

Supplementary material of the manuscript:

“Hierarchical Rough-to-Fine Model for Infant Age Prediction based on Cortical Features”

1. The scatter boxplot of the dataset

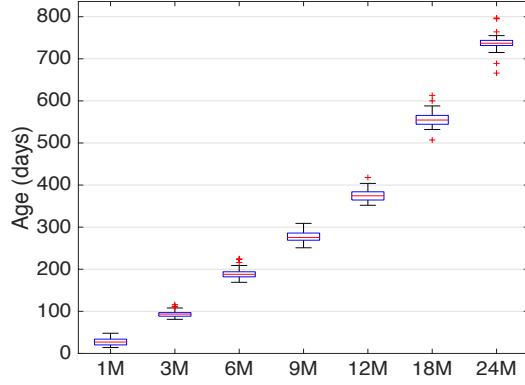


Fig.S1. The scatter boxplot of the dataset. M is short for month.

2. Age range partition learning in HRtoF age prediction model

An optimal partition-nodes set $\{y_1, \dots, y_{k-1}, y_k, y_{k+1}, \dots, y_{K-1}\}$ of the age range $(0, 800)$ will be learned. It leads to the partition as $\{(0, y_1), [y_1, y_2), \dots, [y_{k-1}, y_k), [y_k, y_{k+1}), \dots, [y_{K-1}, 800)\}$ which corresponds to K subgroups of the age range.

Step 1: Find the optimal number K of the partition based on equidistant partition. K is learned by minimizing the MAE of the age prediction model while the partition interval varies from 1 to 300. When two intervals have the same MAE, the bigger interval will be chosen to guarantee the accuracy of the rough prediction and the interpretability of the results. Suppose $\{y'_1, \dots, y'_{k-1}, y'_k, y'_{k+1}, \dots, y'_{K-1}\}$ is the result obtained from this step.

Step 2: Find the optimal endpoints for each age group based on particle swarm optimization algorithm [49]. For the k^{th} age group $[y'_{k-1}, y'_k)$, $k \neq 1, K$, the left endpoint y_{k-1} and the right endpoint y_k will be learned in the restricted scope of $[y'_{k-2}, y'_k]$ and $[y'_{k-1}, y'_{k+1}]$, respectively. When $k = 1$ or K , only the right endpoints or the left endpoint will be learned in the restricted scope of $[0, y'_2]$ or $[y'_{K-1}, 800]$. This restriction will significantly speed up the optimization process. The final partition nodes set $\{y_1, \dots, y_{k-1}, y_k, y_{k+1}, \dots, y_{K-1}\}$ will be obtained when the algorithm is converged.

3. The relative importance of features obtained by the 6 models in fine prediction stage

The relative importance of each feature type in each model of fine prediction stage was computed by summing the values of its related ROI features and normalized by dividing the total relative importance of the 8 feature types in the model obtained by SVR.

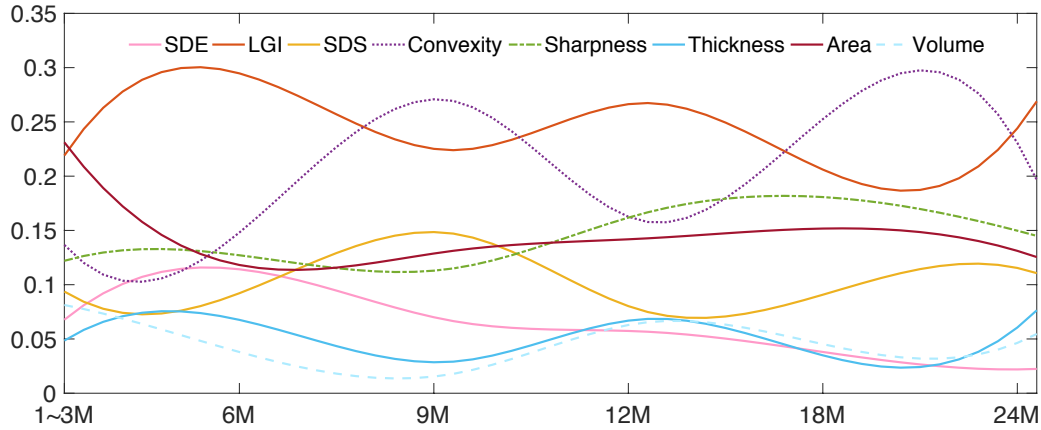


Figure S2. The relative importance of the 8 types of features varies from 1~3 months to 24 months obtained by the fine prediction stage models.

4. The prediction effectiveness of the 6 sub-models in fine prediction stage

To evaluate the prediction effectiveness of these 6 models trained by SVR, they were used for age prediction in their own age group and also neighboring age groups. The correlations between the predicted ages and the chronological ages were shown in Table S1.

TABLE S1. The prediction effectiveness of the 6 fine prediction sub-models. r is the correlation coefficient between the chronological ages and predicted ages obtained by a single fine prediction model. p -value is the corresponding significance value of the correlation. The neighboring age group of 1~3M, 6M, 9M, 12M, 18M and 24 are (1~3M, 6M), (1~3M, 6M, 9M), (6M, 9M, 12M), (9M, 12M, 18M), (12M, 18M, 24M), and (18M, 24M), respectively.

Sub-models	Prediction in its own age group		Prediction in its neighboring age group	
	r	p -value	r	p -value
Model 1~3M	0.9192	0.0000	0.9190	0.0000
Model 6M	0.3735	0.0162	0.3170	0.0435
Model 9M	0.1414	0.4108	0.0968	0.5743
Model 12M	-0.1495	0.3841	-0.1340	0.4360
Model 18M	0.3010	0.0426	0.1287	0.4351
Model 24M	0.4586	0.0278	0.4585	0.0278

5. The relationship between belief threshold θ and the corresponding MAE obtained by HRtoF

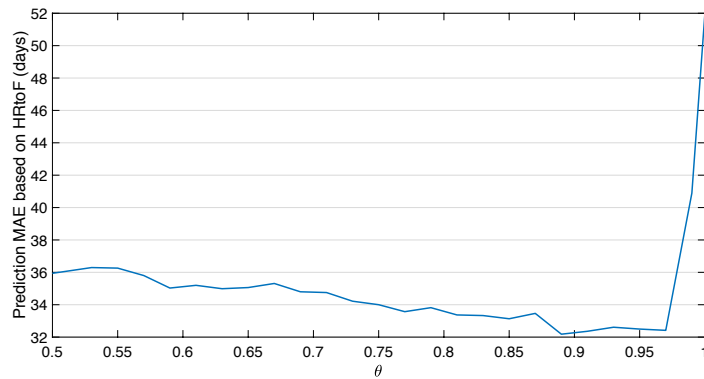


Fig. S3. The prediction MAEs of our proposed HRtoF vary along the belief threshold θ . The minimum value of MAE was obtained at $\theta = 0.89$.

6. The developmental trajectories of the 8 types of global features

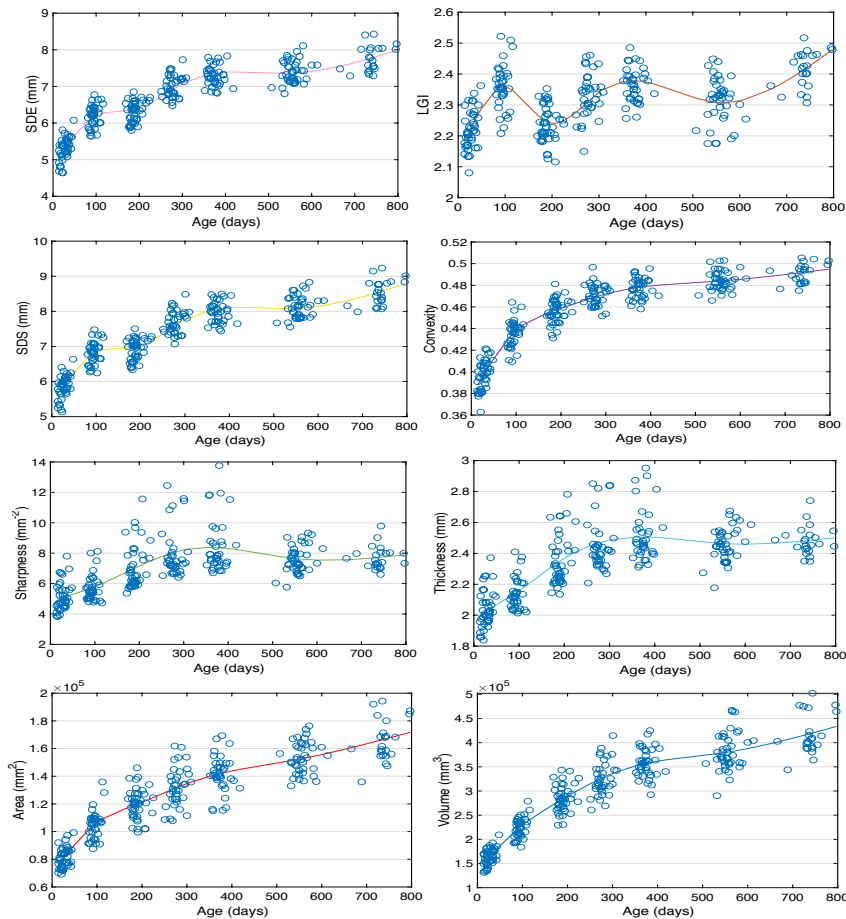


Fig. S4. The developmental trajectories of the 8 types of global features from birth to 2 years of age.