EEG-Based Neonatal Sleep-Wake Classification Using Multilayer Perceptron Neural Network

SAADULLAH FAROOQ ABBASI[1], **JAWAD AHMAD**[2], **(Senior Member, IEEE),**
AHSEN TAHIR[3,4]**, MUHAMMAD AWAIS**[1]**, (Member, IEEE), CHEN CHEN**[1]**,**
MUHAMMAD IRFAN[1]**, HAFIZA AYESHA SIDDIQA**[1]**, ABU BAKAR WAQAS**[1]**,**
XI LONG[5]**, (Member, IEEE), BIN YIN**[6]**, SAEED AKBARZADEH**[1]**, (Member, IEEE),**
CHUNMEI LU[7]**, LAISHUAN WANG**[7]**, AND WEI CHEN**[1,8]**, (Senior Member, IEEE)**

[1]Center for Intelligent Medical Electronics, Department of Electronic Engineering, School of Information Science and Technology, Fudan University, Shanghai 200433, China
[2]School of Computing, Edinburgh Napier University, Edinburgh EH10 5DT, U.K.
[3]School of Computing, Engineering, and Built Environment, Glasgow Caledonian University, Glasgow G4 0BA, U.K.
[4]Department of Electrical Engineering, University of Engineering and Technology at Lahore, Lahore 54890, Pakistan
[5]Department of Electrical Engineering, Eindhoven University of Technology, 5612 AZ Eindhoven, The Netherlands
[6]Connected Care and Personal Health Department, Philips Research, Shanghai 200433, China
[7]Department of Neonatology, Children's Hospital of Fudan University, Shanghai 200032, China
[8]Human Phenome Institute, Fudan University, Shanghai 200433, China

Corresponding author: Wei Chen (w_chen@fudan.edu.cn)

ABSTRACT Objective: Classification of sleep-wake states using multichannel electroencephalography (EEG) data that reliably work for neonates. Methods: A deep multilayer perceptron (MLP) neural network is developed to classify sleep-wake states using multichannel bipolar EEG signals, which takes an input vector of size 108 containing the joint features of 9 channels. The network avoids any post-processing step in order to work as a full-fledged real-time application. For training and testing the model, EEG recordings of 3525 30-second segments from 19 neonates (postmenstrual age of 37 ± 05 weeks) are used. Results: For sleep-wake classification, mean Cohen's kappa between the network estimate and the ground truth annotation by human experts is 0.62. The maximum mean accuracy can reach up to 83% which, to date, is the highest accuracy for sleep-wake classification.

INDEX TERMS Neonatal sleep staging, electroencephalogram, classification, multilayer perceptron, neural network.

I. INTRODUCTION

Sleep is an important human function, which is identified by the sequence of brain alterations. For neonates, they spend most of their time resting in a sleep state. Sleep ontogenesis is an active process for brain maturation and the central nervous system. Clinically, sleep-wake cycling (SWC) is the main hallmark of brain development in neonates [1], [2]. In particular, in a neonatal intensive care unit (NICU), neonatal sleep should be protected and promoted.

Polysomnography (PSG) is considered as the gold standard to monitor sleep and diagnose sleep disorders [3]. In the past decade, many studies have demonstrated the feasibility of automated sleep staging algorithms with PSG signals, among

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

which EEG is considered as the most reliable signal for both adults [4]–[6] and infants [7]–[9].

Hans Berger recorded the first EEG of humans in 1924 [10]. Brain's electrical activity takes place via electrical impulses and can be measured from the scalp of the patient. Electrodes are placed as per the international 10-20 system for electrode placement [11]. Neurologists have established clear EEG patterns in SWC from 30 weeks postmenstrual age [12]. In 1937, Loomis *et al.* proposed the first application of EEG based study of human sleep patterns [16]. After the novel research of Loomis, multiple algorithms have been proposed for adult sleep staging using machine [17-20] and deep learning. Deep learning algorithms for sleep staging include convolutional neural network (CNN) [21], recurrent neural network (RNN) [22], the combination of CNN or RNN [23-24]and Long Short-Term Memory (LSTM) [25,26].

TABLE 1. Bio-insights neonatal sleep and wake stage.

Sleep Stage	EEG pattern	AASM Manual Definition	Description
Sleep (Quiet Sleep)	Trace Alternant (TA)	≥ 3 runs of high voltage (50–150 μV) bursts of 1–3 Hz delta activity alternating with periods of lower amplitude (25–50 μV) 4–7 Hz theta activity	Inter-burst intervals (IBIs) appear. Amplitude of IBIs increases with increasing age.
Sleep (Quiet Sleep)	high voltage slow (HVS)	Continuous high voltage (100–150 μV) 1–3 Hz delta activity, which often has an occipital or central predominance.	Development of autonomous nervous system (ANS) during quiet sleep (QS).
Wake/ rarely Active Sleep	Mixed (M)	Both HVS and LV polyrhythmic components intermingled with little periodicity.	Continuous low amplitude (< 50 μV) 5–7 Hz theta intermixed with (> 100 μV) 2–4 Hz slow waves. M typically seen in sleep after period of Wake. Endogenous neuronal activation occurs during active sleep (AS).
Wake/ Active Sleep	low voltage irregular (LVI)	Continuous low voltage mixed-frequency activity with delta and predominantly theta activity.	Continuous irregular (25–50 μV) 4–7 Hz activity intermixed with 1–3 Hz delta of similarly low amplitude Development of central nervous system (CNS) during active sleep (AS).

Contrasting EEG patterns have been observed for neonates and adults. The neonatal EEG pattern exhibits a smaller magnitude, as compared to the grown-ups. Multiple maturation changes occur within the first three years [27]. For this purpose, multiple automatic neonatal sleep staging algorithms based on EEG have been proposed. To the best of our knowledge, most existing algorithms described in previous studies lack in characterizing 'wake' as a distinct state. Other algorithms classified sleep stages based on different characteristics of EEG signals i.e. low voltage irregular (LVI), Active Sleep II (AS II), high voltage slow (HVS) and Trace Alternant (TA)/Trace Discontinue (TD) [9], [13]–[15]. The brain maturation process initiates during AS and wake, whereas within the existing set of frameworks, amalgamation of 'AS' and 'wake' into LVI state. The bio-insights of sleep and awake stage is illustrated in Table 1.

In this paper, we present a sleep-wake classification algorithm based on multilayer perceptron (MLP) neural network. It attempts to resolve the intermixing of different sleep stages by classifying wake and sleep as un-identical states. Our work is mainly divided into two parts: Feature extraction and Classification. Twelve features were extracted from multichannel EEG and subsequently the MLP was applied for training and testing the neural network. The remaining part of the proposed paper is arranged in the following manner: Section II introduces related work. Section III presents the materials and methods. Sleep-wake classification results using the proposed method are reported and discussed in Section IV and V, respectively. Finally, Section VI concludes the study.

II. RELATED WORK
For EEG-based neonatal sleep stage classification, different features have been proposed to measure, for example, EEG

(dis)continuity [8], frequency content [28], proportional duration of bursts [29], and the frequency content of bursts [30]. Some maturational abnormalities in the neonatal brain are only apparent in QS, reflecting more alterations in brain function [32]–[36]. Turnbull *et al.* [37] detected a specific discontinuous EEG pattern, known as TA as mentioned before. While proving reliable for TA detection, this was shown to be not sufficient to infer QS over a wide age range, as QS also contains HVS signals.

DeWel *et al.* [38] proposed a supervised algorithm for sleep state classification across a wide range of neonates (27-42 postmenstrual age), using an LS-SVM classifier and multi-scale entropy features. Firstly, the proposed algorithm estimates the postmenstrual age (PMA) using four complexity features. Then, a sleep state classifier was developed using these features to identify quiet sleep from neonatal EEG data. In 2017, Dereymaeker *et al.* [7] proposed CLuster-based Adaptive Sleep Staging (CLASS) to automatically detect quiet sleep (QS) [7]. They highlighted the benefit of QS detection in brain maturation. Another algorithm based on SVM with radial basis function was proposed by Koolen *et al.* [8], which can detect QS with an accuracy of 85%.

To classify all stages of neonate sleep, Pillay *et al.* [9] proposed an algorithm based on a generative modeling approach. Hidden Markov models (HMMs) and Gaussian mixture models (GMMs) were trained using EEG features extracted from 16 EEG recordings. For four stages (LVI, TA/TD, AS II, HVS), HMMs were showed to be superior with a Cohen's kappa of 0.62. Recently, a CNN-based algorithm outclasses the state-of-the-art sleep stage classification algorithms with a kappa ranged from 0.66 to 0.70 [39]. However, none of these algorithms considered wake as a separate state in neonates.
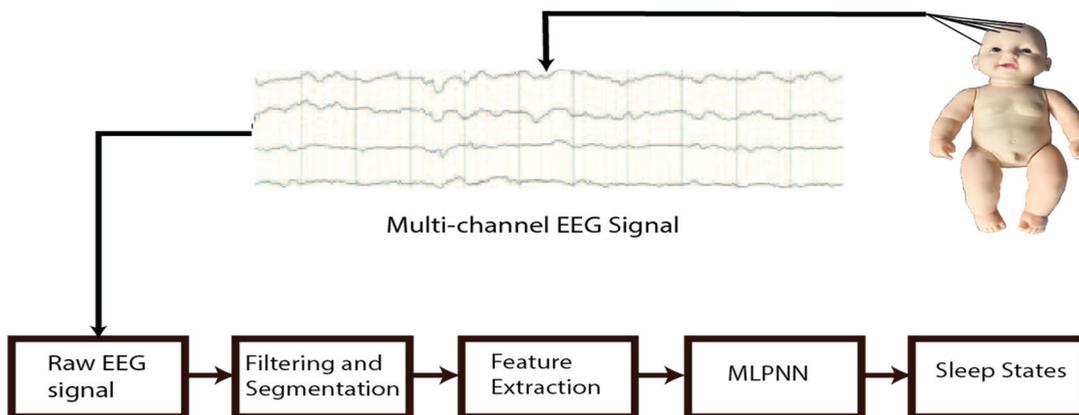
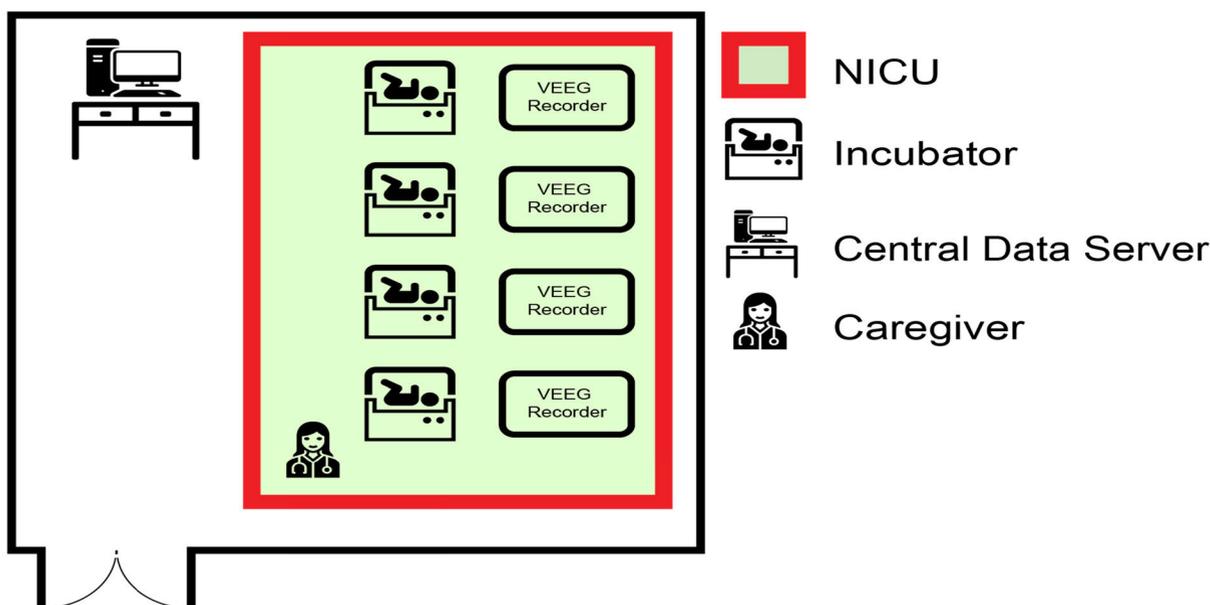**FIGURE 1.** Block diagram of the proposed scheme.

FIGURE 2. Experimental setup.

As stated before, the current method for sleep-wake classification in neonates is limited. Fraiwan and Lwseey [31] proposed an algorithm based on auto-encoders to classify neonatal sleep stages. It explained two main steps: feature extraction and classification. Twelve EEG features were extracted from neonatal EEG for training and testing, yet its accuracy to detect wake state was restricted to merely 17%.

III. MATERIALS AND METHODS

In this section, we present the complete description of the proposed MLP neural network. Figure 1. shows the block diagram of the proposed algorithm.

A. DATASET

Video EEG (VEEG) recordings from 19 neonates were used for testing and training the network [40]. EEG recordings

were obtained from a NICU at Children's hospital affiliated to Fudan University, Shanghai, China. Approval of the study was issued by the Research Ethics Committee of the Children's Hospital affiliated to Fudan University (Approval No. (2017) 89). Figure 2. shows the experimental setup for neonatal sleep monitoring. On average, 120 minutes of data was obtained from each neonate, in which at least one sleep cycle was observed. The full 10-20 system for electrode placement includes "Fp1 - 2," "F3 - 4," "F7 - 8," "C3 - 4," "P3 - 4," "T3 - 4," "T5 - 6," "O1 - 2" and "Cz" (17 electrodes). Among the 19 EEG recordings we recorded, 15 include all the given electrodes except "T5 - 6," "F7 - 8" and "O1 - 2" (11 electrodes). In the remaining 4 recordings "T5 - 6," "F7 - 8," "Cz" and "O1 - 2" were not recorded, leading to 10 electrodes included. NicoletOne EEG system was used to extract multichannel EEG for this study.

B. VISUAL SLEEP SCORING

EEG segments were visually annotated by two trained doctors. The primary rater (CL) labelled the start and end time of each stage i.e. wake, sleep and the artifactual region where clear sleep stage cannot be classified. Whereas the secondary rater (LW) verified the regions annotated by the primary rater. Also, LW annotated the regions where CL were not agreed on. Sleep and wake stage were identified using both EEG and non-cerebral characteristics. Videos of the NICU were also considered during annotation. The annoted regions were divided into three main categories: wake, sleep and artifacts.

C. PRE-PROCESSING

During recording and processing, EEG recordings got contaminated with noise and artifacts. EEG recordings were processed on their original frequency of 500 Hz. Before further processing, these artifacts and noises should be removed from the EEG recordings. Our pre-processing is divided into three parts:

1) Filtering the EEG signal using a FIR filter [41].
2) Segmenting the multichannel EEG into 30 sec epochs and assigning a label to each epoch.
3) Post segmentation, artifacts were removed from the EEG recordings, in case, if it constitutes 20% of the artifact. Artifacts were removed based on the annotation provided by the professional doctors. After eliminating artifacts, we were left with 3525 segments for training and testing.

D. FEATURE EXTRACTION

After the pre-processing, we extracted 12 potential features from each EEG channel and combined them to form an input vector of size 108. The feature set includes 8 time and 4 frequency domain features. Frequency domain features were extracted by taking the Fast Fourier Transform (FFT) of the EEG segments. After FFT, mean frequencies were extracted from each band (alpha, delta, theta, and beta). Table 2. shows the list of the extracted features.

E. MULTILAYER PERCEPTION NEURAL NETWORK

MLP neural network has advantageous properties such as, smaller training set requirements [42], [43], easy implementation and fast operation. It consists of three primary layers: an input layer, hidden layer(s), and an output layer. The hidden layers are responsible for processing and transmitting the input information to the output layer. An MLP is a mapping between two Euclidean spaces ($RR^{n_1}n1, R^{n_2}$).

The mapping is defined as a sequence of Euclidean spaces $R^{n_1}, R^{n_2} \ldots R^{n_L}$ and the mapping (F) connecting them:

$$F_1 : R^{n_1} \rightarrow R^{n_1}, F_2 : R^{n_1} \rightarrow R^{n_1} \ldots F_{L-1} : R^{n_L-1} \rightarrow R^{n_L},$$

where L are the total number of MLP layers. In the MLP neural network, each neuron j in the hidden layer sums its input signals x_i after multiplying them by the strengths of the respective connection weights w_{ij} and computes its output y_j

TABLE 2. Features (per 30-s epoch) extracted from each EEG channel.

Time Domain	Mean
	Median
	Standard Deviation
	Variance
	Kurtosis
	Skewness
	Minimum
	Maximum
Frequency Domain (Mean amplitudes)	Delta Band (0.5-3 Hz)
	Theta Band (3-8 Hz)
	Alpha Band (8-12 Hz)
	Beta Band (12-30 Hz)

as a function of the sum. Mathematically,

$$y_j = f(\sum w_{ij} * O_i) \tag{1}$$

Here w are the weights and are updated according to the gradient descent algorithm whereas O symbolizes the hidden layers.

F. GRADIENT DESCENT ALGORITHM

For a neural network, efficiency of learning is important. One of the efficient algorithms to train the neural network is the standard gradient descent algorithm (GDA) [44]. GDA works by taking the derivative of the error function with respect to the weights at a specific position on the loss function and then updating the weights towards a negative gradient.

Initially, the weights for the GDA are selected randomly. Once the maximum number of iterations are reached, the training algorithm stops immediately. The main objective of the gradient descent algorithm is to minimize the loss function. The loss function is given as:

$$E = \frac{1}{N} \sum_{m=1}^{N} ||q_m - y_m||^2 \tag{2}$$

N is the total number of iterations, whereas q_m and y_m are the expected and desired value, respectively.

G. PROPOSED MLP NEURAL NETWORK ARCHITECTURE

In this study, an efficient yet simple deep MLP neural network is proposed. The deep MLP has a densely connected architecture with an input layer of size 256, 3 hidden layers of size 256-128-64 and an output layer with sigmoid activation function, resulting in a total of 5-layer MLP neural network. RMSProp algorithm [45] was used for training the neural network with a learning rate of 0.001. Two other deep neural

TABLE 3. Deep neural network structures used in this study for performance comparison in neonatal sleep-wake classification.

Network	Input layer	Hidden layer	Output layer	Epochs
MLP	256	256-128-64	1 (Sigmoid)	5
CNN (Convolutional Neural Network)	1*9*12	Conv (50)-maxpool-conv (30)-maxpool-conv (30)-maxpool-MLP	1 (Sigmoid)	30
RNN (Random Neural Network)	108	13	1 (Sigmoid)	20

TABLE 4. Learning algorithm parameters for MLP neural network.

Parameters	
Learning algorithm	RMSProp
Epochs	5
Moving average parameters	0.9
Learning rate	0.001

network algorithms including CNN and RNN were implemented for performance comparison. The network parameters for all the other networks were well tuned and the one giving best accuracy are reported for comparison. MLPs are universal function approximators [46] therefore they can be used for creating mathematical models by regression analysis. This is the main reason behind the success of MLP neural network in this particular case. The network parameters of different neural networks implemented are shown in Table 3.
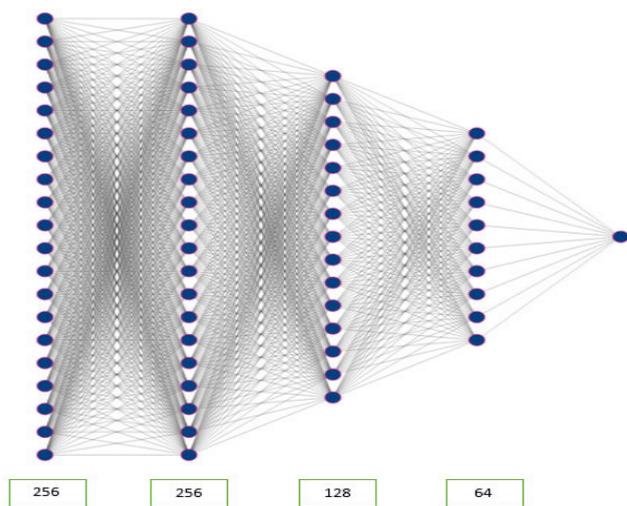

FIGURE 3. MLP neural network architecture with input size 108. The MLP network architecture is 256∗256∗128∗64∗1.

Figure 3 shows the proposed MLP neural network architecture. The input is a combination of features extracted from the 9 bipolar EEG channels. A total of 12 features are extracted from each bipolar EEG recording, consisting of 8-time domain features and 4-spatial domain features as described in Table 2 of the feature extraction subsection. The learning parameters of the proposed MLP neural network are shown in Table 4.

H. EVALUATION

To access the performance of the proposed scheme, multiple performance matrices have been used. These matrices are mean accuracy, mean Cohen's kappa, sensitivity and specificity. For comparison between different NNs, we reported mean kappa and accuracy. Mathematically, these performance parameters are given as:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
$$Sensitivity = \frac{TP}{TP + FN}$$
$$Specificity = \frac{TN}{TN + FP}$$

where TP are true positives, TN are true negatives, FP are false positives and FN are false negatives. In addition to these, the confusion matrix is also reported.

In order to validate the proposed algorithm during the design process, the neonatal sleep data was split into 4 folds. The dataset is divided into 4 subsets and holdout method is repeated 4 times. Each time, one of the 4 subsets is used as the test set and the other 3 subsets are put together to form a training set. Then the average error across all 4 trials is computed. The advantage of this method is that it matters less how the data gets divided. 3525 segments were used from 19 subjects for training and testing the neural network. The final results were obtained by taking the mean of the stratified 4-fold permutation $\pm$ standard deviation.

Learning rate, also known as step size, is one of the most important tuning parameters in an optimization algorithm. Usually, the learning rate has a small value within the range 0-1. It determines the step size of every iteration while moving towards the minimum of the loss function. For this purpose, different learning rates were used to evaluate the performance of the proposed algorithm. Processing time is also an important parameter to access the quality of the proposed algorithm. Computational cost is directly proportional to the processing time. We also calculated and analyzed the time used for training and testing the neural network.

As per our assiduous research, there is only one proposed algorithm which classifies awake as a separate state [31]. For this reason, we applied different machine learning and deep neural network algorithms on our dataset and compared the results accordingly. It is pertinent to mention that the same training and testing datasets were used for each reported algorithm. Also, same features were used as an input for each

reported network. This helps us to have a fair comparison between different algorithms.

IV. RESULTS

All the networks are trained and tested on Intel Core i5-8400, RAM 16GB with GTX 1050ti. The proposed neural network was implemented using Keras and TensorFlow. The features were obtained using MATLAB 2019b. The testing results give the highest accuracy of 82.53 ± 1.63% (standard error = 0.82) for sleep-wake classification using 5-layer MLP neural network. Table 5. Shows the confusion matrix for sleep-wake classification using MLP neural network.

TABLE 5. Confusion matrix for Sleep-Wake classification.

	Positive	**Negative**
Wake (True)	379	78
Sleep (False)	76	349

Table 6. shows the overall test performance alluded by the confusion matrix. Accuracy, kappa, sensitivity and specificity are calculated using the confusion matrix shown in Table 5.

TABLE 6. Results for Sleep-Wake classification.

	Kappa	Accuracy	Sensitivity	Specificity
MLP	0.65	82.53 %	83.29 %	81.73 %

The results of MLP neural network using different number of hidden layers are illustrated in Table 7. The proposed network is optimal with 5 layers having the highest accuracy of 82.53%, which reduces with increment or decrement in the number of layers. If we increase the number of layers the network starts overfitting. Whereas by reducing the number of hidden layers the networks give underfitted results (Table 7).

TABLE 7. Comparison of MLP neural network with different number of neurons and hidden layers.

Network	I	II	III	IV	V
Total layers	3	4	5	6	7
Hidden layers	1	2	3	4	5
Total neurons	513	641	705	737	752
Accuracy (%)	75.6	77	**82.53**	78	65
Kappa	0.515	0.52	**0.65**	0.55	0.317

The accuracy of the neural network changes by changing the learning rate. Table 8. shows the results of the proposed MLP neural network using different learning rates. With the learning rate of 0.001, MLP neural network gives highest reported accuracy. It is important to note that for learning rates 0.1 and 0.01, the kappa is zero i.e. the chances are

TABLE 8. Results of MLP neural network with different learning rates.

Learning rate	Accuracy	Kappa
0.1	49.09	0.0
0.01	51.02	0.0
0.001	**82.7**	**0.65**
0.0001	73.01	0.48
0.00001	66.21	0.33

TABLE 9. Comparison of MLP neural network with other neural networks for Sleep-wake classification.

Results for sleep wake classification	Accuracy (%) Sleep-wake Classification	Kappa	Time (sec)
MLP	**82.53**	**0.65**	**10.9139**
CNN (Convolutional Neural Network)	75.67	0.58	107
RNN (Random Neural Network)	71	0.55	521.489

TABLE 10. MLP neural network with different number of channels.

Number of channels	Channel name	Accuracy	Kappa
1	F4 - C4	71%	0.52
4	F4-C4, C4-T4, F3-C3, C3-T3	74.7%	0.55
9	F4-C4, C4-T4, F3-C3, C3-T3, T3-C3, T4-P4, P3-F3, P4-F4, T3-P3	**82.7%**	**0.65**

TABLE 11. Comparison of MLP with other machine learning algorithms in neonatal sleep-wake classification.

Algorithm	Accuracy
MLP	**82.7%**
SVM (Quadratic kernel)	72.2%
Logistic Regression	65.1%
KNN (K=10)	63.1%
Decision Tree (100 splits)	60.85%

random for sleep-wake classification. Selecting high learning rate caused undesirable divergent behavior in our loss function.

The test results of the proposed network architecture are compared with other neural network architectures. The results are illustrated in table 9. It is very important to note that wake and sleep both contain LVI signals, so it is challenging to distinguish the two stages. Being cognizant of this fact, the proposed MLP neural network achieved very promising results to classify sleep-wake cycling with EEG recordings.

TABLE 12. Overall performance of MLP vs existing algorithms for sleep-wake classification.

Study	Year	Database	Method	Overall Performance (wake classification)	Additional information
Fraiwan et al. [31]	2017	University of Pittsburg recordings	Deep learning autoencoders	17%	Epoch length=60 sec Sensitivity=46.6%, Specificity=96%, Kappa= Random.
This study	2020	Children hospital Fudan University recordings	Multilayer perceptron	82.53%	Epoch length=30 sec Sensitivity=83.29, Specificity=81.73, Kappa= 0.65.

Furthermore, deep neural networks are usually computationally intensive having high computational cost. In this regard, the proposed network architecture is simple, efficient and have low computational time (Table 9). Therefore, MLP neural network has potential to be used for real-time sleep-wake states classification as the network didn't use any post-processing process.

Not every NICU is equipped with large amount of EEG electrodes therefore we also investigated the results of MLP neural network over a smaller number of EEG electrodes. For all the cases, we used bipolar EEG recordings. Table 10. shows the results of MLP neural network using different number of channels. In the proposed network architecture, we used 9 bipolar EEG channels. By reducing the number of channels, the network accuracy decreases. In case of using 4 channels, the accuracy is 74.7% and if using only one channel, the accuracy drops to 71%.

Mostly, Machine learning algorithms like, SVM, K Nearest Neighbors (KNN), and Decision Tree give better results while using handcrafted features. For this purpose, we compared the results of MLP neural network with machine learning algorithms. Table 11 gives the comparison table of MLP neural network vs machine learning algorithms. It has been noticed that MLP neural network performs better as compared to machine learning algorithms. For each reported algorithm, results are shown with the parameters giving the best possible results. Also, it is very important to note that the same training and testing data were used for every algorithm.

V. DISCUSSION
MLP is proven to be a very successful classifier for different applications, but to the best of our knowledge this is the first time MLP has been used for sleep staging. There are number of algorithms which classified neonatal sleep stages but none of those algorithms classify wake as a separate sleep stage. Mostly, wake is combined with AS I to form an LVI stage. This results in the intermixing of two sleep stages. In this paper, we classified sleep and wake as distinct stages with an accuracy of 82.53.

We extracted 12 EEG features from each 30 sec EEG segment. They are divided into two categories: time domain and frequency domain. These features were selected because we noticed highest accuracy with these features. If we change the number of features, the accuracy decreases gradually. The most prominent features are the frequency domain features. These features are calculated by taking the Fourier transform of the EEG segment. After applying Fourier transform, we calculated the mean amplitude of the given bands i.e. alpha, beta, theta and delta. By adding these features, the accuracy increases by 15%.

Mostly, CNN method outperforms other neural networks. To the best of our knowledge, CNN works well with raw data which has a spatial relationship to extract its own features. While, in our case, we engineered our own features which were best suited for sleep-wake classification. The engineered features do not possess any spatial relationship, which a convolutional kernel in a CNN assumes. The use of different weights for each feature compared to same weights of a convolutional kernel assumes no spatial relation, which is truly the case for our features and hence provides better classification performance.

To report the overall performance of the proposed study along with the existing algorithm for sleep-wake classification. Table 12. Provides the complete evaluation for the metrices discussed in section III-H. In addition, Table 12. provides the additional information related to the epoch length selection and other evaluation metrices. From this, we can conclude that MLP performs better as compared to the existing algorithm i.e. deep learning autoencoders.

There are two main limitations of this study that should be taken into consideration: first, the data used in this study consists of 19 subjects which is very small. A larger dataset is likely to increase the performance and concreteness of the proposed algorithm. Second, there is no algorithm with which we can compare our proposed algorithm. For this purpose, we applied different machine and deep learning algorithms on our dataset and compared the results with MLP neural network. MLP neural network outclasses all algorithms.

In the proposed study, artifacts were not considered and were removed manually during preprocessing. These artifacts, in the NICU, can contaminate the EEG recordings and decrease the network performance. It would be advantageous if we have an automatic method for artifact removal so that the proposed study can be used directly in an NICU.

As a future work, we aim to classify further sleep stages: AS, QS, IS and wake. In addition, more data will be used in

the future study to increase the performance. More considerations will be taken during training and testing to improve the performance with a smaller number of EEG channels, as this may help in practical usage.

VI. CONCLUSION

In this study, we proposed a low cost, efficient, and simple deep MLP neural network for sleep wake classification using multichannel EEG signals. 8-time domain and 4-spatial domain features were extracted from neonatal EEG recordings and combined to form an input to the neural network of size 108. This is the first reported automatic sleep wake classification algorithm. The proposed neural network didn't use any post-processing technique which strengthen its candidature to be used for real-time sleep-wake classification. For comparison, different neural networks were applied on the same dataset. It is evident from the results that the proposed MLP neural network shows better results for sleep wake classification. To conclude, we can say that properties like real-time processing, low computational cost and easy implementation makes MLP neural network feasible for neonatal sleep staging. More importantly, this network can help to quantify the abnormalities in neonatal brain development. The future work aims to develop a neural network algorithm for the classification of more sleep stages.

REFERENCES

[1] S. Graven and J. Browne, "Sleep and brain development: The critical role of sleep in fetal and early neonatal brain development," *Newborn Infant Nursing Rev.*, vol. 8, no. 4, pp. 173–179, 2008.

[2] S. M. Ludington-Hoe, "Neurophysiologic assessment of neonatal sleep organization: Preliminary results of a randomized, controlled trial of skin contact with preterm infants," *Pediatrics*, vol. 117, no. 5, pp. e909–e923, May 2006.

[3] W. H. Spriggs, *Essentials of polysomnography*. Burlington, MA, USA: Jones & Bartlett, 2014.

[4] A. R. Hassan and M. I. H. Bhuiyan, "An automated method for sleep staging from EEG signals using normal inverse Gaussian parameters and adaptive boosting," *Neurocomputing*, vol. 219, pp. 76–87, Jan. 2017.

[5] A. R. Hassan and M. I. H. Bhuiyan, "A decision support system for automatic sleep staging from EEG signals using tunable Q-factor wavelet transform and spectral features," *J. Neurosci. Methods*, vol. 271, pp. 107–118, Sep. 2016.

[6] S. Biswal, J. Kulas, H. Sun, B. Goparaju, M. B. Westover, M. T. Bianchi, and J. Sun, "SLEEPNET: Automated sleep staging system via deep learning," 2017, *arXiv:1707.08262*. [Online]. Available: http://arxiv.org/abs/1707.08262

[7] A. Dereymaeker, K. Pillay, J. Vervisch, S. Van Huffel, G. Naulaers, K. Jansen, and M. De Vos, "An automated quiet sleep detection approach in preterm infants as a gateway to assess brain maturation," *Int. J. Neural Syst.*, vol. 27, no. 06, Sep. 2017, Art. no. 1750023.

[8] N. Koolen, L. Oberdorfer, Z. Rona, V. Giordano, T. Werther, K. Klebermass-Schrehof, N. Stevenson, and S. Vanhatalo, "Automated classification of neonatal sleep states using EEG," *Clin. Neurophysiol.*, vol. 128, no. 6, pp. 1100–1108, Jun. 2017.

[9] K. Pillay, A. Dereymaeker, K. Jansen, G. Naulaers, S. Van Huffel, and M. De Vos, "Automated EEG sleep staging in the term-age baby using a generative modelling approach," *J. Neural Eng.*, vol. 15, no. 3, Jun. 2018, Art. no. 036004.

[10] J. Bronzino, "Principles of electroencephalography," Tech. Rep., 1995.

[11] R. W. Homan, J. Herman, and P. Purdy, "Cerebral location of international 10–20 system electrode placement," *Electroencephalogr. Clin. Neurophysiol.*, vol. 66, no. 4, pp. 376–382, Apr. 1987.

[12] M. Andre, "Pesquisas sobre formaço de professores: Contribuiçes delimitaço do campo. Convergncias e tenses no campo da formao e do trabalho docente: Didática, formaço de professores, trabalho docente," Tech. Rep., 2010.

[13] A. H. Parmelee, W. H. Wenner, Y. Akiyama, M. Schultz, and E. Stern, "Sleep states in premature infants," *Develop. Med. Child Neurol.*, vol. 9, no. 1, pp. 70–77, Nov. 2008.

[14] J. A. Hobson, "A manual of standardized terminology, techniques and criteria for scoring of states of sleep and wakefulness in newborn infants," *Electroencephalogr. Clin. Neurophysiol.*, vol. 33, no. 6, pp. 614–615, Dec. 1972.

[15] D. Y. Barbeau and M. D. Weiss, "Sleep disturbances in newborns," *Children*, vol. 4, no. 10, p. 90, Oct. 2017.

[16] A. Loomis, E. Harvey, and G. Hobart, "Cerebral states during sleep, as studied by human brain potentials," *J. Experim. Psychol.*, vol. 21, no. 2, p. 127, 1937.

[17] T. Lajnef, S. Chaibi, P. Ruby, P.-E. Aguera, J.-B. Eichenlaub, M. Samet, A. Kachouri, and K. Jerbi, "Learning machines and sleeping brains: Automatic sleep stage classification using decision-tree multi-class support vector machines," *J. Neurosci. Methods*, vol. 250, pp. 94–105, Jul. 2015.

[18] M. Xiao, H. Yan, J. Song, Y. Yang, and X. Yang, "Sleep stages classification based on heart rate variability and random forest," *Biomed. Signal Process. Control*, vol. 8, no. 6, pp. 624–633, Nov. 2013.

[19] P. Fonseca, N. den Teuling, X. Long, and R. M. Aarts, "Cardiorespiratory sleep stage detection using conditional random fields," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 4, pp. 956–966, Jul. 2017.

[20] S. Gudmundsson, T. P. Runarsson, and S. Sigurdsson, "Automatic sleep staging using support vector machines with posterior probability estimates," in *Proc. Int. Conf. Comput. Intell. Modeling*, 1998, pp. 366–372.

[21] A. Sors, S. Bonnet, S. Mirek, L. Vercueil, and J.-F. Payen, "A convolutional neural network for sleep stage scoring from raw single-channel EEG," *Biomed. Signal Process. Control*, vol. 42, pp. 107–114, Apr. 2018.

[22] Y.-L. Hsu, Y.-T. Yang, J.-S. Wang, and C.-Y. Hsu, "Automatic sleep stage recurrent neural classifier using energy features of EEG signals," *Neurocomputing*, vol. 104, pp. 105–114, Mar. 2013.

[23] H. Phan, F. Andreotti, N. Cooray, O. Y. Chen, and M. De Vos, "Joint classification and prediction CNN framework for automatic sleep stage classification," *IEEE Trans. Biomed. Eng.*, vol. 66, no. 5, pp. 1285–1296, May 2019.

[24] C. Sun, J. Fan, C. Chen, W. Li, and W. Chen, "A two-stage neural network for sleep stage classification based on feature learning, sequence learning, and data augmentation," *IEEE Access*, vol. 7, pp. 109386–109397, 2019.

[25] N. Michielli, U. R. Acharya, and F. Molinari, "Cascaded LSTM recurrent neural network for automated sleep stage classification using single-channel EEG signals," *Comput. Biol. Med.*, vol. 106, pp. 71–81, Mar. 2019.

[26] M. Radha, P. Fonseca, A. Moreau, M. Ross, A. Cerny, P. Anderer, X. Long, and R. M. Aarts, "Sleep stage classification from heart-rate variability using long short-term memory neural networks," *Sci. Rep.*, vol. 9, no. 1, Dec. 2019, Art. no. 14149.

[27] L. V. Marcuse, M. Schneider, K. A. Mortati, K. M. Donnelly, V. Arnedo, and A. C. Grant, "Quantitative analysis of the EEG posterior-dominant rhythm in healthy adolescents," *Clin. Neurophysiol.*, vol. 119, no. 8, pp. 1778–1781, Aug. 2008.

[28] A. Piryatinska, G. Terdik, W. A. Woyczynski, K. A. Loparo, M. S. Scher, and A. Zlotnik, "Automated detection of neonate EEG sleep stages," *Comput. Methods Programs Biomed.*, vol. 95, no. 1, pp. 31–46, Jul. 2009.

[29] K. Palmu, T. Kirjavainen, S. Stjerna, T. Salokivi, and S. Vanhatalo, "Sleep wake cycling in early preterm infants: Comparison of polysomnographic recordings with a novel EEG-based index," *Clin. Neurophysiol.*, vol. 124, no. 9, pp. 1807–1814, Sep. 2013.

[30] N. J. Stevenson, K. Palmu, S. Wikström, L. Hellström-Westas, and S. Vanhatalo, "Measuring brain activity cycling (BAC) in long term EEG monitoring of preterm babies," *Physiological Meas.*, vol. 35, no. 7, pp. 1493–1508, Jun. 2014.

[31] L. Fraiwan and K. Lweesy, "Neonatal sleep state identification using deep learning autoencoders," in *Proc. IEEE 13th Int. Colloq. Signal Process. its Appl. (CSPA)*, Mar. 2017, pp. 228–231.

[32] M. Solc, R. Ronchi, J. Bello-Ruiz, T. Schmidlin, B. Herbelin, F. Luthi, M. Konzelmann, J.-Y. Beaulieu, F. Delaquaize, A. Schnider, A. G. Guggisberg, A. Serino, and O. Blanke, "Heartbeat-enhanced immersive virtual reality to treat complex regional pain syndrome," *Neurology*, vol. 91, no. 5, pp. e479–e489, Jul. 2018.

[33] N. Koolen, A. Dereymaeker, O. RÃsÃnen, K. Jansen, J. Vervisch, V. Matic, M. De Vos, S. Van Huffel, G. Naulaers, and S. Vanhatalo, "Interhemispheric synchrony in the neonatal EEG revisited: Activation synchrony index as a promising classifier," *Frontiers Human Neurosci.*, vol. 8, p. 1030, Dec. 2014.

[34] G. Holmes and C. Lombroso, "Prognostic value of background patterns in the neonatal EEG," *J. Clin. Neurophysiol.*, vol. 10, p. 323, Oct. 1993.

[35] K. Watanabe, "Neurophysiological aspects of neonatal seizures," *Brain Develop.*, vol. 36, no. 5, pp. 363–371, May 2014.

[36] H. Kidokoro, T. Inder, A. Okumura, and K. Watanabe, "What does cyclicity on amplitude-integrated EEG mean?" *J. Perinatol.*, vol. 32, no. 8, pp. 565–569, Aug. 2012.

[37] J. P. Turnbull, K. A. Loparo, M. W. Johnson, and M. S. Scher, "Automated detection of tracé alternant during sleep in healthy full-term neonates using discrete wavelet transform," *Clin. Neurophysiol.*, vol. 112, no. 10, pp. 1893–1900, Oct. 2001.

[38] W. O. De, M. Lavanga, A. Dorado, K. Jansen, A. Dereymaeker, and G. Naulaers, "Complexity analysis of neonatal EEG using multiscale entropy: Applications in brain maturation and sleep stage classification," *Entropy*, vol. 19, p. 516, Oct. 2017.

[39] A. Ansari, H. De Wel, O. Pillay, K. Dereymaeker, A. Jansen, K. Van Huffel, S. Naulaers, G. and M. De Vos, "A convolutional neural network outperforming state-of-the-art sleep staging algorithms for both preterm and term infants," *J. Neural Eng.*, vol. 17, no. 1, Jan. 2019, Art. no. 016028.

[40] M. Awais, C. Chen, X. Long, B. Yin, A. Nawaz, S. F. Abbasi, S. Akbarzadeh, L. Tao, C. Lu, L. Wang, R. M. Aarts, and W. Chen, "Novel framework: Face feature selection algorithm for neonatal facial and related attributes recognition," *IEEE Access*, vol. 8, pp. 59100–59113, 2020.

[41] S. F. Abbasi, M. Awais, X. Zhao, and W. Chen, "Automatic denoising and artifact removal from neonatal EEG," in *Proc. Int. Conf. Biol. Inf. Biomed. Eng.*, Feb. 2019, pp. 1–5.

[42] Y. Kocyigit, A. Alkan, and H. Erol, "Classification of EEG recordings by using fast independent component analysis and artificial neural network," *J. Med. Syst.*, vol. 32, no. 1, pp. 17–20, Feb. 2008.

[43] A. Subasi, "EEG signal classification using wavelet feature extraction and a mixture of expert model," *Expert Syst. Appl.*, vol. 32, no. 4, pp. 1084–1093, May 2007.

[44] D. P. Mandic, "A generalized normalized gradient descent algorithm," *IEEE Signal Process. Lett.*, vol. 11, no. 2, pp. 115–118, Feb. 2004.

[45] G. Hinton, N. Srivastava, and K. Swersky, "Neural networks for machine learning lecture 6a overview of mini-batch gradient descent," *Cited*, vol. 14, no. 8, pp. 1–31, 2012.

[46] G. Cybenko, "Approximation by superpositions of a sigmoidal function," *Math. Control, Signals, Syst.*, vol. 2, no. 4, pp. 303–314, Dec. 1989.

AHSEN TAHIR received the Ph.D. degree from Glasgow Caledonian University. He is currently working as an Assistant Professor with the Department of Electrical Engineering, University of Engineering and Technology, Lahore, Pakistan. He also develops re-configurable embedded hardware accelerators for machine and deep learning applications. His research areas are machine learning and deep learning for healthcare applications.

MUHAMMAD AWAIS (Member, IEEE) received the master's degree from Universiti Teknologi PETRONAS, Malaysia, in 2014 and 2016. He is currently pursuing the Ph.D. degree with the Center for Intelligent Medical Electronics (CIME), Department of Electronic Engineering School of Information Science and Technology, Fudan University, Shanghai. He received a two-year funded position in Universiti Teknologi PETRONAS, and a Grant Scholarship (STRIF) from the Minister of Education, Malaysia, for an Excellent and an Efficient Research, in 2015. His research interests include biomedical image processing, signal processing, health informatics, image processing and pattern recognition, unobtrusive-monitoring, the Internet-of-Things, health care, personalized, and smart environment.

CHEN CHEN received the M.S. degree in embedded system from the Institut Supérieur d'Electronique de Paris (ISEP), in 2013, and the Ph.D. degree in computer science from Université Pierre et Marie Curie (UPMC), in 2016. She is currently a Postdoctoral Researcher with the Centre of Intelligent Medical Electronics, School of Information Science and Technology, Fudan University. Her research interests include biomedical engineering, focusing on electroencephalography monitoring, biomedical signal processing, sleep analysis, wearable sensor systems, and personalized health monitoring.

SAADULLAH FAROOQ ABBASI received the bachelor's degree from Muhammad Ali Jinnah University (MAJU), Islamabad, in 2014, and the M.S. degree in electrical engineering from the National University of Science and Technology (NUST), Islamabad, in 2017. He is currently pursuing the Ph.D. degree with the Centre of Intelligent Medical Electronics, School of Information Science and Technology, Fudan University. His research interests lie in biomedical engineering, focusing on electroencephalography monitoring, biomedical signal processing, sleep analysis, image processing, and neonatal health monitoring.

MUHAMMAD IRFAN received the B.Sc. degree in electronic engineering from the University of Engineering and Technology Taxila, Pakistan, in 2015. He is currently pursuing the master's degree in electronics and communication engineering with Fudan University, Shanghai, China.

From September 2015 to October 2016, he worked with Hunza distillery plant Faisalabad and Pak Agro Tech as Assistant Engineer and a Trainee Engineer, respectively. He also has collaborated in 02 industrial projects with Bear net and Nile intelligent companies, Shanghai. His research interests include the IoT, smart sensor systems, and wireless.

JAWAD AHMAD (Senior Member, IEEE) is currently an Experienced Researcher with more than ten years of cutting-edge research and teaching experience in prestigious institutes, including Edinburgh Napier University, U.K., Glasgow Caledonian University, U.K., Hongik University, South Korea, and HITEC University Taxila, Pakistan. He has coauthored more than 50 research articles, in international journals and peer-reviewed international conference proceedings. He has taught various courses both at bachelor's and master's levels during his career. He regularly organizes timely special sessions and workshops for several flagship IEEE conferences. He is also an invited Reviewer for numerous world-leading high-impact journals (reviewed more than 50 journal articles to date). His research interests include cybersecurity, multimedia encryption, machine learning, and application of chaos theory in cybersecurity.


HAFIZA AYESHA SIDDIQA received the B.Sc. and M.S. degrees in electronic engineering from Hitec University Taxila, Pakistan, in 2015 and 2018, respectively. She is currently pursuing the Ph.D. degree in electronics and communication engineering from Fudan University, Shanghai, China. Her research interest includes neonatal sleep staging.

ABU BAKAR WAQAS received the B.Sc. and M.Sc. degrees in electrical engineering from the University of Engineering and Technology Taxila, Pakistan, in 2011 and 2013, respectively. He is currently pursuing the Ph.D. degree with Fudan University, Shanghai, China. He served his Alma Mater as a Lecturer with Fudan University, for four years. His research interests include quantum communications, biomedical signal and image processing, and optimization of electrical engineering problems.

XI LONG (Member, IEEE) was born in China, in 1983. He received the B.Eng. degree in electronic information engineering from Zhejiang University, Hangzhou, China, in 2006, and the M.Sc. and the Ph.D. *(cum laude)* degrees in electrical engineering from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 2009 and 2015, respectively.

From 2010 to 2011, he was with Tencent, China, on data mining and user research. He has ten years of Research and Development experience in healthcare industry, with Philips Research, Eindhoven. He is currently a Senior Scientist with Philips Research and an Assistant Professor with the Eindhoven University of Technology. His research interests include signal processing and machine learning in healthcare and biomedicine applications. He has published more than 80 articles and reports in these fields and his inventions led to more than ten patent applications. He received the Best Student Paper Award from the IEEE BIBE Conference and the First Runner-up Paper Award from the IEEE BHI Conference, in 2012.
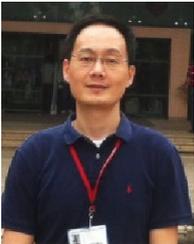
BIN YIN received the Ph.D. degree in electrical engineering from Southeast University, Nanjing, China. He is currently a Principal Scientist leading the Connected Care Team with Philips Research, Shanghai. After a two-year Postdoctoral Research Program with the Eindhoven University of Technology, he joined Philips Research, Eindhoven, The Netherlands. His research interests include monitoring technologies, wearables, signal processing, and intelligent algorithms, with recent applications in chronic disease management and patient engagement. He holds a part-time professorship at the School of Information Science and Technology, Fudan University, Shanghai.

SAEED AKBARZADEH (Member, IEEE) was born in Iran, in 1989. He received the degree in biomedical engineering from the University of Teheran Shomal, Tehran, Iran. He is currently pursuing the Ph.D. degree in biomedical engineering with the Center for Intelligence Medical Electronic (CIME), Fudan University, Shanghai, China. His scientific interests have always been concerned with the design and develop the electrical and mechanical systems, especially in biomedical engineering and neuroscience. His current research activities are particularly focused on the development and characterization of wearable devices and smart sensorized systems for biomedical engineering and the IoT.

CHUNMEI LU received the bachelor's degree from Shanghai Jiao Tong University, in 2014. She is currently working as the Head Nurse with the Children's Hospital of Fudan University. She co-edited two books and published ten articles. Her professional fields include neonatal nursing management, neonatal vascular access management and post-discharge follow-up, and early intervention of high-risk infants. She won the Second Prize of the Science and Technology Award of the Chinese Nursing Association on Research on the Application of Key Treatment Techniques based on NIDCAP Theory in the nursing of premature infants.

LAISHUAN WANG is currently the M.D., the Chief Physician, a Ph.D. Tutor, and the Deputy Director of the Neonatal Department, Fudan University Affiliated Paediatric Hospital. He also serves as the Deputy Director of the Neonatal Division of the Chinese Medical Science Branch, the Standing Committee of the Special Committee of infant and young child mental health of the Chinese Society for Maternal and Child Health, a member of the Perinatal Critical Medicine Group, Eighth Committee of the Perinatal Medical Branch, the Chinese Medical Association, the Secretary of the Disaster Division of the Chinese Medical Science Branch, and the Neonatologist of the Shanghai Paediatric Society. He is a Young Member and the Secretary of the Shanghai Perinatal Medical Association, youth Editorial Board of the *Chinese Journal of Contemporary Pediatrics* and other magazines. He is also undertake the National Ministry of Science and Technology Key Research and Development Program topics and the National Natural Science Foundation and other topics. He is also good at the diagnosis and treatment of neonatal acute and difficult diseases and the neurodevelopmental assessment and early intervention of children with various brain injury. He has won the Ministry of Education Science and Technology Progress Second Prize, the Chinese Medical Science and Technology Award, and the Shanghai Science and Technology Progress Second Prize.

WEI CHEN (Senior Member, IEEE) received the B.Eng. and M. Eng. degrees from the School of Electronics and Information Engineering, Xian Jiaotong University, China, in 1999 and 2002, respectively, and the Ph.D. degree from the Department of Electrical and Electronics Engineering, The University of Melbourne, Parkville, VIC, Australia, in 2007. She worked with Bell Laboratories, Germany, Alcatel-Lucent, Stuttgart, as an Intern, in 2005. She was a Research Assistant with the Department of Electrical and Electronics Engineering, The University of Melbourne, in 2007. From 2007 to 2015, she was an Assistant Professor with the Eindhoven University of Technology, The Netherlands. From 2009 to 2013, she has served as the Chair of the Theme Health Care at the Department of Industrial Design, Eindhoven University of Technology, The Netherlands. Since October 2015, she has been a Full Professor and the Director of the Center for Intelligent Medical Electronics (CIME) with the Department of Electronic Engineering, School of Information Science and Technology, Fudan University. Her research interests include patient health monitoring, medical monitoring system design using wearable sensors, sleep monitoring, brain activity monitoring, wireless body area networks, ambient intelligence, personalized and smart environment, smart sensor systems, and signal processing. She is an Associate Editor of IEEE Transactions on Neural Systems and Rehabilitation Engineering (TNSRE), and IEEE Journal on Biomedical Health Informatics (JBHI), and the Managing Editor of IEEE Reviews in Biomedical Engineering (R-BME).

• • •