A Self-attention-based Ensemble Convolution Neural Network Approach for Sleep Stage Classification with Merged Spectrogram

Chih-En Kuo[1*], Po-Yu Liao[2], and Yu-Syuan Lin[2]
[*1]National Chung Hsing University, Taichung, Taiwan
E-mail: cekuo@nchu.edu.tw Tel: +886-4-2284-0421#404
[2] Feng Chia University, Taichung, Taiwan.
E-mail: 388chris@mail.fcu.edu.tw, a0987209699@gmail.com

Abstract*— Scoring sleep data is a subjective and time-consuming. It takes more than one hour to score a whole night's PSG data. The automatic sleep staging method is needed to reduce clinical manpower. In this paper, an attention-based ensemble convolution neural network approach for sleep stage classification with merged spectrogram was proposed. All-night sleep physiological signals from 19 healthy individuals and 90 insomnia patients were used. First, the all-night polysomnography (PSG) signals including electroencephalogram (EEG), electrooculogram (EOG), and electromyogram (EMG) were segmented into 30-sec segments. Subsequently, each segment was transformed into spectrograms by continuous wavelet transform and a simple merged processing was applied to generate the merged spectrograms with different viewpoint. Next, the three merged spectrogram groups with different viewpoint were utilized as an input of our proposed CNN with self-attention, named merged spectrogram Net (MS-Net). The three trained MS-Net models were used to form an ensemble MS-Net. The experimental results showed that the accuracy, kappa coefficient, and F1 score of the proposed method were 89.83%, 84.82%, and 85.09%, respectively. The results proved that the proposed deep learning approach, ensemble MS-Net, had highly accuracy for sleep PSG spectrogram classification.[1]

I. INTRODUCTION

Sleep plays an important part in the consolidation of memories, learning, physical development, emotion regulation, and quality of life [1]. However, humans may suffer from various sleep disorders such as insomnia, which is the most common specific sleep disorder. In the general population, the prevalence rate of insomnia is 33% approximately [2]. To diagnose sleep disorder, all-night polysomnography (PSG) recordings are usually taken from the subjects, and well-trained staffs scored each epoch (i.e., 30-s data) into wakefulness (Wake), non-rapid eye movement (Non-REM) or rapid eye movement (REM) sleep stage according to the American Academy of Sleep Medicine (AASM) [3] rules.

However, diagnosing sleep disorders is time consuming and has a considerable workload [4]. From the patient perspective, they need to wait at least two months to record sleep signals at the sleep center. From the clinical staff perspective, the scoring process is time intensive, because the length of time to record sleep data is approximately 6 to 8 hours, and they have to manually analyze patients' sleep data for sleep scoring and annotating sleep-related events, which is a process of at least one hour.

In recent years, the deep learning has also been successfully applied to various types of physiological signals. Several automatic or aided sleep stage classification based on deep learning have been proposed, which used the convolution neural networks (CNNs), recurrent neural networks (RNNs) or combined CNN and RNN. The raw signals or their spectrograms can be as input of the CNN or the RNN, and they can automatically extract the pertaining features [5-14].

Thus, we aim to propose a method which used the merged spectrograms with different viewpoint to train three CNNs for sleep stage classification and the three CNN models were used to form an ensemble model for higher accuracy. In this paper, the 109 PSG recordings which are from 19 healthy individuals and 90 insomnia patients are used. First, the continuous wavelet transform was used to convert each 30-s EEG, EOG, and EMG segment into three spectrograms. The three spectrograms were merged as a single spectrogram (called original group) and used it to generate the horizontal-reflection spectrogram group and vertical-reflection spectrogram group, respectively. The three merged spectrogram groups with different viewpoint were utilized as an input of our proposed CNN with self-attention, named merged spectrogram Net (MS-Net). The three trained MS-Net models were used to form an ensemble MS-Net. Finally, the sleep architecture was classified by the ensemble MS-Net.

The main contributions of this work can be described as follows: (1) a simple merged and multi-viewpoint processing was proposed for build the ensemble CNN model. (2) The MS-Net with self-attention was proposed which with high accuracy for sleep PSG spectrogram classification. (3) The ensemble CNN that making decision by using a neural network to optimize the weights of each classifier in the

[1] This work was supported by Ministry of Science and Technology of Taiwan under Grants, MOST 109-2221-E-005 -080 -MY3 and 110-2634-F-006 -010. It was carried out at the AI Biomedical Research Center, Tainan, Taiwan.

ensemble model. (4) The results were improved by the ensemble CNN method with increases the accuracy, kappa coefficient, and F1 score by 6.48%, 9.81%, and 8.55%, respectively.

II. Materials and Methods

A. Materials

The subjects included 19 healthy individuals (sleep efficiency ≥ 85%) and 90 patients with insomnia (sleep efficiency < 85%). The patients with insomnia lasting more than three days per week for one month experienced drowsiness, sleepiness, irritable mood in daytime affecting learning and working. The insomnia patients have SE < 85%, sleep onset time > 15 min. and/or wake after sleep onset time >30 min. The 9 subjects from the healthy group and the 45 subjects from the insomnia group were used to generate the system, and the other 10 subjects from the healthy group and the other 45 subjects from the insomnia group were used for testing.

These measurements were approved by the internal review board of National Cheng Kung University. Subjects were recruited from the public by online advertisements and announcements on notice boards at National Cheng Kung University. Participants must avoid any drug/medication and limit caffeine use. The all-night physiological signals were measured in the sleep laboratory at the cognitive institute of National Cheng Kung University. The physiological signals included two electroencephalogram (EEG) channels (C3-A2 and C4-A1), one electrooculogram (EOG) channels (ROC-LOC), and a chin electromyogram (EMG) channel. The sampling rate was 256 Hz with 16-bit resolution. All 109 PSG sleep recordings were visually scored by two sleep specialists using the AASM guidelines with a 30-s interval (named an epoch). The epoch belonging movement (Mov) was extra annotated at the procedure of the manual scoring, because we want to make our model had an ability to distinguish Mov epoch. Table 1 shows the percentage of sleep stage of the healthy individuals and insomnia patients, respectively.

B. The proposed sleep stage classification

Fig. 1 shows the flowchart of the ensemble CNN model for sleep Staging with merged spectrogram. The process comprises two main steps: (1) Preprocessing and (2) Classification.

1) Preprocessing

The preprocessing process comprises three steps: (1) Segmented into 30-s epochs: An all-night PSG signal from EEG (C3-A2), EOG, and EMG was segmented into 30-s epochs with non-overlap. (2) Continuous wavelet transform (CWT): Each 30-s epoch was converted into three spectrograms by the CWT. Then, the three spectrograms were merged as a single spectrogram. (3) generate the horizontal-reflection spectrogram group and vertical-reflection spectrogram group from the original group.

Table 1
Percentage of sleep stage of the healthy individuals and insomnia patients

Subject	Total epoch	Wake (%)	N1 (%)	N2 (%)	N3 (%)	REM (%)	Mov (%)
Healthy	18,026	2.25	3.18	48.98	15.87	19.32	10.40
Insomnia	73,609	16.25	6.97	39.66	12.64	14.62	9.86
Total	91,635	13.49	6.22	40.52	13.27	15.54	10.96

The CWT has been successfully apply to other biomedical signal analysis [15, 16]. The CWT formula is shown in (1).

$$W(a,b)=\int_{-\infty}^{\infty} x(t)\frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right)dt \tag{1}$$

where $x(t)$ is a 30-s signal; $\psi(t)$ is a mother wavelet; a is a dilation parameter; b is a translation parameter. The spectrogram image size was set 224×224×3, since the input sizes of the proposed CNN. Fig. 2 shows the raw 30-s EEG and corresponding spectrogram at the five different sleep stages. The X-axis and Y-axis in the spectrogram represented time (sec) and frequency (Hz), respectively. The scale of the Y-axis is logarithmic. The range in time is 0-30 sec; the range in frequency is 0.11-113.11 Hz. The warm color regions in the spectrogram are represented which frequency band has more power. The characteristics of each sleep stage are clearly observed.

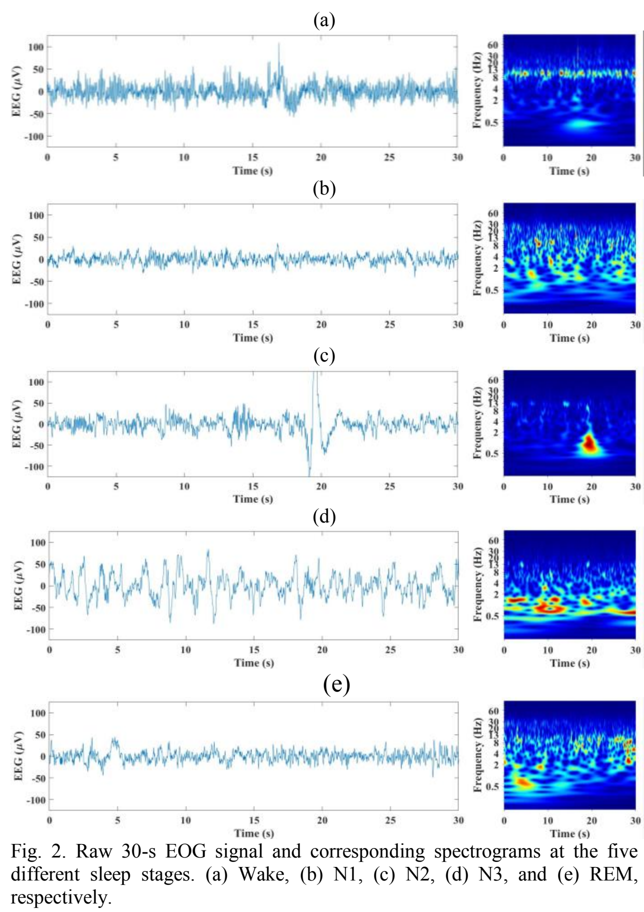

Fig. 2. Raw 30-s EOG signal and corresponding spectrograms at the five different sleep stages. (a) Wake, (b) N1, (c) N2, (d) N3, and (e) REM, respectively.

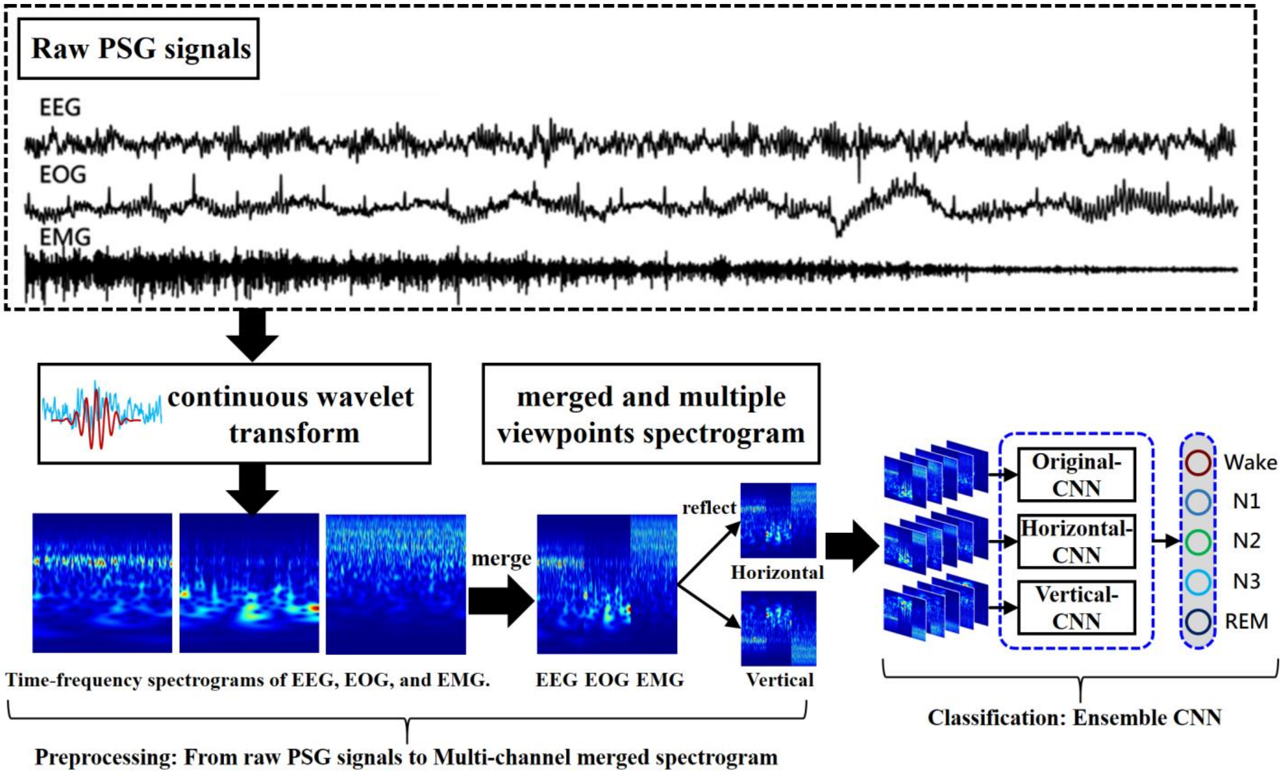

Fig. 1. Flowchart of the proposed ensemble CNN for sleep Staging with merged spectrogram.

2) Merged spectrograms with different viewpoint

Figs. 3, 4 (a) and 4 (b) represent the processing of the original merged spectrogram, horizontal-reflection merged spectrogram, and vertical-reflection merged spectrogram, respectively. In the processing of original merged spectrogram, the widths of EEG, EOG, EMG spectrograms were compressed to one-third of the original and merged as a single spectrogram along the horizontal direction.

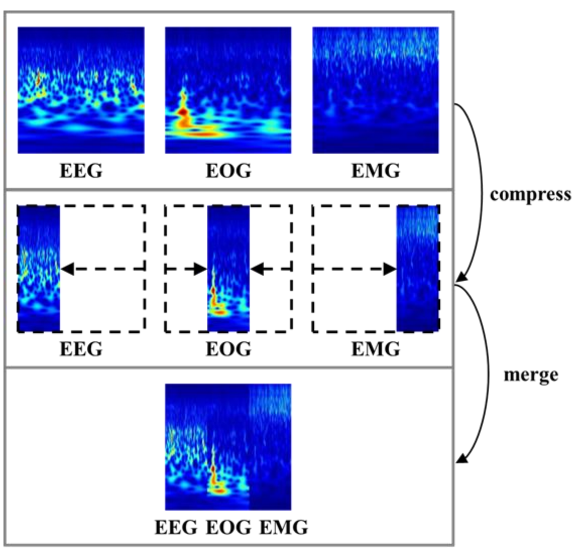

Fig. 3. The processing of the original merged spectrogram.

In the processing of horizontal-reflection merged spectrogram (see Fig. 4(a)), the EEG, EOG, EMG spectrograms were reflected along the horizontal direction firstly. Then, the remaining step was same as the original merged processing. The vertical-reflection merged spectrogram was obtained easily as shown in Fig. 4(b).

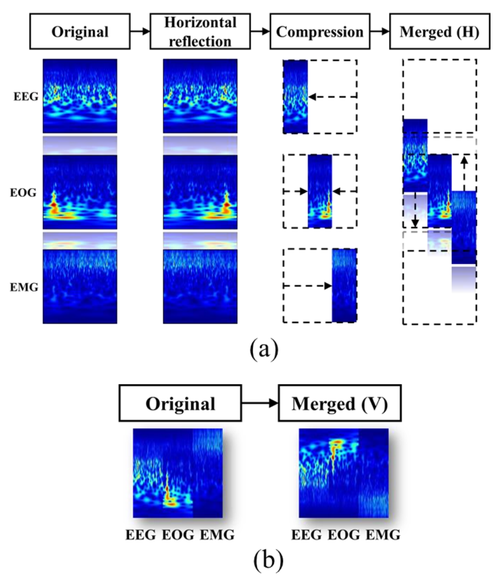

Fig. 4. The processing of (a) the horizontal-reflection merged and (b) vertical-reflection merged spectrogram.

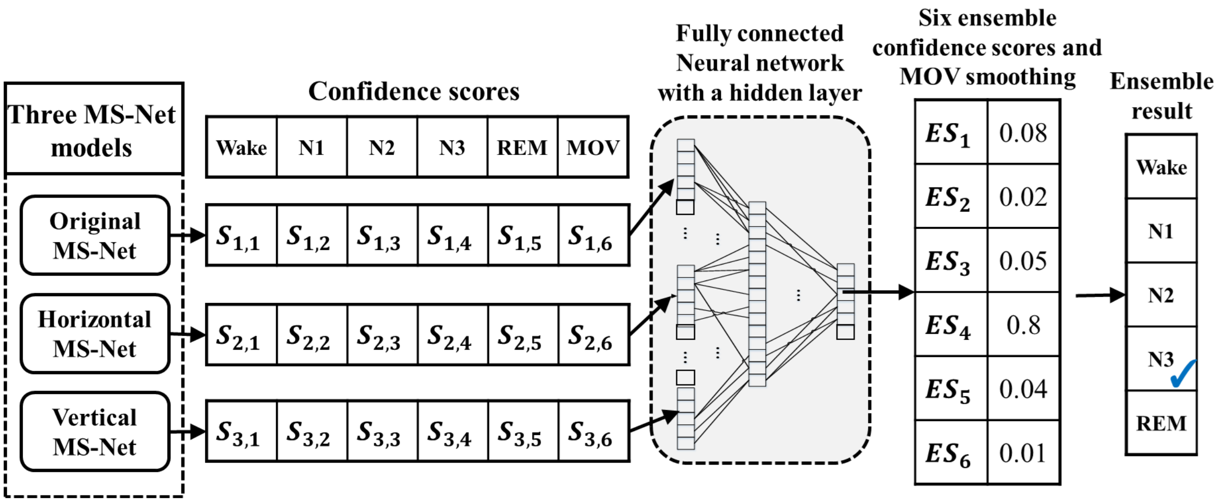

Fig. 6. The flowchart of the proposed ensemble CNN method.

3) Merged spectrogram Net (MS-Net)

Convolutional neural network (CNN) [17] is a class of deep neural network that is widely used for computer vision or analyzing visual imagery. The main advantage of CNN is that it automatically detects the important features without any human supervision or feature engineering. Moreover, CNNs have already been successfully applied to various types of physiological signals, including EEG recordings [7, 8]. CNN is a very good feature extractor for a completely new task or problem. The useful features or attributes were could extracted from an already trained CNN and tune the weights of the trained CNN a bit for the specific task by the new training data from new task or problem.

In this paper, we proposed a CNN model, named merged spectrogram Net (MS-Net) for merged spectrogram classification. The illustration of the proposed MS-Net as shown in Fig. 5. The MS-Net was mainly composed of multiple Residual-GC block. The illustration of the Residual-GC block was shown in Fig. 5(a), which is composed of two modules: residual learning [18] and global context (GC) block [19]. The residual learning and GC module were constructed with reference to ResNet [18] and global context networks [20], respectively. The advantage of GC block is that it has a self-attention mechanism which enable us to get a better feature maps.Fig. 5(b) represents the overall structure of the MS-Net model. The MS-Net contains 8 Residual-GC blocks.

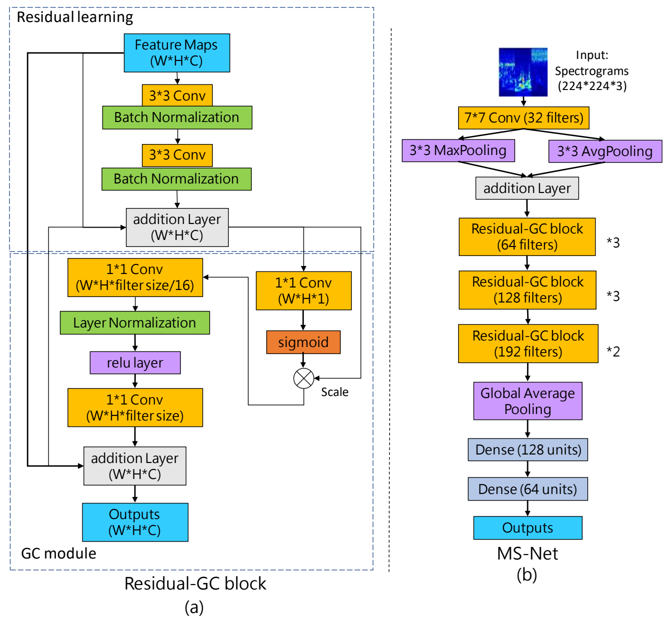

Fig. 5. Illustration of the proposed MS-Net. (a) illustrates the Residual-GC block, which is composed of two modules: residual learning and global context block. (b) represents the overall structure of the MS-Net model. The MS-Net contains 8 Residual-GC blocks.

4) Ensemble CNN method

When inputting a merged spectrogram to the MS-Net, it will output the confidence scores (probability values) that the spectrogram belongs to each sleep stage. Moreover, the sleep stage of the spectrogram will be determined by the highest confidence score. The flowchart of the proposed ensemble CNN method as shown in Fig. 6. The three trained MS-Net CNNs were used to form an ensemble model. The main strategy of the ensemble method is used a total of 18 (3*6) confidence scores from three trained CNNs as the input of the ensemble method and a fully connected neural network with a hidden layer (18 hidden neurons) was adopted to obtain the six ensemble confidence scores that corresponding to six sleep stages. In addition, the movement epochs from classification results were smoothing according to the AASM rules. If the Mov was adjacent to the Wake, Mov was revised to the Wake; If the Mov was adjacent to the non-Wake, Mov was revised to the same stages as its subsequent epochs.

C. Performance evaluation

We have used different metrics to evaluate the performance of the proposed method including, overall accuracy (Acc), sensitivity (Se), and F1 score. These metrics are defined as follows:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

$$Se = \frac{TP}{TP + FN} \tag{3}$$

$$PPV = \frac{TP}{TP + FP} \tag{4}$$

$$F_1 = \frac{2 \times (Se \times PPV)}{Se + PPV} \tag{5}$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative. In the two-class matrix, they indicate correctly classified, correctly rejected, incorrectly classified, and incorrectly rejected, respectively. Sensitivity (Se) represents that the proportion of actual positives that are correctly classified. Positive predictive value (PPV) represents that the proportion of predictive positives that are correctly classified. In addition, Cohen's kappa coefficient [21] was also calculated for each subject to evaluate the agreement between clinical staff and the proposed method. Cohen's kappa coefficient is a statistical measure of interrater agreement among two or more raters.

III. RESULTS

A. Experiment setups

There are two main steps in our experiment: step (1) train and evaluate the proposed MS-Net model; step (2) the three trained MS-Net models to form the ensemble CNN model. Then, the ensemble CNN model was evaluated after the fully connected neural network that in the ensemble CNN model was trained. Half of the subjects (*N*=54) where from healthy and insomnia group were respectively and randomly selected for training single MS-Net and the ensemble CNN, and the others (*N*=55) were used to test single MS-Net and the ensemble CNN performance. Adaptive moment estimation (Adam) optimizer were used to train all models. The learning rate which initial value is 0.001 was piecewise dropped based on a multiplicative factor after each training epoch. The computer program that we used to develop and evaluate the proposed method is MATLAB (version: 2021a including deep learning toolbox).

B. The performance of MS-Net and ensemble MS-Net

The evaluation metrices of the classification result between experts were shown as Table 2. The accuracy, kappa, F1 score, between the three single MS-Net were very close, and their average accuracy, kappa, F1 score were 83.28%, 75.01%, and 76.42%, respectively. The accuracy of single MS-Net was over the average agreement between the experts (82.6%) [22]. This shows that the basic classifier of our ensemble method already has good accuracy. The best accuracy, kappa, F1 score, and Se of each five stage were the ensemble CNN method by using the merged spectrograms with different viewpoint. The accuracy, kappa coefficient, and F1 score of the proposed method were 89.83%, 84.82%, and 85.09%, respectively. The accuracy, kappa, and F1 score improved by an average of 6.55%, 9%, and 8.38% after using ensemble CNN, respectively. It was proved that the ensemble CNN can make the results closer expert scoring.

Table 2
Evaluation metrices representing the classification result between experts and Meraged (original, horizontal, and vertical) and ensemble CNN

	Acc.	Kappa	F1 score	Se Wake	Se N1	Se N2	Se N3	Se REM
Merged (Original)	83.35	75.01	76.54	87.37	40.71	89.45	76.95	83.90
Merged (Horizontal)	83.31	75.01	76.46	87.37	40.30	89.39	77.79	83.45
Merged (Vertical)	83.18	75.01	76.25	87.00	40.57	89.47	76.30	83.72
Merged+ Ensemble	**89.83**	**84.82**	**85.09**	**91.40**	**54.01**	**94.88**	**80.23**	**93.26**

Table 3 shows the literatures related to the automatic sleep scoring in recent years and the proposed ensemble CNN method. These literatures used machine learning and deep learning to classify sleep stages. Pankaj et. al. [23] used SqueezeNet to automatically learn the features from spectrogram and classify. Akara et. al. [10] used the network combined CNN and long short-term memory to learn features from raw data. In addition to Ref. [24, 25] and our proposed method, these automatic sleep scoring systems were developed by using healthy individuals in Table 3. However, the sleep data recorded from the patients with sleep disorder is even more often seen in clinic. The automatic sleep scoring system should be developed by the patients' dataset such as SHHS, instead of only healthy individual. Wu et. al. [26] has a high accuracy, but the validation method is not subject-independent. Therefore, our proposed method was higher robustness and generality than Wu et. al. [26]. In addition, we will also use a larger number of clinical public data sets such as PhysioNet Challenge 2018 [27] to verify the effectiveness of our proposed method.

IV. CONCLUSIONS

In this paper, an automatic sleep scoring that is classified sleep stages using ensemble CNN was developed. First, the merged spectrograms were generated by 30-s PSG signals using the CWT. Next, the proposed MS-Net were trained using 54 sleep data recorded from healthy individuals and insomnia patients for sleep stage classification task. The three trained MS-Net models were used to form an ensemble model with using a neural network to optimize the weights of each classifier in the ensemble model. The accuracy, kappa coefficient, and F1 score of the proposed method were 89.83%, 84.82%, and 85.09%, respectively. The limitation of this study is the proposed method only evaluated by healthy individuals and insomnia patients. The patients with other sleep disorder such as sleep apnea or PLM should need further evaluation in the future.

Table 3
Literatures related to the automatic sleep scoring in recent years and the proposed method

Ref.	dataset	No. of subjects	Method	Validation	Acc (%)	Kappa	F1 (%)	Sensitivity (%)				
								Wake	N1	N2	N3	REM
[10]	Sleep-EDF	20 (H [a])	raw EEG + CNN-LSTM	20-fold	82.0	--	--	83.4	50.1	81.7	94.2	83.9
[28]	Sleep-EDF	20 (H)	raw EEG + 1dCNN-OCRNN	hold-out (9:1)	82.4	--	--	85.2	21.6	82.9	87.5	91.4
[29]	unpublish	265 (D [b])	raw EEG, EOG, and EMG + 1dCNN	hold-out (35:8:57)	83.6	0.77	78.1	86.8	46.7	90.6	39.2	91.7
[30]	MASS	147 (H)	hand-crafted features + raw EEG, EOG, and EMG + Bi-LSTM	LOSO	87.8	0.82	81.8	89.1	43.5	92.7	83.4	93.9
[26]	Sleep-EDF	99 (H)	hand-crafted features + SVM	10-fold	93.1	0.84	73.9	99.1	37.26	88.3	82.9	75.6
[24]	SHHS	5728 (H and D)	DWT + LDA	hold-out (5:2:3)	87.2	0.81	84.0	90.4	34.4	87.1	84.5	84.22
[12]	Sleep-EDF	22 (H)	raw EEG + CNN-LSTM	20-fold	84.3	--	--	90.6	54.5	82.7	88.9	88.7
[23]	Sleep-EDF	42 (H)	time-frequency spectrogram + SqueezeNet	hold-out (7:1:2)	84.7	--	--	91.8	34.7	91.7	78.2	88.9
[2]	unpublish	20 (H)	MSE and AR coefficients of EEG + LDA	hold-out (1:1)	88.1	0.81	--	86.3	28.5	88.1	86.7	97.6
[31]	unpublish	16 (H)	MSE and AR coefficients of EOG + LDA	hold-out (1:1)	84.3	0.77	--	83.8	42.9	83.0	81.3	94.1
[25]	unpublish	32 (H and D)	raw EEG and EMG, hand-crafted features + genetic fuzzy inference system	2-fold	86.4	0.81	--	86.8	35.2	89.6	88.6	88.2
Ours (ensemble CNN)	unpublish	109 (H and D)	time-frequency spectrogram + ensamble CNN	hold-out 1:1	89.8	0.85	85.1	93.2	52.9	94.8	80.9	93.5

[a]H= healthy individuals; [b]D= patients with sleep disorder.

REFERENCES

[1] I. Capellini, P. McNamara, B. T. Preston, C. L. Nunn, and R. A. Barton, “Does sleep play a role in memory consolidation? A comparative test,” *PLoS One,* vol. 4, no. 2, pp. e4609, 2009.

[2] S.-F. Liang, C.-E. Kuo, Y.-H. Hu, Y.-H. Pan, and Y.-H. Wang, “Automatic stage scoring of single-channel sleep EEG by using multiscale entropy and autoregressive models,” *IEEE Transactions on Instrumentation and Measurement,* vol. 61, no. 6, pp. 1649-1657, 2012.

[3] R. B. Berry, R. Brooks, C. E. Gamaldo, S. M. Harding, C. Marcus, and B. V. Vaughn, “The AASM manual for the scoring of sleep and associated events,” *Rules, Terminology and Technical Specifications, Darien, Illinois, American Academy of Sleep Medicine,* vol. 176, 2012.

[4] A. Roebuck, V. Monasterio, E. Gederi, M. Osipov, J. Behar, A. Malhotra, T. Penzel, and G. Clifford, “A review of signals used in sleep analysis,” *Physiological measurement,* vol. 35, no. 1, pp. R1, 2013.

[5] C.-E. Kuo, and G.-T. Chen, “Automatic sleep staging based on a hybrid stacked LSTM neural network: verification using large-scale dataset,” *IEEE Access,* vol. 8, pp. 111837-111849, 2020.

[6] S. Biswal, H. Sun, B. Goparaju, M. B. Westover, J. Sun, and M. T. Bianchi, “Expert-level sleep scoring with deep neural networks,” *Journal of the American Medical Informatics Association,* vol. 25, no. 12, pp. 1643-1650, 2018.

[7] A. Patanaik, J. L. Ong, J. J. Gooley, S. Ancoli-Israel, and M. W. Chee, “An end-to-end framework for real-time automatic sleep stage classification,” *Sleep,* vol. 41, no. 5, pp. zsy041, 2018.

[8] H. Phan, F. Andreotti, N. Cooray, O. Y. Chén, and M. De Vos, “SeqSleepNet: end-to-end hierarchical recurrent neural network for sequence-to-sequence automatic sleep staging,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 27, no. 3, pp. 400-410, 2019.

[9] A. Supratak, and Y. Guo, "TinySleepNet: An Efficient Deep Learning Model for Sleep Stage Scoring based on Raw Single-Channel EEG." pp. 641-644.

[10] A. Supratak, H. Dong, C. Wu, and Y. Guo, “DeepSleepNet: A model for automatic sleep stage scoring based on raw single-channel EEG,” *IEEE Transactions on Neural Systems and Rehabilitation Engineering,* vol. 25, no. 11, pp. 1998-2008, 2017.

[11] L. Fiorillo, M. Wand, I. Marino, P. Favaro, and F. D. Faraci, "Temporal dependency in automatic sleep scoring via deep learning based architectures: An empirical study." pp. 3509-3512.

[12] S. Mousavi, F. Afghah, and U. R. Acharya, “SleepEEGNet: Automated sleep stage scoring with sequence to sequence deep learning approach,” *PloS one,* vol. 14, no. 5, pp. e0216456, 2019.

[13] A. Craik, Y. He, and J. L. Contreras-Vidal, “Deep learning for electroencephalogram (EEG) classification tasks: a review,” *Journal of neural engineering,* vol. 16, no. 3, pp. 031001, 2019.

[14] O. Faust, H. Razaghi, R. Barika, E. J. Ciaccio, and U. R. Acharya, “A review of automated sleep stage scoring based on physiological signals for the new millennia,” *Computer methods and programs in biomedicine,* vol. 176, pp. 81-91, 2019.

[15] M. Wacker, and H. Witte, “Time-frequency techniques in biomedical signal analysis,” *Methods of information in medicine,* vol. 52, no. 04, pp. 279-296, 2013.

[16] R. Boostani, F. Karimzadeh, and M. Nami, “A comparative review on sleep stage classification methods in patients and healthy individuals,” *Computer methods and programs in biomedicine,* vol. 140, pp. 77-91, 2017.

[17] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks." pp. 1097-1105.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition." pp. 770-778.

[19] Y. Cao, J. Xu, S. Lin, F. Wei, and H. Hu, "Global Context Networks," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1-1, 2020.

[20] J. Hu, L. Shen, and G. Sun, "Squeeze-and-excitation networks." pp. 7132-7141.

[21] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and psychological measurement,* vol. 20, no. 1, pp. 37-46, 1960.

[22] R. S. Rosenberg, and S. Van Hout, "The American Academy of Sleep Medicine inter-scorer reliability program: sleep stage scoring," *Journal of clinical sleep medicine,* vol. 9, no. 1, pp. 81-87, 2013.

[23] P. Jadhav, G. Rajguru, D. Datta, and S. Mukhopadhyay, "Automatic sleep stage classification using time–frequency images of CWT and transfer learning using convolution neural network," *Biocybernetics and Biomedical Engineering,* vol. 40, no. 1, pp. 494-504, 2020.

[24] Y. Liu, J. Gao, W. Cao, L. Wei, Y. Mao, W. Liu, W. Wang, and Z. Liu, "A hybrid double-density dual-tree discrete wavelet transformation and marginal Fisher analysis for scoring sleep stages from unprocessed single-channel electroencephalogram," *Quantitative Imaging in Medicine and Surgery,* vol. 10, no. 3, pp. 766, 2020.

[25] S.-F. Liang, C.-E. Kuo, F.-Z. Shaw, Y.-H. Chen, C.-H. Hsu, and J.-Y. Chen, "Combination of expert knowledge and a genetic fuzzy inference system for automatic sleep staging," *IEEE Transactions on Biomedical Engineering,* vol. 63, no. 10, pp. 2108-2118, 2015.

[26] W. Huang, B. Guo, Y. Shen, X. Tang, T. Zhang, D. Li, and Z. Jiang, "Sleep staging algorithm based on multichannel data adding and multifeature screening," *Computer Methods and Programs in Biomedicine,* vol. 187, pp. 105253, 2020.

[27] A. L. Goldberger, L. A. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *circulation,* vol. 101, no. 23, pp. e215-e220, 2000.

[28] F. Zhu, and Q. Liang, "OCRNN: An Orthogonal Constrained Recurrent Neural Network for Sleep Analysis Based on EEG Data," *Ad Hoc Networks*, pp. 102178, 2020.

[29] X. Zhang, M. Xu, Y. Li, M. Su, Z. Xu, C. Wang, D. Kang, H. Li, X. Mu, and X. Ding, "Automated multi-model deep neural network for sleep stage scoring with unfiltered clinical data," *Sleep and Breathing*, pp. 1-10, 2020.

[30] C. Sun, C. Chen, W. Li, J. Fan, and W. Chen, "A Hierarchical Neural Network for Sleep Stage Classification Based on Comprehensive Feature Learning and Multi-Flow Sequence Learning," *IEEE Journal of Biomedical and Health Informatics,* vol. 24, no. 5, pp. 1351-1366, 2019.

[31] S.-F. Liang, C.-E. Kuo, Y.-C. Lee, W.-C. Lin, Y.-C. Liu, P.-Y. Chen, F.-Y. Cherng, and F.-Z. Shaw, "Development of an EOG-based automatic sleep-monitoring eye mask," *IEEE Transactions on Instrumentation and Measurement,* vol. 64, no. 11, pp. 2977-2985, 2015.