RESEARCH ARTICLE

A Deep Learning-Based Brain Age Prediction Model for Preterm Infants via Neonatal MRI

JIAJIE TANG[1,2,3], **PU YANG**[2], **BINGBING XIE**[1], **CONG WEI**[2], **LIANTING HU**[4], **SHOUYI WANG**[2], **YUXUAN JIANG**[1], **RUIZHUO LI**[5,6], **HANG ZHOU**[5,6], **HAIQING XU**[7], **QIRONG WAN**[8], **JIN HAN**[1,3,5,6], **DONGCHI ZHAO**[2], AND **LONG LU**[1,5,6,9,10]

[1]School of Information Management, Wuhan University, Wuhan 430072, China
[2]Children's Health Big Data Research Center, Zhongnan Hospital of Wuhan University, Wuhan 430071, China
[3]Obstetrics and Gynecology Medical Center, Dongguan Kanghua Hospital, Dongguan 523080, China
[4]Medical Big Data Center, Guangdong Provincial People's Hospital, Guangzhou 510317, China
[5]Prenatal Diagnosis Center, Institute of Pediatrics, Guangzhou Medical University, Guangzhou 510180, China
[6]Guangzhou Women and Children's Medical Center, Guangzhou Medical University, Guangzhou 510180, China
[7]Hubei Maternal and Children Hospital, Wuhan 430090, China
[8]Department of Psychiatry, Renmin Hospital of Wuhan University, Wuhan 430064, China
[9]School of Public Health, Wuhan University, Wuhan 430072, China
[10]Center for Healthcare Big Data Research, Big Data Institute, Wuhan University, Wuhan 430072, China

Corresponding authors: Long Lu (whumil@126.com), Dongchi Zhao (dczhao@whu.edu.cn), and Jin Han (hanjin1123@163.com)

ABSTRACT The accurate, quantitative, and objective prediction of the brain age for premature infants will contribute to the exploration of brain maturity and catch-up growth. Traditional approaches rely heavily on a pediatrician's clinical experience, which makes the whole process time-consuming and labor-intensive. To solve this problem, we propose a deep learning-based brain age prediction model for preterm infants via neonatal MRI for this purpose, and it is called as BAPNET for short. First of all, we collected a specific dataset including MR images of 281 preterm infants. Then, a pretraining model (DeepBrainNet) is applied as the main backbone, and transfer learning is utilized to enhance the baseline model by making knowledge transfer from the ImageNet dataset. The proposal can be viewed as a specific prediction model by absorbing knowledge enhancement from peripheral visual features. On a test set of 70 preterm infants held out from the original dataset, 2D-BPANET achieved results with an mean square error (MAE) of 1.15 and the 95% - 95% content tolerance interval for a difference (prediction and ground truth) of [-3.82, 3.39], whereas 3D-BPANET achieved better results with an MAE of 1.8 and a difference of [0.51, 3.09]. Meanwhile, we leverage heatmaps to verify the consistency between hindbrain regions and cortical fold regions outputted by our model and the latest studies of brain development in preterm infants. In conclusion, BPANET demonstrates that deep learning can estimate brain maturity in preterm infants and provides a reference standard for preterm infant brain development, which could be applied as a promising tool.

INDEX TERMS Deep learning, medical image processing, prediction model, MRI, medicine and biology.

I. INTRODUCTION

Preterm infants are newborns weighing 1000-2499g at less than 37 weeks of gestation [1]. With the advancement of reproductive medicine and perinatal medicine, as well as the introduction of neonatal intensive care units, the prevalence of preterm birth has increased in recent years. Related research shows that approximately 10.6% of live births were born preterm and require special care [2], [3]. Preterm infants are susceptible to brain damage such as intracranial hemorrhage, hypoxia, and ischemia due to neurodevelopmental impairment (NDI), which could lead to delay and impairment in

The associate editor coordinating the review of this manuscript and approving it for publication was Alessandra Bertoldo.
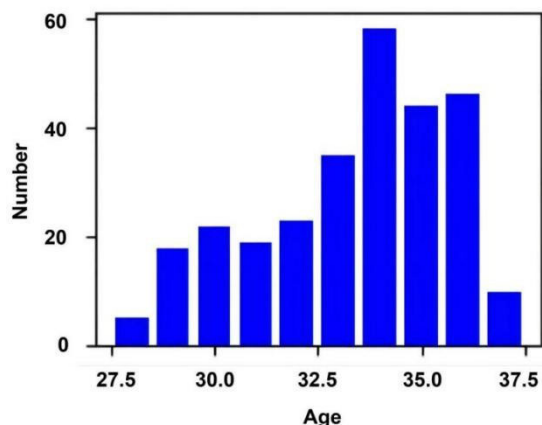
FIGURE 1. Gestational age/brain age distribution (cases per weeks).

brain development. It may cause neurological disorders such as epilepsy and cerebral palsy [4]. Therefore, real-time monitoring for the brain development status of preterm infants, detection of lagging brain development, and early intervention are extremely important to stimulate brain development and improve the survival quality of preterm infants.

In hospital neonatal units, the mainstream method for determining brain age in preterm and neonatal infants is still based on the total maturation scores (TMS) scoring system, which is a simple scoring system for MR images that determines the TMS score by assessing four parameters of brain maturation: myelin formation, cortical folding, glial cell migration, and germinal matrix distribution [5]. Some studies used intracranial structures to determine brain age, such as the width measurement of the superior frontal gyrus in newborns, and others used Magnetic Resonance Imaging (MRI) to measure the volume of each brain structure, such as the volume calculation of the ventricular system, cerebellum, basal ganglia, corpus callosum, amygdala, and hippocampus, and the distribution observation of gray matter and white matter, as well as quantifying cortical [6], [7], [8], [9]. However, it is difficult to guarantee accuracy because of the individual differences in preterm infants and the high requirement, subjectivity, as well as time consumption of the aforementioned methods.

Artificial intelligence (AI) is a critical tool in the ongoing personalization and development of precision medicine [10]. Recently, there has been a huge rise in medical imaging-related machine learning research, which can aid clinicians in diagnosis and prognosis [11], [12], [13]. Many researchers utilized machine learning methods such as Support Vector Machine (SVM) to predict brain age using feature representations of brain white matter, gray matter, cortex, and sulcal gyrus [14], [15], [16], [17].

However, feature extraction and selection need a significant amount of prior knowledge. With the development of computer vision technology, scholars have made significant progress in automatically extracting visual characteristics

for brain age prediction with convolutional neural networks (CNN) [18], [19], [20], [21]. There are a large number of studies focusing on adult and elderly brain age prediction but none of them can be used for preterm infants due to the scarcity and uniqueness of preterm infants' data. Notably, a study found that the CNN model may predict motor outcomes in preterm newborns using brain diffusion MRI [22], which shed light on the possibility of developing a prediction model with CNN for brain age in preterm infants. Until now, there has been no relevant literature for preterm infant brain age forecasting with deep learning based on MRI in Google Scholar and PubMed.

In this paper, we collected 281 MR images from preterm to develop the first deep learning-based automated model named BAPNET. At first, we preprocessed images by SPM12 and extended our initial dataset by data augmentation. For the 2D approach, we applied transfer learning starting from an already existing model (DeepBrainNet) trained on 2D slices. For the 3D approach, we applied transfer learning from ImageNet to pretrain 3D CNN models, and the database of preterms was then exploited for optimization. The 2D-BAPNET and 3D-BAPNET models devised in this study are applicable to 2D images/slices and volumes. They were evaluated on test sets held out from the original dataset. The Grad-CAM [23] visualization technique was finally used to highlight the highly contributing regions.

In this paper, we collected 281 MR images from preterm to develop the first deep learning-based automated model named BAPNET. At first, we preprocessed images by SPM12 and extended our initial dataset by data augmentation. For the 2D approach, we applied transfer learning starting from an already existing model (DeepBrainNet) trained on 2D slices. For the 3D approach, we applied transfer learning from ImageNet to pretrain 3D CNN models, and the database of preterms was then exploited for optimization. The 2D-BAPNET and 3D-BAPNET models devised in this study are applicable to 2D images/slices and volumes. They were evaluated on test sets held out from the original dataset. The Grad-CAM [23] visualization technique was finally used to highlight the highly contributing regions.

II. MATERIALS AND METHODS
A. DATA ACQUISITION AND QUALITY CONTROL
This research utilized T1-weighted images of 281 preterm infants without intracranial injury recruited with informed consent at Zhongnan Hospital of Wuhan University, Wuhan, China, from May 2017 to October 2020. They aged between 27 and 37 weeks and averaged 33.4 weeks of age (**Fig. 1**). The parameters of the equipment are as follows, FoV (Field of View) read: 180mm, FoV phase: 100%, TR (repetition time): 440ms, TE (echo time): 2.68ms. A team of specialists with more than five years of experience conducted all examinations prior to the research. The examination was performed within one week of the preterm infant's birth, and the gestational age was corrected to guarantee that the

gestational age and brain age are equal. The reference standard gestational age (brain age) was estimated from the last menstruation date and the MRI measurements performed by the pregnant mother's referring obstetrician. All preterm infants with underlying brain disorders, psychiatric disorders, neurological disorders, and structural brain abnormalities were excluded. In addition, we excluded (n = 79) from the initial list of 360 cases without satisfactory images and complete clinical data. Thus, our dataset contained 281 T1-weighted images. The MRI datasets was divided into two separate sections: a training set used to optimize the model parameters and tune the model hyperparameters, and a test set used to evaluate the performance of the model relative to the reference standard. The division of the training and test sets is randomized, and we finally obtain a training set (for training and tuning) of 211 cases and a test set (for testing) of 70 cases.

B. MRI PREPROCESSING

The preprocessing of the T1-weighted images was implemented on MATLAB with applications of the Statistical Parametric Mapping program (SPM12). The images of the training set and test set need to be handled according to the following preprocessing steps.

1) DATA CONVERSION

The original Digital Imaging and Communications in Medicine (DICOM) file format of the T1-weighted images needs to be converted into the three-dimensional Neuroimaging Informatics Technology Initiative (NIfTI-1) file format. Since SPM8 uses NIfTI-1 as the image format, this conversion will be convenient for SPM8 to implement the subsequent image preprocessing.

2) RESETTING THE MR IMAGE ORIENTATION AND CORRECTING THE POSITION OF THE AC-PC LINE

There are three views using SPM12 to read images including coronal, axial, and sagittal. First, adjust the roll value (rotate around the y-axis) so that the crosshair passes through the center of the coronal plane, and then adjust the yaw value (rotate around the z-axis) so that the crosshair passes through the center of the axial plane. Finally, in the sagittal plane, find the clearest point of the corpus callosum (point AC) and adjust the pitch value (rotate around the x-axis) to position the crosshair to point AC.

3) SPATIAL NORMALIZATION

Due to the variety of shapes and volumes of preterm infants' brains, the T1-weighted images should be normalized and resampled to Montreal Neurological Institute (MNI) space, and the XYZ values are used to describe a particular coordinate position.

All of the images were resized to a uniform size (79 × 95×79) after data preprocessing and were trained in the 3D-BAPNET network built by this research. To acquire 2D data, we extracted 2D slices from the axial plane and 40 slices

beginning at index 23 and ending at index 63, respectively, because the middle part of the image contains the most information. It is worth noting that the location and number of slices, which were commonly employed for brain development assessment, are determined by the senior physician. Finally, the original 281 cases of 3D data samples were changed into 9721 segmented 2D slices according to the 2D extraction technique. During the training process, each 2D slice was an independent sample. To obtain the final age prediction for a test sample, each of 40 slices of the test scan was input to the trained model independently and the median prediction was calculated as the predicted brain age [24]. The comparison chart of pre-processing is shown in **Fig. 2**.

C. DATA AUGMENTATION

In medical imaging tasks, data augmentation has various advantages like training data distribution enrichment, the model's generalization ability improvement, and overfitting removal in training process. In our experiments, we employed four common methods of data augmentation to enhance the 2D training dataset including: (a) distortion, (b) zoom in and zoom out, (c) tilt, (d) crop.

D. DEVELOPED 2D-BAPNET

The deep brain network (DeepBrainNet) was built using a large (n = 11729) set of MRI scans from a highly diversified cohort spanning different studies, scanners, ages, and geographic locations around the world [21]. The DeepBrainNet model was built using the Inception-Resnet-v2 framework, which combines skip connections and inception modules. It used 2D slices as input and brain age as a label to learn brain images from 3 to 95 years of age, learning the process of brain growth and aging, which is suitable as a pre-training model for brain age prediction in preterm infants. We applied transfer learning starting from an already existing model (DeepBrainNet) trained on 2D slices from preterm infant datasets. For comparison, we used the same datasets to train DenseNet-169 and ResNet-101, which are often used for brain age prediction tasks. Adam Optimizer was used to optimization during training, the initial learning rate was set to 1e-4, the weight decay rate was 1e-7, the batchsize was set to 40, the epoch was set to 200.

E. DEVELOPED 3D-BAPNET

Since 3D-CNN can be capable of learning the information of the whole 3D image, we also developed the 3D-BAPNET to predict the brain age of preterm neonates. For obtaining the best strategy to train 3D-BAPNET, three state-of-the-art CNN architectures (DensNet-169, ResNet-101, and ResNet-152) were investigated in this study. Weights pre-trained for ImageNet classification were employed to initialize the CNN architectures.

(1) Residual module (ResNet)

$$y = \sigma \left(F \left(x, W \right) + x \right) \tag{1}$$
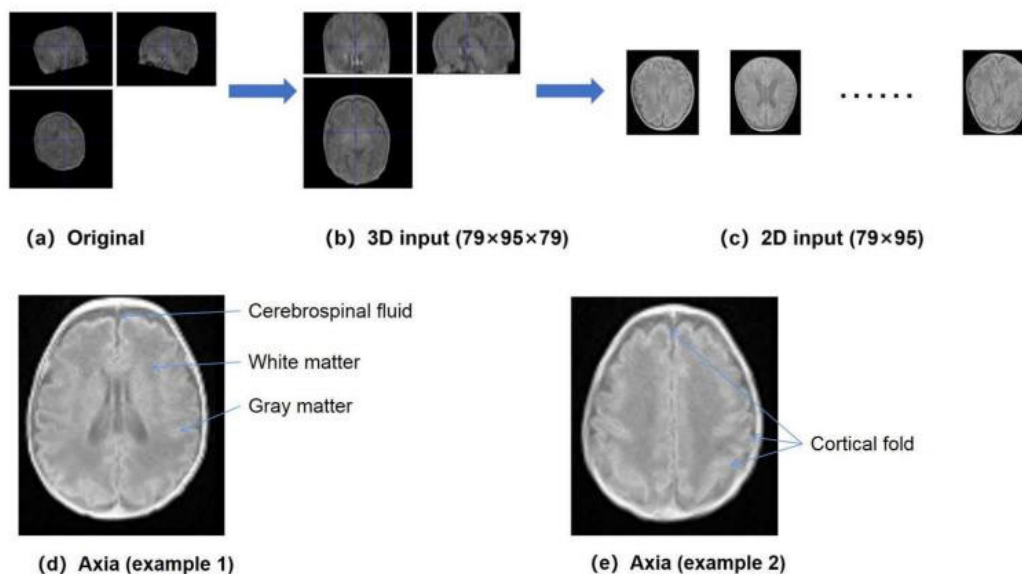
FIGURE 2. Example of data pre-processing: (a) The NIfTI-1 file format of the T1-weighted images. (b) Processed image for 3D-CNN input. (c) Slices for 2D-CNN input. (d) Axia (example 1). (e) Axia (example 2).

y represents the output of residual blocks; $\sigma(\cdot)$ represents the activation function; $F(\cdot)$ represents the residual function; x represents the input and W represents the weights in the residual block.

(2) Dense connection mechanism (DenseNet)

$$x_\iota = H_\iota\left(\left[x_0, x_1, \ldots, x_{\iota-1}\right]\right) \tag{2}$$

ι represents the number of convolutional layers; x_ι represents the output of layer ι, and H_ι represents a nonlinear transformation, which is channel merging.

All the functionality, experiments, and analysis were implemented used Python (NumPy 1.16, for array manipulation; opencv-python 4.1.0 and Pillow 6.0 for image operations; and scikit-learn 0.19.1 for performance quantification) and Google Tensorflow (for the implementation of the deep learning architecture). The overview of the brain age prediction process is shown in **Fig. 3**. The detailed structure of the model is presented in Supplementary Material and Data Availability.

F. STATISTICAL ANALYSIS AND EXPERIMENTAL FRAMEWORK

The deep learning models were developed with TensorFlow, Keras and Python. The original patient data were divided into a training set and a test set in the experiments. TensorFlow, Scikit-learn, and Python were used for statistical analysis. The performance of the BAPNET was evaluated by calculating the MAE, RMSE, and r on the test set. We also calculated the 95% - 95% content tolerance interval and 95% prediction interval for a difference between BAPNET and the ground truth (true value).

III. RESULTS

A. DATASET CHARACTERISTICS

To develop a deep learning model for predicting the brain age of preterm neonates using routine clinical brain MR images, we enrolled 281 preterm infants aged 28 to 37 weeks (**Fig. 1** shows the distribution of participants). This was a retrospective study in which each subject received an MRI scan of the head after birth. The holdout method was employed to randomly divide the 281 MR images into two parts, one part with 211 MR images used for training and tuning, and the other part containing 70 images as a test datasets.

B. PERFORMANCE OF THE 2D-BAPNET

The 2D-BAPNET achieved MAE of 1.15 weeks, RMSE of 1.57 weeks, R of 0.72 in the test set. The MAE of the ResNet101 and DenseNet169 were 1.22 and 1.51, respectively, while the RMSE could only reach 1.60 and 1.96, indicating that transfer of brain knowledge has benefits to predict brain age. MSE, R, and other metrics are shown in **Table 1**. Compared to the reference standard, 2D-BAPNET was considered to achieve high accuracy in predicting the brain age of preterm infants aged 28-37 weeks.

Preterm infants were divided into three groups based on their gestational week of birth including early preterm infants (28-32 weeks), moderately preterm infants (32-34 weeks), and late preterm infants (34-37 weeks). Preterm infants require appropriate treatment protocols at different stages. It is necessary to divide the age into three age groups based on gestational weeks and separately evaluate the prediction results of the three groups. The MAE of Group A, B, and C were 2.60, 0.53, and 1.64, respectively. **Table 2** shows the assessment results from 2D-BAPNET. We also observed that most predictions for moderately preterm infants and late
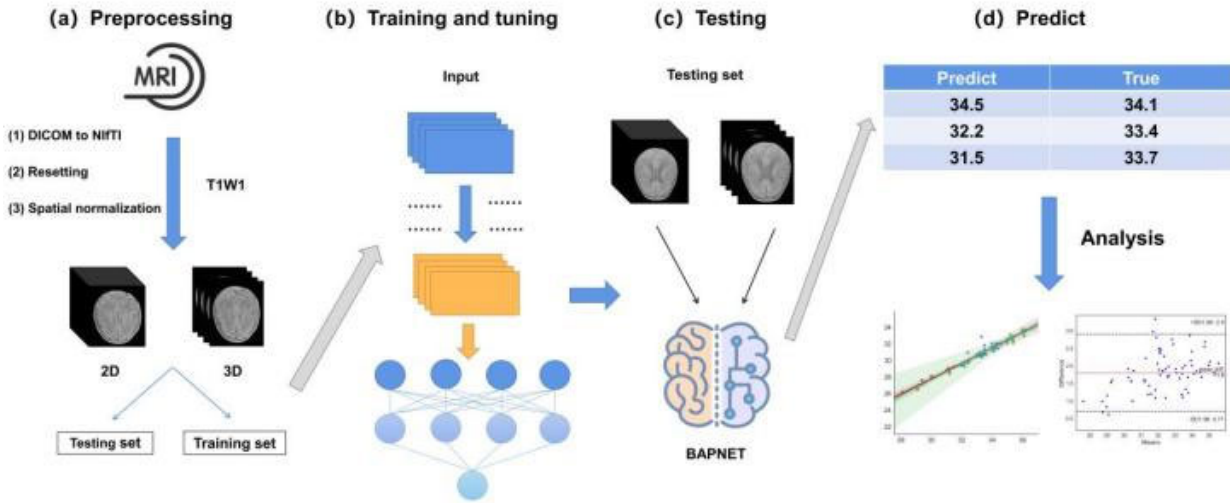
FIGURE 3. Overview of the brain age prediction process. (a) Preprocessing. (b) Training and tuning. (c) Testing. (d) Predict.

TABLE 1. Model performance in predicting the brain age of preterm infants.

		MAE	MSE	RMSE	r
3D architecture	3D-BAPNET	1.80	3.56	1.89	0.96
	DenseNet169	2.46	6.21	2.49	0.98
	ResNet152	2.10	4.62	2.15	0.97
2D architecture	2D-BAPNET	1.15	2.47	1.57	0.72
	ResNet101	1.22	2.55	1.60	0.71
	DenseNet169	1.51	3.85	1.96	0.16

preterm infants are closer to the true values compared with others in **Fig. 4**.

To further assess the reliability of the predicted results, we drew the residual plot (Bland-Altman plot). The residual plot shows the relationship between the mean and the difference between the predicted and actual value, which is shown in **Fig. 5**. We also plot the tolerance interval for the difference, which is shown in **Fig. 6**. The 95%-95% content tolerance interval and 95% prediction interval for a difference in all groups is [-3.82, 3.39] and [-3.37, 2.93].

C. PERFORMANCE OF THE 3D-BAPNET
In 3D-CNN, ResNet101 achieved the optimal performance with MAE of 1.80 weeks, RMSE of 1.88 weeks, R of 0.96 in the test set, which was named as 3D-BAPNET. The MAE of the ResNet152 and DenseNet 169 were 2.10 and 2.46, respectively, while the RMSE could only reach 2.15 and 2.50. To further assess the superiority of our model, we also plotted

the residuals (Bland-Altman plot), as shown in **Fig. 5 (B)**. The 95%-95% content tolerance interval and 95% prediction interval for a difference in all groups is 0.51-3.09, 0.67-2.93, respectively. We were able to find that BAPNET-3D provides upper predictions than the true values for all values.

D. THE RELIABILITY OF 3D-BAPNET
In comparison of the 2D-BAPNET and 3D-BAPNET model, the MAE of 1.15 for 2D-BAPNET are lower than 1.80 for 3D-BAPNET. However, the R of 0.9 for 3D-BAPNET is significantly higher than 0.7 for 2D-BAPNET, and it can be seen that the RMSE/MAE in the 2D-BAPNET are larger than that in the 3D-BAPNET. As shown in **Fig. 5** and **Table 2**, the performance of 2D-BAPNET for predicting early preterm group is obviously insufficient (MAE = 2.60), which is also the reason why RMSE is significantly greater than MAE. To deeply analyze the prediction performance of 2D-BAPNET and 3D-BAPNET, we plotted the Bland-Altman plots. In general, 3D-BAPNET has a smaller 95%-95% content tolerance interval and 95% prediction interval, and has a stable distribution of differences in predicted values for each group. There is a trend in the 2D-BAPNET, it systematically provides over-estimation predictions for low values and under-estimation for high values, 3D-BAPNET looks better. It is worth mentioning that the predicted values in the 2D-BAPNET were concentrated in the interval [32, 35] and had the worst predictions in the early group, while the 3D-BAPNET achieved the best predictions in the early group.

## E. HEATMAPS HIGHLIGHT THE REGIONS OF HINDBRAIN AND CORTICAL FOLD
To visualize the regions contributing to the BAPNET, we generated a heatmap that superimposed a visualization layer at the end of the CNN [23]. We found that heatmaps highlighted the regions of the hindbrain and cortical fold. To verify the
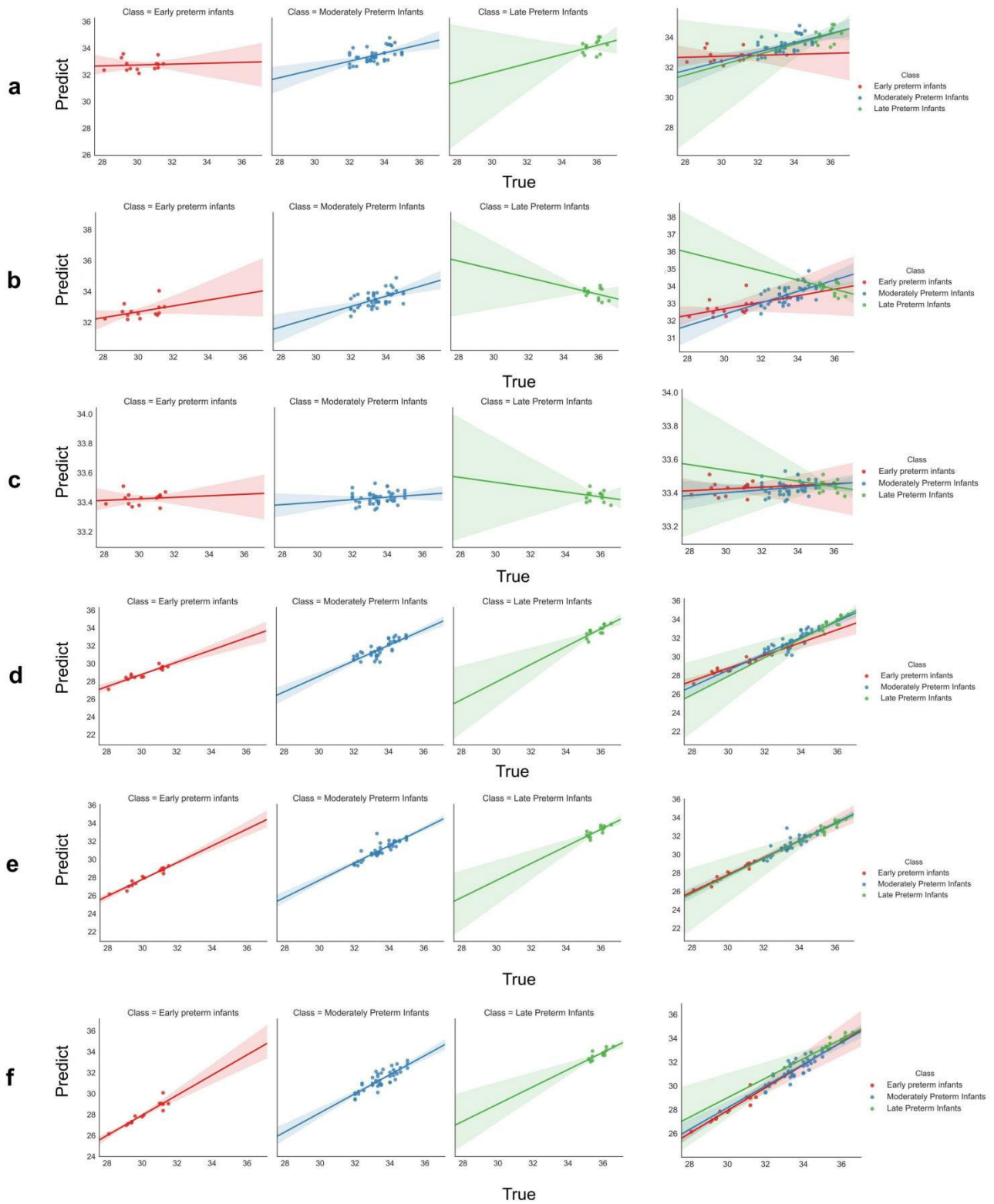
FIGURE 4. Scatterplots and regression lines showing the correlation between predictions from the model (Predict) and the reference standard (True).
(a) 2D-BAPNET. (b) 2D ResNet101. (c) 2D DenseNet169. (d) 3D ResNet152 (e) 3D DensNet169 (f) 3D-BAPNET. Red: Early preterm infants; Blue:
moderately preterm infants; Green: late preterm infants. Six experiments were conducted to compare the regression performances of different
models. The color represents the regression performance of different groups of premature infants.

A   B
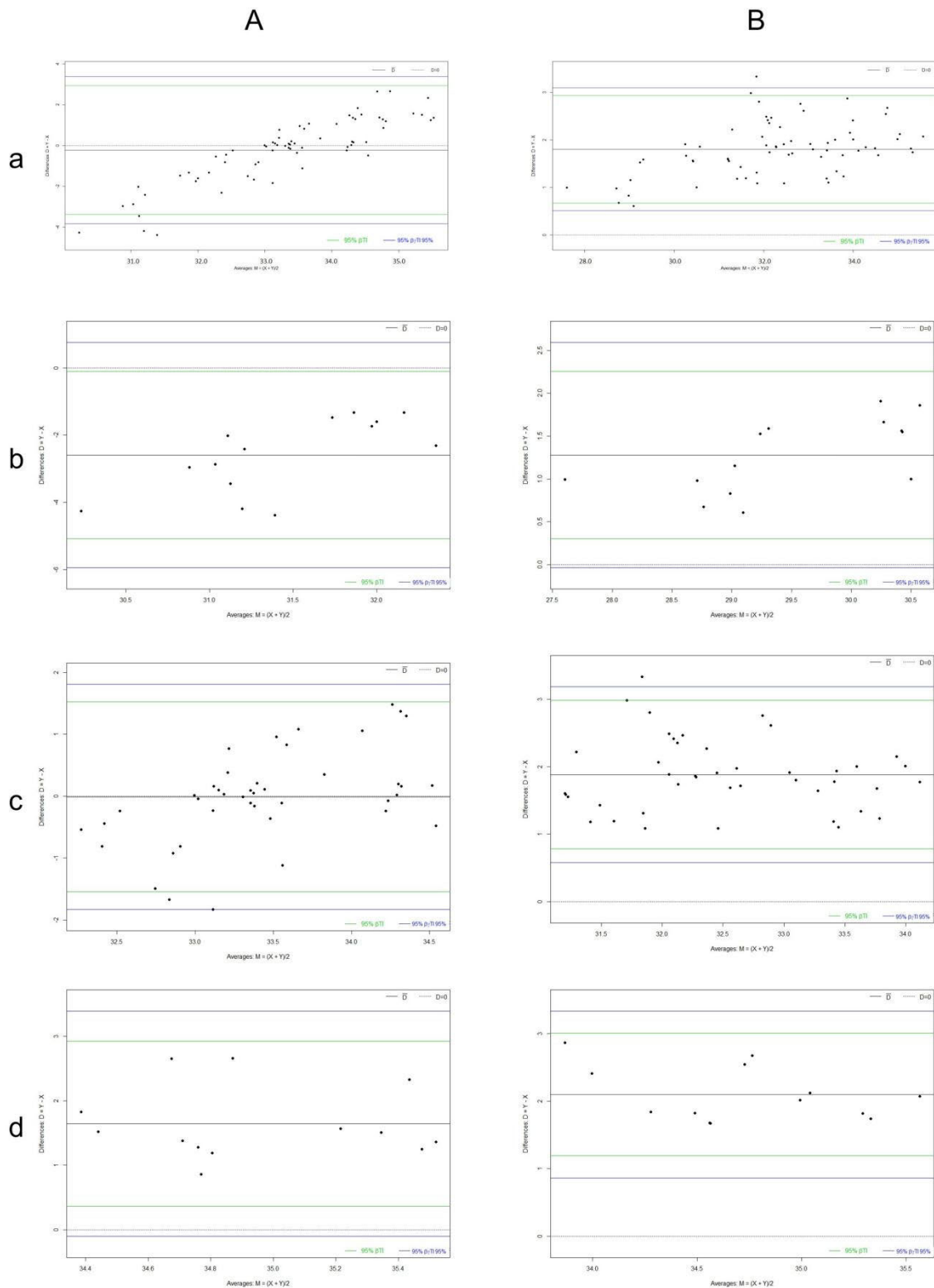
FIGURE 5. Bland-Altman plots compare brain age predictions between the 2D-BAPNET or 3D-BAPNET and the reference standard. The black solid line represent the difference, and the green solid line and blue solid line represents the 95% prediction interval and the 95%-95% content tolerance interval, respectively. "Y" represents the true value and "x" represents the predicted value. table 2 records the value of the 95% prediction interval and the 95% - 95% content tolerance interval. Take b (a) as an example, the 95% prediction interval is [0.67, 2.93], which means that one is 95% confident that the next infant will be measured with a difference between 0.67 and 2.93 between the predictive value and ground truth. And the 95%-95% tolerance interval is [0.51, 3.09], which means that one is 95% confident that at least 95% of the future differences between the predictive values and true values are expected to lie between 0.51 and 3.09. (a) 2D-BAPNET. (b) 3D-BAPNET. (a) all groups. (b) early preterm groups. (c) moderately preterm group. (d) late preterm groups.
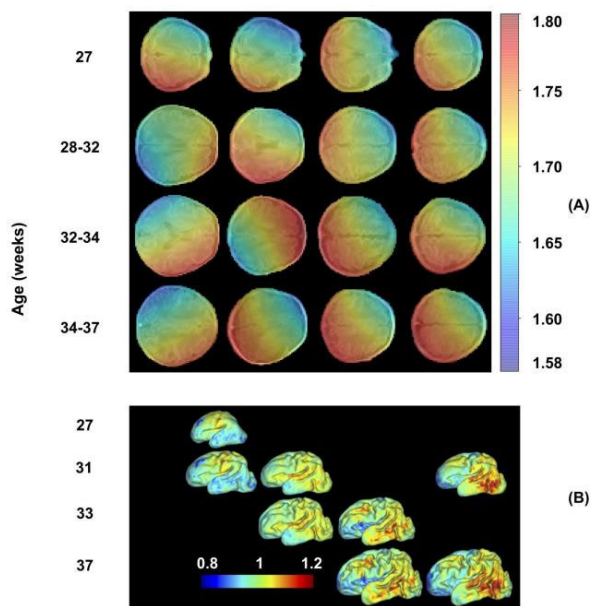
FIGURE 6. Example of heat maps from BAPNET and comparison with the study of Kara et al. [25]. The heat map illustrates the importance of local areas within the image for prediction. The importance value is scaled between 1.58 and 1.80, where a higher number indicates that the area is of higher importance. (A) Heat map from BAPNET. (B) Specific growth maps for individual brains in preterm infant from Kara et al.

accuracy of the model's heat map, we compared it with a recent study that proposed a unique method to estimate local growth from sequential cortical reconstructions in preterm infants [25]. Examples of heat maps are shown in **Fig. 6**.

IV. DISCUSSION

In this study, deep learning was used to predict the brain age of preterm infants. To the best of our knowledge, this is the first study to focus on the use of CNN for predicting the brain age of preterm infants. Extensive experiments demonstrate that our deep learning neural networks-based model could predict the brain age of preterm infants accurately. In our internal test datasets consisting of brain images of preterm infants, the 3D-BAPNET is more reliable in predicting the brain age of preterm infants in all groups, the MAE and RMSE were 1.80 and 1.88, respectively. Although the MAE of 2D-BAPNET was only 1.15, the model is less stable because the MAE of early group is high as 2.60. Generally speaking, our BAPNET model shows practicality for using AI technologies to automatically estimate the brain age of preterm infants, with important implications for neonatal care and assessment of brain maturation in preterm infants.

As the above references report, Toews et al. constructed a brain age prediction model for infants between 8 and 590 days, with a final MAE of 72 days [26]. Hong et al. constructed a 3DCNN model using MRI of infants and children up to 5 years of age, and the MAE of the predicted outcome was 67.6 days [27]. Hu et al. built a deep learning model for infants aged 14-797 days with a model MAE of

TABLE 2. The performance of BAPNET in predicting the brain age of preterm infants in different groups.

		MAE	MSE	RMSE	Mean Difference (95%-95% Content Tolerance Interval)	Mean Difference (95% Prediction Interval)
3D	All groups	1.80	3.56	1.89	1.80 (0.51, 3.09)	1.80 (0.67, 2.93)
	Early group	1.28	1.81	1.35	1.28 (-0.04, 2.60)	1.28 (0.30, 2.25)
	Moderately group	1.88	4.55	1.96	1.88 (0.58, 3.18)	1.88 (0.78, 2.98)
	Late group	2.10	4.55	2.13	2.10 (0.86, 3.33)	2.10 (1.19, 3.01)
2D	All groups	1.15	2.47	1.57	-0.22 (-3.82, 3.39)	-0.22 (-3.37, 2.93)
	Early group	2.60	7.88	2.81	-2.60 (-5.95, 0.76)	-2.60 (-5.08, -0.11)
	Moderately group	0.53	0.55	0.74	-0.01 (-1.81,1.83)	-0.01 (-1.54, 1.52)
	Late group	1.65	3.00	1.73	1.65 (-0.10, 3.39)	1.65 (0.37, 2.93)

Note: Early group = Early preterm infants, Moderately group = Moderately preterm infants, Late group = Late preterm infants.

32.1 days [28]. However, since the brain development of preterm infants differs from that of infants and toddlers, the results of related studies can only be used as a reference. We used DeepBrainNet's pre-training parameters for transfer learning and obtained 2D-BAPNET. This is the first 2D-CNN model for assessing brain maturity in preterm infants, which was trained on the preterm infant data set for the brain age prediction task, and achieved the best results compared to other 2D networks. In particular, we developed three 3D-CNN models (ResNet and DenseNet), all used transfer learning. These models were initialized with network weights of the ImageNet [29], and further trained and tested for the same data set with 2D-BAPNET.

We also created heat maps by the Grad-CAM technique to visualize the regions that contributed the most to the prediction of the model [23]. As shown in **Fig. 6**, BAPNET highlights the hindbrain region, especially the cortical folding region. In some slices, model attention is focused on the volume and shape of the entire brain. The latest research shows that the human brain exhibits complex folding patterns that emerge during the third trimester of fetal development. In addition, the structural data suggest that growth might vary in both space (by region on the cortical surface) and time. We compared the heat map with a study that visualized cortical folding in preterm infants [25]. The results show that the regions of interest of BAPNET largely match with related studies, which implies that our model can also visualize the focal regions of brain development in preterm infants.

In this research domain, potential applications of BAPNET may include: (1) Study of brain development in preterm infants, with special reference to the temporal and spatial distribution of cortical folding; (2) Analysis of catch-up growth

in preterm infants; and (3) Research on brain disorders in infants and children, with special reference to brain degeneration. In the clinical domain, potential applications in the future could include: (1) It can be packaged as an easy-to-deploy brain analysis software for clinical in brain maturation assessment of preterm infants; (2) Deployed as a rapid, low-cost treatment tool to primary care institutions for graded care; (3) Large-scale diagnosis in backward areas.

Although the present study has proven the potential of BAPNET in predicting the brain age of preterm infants, this is the first step in the application of artificial intelligence technology to brain maturity assessment in preterm infants. Our model has several limitations which we wish to address in the near future: (1) Due to the uniqueness of the data on preterm infants and the difficulty of MRI image acquisition, only 281 MRI images from a single-center were included in our study. This is the reason for the poor stability of the 2DBAPNET model. In the future, we hope to conduct multicenter research. (2) Due to the unique dataset of preterm infants, it is not possible to rely on the existing brain atlas for segmentation of white matter, gray matter, and cerebrospinal fluid, which is why only T1-MRI was used in our study instead of multimodality data. (3) In this experiment, all preterm infants with underlying brain disorders, psychiatric disorders, neurological disorders, and structural brain abnormalities were excluded. In the subsequent research, we will focus on the improvement of data quantity and quality, as well as the improvement of the model performance.

V. CONCLUSION
To obtain a low-cost, non-invasive, robust, and deployable approach for accurate, quantitative, and objective prediction of the brain age for premature infants, we developed the first brain age prediction model for preterm infants, BPANET, to assess brain development and highlight the most significant locations of growth and development of preterm infants. It demonstrated that deep learning could precisely predict brain age from T1-weighted images of preterm infants. Prediction of brain age with deep learning has significant implications for the care and treatment of preterm infants. Our BAPNET model has the potential to be scalable into a quantitative tool for brain maturity estimation in preterm infants and is expected to be an objective reference for the catch-up growth of preterm infants.

REFERENCES
[1] A. Wilczyńska, T. Niemiec, and M. Troszyński, "Assessment of three-level selective perinatal care based on the analysis of early perinatal death rates and cesarean sections in Poland in 2008," *Ginekologia Polska*, vol. 80, no. 9, pp. 667–670, 2009, doi: 10.1055/s-0029-1186015.

[2] A. C. C. Lee, H. Blencowe, and J. E. Lawn, "Small babies, big numbers: Global estimates of preterm birth," *Lancet Global Health*, vol. 7, no. 1, pp. e2–e3, 2019, doi: 10.1016/S2214-109X(18)30484-4.

[3] *Born Too Soon: The Global Action Report on Preterm Birth*, World Health Org., Geneva, Switzerland, 2012.

[4] M. Carrasco and C. E. Stafstrom, "How early can a seizure happen? Pathophysiological considerations of extremely premature infant brain development," *Develop. Neurosci.*, vol. 40, nos. 5–6, pp. 417–436, 2018, doi: 10.1159/000497471.

[5] J. Gao, Q.-L. Sun, Y.-M. Zhang, Y.-Y. Li, H. Li, X. Hou, B.-L. Yu, X.-H. Zhou, and J. Yang, "Semi-quantitative assessment of brain maturation by conventional magnetic resonance imaging in neonates with clinically mild hypoxic-ischemic encephalopathy," *Chin. Med. J.*, vol. 128, no. 5, pp. 574–580, Mar. 2015, doi: 10.4103/0366-6999.151646.

[6] B. S. Peterson, B. Vohr, L. H. Staib, C. J. Cannistraci, A. Dolberg, K. C. Schneider, K. H. Katz, M. Westerveld, S. Sparrow, A. W. Anderson, C. C. Duncan, R. W. Makuch, J. C. Gore, and L. R. Ment, "Regional brain volume abnormalities and long-term cognitive outcome in preterm infants," *J. Amer. Med. Assoc.*, vol. 284, no. 15, pp. 1939–1947, 2000, doi: 10.1001/jama.284.15.1939.

[7] M. Giménez, M. J. Miranda, A. P. Born, Z. Nagy, E. Rostrup, and T. L. Jernigan, "Accelerated cerebral white matter development in preterm infants: A voxel-based morphometry study with diffusion tensor MR imaging," *NeuroImage*, vol. 41, no. 3, pp. 728–734, Jul. 2008, doi: 10.1016/j.neuroimage.2008.02.029.

[8] P. Mukherjee, J. H. Miller, J. S. Shimony, J. V. Philip, D. Nehra, A. Z. Snyder, T. E. Conturo, J. J. Neil, and R. C. McKinstry, "Diffusion-tensor MR imaging of gray and white matter development during normal human brain maturation," *Amer. J. Neuroradiol.*, vol. 23, no. 9, pp. 1445–1456, 2002, doi: 10.1097/00002093-200210000-00011.

[9] E. Biagioni, M. F. Frisone, S. Laroche, B. A. Kapetanakis, D. Ricci, M. Adeyi-Obe, H. Lewis, N. Kennea, G. Cioni, F. Cowan, M. Rutherford, D. Azzopardi, and E. Mercuri, "Maturation of cerebral electrical activity and development of cortical folding in young very preterm infants," *Clin. Neurophysiol.*, vol. 118, no. 1, pp. 53–59, 2007, doi: 10.1016/j.clinph.2006.09.018.

[10] K. B. Johnson, W. Wei, D. Weeraratne, M. E. Frisse, K. Misulis, K. Rhee, J. Zhao, and J. L. Snowdon, "Precision medicine, AI, and the future of personalized health care," *Clin. Transl. Sci.*, vol. 14, no. 1, pp. 86–93, Jan. 2021, doi: 10.1111/cts.12884.

[11] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," in *Classification in BioApps*. 2018, pp. 323–350, doi: 10.1007/978-3-319-65981-7_12.

[12] F. Jiang, X. Yu, C. Wu, M. Wang, K. Wei, J. Wang, and G. Zhou, "A simple-to-use nomogram for predicting survival in children with acute myeloid leukemia," *Biomed. Res. Int.*, vol. 11, Mar. 2021, Art. no. 7264623, doi: 10.1155/2021/7264623.

[13] F. Jiang, K. Wei, W. Lyu, and C. Wu, "Predicting risk of insulin resistance in a Chinese population with polycystic ovary syndrome: Designing and testing a new predictive nomogram," *Biomed. Res. Int.*, vol. 13, Oct. 2020, Art. no. 8031497, doi: 10.1155/2020/8031497.

[14] Y. Taki, B. Thyreau, S. Kinomura, K. Sato, R. Goto, R. Kawashima, and H. Fukuda, "Correlations among brain gray matter volumes, age, gender, and hemisphere in healthy individuals," *PLoS ONE*, vol. 6, no. 7, Jul. 2011, Art. no. e22734, doi: 10.1371/journal.pone.0022734.

[15] B. Wang and T. D. Pham, "MRI-based age prediction using hidden Markov models," *J. Neurosci. Methods*, vol. 199, no. 1, pp. 140–145, Jul. 2011, doi: 10.1016/j.jneumeth.2011.04.022.

[16] J. Wang, "Human age estimation with surface-based features from MRI images," in *Proc. Int. Workshop Mach. Learn. Med. Imag.*, 2012, pp. 111–118.

[17] J. H. Cole, R. Leech, and D. J. Sharp, "Prediction of brain age suggests accelerated atrophy after traumatic brain injury," *Ann. Neurol.*, vol. 77, no. 4, pp. 571–581, Apr. 2015, doi: 10.1002/ana.24367.

[18] N. K. Dinsdale, E. Bluemke, S. M. Smith, Z. Arya, D. Vidaurre, M. Jenkinson, and A. I. L. Namburete, "Learning patterns of the ageing brain in MRI using deep convolutional networks," *NeuroImage*, vol. 224, Jan. 2021, Art. no. 117401, doi: 10.1016/j.neuroimage.2020.117401.

[19] B. A. Jonsson, G. Bjornsdottir, T. E. Thorgeirsson, L. M. Ellingsen, G. B. Walters, D. F. Gudbjartsson, H. Stefansson, K. Stefansson, and M. O. Ulfarsson, "Brain age prediction using deep learning uncovers associated sequence variants," *Nature Commun.*, vol. 10, no. 1, pp. 1–10, Nov. 2019, doi: 10.1038/s41467-019-13163-9.
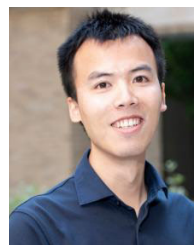
[20] P. Sturmfels, S. Rutherford, M. Angstadt, M. Peterson, C. Sripada, and J. Wiens, "A domain guided CNN architecture for predicting age from structural brain images," in *Proc. Mach. Learn. Healthcare Conf.*, 2018, pp. 295–311.

[21] P. K. Lam, V. Santhalingam, P. Suresh, R. Baboota, A. H. Zhu, S. I. Thomopoulos, N. Jahanshad, and P. M. Thompson, "Accurate brain age prediction using recurrent slice-based networks," in *Proc. 16th Int. Symp. Med. Inf. Process. Anal.*, vol. 11583, Nov. 2020, pp. 11–20.

[22] S. Saha, A. Pagnozzi, P. Bourgeat, J. M. George, D. Bradford, P. B. Colditz, R. N. Boyd, S. E. Rose, J. Fripp, and K. Pannek, "Predicting motor outcome in preterm infants from very early brain diffusion MRI using a deep learning convolutional neural network (CNN) model," *NeuroImage*, vol. 215, Jul. 2020, Art. no. 116807, doi: 10.1016/j.neuroimage.2020.116807.

[23] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, and D. Batra, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 618–626, doi: 10.1109/ICCV.2017.74.

[24] V. M. Bashyam et al., "MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide," *Brain*, vol. 143, no. 7, pp. 2312–2324, Jul. 2020, doi: 10.1093/brain/awaa160.

[25] K. E. Garcia, E. C. Robinson, D. Alexopoulos, D. L. Dierker, M. F. Glasser, T. S. Coalson, C. M. Ortinau, D. Rueckert, L. A. Taber, D. C. Van Essen, C. E. Rogers, C. D. Smyser, and P. V. Bayly, "Dynamic patterns of cortical expansion during folding of the preterm human brain," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 12, pp. 3156–3161, Mar. 2018, doi: 10.1073/pnas.1715451115.

[26] M. Toews, W. M. Wells, and L. Zöllei, "A feature-based developmental model of the infant brain in structural MRI," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervent.*, 2012, pp. 204–211, doi: 10.1007/978-3-642-33418-4_26.

[27] J. Hong, Z. Feng, S.-H. Wang, A. Peet, Y.-D. Zhang, Y. Sun, and M. Yang, "Brain age prediction of children using routine brain MR images via deep learning," *Frontiers Neurol.*, vol. 11, Oct. 2020, Art. no. 584682, doi: 10.3389/fneur.2020.584682.

[28] D. Hu, Z. Wu, W. Lin, G. Li, and D. Shen, "Hierarchical rough-to-fine model for infant age prediction based on cortical features," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 1, pp. 214–225, Jan. 2020, doi: 10.1109/JBHI.2019.2897020.

[29] S. Kornblith, J. Shlens, and Q. V. Le, "Do better ImageNet models transfer better?" in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 2661–2671.

BINGBING XIE received the master's degree. He is currently pursuing the Ph.D. degree with the School of Computing, Macquarie University, and the School of Information Management, Wuhan University. His search interests include data mining, machine learning, and deep learning.
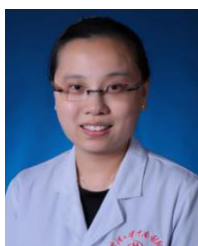

CONG WEI was born in 1987. He received the bachelor's and master's degrees in clinical medicine from the Second Clinical College, Wuhan University, where he is currently pursuing the Ph.D. degree with the Department of Pediatrics. Since 2011, he has been with the Department of Pediatrics, Wuhan University Central South Hospital, mainly engaged in the management of newborn preterm infants and critical care, and was appointed as an Attending Physician, in 2005. Currently, he is the Deputy Director of the Youth Academic Committee of Hubei Pediatric Medical Union; and a member of Pediatric Critical Care Committee, Hubei Pathophysiology Society, and the Pediatric Infectious Branch, Wuhan Medical Association.

JIAJIE TANG received the master's degree from the Wuhan University of Science and Technology. He is currently pursuing the Ph.D. degree with the School of Information Management, Wuhan University. His research interest includes medical big data analysis and application.

LIANTING HU received the bachelor's degree in industrial engineering and the master's degree in mechanical engineering from the Wuhan University of Science and Technology, in June 2015 and June 2018, respectively, and the Ph.D. degree in management science and engineering from Wuhan University, in June 2021. He has been a Postdoctoral Fellow with the Medical Big Data Center, Guangdong Provincial People's Hospital, since July 2021, where he focuses on exploring the application value of artificial intelligence technology and multimodal medical data in medical image lesion segmentation, diagnosis, and prognosis.

PU YANG was born in Anyang, Henan, China, in 1983. She received the M.S. degree in clinical medicine and the Ph.D. degree in pediatrics from Wuhan University, in 2009 and 2015, respectively. In 2011, she completed the Residency Training of pediatrics and neonatology from the Affiliated Hospital of Paris Diderot University, France. She was a Resident with the Department of Pediatrics, Zhongnan Hospital of Wuhan University, from 2009 to 2016, where she has been an Attending Physician of pediatrics, since 2016. Her research interests include critical medicine of premature infants and pediatric infectious diseases.

SHOUYI WANG has engaged in pediatric clinical work for more than ten years. She is good at the diagnosis and treatment of various difficult and critical diseases in children and newborns. She is especially good at the diagnosis and treatment of hematology, tumors, and immune diseases in children, and the diagnosis of neurodevelopmental diseases in prematurity.

YUXUAN JIANG was born in Qinhuangdao, China, in 1996. She received the bachelor's degree in art from Zhejiang Normal University, in 2019, specializing in healthy architectural design for human settlements. In 2020, she studied with the Wuhan University of Technology, focusing on digital humanities and image recognition. From 2021 to 2023, she was an Intern with the Data Center of Guangzhou Women and Children's Medical Center, Guangzhou.

QIRONG WAN received the M.D. degree from the Mental Health Center, Renmin Hospital of Wuhan University. She is with the Mental Health Center, Renmin Hospital of Wuhan University, as a Chief Physician. She specialized in the diagnosis and treatment of schizophrenia, bipolar disorder, depression, anxiety disorders, and other disorders. She has published more than 20 academic articles in domestic and international journals and has participated in three books. She has presided over or participated in many scientific research projects above the provincial level.

RUIZHUO LI completed the joint Doctoral Training with the Guangzhou Women and Children's Hospital, South China University of Technology, and the Standardized Radiological Resident Training with the Xiangya Second Hospital. She received the Certificate of Radiology. Her main research interests include obstetrics and gynecological ultrasonography and congenital structural anomalies.

JIN HAN received the master's degree in medicine from the Prenatal Diagnosis Center and the Guangzhou Women and Children's Medical Center. She is a Chief Physician with the Prenatal Diagnosis Center and the Guangzhou Women and Children's Medical Center. She specialized in prenatal consultation, interventional prenatal diagnosis, down's syndrome screening, various prenatal diagnostic procedures, eugenics, and diagnosis of various genetic diseases, especially in fetal ultrasonography and prenatal screening for down's syndrome.

HANG ZHOU received the degree in prenatal diagnosis and fetal medicine from Guangzhou Medical University, in 2020. He has participated in one national natural project, two provincial projects, and two municipal projects.

DONGCHI ZHAO received the M.D. degree. He is a Chief Physician, a Professor, and the Director of the Department of Pediatrics. He is also the Ph.D. Supervisor and the Deputy Director of the Hubei Pediatric Society; the Deputy Director of the Wuhan Pediatric Society; and a member of the Infection Group, Pediatrics Branch, Chinese Medical Association. In recent years, he has carried out the research on correlation between respiratory syncytial virus and childhood asthma. He received three projects funded by the National Natural Science Foundation of China.

HAIQING XU is the Chief Physician (Professor II) and the Chief Expert of the Hubei Provincial Maternal and Child Health Hospital, a Special Allowance Expert of the State Council, a Young and Middle-Aged Expert with Outstanding Contribution of Hubei Province, and a part-time Professor of the Tongji Medical College, Huazhong University of Science and Technology.

LONG LU received the B.S. degree from Peking University, and the Ph.D. degree from the School of Medicine, Washington University, St. Louis. He is a Postdoctoral Fellow with Yale University; and a Professor and the Doctoral Supervisor with the School of Information Management, Wuhan University. He is a National-Level Talent title holder.

• • •