the performance of a single feature was not consistent due to large inter-participant variabilities, and the combination of several features provided better performance in stress detection [10]. This indicated that the response of people to stress might vary and had multiple manifestations in HRV. The extracted features were then sent to the classifier for detection. Traditional classifiers, such as support vector machine (SVM) [12], linear discriminant analysis (LDA). [13], were commonly used in the literature.

Convolutional neural network (CNN) is a class of artificial neural network (ANN), typically containing convolutional layer in its hidden layers [14]. It was initially applied in computer vision tasks, and gradually attracted interests across a variety of domains, including biosignal classification [15], [16], such as electromyogram (EMG) signal classification for gesture recognition [17], electroencephalogram (EEG) pattern identification for assistive machine control [15]. Specifically, in ECG signal classification, CNN was successfully applied in arrhythmia detection [18], signal component identification [19], biometric recognition [20], etc. These studies demonstrated the power of CNN in biosignal classification, implying the potential of CNN in ECG based stress detection. However, to the best knowledge of the authors, this is the first study employing CNN in real-time mental stress detection with ECG signals, i.e., inferring mental stress only from a super-short (10 s) segment of ECG signals. Real-time stress detection was the demand of some practical applications, especially in the scenario of acute cognitive stress detection, which required decisions generated in a super short window. However, with the super-short windows, the performance of conventional HRV features might be limited due to the decrease of window length [21]. The fundamental advantage of CNN was its ability to learn features of the given task, *i.e.* it could capture the information relevant to the task automatically from the input. As such, for stress detection, with CNN, it was not necessary to take feature extraction from HRV measurement, which might explore more information for performance improvement in the super-short window scenario.

This study focused on acute cognitive stress detection within a 10 s window. The acute cognitive stress was induced by a well-established method, mental arithmetic calculation [22], [23]. We compared the performance of the proposed CNN with traditional HRV based stress estimation methods and explored the information CNN used for stress detection. The goal of this study was to promote the practical application of HRV based stress estimation by providing faster, more reliable and real-time results.

II. METHOD

A. EXPERIMENTAL PROTOCOL AND DATA PREPROCESSING

Twenty healthy subjects, eight females and twelve males, aging from 18 to 35, participated in this study. The inclusion criteria were: 1) no history of neurological and heart disease; 2) no allergy to adhesive or rubbing alcohol; 3) receiving secondary school education or higher, capable of performing two-digit arithmetic calculation mentally. Written informed consent was obtained before the experiment. The experimental protocol was in accordance with the Declaration of Helsinki and approved by the Research Ethics Committee of the University of Waterloo (#22188).

Mental calculation method was adopted in this study to induce the acute cognitive stress for the participants. Before the experiment, three ECG electrodes were separately attached on the LA, RA, and RL of each participant (Fig. 1), for the collection of standard lead II ECG signal. The raw data was acquired by a custom-made biomedical signal collection device with a sampling frequency at 500 Hz. The chip of analog to digital converter (ADC) is ADS1298IPAG (Texas Instruments Inc., USA), which are mainly applied in medical instrumentation, such as EEG, EMG and ECG measurement. Its common mode rejection ration (CMRR) is −115 dB. During the experiment, the participants were seated in a height adjustable chair with a headset, around 1 m away from the monitor. There were five sessions for each participant. The first, third and fifth session was the rest session, in which the participant was asked to sit still and relax. Meanwhile, the headset would play comfortable music to help the participants loosen up as much as possible. The duration of the three sessions was 120 s, 60 s and 120 s, respectively. The middle

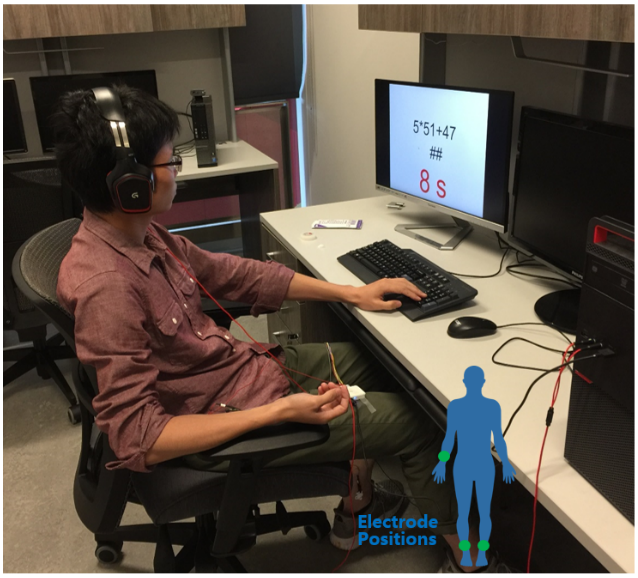


FIGURE 1. Experimental setup. The electrodes were positioned on the right arm, right leg and left leg, respectively. To avoid introducing the noise from activities of muscles, only left hand was allowed to press the keys of the keyboard.

rest session was added to avoid mental fatigue from mental tasks.

The second and fourth session was the mental task session. The participants were asked to mentally perform a series of mathematical calculations, which were displayed in the front monitor. In the second session (the first mental task session), the question consisted of two operations, the multiplication of a one-digit and two-digit number, and the addition/subtraction of the product and a one-digit number. There were 10 questions. For each question, the participant had 10 s to input the answer. In the fourth session (the second mental task session), the question difficulty was increased. The addition/subtraction operation was conducted between the product of a one-digit and two-digit number and a two-digit number, instead of a one-digit number in the second session. There were 30 questions, and the time limit for each question was increased to 20 s. As there was no electrode attached on the left hand, to avoid the interference from muscle contractions, it was the only hand allowed to press the keys for submitting the answers. There were several settings added to increase the mental stimulation level and engage the participants in the calculations. During the question and answer period, the headset would play road noise, such as the footsteps from the pedestrians, engine sound from the vehicles, to interference with the participants. Meanwhile, a stopwatch was positioned below the mathematical question, displaying the remaining time, along with the beeping sound. The pitch of the beeping sound would increase as the time went by. In addition, two or three people would stand behind the participant, of which he/she was aware. Once the answer was submitted, or time run out, the feedback, 'correct' or 'incorrect', would be given immediately. For the correct answers, the sound of doorbell would play, while for the incorrect answers, a loud beep would play. The image of the result would be kept for 5 s. All the participants reported that the stress level of these two sessions evaluated by their subjective experience was higher than that of the rest sessions.

This study mainly focused on the binary classification between stress and rest status. The second rest session and the first 30 s of the last session were abandoned to avoid the effect of residual stress. As such, the data of rest status consisted of the first session and the last 90 s of the last session. To make the data volume even between two classes, the data of stress status consisted of half of the fourth session, which was around 210 s. Before classification, the ECG signals were first band-pass filtered between 1 and 50 Hz to remove the noise from the respiration and muscle contractions. Then the signal was windowed within 10 s for feature extraction. The step size between two consecutive windows was 0.1 s (the overlap was 9.9 s).

B. CNN CLASSIFICATION

To increase the robustness of the system against noise, the spectrum derived from the positions of R peaks was used for the input of CNN. For ECG signal within each window, the R peaks were first extracted with the classic QRS detection algorithm [24], and visual inspection was performed to correct the false acceptance and false rejection of R peaks. Next, the abnormal RR interval was removed based on the criteria in [10]. The remaining interval was termed as normal-to-normal (NN) interval, which was used to construct the zero-one sequence, where the positions of R peaks were set to one, and the others were set to zero. Then the Lomb Periodogram was employed to obtain the spectrum for its simplicity and common in spectral density estimation of HRV data [9], [10], [25] and no tuning parameters compared to parametric methods. The band from 0.04 to 20 Hz was adopted as the input of CNN. The procedures are illustrated in Fig. 2.

A simple CNN structure with ten layers was used in this study. The first layer was an image input layer. Its size was $799 \times 1 \times 1$, which was equal to the size of the input frequency band. The second layer was a convolutional layer, which consisted of six filters of $4 \times 1 \times 1$ with the stride 1-by-1. It was followed by a batch normalization layer, a ReLU layer, and a dropout layer with a probability of 0.5. The next three layers were two fully connected layers and a batch normalization layer between them. The number of units for the two fully connected layers was ten and two, respectively. The last two layers were a softmax layer and a classification output layer, which both had two units.

For the training options of CNN, the stochastic gradient descent with momentum (SGDM) optimizer was used. The maximum number of epochs was set to 30. The learning rate started at 0.001 and was reduced by a factor of 0.1 every 10 epochs. To assess the model performance accurately, a four-fold cross-validation scheme, which was a common method to test the ability of model prediction in machine learning [15], [26], [27], was adopted in this study. The

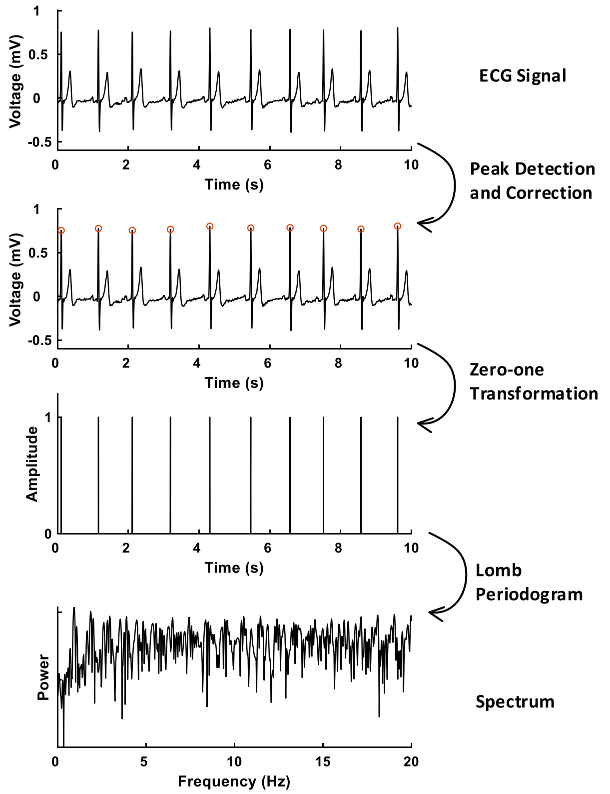


FIGURE 2. Signal transformation from time domain to spectral domain. The frequency band between 0.04 and 20 Hz is used as the input of CNN. The vertical axe of spectrum plot is in log scale.

dataset was divided into four partitions, three of which were used for training and the remaining one was for testing, *i.e.* the samples used for training and testing of each class were 1278 and 426, respectively. This process was repeated four times, so each partition was used the testing set once. The average accuracies of the four folds was used to quantify algorithm performance.

C. CONVENTIONAL CLASSIFICATION

In this study, four types of features based on HRV, plus HR, were investigated. One HRV feature was the standard deviation of NN intervals, termed as SDNN. Two HRV features were derived from the Poincare plot [28], where the interval of two R peaks was plotted against its following. Poincare plot measured the change of NN interval, which was one of the symptoms caused by stress. To quantify the geometry of the Poincare plot, an ellipse was used to fit the plot shape. Two descriptors, termed as SD1 and SD2, were extracted, representing the minor and major semi-axis of the fitted ellipse, respectively. The two HRV features were SD2 and pQ, which was the ratio of SD1 to SD2. Another HRV features were derived from the frequency domain. The spectrum was calculated from the zeros-and-ones sequence, the same in CNN classification. Two frequency bands were identified for their relation to cognitive stress: 1) low-frequency band from 0.04 to 0.15 Hz, and 2) high-frequency band from 0.15 to 0.4 Hz. The ratio of the power of the two bands (LH) [10] was reported to reflect the balance change between the sympathetic and parasympathetic branch of ANS.

As the validity of HRV features with super-short (10 s) windows is still controversial for different statistical methods and stress protocols in literatures [29], [30], SD2, SDNN and pQ were chosen for their good performance with ultra-short windows (usually in minutes) [10]. The selection of, HR and LH was for their prevalence in HRV based stress detection [8], [10], [21]. In addition, the combination of these five features (Comb) were also investigated in this study.

The extracted features were then fed into a classifier to discriminate stress status from rest status. Two common classifiers, LDA and SVM, were separately tested with each feature set. For the setting of SVM, the dot product (linear kernel) was adopted to map the data into kernel space. To help converge in SVM, 0.5% of the variables were allowed to violate the Karush-Kuhn-Tucker (KKT) conditions. Same as the CNN classification, a four-fold cross-validation scheme was adopted. In each run, 75% of the data was used for training and the remaining 25% of the data for testing in four runs. All the data processing was performed on the platform of Matlab R2018b (The Mathworks, Inc., Natick, MA).

D. PERFORMANCE EVALUATION AND STATISTICAL ANALYSIS

For performance evaluation, three metrics were used in this study: error rate (ER), false acceptance rate (FAR), and false rejection rate (FRR). ER quantified the overall performance, which was the ratio between the incorrectly classified samples and the total number of samples. FAR quantified the performance of rest status classification, which was the ratio between the incorrectly classified rest class samples and the total number of the rest class samples. FRR quantified the performance of stress status classification, which was the ratio between the incorrectly classified stress class samples and the total number of the stress class samples. It could be induced that ER, FAR, and FRR was equal to the difference between one and accuracy, sensitivity and specificity [21], respectively.

To compare the performance of detecting stress using various methods, one-way analysis of variance (ANOVA) was used in this study. Each of the three metrics, ER, FRR and FAR was the response variable, respectively. The main factor was the Methods, including the six feature sets (HR, LH, pQ, SD2, SDNN, Comb) with LDA and SVM, respectively, and CNN. In addition, to illustrate the performance improvement from CNN, the activation difference between rest and stress samples of the first five layers, image input, convolution, batch normalization, ReLU, and dropout, was calculated. A student t-test was implemented on the difference of each unit. Further, to investigate the effect of frequency bands on stress detection, a one-way ANOVA was employed on the ER values of CNN with the inputs of different frequency bands. The main factor was the higher limit of the frequency band used: 0.4, 1, 2, 6, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 Hz.

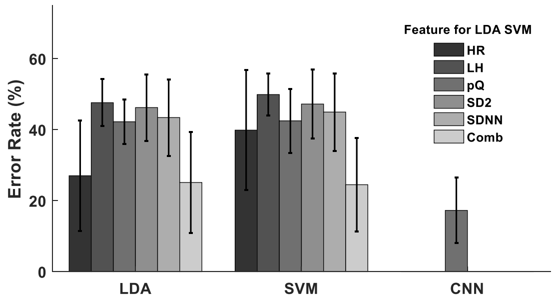


FIGURE 3. Comparison of classification performance in overall error rate between CNN and the conventional methods, which includes six feature sets, heart rate (HR), power ration of low frequency to high frequency (LH), standard deviation of NN intervals (SDNN), minor axis and major to minor axis ratio from Poincare Plot (SD2, pQ), and their combinations, separately being classified by linear discriminant analysis (LDA) and support vector machine (SVM). CNN significantly outperforms the conventional methods.

TABLE 1. Error rate (ER) of the three classifiers for each subject (%).

	LDA	SVM	CNN
S1	5.2 ± 1.3	5.8 ± 3.7	1.9 ± 2.4
S2	35.4 ± 20.2	34.7 ± 20.9	20.7 ± 15.4
S3	23.4 ± 10.3	23.7 ± 10.7	13.0 ± 10.6
S4	40.8 ± 8.5	40.2 ± 6.6	22.6 ± 6.3
S5	39.7 ± 15.3	34.6 ± 11.3	9.4 ± 3.0
S6	3.6 ± 3.1	4.3 ± 4.0	8.0 ± 8.9
S7	38.7 ± 19.9	39.5 ± 18.1	33.4 ± 15.6
S8	39.4 ± 9.8	37.5 ± 8.6	29.5 ± 5.3
S9	29.6 ± 11.1	29.1 ± 11.2	28.0 ± 13.1
S10	47.0 ± 11.1	40.2 ± 6.7	24.1 ± 13.0
S11	12.5 ± 13.1	13.0 ± 17.0	12.9 ± 17.6
S12	36.2 ± 19.3	36.5 ± 17.7	22.8 ± 15.5
S13	28.9 ± 17.3	29.5 ± 17.5	25.5 ± 14.3
S14	18.2 ± 5.9	17.0 ± 5.5	11.4 ± 8.8
S15	33.8 ± 6.4	34.0 ± 6.3	23.9 ± 14.7
S16	5.9 ± 7.5	5.5 ± 6.5	10.4 ± 13.1
S17	33.0 ± 14.7	32.6 ± 14.3	25.7 ± 6.6
S18	9.2 ± 12.4	10.6 ± 17.1	8.2 ± 13.7
S19	12.0 ± 9.5	13.0 ± 10.2	6.9 ± 6.7
S20	9.8 ± 9.4	8.8 ± 9.7	7.2 ± 10.7
Average	25.1 ± 14.2	24.5 ± 13.2	17.3 ± 9.2

The feature set for LDA and SVM is the combination of HR, LH, pQ, SD2 and SDNN.

The lower bound of the frequency band remained at 0.04Hz. If the significant difference was indicated, a post hoc analysis with Bonferroni correction was conducted. The significance level of all the tests was set to 0.05.

III. RESULTS

A. PERFORMANCE COMPARISON

In general, CNN outperformed all the conventional methods (Fig. 3), among which the combination of all the features (Comb) achieved the best performance. Specifically, CNN achieved the lowest ER at 17.3%, which was 7.2% and 32.6% lower than the best and worst performance of the conventional methods, Comb with SVM and LH with SVM, respectively. For each subject, the ER value with three methods, CNN and two best conventional methods, Comb with SVM and LDA, respectively, was listed in Table 1. CNN reduced ER for most subjects. The reduction even reached 25.2% for S5, making the ER value below 10% for this participant.

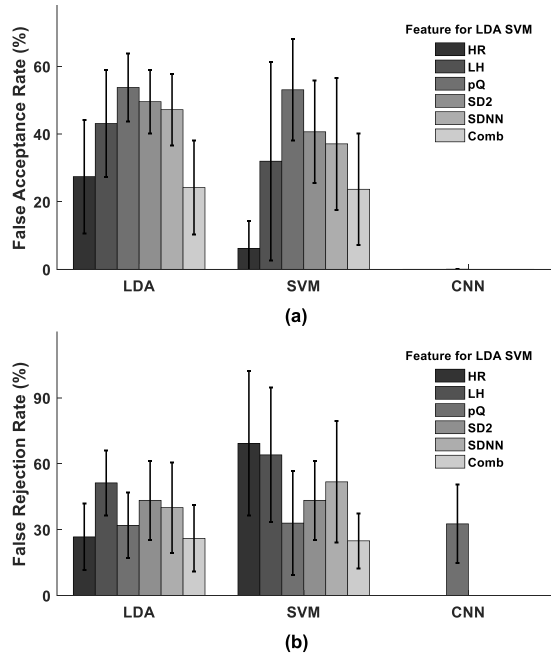


FIGURE 4. The comparison of classification performance in (a) false acceptance rate (FAR), and (b) and false rejection rate (FRR) between CNN and the conventional methods. The conventional methods were the combination of six feature sets (HR, LH, SDNN, pQ, SD2 and Comb) and two classifiers (LDA and SVM). CNN significantly outperforms the conventional methods on FAR (note the average FAR value of CNN is 0.1%, almost invisible in (a)). For FRR, there was no significant difference between CNN and the best conventional method, Comb with SVM.

When dividing the ER into FAR and FRR, it indicated that the ER reduction in CNN was mainly due to the drastic reduction of FAR (Fig. 4). The FAR value of CNN was below 0.1%, close to zero. while the lowest of the conventional methods was 6.2% (HR with SVM), and the FAR value of Comb with SVM, which had the best ER among the conventional methods, was 23.7%. The FRR of CNN was 32.1%, and the range of the conventional methods were between 24.9% (Comb with SVM) and 69.2% (HR with SVM). For the results of statistical analysis, the performance of CNN was significantly better than the performance of the conventional methods in ER ($p < 0.01$) and FAR ($p < 0.001$). In terms of FRR, there was no significant difference between the values of CNN and the best conventional method, Comb with SVM.

B. ACTIVATION DIFFERENCE BETWEEN REST AND STRESS

To investigate the information CNN extracted for stress detection, the activation difference of each layer, *i.e.* the output difference of each layer with respective to rest and stress

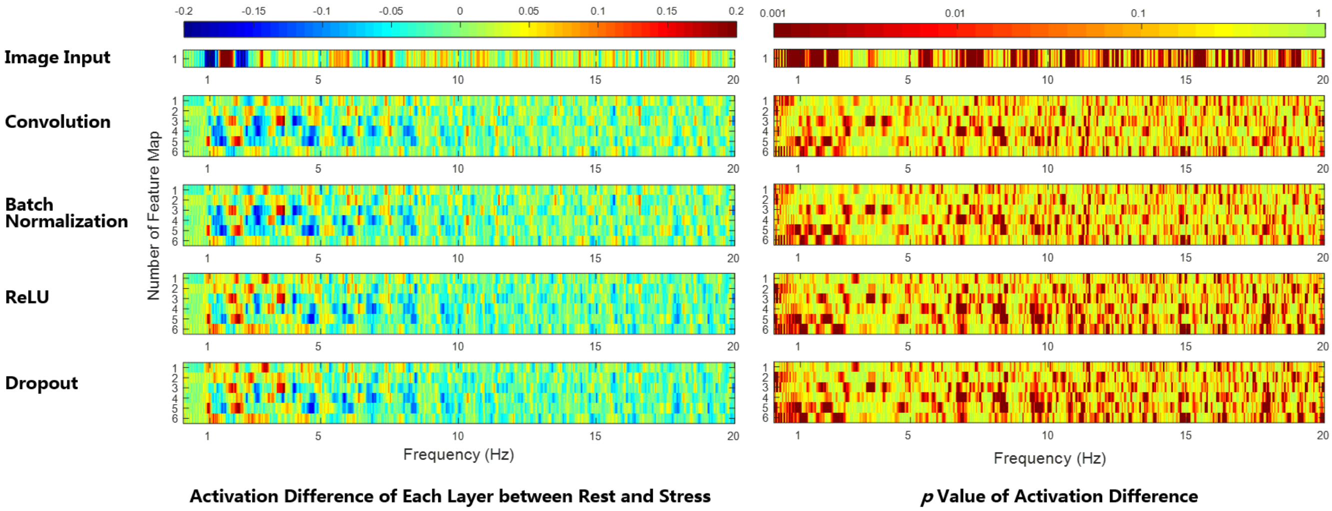


FIGURE 5. Activation difference between rest and stress for first five layer of CNN. The data is normalized by the maximum value of each feature map and averaged across the subjects. For the Convolution to Dropout layer, their activations of each feature map are resized with cubic interpolation to the dimension of image input layer. It indicated that the difference between rest and stress detected by CNN existed in the entire frequency band.

classes, was displayed in Fig. 5. The values were normalized within each layer and averaged across all the subjects. For the last four layers, cubic interpolation was adopted to resize the length of each feature map to fit the dimension of the first layer, image input. The biggest activation difference mainly located around 1.5 Hz, which was expected for it was corresponding to the frequency of R peaks, the main power of ECG signals. However, the differences were also observed in high frequency band. In the results of statistical analysis, it indicated that the significant difference between rest and stress was not only limited to the frequency band < 2 Hz, of which the information was usually used for feature extraction in conventional methods, such as HR and LH, but also existed in the range > 2 Hz, which might contribute to the improvement from CNN.

C. EFFECTS OF THE FREQUENCY BAND

The results of section B indicated that the information extracted from higher frequency content beyond 2 Hz might be the reason for the low ER value of CNN on acute stress detection. To further investigate this point, the CNN was tested with the inputs of different frequency bands. The starting frequency points of these bands were the same, which was 0.04 Hz, while the ending frequency point was different, ranging from 0.4 to 100 Hz. As displayed in Fig. 6, the ER value decreased with the increase of the higher end of the frequency band from 0.4 to 6 Hz and plateaued from there on. The statistical analysis showed that there was a significant difference among the ER values of 0.4, 1 and 2 Hz, but the difference between the ER values of 2 Hz and 6 Hz, as well as the other bands, was not significant.

IV. DISCUSSION

This study demonstrated that CNN can significantly improve the real-time detection of acute cognitive stress compared to the conventional HRV based methods ($p < 0.01$). To investigate the information CNN exploited for stress estimation, the activation differences between the stress state and the rest state of the first five layers were presented (Fig. 5), and the classification performance with the input of different frequency bands was calculated (Fig. 6). The activation difference map showed that there was significant difference between stress and rest in the output corresponding to the frequency band above 2 Hz ($p < 0.05$). The performance comparison showed that there were differences when the upper bound of the frequency band was higher than 2 Hz and above. However, the differences were not significant ($p > 0.05$). This indicated that the information in frequencies higher than 2 Hz might play a role in stress estimation of the CNN used in this study, but it was not critical. The main information CNN captured was still mainly from the frequency band below 2 Hz. Though frequency band above 2 Hz was not emphasized in traditional methods, based on the results of this study, its importance in stress estimation still needed to be further explored.

In order to help the participants relax themselves in the rest session, listening to music was adopted for the strong effect on increasing relaxation [31], [32]. The quick relaxation before and after the experiment would reduce the bias of the data in rest sessions, and hence increased the reliability of the results in this study. The employment of CNN also provided a new way to learn the characteristics of stress. For the conventional methods, the features used mainly depended on the well-established physiological theory of the effect of stress, such as the balance between sympathetic and parasympathetic nervous systems. Because CNN has the automatic feature learning ability, it enabled the researchers to explore the characteristics of stress from the information CNN captured, which may conversely help the researchers, such as biologists and physiologists to acquire more insights the physiological effect of stress. In addition, the CNN structure used in this study was simple, only containing one convolutional layer,

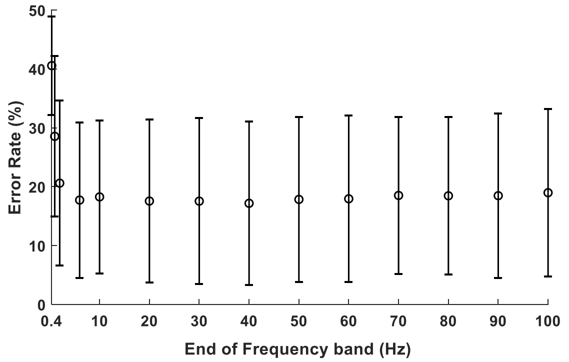


FIGURE 6. Error rate of CNN with different frequency bands as the input. The lower bound of the bands was fixed at 0.04 Hz. The higher bound of the frequency band was 0.4, 1 ,2, 6, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 Hz, respectively. The error rate of 0.4 Hz was significantly higher than that of 1 Hz, and they were both significantly higher than that of 2 Hz, indicating discriminative information exists between 0.4 and 2 Hz. Though the error rate of 2 Hz was higher than that of 6 Hz, there was no significant difference between them, as well as among the values from 6 to 100 Hz.

for consideration of computation load and the training time. However, as the structure would also affect the classification performance, complex structures could be explored in future to improve its detection performance further.

Through breaking down the ER into FAR and FRR, it showed that the main improvement from CNN lied on the reduction of FAR, while FRR of CNN was still comparable to the conventional methods. This means that many stress data were incorrectly classified into rest class. The imbalance between FRR and FAR could be remedied by adjusting the probability threshold. The softmax layer of CNN provided the probability of the sample assigned to each class. By default, the threshold of the probability was set to 0.5. Through optimizing the probability threshold in the training phase, the classification performance of stress data, consequently the overall stress detection performance, could further be improved.

The spectrum derived from the positions of R peaks was used for the input of CNN. As the high amplitude of R peaks, its position detection would be hard to be disrupted by noise from the muscles. On the other hand, the R peaks could also be captured in some non-standard ECG settings, such as arm and ear ECG [27]. These two points made the method robust and applicable in non-ideal conditions. On the other hand, in addition to ECG, there were two other signal modalities, photoplethysmography (PPG) and photoplethysmographic imaging (PPGI), which could also be used for HRV measurement, both of which were non-invasive. PPG was light-based, and the sensor was usually mounted on the mounted on the finger, ear or toe [33]. PPGI was a non-contact measure, collecting signals with a light source and a light detector [34]. The broad spectrum of HRV measurement expanded its application on stress detection. The sensors could be integrated with intelligent mobile devices, such as the smartphone, to make stress estimation available with little limitation of time and place.

V. CONCLUSION

This study compared the performance of CNN with six conventional HRV-based methods in acute cognitive stress detection, using only 10s of ECG data. Stress was induced by mental calculation for twenty participants. Different from previous studies with minutes long windows, super-short temporal windows were used for it was highly desirable in many practical applications, where the real-time stress monitoring was an important feature. The results showed that the performance of CNN was significantly better than the conventional methods ($p < 0.01$), and the improvement was at least 7.2%. The activation map of each layer showed the discriminative information regarding stress versus rest existed in the frequency range higher than the conventional 0.4 Hz, and highly concentrated between 0.4 and 2 Hz. The outcome of this study demonstrated that it was possible to perform real-time detection of acute cognitive stress from HRV, and CNN had potential in practical applications of HRV based stress detection.

REFERENCES

[1] Y. M. Ulrich-Lai and J. P. Herman, "Neural regulation of endocrine and autonomic stress responses," *Nature Rev. Neurosci.*, vol. 10, no. 6, pp. 397–409, 2009.
[2] American Psychological Association. (2015). *Stress in America: Paying With Our Health*. Accessed: Nov. 30, 2018. [Online]. Available: https://www.apa.org/news/press/releases/stress/2014/stress-report.pdf
[3] S. Betti *et al.*, "Evaluation of an integrated system of wearable physiological sensors for stress monitoring in working environments by using biological markers," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 8, pp. 1748–1758, Aug. 2018.
[4] R. Blaug, A. Kenyon, and R. Lekhi, "Stress at work: A report prepared for The Work Foundation's principal partners," Work Found., London, U.K., Tech. Rep., 2007. [Online]. Available: https://westminsterresearch.westminster.ac.uk/item/91wvv/stress-at-work-a-report-prepared-for-the-work-foundation-s-principal-partners
[5] G. E. Billman, H. V. Huikuri, J. Sacha, and K. Trimmel, "An introduction to heart rate variability: Methodological considerations and clinical applications," *Front. Physiol.*, vol. 6, p. 55, Feb. 2015.
[6] F. Lombardi and P. K. Stein, "Origin of heart rate variability and turbulence: An appraisal of autonomic modulation of cardiovascular function," *Front. Physiol.*, vol. 2, p. 95, Dec. 2011.
[7] C. Garcia, *12-Lead ECG: The Art Of Interpretation*, 2nd ed. Burlington, MA, USA: Jones & Bartlett Learning, 2013.
[8] R. Castaldo, P. Melillo, U. Bracale, M. Caserta, M. Triassi, and L. Pecchia, "Acute mental stress assessment via short term HRV analysis in healthy adults: A systematic review with meta-analysis," *Biomed. Signal Process. Control*, vol. 18, pp. 370–377, Apr. 2015.
[9] D. McDuff, S. Gontarek, and R. Picard, "Remote measurement of cognitive stress via heart rate variability," in *Proc. IEEE 36th Annu. Int. Conf. Eng. Med. Biol. Soc.*, Aug. 2014, pp. 2957–2960.
[10] R. T. Ribeiro and J. P. S. Cunha, "A regression approach based on separability maximization for modeling a continuous-valued stress index from electrocardiogram data," *Biomed. Signal Process. Control*, vol. 46, pp. 33–45, Sep. 2018.
[11] R. McCraty and M. A. Zayas, "Cardiac coherence, self-regulation, autonomic stability, and psychosocial well-being," *Front. Psychol.*, vol. 5, p. 1090, Sep. 2014.
[12] M. Kusserow, O. Amft, and G. Tröster, "Modeling arousal phases in daily living using wearable sensors," *IEEE Trans. Affective Comput.*, vol. 4, no. 1, pp. 93–105, Jan./Mar. 2013.
[13] P. Melillo, M. Bracale, and L. Pecchia, "Nonlinear Heart Rate Variability features for real-life stress detection. Case study: Students under stress due to University examination," *BioMed. Eng. OnLine*, vol. 10, p. 96, Nov. 2011.
[14] L. Deng and D. Yu, "Deep learning: Methods and applications," *Found. Trends Signal Process.*, vol. 7, nos. 3–4, pp. 197–387, Jun. 2014.
[15] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *J. Neural Eng.*, vol. 14, no. 1, 2017, Art. no. 016003.

[16] X. Zhai, B. Jelfs, R. H. M. Chan, and C. Tin, "Self-recalibrating surface EMG pattern recognition for neuroprosthesis control based on convolutional neural network," *Frontiers Neurosci.*, vol. 11, p. 379, Jul. 2017.

[17] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, "Gesture recognition by instantaneous surface EMG images," *Sci. Rep.*, vol. 6, Nov. 2016, Art. no. 36571.

[18] J. T. Ruiz, J. D. B. Pérez, and J. R. B. Blázquez, "Arrhythmia detection using convolutional neural models," in *Proc. Int. Symp. Distrib. Comput. Artif. Intell.*, Avila, Spain, vol. 800. Cham, Switzerland: Springer, Jul. 2018, pp. 120–127. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-319-94649-8_15#citeas

[19] Y. Xiang, Z. Lin, and J. Meng, "Automatic QRS complex detection using two-level convolutional neural network," *BioMed. Eng. OnLine*, vol. 17, no. 1, p. 13, 2018.

[20] R. D. Labati, E. Muñoz, V. Piuri, R. Sassi, and F. Scotti, "Deep-ECG: Convolutional Neural Networks for ECG biometric recognition," *Pattern Recognit. Lett.*, to be published. doi: 10.1016/j.patrec.2018.03.028.

[21] R. Castaldo, L. Montesinos, P. Melillo, C. James, and L. Pecchia, "Ultrashort term HRV features as surrogates of short term HRV: A case study on mental stress detection in real life," *BMC Med. Inf. Decis. Making*, vol. 19, no. 1, p. 12, 2019.

[22] N. Hjortskov, D. Rissén, A. K. Blangsted, N. Fallentin, U. Lundberg, and K. Søgaard, "The effect of mental stress on heart rate variability and blood pressure during computer work," *Eur. J. Appl. Physiol.*, vol. 92, no. 1, pp. 84–89, Jun. 2004.

[23] A. Moriguchi *et al.*, "Spectral change in heart rate variability in response to mental arithmetic before and after the beta-adrenoceptor blocker, carteolol," *Clin. Autonomic Res.*, vol. 2, no. 4, pp. 267–270, 1992.

[24] J. Pan and W. J. Tompkins, "A real-time QRS detection algorithm," *IEEE Trans. Biomed. Eng.*, vol. BME-32, no. 3, pp. 230–236, Mar. 1985.

[25] C. L. Shultis, "Effects of music therapy vs. music medicine on physiological and psychological parameters of intensive care patients: A randomized controlled trial," *Diss. Abstr. Int., B, Sci. Eng.*, vol. 73, no. 10-B(E), pp. 83–88, 2013.

[26] M. A. Oskoei and H. Hu, "Support vector machine-based classification scheme for myoelectric control applied to upper limb," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 8, pp. 1956–1965, Aug. 2008.

[27] Q. Zhang and D. Zhou, "Deep arm/ear-ECG image learning for highly wearable biometric human identification," *Ann. Biomed. Eng.*, vol. 46, no. 1, pp. 122–134, 2018.

[28] M. Brennan, M. Palaniswami, and P. Kamen, "Do existing measures of Poincare plot geometry reflect nonlinear features of heart rate variability?" *IEEE Trans. Biomed. Eng.*, vol. 48, no. 11, pp. 1342–1347, Nov. 2001.

[29] L. Pecchia, R. Castaldo, L. Montesinos, and P. Melillo, "Are ultra-short heart rate variability features good surrogates of short-term ones? State-of-the-art review and recommendations," *Healthcare Technol. Lett.*, vol. 5, no. 3, pp. 94–100, 2018.

[30] L. Salahuddin, J. Cho, M. G. Jeong, and D. Kim, "Ultra short term analysis of heart rate variability for monitoring mental stress in mobile settings," in *Proc. IEEE 29th Annu. Int. Conf. Eng. Med. Biol. Soc.*, Aug. 2007, pp. 4656–4659.

[31] C. L. Pelletier, "The effect of music on decreasing arousal due to stress: A meta-analysis," *J. Music Therapy*, vol. 41, no. 3, pp. 192–214, 2004.

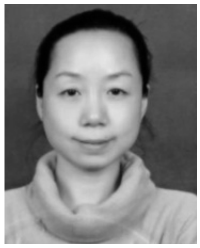
[32] J. C. Smith and C. A. Joyce, "Mozart versus new age music: Relaxation states, stress, and ABC relaxation theory," *J. Music Therapy*, vol. 41, no. 3, pp. 215–224, 2004.

[33] N. Selvaraj, A. Jaryal, J. Santhosh, K. K. Deepak, and S. Anand, "Assessment of heart rate variability derived from finger-tip photoplethysmography as compared to electrocardiography," *J. Med. Eng. Technol.*, vol. 32, no. 6, pp. 479–484, 2008.

[34] R. Amelard *et al.*, "Feasibility of long-distance heart rate monitoring using transmittance photoplethysmographic imaging (PPGI)," *Sci. Rep.*, vol. 5, Oct. 2015, Art. no. 14637.

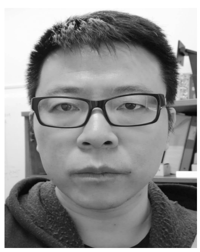
JIAYUAN HE received the bachelor's degree in mechanical engineering and automation from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2010, and the Ph.D. degree in mechanical engineering from Shanghai Jiao Tong University, Shanghai, China, in 2016. He is currently a Senior Research Scientist with the Engineering Bionics Lab, University of Waterloo, Canada. His research interests include biomedical signal processing and myoelectric prosthesis control.

KE LI received the B.S. degree in medical informatics from Tongji Medical University, Wuhan, China, in 1991, and the M.Sc. and Ph.D. degrees in engineering from the University of Electronic Science and Technology of China (UESTC), in 2002 and 2009, respectively. He was an Engineer at the Center of Medical Information Technology, Chengdu, China, from 1991 to 2002. He was a Research Associate Professor with UESTC, from 2002 to 2006. He was a Visiting Scholar with the Department of BME, Case Western Reserver University, Cleveland, OH, USA, from 2007 to 2008. He was the Vice-Dean of the School of Life Science and Technology, UESTC, from 2008 to 2014, where he has been a Professor, since 2009. His research interests include medical informatics, medical signal processing, and brain–computer interfaces.



XIAOLI LIAO received the B.S. degree in biomedical engineering from Chongqing University, Chongqing, China, in 1995, the M.S. degree in biomedical engineering from the Electronic Science and Technology University of China, Chengdu, China, in 2002, and the Ph.D. degree in biomedical engineering from Sichuan University, Chengdu, in 2011. Since 2000, she has been a Lecturer with the Department of Biomedical Engineering, Electronic Science and Technology University of China. Her research interests include biomedical signal measurement and medical instrument developing.

PING ZHANG received the B.S. degree in biomedical engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2010, and the M.S. degree in biomedical engineering from the University of Electronic Science and Technology of China, Chengdu, in 2014, where he has been a Teacher, since 2010. His research interests include medical informatics, medical signal processing, and big date medical.



NING JIANG received the B.S. degree in electrical engineering from Xi'an Jiaotong University, Xi'an, China, in 1998, and the M.Sc. and Ph.D. degrees in engineering from the University of New Brunswick, Fredericton, NB, Canada, in 2004 and 2009, respectively. He was a Research Assistant Professor with Aalborg University, Denmark, from 2009 to 2010, a Marie Curie Fellow of the Strategic Technology Management, Otto Bock Healthcare GmbH, Germany, from 2010 to 2012, and a Research Scientist with the Department of Neurorehabilitation Engineering, University Medical Center Göttingen, Göttingen, Germany, from 2012 to 2015. Since 2015, he has been the Director of the Engineering Bionics Lab, and also an Assistant Professor with the University of Waterloo, Waterloo, Canada. His research interests include signal processing of electromyography and electroencephalogram, advanced prosthetic control, neuromuscular modeling, and brain–computer interfaces for neurorehabilitation. He is currently an Associate Editor of the IEEE Journal of Biomedical and Health Informatics, the *Brain–Computer Interface*, and the *Frontiers in Neuroscience*.