GeneMorphFormer: Transformer-Driven Cross-Scale Mapping from Gene Expression to Cortical Morphology

Xiao Li[1], Han Zhang[2], Qitai Sun[2], Chenjie Jia[1], Xiaowei He[1,2(✉)], and Yudan Ren[1,3]

[1] School of Information Science & Technology, Northwest University, Xi'an, China
[2] School of Network and Data Center, Northwest University, Xi'an, China
`hexw@nwu.edu.cn`
[3] Yunnan Key Laboratory of Plateau Food Advanced Manufacturing, Kunming, China

Abstract. We present GeneMorphFormer, a Transformer-based model to decode nonlinear interactions between gene expression and cortical morphology. We align expression maps with gray matter and white matter boundary curves through spatial normalization by leveraging marmoset in situ hybridization (ISH) data. Our model employs multi-head self-attention to model global dependencies across 1024 gene features, optimized by a hybrid loss (MSE and Hausdorff distance) balancing local precision and global shape fidelity. SHapley Additive exPlanations (SHAP) analysis is subsequently employed to quantify the contribution of genes to morphological shape. Wavelet-based clustering further reveals distinct gene sets governing smooth versus fluctuating morphologies, suggesting hierarchical genetic regulation. Experimental results demonstrate that GeneMorphFormer outperforms traditional networks in both global shape matching and local precision. This work proposed a biologically interpretable Transformer architecture for cross-scale gene-morphology mapping and enables systematic exploration of genetic drivers in cortical morphology malformations. Our code is publicly available at https://github.com/Leveup/GeneMorphFormer.

Keywords: Cortical morphology · Gene expression · Transformer.

1 Introduction

Primate cerebral cortex undergoes a large fold expansion in surface area from gestation to adolescence, driven by spatially heterogeneous gene expression [9]. Many genes were observed to exhibit expression variation across different cortical regions, inducing mechanical forces that orchestrate gyral folding patterns [11]. This gene-morphology interplay establishes functional hierarchies that high-gyrification regions are closely related to brain activities and interactions showing distinct thickness [8].

While several hypotheses have been proposed to describe folding mechanics, transcriptomic regulation is considered as the root driver. Gene Trnp1 influences neuronal differentiation and cortical growth, promoting the regionalized folding pattern [10], and TMEM14B explains primate-specific sulcal length variance through radial glia expansion [5]. However, cortical expansion is a complex process driven by synergistic interaction of thousands of genes, which is undiscovered in these studies.

Many researchers use linear models and deep learning methods to capture the potential relationship between gene expression and phenotypic morphology. For example, Zhou applied partial least squares regression and enrichment analysis to examine the relations between spatial pattern of normal cortical thinning and genes predominantly expressed in various neurons in children and adolescents [15]. Similarly, Venugopalan utilized stacked denoising auto-encoders to extract features from clinical and genetic data, and 3D-convolutional neural networks (CNNs) to predict Alzheimer's disease by integrating intermediate-level features [13]. These methods often rely on feature concatenation rather than modeling the complex distributional relationships within the data, which limits their potential to uncover deeper feature interactions.

Recently, Transformer architecture has demonstrated remarkable success in capturing long-range dependencies between data elements, excelling in sequence generation and image processing. It has also shown unique advantages in the field of neuroscience. For instance, Sun et al. applied Transformers to fMRI data, successfully capturing temporal dependencies and accurately predicting brain activity over short time intervals (predicting 5.04 seconds of activity from 21.6 seconds of data). This demonstrates the effectiveness of Transformers in modeling the spatiotemporal dynamics of brain function [12].

Moreover, in situ hybridization (ISH) enables cellular resolution gene mapping, bridging gene expressions to particular brain regions, even cortical morphology. Li et al. utilized sparse coding on ISH data to identify spatial expression patterns of genes across different brain regions, revealing the transcriptomic architecture of mouse brain [4]. Li et al. also used ISH data from marmoset brains to look for genes corresponding to convex and concave regions in cortical surface by a support vector machine model [3].

Therefore, we propose GeneMorphFormer, a novel architecture specifically designed for coordinate prediction. By standardizing the spatial positions and scales of cortical curves, our model extracts and compares features within a consistent reference framework. It predicts gray matter and white matter boundary morphology from gene expression data, capturing the nonlinear relationship between genetic features and cortical shape. Furthermore, by quantifying genes' contribution by SHapley Additive exPlanations (SHAP) analysis, our approach identifies key genetic factors driving cortical development and folding, offering new insights into the genetic mechanisms underlying cerebral cortex formation.
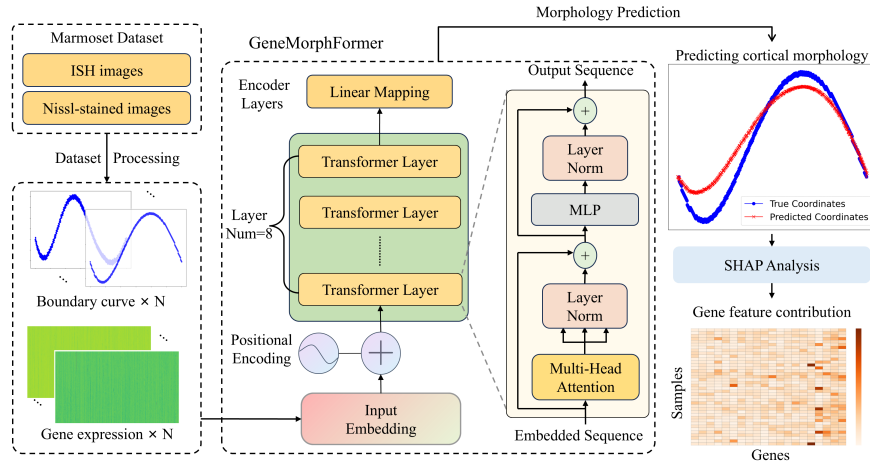
Fig. 1. Overview of the cortical morphological gene analysis workflow.

2   Method

The comprehensive methodological pipeline is illustrated in Fig. 1. The workflow begins with extracting cortical boundary curves and transcriptomic data. Subsequently, a transformer-based predictive modeling framework, we called GeneMorphFormer, is implemented to infer cortical morphological features from gene expression signatures. Finally, SHAP analysis is employed to quantitatively assess feature importance and elucidate key molecular determinants underlying cortical development.

2.1   Generating boundary curves and corresponding gene expression matrices

We selected 1024 ISH images from the Brain/MINDS marmoset gene atlas dataset [7], which also includes Nissl-stained and MRI anatomical images. Both ISH and Nissl-stained images are composed of 60 sections with a resolution of 96 dpi and dimensions of 7200×8400 pixels. The grayscale values in the ISH images correspond to the gene expression levels at each point

As illustrated in Fig. 2, our computational pipeline begins with a CBAM-enhanced TransUnet [14, 1] trained on MRI data to segment the cortical gray matter on Nissl-stained templates. The extraction process involves identifying the contours of the segmented gray matter, followed by selecting the inner boundary curves. These curves were discretized into 600-point segments, resulting in 379 segments. A subsequent 2D affine transformation from ISH images to Nissl-stained templates enabled gene expression mapping onto the boundary curves, generating a 600×1024 gene expression matrix per segment.
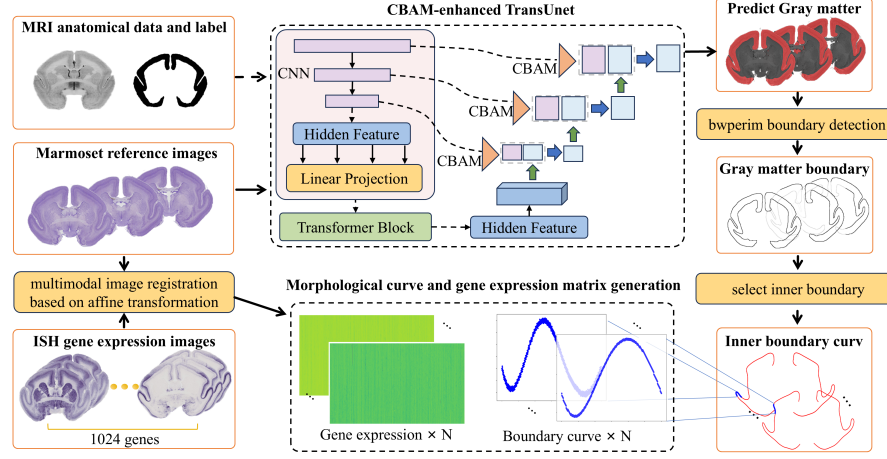
Fig. 2. Pipline for generating morphological curves and gene expression matrices

Prior to model training, cortical segments were spatially normalized using sequential rigid transformations to eliminate geometric confounders. The normalization process included: (1) translate initial points to the origin (0,0); (2) rotate curves to ensure the terminal points were collinear along the x-axis; and (3) isotropic scaling to fix the terminal coordinates at (600,0). This standardization established a unified coordinate system, removing variability in position, orientation, and scale. The resulting framework enables deep learning models to focus on intrinsic gene-morphology relationships without spatial transformation artifacts.

To enhance model generalization and mitigate overfitting, we applied horizontal flipping to each cortical curve as a data augmentation strategy, resulting in 758 samples. These were divided into 683 for training and 75 for testing.

2.2   GeneMorphFormer Model

Then GeneMorphFormer is employed to predict the 2D coordinates of boundary curve segments from corresponding gene expression data. The model takes a gene expression matrix $X \in \mathbb{R}^{600 \times 1024}$ as input, where each row represents a point along the cortical curve, and each column corresponds to gene expressions. The input features are first mapped to a higher-dimensional embedding space via a linear transformation:

$$X_{\text{emb}} = X W_{\text{in}} + b_{\text{in}}, \quad W_{\text{in}} \in \mathbb{R}^{1024 \times d}, \quad b_{\text{in}} \in \mathbb{R}^{d} \tag{1}$$

where $X_{\text{emb}} \in \mathbb{R}^{600 \times d}$ is the embedded gene expression matrix, and d is the embedding dimension. To retain the positional information of each point, we add position embeddings P_{emb} based on position indices:

$$P_{\text{emb}} = \text{PositionEmbedding}(p), \quad p \in \{0, 1, 2, \ldots, 599\} \tag{2}$$

Then we added these positional embeddings to learn morphological information correctly:

$$X_{\text{input}} = X_{\text{emb}} + P_{\text{emb}} \tag{3}$$

The encoder consists of 8 layers of Transformer Encoder blocks, each incorporating a Multi-Head Self-Attention (MSA) mechanism and a Feed-Forward Network (FFN). The self-attention mechanism captures global dependencies among gene expression values across different points. The output of the attention mechanism is processed through the feed-forward network.

The final encoded representation is passed through a fully connected layer to predict the 2D spatial coordinates of each point:

$$Y_{\text{pred}} = X_{\text{enc}} W_{\text{out}} + b_{\text{out}}, \quad W_{\text{out}} \in \mathbb{R}^{d \times 2}, \quad b_{\text{out}} \in \mathbb{R}^2 \tag{4}$$

To maintain the consistency of the morphology, the starting and ending points of the curve are fixed.

2.3   Loss Function

To accurately predict points coordination and line shape, we use a weighted sum of Mean Squared Error (MSE) and Hausdorff distance as loss function:

$$Loss = \alpha \times MSE + \beta \times Hausdorff \tag{5}$$

where α and β are the weighting factors for MSE and Hausdorff distance, respectively. The MSE metric computes the average squared differences between corresponding points on predicted and ground truth curves, providing a fine-grained measure of local alignment accuracy at each sampled point. In contrast, the Hausdorff distance, defined as the maximum of the directed distances between two curves, evaluates global shape correspondence by identifying the worst-case deviation between the predicted and actual curves. After experimentation, we set $\alpha = 0.8$ and $\beta = 0.6$ as the optimal parameters. The Adam optimizer [2] with an initial learning rate of 10^{-4} was used for 2000 epochs, with StepLR halving the learning rate every 250 epochs to improve convergence and stability.

2.4   SHAP Analysis

In this study, we use SHAP (SHapley Additive exPlanations [6]) to quantify the contribution of gene expression features to predictions. We apply the Gradient-Explainer method, which estimates feature contributions by computing gradients of the model's output with respect to each input feature. This method integrates the gradients into a reference distribution to assess the impact of each gene on cortical morphology prediction. Specifically, for an input sample x, the SHAP value is approximated as:

$$ExpectedGradients_i(x) = E_{\substack{x' \sim D \\ \alpha \sim U(0,1)}} \left[(x_i - x_i') \times \frac{\partial f(x' + \alpha(x - x'))}{\partial x_i} \right] \tag{6}$$

where x' is a reference sample from the data distribution D, α is a scaling factor sampled from $[0, 1]$, and $\frac{\partial f}{\partial x_i}$ is the model's gradient with respect to feature i. The formula captures how the input gene expression signatures affect the morphological prediction of the output, enabling accurate attribution of gene contributions.

2.5   Cortical Curve Classification

To further investigate the relationship between cortical shape and gene expression, we apply discrete wavelet transform (DWT) with Daubechies 4 (db4) wavelet for fourth-order decomposition, extracting multi-scale shape features. The resulting approximation and detail coefficients form feature vectors representing each curve's morphology. Then we employ K-Means and determine the optimal number of clusters K=2 using the silhouette coefficient. Post-hoc inspection of the clustering results revealed that the two groups corresponded well to curves with smoother trajectories and those exhibiting more pronounced local variations. After that, we link SHAP values for all genes with specific curve clusters.

3   Result

3.1   Cortical morphology prediction performance

We conducted a detailed comparison between the predicted results and the true coordinates. We used the Mean Squared Error (MSE) and Mean Absolute Error (MAE) to quantify the overall deviation of the predicted values at each coordinate point. At the same time, we utilized the Hausdorff distance to measure the deviation of the overall shape. We compared the prediction performance of mainstream deep learning neural networks on all test samples. The specific results are shown in Table 1. The experimental results clearly show that compared with these models, our method has more advantages in accuracy and global shape matching.

Table 1. Model Evaluation Indicators

Model	Evaluation Indicators		
	MSE	**MAE**	**Hausdorff distance**
CNN	0.200	0.153	0.788
ResNet	0.197	0.178	0.674
GCN	0.255	0.218	1.157
Our Model	0.182	0.207	0.554

In addition, we selected 5 samples from the original curves for visual presentation (see Fig. 3) to clearly demonstrate the performance of each model in
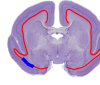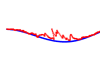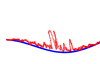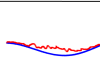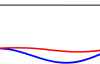
Fig. 3. Predicted curves and ground truth curves of 5 selected segments.

the cortical morphology prediction task. Here, red curves represent the cortical morphology predicted by each model, while blue curves represent the actual morphology of the cortical segments. It can be observed that in multiple samples, prediction curves generated by our model have a high degree of fit with the actual curves, indicating that the predicted morphology is relatively close to the real one. This visual comparison further illustrates the superiority of our model in cortical morphology prediction.

3.2 SHAP Analysis Result

In this study, with the help of SHAP analysis, we screened out the top 20 genes that contribute the most to the prediction of cortical morphology. To investigate gene contribution levels among different curve samples, we finally determined to divide all segments into two categories, presented in the upper-right corner of Fig. 4. by calculating the silhouette coefficients under different numbers of clusters. It can be clearly seen that these two types of curves have significant morphological differences. One type shows a smooth shape, while the other exhibits a more complex fluctuating shape.

For each type of curve, we further investigated the contribution of different genes to the shape of the curve. Again, using SHAP analysis, we both selected top 20 genes with the highest contribution levels on two types. As a result, we obtained 12 common genes: CACNA2D3, FBXO32, IGFBP4, MPPED1, NFE2L3, NFIB, NFIX, PJA2, PPP1R1B, PTPRF, SV2A, and TRIM16. These results are shown in the bottom right corner of the graph, with crossed genes marked with asterisks.

Fig. 4. SHAP analysis results. (a) The bar chart illustrates the contribution of genes to cortical morphology prediction, and the accompanying scatter plot shows gene expression levels across samples. (b) Clustering results based on cortical curve characteristics. (c) SHAP-based gene contribution analysis for each cluster. Genes marked with an asterisk (*) exhibit high importance in both clusters.

Our research identified multiple genes that play a crucial role in curves of different morphologies. These genes provide important clues for exploring the universal regulatory mechanisms of curve morphologies. At the same time, by separately screening genes with high contribution in a single type of curve, we have successfully revealed the unique regulatory genes for specific curve morphologies. This series of achievements is of great significance for in-depth exploration of the complex relationship between genes and curve morphologies.

4   Discussion

Our Study presents a novel Transformer-based framework for decoding the relationship between gene expression and cortical morphology. This framework demonstrates superior performance in coordinate regression tasks compared to CNN and GCN. Meanwhile, our study reveals several genes as key contributors to cortical morphology changes by SHAP analysis, aligning with prior findings of their roles in biological process. For example, genes NFIX and CACNA2D3 with high contribution values are reported to be essential for neuronal migration and calcium signaling. Notably, our wavelet-based clustering finds gene sets associated with smooth versus convoluted morphologies, which suggests a hierarchical genetic regulation mechanism: different cortical morphologies are driven by different genes.

In conclusion, this study reveals the complex relationship between cortical morphology and gene expression and provides an in-depth exploration of gene contributions in different curve morphologies using wavelet transform and SHAP analysis methods. Our research not only offers new insights into cortical morphological development but also lays the foundation for further exploration of how genes regulate morphological changes in the brain cortex.

Disclosure of Interests. The authors have no competing interests to declare that they are relevant to the content of this article.

References

1. Chen, J., Lu, Y., Yu, Q., Luo, X., Adeli, E., Wang, Y., Lu, L., Yuille, A.L., Zhou, Y.: Transunet: Transformers make strong encoders for medical image segmentation. arXiv preprint arXiv:2102.04306 (2021)
2. Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 (2014)
3. Li, X., Liu, T., Li, Y., Li, Q., Wang, X., Hu, X., Guo, L., Zhang, T., Liu, T.: Marmoset brain ish data revealed molecular difference between cortical folding patterns. Cerebral Cortex **31**(3), 1660–1674 (2021)
4. Li, Y., Chen, H., Jiang, X., Li, X., Lv, J., Li, M., Peng, H., Tsien, J.Z., Liu, T.: Transcriptome architecture of adult mouse brain revealed by sparse coding of genome-wide in situ hybridization images. Neuroinformatics **15**, 285–295 (2017)
5. Liu, J., Liu, W., Yang, L., Wu, Q., Zhang, H., Fang, A., Li, L., Xu, X., Sun, L., Zhang, J., et al.: The primate-specific gene tmem14b marks outer radial glia cells and promotes cortical expansion and folding. Cell stem cell **21**(5), 635–649 (2017)
6. Lundberg, S.M., Lee, S.I.: A unified approach to interpreting model predictions. Advances in neural information processing systems **30** (2017)
7. Shimogori, T., Abe, A., Go, Y., Hashikawa, T., Kishi, N., Kikuchi, S.S., Kita, Y., Niimi, K., Nishibe, H., Okuno, M., et al.: Digital gene atlas of neonate common marmoset brain. Neuroscience research **128**, 1–13 (2018)
8. Shin, J., French, L., Xu, T., Leonard, G., Perron, M., Pike, G.B., Richer, L., Veillette, S., Pausova, Z., Paus, T.: Cell-specific gene-expression profiles and cortical thickness in the human brain. Cerebral cortex **28**(9), 3267–3277 (2018)
9. Singh, A., Del-Valle-Anton, L., de Juan Romero, C., Zhang, Z., Ortuño, E.F., Mahesh, A., Espinós, A., Soler, R., Cárdenas, A., Fernández, V., et al.: Gene regulatory landscape of cerebral cortex folding. Science Advances **10**(23), eadn1640 (2024)
10. Stahl, R., Walcher, T., Romero, C.D.J., Pilz, G.A., Cappello, S., Irmler, M., Sanz-Aquela, J.M., Beckers, J., Blum, R., Borrell, V., et al.: Trnp1 regulates expansion and folding of the mammalian cerebral cortex by control of radial glial fate. Cell **153**(3), 535–549 (2013)

11. Sun, T., Hevner, R.F.: Growth and folding of the mammalian cerebral cortex: from molecules to malformations. Nature Reviews Neuroscience **15**(4), 217–232 (2014)
12. Sun, Y., Cabezas, M., Lee, J., Wang, C., Zhang, W., Calamante, F., Lv, J.: Predicting human brain states with transformer. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 136–146. Springer (2024)
13. Venugopalan, J., Tong, L., Hassanzadeh, H.R., Wang, M.D.: Multimodal deep learning models for early detection of alzheimer's disease stage. Scientific reports **11**(1), 3254 (2021)
14. Woo, S., Park, J., Lee, J.Y., Kweon, I.S.: Cbam: Convolutional block attention module. In: Proceedings of the European conference on computer vision (ECCV). pp. 3–19 (2018)
15. Zhou, Z., Wei, D., Liu, W., Chen, H., Qin, S., Xu, P., Zuo, X.N., Luo, Y.J., Qiu, J.: Gene transcriptional expression of cortical thinning during childhood and adolescence. Human Brain Mapping **44**(10), 4040–4051 (2023)