

# INTERPRETING AGE EFFECTS OF HUMAN FETAL BRAIN FROM SPONTANEOUS fMRI USING DEEP 3D CONVOLUTIONAL NEURAL NETWORKS

Xiangrui Li<sup>1</sup>    Jasmine Hect<sup>2</sup>    Moriah Thomason<sup>3</sup>    Dongxiao Zhu<sup>1</sup>

<sup>1</sup> Department of Computer Science, Wayne State University

<sup>2</sup> Department of Psychology, Wayne State University

<sup>3</sup> Department of Child and Adolescent Psychiatry, New York University

## ABSTRACT

Understanding human fetal neurodevelopment is of great clinical importance as abnormal development is linked to adverse neuropsychiatric outcomes after birth. With the advances in functional Magnetic Resonance Imaging (fMRI), recent studies focus on brain functional connectivity and have provided new insight into development of the human brain before birth. Deep Convolutional Neural Networks (CNN) have achieved remarkable success on learning directly from image data, yet have not been applied on fetal fMRI for understanding fetal neurodevelopment. Here, we bridge this gap by applying a novel application of 3D CNN to fetal blood oxygen-level dependence (BOLD) resting-state fMRI data. We build supervised CNN to isolate variation in fMRI signals that relate to younger *v.s.* older fetal age groups. Sensitivity analysis is then performed to identify brain regions in which changes in BOLD signal are strongly associated with fetal brain age. Based on the analysis, we discovered that regions that most strongly differentiate groups are largely bilateral, share similar distribution in older and younger age groups, and are areas of heightened metabolic activity in early human development.

**Index Terms**— Fetal neurodevelopment, fMRI, CNN, Model Interpretation.

## 1. INTRODUCTION

The human brain undergoes rapid development in the prenatal period. Previous studies show significant links between abnormal brain development *in utero* and adverse postnatal outcomes [1, 2]. As a result, understanding fetal neurodevelopment, particularly timescales and patterns for maturation of functional systems, is of great clinical importance. Recent methodological advances in fetal fMRI have attracted much attention [2, 3]. For fetal neurodevelopmental studies, existing works focus on analysis of the functional connectivity in fetal brain, such as discoveries of functional networks and identification of highly connected hubs [4]. Those studies are based on functional connectivity measures (i.e. adjacency matrix of Fisher z-scores) extracted from the spatio-temporal 4D resting-state fMRI. However, from the perspective of

computational methodology, those approaches rely heavily on the data processing steps for extracting z-scores and only utilize temporal information in the fMRI images, which miss the spatial information in the 2D or 3D fMRI images.

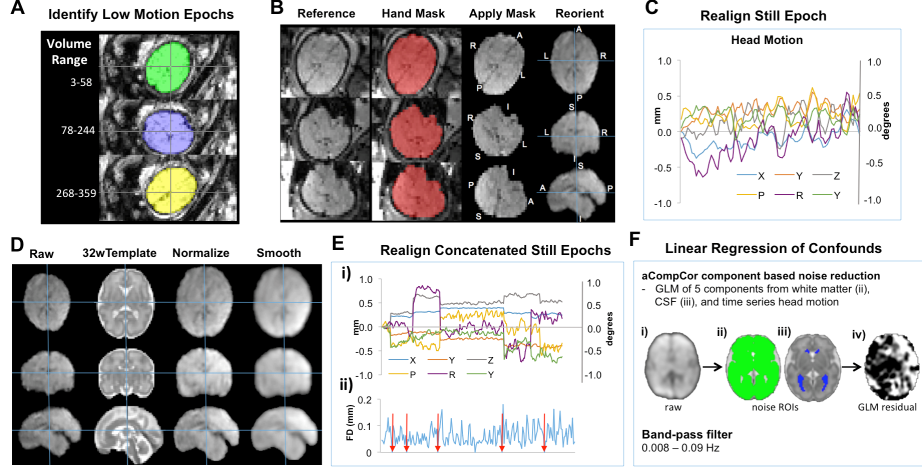
There is precedent for direct learning from MRI images in children and adults. For example, CNN has been used to classify neurological disorders such as Autism Spectrum Disorder (ASD) [5] and Alzheimer’s disease [6], and to predict age from fMRI data [7]. Probably due to very limited availability of fetal fMRI data, CNN models have not been applied for understanding fetal neurodevelopment.

To bridge the methodological gap, we propose a novel application of deep CNN in conjunction with CNN interpretation techniques on the analysis of fetal brain fMRI data. To the best of our knowledge, this work represents the first application of CNN models as a data-driven approach in understanding fetal developmental processes. Specifically, as age and brain development are closely linked, we utilize the age group as the surrogate target (younger *v.s.* older) and train a supervised CNN directly on the residual fMRI images to capture associations between age effect and blood oxygen-level dependence (BOLD). The trained model achieves superior predictive performance over traditional machine learning algorithms. We interpret the trained CNN to highlight brain regions that are strongly influential in model predictions, which are clinically compelling that match with existing studies. The trained CNN model also proposes novel clinically sensible regions that are potentially informative in characterizing brain development in fetuses, which can potentially enhance our understanding of fetal brain development.

## 2. METHOD

### 2.1. Data Acquisition and Preprocessing

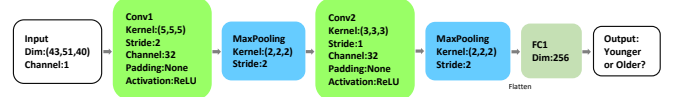
Subjects in our analysis were subset of a cohort in a study of longitudinal fetal brain development. All women provided written informed consent before undergoing MRI examination. Participation was approved by Human Investigation Committee of Wayne State University. The cohort consists of 148 pregnant women that underwent MRI with gestational



**Fig. 1.** Data preprocessing pipeline from raw MRI image to residual fMRI data.

age (GA) between 24 and 40 weeks. We manually checked subjects according to data quality (high motion, artifacts et al.) and inclusionary criteria (gestational age, born without nervous system abnormality et al.), leaving 75 qualified subjects: 30 subjects with GA between 26 and 29 weeks (younger group), and 45 between 34 and 37 weeks (older group).

MRI data were acquired on a Siemens Verio 70-cm open-bore 3T MR system. Resting-state fMRI data were acquired using a gradient echo planar imaging sequence: TR/TE 2000/30 ms, flip angle  $80^\circ$ , 360 frames, axial 4 mm slice thickness, voxel size  $3.4 \times 3.4 \times 4 \text{ mm}^3$ , repeated twice. Between 12 to 24 minutes of fetal resting-state fMRI data were collected per subject. Extraction of fetal brain fMRI data requires multiple steps in preprocessing raw resting state fMRI images. In brief, periods of fetal quiescence were manually identified using FSL image viewer, wherein individual segments must consist of at least 20 seconds (10 frames) of low motion ( $<2 \text{ mm}$  translation and/or 3 degrees rotational movement) (A). After motion censoring, fetal brain masks were then created separately for each low-motion epoch from a single reference image using Brainsuite [8] (B). After masking, each temporal segment was reoriented manually, realigned to the mean BOLD volume, resampled to 2 mm isotropic voxels, and normalized to a 32-week fetal brain template [9] using SPM8 [10] (C). All normalized images from each segment were then concatenated into one run, realigned to the mean BOLD volume, and smoothed with a 4 mm FWHM Gaussian kernel (D-E). Further preprocessing was performed in CONN toolbox (v14n) [11] including linear detrending, nuisance regression using aCompCor of five principal components extracted from a 32-week fetal atlas white matter and CSF mask, six head motion parameters, and band-pass filtering at 0.008 to 0.09 Hz (F). Fig. 1 shows the preprocessing pipeline from raw MRI to residual fMRI used in our model.



**Fig. 2.** CNN architecture for age group prediction.

## 2.2. 3D CNN for Age Effect on Neurodevelopment

For each subject, the data is a time series of 3D fMRI volumes of dimension  $43 \times 51 \times 40$ . To exploit the spatio-temporal information in fMRI, we take a sliding window approach [5]. Specifically, starting from the first 3D frame of 4D fMRI whose length is  $T$ , we take the mean 3D fMRI image in a window of size  $m$  over the time axis; as the window slides over the time axis with stride  $s$ , a total of  $\lfloor \frac{T-m}{s} \rfloor + 1$  3D fMRI images are generated, which significantly increases the sample size in training data. After initial experiments,  $m = 2$  (or 3) and  $s = 1$  achieve good performance and increasing window size leads to over-smoothing and degrade model performance. As mentioned in Section 1, younger ( $26 < \text{GA} < 29$  weeks) and older ( $34 < \text{GA} < 37$  weeks) age groups are used as the surrogate target. For each 3D fMRI generated from sliding window, we label it with the subject's age group.

To learn fetal neurodevelopmental processes by classifying 3D fMRI, we build an effective 3D CNN to capture the association between residual fMRI and age group. The proposed CNN along with its architectural parameters (i.e. kernel size, stride, channel et al.) is shown in Fig. 2. It has two convolutional layers with ReLU non-linear activation and one fully connected layer, where each convolutional layer is followed by a max pooling layer. Model predictions are made with sigmoid activation on the output layer after the fully connected layer. The objective function is hence the binary cross

entropy plus  $L_2$  regularization on the network weights:

$$L(\theta) = - \sum_i y_i \log(f(\mathbf{x}_i)) + (1 - y_i) \log(1 - f(\mathbf{x}_i)) + \lambda \|\theta\|_2^2,$$

where  $y_i = 1$  (older group) or 0 (younger group) is the true label of  $\mathbf{x}_i$ ,  $f(\mathbf{x}_i)$  is the probability  $p(y = 1|\mathbf{x}_i)$  modeled by CNN, and  $\lambda$  is the tuning parameter of  $L_2$  regularization.

### 2.3. Model Interpretation

As our primary goal is to understand the age effect on fetal brain development, model interpretation is critical to our CNN approach. Once CNN is trained, we perform sensitivity analysis (SA) [12, 13] on leave-out testing images to identify important regions of interests (ROIs). The sensitivity score for each image pixel is the squared gradient with respect to the CNN output (i.e. probability of being younger or older), which measures how sensitive the prediction is to the change of voxel values. The calculation of sensitivity score only needs one pass of backpropagation with respect to the input image. For fetal fMRI images, a region (i.e. multiple voxels in a neighborhood) in the brain with larger sensitivity scores is more influential, hence more important in predicting age group.

### 2.4. Training Setup

The CNN is trained with classic stochastic gradient descent algorithm (SGD) with momentum set to 0.8. The learning rate is set to 0.1 initially and decreased by multiplying a factor 0.2 for every 7 epochs.  $L_2$  regularization with  $\lambda = 0.001$  is used to prevent overfitting. The size of mini-batch in each epoch is set to 128. We apply early stopping as another regularization in model validation. In our practice, training usually is stopped at 15 epochs. Model parameters are initialized with uniform distribution on  $(-\sqrt{u}, \sqrt{u})$  where  $u$  is reciprocal of the number of weights in each layer. We implement our CNN in Pytorch [14]<sup>1</sup>.

To avoid model seeing images of the same subject in testing and training (since each subject generates multiple 3D fMRI images), we split the dataset *at subject level*. In the experiments, data are randomly divided with stratification into training, validation and testing sets by 80%/10%/10%. All images are normalized by subtracting mean image and then divided by the maximal absolute intensity value, both calculated from training data. The splitting procedure results in about 9300 images in training, 1200 in validation and testing. We repeat this procedure 10 times in experiments. To evaluate model performance, we use F1 score which is calculated as  $2(\text{precision} \times \text{recall}) / (\text{precision} + \text{recall})$  and Area under ROC curve (AUC) as the evaluation metric, and the average F1 and AUC over 10 runs are reported.

**Table 1.** Predictive performance (standard deviation) on testing subjects. Note that Fisher z-score as functional connectivity measure is calculated based on brain parcellation and not applicable for CNN.

	Fisher z-score		fMRI Image	
	F1	AUC	F1	AUC
CNN	-	-	<b>0.84</b> <b>(0.05)</b>	<b>0.91</b> <b>(0.06)</b>
RF	0.76 (0.03)	0.69 (0.18)	0.79 (0.06)	0.77 (0.15)
GBM	0.73 (0.09)	0.69 (0.17)	0.81 (0.06)	0.77 (0.18)
LR	0.77 (0.06)	0.75 (0.18)	0.77 (0.01)	0.67 (0.14)

## 3. EXPERIMENTAL RESULTS

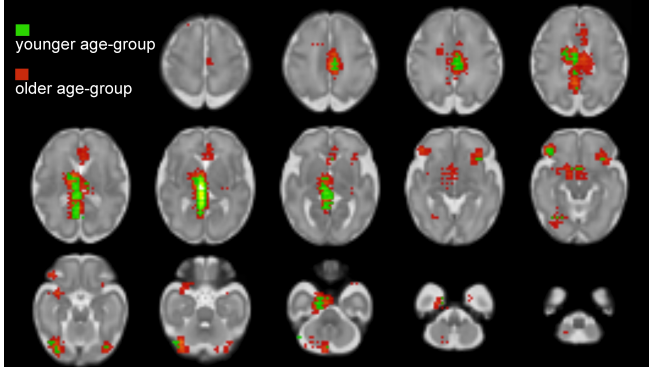
### 3.1. Predictive Performance

For performance comparison, we also test several baseline classification algorithms, including random forest (RF), gradient boosting machine (GBM) and logistic regression (LR) with  $L_2$  regularization. These alternatives were tested on two types of data: (1) fetal brain functional connectivity matrices extracted by correlating resting-state fMRI time series data across 100 brain regions (Fisher z-score), (2) fetal BOLD fMRI images. For fMRI, due to the high dimensionality ( $\text{dim} = 39984$ )<sup>2</sup> and strong correlations among neighboring voxels in fMRI, we use principal component analysis (PCA) for dimension reduction before testing on baselines. We select the optimal number of PCA components as well as optimal parameters of classification models using the validation data. In our experiments, 100 components achieves good performance and including more components does not necessarily lead to performance gain.

The predictive performance (F1 score and AUC) on testing data at subject level is shown in Table 1. Note that subject level prediction is made with soft voting: probability for each subject is averaged across its corresponding 3D images. The probability threshold for classification is chosen as 0.5 for calculating F1 score. We see from Table 1 that CNN trained directly on fMRI images achieves the best and robust performance (F1 0.84 with standard deviation 0.05, AUC 0.91 with 0.06), due to CNN’s merit of capturing spatial information of fMRI images in classification. The improvements demonstrate that CNN can effectively capture discriminative information that could reveal associations between age effect and brain development.

<sup>1</sup><https://github.com/Dichoto/FetalAge>

<sup>2</sup>We remove zero columns after flattening original 3D fMRI into a vector of size  $43 \times 51 \times 40 = 87720$



**Fig. 3.** Sensitivity analysis for younger v.s. older age groups. Important regions of high sensitivity scores that are associated with age effects are noted in green for younger fetuses and red for older fetuses.

### 3.2. Inference About Fetal Neurodevelopment from CNN Classification

It was possible to discriminate older versus younger fetuses on the basis of spontaneous baseline resting-state BOLD measurements. BOLD reflects changing regional blood concentrations of oxy- and deoxy-hemoglobin, which are influenced at least in part by regional metabolic activity [15]. Having observed that fetal age can be differentiated on the basis of these fluctuating signals suggests that patterns in baseline BOLD during resting-state reflect aspects of brain maturation.

Sensitivity analysis was used to identify regions within each age group where change in voxel values most altered strength of the prediction. Important regions (identified with large sensitivity scores, see Section 2.3) for both younger and older fetus groups are shown in Fig. 3. We found that sensitive regions (1) have a high degree of spatial overlap across both age groups, (2) are largely bilaterally distributed, (3) encompass brain regions that have been identified as having high baseline metabolic activity in positron emission tomography (PET) studies in human newborns and infants, including subcortex, thalamus, and medial temporal lobe [16].

Specifically, across groups we observe high sensitivity in bilateral occipital cortex, bilateral ventrolateral prefrontal cortex (vlPFC), sub-cortex, posterior cingulate cortex (PCC), medial temporal lobe (MTL), thalamus, and cerebellum. In addition, baseline BOLD in anterior cingulate cortex (ACC), hypothalamus, and insula regions are important for group classification in older fetuses.

## 4. CONCLUSIONS

In this paper, we propose a novel application of CNN for interpretation of fetal brain age effects directly from fMRI images. The predictive performance of CNN demonstrate that it can well capture associations between age and varia-

tion in BOLD signal. To better understand the relevance of our predictive CNN to fetal development, we use sensitivity analysis to isolate regions critical for CNN performance, and discovered that our most sensitive regions were regions that are high in metabolic activity in early human brain development. The experimental results reveal compelling associations and demonstrate potential promise of CNN applied to spontaneous BOLD activity as a data-driven approach to understand fetal developmental processes.

With those identified important ROIs, in our future studies, we plan to perform seed-based analyses to examine how functional connectivities and networks of those ROIs are associated with whole-brain age effects.

## 5. REFERENCES

- [1] Oualid M Benkarim, et al., “Toward the automatic quantification of in utero brain development in 3d structural mri: a review,” *Human brain mapping*, vol. 38, no. 5, pp. 2772–2787, 2017.
- [2] Marion I van den Heuvel and Moriah E Thomason, “Functional connectivity of the human brain in utero,” *Trends in cognitive sciences*, vol. 20, no. 12, pp. 931–939, 2016.
- [3] Lindsey Konkell, “The brain before birth: Using fmri to explore the secrets of fetal neurodevelopment,” 2018.
- [4] Marion I van den Heuvel, et al., “Hubs in the human fetal brain network,” *Developmental cognitive neuroscience*, vol. 30, pp. 108–115, 2018.
- [5] Xiaoxiao Li, et al., “Brain biomarker interpretation in asd using deep learning and fmri,” in *MICCAI*. Springer, 2018, pp. 206–214.
- [6] Sergey Korolev, et al., “Residual and plain convolutional neural networks for 3d brain mri classification,” in *ISBI*, 2017, pp. 835–838.
- [7] Hongming Li, Theodore D Satterthwaite, and Yong Fan, “Brain age prediction based on resting-state functional connectivity patterns using convolutional neural networks,” in *ISBI*, 2018, pp. 101–104.
- [8] David W Shattuck and Richard M Leahy, “Brainsuite: an automated cortical surface identification tool,” *Medical Image Analysis*, vol. 6, no. 2, pp. 129–142, 2002.
- [9] Ahmed Serag, et al., “Construction of a consistent high-definition spatio-temporal atlas of the developing brain using adaptive kernel regression,” *NeuroImage*, vol. 59, no. 3, pp. 2255–2265, 2012.
- [10] William D Penny, Karl J Friston, John T Ashburner, Stefan J Kiebel, and Thomas E Nichols, *Statistical parametric mapping: the analysis of functional brain images*, Elsevier, 2011.
- [11] Susan Whitfield and Alfonso Nieto, “Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks,” *Brain connectivity*, vol. 2, no. 3, pp. 125–141, 2012.
- [12] Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman, “Deep inside convolutional networks: Visualising image classification models and saliency maps,” *arXiv preprint arXiv:1312.6034*, 2013.
- [13] Johannes Rieke, et al., “Visualizing convolutional networks for mri-based diagnosis of alzheimer’s disease,” in *Understanding and Interpreting Machine Learning in Medical Image Computing Applications*, pp. 24–31. Springer, 2018.
- [14] Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer, “Automatic differentiation in pytorch,” 2017.
- [15] Masaki Fukunaga, et al., “Metabolic origin of bold signal fluctuations in the absence of stimuli,” *Journal of Cerebral Blood Flow & Metabolism*, vol. 28, no. 7, pp. 1377–1387, 2008.
- [16] Harry T Chugani, “Imaging brain metabolism in the newborn,” *Journal of child neurology*, vol. 33, no. 13, pp. 851–860, 2018.