# Disentangled Geometric Variational Learning of Normative Pediatric Development from Anatomical Surface Representations of Patients with Pathology

Joseph Nagel[1], Ines A. Cruz-Guerrero[1,2], Brooke French[3], and Antonio R. Porras[1,2]

[1] Department of Biostatistics and Informatics, Colorado School of Public Health, University of Colorado Anschutz Medical Campus, Aurora, CO 80045, USA
[2] Department of Neurological Surgery, The University of Chicago, Chicago, IL 60637, USA
[3] Department of Pediatric Plastic and Reconstructive Surgery, Children's Hospital Colorado, Aurora, CO 80045, USA
joseph.nagel@cuanschutz.edu, alejandro.cruz@bsd.uchicago.edu

**Abstract.** The objective quantification of pediatric craniofacial anomalies is essential for diagnosis of pathology and treatment planning. However, clinical evaluation still relies on subjective assessment or metrics with low sensitivity because existing craniofacial anatomy models have been mostly developed for adult applications using methods that typically fail to capture rapid anatomical changes during pediatric growth given the unavailability of large normative pediatric datasets. Moreover, they normally require extensive pre-processing and manual annotations to create standardized anatomical surface meshes and often impose strong statistical assumptions. We introduce a novel geometric learning model that operates directly on diverse anatomical surface representations of pediatric patients with pathology and does not require pre-processing. Our model infers the normative distributions in the population by disentangling anatomical representations from pathology, age and sex information. First, it learns a probabilistic mapping between input surfaces and a uniform, geometrically consistent latent space. Then, a new variational-adversarial training framework removes pathology, age and sex information to learn population-wide statistical distributions, which enables either generating personalized normative references to characterize pathology or realistic synthetic data. Trained on 2,696 3D photograms with diverse craniofacial pathologies and evaluated on normative anatomical regions, our model achieved a point-to-surface distance of $1.50 \pm 0.72$ mm. Moreover, our personalized normative references allowed identifying phenotypes of cranial pathology with 96% accuracy. Finally, synthetically generated data showed no significant volumetric differences compared to real data during pediatric development ($p = 0.13$). Our results show that our method can generate personalized normative anatomical references and synthesize realistic pediatric datasets.

**Keywords:** Geometric Learning · Variational Inference · Personalized Normative References · Synthetic Data Generation.

## 1 Introduction

Pediatric craniofacial development anomalies often indicate the presence of pathology [16], and their early detection is paramount for children to receive prompt care. Beyond diagnosis, the morphological assessment of craniofacial anomalies is essential for surgical planning and treatment outcomes evaluation. However, their traditional clinical assessment relies on subjective expert evaluation and/or simple landmark-based metrics [1] that lack sensitivity to characterize pathology [6] and often suffer from ambiguity and reproducibility issues [7].

Although 3D Morphable Models (3DMMs) have been utilized to quantitatively assess craniofacial morphologies [9, 14], their use has largely focused on adult data because of the challenges of modeling the rapid anatomical changes during pediatric development [8], especially given the typically small available pediatric image datasets. Moreover, most craniofacial 3DMMs have focused on modeling either the face or the cranium, motivated by their distinct developmental pathways and growth profiles [13]. Although recent methods have addressed some of these challenges [8, 15], they typically require significant pre-processing or annotations to standardize anatomical mesh representations, which limits their scalability, reproducibility, and integration into clinical pipelines [9, 11, 14]. Moreover, they often impose strong statistical assumptions that are ill-suited for capturing the complex and non-linear morphological changes during pediatric development, which present distinctive patterns between the face and cranium [2]. In addition, since medical images of healthy children are rare, existing models are either derived from small datasets [11] or focused on specific pathologies [15].

We present a novel variational deep learning framework to model the normative development of the craniofacial complex using head surface representations from children with diverse craniofacial pathologies. Our model builds upon a recent mesh standardization network [3] that creates uniform anatomical representations of clinical 3D photograms with variable pose, resolution and connectivity while avoiding pre-processing or manual annotations. Our method then incorporates a novel adversarial pathology-disentangled encoding approach that removes pathology information from patient-specific phenotype encodings, in addition to age and sex information, to obtain personalized normative anatomical references. These encodings are then used by a decoding architecture together with age and sex information to reconstruct craniofacial surface meshes representing personalized normative anatomical references, which can be used to quantitatively characterize pathology. Importantly, our variational approach also enables learning the statistical distributions of craniofacial anatomy in the normative population independently from age and sex, and the generation of craniofacial meshes of synthetic normative children with variable age and sex.

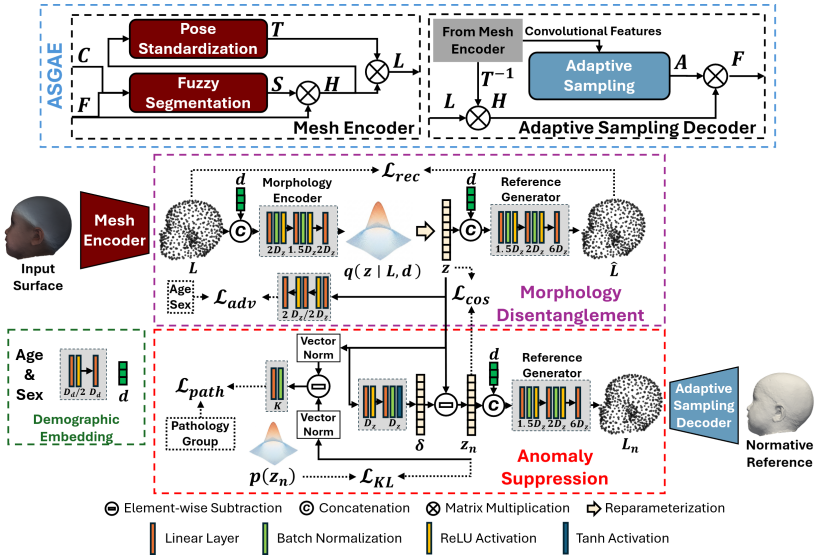


**Fig. 1.** Overview of the proposed model architecture.

## 2 Materials and Methods

Fig. 1 presents an overview of the proposed architecture. In summary, a Mesh Encoder creates pose-standardized and uniform latent representations of geometrically consistent pseudo-landmarks identified in an unsupervised manner from anatomical surfaces with variable pose, resolution, number of vertices and connectivity. The Morphology Disentanglement module isolates a patient's specific phenotype from demographic factors (age and sex). Next, the Anomaly Suppression module decomposes the phenotype into a combination of a normative representation and its modifications attributable to pathology, and generates a patient-specific normalized pseudo-landmark representation. Finally, the Adaptive Sampling Decoder module reconstructs that representation into a full-resolution surface mesh that represents a personalized normative anatomical reference.

### 2.1 Surface Mesh Encoder and Adaptive Sampling Decoder

The Mesh Encoder and Adaptive Sampling Decoder are based on the Adaptive Sampling Graph Autoencoder (ASGAE) [3], an unsupervised geometric model designed to create standardized, explainable pseudo-landmark representations of 3D surface meshes with variable pose, node distribution, and connectivity. The Mesh Encoder (see Fig. 1) receives a surface mesh represented as an undirected graph $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathbf{F}, \mathbf{C})$, where $\mathcal{V}$ is a set of $N$ nodes, $\mathcal{E}$ is a set of edges, $\mathbf{F} \in \mathbb{R}^{N\times 3}$ is a matrix of Euclidean coordinates, and $\mathbf{C} \in \mathbb{R}^{N\times 3}$ is a matrix

of RGB values. As in ASGAE, this module segments in an unsupervised manner a set of $D_{\mathbf{L}}$ representative Euclidean points that form an interpretable and ordered point cloud $\mathbf{L} \in \mathbb{R}^{D_{\mathbf{L}} \times 3}$ of geometrically consistent pseudo-landmarks across input surfaces [3]. This segmentation is performed using two blocks: Fuzzy Segmentation and Pose-Standardization. The Fuzzy Segmentation block applies geometric convolutions to learn a normalized probabilistic linear transformation $\mathbf{S} \in \mathbb{R}^{N \times D_{\mathbf{L}}}$ between the surface mesh space and the pseudo-landmark space, operating directly on $\mathbf{F}$ and $\mathbf{C}$. Hence, the initial pseudo-landmark representation of an input surface is defined as $\mathbf{H} = \mathbf{S}^{\top}\mathbf{F} \in \mathbb{R}^{D_{\mathbf{L}} \times 3}$. The Pose-Standardization block removes pose from $\mathbf{H}$ using a rigid transformation matrix $\mathbf{T} \in \mathbb{R}^{4 \times 4}$, estimated via Ordinary Procrustes Analysis [4], to obtain the pose-normalized pseudo-landmark representation $\mathbf{L} = \mathbf{HT}$.

For the generation of full resolution meshes, we used ASGAE's Adaptive Sampling Decoder [3] (see Fig. 1), which reconstructs a surface mesh from its pseudo-landmark representation. Given a pseudo-landmark representation $\mathbf{L}$, it restores the pose via $\mathbf{H} = \mathbf{LT}^{-1} \in \mathbb{R}^{D_{\mathbf{L}} \times 3}$. The Adaptive Sampling Decoder leverages the learned convolutional features from the Fuzzy Segmentation block and an attention block to learn a non-linear transformation matrix $\mathbf{A} \in \mathbb{R}^{N \times D_{\mathbf{L}}}$ that converts $\mathbf{H}$ back into a mesh feature matrix as $\mathbf{F}' = \mathbf{AH} \in \mathbb{R}^{N \times 3}$, which is used to reconstruct the original surface mesh from its pseudo-landmark representation as $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathbf{F}')$. Note, in this work, RGB values were only used to facilitate learning, but were not encoded or reconstructed since they are not relevant to anatomy. Additional details can be found in [3].

### 2.2 Demographic-Independent Pathology Disentanglement

Using the standardized pseudo-landmark representations $\mathbf{L}$ produced by the Mesh Encoder, we propose a variational framework for separating pathological from normative morphology, and for modeling the normative anatomical distributions in the population independently from age and sex. This independence is essential to create a generative model of development using only cross-sectional data [10]. Our approach consists of two modules: Morphology Disentanglement (MD) and Anomaly Suppression (AS). The MD module attempts to learn patient phenotype representations that are independent from age and sex. The AS module aims to decompose these phenotype representations into two components: a normative phenotype and the changes induced by pathology. Then, it models the statistical anatomical distributions in the normative population using standard normal distributions. Note that normally distributed latent features do not necessarily translate into normally distributed Euclidean surface mesh representations.

**Morphology Disentanglement.** This module uses variational training to learn the mean and variance representing the confidence interval of a set of encoding features from $\mathbf{L}$ through a Morphology Encoder (ME) block. The ME block

receives the concatenation of the pseudo-landmark representation $\mathbf{L}$ and a demographic embedding of age and sex $[\text{vec}(\mathbf{L}); \mathbf{d}] \in \mathbb{R}^{3D_{\mathbf{L}}+D_{\mathbf{d}}}$ learned by a demographic embedding block (Fig. 1). Then, it learns the latent representation of specific patient phenotypes conditioned on age and sex as $q(\mathbf{z} \mid \mathbf{L}, \mathbf{d}) \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, where $\boldsymbol{\mu} \in \mathbb{R}^{D_{\mathbf{z}}}$ is the predicted mean vector and $\boldsymbol{\Sigma} \in \mathbb{R}^{D_{\mathbf{z}} \times D_{\mathbf{z}}}$ is a diagonal covariance matrix. In our implementation, the ME block learns the logarithm of the diagonal of $\boldsymbol{\Sigma}$ to facilitate training. A latent phenotype vector $\mathbf{z} \in \mathbb{R}^{D_{\mathbf{z}}}$ is obtained by sampling from $q(\mathbf{z} \mid \mathbf{L}, \mathbf{d})$ via reparameterization [7], which allows for stochastic sampling while preserving differentiability. Then, a multi-layer perceptron network predicts age ($\widehat{a}$) and sex ($\widehat{s}$) from $\mathbf{z}$ and promotes demographic independence via gradient reversal [5]. See layer details in Fig. 1.

**Anomaly Suppression.** This module models the population-wide normative distribution of anatomy as $p(\mathbf{z}_n) \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{D_{\mathbf{z}}})$, and decomposes each patient's phenotype into a normative anatomy vector $\mathbf{z}_n \in \mathbb{R}^{D_{\mathbf{z}}}$ modeled as an observation from $p(\mathbf{z}_n)$, and a pathological deviation vector $\boldsymbol{\delta} \in [-\varepsilon, \varepsilon]^{D_{\mathbf{z}}}$, where $\varepsilon$ is a hyperparameter representing the number of standard deviations away from normality that the model allows. Thus, the phenotype $\mathbf{z}$ of a patient is regarded as $\mathbf{z} = \mathbf{z}_n + \boldsymbol{\delta}$, where $\boldsymbol{\delta}$ is estimated from $\mathbf{z}$ by a fully connected linear layer with ReLU activation followed by a second linear layer, batch normalization and hyperbolic tangent activation (see Fig. 1). We include a classification block consisting of batch normalization and a linear layer. It receives the difference between the normalized $\mathbf{z}$ and $\mathbf{z}_n$, and outputs class probabilities $\widehat{\mathbf{p}} \in [0, 1]^K$ that $\mathbf{z}$ belongs to the $k^{\text{th}}$ pathology group. This design encourages pathological variation to be captured in $\boldsymbol{\delta}$, disentangling it from $\mathbf{z}_n$.

Finally, a Reference Generator (RG) block learns to convert $[\mathbf{z}_n; \mathbf{d}] \in \mathbb{R}^{D_{\mathbf{z}}+D_{\mathbf{d}}}$ to pseudo-landmarks to produce a patient-specific normative pseudo-landmark reference $\mathbf{L}_n \in \mathbb{R}^{D_{\mathbf{L}} \times 3}$ conditioned on $\mathbf{d}$, which is subsequently decoded into a full-resolution surface mesh representing a personalized normative anatomical reference mesh as $\mathcal{G}(\mathcal{V}, \mathcal{E}, \mathbf{F}_n)$, where $\mathcal{V}$ and $\mathcal{E}$ are the input nodes and edges and $\mathbf{F}_n$ represents the Euclidean coordinate feature matrix derived from the geometry of $\mathbf{L}_n$. Importantly, this framework also enables the generation of synthetic normative anatomies by sampling $\mathbf{z}'_n \sim p(\mathbf{z}_n)$ and combining the sampled vector $\mathbf{z}'_n$ with a demographic embedding $\mathbf{d}'$ of the desired age and sex.

### 2.3 Optimization

We initially trained the two modules from ASGAE (Surface Encoder and Adaptive Sampling Decoder) as described in [3] to generate standardized pseudo-landmark representations. Then, for the MD and AS modules to separate normative from pathological phenotypes, we used five loss terms. First, we encouraged accurate reconstruction of pseudo-landmarks $\widehat{\mathbf{L}}$ from the learned latent representation $\mathbf{z}$ obtained from $\mathbf{L}$ using the loss function $\mathcal{L}_{rec} = \text{MSE}(\mathbf{L}, \widehat{\mathbf{L}})$. Second, we promoted demographic (age and sex) invariance with the adversarial objective function

$$\mathcal{L}_{adv} = \max_{\text{ME}} \min_{\text{DA}} \left( \lambda_{age} \, \text{MSE}(a, \widehat{a}) + \text{BCE}(s, \widehat{s}) \right), \tag{1}$$

where $\widehat{a}$ and $\widehat{s}$ are the predicted age and sex, and $a \in \mathbb{R}$ and $s \in \{0, 1\}$ denote ground-truth, BCE$(\cdot)$ denotes binary cross-entropy, and $\lambda_{age}$ is a hyperparameter that balances the contribution of the demographic variables. Third, we promoted population-wide normative distributions for each feature in $\mathbf{z}_n$ using a KL divergence loss as

$$\mathcal{L}_{KL} = \frac{1}{D_{\mathbf{z}}} D_{KL}\left(\mathcal{N}\left(\widehat{\boldsymbol{\mu}}_{\mathbf{z}_n}, \widehat{\boldsymbol{\Sigma}}_{\mathbf{z}_n}\right) \parallel \mathcal{N}\left(\mathbf{0}, \mathbf{I}_{D_{\mathbf{z}}}\right)\right), \tag{2}$$

where the mean $\widehat{\boldsymbol{\mu}}_{\mathbf{z}_n} \in \mathbb{R}^{D_z}$ and covariance $\widehat{\boldsymbol{\Sigma}}_{\mathbf{z}_n} \in \mathbb{R}^{D_{\mathbf{z}} \times D_{\mathbf{z}}}$ are computed over the phenotype vectors $\mathbf{z}_n$ and $\mathcal{N}\left(\mathbf{0}, \mathbf{I}_{D_{\mathbf{z}}}\right)$ is a multivariate standard normal distribution. Fourth, we promoted the separation of pathological phenotype traits via a cross-entropy loss $\mathcal{L}_{path} = \text{CE}(y, \widehat{\mathbf{p}})$, where $y \in \{1, \ldots, K\}$ is the true pathology label for our patients (cranial vs. facial pathology). Finally, to preserve similar representations between the patient phenotype $\mathbf{z}$ and its normalized version $\mathbf{z}_n$, we used a cosine distance loss $\mathcal{L}_{cos} = 1 - \frac{\mathbf{z}^\top \mathbf{z}_n}{\|\mathbf{z}\| \|\mathbf{z}_n\|}$. The total loss to train the MD and AS modules was

$$\mathcal{L} = \lambda_{rec}\mathcal{L}_{rec} + \lambda_{adv}\mathcal{L}_{adv} + \lambda_{KL}\mathcal{L}_{KL} + \lambda_{path}\mathcal{L}_{path} + \lambda_{cos}\mathcal{L}_{cos}, \tag{3}$$

where $\lambda_{rec}, \lambda_{adv}, \lambda_{KL}, \lambda_{path}$, and $\lambda_{cos}$ are hyperparameters that balance the relative contribution from each term.

## 3 Experiments and Results

### 3.1 Data Description

After approval by the University of Colorado Anschutz Medical Campus IRB (protocol #20-1563), we collected a retrospective dataset of 2,696 3D photograms acquired with the 3dMD Head System (3dMD, Atlanta, GA) from 1,194 children (844 male, 350 female, aged $2.0 \pm 2.0$ years, range 0–10 years) with diverse craniofacial pathologies, which was used to train and evaluate our proposed model. The dataset included 427 patients with cleft lip and/or palate, 646 patients with craniosynostosis, and 121 with positional plagiocephaly. Each photogram had a variable pose, node distribution and connectivity ($56{,}154 \pm 9{,}338$ nodes and $168{,}123 \pm 27{,}999$ edges). A subset of 375 3D photograms corresponded to 280 children with single suture craniosynostosis and were acquired prior to their surgical treatment. These patients had the following phenotypic distribution: 60 metopic, 186 sagittal, 7 left coronal, 19 right coronal, 5 left lambdoid, and 3 right lambdoid. This subset was used to demonstrate the ability of our model to characterize different phenotypes of pathology.

### 3.2 Training Details

Patients were grouped into cranial (craniosynostosis and plagiocephaly) or facial (cleft lip and/or palate) pathology groups. The dataset was then split at the

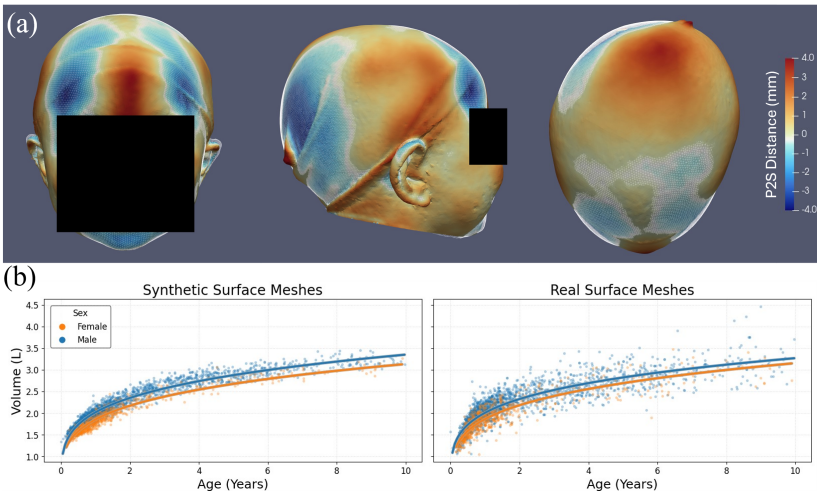


**Fig. 2.** (a) Signed P2S distance map (in mm) to normative reference (white wireframe) for a patient with metopic craniosynostosis. (b) Age- and sex-regressed volumes for the synthetic and real datasets.

patient level into training (80%), validation (10%), and test (10%) sets stratified by pathology group. The training dataset was used to train ASGAE as in [3]. The pseudo-landmark space dimension $D_{\mathbf{L}}$ was empirically set to 512. Training was performed until convergence on the validation dataset, defined as 20 consecutive epochs without loss improvements, using an Adam optimizer with a learning rate of 1E-4, cosine annealing decay to 1E-6, and batch size of 3.

Using the pretrained ASGAE model, we computed pseudo-landmark representations $\mathbf{L}$ for all surfaces, which were used to train the remainder of the model using Eq. 3. Hyperparameters were heuristically selected as $\lambda_{rec} = 1$, $\lambda_{KL} = 1$, $\lambda_{path} = 5$, $\lambda_{adv} = 1$, $\lambda_{age} = 3$, $\lambda_{cos} = 1$, $D_{\mathbf{z}} = 256$ and $\varepsilon = 4$. Training until convergence on the validation dataset was performed using the Adam optimizer with a learning rate of 1E-4, cosine annealing learning rate decay to 5E-5, weight decay of 0.1, and batch size of 256. All modules were trained on an NVIDIA RTX A6000.

### 3.3 Performance Evaluation

We first evaluated ASGAE's performance at surface mesh reconstruction on the testing dataset, which provided a point-to-surface (P2S) distance of $0.25 \pm 0.05$ mm. Next, we evaluated the accuracy of the proposed framework to generate surface meshes that are personalized. We first calculated the P2S distance between the input surfaces of our patients and the personalized normative references, which was $1.50 \pm 0.72$ mm. Then, we calculated the P2S distance between the input surfaces and randomly generated normative anatomies with the age and

sex of the input patients, obtained by replacing their personalized patient phenotypes $\mathbf{z}_n$ with random noise sampled from the learned normal distribution $p(\mathbf{z}_n)$. We achieved a P2S distance of $3.51 \pm 1.07$ mm, which was significantly higher than the distance to personalized references (Wilcoxon Signed Rank test; $p < 0.001$).

Separately, we evaluated the utility of the personalized normative reference to quantify anatomical anomalies that can characterize diverse phenotypes of pathology. Specifically, we quantified anatomical anomalies at every pseudo-landmark in $\mathbf{L}$ as their signed Euclidean distance to their normalized version in $\mathbf{L}_n$. Fig. 2 (a) shows an example of these anomalies projected on the head surface of a patient with metopic craniosynostosis and right posterior positional plagiocephaly. Then, these pseudo-landmark distances were used to train a logistic regression classifier to distinguish the diverse phenotypes in our curated subset of pre-surgical 3D photograms of 280 patients with single suture craniosynostosis (see Section 3.1). Evaluated using stratified 10-fold cross-validation stratified by patient and phenotype, and incorporating sample weighting to account for class imbalance, our classifier provided an average area under the receiver operating characteristic curve of $0.97 \pm 0.05$ and an average classification accuracy of $96\% \pm 3\%$ across phenotypes.

Finally, we evaluated that the synthetic data generated by our framework was plausible for the provided age and sex. Specifically, we randomly generated normative phenotypes $\mathbf{z}_n$ by sampling the learned statistical distributions. These were used by the RG together with age and sex to create pseudo-landmark representations of normative synthetic subjects, which were decoded into full surface meshes using ASGAE's Adaptive Sampling decoder with the nodes and connectivity of a randomly selected reference mesh. We generated as many synthetic subjects as available in our datasets following the same age and sex distributions. Then, we calculated the head volumes of both our synthetic normative surface meshes and real patients, and we regressed these volumes as a function of age, sex, and origin (synthetic or real) as:

$$\begin{aligned} \ln(volume) = \beta_0 + \beta_1 \ln(age) + \beta_2 sex + \beta_3 origin + \beta_4 \ln(age) \cdot sex \\ + \beta_5 \ln(age) \cdot origin + \beta_6 sex \cdot origin + \beta_7 \ln(age) \cdot sex \cdot origin \end{aligned} \tag{4}$$

where *origin* is a binary indicator of real (1) or synthetic (0). No significant effect of *origin* was identified (Wald test; $p = 0.13$), indicating no differences in the volume trajectories between the real and synthetic meshes, as shown in Fig. 2 (b).

## 4 Discussion and Conclusion

We introduced a novel geometric learning model that can infer the normative anatomical distributions in the population using anatomical surfaces of patients with diverse conditions and create personalized normative references to characterize pathology. Unlike state-of-the-art methods [8, 9, 11, 15], our method does

not require manual annotations or pre-processing for standardization. Additionally, our variational framework models anatomical distributions independently from age and sex without restrictive statistical assumptions in the Euclidean space, which is essential to learn anatomical variability and its significant changes during pediatric development using cross-sectional data. This is achieved through our novel MD and AS modules that disentangle pathological morphology, age, and sex from the normative anatomy, allowing for the generation of personalized references of normality or synthetic anatomical meshes.

Our extensive experiments show that the generated anatomical references are personalized, since they provide significantly higher resemblance with the input patient anatomies than synthetically generated data. We also showed that our personalized normative references are useful to characterize pathology. Indeed, the quantified anomalies using these references were not only easy to interpret (see Fig. 2 (a)) but allowed identifying sub-phenotypes of pathology with state-of-the-art accuracy [4, 12]. Finally, our framework was able to generate craniofacial surfaces that resembled the growth patterns observed in the population conditioned on age and sex.

To our knowledge, this is the first framework to correct anomalies from anatomical surface references without requiring normative data, infer population-wide anatomical distributions independent of age and sex, and generate age- and sex-specific normative synthetic datasets from cross-sectional data of patients with pathology without manual annotations or extensive pre-processing. While validation with independent healthy cohorts is the next logical step, our results provide strong evidence given the scarcity of pediatric normative data. Future work also includes the study of the generalization of our framework to other anatomical structures.

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

**Data Statement.** This study used a retrospective dataset of 3D craniofacial photograms of pediatric patients, approved by the University of Colorado Anschutz Medical Campus Institutional Review Board (protocol #20-1563). All patient surfaces displayed in the paper are either de-identified or synthetically generated. The dataset is not publicly available owing to privacy and ethical restrictions on identifiable clinical data.

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

## References

1. Al-Omari, I., et al.: Methods of assessment of cleft-related facial deformity: a review. Cleft Palate Craniofac. J. **42**(2), 145–156 (2005). https://doi.org/10.1597/02-149.1

2. Albert, A., Wright, C.: The progression of craniofacial growth and development: An anthropological study applicable to the forensic and identity sciences. Global Journal of Anthropology Research **3**(1), 16–21 (2016)
3. Cruz-Guerrero, I., et al.: Unsupervised adaptive sampling graph autoencoder for 3D surface encoding and mesh representation transfer. Medical Image Analysis **114**, 104228 (2026). https://doi.org/10.1016/j.media.2026.104228
4. Elkhill, C., et al.: Geometric learning and statistical modeling for surgical outcomes evaluation in craniosynostosis using 3D photogrammetry. Comput. Methods Programs Biomed. **240**, 107689 (2023). https://doi.org/10.1016/j.cmpb.2023.107689
5. Ganin, Y., Lempitsky, V.: Unsupervised domain adaptation by backpropagation (2015). https://doi.org/10.48550/arXiv.1409.7495
6. Hammond, P., et al.: Discriminating power of localized three-dimensional facial morphology. Am. J. Hum. Genet. **77**(6), 999–1010 (2005). https://doi.org/10.1086/498396
7. Kouchi, M., Mochimaru, M.: Errors in landmarking and the evaluation of the accuracy of traditional and 3D anthropometry. Appl. Ergon. **42**(3), 518–527 (2011). https://doi.org/10.1016/j.apergo.2010.09.011
8. Liang, C., et al.: Normal human craniofacial growth and development from 0 to 4 years. Sci. Rep. **13**(1), 9641 (2023). https://doi.org/10.1038/s41598-023-36646-8
9. Liu, J., et al.: Data-driven normative reference of pediatric cranial bone development. Plast. Reconstr. Surg. Glob. Open **10**(8), e4457 (2022). https://doi.org/10.1097/GOX.0000000000004457
10. Liu, J., et al.: Population-driven synthesis of personalized cranial development from cross-sectional pediatric CT images. IEEE Trans. Biomed. Eng. **72**(9), 2732–2741 (2025). https://doi.org/10.1109/TBME.2025.3550842
11. Morales, A., et al.: BabyFM: Towards accurate 3D baby facial models using spectral decomposition and asymmetry swapping. Computers in Biology and Medicine **186**, 109652 (2025). https://doi.org/10.1016/j.compbiomed.2025.109652
12. Porras, A., et al.: Quantification of head shape from three-dimensional photography for presurgical and postsurgical evaluation of craniosynostosis. Plast. Reconstr. Surg. **144**(6), 1051e–1060e (2019). https://doi.org/10.1097/PRS.0000000000006260
13. Roth, D., et al.: Craniofacial development: Neural crest in molecular embryology. Head Neck Pathol. **15**(1), 1–15 (2021). https://doi.org/10.1007/s12105-021-01301-z
14. Schnabel, T., et al.: Large-scale 3D infant face model. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2024. LNCS, vol. 15003, pp. 217–227. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-72384-1_21
15. Schnabel, T., et al.: Multi-linear 3D craniofacial infant shape model. In: Medical Image Computing and Computer Assisted Intervention – MICCAI 2025. 28th International Conference, Daejeon, South Korea, September 23–27, 2025, Proceedings, Part X. pp. 338–348. Springer, Berlin, Heidelberg (2025). https://doi.org/10.1007/978-3-032-05127-1_33
16. Winter, R.: What's in a face? Nat. Genet. **12**(2), 124–129 (1996). https://doi.org/10.1038/ng0296-124