

Applying Time-Frequency Image of Convolutional Neural Network to Extract Feature on Long-Term EEG Signals to Predict Depth of Anesthesia

Yu-Po Huang, Jerry Chen, Shou-Zen Fan, Maysam F. Abbod, Jiann-Shing Shieh and Yu-Chen Kung

Abstract—In this study, EEG signals were converted using continuous Wavelet transform (CWT) and short-term Fourier transform (STFT) into time-frequency images as input to the convolutional neural network. According to Bi-Spectral (BIS) index and signal quality indicator (SQI) of commercial machines, anesthetic state can be classified as anesthetic light (AL), anesthetic ok (AO), anesthetic deep (AD), and Noise. The EEG signal is converted into an image every 5 seconds as well as 2 minutes period. The 5 seconds images dataset was generated from 13 patients as reported in a previous study which is compared to current study that is based on 2 minutes images dataset generated 55 patients. As a result, the 5 seconds EEG CWT image model predicts an accuracy of the individual categories of: AL is 69%, AO is 75%, AD is 73%, and Noise is 50%. The overall accuracy of the model is 72.13%. However, the 2 minutes EEG CWT images model predicts an accuracy of the individual categories of: AL is 81%, AO is 86%, AD is 91%, and Noise is 59%. The overall accuracy of the model is 85.62%. In addition, the 2 minutes EEG STFT image model predicts the accuracy of individual categories of AL is 82%, AO is 85%, AD is 92%, and Noise is 52%. The overall accuracy of the model is 84.71%. The result shows that the 2 minutes images model is better than the 5 seconds images model. Therefore, ten patients were randomly selected from the data of 55 patients as test data. The test results show an overall accuracy of 92.5% and 87.85% for the CWT image model and the STFT image model. In conclusion, the 2 minutes EEG CWT image model is the best model for this study.

Index Terms—Electroencephalogram (EEG), Continuous wavelet transform (CWT), short-term Fourier transform (STFT), Convolutional neural networks (CNN), Anesthesia,

I. INTRODUCTION

CONVOLUTION neural network (CNN) have been widely used in various fields in recent years. One of the most important advances in artificial intelligence learning is

Manuscript received March 12, 2019; revised April 03, 2019.

Yu-Po Huang, Jerry Chen, and Jiann-Shing Shieh are with the Department of Mechanical Engineering, Yuan Ze University, Taoyuan, Chung-Li 32003 Taiwan (corresponding author to provide phone: +886-3-4638800 ext. 2470; fax: +886-3-4558013; email: agg2f6fg@gmail.com, jerry423324@gmail.com, and jsshieh@saturn.yzu.edu.tw).

Shou-Zen Fan was with Department of Anesthesiology, College of Medicine, National Taiwan University, Taipei 100, Taiwan (e-mail: shouzen@gmail.com).

Maysam F. Abbod was with Department of Electronic and Computer Engineering, Brunel University London, Uxbridge, UB8 3PH, UK (e-mail: maysam.abbod@brunel.ac.uk).

Yu-Chen Kung is with Lenovo Global Technology Ltd, Taipei 115, Taiwan (e-mail: ekung@lenovo.com).

deep learning. In addition, deep learning has been found to be mainly applied to dimensionality learning and multi-layer feature learning [1]. Therefore, CNN is a method based on multi-layer multi-dimensional convolution of images and extraction of specific features from images.

EEG-related CNN methods in biomedical engineering are mainly focused on epilepsy [2], emotion recognition [3], sleep [4], and motor imagining [5]. However, the CNN method is rarely used in EEG studies to assess the depth of anesthesia (DOA) in patients. It is important for patient DOA monitoring during general anesthesia surgery. If the degree of anesthesia is too shallow during the procedure, the patient will have a slight awareness or feel a slight pain resulting in some postoperative memory impairment [6]. Moreover, long-term maintenance of deep anesthesia can lead to other complications in patients, so anesthesia management is very important [7]. Currently, many products for DOA monitoring have been developed on the market. For instance, Mid-latency auditory evoked potential (MLAEP) [8], a method of stimulating auditory response under general anesthesia and then evaluating the state of EEG, spectral entropy (SpE) [9] monitors DOA by calculation the state entropy (SE) and response entropy (RE) of the patient's EEG, or bi-spectral (BIS) [10] is an indicator obtained by calculating image bit coupling of the frequency of an EEG signal. However, these methods are calculated by complex numerical algorithms.

In the past research, processing techniques from numerical conversion to image have been developed and can be divided into time domain, frequency domain and time-frequency domain images. EEG inputs as CNN typically use time-domain and time-frequency images, where time-domain images are usually drawn with raw data, recurrent plot, and brain computer interface (BCI), while time-frequency domain images typically use continuous wavelet transforms (CWT), short-time Fourier transform (STFT), Hilbert-Huang transform (HHT) and Wigner-Ville distribution (WVD). Time-frequency techniques are a suitable method for EEG such nonlinear and unsteady signals, while providing instantaneous information about the frequency and intensity of brainwave activity [11]. Studies have shown that brainwave activity at different frequencies represents a different phenomenon [12]. Anesthetics reduce the activity of high frequency beta and alpha bands during induction of anesthesia and increased activity in the low frequency band during deep anesthesia [13]. Therefore, it can be observed

that EEG shows changes in signal characteristics and activity intensity during the anesthesia induction period, the maintenance period and the recovery period in the time-frequency domain image.

Although the instantaneous frequency and intensity of the EEG can be obtained from time-frequency images, the EEG waveform do not have significant waveform characteristics of physiological signals such as ECG and PPG, and thus do not have clear and highly reproducible identifiable patterns. In this study, CWT and STFT are used to convert EEG signals to time-frequency images as CNN input images. In addition, the CWT method will have short-term and long-term EEG signals converted into images for training. The images will be classified according to the BIS values given by the Philips patient monitor. The purpose of the study was to evaluate the depth of anesthesia in a simpler way by CNN.

II. MATERIALS AND METHODS

A. Anesthesia dataset

In this study, two sets of data were prepared. The patients were all from the National Taiwan University Hospital (NTUH) in Taiwan. Before the surgery began, patients were injected with propofol to make the patient into an unconscious state. The first group used desflurane anesthesia from 13 patients during anesthesia maintenance in our previous study [14] and the second group used 55 patients with sevoflurane anesthetics during anesthesia maintenance. 55 patient's datasets surgery type selection has surgery for uterus and ovarian diseases. The signal channel 128 Hz EEG signal is recorded by the BIS™ Quatro Sensor of the MP60 as well as record BIS and the signal quality indicators (SQI) every 5 seconds. The first group will convert the image every 5 seconds, so it is divided into four categories based on the value of the machine every 5 seconds: Anesthetic Light (AL) ($100 \geq \text{BIS} \geq 60$, $\text{SQI} \geq 50$), Anesthetic Ok (AO) ($60 > \text{BIS} \geq 40$, $\text{SQI} \geq 50$), Anesthetic Deep (AD) ($40 \geq \text{BIS} \geq 0$, $\text{SQI} \geq 50$), and Noise ($\text{SQI} < 50$). The second group will convert the

image every 2 minutes but the BIS and SQI will have 24 record points. Therefore, the classification decision for this category will determine the BIS level based on the median, while the SQI is less than 50, and more than 25% of the total record point will be classified into the Noise category. As CNN training, validation and testing, the first and second groups will be 70%-15%-15% and 70%-20%-10%, respectively. The data on anesthesia maintenance period during general anesthesia were the most abundant, as shown in Table I. The AO and AD of the first group accounted for 45% and 44% of the data, while the AO and AD of the second group accounted for 47% and 36%. There is very little data that can be received during the anesthesia induction period and the anesthesia recovery period, with the first group of patients and the second group of patients accounting for only 6% and 10%, respectively. The phenomenon of disturbance during the operation is not frequent, so the first group and the second group account for only 5% and 7% of the noise category.

B. Data Preprocessing

Since 2D CNN was applied to our research, the raw EEG signal was converted to an image and reflected the state of the DOA. First, the first group of patients' data is converted into images by CWT every 5 seconds as shown in our previous study [14]. The data of the second group of patients was converted into images by CWT and STFT every two minutes. The second group of data is updated once every 30 seconds in the processing of time, so there will be 75% overlap of the images before and after. The 2 min CWT image is the same as the 5s image processing method. The original EEG data converted as STFT will only be filtered once to retain the signal in the frequency range of 0.5-30 Hz. 5 s CWT image, 2 min CWT, and STFT image AL category are shown Figs. 1 (a), (e), and (i). AO category are shown in Figs. 1 (b), (f), and (j). AD categories are shown in Fig. 1 (c), (g), and (k). Noise category are shown in Figs. 1 (d), (h), and (l). Noise images usually occur when the patch is lost, or people touch the signal transmission line.

Table I. Data distribution for each category.

13 patients				
Category	Training (70%)	Validation (15%)	Testing (15%)	Data distribution
Anesthetic Light (AL)	812	173	173	6%
Anesthetic Ok (AO)	5777	1238	1238	45%
Anesthetic Deep (AD)	5715	1224	1224	44%
Noise	689	147	147	5%
Total	12993	2782	2782	100%
55 patients				
Category	Training (70%)	Validation (20%)	Testing (10%)	Data distribution
Anesthetic Light (AL)	1149	328	164	10%
Anesthetic Ok (AO)	5404	1544	772	47%
Anesthetic Deep (AD)	4174	1192	596	36%
Noise	820	232	116	7%
Total	11547	3296	1648	100%

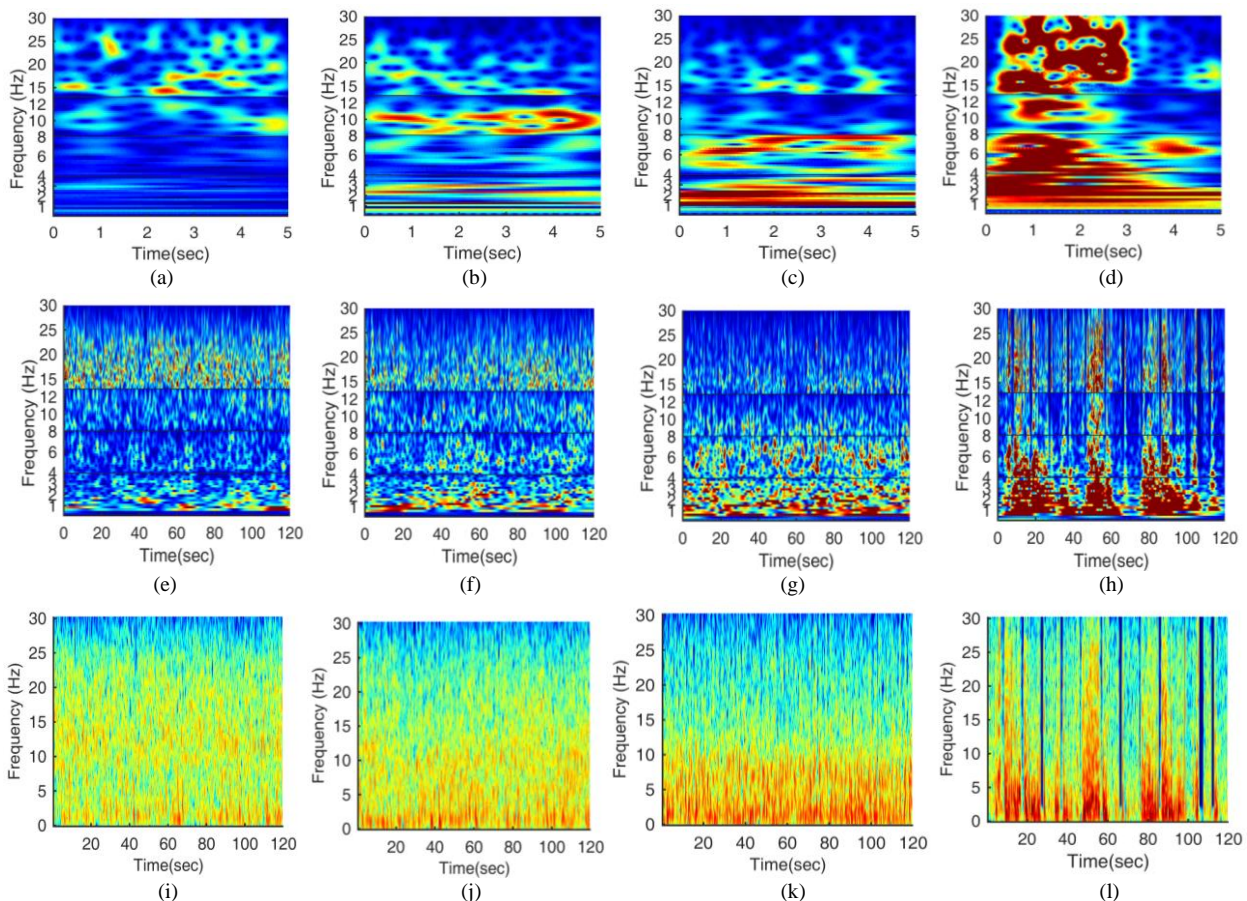


Fig. 1. Time-frequency image: (a) 5 seconds CWT image of Anesthetic Light, (b) 5 seconds CWT image of Anesthetic Ok, (c) 5 seconds CWT image of Anesthetic Deep, (d) 5 seconds CWT image of Noise, (e) 2 minutes CWT image of Anesthetic Light, (f) 2 minutes CWT image of Anesthetic Ok, (g) 2 minutes CWT image of Anesthetic Deep, (h) 2 minutes CWT image of Noise, (i) 2 minutes STFT image of Anesthetic Light, (j) 2 minutes STFT image of Anesthetic Ok, (k) 2 minutes STFT image of Anesthetic Deep, (l) 2 minutes STFT image of Noise

C. Model and Operating Environment

CNN is a deep learning algorithm that uses multiple layers of convolution to extract image features. In this study, we cooperate with Lenovo's branch office in Taiwan. They provide high-performance computer (HPC) equipment for us to use. The server system and GPU devices are Lenovo ThinkSystem SD530 and two Volta 100 GPUs as shown in Fig. 2. Firstly, you need to login from your local computer and connect to Lenovo's HPC input command training model. Finally, you can download the model to your local computer. When choosing the CNN model structure, we chose the 8-layer shallow Alexnet model that won the championship in the ImageNet competition. The model structure is as shown in Fig. 3. No studies have shown which models and hyperparameter adjustments are appropriate for EEG studies, even if there are many EEG-related studies. In addition, considering that brain waves are quite complicated signals, each patient's brain waves for anesthetic drugs and other drugs (i.e. Ketamine, painkiller) will show different responses. The number of patient samples is not enough, so do not choose more complicated and deeper network structure. In

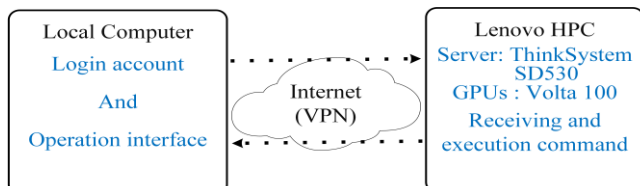


Fig. 2. Operating procedures and hardware devices

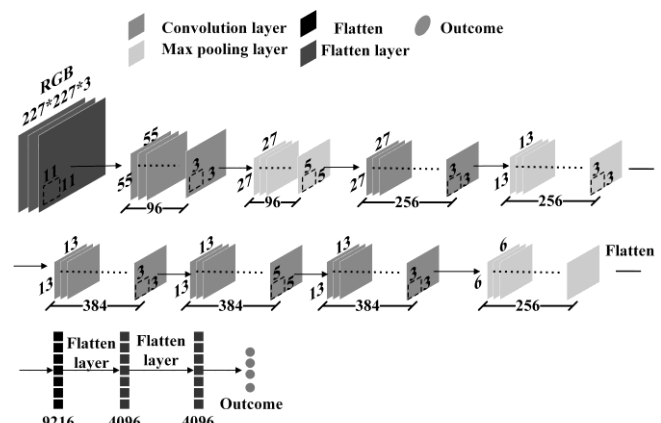


Fig. 3. Alexnet structure.

this study, the second group of model tests, in addition to using the original test set to test the model. Randomly selected 10 patients from 55 patients to test the model. The reason is to test the model prediction of the patient's EEG in real surgery.

III. RESULTS

The first group of data is entered at the beginning of CNN batch size 128 and 100 epoch in our previous study [14]. The second group of data sets the batch size 128 and 150 epoch. The first group and the second group of data were trained 12,993 and 11,547 images respectively compressed to 227×227 as the initial input. It takes 3.5 hours for the 5 s CWT image model trained by the first group of data. The 2 min

CWT and STFT image models trained by the second set of data takes 5.2 hours and 4.7 hours. The model test results are shown in Table II. The results show that the accuracy of the 5s CWT image model in AL, AO, AD, and Noise is 69%, 75%, 73%, and 50%, respectively. The overall accuracy of the model is 72.36%. Moreover, it can be seen that the 5s CWT image model predicts AL images and is predicted to be AD only 2 images are predicted. The results of the 2min CWT image model and the STFT image model show that the accuracy in AL, AO, AD, and Noise are 81%, 86%, 91%, 59% and 82%, 85%, 92%, 52%. The 2min CWT image model and STFT image overall accuracy is 85.62% and 84.71%. In the 2min image model, only one image was predicted to be AD when AL was predicted, and no image was predicted as AL when AD was predicted. The results show that the 2min Image model not only has accuracy in all categories, but also overall accuracy is higher than 5s CWT model. The overall accuracy of the 2min CWT image model is slightly higher than the STFT image model.

In order to test the prediction results of the 2min image model in the complete surgical procedure, 10 patient's data test models were randomly selected from 55 patients. Table III shows that the accuracy of the 2 min CWT image model in AL, AO, AD, and Noise is 96%, 90%, 97%, and 74%. The accuracy of the 2min STFT model in AL, AO, AD, and Noise is 92%, 85%, 96%, and 74%. The 2 min CWT image model and STFT model overall accuracy is 92.5% and 87.85%. It can be seen from Table III that when the AL is recognized, no image is predicted as AD, and when AD is recognized, only one image is predicted to be AL. The 2min CWT image model overall accuracy is significantly higher than the STFT image model. In addition, the 2 min CWT image prediction for Noise is much higher than the STFT model.

Table II. Model testing result.

5s CWT image model						
		Predict result				Individual accuracy
		AL	AO	AD	Nosie	
True data	AL	120	37	2	15	69%
	AO	66	927	193	51	75%
	AD	25	240	892	67	73%
	Noise	23	27	23	74	50%
Overall accuracy is 72.36%						
2min CWT image model						
		Predict result				Individual accuracy
		AL	AO	AD	Nosie	
True data	AL	133	27	1	3	81%
	AO	22	667	73	10	86%
	AD	0	49	543	4	91%
	Noise	13	21	14	68	59%
Overall accuracy is 85.62%						
2min STFT image model						
		Predict result				Individual accuracy
		AL	AO	AD	Nosie	
True data	AL	134	24	1	5	82%
	AO	28	655	77	12	85%
	AD	0	43	547	6	92%
	Noise	15	25	16	60	52%
Overall accuracy is 84.71%						

Table III. Whole surgical procedure data test results for 2 min image model.

2min CWT image model						
		Predict result				Individual accuracy
		AL	AO	AD	Nosie	
True data	AL	510	19	0	3	96%
	AO	44	1211	75	15	90%
	AD	1	30	1172	5	97%
	Noise	14	36	5	153	74%
Overall accuracy is 92.50%						
2min STFT image model						
		Predict result				Individual accuracy
		AL	AO	AD	Nosie	
True data	AL	490	37	1	4	92%
	AO	54	1144	137	10	85%
	AD	0	41	1159	8	96%
	Noise	21	64	23	100	48%
Overall accuracy is 87.85%						

IV. DISCUSSION AND CONCLUSION

The 5s image model test results show that the categories are confused, and the 2min image model can distinguish between the AL and AD categories. That's because short-term EEG images do not contain EEG changes in the previous period of time and long-term EEG images contain EEG changes in patients from previous periods. Furthermore, long-term EEG-converted images provide long-term stable phenomena characteristic on different anesthesia stages. However, the 2min image model still cannot effectively distinguish the images of the AO and Noise category. Previous tests were based on a scattered data test model. However, in real surgical cases, there is only one patient in the moment, and the data has continuous complete data rather than scattered and lost data. 10 patients randomly were selected from the second group of data were trained with 70% of the images but 30% of the images were still not trained. From the results, the STFT model is close to 90% and the CWT model has exceeded 90%. In future work, we need to perform cross-validation to test the generalization of the model. The process of brain wave changes during anesthesia is not too fast so the model categories will be divided into ten categories in a more subtle way. The structural changes in the model can be added to the recurrent convolutional neural network (RCNN) or CNN-LSTM. The network structure that takes into account the concept of time allows the model to take into account the anesthetic state of the previous period and improve the accuracy.

ACKNOWLEDGMENT

This research was financially supported by the Lenovo Technology B.V. Taiwan Branch. Also, it was supported by Ministry of Science and Technology (Grant number: MOST 107-2221-E-155 -009 -MY2).

REFERENCES

- [1] G. Hinton and R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science Magazine*, 2006, vol. 313, pp. 504-507.
- [2] U.R. Acharya, S. L. Oh, Y. Hagiwara, J.H. Tan and H. Adeli, "Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals." *Comput. Biol. Med.*, Sep. 2017, vol. 100, pp. 270–278.
- [3] S. Tripathi, S. Acharya, R.D. Sharma, S. Mittal and S. Bhattacharya, "Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset," in *Proc. IAAI*, 2017, pp. 4746–4752.
- [4] O. Tsinalis, P.M. Matthews and Y. Guo, "Automatic sleep stage scoring using time-frequency analysis and stacked sparse autoencoders," *Ann. Biomed. Eng.*, 2016, vol. 44, no. 5, pp. 1587–1597.
- [5] Y.R. Tabar and U. Halici, "A novel deep learning approach for classification of EEG motor imagery signals," *Journal of Neural Engineering*, 2017, vol. 14, no. 1, 016003.
- [6] G. Kotsovolis and G. Komninos, "Awareness during anesthesia: how sure can we be that the patient is sleeping indeed?" *Hippokratia*, 2009, vol. 13, pp. 83–9.
- [7] A. Petsiti, V. Tassoudis, G. Vretzakit, D. Zacharoulis, K. Tepetes, G. Ganeli, et al. "Depth of anesthesia as a risk factor for perioperative morbidity," *Anesthesiol Res Pract*, 2015, 2015:829151.
- [8] D. Schwender, A. Kaiser, S. Klasing, K. Peter and E. Poppel, "Midlatency auditory evoked potentials and explicit and implicit memory in patients undergoing cardiac surgery," *Anesthesiology*. 1994, vol. 80, no. 3, pp. 493–501.
- [9] H. Viertiö-Oja, V. Maja, M. Särkelä, P. Talja, N. Tenkanen, H. Tolvanen-Laakso, M. Paloheimo, A. Vakkuri, A. Yli-Hankala, and P. Meriläinen, "Description of the Entropy™ algorithm as applied in theDatex-Ohmeda S/5™ entropy module *Acta Anaesthesiol.*" *Scand*, 2004, 48, pp.154–61.
- [10] J.C. Sigl, and N.G. Chamoun, "An introduction to bispectral analysis for the electroencephalogram," *Journal of Clinical Monitoring*, vol. 10, no. 6, 1994, pp.392–404.
- [11] B. Boashash, M. Mesbah and P. Golditz, "Time-Frequency Detection of EEG Abnormalities.," Amsterdam, The Netherlands: Elsevier, 2003, pp. 663–669, ch15.
- [12] S.A. Taywade, and R.D. Raut, "A review: EEG signal analysis with different methodologies," in *Proceedings of the National Conference on Innovative Paradigms in Engineering and Technology (NCIPET '12)*, 2014, pp. 29–31.
- [13] K. Kuizenga, JM. Wierda, and CJ. Kalkman, "Biphasic EEG changes in relation to loss of consciousness during induction with thiopental, propofol, etomidate, midazolam or sevoflurane," *Br J Anaesth*, 2001, vol. 86, pp. 354–360.
- [14] Yu-Po Huang, Jerry Chen, Shou-Zen Fan, Maysam F. Abbod, Jiann-Shing Shieh and Yu-Chen Kung, "Applying CNN Model for time-frequency analysis of EEG to assess the depth of anesthesia based on BIS value", 24th International Symposium on Artificial Life and Robotics (AROB 24th 2019), Jan. 23~25, B-Con Plaza, Beppu, Oita, Japan, 2019, pp. 39-42.