

# Personalized phenotype encoding and prediction of pathological head development from cross-sectional images

Connor Elkhill\*, Ines. A Cruz-Guerrero, PhD\*, Jiawei Liu, PhD\*, Marius George Linguraru, DPhil†,  
Allyson Alexander, MD§, Brooke French, MD¶, Antonio R. Porras, PhD\*§¶||

\*Department of Biostatistics and Informatics, Colorado School of Public Health, Aurora, CO

†Sheikh Zayed Institute for Pediatric Surgical Innovation, Children's National Hospital, Washington, DC

§Department of Pediatric Neurosurgery, Children's Hospital Colorado, Aurora, CO

¶Department of Pediatric Plastic and Reconstructive Surgery, Children's Hospital Colorado, Aurora, CO

||Departments of Pediatrics, Surgery and Biomedical Informatics, University of Colorado School of Medicine, Aurora, CO

Email: connor.2.elkhill@cuanschutz.edu

**Abstract**—Prediction of anatomical development plays a crucial role in the selection and planning of pediatric surgical treatments. However, the rapid growth of young patients and the potential involvement of pathology makes prediction of anatomical changes challenging. We present a novel deep learning architecture to make personalized predictions of normative and pathologic head development using only cross-sectional data. We designed a novel phenotype encoder that uses domain adversarial training to create age- and sex-independent representations of patient phenotypes and growth predictor that learns to generate the head shape of patients given the anatomical effects of age, sex and pathology. These modules cooperate to instantiate patient anatomies to any age, enabling personalized predictions conditioned to the presence of pathology. We trained our model using head segmentations from cross-sectional CT images and 3D photograms and evaluated model performance on an independent longitudinal dataset. The proposed model achieved a head surface growth prediction error of  $4.93 \pm 2.29$  mm and a volumetric error  $0.16 \pm 0.11$  L in patients with cranial pathology, and  $4.61 \pm 3.28$  mm and  $0.27 \pm 0.19$  L for normative subjects, demonstrating high accuracy. Our method is the first to create age- and sex-agnostic patient phenotype representations, and to enable personalized predictions of pathological development without requiring longitudinal data for training.

**Index Terms**—craniosynostosis, generative adversarial network, domain adversarial training, pediatric development

## I. INTRODUCTION

Children undergo a rapid cranial growth during the first few years of life that is crucial for their cognitive development [1]. Craniosynostosis is a condition in which one or more of the cranial sutures fuse prematurely, constraining cranial and brain development and often requiring surgical intervention [2]. Although predicting head development in these patients is essential to plan treatments and decide their optimal timing, existing growth charts based on simple clinical metrics such as head circumference or intracranial volume [3] cannot

This work was supported by the National Institute of Dental & Craniofacial Research (NIDCR) of the National Institutes of Health under Award Number R01DE032681. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

adequately characterize or make personalized predictions of development [4]. Therefore, a method to predict local head development in these patients would be invaluable to assess pathology in young children.

Given the scarcity of pediatric longitudinal imaging datasets, several methods to predict cranial growth using only cross-sectional training datasets have recently been proposed. Liu et al [5] created a normative reference of cranial growth based on age and sex using principal component analysis and temporal regression. However, this method only identified average growth trajectories in the normative pediatric population. Porras et al [6] presented a personalized predictive model of cranial growth, but the optimization method used is not computationally feasible for large datasets with high temporal resolution. Recently, Liu et al [7] presented a data-driven model of cranial suture growth trained on cross-sectional data to predict pathological development in patients with craniosynostosis. Despite high predictive accuracy, this method requires the observation of the cranial sutures from computed tomography (CT) images, an image modality that results in harmful radiation exposure and is often avoided in young children [8]. In a different domain, Xia et al [9] utilized generative adversarial networks to learn subject-specific trajectories from a cross-sectional dataset of magnetic resonance (MR) images to predict changes in the brain morphology of patients with Alzheimer's disease. While providing satisfactory results in adults, their paired training scheme does not allow capturing significant anatomical structural changes, which is not realistic when modeling pediatric development [3].

While CT imaging has traditionally been used to evaluate pediatric cranial malformations, 3D photogrammetry has become a popular radiation-free and low-cost clinical alternative [10]. However, 3D photogrammetry can only image the head surface, rendering prior personalized methods [7] that rely on the identification of the cranial sutures unavailable. We present a novel deep learning architecture designed to predict normative and pathologic head development trained

using only head surface information from both cross-sectional CT images and 3D photographs, thus enabling a radiation-free personalized prediction of pediatric development in a clinical setting. Specifically, we designed a phenotype encoder that utilizes domain adversarial training to generate age- and sex-agnostic latent representations of patient phenotypes. Then, we propose a temporal growth predictor that instantiates these representations to specific ages, using learned age- and sex-specific anatomical distributions in presence or absence of pathology in the pediatric population. The model was trained using only cross-sectional data from normative children and patients with craniosynostosis younger than 10 years and was evaluated using an independent longitudinal dataset.

## II. MATERIALS AND METHODS

### A. Data

After IRB approval at University of Colorado Anschutz Medical Campus (IRB #20-1563), we collected two independent retrospective datasets: (1) a cross-sectional dataset used for model training; and (2) a longitudinal dataset used for performance evaluation. Our datasets include CT images, which are traditionally available to evaluate pediatric cranial malformations, and 3D photographs, which are often used clinically as a radiation-free alternative. *Dataset 1* includes 2,672 cross-sectional CT images ( $N=2,262$ ) and 3D photographs ( $N=410$ ) of two patient populations: 2,020 normative subjects without cranial pathology (1081 male, 939 female, age  $3.14 \pm 3.05$  years, range 0-10 years) and 652 patients with craniosynostosis before any corrective surgery (384 male, 268 female, age  $0.64 \pm 1.04$  years, range 0-8.8 years). *Dataset 2* includes two groups of longitudinal images: 61 CT image pairs from 51 normative subjects (28 male, 23 female) with average age at first image  $2.24 \pm 2.22$  years and age at second image  $3.55 \pm 2.71$  years (ten subjects had two image pairs); and 75 pairs of 3D photographs from 75 patients with craniosynostosis before any corrective surgery (58 male, 17 female), with age of  $0.61 \pm 1.07$  years at the first image and  $0.85 \pm 1.18$  years at the second image.

### B. Standard representation of the head surface

We created standardized 2D anatomical representations from both CT images and 3D photographs using our previous methods [6]. In the CT images, four cranial landmarks were automatically identified at the glabella, two temporal processes of the dorsum sellae and the opisthion (Fig. 1a). In 3D photographs, homologous anatomical landmarks (Fig. 1b) were automatically identified on the head surface using our previous methods [11]. Each image was aligned to a manually annotated template using these landmarks. Then, the surface around the calvaria was segmented from the rest of the head at the cranial base (Fig. 1c). Finally, the head surface was sampled in spherical coordinates to create a standard 2D representation [6] where every pixel contains the corresponding coordinates on the head surface, as presented in Fig. 1d.

### C. Model Architecture

We propose a neural network architecture composed of three main modules as presented in Fig. 2.

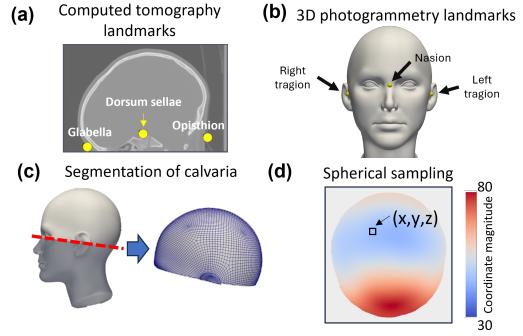


Fig. 1. (a) CT image with anatomical landmarks. (b) 3D photograph with anatomical landmarks. (c) Segmented head surface using the nasion-tragion plane. (d) 2D head surface representation using spherical sampling, where each pixel contains the corresponding (X, Y, Z) coordinates on the head surface.

**Personalized phenotype encoder (PPE).** This module creates a vector representation of the head phenotype of a patient independent from age and sex. The PPE employs the convolutional architecture presented in [12] to generate a latent patient-specific phenotype representation  $l$  based on a real head shape observation  $\mathbf{X}$ . We then incorporate a phenotype discriminator (PD, purple in Fig. 2) and a domain adversarial training scheme [13] (see Eq. (1)) to promote a phenotype representation  $l$  that is independent from age or sex. Hence, the encoder will learn to generate latent phenotype representations  $l$  that cannot be used by the PD to identify the age and sex of the patient.

**Growth predictor (GP).** The GP is trained to combine the latent representation of a patient's head phenotype  $l$  (from the PPE) with age, sex, and pathology information (i.e., what sutures are fused), to predict a personalized head anatomy at the desired age, sex, and pathological status. The GP follows the convolutional architecture from [12] but modified to include the conditions of patient age, sex (binary encoded) and suture fusion status (one-hot encoded).

**Growth discriminator (GD).** This module (Fig 2, in green) enables adversarial training of both the GP and PPE. The GD also utilizes the convolutional architecture proposed in [12] but modified to include the conditions of age, sex, and pathology. This module is trained to distinguish between real head shapes and head shapes reconstructed from  $l$ .

### D. Optimization

We defined the prediction of age and sex from the PD as  $S', A' = PD(l; \theta_{PD})$ , where  $\theta_{PD}$  are the learned parameters of the PD and  $l$  is the latent representation generated from the PPE using a real input image  $\mathbf{X}$ . We computed the cross entropy between  $S$  and  $S'$  and KL divergence between the continuous distributions of  $A$  and  $A'$ , where  $A$  and  $S$  are the true patient age and sex. The PPE and PD modules are trained using the following adversarial loss function:

$$L_{PPPE} = \max_{PPE} \min_{PD} S \log(S') + (1-S) \log(1-S') + A \log \frac{A}{A'}. \quad (1)$$

To train the GP and GD modules, we used an adversarial scheme with the following Wasserstein loss function [14], conditioned by  $A, S$  and pathology status  $C$ :

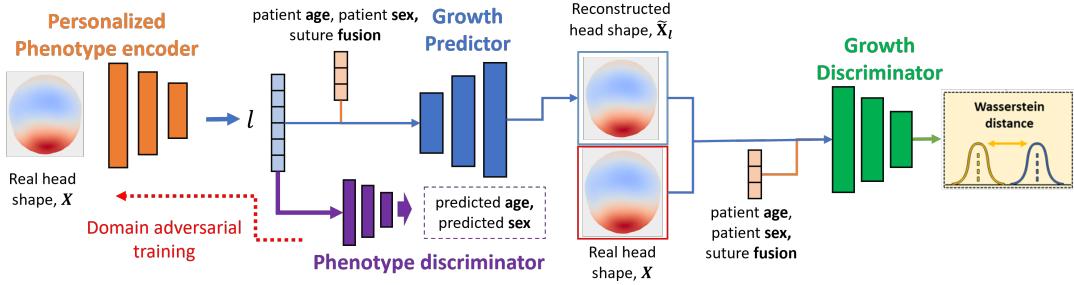


Fig. 2. **Proposed architecture:** The phenotype encoder creates a patient-specific latent representation from head shape, independent of age or sex. A growth predictor, conditioned on age, sex, and pathology, reconstructs head anatomies. The trained model generates personalized predictions at any age.

$$L_{GP} = \mathbb{E}_{\tilde{\mathbf{X}}_l \sim P_l | \mathbf{X} \sim P_x} GD(\mathbf{X}_l | A, S, C) - GD(\mathbf{X} | A, S, C) + \lambda_{grad} \mathbb{E}_{\tilde{\mathbf{X}} \sim P_{\tilde{\mathbf{X}}}} [\|\nabla_{\tilde{\mathbf{X}}} GD(\tilde{\mathbf{X}}, A, S, C, \Theta_{GD})\|_2 - 1]^2, \quad (2)$$

where  $\mathbb{E}_{\tilde{\mathbf{X}}_l \sim P_l | \mathbf{X} \sim P_x}$  is the expected value of the output of  $GD$  given the probability distributions  $P_x$  and  $P_l$  of the real images  $\mathbf{X}$  and reconstructed images from the GP  $\mathbf{X}_l$ , respectively. The second term of Eq (2) is a penalty to the gradient with respect to the model parameters, aimed to optimize discriminator performance and enable calculation of Wasserstein distance [14]. For each set of input images  $\mathbf{X}$  and  $\mathbf{X}_l$  to  $GD$ , we uniformly sampled points  $\hat{\mathbf{X}}$  from the distribution  $P_{\hat{\mathbf{X}}}$  between the real input image distribution  $P_x$  and reconstructed image distribution  $P_{\mathbf{X}_l}$  and computed the L2-norm of the gradient with respect to the model parameters  $\Theta_{GD}$ . This penalty is weighted in the loss using the hyperparameter  $\lambda_{grad}$ . To train the entire model, the two loss functions presented in Eq. (1) and Eq. (2) were combined with an additional term quantifying the reconstruction error between the original image  $\mathbf{X}$  and the reconstructed image  $\tilde{\mathbf{X}}_l$  to promote accurate image predictions. The loss function used during training was:

$$L = L_{GP} + \lambda_{PPE} L_{PPE} + \lambda_{rec} \|\tilde{\mathbf{X}}_l - \mathbf{X}\|_2, \quad (3)$$

where  $\lambda_{PPE}$  and  $\lambda_{rec}$  are hyperparameters balancing the contribution of the domain-adversarial and reconstruction losses, respectively.

### III. EXPERIMENTS AND RESULTS

#### A. Image pre-processing and training details

We randomly divided *Dataset 1* into training (90%, N=1,712) and validation (10%, N=190) sets. We processed each image as described in Section 2.2 and represented them using an image resolution of  $64 \times 64$  pixels, to be compatible with [12]. We normalized all image pixel components to the range [-1,1] and age to the range [0,1]. Sex was encoded as 1 for male and 0 for female and suture fusion status was one-hot encoded for each of the following sutures: sagittal, metopic, right coronal, and left coronal. Patients with lambdoid suture fusion were excluded due to insufficient data from this rare form of craniosynostosis. We facilitated convergence by initializing the GP and GD parameters with synthetic representations  $l$  using randomly sampled noise from a standard normal distribution, training only with the loss function from Eq. 2. Once

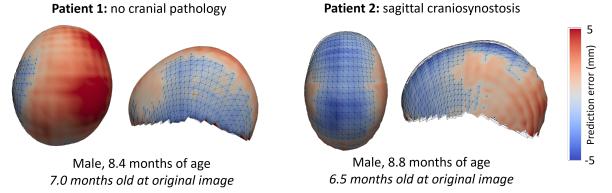


Fig. 3. Growth predicted from the original image of a normative subject (left) and a patient with sagittal craniostenosis (right). Colors represent the prediction error in mm, and the mesh overlay represents the true image.

initialized, we trained the entire network using Eq. 3 and the Adam optimizer for no more than 5,000 epochs, stopping upon convergence of the validation loss or no improvements after 100 consecutive epochs. Diagrams describing the parameters of our model architecture, the fully trained network, and additional hyperparameters used for training can be found at: <https://github.com/cspf-mi2/sipaim-2024>.

#### B. Performance Evaluation

We organized each longitudinal image pair in *Dataset 2* in a bidirectional fashion, resulting in 272 total pairs. We computed the surface prediction error as the mean point-wise Euclidean distance in millimeters between the predicted and true images and computed difference in volume between the predicted and true head shape. Table I shows the performance of our proposed network and compared to alternative methods, including: a PCA-based regression model of normative head development in the pediatric population [5], a conditional DCGAN [12], and the model of brain aging available from [9]. We also included two ablation studies: our method without the use of domain-adversarial loss and without the condition of pathology to make predictions. All models in Table I were trained on *Dataset 1* and evaluated using *Dataset 2*.

Finally, Fig. 3 shows an example of longitudinal predictions using our proposed model of two patients: a normative subject without cranial pathology and a patient with craniostenosis.

### IV. DISCUSSION

We present a novel deep learning network trained using only cross-sectional data to make personalized predictions of anatomical development for patients with and without pathology. The results detailed in Table I demonstrate that our proposed network can infer anatomical development with high accuracy.

TABLE I  
MEAN SURFACE PREDICTION AND VOLUMETRIC ERRORS FOR PAIRS OF IMAGES FROM PATIENTS WITH CRANIOSYNOSTOSIS (N=75) AND NORMATIVE SUBJECTS (N=61). P-VALUES ESTIMATED USING A WILCOXON SIGNED-RANK TEST BETWEEN THE PROPOSED METHOD AND OTHER METHODS.

Model	Surface prediction error (mm)		Volumetric Error (L)	
	Craniosynostosis	Normative	Craniosynostosis	Normative
Proposed method	<b>4.93 ± 2.29</b>	4.61 ± 3.28	<b>0.16 ± 0.11</b>	0.27 ± 0.19
w/o domain adversarial loss	5.22 ± 2.34 (p=0.04)	5.22 ± 3.44 (p<0.005)	0.17 ± 0.12 (p=0.21)	0.29 ± 0.21 (p=0.09)
w/o condition	6.15 ± 2.02 (p<0.005)	6.90 ± 3.48 (p<0.005)	0.17 ± 0.12 (p=0.35)	0.31 ± 0.19 (p<0.005)
PCA Model [5]	5.06 ± 4.24 (p=0.20)	<b>3.77 ± 2.80</b> (p=0.144)	0.28 ± 0.24 (p<0.005)	0.26 ± 0.20 (p=0.06)
Conditional-DCGAN [12]	5.39 ± 2.28 (p<0.005)	5.29 ± 3.51 (p=0.01)	0.17 ± 0.14 (p=0.80)	<b>0.22 ± 0.19</b> (p=0.03)
BrainAgeing [9]	6.63 ± 3.81 (p<0.005)	7.20 ± 5.57 (p<0.005)	0.26 ± 0.18 (p<0.005)	0.29 ± 0.27 (p=0.93)

Our proposed method achieved the lowest surface prediction error in patients with craniosynostosis in comparison to all other methods. We improved prediction accuracy using the proposed domain-adversarial loss function and found that removing the condition of pathological status reduces performance, supporting the importance of pathology as a mediator of pediatric development [4]. Our model also outperformed a Conditional-DCGAN, which only uses non-personalized, randomly sampled noise, and the Brain Ageing GAN, which is trained to model small anatomical changes between longitudinal observations and cannot predict the significant anatomical changes present in pediatric development.

However, we found that the PCA-based regression model of normative development provides the lowest average surface prediction error in normative subjects compared to other methodologies, which we hypothesize is due to the substantially higher image resolution employed by this model ( $500 \times 500$  pixels) [5]. Despite the lower image resolution of our proposed method ( $64 \times 64$  pixels), the differences in surface prediction error between our proposed method and the PCA-based regression model are not significant.

Finally, our proposed method produces significantly lower volumetric errors in patients with craniosynostosis than the PCA-based regression normative model which is trained only on normative subjects and cannot predict pathological development. We also found our proposed method provides similar volumetric errors to Conditional-DCGAN in patients with craniosynostosis, likely due to the shared training procedure and network architecture in both models. One limitation of our study is the small size of our longitudinal evaluation datasets, as these datasets are rare in children.

## V. CONCLUSION

We presented a novel deep learning network to predict pediatric anatomical development that, in addition to CT, can be used with 3D photogrammetry, a radiation-free modality. We designed a novel phenotype encoder to generate age- and sex-independent latent phenotype representations and a growth predictor that learns age-, sex-, and pathology-specific anatomical distributions of head shape. Our network demonstrated high predictive accuracy evaluated on an independent longitudinal dataset of patients with and without cranial pathology. Our age- and sex-independent phenotype representations could also be leveraged to study associations between patient phenotype and pathology in diverse pediatric datasets.

## REFERENCES

- [1] M. Cao, H. Huang, and Y. He, "Developmental connectomics from infancy through early childhood," *Trends in Neurosciences*, vol. 40, pp. 494–506, 8 2017.
- [2] I. M. Mathijssen, "Updated guideline on treatment and management of craniosynostosis," *Journal of Craniofacial Surgery*, vol. 32, pp. 371–450, 1 2021.
- [3] P. Meyer-Marcotty, H. Bohm, C. Linz, J. Kochel, A. Stellzig-Eisenhauer, and T. Schweitzer, "Three-dimensional analysis of cranial growth from 6 to 12 months of age," *The European Journal of Orthodontics*, vol. 36, pp. 489–496, 10 2014.
- [4] R. Breakey, P. G. Knoops, A. Borghi, N. Rodriguez-Florez, J. O'Hara, G. James, D. J. Dunaway, S. Schievano, and N. O. Jeelani, "Intracranial volume and head circumference in children with unoperated syndromic craniosynostosis," *Plastic and reconstructive surgery (1963)*, vol. 142, no. 5, pp. 708e–717e, 2018. publisher-place: United States publisher: United States: American Society of Plastic Surgeons.
- [5] J. Liu, C. Elkhill, S. LeBeau, B. French, N. Lepore, M. G. Linguraru, and A. R. Porras, "Data-driven normative reference of pediatric cranial bone development," *Plastic and Reconstructive Surgery - Global Open*, vol. 10, p. e4457, 8 2022.
- [6] A. R. Porras, R. Keating, J. Lee, and M. G. Linguraru, "Predictive statistical model of early cranial development," *IEEE Transactions on Biomedical Engineering*, vol. 69, pp. 537–546, 2 2022.
- [7] J. Liu, J. H. Froelicher, B. French, M. G. Linguraru, and A. R. Porras, "Data-driven cranial suture growth model enables predicting phenotypes of craniosynostosis," *Scientific Reports*, vol. 13, p. 20557, 11 2023.
- [8] T. Schweitzer, H. Böhm, P. Meyer-Marcotty, H. Collmann, R.-I. Ernestus, and J. Krauß, "Avoiding ct scans in children with single-suture craniosynostosis," *Child's Nervous System*, vol. 28, pp. 1077–1082, 7 2012.
- [9] T. Xia, A. Chartsias, C. Wang, and S. A. Tsafaris, "Learning to synthesise the ageing brain without longitudinal data," *Medical Image Analysis*, vol. 73, p. 102169, 10 2021.
- [10] M. S. Kurniawan, P. A. Tio, T. Abdel Alim, G. Roshchupkin, C. M. Dirven, M. M. Pleumeeckers, I. M. Mathijssen, and M.-L. C. van Veelen, "3D Analysis of the Cranial and Facial Shape in Craniosynostosis Patients: A Systematic Review," *Journal of Craniofacial Surgery*, vol. 35, no. 3, 2024.
- [11] C. Elkhill, J. Liu, M. G. Linguraru, S. LeBeau, D. Khechoyan, B. French, and A. R. Porras, "Geometric learning and statistical modeling for surgical outcomes evaluation in craniosynostosis using 3d photogrammetry," *Computer Methods and Programs in Biomedicine*, vol. 240, p. 107689, 10 2023.
- [12] A. Radford, L. Metz, and S. Chintala, "Unsupervised representation learning with deep convolutional generative adversarial networks," *arXiv preprint arXiv:1511.06434*, 1 2016.
- [13] Y. Ganin, E. Ustinova, H. Ajakan, P. Germain, H. Larochelle, F. Laviolette, M. Marchand, and V. Lempitsky, "Domain-adversarial training of neural networks," *arXiv.org*, 2016. publisher-place: Ithaca publisher: Ithaca: Cornell University Library, arXiv.org.
- [14] I. Gulrajani, F. Ahmed, M. Arjovsky, V. Dumoulin, and A. C. Courville, "Improved training of wasserstein gans," *Advances in Neural Information Processing Systems*, vol. 30, 2017.