

# Emotional Changes Detection for Dementia People with Spectrograms from Physiological Signals

Zeng Fangmeng<sup>1</sup>, Liao Peijia<sup>2</sup>  
Graduate School of Engineering and  
Science  
Kyoto Institute of Technology  
Kyoto, Japan

Miyuki Iwamoto<sup>3</sup>  
Department of Intelligence Science  
and Technology  
Kyoto University  
Kyoto, Japan

Noriaki Kuwahara<sup>4</sup>  
Graduate School of Science and  
Technology  
Kyoto Institute of Technology  
Kyoto, Japan

**Abstract**—Due to aging society, there has recently been an increasing percentage of people with serious cognitive decline and dementia around the world. Such patients often lose their diversity of facial expressions and even their ability to speak, rendering them unable to express their feelings to their caregivers. However, emotions and feelings are strongly correlated with physiological signals, detectable with EEG and ECG etc. Therefore, this research develops an emotion predicting system for people with dementia using bio-signals to support their interaction with their caregivers. In this paper, we focused on a previous study for binary classification of emotional changes using spectrograms of EEG and RRI by CNN, verifying the effectiveness of the method. Firstly, the participants were required to watch simulating videos while collecting their EEG and ECG data. Then, STFT was performed, processing the raw data signals by extracting the time-frequency domain features to get the spectrograms. Finally, deep learning was used to detect the emotional changes. CNN was used for arousal classification, with an accuracy of 90.00% with EEG spectrograms, 91.67% with RRI spectrograms, and 93.33% with EEG and RRI spectrograms.

**Keywords**—Emotion classification; people with dementia; EEG and ECG; spectrograms; CNN

## I. INTRODUCTION

### A. Background

According to report from World Population Prospects 2017, the number of people aged 60 or over, is expected to rise from 962 million globally in 2017 to 2.1 billion in 2050 and 3.1 billion in 2100 [1]. Furthermore, rapidly aging populations worldwide result in an increase of people living with age-related illnesses like dementia, causing considerable concern for future health and social care provision [2]. It's estimated that the proportion of the general population aged 60 and over with dementia is between 5 to 8 percent, meaning that 50 million people were living with dementia in 2017, while the total number of people with dementia is predicted to approach 75 million in 2030 and almost triple by 2050 to 132 million [3]. Unfortunately, no treatment is currently available to cure dementia or to alter its progressive course. Therefore, early diagnosis and the promotion of optimal health management, along with the care provided by family and friends are of the utmost importance in caring for people with dementia.

As the illness gradually develops, cognitive impairment and decreased ability to communicate as dementia becomes more

severe [4], resulting in an increasing physical, emotional and financial stress for caregivers. As cognitive decline progresses into its later stages, it can become difficult for patients to express their feelings appropriately, making communication with caregivers increasingly difficult. Therefore, in order to support care for such dementia patients in daily life, the need for emotion predicting systems has been put forward in the field. Generally speaking, that is the need to detect the emotions and feelings of dementia patients to support caregivers by helping them to understand what the patients are experiencing.

### B. Related Works

In general, the main method used for detecting the emotional state of a subject is through their emotional expressions [5], like speech, facial expression, gesture, and/or physiological signals. There is already a body of work focusing on emotional expressions: speech signal, facial expressions and gestures [6][7][8]. However, physiological signals also reflect internal emotional information, and have increasingly received more attention recently. They are comprised of the signals originating from the central nervous system (CNS): electroencephalogram (EEG), and the autonomic nervous system (ANS): electrocardiogram (ECG), galvanic skin response (GSR), electromyogram (EMG) etc. Recently presented studies have shown that ANS, which includes the sympathetic nervous system and the parasympathetic nervous system, is viewed as a vital component of emotional response [9].

Among the most common methods for measuring physiological signals for emotion recognition are EEG and ECG, which have been shown by cognitive theory and psychological experiments to indicate possible underlying emotional states [10]. EEG analysis is mainly comprised of four frequency bands: delta waves ( $\delta$ : 0.5-4Hz), usually appearing during deep sleep; theta waves ( $\theta$ : 4-8Hz), always present during drowsiness; alpha waves ( $\alpha$ : 8-13Hz), associated with relaxation and super-learning; beta waves ( $\beta$ : 13-30Hz), related with active thinking and attention [11]. The partial correlations of cortical functions with EEG channels are shown in Fig. 1 [12]. ECG, on the other hand, records the electrical activity of the heart. Based on feature extraction of ECG signals, heart rate variability (HRV) is determined by heart rate (HR) or RR interval, calculated from R peaks. A low HRV can indicate a relaxed state, whereas a high one represents a

potential state of stress or depression. While signal power [13] in low frequency (LF: 0.04-0.15Hz) and high frequency (HF: 0.15-0.4Hz) bands also appear to have a close relationship with emotions.

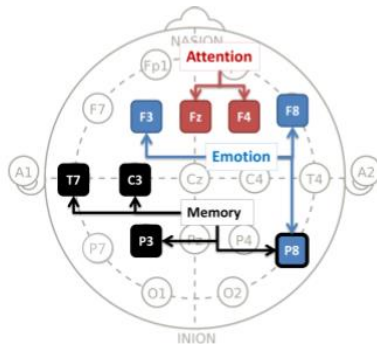


Fig. 1. [12] The Partial Correlations of Cortical Functions with EEG Channels.

While speech and facial appearance might be disguised by a person intentionally or consciously [14], the physiological signs are natural reactions of the body, and as such are very difficult to mask. Since it is difficult to detect the emotions of a dementia patient by speech or facial expression, these signals may be a more reliable way to access that information. This study concentrates on emotion classification based on EEG and ECG, one of which represents the CNS and the other the ANS parameter [15].

### C. Convolutional Neural Network

Over these years, there has been a lot of works in the field of affective computing, recognizing emotions from psychological signals by different classification algorithms. The team led by Prof. Picard [16] firstly performed the research on the emotion detection from signals: ECG, GSR and EMG. They extracted the time domain features, like mean value, variance value, first order differential mean value, etc. The Sequential Floating Forward Selection (SFFS) and Fisher Projection (FP) were chosen to select features for classification of 8 kinds. The best accuracy reached 82.5%. Murugappan et al [17], applied KNN and linear discriminate analysis to get an accuracy of 77%-88% using HRV features extracted from ECG. Mohammadi et al [18], utilized the frequency domain features for ECG, and classified by KNN and SVM. Nasehi et al [19] obtained an accuracy of 64.78% from EEG by neural network classifier.

Upon above research and results, it could be concluded that it's feasible to detect emotion state from physiological signals. However, most of them focused on time domain or frequency domain features, few used spectrograms created from time-frequency domain features as image classification using CNN. Also, there were not many studies combing the features together drawn from EEG and ECG data for emotion classification. The paper opted another perspective for feature extraction to get spectrogram images from EEG and ECG. We intended to sort different emotions using CNN conveniently to get good accuracy simultaneously.

Currently, the application of deep learning in the field of image classification is increasingly popular, especially CNNs (convolutional networks), which have been proven extremely effective in computer vision and are perhaps the most widely used approach in the class of machine learning algorithms [20]. Their specific characteristic is that they capture features automatically, which is the integration of feature extraction and classification into a single algorithm only with the "raw" data (like pixels) as inputs [21]. More recently, applying CNNs for bio-signal related problems, such as EEG and ECG, has attracted growing attention from psychologists and researchers [22]. In this paper, we adopt CNNs to process spectrograms (2-D time-frequency data) with the aim to classify emotions, in order to address the above stated needs of dementia patients. With this method, we avoid the need to select frequency bands or channels in the process of feature selection.

## II. EXPERIMENT AND METHODS

### A. Emotional Data Acquisition

There are many different ways to elicit the target emotions, such as pictures, film clips, music, emotional behaviors, imaging different scenarios and telling stories [23][24][25][26]. Film-clips have been found to elicit target emotions better compared with other methods. In this experiment, six different short video clips [27][24][28] were chosen in advance according to other articles and YouTube rankings to stimulating different emotions (scared, excited, quiet, bored, sad, happy). Each video clip lasts around 2.5-4 minutes and was sandwiched between "quiet time" for 10 seconds, used to provide a short time gap for a smoother transition between emotional states. The entire process lasted for about 20 minutes.

### B. Participants and Experiment Protocol

There were 11 healthy young participants, 5 males and 6 females, aged between 24~30 years old, participated in the experiment. Prior to the experiment, each subject was asked to sign a consent form and filled out a questionnaire about their recent mental state. Next, a presentation was given to them about the meaning of the experiment protocol, and scores used for evaluation. We arranged an experimenter available for any questions. After the participants understood about the experiment contents, they were asked to sit down and prepare for the experiment. Then the electrodes were placed on the right position on the skin of the subject. The signals should be checked for stability, after that, there was a practical trial for the participants to get familiar with the whole system. The subjects were requested to relax, minimize movement and concentrate on the videos. After that, there were a few minutes for them to calm down, and then reset their physiological signals during a 1 min quiet state, followed by the each video for a total of 20 minutes. The whole emotion inducing process is as shown Fig. 2.

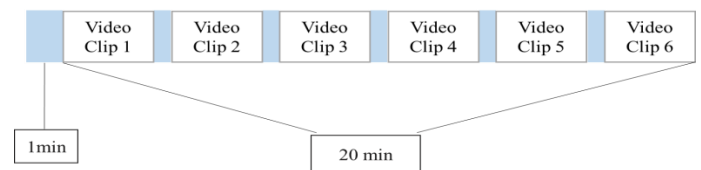


Fig. 2. The Emotion Inducing Process.

### C. Materials and Setup

The experiments were conducted in a laboratory environment with suitable conditions ( $T = 25^{\circ}\text{C}$ ,  $\text{RH} = 50\%$ ). EEG and ECG signals were collected on an equipped recording PC (Dell Precision M3800). Stimulating videos were performed on a PC (HP Touch Smart). The subjects were seated approximately 70 cm from the screen. The speaker volume was set at a relatively loud level; however, each subject was asked whether the volume was comfortable or not before the experiment and it was adjusted when necessary.

EEG and ECG signals were recorded simultaneously using Polymate II AP216 [29], an all-in-one type bio-signal recording device, at a sampling rate of 1000Hz using 7 active electrodes. Fig. 3 illustrates the whole experiment of one subject, (a) shows the monitor instrument set: Polymate II AP216 Dell Precision M3800; (b) shows one subject of the experiment; (c) describes the electrode placement on the subject's body, REF.E channel used as the main electrodes, electrode 1 and 2 set to get ECG, electrode 3, 4, 5 used for EEG with two channels (4, 5 two dipole electrodes and 3 one ground); (d) illustrates the data collecting page on the measurement software AP Monitor [29], which can set up measurement conditions, the information of subjects, also, can monitor the waveform as standard during measurement.

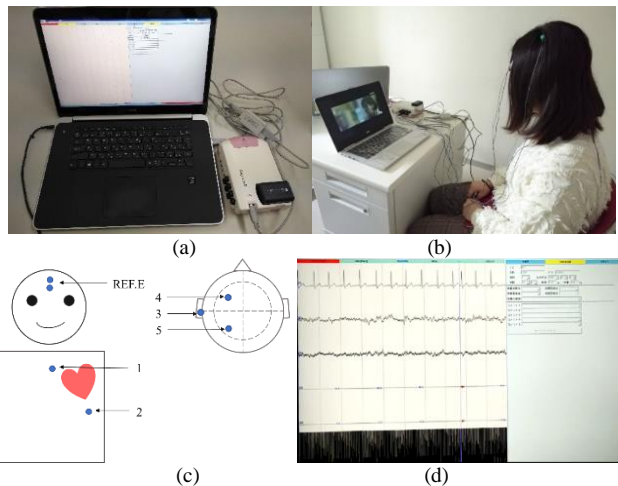


Fig. 3. The whole Experiment Setup. (a) the Monitor Instrument Set: Polymate II AP216 and Dell Precision M3800; (b) the Experiment of One Subject; (c) the Electrode Placement; (d) the Signals Collecting Page.

### III. DATA ANALYSIS

The raw ECG and EEG data measured by the Polymate II AP216 had the noise due to muscle tone, blinking, and other small movements removed by the display software AP Viewer [29]. The software can redisplay the waveform of data recorded with the instrument, and export specified sections of the file during a given period into CSV format.

#### A. EEG Spectrogram

For EEG feature extraction, the spectrogram analysis of EEG was performed in a time-frequency domain method, using the short time Fourier transform (STFT) and applying a hamming window. The window size for spectrogram analysis was 4s with 3.9s overlapping, sliding window was 0.1s. We

recorded the amplitude of each frequency band, consisting of delta wave, theta wave, alpha wave, and beta wave frequencies. The delta band (0-4Hz) was removed to eliminate the remaining noise from pulses, neck movement, and eye blinking [30]. Our research focused on the alpha (8-14Hz) and beta (14-30Hz) frequencies. Alpha waves are typical for relaxed mental state and are most visible in normal daily lives. Beta waves are related to an activated state of mind appeared during intensive mental activity [31]. The analysis method is reported as Fig. 4, which was realized by the API "signal processing" [32] through "SciPy", a Python-based ecosystem of open-source software for mathematics, science, and engineering to render many kinds of functions and algorithms.

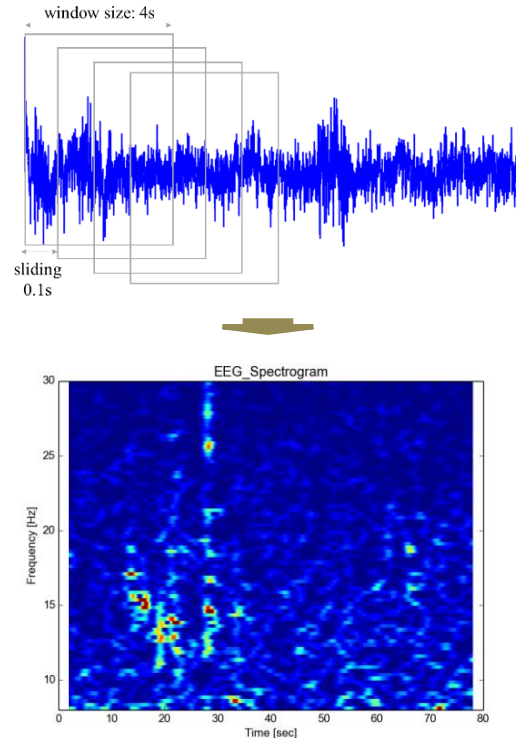


Fig. 4. The Spectrogram Process of EEG Data.

#### B. ECG Analysis

The whole process of ECG is shown in Fig. 5. To extract features from ECG, the peaks of R-waves on the ECG plot were detected, as Fig. 5(a). After R-peaks detection, the R-R intervals (RRI) were calculated by subtracting the peak time from the previous timing to form the waveform which shows the temporal alternation of the peak timings, as Fig. 5(b). Due to this process, the RRI are sampled sparsely and unevenly. To apply a time-frequency domain analysis, the time intervals need to be resampled by 4Hz to get the evenly data graph, as Fig. 5(c). In the analysis, we used a window size of 60s and a step size of 1s, to get the RRI spectrogram. And the LF (0.04-0.15Hz) and HF (0.15-0.4Hz) amplitudes were obtained from the spectrogram, arranging the frequency from 0-0.5Hz, as Fig. 5(d). To get the RRI, the package "BioSPPy" [33], a toolbox for bio-signal processing written in Python, was adopted. Besides, the RRI spectrograms were achieved by the API "signal processing" [32] through "SciPy".

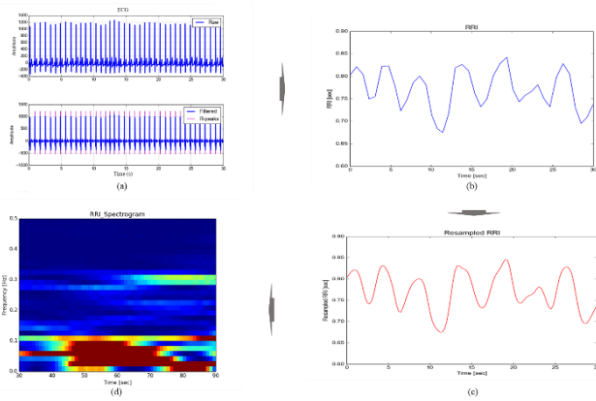


Fig. 5. The Processing Steps of ECG. (a) R-Peaks Detecting; (b) RRI Calculating; (c) RRI Resampling; (d) RRI Spectrogram Processing.

### C. Prepare Dataset

As the measured data of 1 subject was incomplete due to her exceeding reaction to the scared scenes, her data was excluded. The other 10 participants, 5 males and 5 females were taken into account for emotion changes detection. Considering that the emotion of participants may change during one video and the adequate information of RRI spectrograms, the spectrogram images were cut into 15s pieces. And the subjects were asked to make an evaluation about the videos per 15s. Then based on their evaluations, the video pieces per 15s were labeled with the scores. However, the evaluations on the fifth video eliciting boredom from subjects showed considerable variation, and so the data was discarded. The video showed somebody playing games by himself because nobody would celebrate his birthday with him. It seemed his words and accents were difficult to understand, so the subjects didn't know what was happening.

The problem of binary classification of arousal for changes in emotion was considered, where a given set of signals is classified by the subject according to how much the emotion was induced by each video. The subjective rating has a 9-point rating scale, so we defined the scores ranging from 1 to 5 as the low arousal class, labeled as 0, and those from 6 to 9 as the high arousal class [33], labeled as 1, as illustrated in Table I. The representative EEG and RRI spectrograms of two states: the high arousal emotion and the low arousal emotion were shown below, as Fig. 6. For the low arousal emotion as in Fig. 6(a), there's no sign of beta waves, while the high arousal emotion as in Fig. 6(b), beta waves are a prominent component of the signal. When feeling relaxed and experiencing no big emotional changes, HF amplitudes are prominent, as in Fig. 6(c). On the other hand, when emotional arousal was higher, the pressured moment may persist, and the LF amplitudes become more predominant, as shown in Fig. 6(d).

TABLE I. THE SCORES OF VIDEOS AND THE CLASSIFICATION LABEL

Scores	Class	Label
1-5	Low arousal	0
6-9	High arousal	1

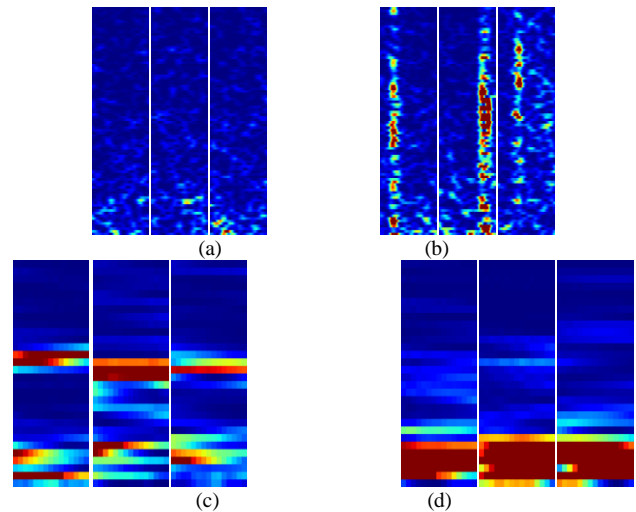


Fig. 6. The Representative EEG and RRI Spectrograms of the Two States. (a) the EEG Spectrograms of the Low Arousal Class; (b) the RRI Spectrograms of the High Arousal Class; (c) the RRI Spectrograms of the Low Arousal Class; (d) the RRI Spectrograms of the High Arousal Class.

### D. Results

CNN classification was constructed to detect emotional changes. The network adopted in this study was implemented in "Keras" (Python), and mainly constructed of convolution and pooling layers. The networks for EEG and RRI spectrograms were adjusted according to the different features.

First of all, we tried to apply VGGNet-16 [34], an extensive architecture built for the ImageNet image classification challenge. Both EEG and RRI spectrograms were processed separately with 300 epochs, this took about 34 hours, with 405s per epoch. The accuracy of EEG was 93.33%, while RRI's was 91.67%, this verified the stability and reliability of the dataset.

Then we designed simpler CNN structures for EEG and RRI, these required much less time, at just 40s per epoch. The CNN1 used for EEG had three convolutional layers, three max-pooling layers. A fully connected layer with 256 hidden nodes was added after the last max-pooling layer to obtain the output defined as 0 for the low arousal class and 1 for the high arousal class. The size of all the filters in the convolutional layers was set to 3x3, so the frequency resolution of EEG spectrograms is smaller and the small scale features are considered. We used a dropout probability of 0.3 after the fully connected layer to avoid overfitting.

Next, CNN1 was applied to the RRI spectrograms; however, this seemed to produce greater over-learning. We believe this is because the scales and characteristics of their features are different. Therefore, CNN2 was set for RRI by adjusting some parameters of CNN1. CNN2 had two convolutional layers, and two max-pooling layers, followed by fully connected layers. The size of all the filters in convolutional layers was set to 5x5, to bring out the bigger features of the image. The dropout was used after two max-pooling layers and full connected layer. Dropout regularization has been proved to be an effective way to reducing the overfitting in deep learning [35].

The ReLU activating function for the convolutional layers of both networks was used. Also, the optimizer used for both CNNs was a SGD (Stochastic gradient descent), with a smaller learning rate of 0.0001. After 300 epochs, the accuracy of CNN1 with EEG spectrograms was 90.00%, and the CNN2 for RRI spectrograms was 91.67%. The accuracy curve for each network is as follows in Fig. 7 and Fig. 8.

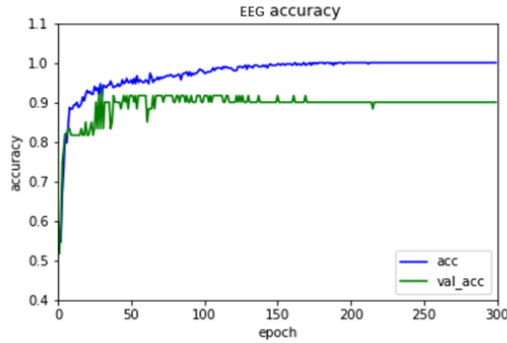


Fig. 7. The Accuracy Curves of CNN1 for EEG Spectrograms.

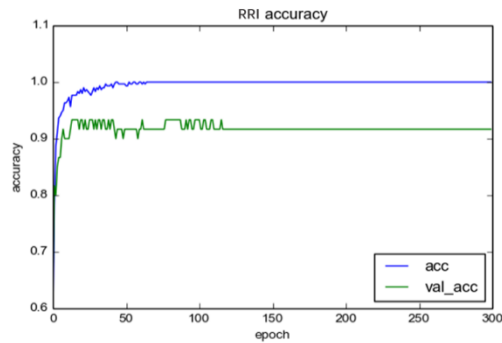


Fig. 8. The Accuracy Curves of CNN2 for RRI Spectrograms.

TABLE II. THE MERGED STRUCTURE OF CNN1 AND CNN2

Input 1	Input 2
EEG spectrograms	RRI spectrograms
CNN1	CNN2
Conv (32, (3×3), ReLU)	Conv (32, (5×5), ReLU)
Max-pooling (2×2)	
	Dropout (0.2)
Conv (32, (3×3), ReLU)	Conv (32, (5×5), ReLU)
Max-pooling (2×2)	
	Dropout (0.2)
Conv (32, (3×3), ReLU)	
Max-pooling (2×2)	
Flatten	
FC-256	
Dropout (0.3)	
FC-2	
<b>Concatenate</b>	
Softmax	
SGD (lr = 0.0001)	

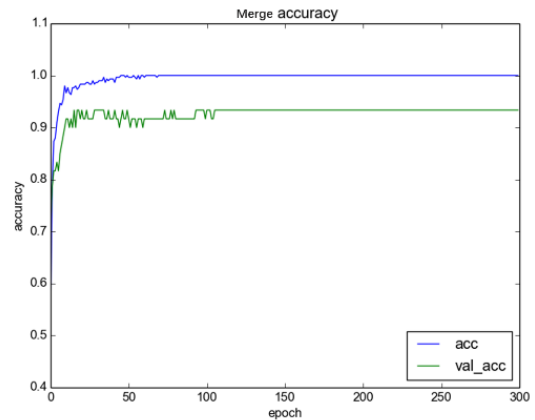


Fig. 9. The Merge Accuracy of CNN1 and CNN2 with Spectrograms of both EEG and ECG.

After that, both EEG and RRI together are considered as features for the binary classification. The merged network structure according to the “concatenate” function from “Merge layers” in “Keras” [36], is shown in Table II. It takes the spectrograms of EEG and RRI as inputs for CNN1 and CNN2 respectively, then merge the two outputs to get the final result. The overall accuracy reached 93.33% after 300 epochs, as illustrated in Fig. 9.

A complete comparison of accuracies between EEG and RRI (Table III), demonstrated that combining features of both EEG and RRI for classification of emotional changes performed better than only EEG or RRI features separately.

TABLE III. COMPARISON OF CLASSIFICATION ACCURACIES BETWEEN EEG, RRI, EEG AND RRI

Signals	Networks	Accuracy
EEG	CNN1	90.00%
RRI	CNN2	91.67%
EEG+RRI	CNN1+CNN2	93.33%

#### IV. CONCLUSIONS

In this paper, we strove to build upon prior research in the field of emotion detection for people with dementia and explored new techniques for increasing the effectiveness of deep learning models for EEG and ECG signals. It was proven that with EEG and RRI spectrograms from young people, binary emotional changes classification could be performed well by CNN. This allows for the convenience of automatic feature extraction. As a result, we are hopefully that the emotion detection system supposed in the study could be implemented in practice in caring for the elderly.

However, there are still some limitations. Our results show only binary classification for emotional changes; however, for real applications, recognition of the discrete emotions (e.g. sad, scary, angry, excited etc.) will need to be achieved in the future. Besides that, we found some factors that influence accuracy, including placement of electrodes, number, gender, educational and cultural backgrounds of subjects, number of emotions considered, and type and effect of emotion induction

[37]. Therefore, in further studies we'd like to invite more participants of a similar cultural background to those already involved in our experiment, then try several different stimuli for triggering target emotions to obtain more sample data with the aim of working towards recognition of discrete emotions. Furthermore, we should try to perform the experiment with elderly participants in the future. Taking the accessibility and convenience of signals measurement of the elderly into account, attention would be placed on designing garment-type sensors for physical data collection instead of using active electrodes. We intend to incorporate a silver fiber in the fabric of garments to make a chest belt and head belt for more convenient and flexible data measurement.

#### ACKNOWLEDGMENT

This work was supported by JSPS KAKENHI (Grants-in-Aid for Scientific Research) Grant Number 15H01698.

#### REFERENCES

- [1] <http://www.un.org/en/sections/issues-depth/ageing/index.html>.
- [2] Alzheimer's Disease International. World Alzheimer report 2015: the global impact of dementia: an analysis of prevalence, incidence, cost and trends, 2015.
- [3] [http://www.who.int/mental\\_health/neurology/dementia/en/](http://www.who.int/mental_health/neurology/dementia/en/).
- [4] M. Poole, C. Bamford, E. McLellan, R. P. Lee, C. Exley, J. C. Hughes, .etl, "End-of-life care: A qualitative study comparing the views of people with dementia and family carers," *Palliat Med.*, vol. 32(3), pp. 631-642, March 2018.
- [5] S. Koelstra, C. Muhl, M. Soleymani, J. S. Lee, A. Yazdani, T. Ebrahimi, .etl, "DEAP: a database for emotional analysis using physiological signals," *IEEE Transactions on Affective Computing*, vol. 3(1), pp. 18-31, January-March 2012.
- [6] A. M. Badshah, J. Ahmad, N. Rahim, S. W. Baik, "Speech emotion recognition from spectrograms with deep convolutional neural network," 2017 International Conference on Platform Technology, February 2017.
- [7] S. E. Kahou, C. pal, X. Bouthillier, P. Froumenty, C. Gülçehre, R. Memisevic, "Combining modality specific deep neural networks for emotion recognition in video," *ICMI '13 Proceedings of the 15th ACM on International conference on multimodal interaction*, pp. 543-550. December, 2013.
- [8] H. Gunes, M. Piccardi, "A bimodal face and body gesture database for automatic analysis of human nonverbal affective," *IEEE 18th International Conference on Pattern Recognition*, August 2006.
- [9] Y. L. Hsu, J. S. Wang, W. C. Chiang, C. H. Hung, "Automatic ECG-based emotion recognition in music listening," *IEEE Transactions on Affective Computing*, pp. 1-1, December 2017.
- [10] H. J. Yoon, S. Y. Chung, "EEG-based emotion estimation using Bayesian weighted-log-posterior function and perceptron convergence algorithm," *Computers in biology and medicine*, vol. 43(12), pp. 2230-2237, 2013.
- [11] S. K. Hadjidimitriou, L. J. Hadjileontiadis, "Toward an EEG-based recognition of music liking using time-frequency analysis," *IEEE Trans Biomed Eng*, September 2012.
- [12] D. K. Prasad, S. J. Liu, S. H. A. Chen, C. Quek, "Sentiment analysis using EEG activities for suicidology," *Expert Systems with Applications*, vol. 103, pp. 206-217, August, 2018.
- [13] G. Acerbi, E. Rovini, S. Betti, A. Tirri, J. F. Rónai, A. Sirianni, "A wearable system for stress detection through physiological data analysis," *Italian Forum of Ambient Assisted Living*, vol. 426, pp. 31-50, 2016.
- [14] K. Schaaff, T. Schultz, "Towards an EEG-based emotion recognizer for humanoid robots," *The 18th IEEE International Symposium on Robot and Human Interactive Communication*, November 2009.
- [15] B. R. Lee, D. O. Won, K. S. Seo, H. J. Kim, S. W. Lee, "Classification of wakefulness and anesthetic sedation using combination feature of EEG and ECG," 2017 5th International Winter Conference on Brain-Computer Interface (BCI), February 2017.
- [16] R. W. Picard, E. Vyzas, J. Healey. "Toward machine emotional intelligence: Analysis of affective physiological state," *IEEE Transactionas on Pattern Analysis and Machine Intelligence*, vol. 23, pp. 1175-1191, 2001.
- [17] M. Murugappan, S. Murugappan, B. S. Zheng. "Frequency band analysis of electrocardiogram (ECG) signals for human emotional state classification using discrete wavelet transform (DWT)," *Journal of Physical Therapy Science*, vol. 25, pp. 753-759, 2013.
- [18] Z. Mohammadi, J. Frounchi, M. Amiri. "Wavelet-based emotion recognition system using EEG signal," *Neural Computing & Applications*, pp. 1-6, 2016.
- [19] S. Nasehi, H. Pourghassem. "An optimal EEG-based emotion recognition algorithm using gabor features," *Wseas Transactions on Signal Processing*, vol. 8, pp. 87-99, 2012.
- [20] Y. LeCun, Y. Bengio, G. Hinton, "Deep learning," *Nature*, vol. 521, pp. 436-444, 2015.
- [21] O. Tsinalis, P. M. Matthews, Y. Guo, S. Zafeiriou, "Automatic sleep stage scoring with single-channel EEG using convolutional neural networks," *arXiv.org Statistics Machine learning*, October 2016.
- [22] Y. F. Ren, Y. Wu, "Convolutional deep belief networks for feature extraction of EEG signal," 2014 International Joint Conference on Neural Networks (IJCNN), September 2014.
- [23] F. Agrafioti, D. Hatzinakos, A. K. Anderson, "ECG pattern analysis for emotion detection," *IEEE Transactions on Affective Computing*, vol. 3, pp. 102-115, August 2011.
- [24] M. Soleymani, J. Lichtenauer, T. Pun, M. Pantic, "A multimodal database for affect recognition and implicit tagging," *IEEE Transactions on Affective Computing*, vol. 3(1), pp. 42-55, Jan.-March 2012.
- [25] Y. L. Hsu, J. S. Wang, W. C. Chiang, C. H. Hung, "Autonomic ECG-Based Emotion Recognition in Music Listening," *IEEE Transactions on Affective Computing*, December 2017.
- [26] O. Martin, I. Kotsia, B. Macq, "The eNTERFACE'05 audio-visual emotion database," 22nd International Conference on Data Engineering Workshops, April 2006.
- [27] J. J. Gross, R.W. Levenson, "Emotion eliciting using films," *Cognition and Emotion*, vol. 9, pp. 87-108, 1995.
- [28] M. Soleymani, S. A. Esfeden, Y. Fu, M. Pantic, "Analysis of EEG signals and facial expressions for continuous emotion detection," *IEEE Transactions on Affective Computing*, vol. 7, pp. 17-28, Jan.-March 2016.
- [29] <https://datarecorder.jp/products/me/ap216.html>.
- [30] K. E. Ko, H. C. Yang, K. B. Sim, "Emotion recognition using EEG signals with relative power values and Bayesian network," *International Journal of Control, Automation and Systems*, October 2009.
- [31] D. O. Bos, "EEG-based emotion recognition," *The Influence of Visual and Auditory Stimuli*, 2006.
- [32] <https://docs.scipy.org/doc/scipy/reference/signal.html>.
- [33] S. E. Moon, S. Jang, J. S. Lee, "Learning cross-regional dependence of EEG with convolutional neural networks for emotion classification," *ICLR 2018 Workshop Acceptance Decision*, February 2018.
- [34] K. Simonyan, A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," *In ICLR*, pp: 1-14, April, 2015.
- [35] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, pp: 1097-1105, 2012.
- [36] <https://keras.io/layers/merge/#concatenate>.
- [37] J. Selvaraj, M. Murugappan, K. Wan, S. Yaacob, "Classification of emotional states from electrocardiogram signals: a non-linear approach based on hurst," *Biomed Eng Online*, vol. 16, pp. 12-44, May 2013.