

DELFT UNIVERSITY OF TECHNOLOGY

MSC THESIS IN BIOMEDICAL ENGINEERING

---

# Quantitative Analysis of 3D Cranial Morphology in Craniosynostosis Using Photogrammetry

---

*Author:*

Han Zhang (5782449)



# Quantitative Analysis of 3D Cranial Morphology in Craniosynostosis Using Photogrammetry

by

Han Zhang

**in partial fulfillment of the requirements for the degree of**

**Master of Science**

**in Biomedical Engineering**

**Track Medical Device**

**at the Delft University of Technology,**

**to be defended publicly on Monday August 26, 2024 at 10.45 AM.**

<b>Thesis Committee:</b>	Dr. F.M. Vos, chair	TU Delft
	Dr. Gennady Roshchupkin	Erasmus MC
	Dr. Tareq Abdel Alim	Erasmus MC

# Quantitative Analysis of 3D Cranial Morphology in Craniosynostosis Using Photogrammetry

## Abstract

Subjective assessment of craniosynostosis severity often leads to inconsistent treatment outcomes. This study introduces and evaluates innovative quantitative methods, including distance-based, spectral, shape descriptor, and deep learning view-based approaches, to objectively assess the severity of sagittal craniosynostosis. The effectiveness of these quantitative scores is determined by their correlation with expert clinical ratings, with the goal of improving the consistency and accuracy of severity evaluations. Furthermore, we applied the methods developed with synthetic data to real clinical datasets, focusing on pre- and post-surgery severity assessments. This study showcases how our quantitative approach effectively guides and optimizes surgical treatments, underscoring its potential utility in clinical practice.

## 1. Introduction

### 1.1. Background

Craniosynostosis is a congenital condition characterized by the early fusion of one or more cranial sutures, the narrow seams of undifferentiated mesenchyme between the bones of the skull [32]. This premature closure results in an abnormal head shape due to restricted growth perpendicular to the fused suture and compensatory overgrowth at the non-fused sutures [27]. Various sutures can be affected, including the sagittal, coronal, metopic, and lambdoid sutures, leading to distinct head shapes: scaphocephaly (long and narrow head), trigonocephaly (triangular forehead), brachycephaly (broad and flattened head), and plagiocephaly (skewed head) [18], the classification of craniosynostosis is shown in Figure 1. Scaphocephaly, also known as sagittal craniosynostosis, is the most prevalent subtype among all craniosynostosis cases, accounting for 44% of instances [20]. This condition occurs in approximately one in every 2100 live births [33, 70], often manifesting as cranial malformation within the first year of life.

Craniosynostosis occurs across all racial groups, with over 85% of cases being non-syndromic, an isolated

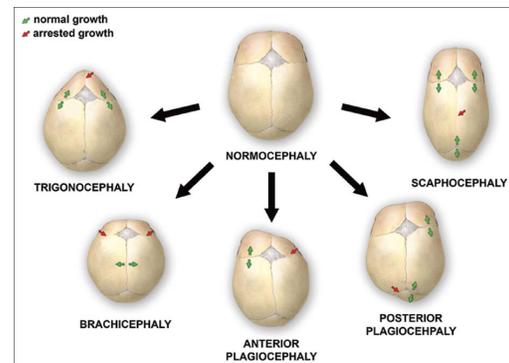


Figure 1. Classification of craniosynostosis. Adapted from [34].

anomaly [52]. Non-syndromic craniosynostosis is thought to have a strong genetic component, potentially involving gene-gene or gene-environment interactions that have yet to be fully identified [15]. This study focuses on the most prevalent subtype, non-syndromic sagittal synostosis.

Craniosynostosis can cause elevated intracranial pressure (ICP) [19], leading to headaches, irritability, vomiting, and brain damage in severe cases, which can result in cognitive and developmental delays [19]. The abnormal skull shape can also cause vision problems such as strabismus, optic nerve compression, and exposure keratitis [35], as well as conductive hearing loss due to structural abnormalities in the ear. Additionally, abnormal facial and skull development can lead to obstructive sleep apnea and other breathing difficulties, necessitating interventions like CPAP support or surgery [39].

3D photography, or 3D photogrammetry, captures the three-dimensional shape of objects by taking multiple 2D photos from different angles and creating a detailed 3D model with specialized software [1]. This technology is used in medicine, engineering, architecture, and entertainment. In medical diagnostics, it has several advantages: it does not use harmful radiation, making it safer for repeated use in children [16,28,49], and usually doesn't require sedation, which is helpful for young patients. It reduces human error, provides accurate, consistent measurements for tracking growth and surgical results [40], and creates detailed

images of cranial and facial structures. It measures parameters like volume and asymmetry and is quick and convenient for both patients and clinicians [4]. Portable 3D scanners offer flexibility in various clinical settings [38], and techniques like heat maps and curvature analyses provide valuable insights for clinical and research purposes [4,41]. The visualization of curvature analysis is shown in Figure 2.

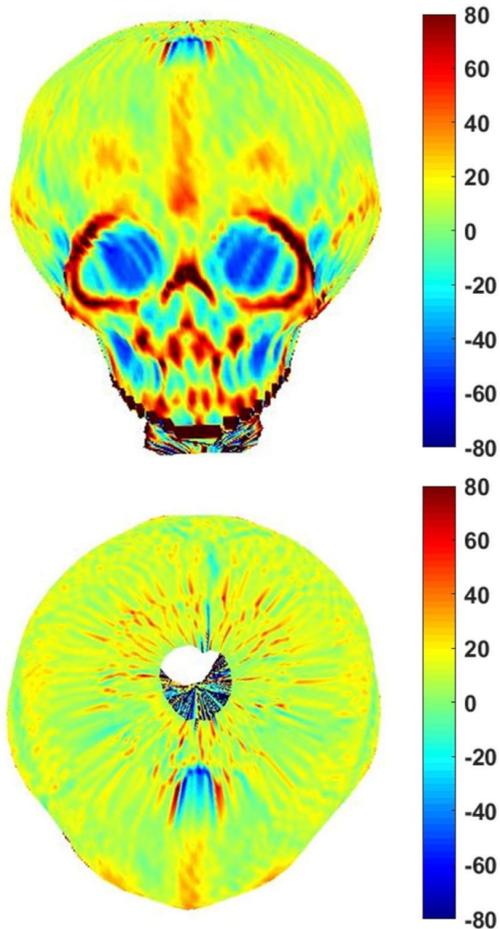


Figure 2. The curvature analysis illustrated by the heatmap. Adapted from [20].

Diagnosing craniosynostosis is generally straightforward, but assessing its severity is more challenging. While clinicians can typically distinguish craniosynostosis from conditions like positional plagiocephaly and reassure parents about its self-limiting nature [14], accurately gauging its severity is crucial for determining appropriate treatment [14]. Historically, severity assessments have been subjective, relying on visual evaluations and 2D photo scores, leading to varied treatment outcomes and inconsistent post-operative evaluations [76]. There is a pressing need for objective metrics to capture the intricate details of cranial abnormalities in three dimensions, aiding in more effective evaluation and clinical intervention [25]. Variability

in expert assessments arises from personal interpretations of craniosynostosis characteristics [11], differences in training and experience [2], and a reliance on visual inspection rather than objective measures [77]. Integrating objective, three-dimensional metrics into clinical practice is essential for achieving consistent, reliable assessments of craniosynostosis severity [26].

Assessing the severity of craniosynostosis is essential due to the risks of surgery. Mild cases might not improve quality of life enough to justify the potential harm and complications of surgery, such as brain injury from invasive ICP monitoring or major reconstruction complications like infection and bleeding [74]. Some patients may even develop new symptoms or see a decline in their condition or cognitive outcome after surgery. The long-term effects on cognitive development and overall health are not fully known, and unnecessary surgeries could lead to unexpected negative outcomes.

Recent research has focused on developing improved metrics for quantifying craniosynostosis severity beyond traditional methods like head circumference, cephalic index, and intracranial volume, which have significant limitations [12,65,70]. For example, the cephalic index often fails to distinguish between different types of craniosynostosis, and intracranial volume measurements can fall within normal ranges despite the presence of cranial abnormalities. To address these issues, more complex anthropometric methods have been introduced, such as the Interfrontal Angle Method, which uses landmark detection to measure various cranial angles [16]. Other approaches include using the interparietal to intercoronal distance ratio, recording ray lengths from key anatomical points, and assessing angles related to trigonocephaly severity [1,9,40,66]. Curvature analysis and statistical modeling techniques, such as Principal Component Analysis (PCA) and Statistical Shape Modeling (SSM), have also been applied to quantify cranial shape and deformation more accurately [10,25,28,38]. Methods like the Head Shape Anomaly(HSA) index and various asymmetry measures, including (Anterior Cranial Asymmetry Index(ACAI) and Posterior Cranial Asymmetry Index(PCAI), provide additional ways to assess severity based on anatomical and statistical data [2,25,50]. These advancements aim to offer a more objective and detailed understanding of cranial dysmorphologies, improving diagnosis and treatment planning.

## 1.2. Problem Statement and Overview

This study explores novel quantitative methods for generating descriptors and assessing the severity of sagittal craniosynostosis (scaphocephaly), primarily through dissimilarity measures compared to a control skull (Figure 3). These methods include distance-related techniques, spectral methods, the Zernike shape descriptor, and view-based deep

learning techniques. By expanding the application of these descriptors, the research aims to enhance the accuracy and consistency of severity assessments in clinical practice.

The evaluation involves correlation tests between expert ratings (ranging from 0 to 3, which are categorical variables) and the resulting severity scores to determine the effectiveness of these quantitative methods. The Spearman Rank Correlation Coefficient is employed to assess the monotonic relationship between the computed scores and the ordinal clinical severity scores. A significance level of less than 0.05 is considered significant for all statistical analyses. Additionally, if box plots and swarm plots show that the severity levels ranging from 0 to 3 have distinct and separate scores, this indicates the most ideal situation. In other words, clear separation in scores across these severity levels would suggest that the analysis or method is effectively differentiating between different degrees of severity, achieving the best possible discrimination among the categories.

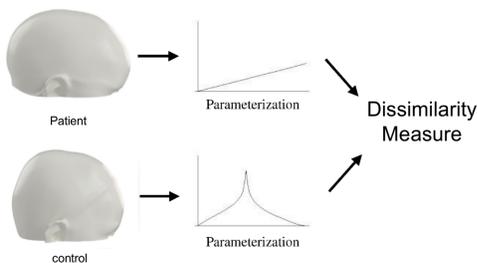


Figure 3. Dissimilarity from the control skull as the measure of severity

## 2. Dataset and Data Reprocessing

### 2.1. Description of Craniosynostosis Data

The synthetic 3D head models utilized for this project were derived from a publicly accessible craniosynostosis dataset [3]. This dataset was created using a statistical shape model specifically tailored to children under 1.5 years of age, with detailed age distribution information provided in the original publication [64]. The dataset comprises a total of 200 head models, equally divided into control (n=100) and sagittal craniosynostosis (n=100) categories, and is available in .ply file format.

### 2.2. Point correspondence Establishment

During pre-processing, the data were clipped along the nasion tragus plane and a template mesh was deformed to match the shape of each 3D surface using a variant of the non-rigid iterative closest point algorithm [7]. This process preserves the topology of the template and guarantees con-

sistent point correspondences throughout the entire collection of 3D data samples [3].

### 2.3. Expert Scoring for Synthetic Data

Expert ratings are correlated with computed severity scores using Spearman Rank Correlation to test the effectiveness of quantitative methods, with a significance level of less than 0.05. Two craniofacial experts, Irene M. J. Mathijssen and Marie-Lise C. van Veelen, evaluated the overall phenotypic severity using a randomized subset of 60 synthetic head models, including 30 that were scaphocephalic [3]. After eliminating inconsistent ratings, our research specifically focuses on these 28 cases of scaphocephaly. They employed a 4-point scale, where '0' indicates normal, '1' stands for mild, '2' signifies moderate, and '3' denotes severe, following published craniosynostosis photoscore guidelines (Gaillard, 2023a, 2023b). Head shapes with discrepancies in scores between the two experts were excluded from the correlation analysis.

### 2.4. Normative Skull Selection and Creation

In past research, it has been common to use normal patients' skulls as a control group. However, this study determines the severity score by comparing the differences between the skulls of normal individuals and those in the control group. There are two limitations to using normal patients' skulls as a control group. First, existing methods that quantify severity, such as those relying on intracranial volume and skull size, are influenced by skull size, which then affects the severity score results.

Normal patients' skulls may contain moderate flatness and some degree of asymmetry, which are critical geometric features for assessing severity. If we select any random normal patient as the control group, we risk losing some of these areas of interest, leading to biased severity scores.

Inspired by previous research, a triaxial ellipsoid has been described as the ideal shape to represent a cranium [10]. The ellipsoid fitting method is valuable because it captures abnormal flatness, which is crucial for distinguishing between typical and atypical cranial shapes. This method employs the difference between the fitted ellipsoid and the skull to identify flattened and rounded areas, utilizing the 'non-ellipsoid' and ellipsoid regions for analysis. We used an ellipsoid fitting approach to determine the minimum-volume ellipsoid that encompasses all the given points, which then serves as the normative data for each specific skull. By using this personalized ellipsoid, we can better capture the area of interest and eliminate the influence of the ellipsoid on our analysis.

However, ellipsoid fitting can result in the loss of significant rendering information, which is crucial in view-based methods. In most of our experiments, we focus solely on the positional information of individual points for each skull,

neglecting depth information from rendering and connectivity data from triangulation techniques in computer graphics. For view-based methods, multi-view images generated from the 3D mesh represent depth and roundness in the 2D projected image [78], and ellipsoid fitting can compromise this information. Therefore, for view-based methods, we use a skull from the control group as the normative reference to preserve this crucial information.

## 2.5. Implementation

The ellipsoid fitting algorithm fits an ellipsoid to a set of 3D points by determining the ellipsoid's parameters that best represent the points' distribution using Linear Least Squares (LLSQ). The algorithm begins by extracting the  $x$ ,  $y$ , and  $z$  coordinates from the 3D points. Given the 3D points from the test skull, we can determine the coordinates of the origin. By setting the centroid of these points as the new origin in spherical coordinates, we can simplify the problem of fitting an ellipsoid. Assuming the parameters of the ellipsoid are defined in spherical coordinates, the radial distance between any known point  $A$  on the test skull and the centroid  $O$  (the new origin) can be easily computed. Consider the line passing through  $A$  and  $O$ . Let  $B$  be the point where this line intersects the ellipsoid described by the parameters shown in Figure 4. The distance  $AB$  is then considered the radial distance. By applying this approach to all points on the test skull, we can obtain a set of such radial distances. The best-fitting ellipsoid is determined when the sum of all these radial distances is minimized. This method effectively transforms the 3D surface fitting problem into a 1D fitting problem. Linear Least Squares can then be used to find the closed-form solution for this optimization problem.

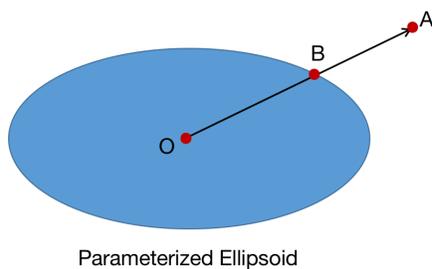


Figure 4. Dissimilarity from the control skull as the measure of severity

After establishment of the above process, the final ellipsoid parameters are extracted, including the center, radii, and semi-axes direction. The center represents the centroid of the ellipsoid, indicating the average location of the data points. The radii describe the lengths of the semi-axes of the

ellipsoid, while semi-axes direction representing its extent in each principal direction. For ease of use and comparison with synthetic skulls, the ellipsoid's spherical coordinates are generated from this parameterized expression and transformed into Cartesian coordinates. Finally, the ellipsoid is plotted in a three-dimensional space, showing its alignment and extent relative to the data points.

The Figure 5 below show a visualization of a test skull `sagittal_06`(grey mesh with shading) in dataset alongside its personalized normative skull(black discrete points).

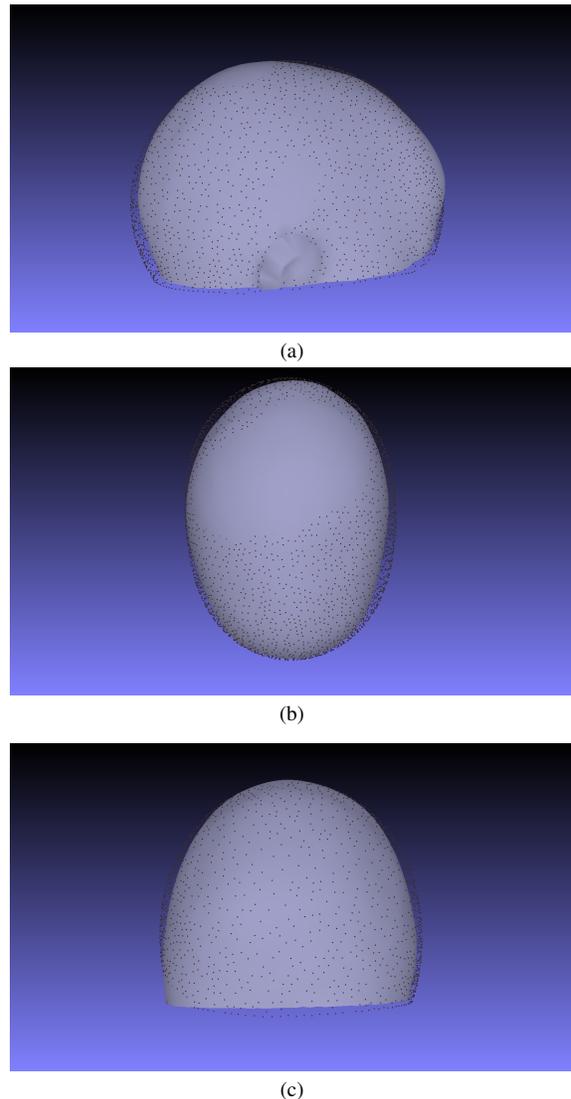


Figure 5. The visualization of the test skull and its fitted ellipsoid shape (discrete points) from three directions.

## 2.6. Quantification Function of the Difference Between Two Descriptors

Most of the methods discussed in this paper require a control group. The severity score is obtained by measuring the difference between the descriptor of the test shape and the normative shape, which necessitates a function to quantify the differences between the test skull and the normative skull. In this paper, most of the descriptors are in a vector-like form, with a few being scalar.

For vector-like descriptors, four loss functions are used to quantify the difference between the descriptor of the normative skull and the test skull. These loss functions are the Mean Absolute Error (MAE), the Mean Squared Error (MSE), Eigenvalue Decomposition Distance (EVD), and Jensen-Shannon Divergence.

The Mean Absolute Error (MAE) is the average of the absolute differences between the components of the vector:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |x_i - y_i| \quad (1)$$

The Mean Squared Error (MSE) is the average of the squared differences between the components of the vector:

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (x_i - y_i)^2 \quad (2)$$

Eigenvalue Decomposition Distance (EVD) measures the difference between two distributions based on their eigenvalues. The formula for EVD when the inputs are the eigenvalue spectra is:

$$\text{Dist}_{\text{EVD}}(P, Q) = \frac{1}{2} \sum_{i=1}^N \frac{\left[ |\lambda_i^P|^{\frac{1}{2}} - |\lambda_i^Q|^{\frac{1}{2}} \right]^2}{|\lambda_i^P|^{\frac{1}{2}} + |\lambda_i^Q|^{\frac{1}{2}}} \quad (3)$$

where:  $\lambda_i^P$  and  $\lambda_i^Q$  are the  $i$ -th eigenvalues of the covariance matrices of distributions  $P$  and  $Q$ . The sum is taken over the first  $N$  eigenvalues.

If the inputs are not eigenvalue spectra, the generic formula for Eigenvalue Decomposition Distance (EVD) using descriptors  $f$  and  $g$  is:

$$\text{Dist}_{\text{EVD}}(f, g) = \frac{1}{2} \sum_{i=1}^N \frac{\left[ |f_i|^{\frac{1}{2}} - |g_i|^{\frac{1}{2}} \right]^2}{|f_i|^{\frac{1}{2}} + |g_i|^{\frac{1}{2}}} \quad (4)$$

where:  $f_i$  and  $g_i$  are the  $i$ -th entries of the vectors  $f$  and  $g$ .  $N$  is the length of the input vectors  $f$  and  $g$ .

The requirement for using this function is that the entries of  $f$  and  $g$  must be non-negative.

Jensen-Shannon Divergence (JSD) is used to measure the similarity between two probability distributions. It is

based on the Kullback-Leibler Divergence (KLD) and is symmetric and finite. The formula for JSD is:

$$\text{JSD}(P \parallel Q) = \frac{1}{2} D_{\text{KL}}(P \parallel M) + \frac{1}{2} D_{\text{KL}}(Q \parallel M) \quad (5)$$

where:  $P$  and  $Q$  are the two probability distributions.  $M = \frac{1}{2}(P + Q)$  is the average of  $P$  and  $Q$ .  $D_{\text{KL}}$  is the Kullback-Leibler Divergence:

$$D_{\text{KL}}(P \parallel Q) = \sum_i P(i) \log \frac{P(i)}{Q(i)} \quad (6)$$

The JSD combines the KLD of each distribution with respect to their average distribution.

When the inputs are two scalars, the forms and formulas for the Mean Absolute Error (MAE) and Mean Squared Error (MSE) are as follows:

The Mean Absolute Error (MAE) between two scalars  $a$  and  $b$  is given by the absolute difference between the two values:

$$\text{MAE}(a, b) = |a - b| \quad (7)$$

The Mean Squared Error (MSE) between two scalars  $a$  and  $b$  is the average of the squared difference between the two values. Since we are dealing with only two values, it simplifies to:

$$\text{MSE}(a, b) = \frac{1}{2} \cdot (a - b)^2 = (a - b)^2 \quad (8)$$

Regarding the Eigenvalue Decomposition Distance and Jensen-Shannon Divergence, these methods are not applicable to scalar values. They are designed for comparing matrices or probability distributions, respectively, and do not have meaningful formulations for scalar inputs.

## 3. Methodology

### 3.1. Mean Facial Asymmetry Index

Traditional methods for assessing facial asymmetry, such as the asymmetry index outlined by Huang, Liu, and Chen [66], typically measure the variation in distances between landmarks on the left and right sides of the face relative to a reference plane. Their study focused on specific facial landmarks, including the glabella, nasion, pronasale, subnasale, labial superius, stomion, labial inferius, menton, exocanthion, endocanthion, alar curvature, and cheilion.

These methods require solely on landmark point correspondence. Our dataset not only ensures the correspondence of these landmark points but also allows us to explore and utilize comprehensive positional information beyond just key points. The mean facial asymmetry (MFA) index is the mean distance between each vertex on one half

of the face to its corresponding vertex on the other half (using a mirrored reflection). It enables a more detailed capture of shape information and provides a more informative quantification of shape differences.

By using mesh registration for facial analysis, we can calculate the mean distance between each vertex on one half of the face and its corresponding vertex on the other half (via mirrored reflection). Additionally, the mean facial asymmetry (MFA) index provides a quantitative measure of the overall asymmetry observed in the face.

This method does not depend on the presence of a control group for comparison. It is designed to be invariant to both rotation and translation, meaning it can accurately assess facial asymmetry regardless of the orientation or position of the face. This invariance ensures that the measurements are consistent and reliable, even if the face is rotated or shifted, providing a robust analysis of facial shape and symmetry.

The visualization using a heatmap is implemented by CraniumPy [5], a software developed for the registration and craniofacial analysis of 3D images. This produces a heatmap (in mm) that highlights areas of varying symmetry in Figure 6.

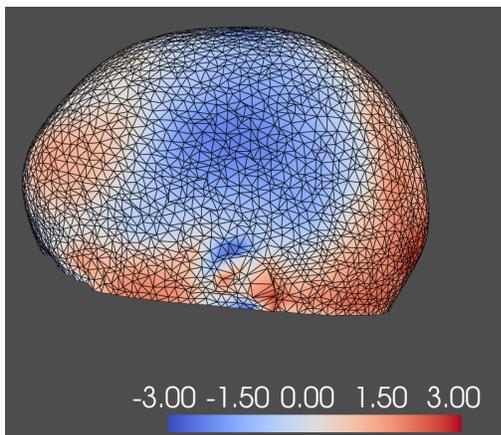


Figure 6. Visualization of Asymmetry Compared to the Other Hemisphere

### 3.2. Difference Between Corresponding Coordinates

Coordinates are the most direct descriptors for shape analysis. Our dataset includes precise point correspondences within each test skull and between each test skull and the normative skull.

We use the coordinates of corresponding points between each test skull and the fitted ellipsoid normative skull as shape descriptors. Arrays generated from these correlated coordinates are directly used as shape descriptors. Quantification functions such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and generalized eigenvalue

distance are employed to quantify the differences between these arrays.

As previously discussed, the Jensen-Shannon Divergence (JSD) is primarily used to measure the similarity between two probability distributions or histograms. Since the coordinate arrays cannot be interpreted or generalized in this manner, we do not use JSD to quantify the differences between the shape descriptors of the test skull and the normative skull.

### 3.3. D2 Shape Descriptor

This method represents a geometric histogram-based shape descriptor. This method combines SSM with distance-based metrics. To be more specific, SSM maps shapes into histograms, and distance is crucial because it provides a visually intuitive metric for surgeons. Unlike many existing methods that focus on distance of specific directions or certain distance between two landmarks, like Frontal bossing index/cite51, our approach uses the D2 shape descriptor to integrate all distance information, rather than relying on a single direction of data.

The geometric histogram-based shape descriptor method involves representing the shape signature of a 3D model as a probability distribution sampled from a shape function, which measures geometric or other properties of the model. This approach generalizes geometric histograms into what is known as a shape distribution. The key idea is to transform an arbitrary 3D model into a parameterized function that can be easily compared with others, simplifying shape analysis to the sampling, normalization, and comparison of probability distributions [59]. The shape descriptor captures global geometric characteristics, such as distance, angle, area, and volume measurements, across random points on the surface.

D2 is a distribution function that measures the distance between two random points on the surface, as shown in Figure 7. The method involves mapping 3D surfaces to a common parameterization through random sampling of shape functions [58]. As illustrated in Figure 8, 3D models like a ball and a cup are sampled to transform their complex geometries into a more manageable format. The shape functions then generate D2 shape distributions, which are represented as graphs plotting probability against distance. These distributions probabilistically capture the geometric properties of the 3D models. A similarity measure is applied to compare these distributions, determining the degree of similarity or difference between the shapes [59].

The approach allows for standardized comparison of different 3D shapes using their D2 shape distributions, simplifying shape analysis and comparison. This method effectively quantifies similarity between complex 3D models. A good shape descriptor should be concise, quick to compute, and invariant to various transformations. The D2

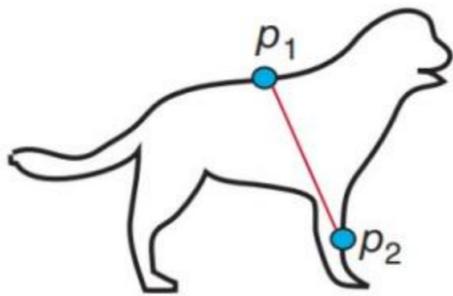


Figure 7. Illustration of D2 Descriptor. Adapted from [42]

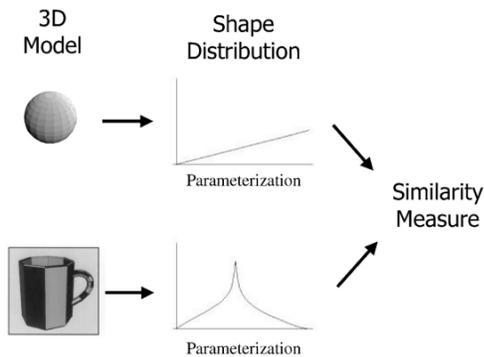


Figure 8. Illustration of Similarity Measure by Shape Histogram. Adapted from [59]

shape descriptor maintains its characteristics regardless of the shape’s position, orientation, mirroring, or scale, ensuring accurate and efficient shape analysis in applications like computer vision and 3D modeling.

Compared to the mean corresponding distance, which measures the distance between corresponding points, our method provides end-to-end generation for both the test skull and the normative skull. It captures the intrinsic relationships and positional information between each point within the skull, offering a more comprehensive analysis.

### 3.3.1 Implementation

To generate a D2 shape descriptor histogram, follow these steps:

1. Downsample the Shape: Reduce the shape to 499 points.
2. Calculate Pair-Wise Distances: Compute the distances between each pair of these 499 points, resulting in  $\frac{499 \times 498}{2}$  (or 124,751) distance values.
3. Create the Histogram: Using 17 predefined bins, create a histogram to represent the shape based on these distances.

4. Result: This histogram of each skull serves as shape descriptor.

The visualization of the D2 histogram of a test skull sagittal\_06 is shown in Figure 9.

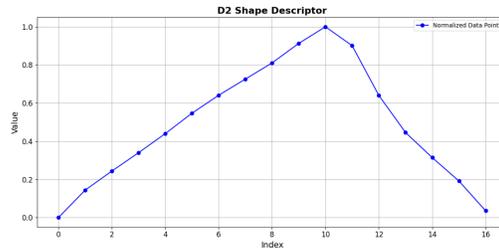


Figure 9. Visualization of D2 Histogram

### 3.4. Spectral Embedding from Affinity Matrix

The size-independent nature of spectral analysis on shapes makes it an attractive choice, leading to its recent potential utilization in craniosynostosis research. The paper describes a method for constructing spectral embeddings from 3D triangle meshes to achieve invariance to bending and rigid-body transformations [29].

Each shape is represented by the eigenvectors of an appropriately defined affinity matrix, forming a spectral embedding [30] that is invariant against rigid-body transformations, uniform scaling, and shape articulation (i.e., bending).

To be more specific, given a 3D triangle mesh with  $n$  vertices, we form an  $n \times n$  affinity matrix  $A$  such that the  $ij$ -th entry of  $A$  is the affinity between the  $i$ -th and the  $j$ -th mesh vertices; several possible choices for the affinities are discussed in subsequent sections. Matrix  $A$  is then eigen-decomposed as  $A = V\Lambda V^T$ , where  $\Lambda$  is a diagonal matrix with eigenvalues  $\lambda_1 \geq \dots \geq \lambda_n$  along the diagonal and  $V = [\mathbf{v}_1 | \dots | \mathbf{v}_n]$  is an  $n \times n$  matrix with  $\mathbf{v}_1, \dots, \mathbf{v}_n$  as the corresponding eigenvectors.

The eigenvectors are scaled by the square root of their corresponding eigenvalues to normalize variations due to mesh size. The first  $k$  scaled eigenvectors form a  $k$ -dimensional spectral embedding, which is represented below:

$$\hat{V}_k = [\hat{v}_1 | \hat{v}_2 | \dots | \hat{v}_k] \quad (9)$$

where

$$\hat{v}_i = v_i \sqrt{\lambda_i}. \quad (10)$$

While the eigenvectors of the affinity matrix form a spectral embedding that provides a normalized representation of the shape, the eigenvalues indicate the variation of the shape along the axes defined by the corresponding eigenvectors [31]. Therefore, we consider eigenvalues as spectral shape descriptors. However, eigenvalues are influenced by

mesh sizes, and typical shape databases contain shapes with different numbers of vertices. To address this, the eigenvectors are scaled by the square root of their corresponding eigenvalues to normalize variations due to mesh size. Thus, the eigenvalues of the original affinity matrices cannot be used for direct shape comparison.

For our research, with the same number of samples taken from each shape, the eigenvalues of the sampled affinity matrices become comparable. Finally, the output spectrum, consisting of a list of eigenvalues, is normalized to produce comparable descriptors.

### 3.4.1 Implementation

For the Affinity Matrix generation, we use Gaussian affinities: with a robust method for estimating geodesic distances, we define the affinity matrix  $A$  as  $A_{ij} = \exp(-d_{ij}^2/2\sigma^2)$ , where  $d_{ij}$  is the distance between the  $i$ th and  $j$ th vertices of the mesh, and  $\sigma$  is the Gaussian width [8]. This approach shows that the affinity between two vertices is inversely related to the Euclidean distance between them.

To implement the spectral shape descriptor, known as EVD (Eigenvalue Decomposition Distance), we use the eigenvalues of a  $500 \times 500$  sampled affinity matrix for each shape, with 500 sample points selected experimentally. The resulting spectrum is normalized to the range  $[0, 1]$ . Upon examining all output embeddings, it was observed that eigenvalues ranked beyond the 20th largest are all less than 0.0001. Therefore, we use only the 20th largest eigenvalue as our shape descriptor and visualization is given in Figure 10.

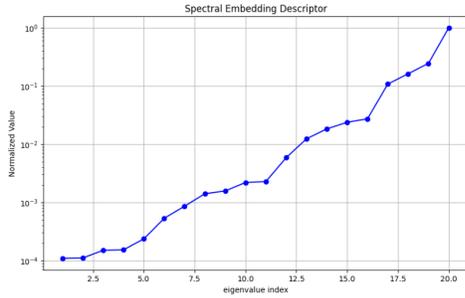


Figure 10. Visualization of spectral embedding descriptor

## 3.5. LBO Spectrum

The Laplace–Beltrami spectrum can be regarded as the set of squared frequencies [61] (the so-called natural or resonant frequencies) that are associated with the eigenmodes of a generalized oscillating membrane defined on the manifold [61].

Let  $f$  be a real-valued function, with  $f \in C^2$ , defined on a Riemannian manifold  $M$  (a differentiable manifold with

a Riemannian metric) [53]. The Laplace–Beltrami operator  $\Delta$  is defined as:

$$\Delta f = \nabla^i \nabla_i f \quad (11)$$

where  $\nabla$  is the Levi-Civita connection associated with the Riemannian metric, and the repeated index  $i$  indicates a summation over all coordinates (Einstein summation convention).

Regarding the original mathematical expressions, the following three forms are also commonly found in papers:

$$\Delta f = \text{div}(\nabla f) \quad (12)$$

$$\Delta f = \nabla \cdot \nabla f \quad (13)$$

$$\Delta f = \text{div}(\text{grad } f) \quad (14)$$

From the above formula, the Laplace–Beltrami operator is a linear differential operator that can be calculated in local coordinates. Unlike the divergence and gradient operators we encounter in Euclidean space, the Laplace–Beltrami operator is defined on a manifold. It can be understood as a generalization of the differential operators in Euclidean space to curved surfaces.

### 3.5.1 Matrix Representation of LBO and Relationship between Eigenvalues and Shape

Local parameterization is defined for a manifold as below:

$$\psi : \mathbb{R}^n \rightarrow \mathbb{R}^{n+k} \quad (15)$$

Defining the intermediate matrix  $G$  and variable  $W$  [79]:

$$g_{ij} = \langle \partial_i \psi, \partial_j \psi \rangle \quad (16)$$

$$G = (g_{ij}) \quad (17)$$

$$W = \sqrt{\det G} \quad (18)$$

$$(g^{ij}) = G^{-1} \quad (19)$$

The Laplacian Beltrami Operator can be written as:

$$\Delta f = \frac{1}{W} \sum_{i,j} \partial_i (g^{ij} W \partial_j f) \quad (20)$$

This can be also written as matrix representation [79]:

$$\Delta f = \nabla \cdot \nabla f = \frac{1}{\sqrt{G}} \sum_{i,j=1}^n g^{ij} \frac{\partial}{\partial x^i} \left( \sqrt{G} g^{ij} \frac{\partial f}{\partial x^j} \right) \quad (21)$$

where  $i, j = 1, \dots, n$ ,  $n$  is the number of vertices in the mesh.

Observing above formula, if the mesh  $M$  degraded to Euclidean plane, the LBO will degrade to well-known Laplacian operator.

$$\Delta f = \frac{\partial^2 f}{(\partial x)^2} + \frac{\partial^2 f}{(\partial y)^2} \quad (22)$$

Recalling the Wavefield theory, considering the wave equation, which gives the differential relationship between the spatial component and temporal component within the wave propagation.

$$\left( \nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) u(\mathbf{r}, t) = 0 \quad (23)$$

where the  $u$  is the wave, like electromagnetic radiation, seismology, and acoustics. Separating the temporal and spatial part of above equation:

$$u(\mathbf{r}, t) = A(\mathbf{r})T(t) \quad (24)$$

Based on above motivation, substituting this separated form into the wave equation and then simplifying, the following equation is obtained.

$$\frac{\nabla^2 A}{A} = \frac{1}{c^2 T} \frac{d^2 T}{dt^2} \quad (25)$$

This is one form of Helmholtz equation. The solution of Helmholtz equation is given:

$$\nabla^2 A = -k^2 A \quad (26)$$

The wave number  $k$  characterizes the spatial frequency of the wave. To understand this formula,  $A$  is the spatial part of the wave function, and  $f$  is derived by parameterizing the manifold, which can be analogous to a spatial wave. For the spatial component  $A$ , applying the Laplacian operator yields information related to spatial frequency. Thus, applying an LBO operation to  $f$  will also provide similar spatial frequency information. This is why LBO can be used as a shape descriptor. We refer to the output of this process as the LBO spectrum.

The Laplace–Beltrami operator is an intrinsic quantity of the surface, meaning it is invariant under isometric deformations of the manifold [13]. The spectrum, which depends only on the gradient and divergence defined by the Riemannian structure of the manifold (i.e., the intrinsic geometry), is also isometry invariant [60]. This property underscores the benefit of normalized LBO spectrum as shape descriptors: shapes can be compared regardless of the object’s scale and position [44].

### 3.5.2 Implementation

In the last part, we define the spectrum in equation(21) in the context of a manifold. To make this spectrum more applicable to meshes in practice, we need to use a discretized version. We approximate the derivative operation by using

the neighboring points of a given point as shown in Figure 11 [80].

$$\Delta f(v_i) = \frac{1}{2} \sum_{v_j \in Nei(v_i)} (\cot \alpha_{ij} + \cot \beta_{ij}) |f(v_j) - f(v_i)| \quad (27)$$

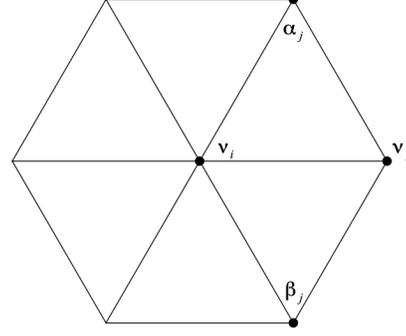


Figure 11. Implementation of LBO spectrum at each vertex

In above diagram,  $f(v_i)$  is the value of the function  $f$  at vertex  $v_i$ , in this project it’s the coordinate.  $N_i$  is the set of neighboring vertices of  $v_i$ .  $\alpha_{ij}$  and  $\beta_{ij}$  are the angles opposite to the edges connecting  $v_i$  and  $v_j$ .

The skull data has been downsampled to 500 points and retriangulated. The resulting spectrum (sagittal.06) is shown in Figure 12:

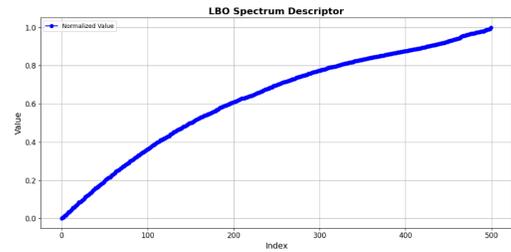


Figure 12. Visualization of LBO spectrum

## 3.6. Laplacian Heat Kernel Descriptor

### 3.6.1 Mesh Laplacian Matrix

The discrete Laplacian matrix is a fundamental tool in graph theory, representing the structure and properties of a graph. It is particularly valuable for analyzing various characteristics of the graph, such as connectivity, clustering, and spectral properties. When applied to mesh data, which can be effectively modeled as a graph, the discrete Laplacian matrix becomes instrumental in a range of mesh processing tasks. The implementation of the discrete Laplacian matrix

is mathematically similar to the discrete Laplace-Beltrami operator (LBO). Assuming the mesh is represented in Figure 13, the discrete Laplacian matrix is given by the following formula [5, 81]:

$$\mathbf{L}_{ij} = \begin{cases} w_{ij} = \frac{1}{2} (\cot \alpha_{ij} + \cot \beta_{ij}) & \text{if } j \text{ adjacent to } i \\ -\sum_{j \in \mathcal{N}(i)} w_{ij} & \text{when } j = i \\ 0 & \text{otherwise} \end{cases} \quad (28)$$

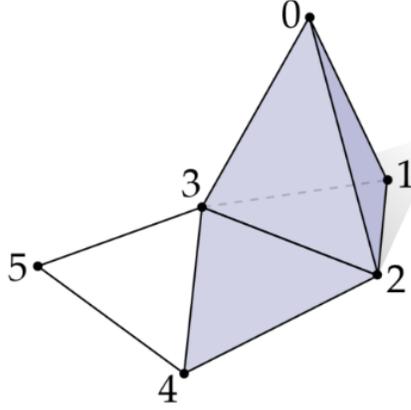


Figure 13. One example subset of mesh data. Adapted from [71]

### 3.6.2 Heat Kernel Signature Histogram

Given an initial heat distribution  $u_0(x)$  over the surface, the heat diffusion equation on a compact Riemannian manifold can be expressed as:

$$\left( \Delta - \frac{\partial}{\partial t} \right) u(x, t) = 0 \quad (29)$$

In this context,  $\Delta$  denotes the Laplace-Beltrami operator, and  $u(x, t)$  represents the heat distribution at a point  $x$  at time  $t$ . The solution can be expressed as follows:

$$u(x, t) = \int h_t(x, y) u_0(y) dy \quad (30)$$

$h_t(x, y)$  represents the heat transferred from a point  $x$  to another point  $y$  at time  $t$ . The eigen decomposition of the heat kernel is expressed as [69].

$$h_t(x, y) = \sum_{i=0}^{\infty} \exp(-\lambda_i t) \phi_i(x) \phi_i(y) \quad (31)$$

where  $\lambda_i$  and  $\phi_i$  denote the  $i$ -th eigenvalue and eigenfunction of the Mesh Laplacian matrix [69]. When restricting

the heat kernel to the temporal domain by replacing  $y$  with  $x$ , the expression simplifies to:

$$h_t(x, x) = \sum_{i=0}^{\infty} \exp(-\lambda_i t) \phi_i^2(x) \quad (32)$$

The above equation provides a local shape descriptor for a vertex, known as the Heat Kernel Signature (HKS). Once we have such a local shape descriptor, a technique similar to the D2 shape descriptor is applied to create the Heat Kernel Signature Distance Distribution (HKSDD) [79].

Since the number of vertices in our skull is finite, the eigenvalues and eigenfunctions obtained from the discretization are also finite, with a maximum count equal to the number of vertices. Let this number be denoted as  $N$ . Therefore, the above equation can be written as:

$$h_t(x, x) = \sum_{i=0}^{N-1} \exp(-\lambda_i t) \phi_i^2(x) \quad (33)$$

The distance of pairs for this Heat Kernel Signature is expressed as:

$$d_h^2(x, y) = \sum_{i=1}^N e^{-2t\lambda_i} (\phi_i(x) - \phi_i(y))^2 \quad (34)$$

### 3.6.3 Implementation

1. Compute the Laplacian matrix according to the method detailed in Section 3.6.1.
2. Perform an eigen decomposition of the Laplacian matrix to obtain its eigenvalues and eigenvectors.
3. Determine the distance of the heat kernel signature values at any two points on skull data according to in Equation 34.
4. Generate a distance histogram as outlined in Section 3.6.2, the visualization of one test skull(sagittal\_06) is shown in Figure 14.

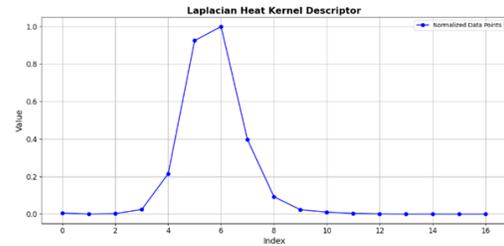


Figure 14. Visualization of Laplacian Heat Kernel Descriptor

### 3.6.4 Verification Using Reconstructiton

This procedure is primarily intended to validate the accuracy of the calculations involved in the complex computation of the Laplacian matrix, rather than being directly related to skull shape analysis.

To reconstruct a 3D mesh using a specified number of eigenvectors obtained from the Laplacian Matrix, the following method is employed:

Initially, the eigenvector matrix is computed from the eigenvalue decomposition of the Laplacian matrix. This matrix is then truncated to include only the eigenvectors corresponding to indices from 0 to 250, while the rest are discarded. Subsequently, each coordinate ( $x$ ,  $y$ ,  $z$ ) of the mesh vertices is projected using the transposed truncated eigenvector matrix, resulting in transformed coordinates. These coordinates are then used to reconstruct the original mesh vertices by multiplying the full eigenvector matrix with the transformed coordinates. The final step assembles these reconstructed coordinates to form the complete set of vertices for the reconstructed mesh. The Figure 15 shows the comparison of original and reconstructed skull shapes.



Figure 15. Comparison of original and reconstructed shapes.

Based on the visualization, the ear region appears slightly blurred; however, overall, the shapes are almost identical. This indicates that our computation of the Laplacian matrix and the capture of eigenvalues and eigenvectors are accurate.

### 3.7. 3D Zernike Shape Descriptor

We are inspired by recent research in protein database retrieval [36, 62, 72], which focuses on quantifying differences between protein shapes to identify the most similar ones. This approach parallels our goal of measuring differences between test and normative skulls. Given the subtle and complex features in protein datasets, highly discriminative shape descriptors are required, which surpass those designed for standard 3D shape classification. Therefore, we applied the high-performing 3D Zernike shape descriptor, commonly used in protein shape retrieval, to enhance our craniosynostosis research.

#### 3.7.1 Zernike Polynomials

In mathematics, the 3D Zernike polynomials are a sequence of polynomials that are orthogonal on the sphere. These polynomials are defined as below in terms of polar coordinates, specifically the azimuthal angle  $\phi$ , the radial distance  $\rho$ , and the polar angle  $\theta$ . The 3D Zernike functions are given by [57]:

$$Z_{nl}^m(r, \theta, \phi) = R_{nl}(r)Y_l^m(\theta, \phi) \quad (35)$$

where  $Y_{lm}(\theta, \phi)$  are the spherical harmonics defined in spherical coordinates  $(r, \theta, \phi)$ . The integers  $n$ ,  $l$ , and  $m$  satisfy  $|m| \leq l$  and  $n - l$  is even.  $R_{nl}(r)$  represents the orthogonal radial polynomials.

#### 3.7.2 The Orthogonality of Zernike Polynomials

According to above formula, we can verify the orthogonality within the unit sphere by below formula:

$$\frac{3}{4\pi} \int_{\|\mathbf{x}\| \leq 1} Z_{nl}^m(\mathbf{x}) \overline{Z_{n'l'}^{m'}(\mathbf{x})} d\mathbf{x} = \delta_{nn'} \delta_{ll'} \delta^{mm'} \quad (36)$$

The prerequisite for serving as a basis function to represent arbitrary functions is fulfilled by the 3D Zernike functions. Based on the above discussion, the orthogonality of Zernike polynomials  $Z_{nlm}$  ensures that they can form a complete orthonormal system [48]. This property allows for the approximation of any smooth real-valued function defined on the unit sphere using the Zernike coefficients.

#### 3.7.3 The Rotation Invariance of Zernike Polynomials

In the context of 3D Zernike functions, there exists an analogous invariance relation to that of spherical harmonics under arbitrary rotation represented by the rotation matrix  $P$ .

This means that when the Zernike functions undergo rotation by  $P$ , they exhibit a consistent relationship similar to the rotational invariance observed in spherical harmonics [54].

$$\mathbf{Z}_{nl}(\mathbf{P}\mathbf{x}) = \mathbf{o}_l(\mathbf{P})\mathbf{Z}_{nl}(\mathbf{x}) \quad (37)$$

### 3.7.4 Zernike Transform

Any sufficiently smooth real-valued phase field over the unit sphere  $G(\rho, \theta, \phi)$  can be represented in terms of its Zernike coefficients [6], much like how periodic functions have an orthogonal representation using the Fourier series [45, 55].

$$f(\mathbf{x}) = \sum_n \sum_l \sum_m \Omega_{nl}^m Z_{nl}^m(\mathbf{x}) \quad (38)$$

The function  $f$  is defined or transformed onto the unit sphere, and  $F_{nl}$  represents the Zernike moment. Given a 3D shape function  $f(x)$ , the Zernike moments are the projections of this shape function onto a set of orthogonal basis functions [46]. For an order  $n$ , Zernike moments are expressed as an integral weighted by  $f$ :

$$F_{nl}^m = \frac{3}{4\pi} \int_{\mathbb{R}^3} f(x) Z_{nl}^m(x) dx \quad (39)$$

The above derivation is similar to the Fourier Transform and its inverse transform. By analogy, Zernike moments  $F_{nl}$  are similar to the Fourier spectrum in a 1D signal, which contains frequency information. Similarly, Zernike moments also encapsulate frequency information.

### 3.7.5 Implementation of Zernike Transform

As regards the calculation the Zernike moment  $F_{nl}$ , we use an established algorithm instead of using the definition.

denotes the geometrical moment of the object. Introducing an intermediate variable  $\chi$ , which is defined based on the geometrical moment expression of Zernike Polynomials.

$$Z_{nl}^m(\mathbf{x}) = \sum_{r+s+t \leq n} \chi_{nlm}^{rst} x^r y^s z^t \quad (40)$$

Define  $M_{rst}$  as the geometrical moment of the object:

$$M_{rst} = \int_{|\mathbf{x}| \leq 1} f(\mathbf{x}) x^r y^s z^t d\mathbf{x} \quad (41)$$

Combing the above Equation 41, 40, 39, the Zernike Moment is given as:

$$F_{nl}^m = \frac{3}{4\pi} \sum_{r+s+t \leq n} \overline{\chi_{nlm}^{rst}} M_{rst} \quad (42)$$

Moving forward, our calculation will focus on  $F_{nl}$ , which will serve as our feature vector for shape description. We collect the moments  $F_{nl}$  into  $(2l + 1)$ -dimensional

vectors. After generating the 3D Zernike moments, taking the norm of these moments ensures rotation invariance. This process involves aggregating the moments into  $(2l + 1)$ -dimensional vectors, denoted as  $\mathbf{F}_{nl} = (F_{nl1}, F_{nl2}, F_{nl3}, \dots, F_{nl(2l+1)})$ .

### 3.7.6 Implementation

According to the previous research, any sufficiently smooth real-valued phase field over the unit sphere  $G(\rho, \theta, \phi)$  can be represented using its Zernike coefficients. However, our skull data is defined by a mesh, not a unit sphere. To fit an unstructured surface mesh inside a unit ball centered at the origin, we follow these steps and illustration is shown in Figure 16:

1. Translate the vertices  $A$  so that their average position aligns with the origin  $O$ .
2. Normalize the distances by calculating each vertex's distance from the newly established origin  $O$  and normalizing these distances.
3. Map the target mesh onto a function defined on a unit sphere, where the value represents the normalized distance.

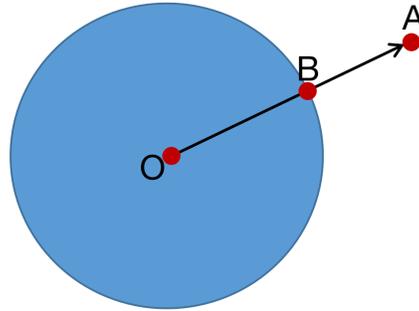


Figure 16. Illustration of mapping a 3D shape onto a unit sphere,  $B$  is the corresponding point of  $A$  under this mapping, where its value is the normalized distance  $|OA|$ .

Once we have established a compatible function defined in the unit sphere for each skull in our dataset, we calculate the Zernike moment vectors for both the test skulls and their corresponding normative skulls across orders from 1 to 30, with each vector having a length of  $2l + 1$ . The most computationally intensive part of this process is the calculation of the geometrical moments. To efficiently handle this, we employ an optimized method developed by Mateusz Banach [9], which uses a fast Python implementation

of the PK algorithm. This approach leverages Numba to optimize computationally demanding functions and loops, resulting in approximately 200 times faster performance at Zernike moment order 20 compared to traditional NumPy-based implementations.

The visualization of Zernike moment of a specified order is given in Figure 17.

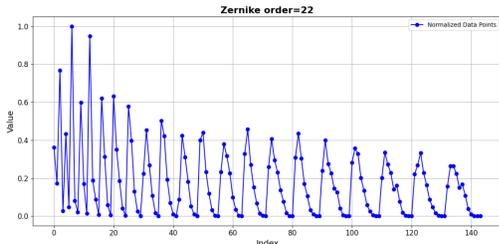


Figure 17. Visualization of Zernike moment of a specified order.

### 3.8. View-based Method Descriptor

View-based method in shape analysis is a kind of straightforward approaches to comparing 3D models is to represent them as a collection of images rendered from various viewpoints around the object, known as "viewers", shown in Figure 18. This method is motivated by several factors: first, the human visual system processes 3D objects as a series of 2D views rather than as a complete 3D representation. Second, similar 3D models will appear alike from all viewing angles. Lastly, prior to the advent of widespread 3D technology, significant progress was made in 2D image analysis, leading to efficient techniques for indexing, classification, and retrieval. By representing 3D models through 2D projections, one can capitalize on the rich body of work in 2D image analysis to advance 3D shape analysis.

View-based techniques for 3D shape analysis address the parameterization issue by projecting the 3D models onto 2D images, thereby simplifying the shape analysis task to a 2D (image) analysis problem. The fundamental concept of this series of methods revolves around the notion that two visually similar 3D models will exhibit similarity from all viewing perspectives.

This idea originates from the Light Field Descriptor (LFD) [22]. In the original paper, a set of 60 light field descriptors is proposed for a 3D model. Each descriptor is obtained using a system of 60 cameras positioned at the vertices of a regular dodecahedron's half hemisphere. Consequently, the dissimilarity between two 3D models can be measured using the  $L_1$  norm between their 2D shape descriptors.

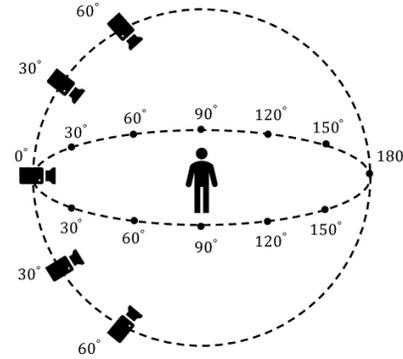


Figure 18. The illustration of basic idea of multi-view recognition

#### 3.8.1 Multi-view CNN

In original view-based method, LFD, authors don't directly use a global descriptor for 3D shape dissimilarity quantification. Instead, they find the minimum distance over 60 possible rotations of the camera system [22]. This approach effectively uses only a selected image to represent the shape for comparison with other shapes, not fully leveraging information from all projection images. Considering the advances in neural network for information fusion, a more effective and informative strategy could involve using neural networks to aggregate the information extracted from each 2D image.

Additionally, many pre-trained CNN-based networks, such as those available through the torchvision.models API in the PyTorch library, like AlexNet trained on the ImageNet dataset, can be utilized. Leveraging these advances in image descriptors and large image databases allows for the pre-training of CNN architectures. This pre-training enables the network to learn generic features for 2D image categorization, which can then be fine-tuned for specific tasks.

There is significant potential to integrate deep learning methods for 3D shape recognition. A Multi-View Convolutional Neural Network (MVCNN) has been developed for this purpose [67]. The MVCNN processes multiple 2D rendered views of a 3D shape through a unified Convolutional Neural Network (CNN). The network consists of five convolutional layers (conv1 to conv5), corresponding to the architecture of pretrained AlexNet in torch.model up to the flattening layer. This is followed by a View-Pooling Layer that applies an element-wise maximum operation across the views. Subsequently, there are three fully connected layers (fc6 to fc8), and the architecture concludes with a softmax classification layer. The architecture is illustrated in Figure 19:

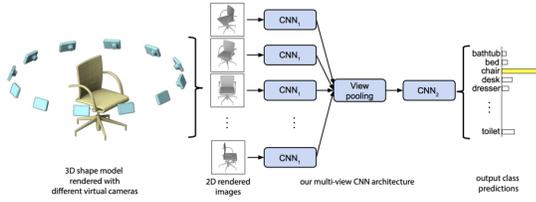


Figure 19. The architecture of MVCNN. Adapted from [68]

### 3.8.2 The Relationship between Shape Descriptor and Neural Network Classifier

During this time frame, various research efforts have highlighted the promise of Artificial Intelligence (AI) in identifying different subtypes of craniosynostosis through the analysis of 3D data [17, 23, 24, 47, 51, 63, 64]. Although these innovative studies are crucial for pushing the boundaries of the field and showcasing AI’s potential in 3D shape analysis, simply distinguishing between subtypes of craniosynostosis has limited clinical impact [3].

In neural network-related research, classification is a significant topic. The MVCNN model focuses on classification tasks, whereas our project is not about classification but on descriptors—the expression of task-related features of a shape. Despite this difference, certain aspects can be used across both areas: information flows from high-dimensional to low-dimensional spaces in neural network architectures, regardless of the complexity of the classifier structure. The intermediate layers function as structures for feature extraction and information aggregation, and their outputs can serve as effective descriptors for task-related features.

To be more specific, a Convolutional Neural Network (CNN) consists of several key components: an input layer for raw pixel values, convolutional layers for feature extraction, ReLU layers for activation, pooling layers for down-sampling, and fully-connected layers for classification.

When an input image is fed into a CNN, it passes through these various layers, ultimately producing a set of scores (one per neuron in the last layer). In image classification, these scores represent the probabilities of the image belonging to each class, such as chairs, horses, or cars. Each class is a collection of different geometric features. The goal of the training process is to learn the weights of the filters at various layers of the CNN. Often, the output of one of the layers before the last fully-connected layer can be used as a global descriptor for the input image, shown in Figure .

Thus, while MVCNN is designed for classification tasks, the intermediate layers involved in feature extraction and aggregation can be leveraged to obtain effective descriptors for shape analysis in our project. Considering the classifier’s ability to represent shapes, we can train a classifier using a standard dataset and use the output from the inter-

mediate layers of this model to represent our skull data.

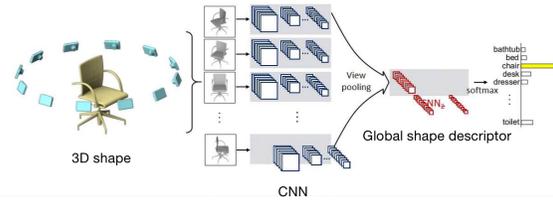


Figure 20. The relationship between global shape descriptor and the network classifier(MVCNN). Adapted from [68]

### 3.8.3 View-GCN

MVCNN exhibit limitations in 3D shape recognition tasks. One major drawback is ignoring view relationships: MVCNN’s use of max-pooling treats each view independently, neglecting the spatial and relational context between different views. Additionally, there is a loss of information: max-pooling can aggregate features in a way that discards valuable information, which could otherwise be captured by understanding the relationships among views.

View-GCN addresses these limitations by leveraging a view-graph representation and a hierarchical GCN architecture. This approach enables the effective aggregation of multi-view features while preserving and utilizing the relationships among views, leading to improved performance in 3D shape recognition tasks.

**View-Graph Construction** The View-Graph Construction process begins with defining nodes as the camera views of the 3D object, represented by coordinates  $v_i$ . Edges between these nodes are established using the k-nearest neighbor (kNN) algorithm based on the spatial relationships among views. The adjacency matrix  $S_{ij}$  is defined as:

$$S_{ij} = \Phi(g_{ij}; \theta_s) \quad (43)$$

where  $g_{ij} = [v_i, v_j, v_i - v_j, \|v_i - v_j\|^2]$  and  $\Phi$  is a non-linear embedding function.

**View-GCN Architecture** In the View-GCN Architecture, initial features  $\{f_i^0\}$  are extracted from 2D images using a pre-trained network such as ResNet-18. Local graph convolution updates node features by considering neighboring nodes:

$$F^l = \Psi(A^l F_{in}^l W^l; \theta_c^l) \quad (44)$$

where  $A^l$  is the learnable adjacency matrix,  $W^l$  is the weight matrix, and  $\Psi$  is a non-linear transformation. Non-local message passing captures long-range relations among

nodes:

$$m_{ij}^l = \Gamma([f_i^l, f_j^l]; \theta_m^l) \quad (45)$$

$$\hat{f}_i^l = \Omega([f_i^l, r_i^l]; \theta_f^l) \quad (46)$$

where  $\Gamma$  and  $\Omega$  are relation and fusion functions, respectively. Selective view-sampling for graph coarsening uses a novel strategy to sample representative views, enhancing diversity and representativeness:

$$v_j^{l+1} = \arg \max_{v_q \in N(v_j)} \left( \max(V(\hat{f}_{v_q}^l; \theta_v^{l,j})) \right) \quad (47)$$

where  $V$  is the view selector function.

A single level of View-GCN consists of local graph convolution, non-local message passing, and selective view-sampling, as shown in Figure 21.

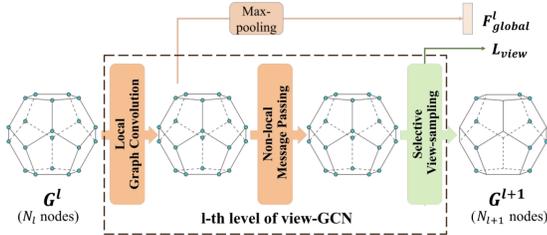


Figure 21. View-GCN layer architecture. Adapted from [75]

**Hierarchical Network Structure** The hierarchical network structure consists of multiple levels of view-GCN, with each level coarsening the graph and updating node features. Features from all levels are fused to form a comprehensive global shape descriptor like the analysis in the MVCNN in section.

As shown in Figure 22, the CNN (blue block) represents the view-graph construction and also performs multi-view image feature extraction. The view-GCN (orange block) represents graph information aggregation.

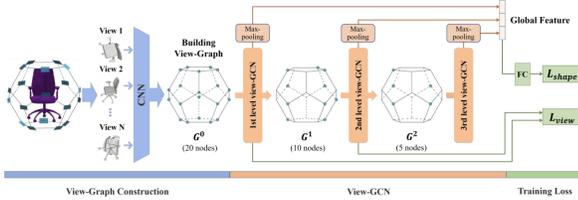


Figure 22. Architecture of entire network of View-GCN

### 3.8.4 MVCNN Implementation

**ModelNet40 description** The ModelNet40 dataset is a widely recognized benchmark for point cloud analysis, containing synthetic object point clouds. Point clouds are sets

of data points in a three-dimensional coordinate system that represent the external surface of objects. Each point in a point cloud corresponds to a specific location on an object’s surface, which allows for detailed 3D shape analysis. The dataset features 12,311 CAD-generated meshes categorized into 40 distinct classes, such as airplane, car, plant, and lamp. Of these, 9,843 meshes are used for training, while the remaining 2,468 are reserved for testing. The point cloud data is uniformly sampled from the surfaces of these meshes and preprocessed by centering it at the origin and scaling it to fit within a unit sphere, ensuring consistency and facilitating effective analysis.

For classifier training of MVCNN and View-GCN, we utilize the ModelNet40 dataset in conjunction with the Blender script provided by Jongchysu (available at [https://github.com/jongchysu/mvcnn\\_pytorch](https://github.com/jongchysu/mvcnn_pytorch)). This GitHub repository also includes a pre-established multi-view image dataset, which facilitates the training process.

**Traning Process** All the experiments are run remotely through Vast.ai, a market leader in low-cost cloud GPU rental. The system utilized is equipped with an NVIDIA RTX 4090 GPU, offering 82.6 TFLOPS of performance and 24 GB of memory, and features an AMD Ryzen Threadripper PRO 3955WX processor with 16 cores.

The training process of the MVCNN is conducted in two stages, both focusing on classifier training. In the first stage, the pretrained AlexNet is fine-tuned with 40 output classes. In the second stage, the entire MVCNN architecture is trained from scratch. Both stages utilize the same training dataset, consisting of multiple views of each shape from the ModelNet40 dataset, with no data augmentation for scaling or rotation.

The loss function of two stages used is CrossEntropy-Loss. The network aggregates features from multiple views and produces a final class probability distribution using a softmax layer. The cross-entropy loss is then computed based on these predicted probabilities and the true class labels of the 3D shapes.

$$L = - \sum_{c=1}^C y_c \log(p_c) \quad (48)$$

where:  $C$  is the number of classes,  $y_c$  is a binary indicator (0 or 1) if class label  $c$  is the correct classification for the given observation,  $p_c$  is the predicted probability that the observation belongs to class  $c$ .

The training is performed for 30 epochs with a batch size of 12. The Adam optimizer is employed with a learning rate of 0.001 and a weight decay of 0.0005 to adjust the model parameters effectively during training.

### 3.8.5 View-GCN Implementation

Regarding the View-GCN, the training process follows a standard classifier training approach. It is conducted in two steps: first, fine-tuning the pre-trained view feature extractor (e.g., ResNet-18) on multi-view 2D images, and second, performing end-to-end training of the complete View-GCN architecture using the total training loss. Both stages utilize the same training dataset, which comprises 12 multiple views of each shape from the ModelNet40 dataset, with no data augmentation for scaling or rotation.

The training loss for View-GCN consists of two components: shape loss  $L_{\text{shape}}$  and view loss  $L_{\text{view}}$ . The shape loss  $L_{\text{shape}}$  is a cross-entropy loss based on the global shape descriptor  $F$ , used to classify the shape:

$$L_{\text{shape}} = \text{CrossEntropy}(F, y) \quad (49)$$

where  $y$  is the class label of the shape. The view loss  $L_{\text{view}}$  is a cross-entropy loss applied to view-selectors to ensure they can discriminate the shape category based on local view features:

$$L_{\text{view}} = \sum_{l=0}^{L-1} \sum_{j=1}^{N_l} \sum_{v_q \in N(v_j)} \text{CrossEntropy}(V(\hat{f}_{v_q}^l; \theta_v^{l,j}), y) \quad (50)$$

where  $V$  is the view selector function,  $\hat{f}_{v_q}^l$  is the updated feature of node  $v_q$  at level  $l$ , and  $\theta_v^{l,j}$  are the parameters of the view selector. The total training loss is the sum of these two losses:

$$L_{\text{total}} = L_{\text{shape}} + L_{\text{view}} \quad (51)$$

**The Training Process** The training of View-GCN involves two main steps. First, the pre-trained 2D image classification network (e.g., ResNet-18) is fine-tuned on multi-view 2D images of all training 3D objects using SGD with momentum, with a batch size of 30, an initial learning rate of  $10^{-2}$  reduced by half every 10 epochs, and a total of 400 epochs.

Second, the entire View-GCN architecture is trained end-to-end on 3D shapes for shape recognition using SGD with momentum. This step uses an initial learning rate of  $10^{-3}$  for 15 epochs, with a batch size of 20 shapes (400 views) for the 20-views version and 32 shapes (384 views) for the 12-views version. The learning rate is linearly increased from 0 to  $10^{-3}$  in the first epoch and then follows a cosine quarter-cycle reduction to 0.

### 3.8.6 Classification Accuracy

For the classification on the validation dataset, the accuracy is 0.9404 for MVCNN and 0.9762 for View-GCN. This indicates that the trained classification models were able to correctly classify each of the 40 classes in the standard

dataset. This suggests that these models can extract useful information from the input 3D shapes and the output values from the intermediate layers have the potential to be used as descriptors.

### 3.8.7 Multi-view Image Generation for Skull Data

Once we have a classifier trained on a standard dataset, we need to set the well-trained classifier to test mode to ensure the parameters remain unchanged. The next step is to input the multi-view images of our skull data and inspect the output values of certain layers. The first task, therefore, is to generate multi-view images. We generate multi-view images for each sagittal skull using PyVista. PyVista is a visualization library built on VTK (The Visualization Toolkit) that provides a user-friendly, "Pythonic" interface to simplify working with VTK. The 'pyvista.Plotter' class is a versatile tool for displaying VTK meshes or NumPy arrays and is especially useful for automated screenshots and visualizations. In this class, the `camera_position` attribute allows for capturing screenshots from various directions by specifying the camera's position relative to the object.

This attribute defines the camera position using three points that determine the projection plane, enabling the generation of multi-view images from different viewpoints.

As regards the MVCNN, The input shapes are oriented upright along vertical axis, with all cameras set in the central plane of the test skull. Twelve virtual cameras are placed around the mesh at 30-degree intervals, resulting in 12 rendered views. For all skulls with sagittal craniosynostosis, we generate 12 multi-view images with camera setting shown in Figure 23a. According to the discussion in 2.4, ellipsoid fitting can distort rendering information, which is critical in view-based methods. Therefore, we select a normal synthetic skull from the database as the normative skull, indexed as "control\_inst.001\_CN". Multi-view images for the normative skull are generated in the same manner.

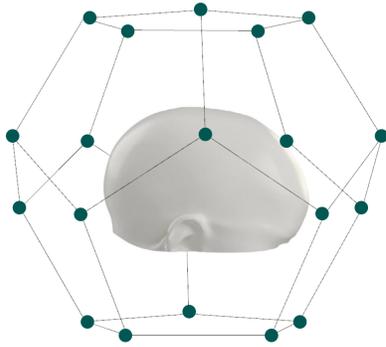
In View-GCN, the camera positions are defined based on an icosahedral configuration, providing a uniform distribution of 20 viewpoints around the skull, shown in Figure 23b.

### 3.8.8 Descriptor Extraction

**MVCNN** We utilized forward hooks in PyTorch to extract intermediate layer outputs from our neural network. Forward hooks are functions registered on specific layers to capture their inputs or outputs during the forward pass. By attaching these hooks to selected layers, we were able to store and analyze the activations generated at various stages of the network. This approach provided valuable insights into the feature extraction and representation learning pro-



(a) Circular camera configuration for MVCNN



(b) Polygon camera configuration for View-GCN

Figure 23. Camera configuration for Image Generation

cesses, enabling a deeper understanding of how the network processes and transforms data through its layers.

As regards the MVCNN, the global shape descriptor is represented by four types of embeddings in the CNN. The first descriptor is the output of the first Linear Layer of the aggregation CNN, with a dimension of 4096. The second descriptor is the input to the second Linear Layer, also with a dimension of 4096; this descriptor has passed through an activation layer (ReLU) compared to the first. The third descriptor is the output of the second Linear Layer after passing through the activation layer, maintaining a dimension of 4096. The final descriptor is the output probability of the neural network with a dimension of 40, representing the probability distribution of the skull data classified into the standard categories of the ModelNet40 dataset in our test mode.

The visualization of descriptors of `sagittal_006` is shown in Figure 24.

**View-GCN** Regarding the View-GCN, the global shape descriptor is not only represented by embeddings in the CNN but also includes  $F^1$  and  $F^2$ . These feature matrices, at different hierarchical levels of the View-GCN, represent

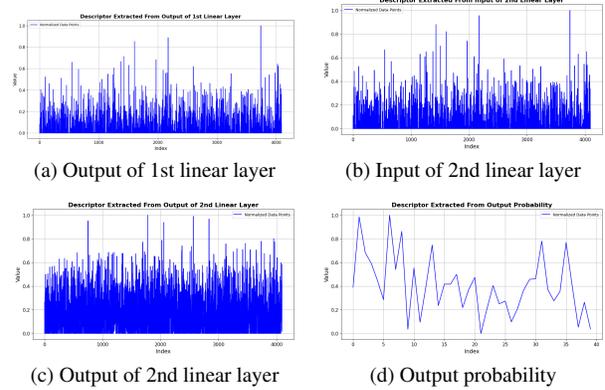


Figure 24. Visualization of different levels of shape descriptor extracted by well-trained MVCNN

progressively coarsened and enriched multi-view features used to construct a global shape descriptor for 3D shape recognition. These matrices are denoted as the input to the `Lview` block in the diagram presented in Section 3.8.3. Based on previous discussions, these matrices are integral to the final loss function of the classifier, making it reasonable to consider these two tensors as shape descriptors. For comparison purposes,  $F^1$  and  $F^2$  are flattened into one-dimensional arrays, with  $F^1$  having a length of 1600 and  $F^2$  having a length of 800. Additionally, the descriptors extracted from the output and input of the second linear layer each have a length of 512, while the descriptor extracted from the output of the first linear layer has a dimensionality of 1536. The visualization of different levels of shape descriptors is shown in Figure 25.

## 4. Result

In this section, we present the findings from our analysis, which involved visualizing the data using box plots and swarm plots, and evaluating the relationship between our various severity scores in the methodology part and the Surgeon Rating index using Spearman rank correlation. The distribution of subjects according to the surgeons' 'overall' severity score is as follows: severity score 0 includes 4 subjects ( $n=4$ ), severity score 1 includes 7 subjects ( $n=7$ ), severity score 2 includes 9 subjects ( $n=9$ ), and severity score 3 includes 8 subjects ( $n=8$ ). These scores represent the surgeons' assessments of the overall severity of the cases.

Figures in this part show the distribution of the Surgeon Rating index across different clinical phenotype scores. The mean severity score for each level, represented by the line within each box, varies across the scores. The interquartile ranges (IQRs), depicted by the boxes, and the whiskers indicate substantial overlap, suggesting significant variability within each clinical phenotype score category. Outliers, marked as points outside the whiskers, are present in

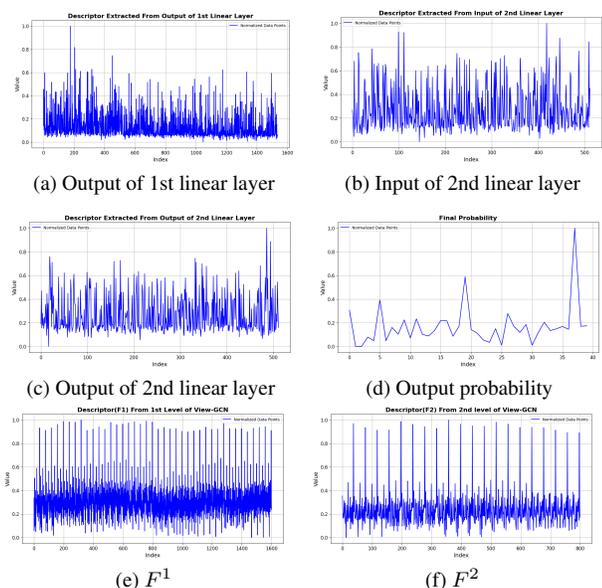


Figure 25. Visualization of different levels of shape descriptor extracted by well-trained View-GCN

multiple categories, indicating some extreme values in the dataset.

The Spearman rank correlation analysis provides further insights into the relationship between clinical phenotype scores and the our severity score. The correlation coefficient and the corresponding p-value are presented to quantify the strength and significance of this relationship.

#### 4.1. Assymetry Index

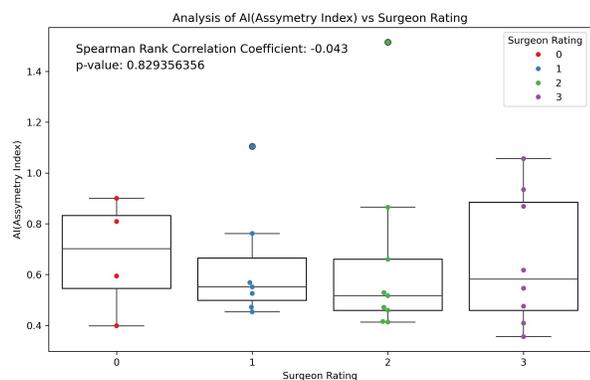


Figure 26. Box plot comparing computed severity scores(AI) for sagittal head models against clinical severity scores

#### 4.2. Difference Between Corresponding Coordinates

From Figure 26, the Spearman rank correlation coefficient is -0.043, with a p-value of 0.829. This low corre-

lation coefficient and high p-value indicate no significant monotonic relationship between the Asymmetry Index and surgeon ratings, suggesting that this method is not effective in correlating Asymmetry Index with the surgeon’s assessment of severity. The mean Asymmetry Index values for surgeon ratings 0, 1, 2, and 3 are approximately 0.65, 0.6, 0.55, and 0.75, respectively. The interquartile ranges (IQRs) show considerable overlap between the groups, further indicating that the Asymmetry Index does not effectively distinguish between different levels of severity.

Overall, the analysis reveals the following insights: The Spearman Rank Correlation Coefficient indicates a strong positive correlation between Surgeon Rating and Mean Absolute Error (MAE); as the Surgeon Rating increases, the MAE tends to increase. Additionally, the p-value confirms that this correlation is statistically significant. These observations collectively validate the effectiveness of all three quantification methods.

The box plot of this method is shown in Figure 27 and summarized in Table 1.

For cases with a Surgeon Rating of 0, the MAE values are tightly clustered around a lower mean distance, suggesting good agreement and lower error for these less severe cases. This indicates that the method performs accurately for such cases.

In cases with a Surgeon Rating of 1, the results are generally satisfactory. When using MSE and MAE as quantification functions, the interquartile ranges are distinct. However, under the EVD condition, there is some overlap between the sections of Surgeon Ratings 0 and 1. Conversely, EVD effectively distinguishes between Surgeon Ratings 1 and 2, with no overlap, highlighting its ability to differentiate between these two severity levels of craniosynostosis. In contrast, MAE and MSE show some overlap between these ratings.

For cases with a Surgeon Rating of 2, the interquartile range under EVD nearly overlaps entirely with that of severity 3.

Notably, all plots for Surgeon Rating 2 exhibit significant outliers, indicating that our severity assessment method does not classify these cases well.

Table 1. Comparison of corresponding Spearson Rank Correlation Coefficient and p-value

Quantification Function	Coefficient	p-value
MAE	0.756	0.00000321
MSE	0.759	0.00000280
EVD	0.722	0.00001420

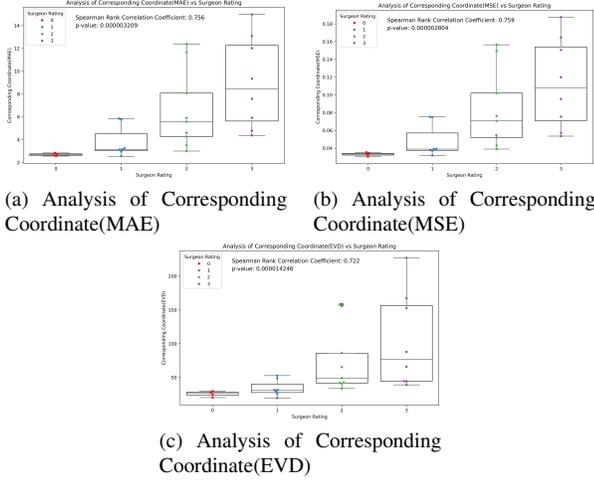


Figure 27. Box plot comparing computed severity scores(Corresponding Coordinate) for sagittal head models against clinical severity scores

### 4.3. D2 Histogram

From Figure 28, we observed a significant overlap in the interquartile ranges. According to statistical principles, if the p-value is smaller than 0.05, we can reject the null hypothesis, indicating that the result is statistically significant. However, in the quantification of EVD, the p-value is greater than 0.05, indicating that the result is not effective. Among the remaining three quantification methods, MSE performs the best in terms of both p-value and correlation coefficient. Nonetheless, overall, they all perform worse than using the Difference Between Corresponding Coordinates as the descriptor. The summary of Spearman rank correlation coefficient is shown in Table 2.

Table 2. Comparison of D2 histogram metrics with their correlation coefficients and p-values

Quantification Function	Coefficient	p-value
MAE	0.383	0.0444
MSE	0.425	0.0242
EVD	0.359	0.0605
JSD	0.375	0.0492

### 4.4. Spectral Embedding

When comparing the four quantification methods, JSD and EVD demonstrate significantly better performance than MSE and MAE in terms of both correlation value and p-value, which is shown in 3. This highlights that the EVD function is more effective for spectrum-related dissimilarity comparisons. Unlike previous methods, which show

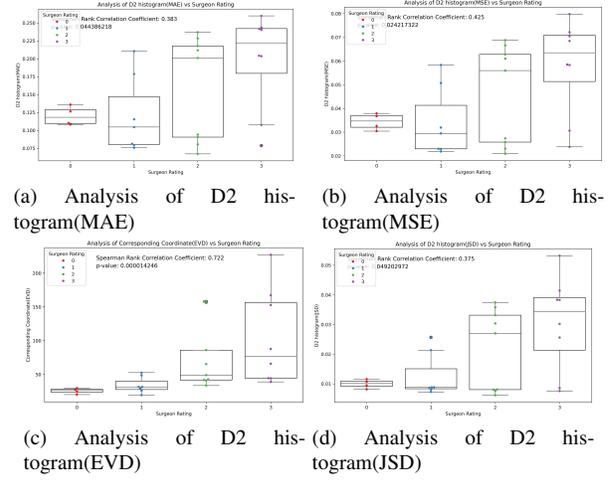


Figure 28. Box plot comparing computed severity scores(D2 histogram) for sagittal head models against clinical severity scores

positive correlation coefficients, this approach uses a negative correlation. This is because it calculates the eigenvalues of the affinity matrix,  $A$ , which is defined as  $A_{ij} = \exp(-d_{ij}^2/2\sigma^2)$ , where  $d_{ij}$  represents the distance between the  $i$ th and  $j$ th vertices of the mesh. This matrix exhibits a negative relationship with point similarity, contributing to the effectiveness of the EVD approach.

Specifically, our severity score using the EVD of the spectrum embedding method accurately reflects the differences between ratings of 2 and 3. However, for distinguishing between ratings of 3 and 4, the JSD score provides a better reflection of the differences, which is shown in Figure 29.

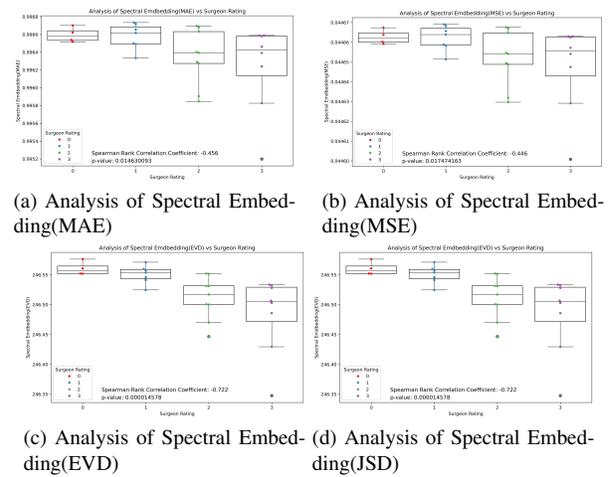


Figure 29. Box plot comparing computed severity scores(Spectral Embeddings) for sagittal head models against clinical severity scores

Table 3. Comparison of Spectral Embedding metrics with their correlation coefficients and p-values

Quantification Function	Coefficient	p-value
MAE	-0.456	0.0146
MSE	-0.446	0.0175
EVD	-0.722	< 0.0001
JSD	-0.748	< 0.0001

### 4.5. LBO Spectrum

In these tests, the p-value for EVD is greater than 0.05, indicating that it is not statistically significant. The remaining three methods are acceptable; according to our principles, the severity scores generated by these methods show some level of correlation with expert ratings. The best performance is achieved using JSD as the quantification function between the normative and test skulls. A careful examination of the interquartile ranges reveals that, among the four groups, those classified as 0 and 1 (relatively low severity) and 2 and 3 (relatively high severity) are clearly distinguishable.

The summary is shown by Table 4 and Figure 30.

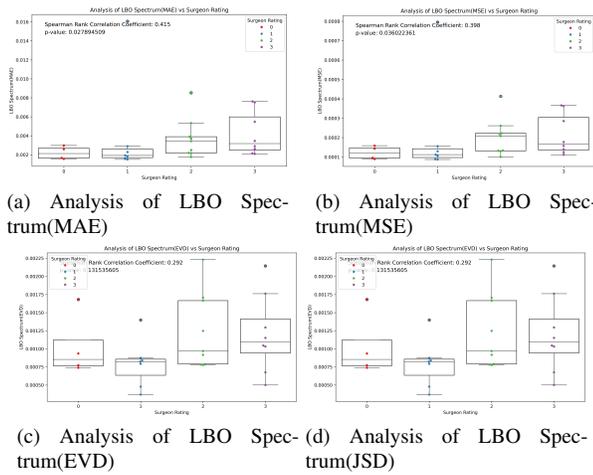


Figure 30. Box plot comparing computed severity scores(LBO Spectrum) for sagittal head models against clinical severity scores

### 4.6. Laplacian Heat Kernel Descriptor

It was observed that there are a large number of zeros in the descriptors, making the use of generalized eigenvalue distance for quantifying differences unsuitable. The p-values for MSE and MAE were notably poor, shown in Table 5. Although JSD is acceptable in terms of both correlation coefficient and p-value, and is statistically significant, the box plots reveal that it is still challenging to dis-

Table 4. Comparison of LBO Spectrum metrics with their correlation coefficients and p-values

Quantification Function	Coefficient	p-value
MAE	0.415	0.0279
MSE	0.398	0.0360
EVD	0.292	0.1315
JSD	0.482	0.0094

tinguish between groups. Notably, expert ratings of 2 dominate nearly 70% of the plot in Figure 31, indicating limited discrimination among the groups.

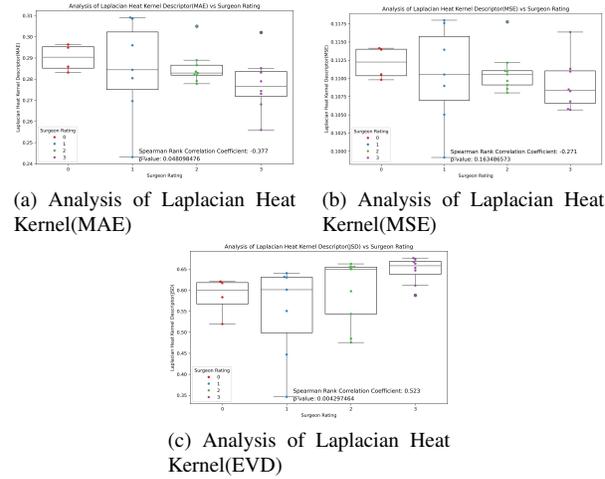


Figure 31. Box plot comparing computed severity scores(Laplacian Heat Kernel) for sagittal head models against clinical severity scores

Table 5. Comparison of Laplacian Heat Kernel Descriptor metrics with their correlation coefficients and p-values

Quantification Function	Coefficient	p-value
MAE	-0.377	0.0481
MSE	-0.271	0.1635
JSD	0.523	0.0043

### 4.7. Zernike Shape Descriptor

After obtaining the Zernike coefficients, we calculated and plotted the correlation coefficient curve with respect to the order for four different quantification functions shown in Figure 32.

Based on above images, we determined that the optimal order and quantification function are an order of 22 and MSE as the quantification metric. While there are a few outliers where ratings of 2 and 3 cannot be distinguished due to

their similar values, the remaining interquartile range is distinctly separated, shown in Figure 33. In terms of p-value and correlation coefficient, this approach proves to be the most effective.

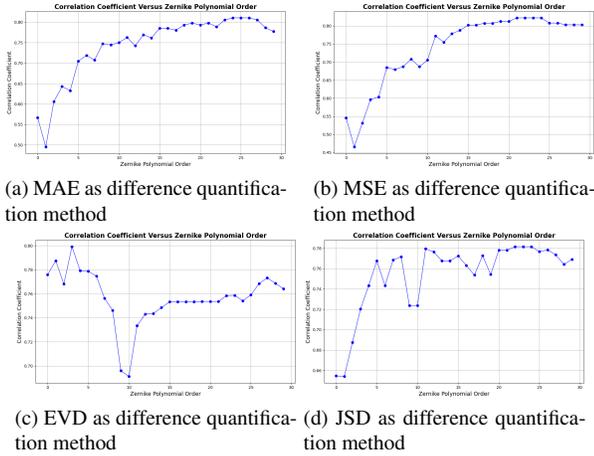


Figure 32. Box plot comparing computed severity scores(Laplacian Heat Kernel) for sagittal head models against clinical severity scores

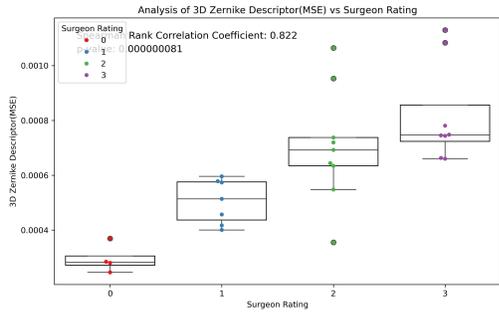


Figure 33. Box plot comparing computed severity scores(3D Zernike Descriptor(order=22)) for sagittal head models against clinical severity scores

## 4.8. View-based Method

### 4.8.1 MVCNN-based Descriptor

From the result image, it's common for some nodes in the hidden layers of a neural network to have values of zero. When these values are used as shape descriptors, the presence of many zeros makes it unsuitable to use generalized eigenvalue distance as a measure of difference.

Based on the p-values for all the test results, none of the outcomes are statistically significant. However, the 1st linear output shows the best performance, with p-values around 0.07 for JSD(36), MAE(34), and MSE(35). This is

in contrast to the final aggregated descriptor of output probability.

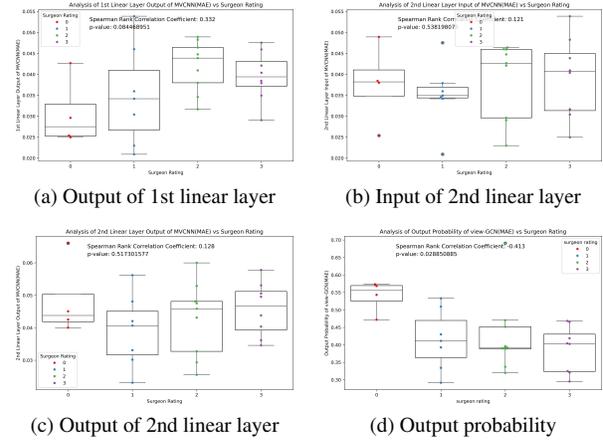


Figure 34. Box plot comparing computed severity scores(MVCNN related, MAE) for sagittal head models against clinical severity scores

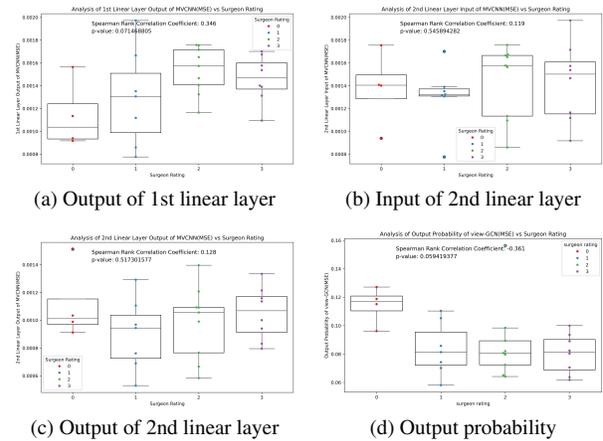


Figure 35. Box plot comparing computed severity scores(MVCNN related, MSE) for sagittal head models against clinical severity scores

### 4.8.2 View-GCN-based Descriptor

By observing the p-values, we can see that only a few severity assessment methods are statistically significant. We provide a summary table of these statistically significant methods. Additionally, in the plots, the interquartile ranges show a considerable amount of overlap.

The figures of boxplot are shown in Figure37,38,39 and summary of correlation coefficient is shown in Table 6.

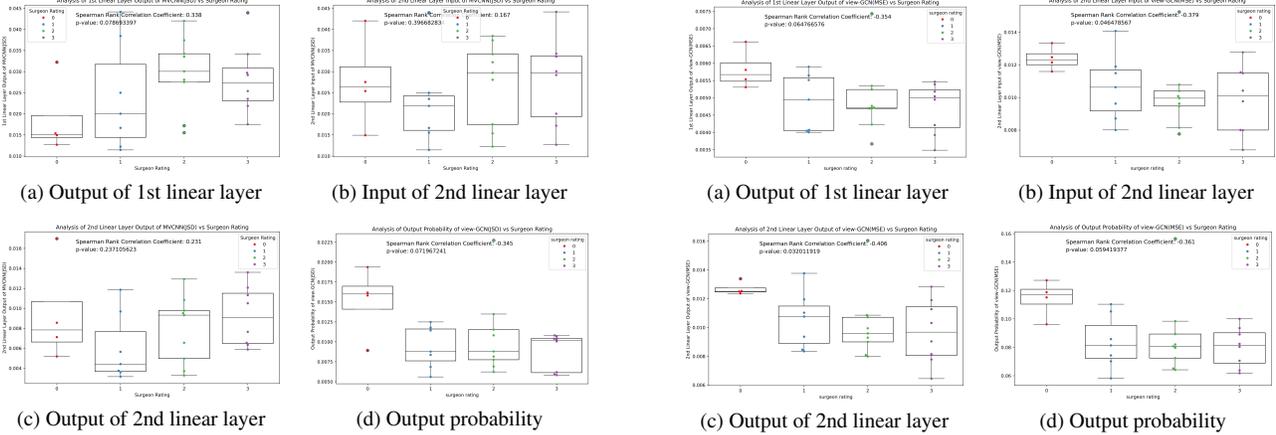


Figure 36. Box plot comparing computed severity scores(MVCNN related, JSD) for sagittal head models against clinical severity scores

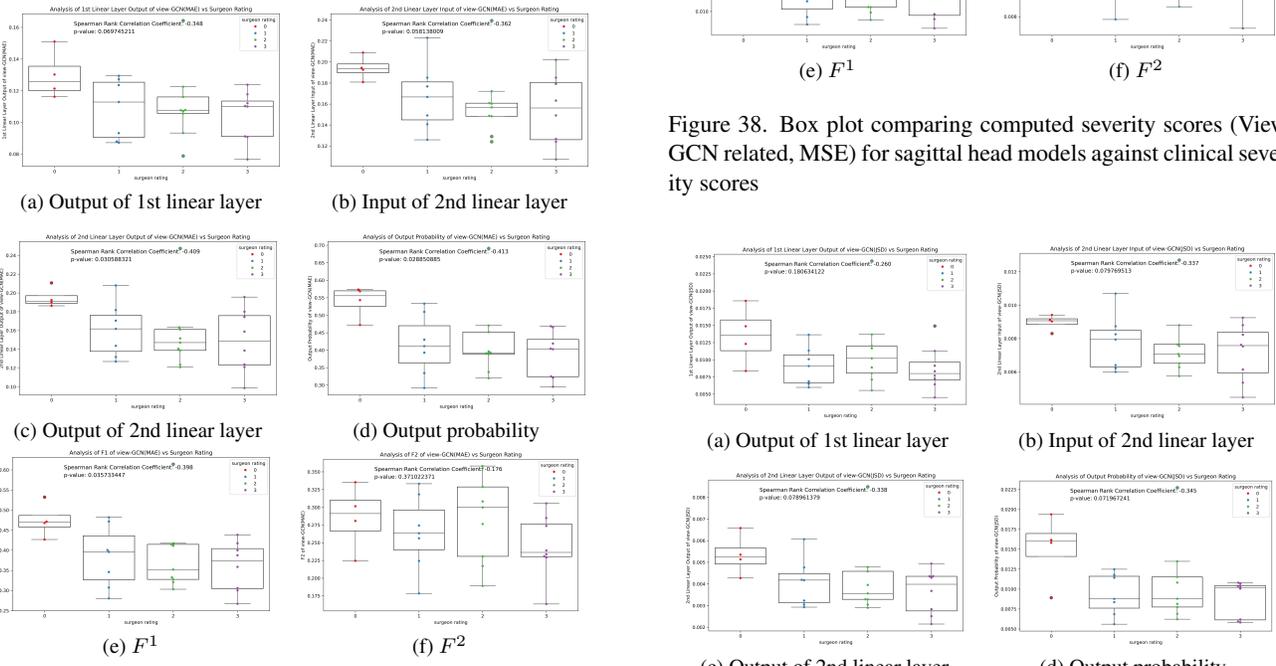


Figure 37. Box plot comparing computed severity scores(View-GCN related, MAE) for sagittal head models against clinical severity scores

## 5. Discussion

### 5.1. Asymmetry Index

Regarding the Severity score of AI (Asymmetry Index), the median Asymmetry Index values for surgeon ratings 0, 1, 2, and 3 are approximately 0.65, 0.60, 0.55, and 0.75, respectively. There is no apparent monotonic trend in the Asymmetry Index values relative to expert ratings, and the large p-values suggest that these differences are not statistically significant.

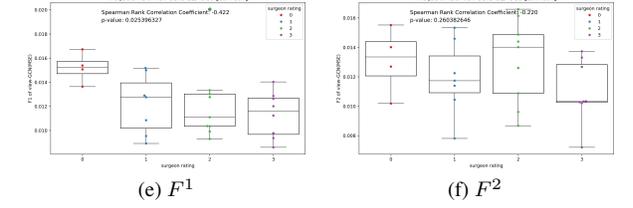


Figure 38. Box plot comparing computed severity scores (View-GCN related, MSE) for sagittal head models against clinical severity scores

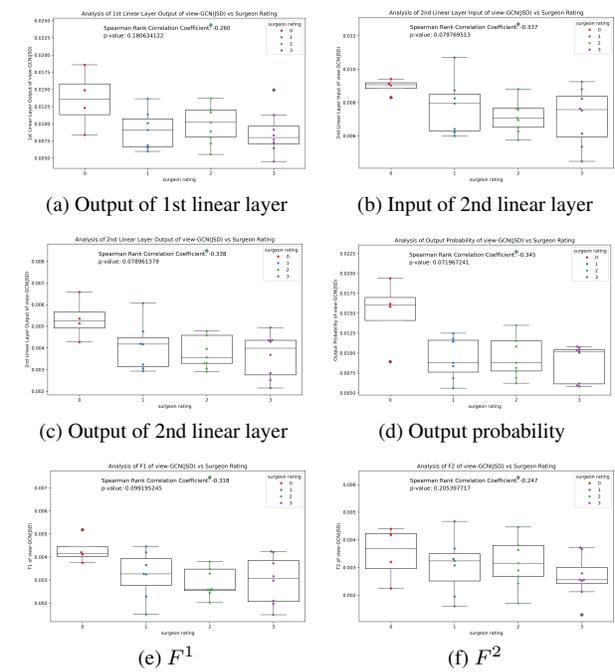


Figure 39. Box plot comparing computed severity scores (View-GCN related, JSD) for sagittal head models against clinical severity scores

Upon further investigation, the visualizationally significant.

Table 6. Test Results and Statistical Significance

Description	Coefficient	p-value
2nd Linear Layer output(MAE)	-0.409	0.0306
Output probability(MAE)	-0.413	0.0289
$F^1$ (MAE)	-0.398	0.0357
$F^2$ (MAE)	-0.176	0.3710
$F^2$ (MSE)	-0.220	0.2600
$F^1$ (MSE)	-0.422	0.0254
2nd Linear Layer Output(MSE)	-0.406	0.0320
2nd Linear Layer Input(MSE)	-0.379	0.0465

tion of a control skull and a sagittal skull regarding its asymmetry index is created in Figure 40, 41. The control skull has a higher level of asymmetry index shown by a deeper color in the heat map and a higher Mean Asymmetry Index overall. This suggests that while the asymmetry index might be an important shape feature, it does not appear to be clinically relevant for diagnosing severity.

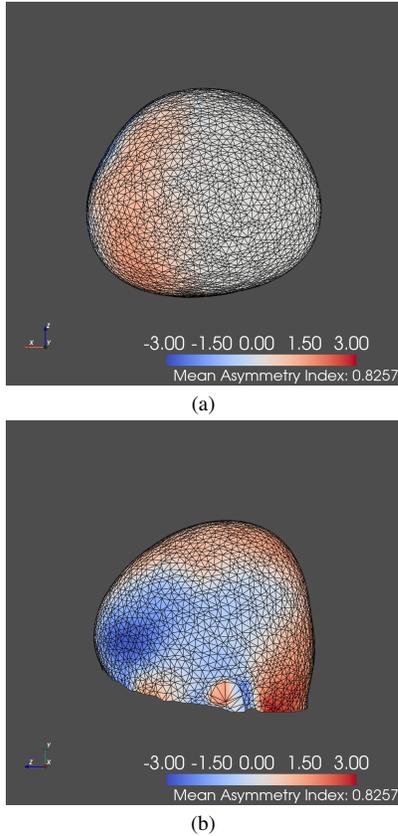


Figure 40. The visualization of Asymmetry Index of a normal skull

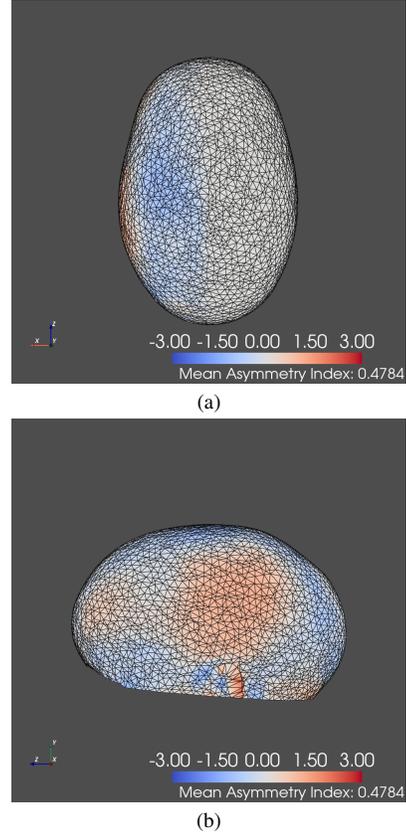


Figure 41. The visualization of Asymmetry Index of a sagittal skull

## 5.2. Corresponding Coordinate

Regarding the corresponding coordinate difference, this method works well by calculating the coordinate difference between each point and its own normative fitted ellipsoid, achieving a high level of distinctiveness. By leveraging the continuous curvature characteristics of the fitted ellipsoid, we perfectly capture the clinically significant regions that differ in the test skull. Compared with using coordinates directly, the greatest drawback of local shape descriptors, especially those using coordinate-derived methods, is their ineffectiveness in providing invariance to transformations such as translation, rotation, and scaling. To address this, we build a normative skull for each test skull, eliminating the issue of coordinates being affected by rotation and translation, which significantly enhances the applicability of our method in clinical settings. Additionally, in the context of 3D photography for clinical data, standards are relatively inconsistent regarding point number, point establishment, and frequent rotations. By employing an ellipsoid-fitted skull as the normative reference, we effectively address the point establishment requirement for quantifying differences between the control skull and the test skull.

### 5.3. Histogram-based Method

Regarding the two histogram-based methods, namely the D2 shape descriptor and the Laplacian heat kernel descriptor (coefficient of D2 histogram: 0.425, p-value = 0.0242; coefficient of Laplacian heat kernel descriptor: 0.523, p-value = 0.0043), both produce normalized histograms with a length of 17. However, these histograms may have limited discriminative power when distinguishing between relatively similar shapes, such as skulls with varying levels of severity. This low discriminative power arises because histogram-based methods aggregate shape information into bins, potentially losing details necessary for distinguishing subtle variations [43], verified by Figure 42. When shapes are very similar, the aggregated histogram features can appear nearly identical, making it challenging to capture minor but clinically significant differences.

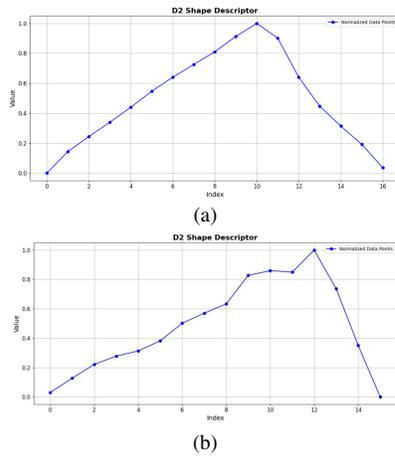


Figure 42. The D2 histogram of severity of 3 and the severity of 1 both exhibit a trend of initially increasing and then decreasing.

### 5.4. Spectrum-based Method

Comparing spectral embedding using the affinity matrix and the LBO spectrum, the former significantly outperforms the latter (coefficient of spectral embedding:  $-0.748$ , p-value =  $4.639 \times 10^{-6}$ ; coefficient of LBO: 0.482, p-value = 0.00939). Both descriptors use spectral related information to capture the variational patterns. The affinity matrix in spectral embedding measures pairwise similarities between points or features on the shape, offering greater robustness to noise and artifacts. By directly assessing how similar or close two points are, it maintains meaningful shape information despite noise.

In contrast, the LBO spectrum relies on the Laplacian-Beltrami operator, which is sensitive to noise and small variations in surface geometry due to its dependence on local

curvature and differential properties. This sensitivity can reduce the reliability of shape recognition when noise is present. Also, the derivative component of LBO operator makes it focus on local information, encoding the shape's connectivity. While these edge-level details are important for graphical rendering, they are less crucial for our shape analysis needs.

Although LBOs excel at capturing local shape features, affinity matrices provide a more comprehensive and flexible approach to shape representation. This makes spectral embedding using the affinity matrix more effective for our purposes, as it better preserves overall shape characteristics and is more resilient to noise and variability in the data.

Compared to the LBO Spectrum, the Laplacian Heat Kernel Descriptor performs better, achieving a Spearman rank correlation coefficient of 0.523, which is slightly higher than that of the LBO Spectrum. This improvement is likely because the Heat Kernel Descriptor integrates global information through the Heat Kernel Signature (HKS) and distance histograms, rather than relying solely on local shape features as in the LBO Spectrum. However, its performance is still significantly lower than that of the Spectral Embedding.

### 5.5. 3D Zernike Descriptor

Among all the methods evaluated, Zernike moment perform the best, as evidenced by their superior separation of interquartile ranges, higher correlation coefficients, and more significant p-values. Zernike moment offer greater discriminative power for circular and radial data (e.g., unit sphere distance mapping discussed in Section 3.7.6) due to their tailored basis functions compared to spectral or eigenvalue methods.

By examining the Zernike descriptor severity score across different orders, we observe that performance is optimal between orders 20 and 25 under four kinds of quantification functions. This finding suggests that the Zernike descriptor captures detailed and clinically relevant information about skull severity, highlighting that important shape-related details are present at a fine level, which distance-based or spectral methods may miss. The precision of double-precision floating-point numbers is surpassed at order 9 [56]. Additionally, research in protein shape analysis, which demands high precision and discriminative power beyond that required for standard object classification, indicates that moments up to order 20 are necessary for effective descriptors [21, 37, 73]. This reinforces the importance of capturing detailed information for accurate severity assessment in further research.

### 5.6. MVCNN

For MVCNN (Multi-View Convolutional Neural Networks), we observed no statistically significant features in

our analysis. This lack of significant findings could be attributed to the specific viewpoints chosen for capturing the images, which may not have included the relevant features needed for effective analysis. Alternatively, the selected viewpoints might not have been discriminative enough to highlight meaningful differences. Furthermore, the cameras used in the study were all positioned on the same plane, which imposed inherent limitations on the captured multi-view images. This uniform camera positioning restricted the diversity of perspectives and angles, potentially leading to less informative and less varied images. Such limitations can hinder the ability of MVCNN to extract and learn from distinctive features that could have improved feature detection and severity clustering performance.

### 5.7. View-GCN

View-GCN performed significantly better compared to MVCNN, which may be attributed to the fact that some image angles, such as the oblique view, are not crucial for assessing severity. However, its performance still lags behind spectrum-based methods. While images from certain views, like the top view, effectively capture asymmetry information and classify shape based on geometric features, they do not significantly impact the severity evaluation for sagittal craniosynostosis. When aggregating features, we cannot adequately down-weight these less relevant angles because they are important for classification. Consequently, our analysis included too much clinically irrelevant but shape classification relevant information.

In training neural networks, it is often necessary to fine-tune hyperparameters to improve performance. One challenge with applying neural network methods to this project is that, even if we utilize the feature representation capacity of a shape classifier, the well-trained classifier with a high level of test accuracy will capture shape-related information. This information is highly related to recognition. However, while these features can classify the shape of the skull in my test, they are not entirely clinically related to severity. This means that the optimal severity quantification performance for the skull does not correspond to the parameter values that yield the highest classification accuracy. This makes fine-tuning the network directionless.

### 5.8. Potential Application

#### 5.8.1 Guidance on Whether Surgery is Required: Objective Recommendations

The mesh data we tested is within a case that has been diagnosed as requiring surgery. We use the three best-performing methods described earlier—Spectral Embedding (JSD), Corresponding Coordinate (MSE), and 3D Zernike Coefficient (order = 22, MSE)—to assess the severity of this case and provide objective guidance. The red indicates the severity score patient corre in following figures.

**Corresponding Coordinate (MSE)** The average severity scores for surgeon ratings 0, 1, 2, and 3 are approximately 0.03, 0.04, 0.07, and 0.105, respectively. The severity score for the case in the image is 0.164. According to this method, a score of 0.164 would be classified as a severity rating of 3 in Figure 43.

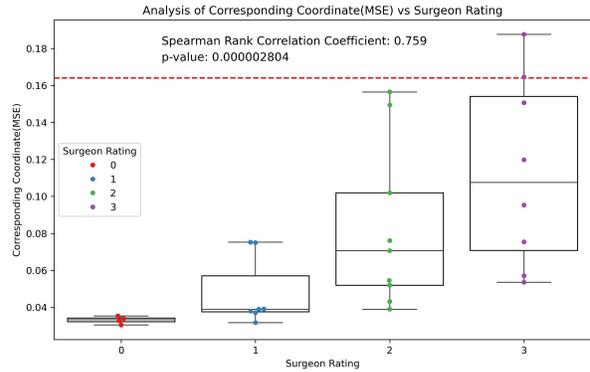


Figure 43. Suggestion from Corresponding Coordinate(MSE)

**Spectral Embedding (JSD)** The severity score from Spectral Embedding (JSD) is negatively correlated with surgeon ratings. The average severity scores for surgeon ratings 0, 1, 2, and 3 are approximately 0.651675, 0.6516, 0.65145, and 0.65135, shown in Figure 44 respectively. The severity score for the case in the image is 0.65125. Based on this method, the severity would be categorized as either rating 2 or 3. Considering the results shown in the figure section, the score falls within the interquartile range of severity rating 3 and not within that of rating 2. Thus, it is most likely a severity rating of 3.

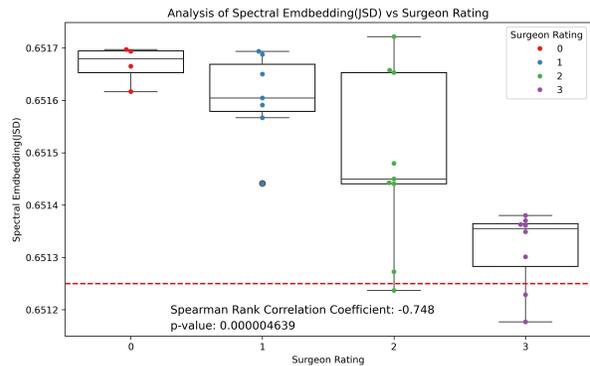


Figure 44. Suggestion from Spectral Embedding(JSD)

**3D Zernike Coefficient (order = 22, MSE)** The average severity scores for surgeon ratings 0, 1, 2, and 3 are approximately 0.0002, 0.0005, 0.0007, and 0.00075 shown in

Figure 45, respectively. The severity score for this case is 0.00098, which falls entirely within the interquartile range for severity rating 3. Therefore, it is highly likely that this case corresponds to severity level 3.

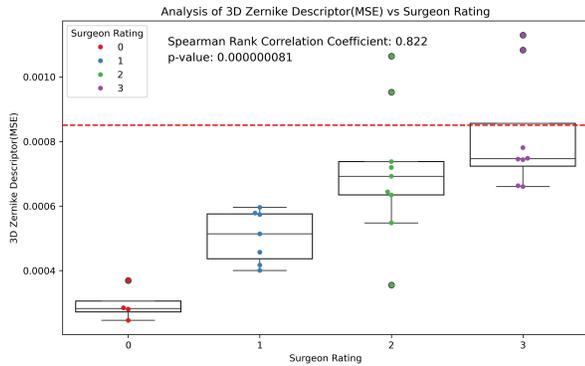


Figure 45. Suggestion from 3D Zernike Descriptor(MSE)

**Summary of the Three Evaluation Methods** Although not all methods can precisely differentiate between severity ratings 2 and 3, the overall conclusion is that the case is indicative of severe craniosynostosis. This conclusion supports the surgeon’s recommendation for surgery. Therefore, this research offers a valuable tool for providing objective support in the surgeon’s diagnostic process.

### 5.9. Recovery Situation Analysis: Pre- and Post-Surgery

We have pre- and post-surgery mesh data for seven anonymous patients who have undergone surgery. Given that our severity scores are not very intuitive due to being small decimal numbers, we applied a linear mapping, specifically mapping the minimum and maximum values of the y-axis in the box plot to a range of 0-100. We calculated the severity scores for each patient using the Zernike Shape Descriptor, identified as the best-performing method in terms of correlation coefficient and p-value, and mapped these scores to the 0-100 range.

The Figure 46, 47 illustrate these results, we can see that our descriptor effectively captures the improvement in shape information in most of cases aside from the third one, demonstrating its potential for assessing the efficacy of the surgery.

### References

[1] Tareq Abdel-Alim, Rik Iping, Eppo B Wolvius, Irene MJ Mathijssen, Clemens MF Dirven, Wiro J Niessen, Marie-Lise C van Veelen, and Gennady V Roshchupkin. Three-dimensional stereophotogrammetry in the evaluation of craniosynostosis: current and potential use cases. *Journal of Craniofacial Surgery*, 32(3):956–963, 2021. 1, 2

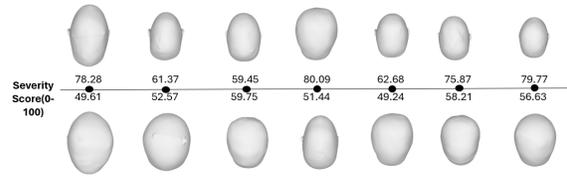


Figure 46. Severity Score Comparison between Pre- and Post-surgery

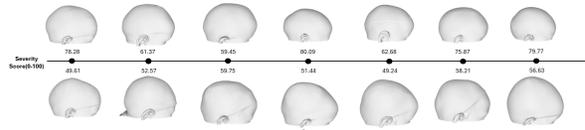


Figure 47. Severity Score Comparison between Pre- and Post-surgery

[2] Tareq Abdel-Alim, Rik Iping, Eppo B Wolvius, Irene MJ Mathijssen, Clemens MF Dirven, Wiro J Niessen, Marie-Lise C van Veelen, and Gennady V Roshchupkin. Three-dimensional stereophotogrammetry in the evaluation of craniosynostosis: current and potential use cases. *Journal of Craniofacial Surgery*, 32(3):956–963, 2021. 2

[3] Tareq Abdel-Alim, Franz Tapia Chaca, Irene MJ Mathijssen, Clemens MF Dirven, Wiro J Niessen, Eppo B Wolvius, Marie-Lise C van Veelen, and Gennady V Roshchupkin. Quantifying dysmorphologies of the neurocranium using artificial neural networks. *Journal of Anatomy*, 2024. 3, 14

[4] Tareq Abdel-Alim, Pauline Tio, Melissa Kurniawan, Irene Mathijssen, Clemens Dirven, Wiro Niessen, Gennady Roshchupkin, and Marie-Lise van Veelen. Reliability and agreement of automated head measurements from 3-dimensional photogrammetry in young children. *Journal of Craniofacial Surgery*, 34(6):1629–1634, 2023. 2

[5] Tareq Abdel-Alim, Pauline Tio, Melissa Kurniawan, Irene Mathijssen, Clemens Dirven, Wiro Niessen, Gennady Roshchupkin, and Marie-Lise van Veelen. Reliability and agreement of automated head measurements from 3-dimensional photogrammetry in young children. *Journal of Craniofacial Surgery*, 34(6):1629–1634, 2023. 6, 10

[6] Piers Allen, Antonio Calcagni, Anthony G Robson, and Ela Claridge. Investigating the potential of zernike polynomials to characterise spatial distribution of macular pigment. *Plos one*, 14(5):e0217265, 2019. 12

[7] Brian Amberg, Sami Romdhani, and Thomas Vetter. Optimal step nonrigid icp algorithms for surface registration. In *2007 IEEE conference on computer vision and pattern recognition*, pages 1–8. IEEE, 2007. 3

[8] Fouzul Atik and Pratima Panigrahi. On the distance and distance signless laplacian eigenvalues of graphs and the smallest gersgorin disc. *The Electronic Journal of Linear Algebra*, 34:191–204, 2018. 8

- [9] Mateusz Banach. Structural outlier detection and zernike–canterakis moments for molecular surface meshes—fast implementation in python. *Molecules*, 29(1):52, 2023. **2, 12**
- [10] Inés Barbero-García, José Luis Lerma, Ángel Marqués-Mateu, and Pablo Miranda. Low-cost smartphone-based photogrammetry for the analysis of cranial deformation in infants. *World neurosurgery*, 102:545–554, 2017. **2, 3**
- [11] Charlotte L Bendon, Hayley P Johnson, Andrew D Judge, Steven A Wall, and David Johnson. The aesthetic outcome of surgical correction for sagittal synostosis can be reliably scored by a novel method of preoperative and postoperative visual assessment. *Plastic and reconstructive surgery*, 134(5):775e–786e, 2014. **2**
- [12] Franz Boas. The cephalic index. *American Anthropologist*, 1(3):448–461, 1899. **2**
- [13] Alexander I Bobenko and Boris A Springborn. A discrete laplace–beltrami operator for simplicial surfaces. *Discrete & Computational Geometry*, 38(4):740–756, 2007. **9**
- [14] Joosje Bommer-Skogstad, Marjolein Dremmen, Irene Mathijssen, and Robin Smithuis. Craniosynostosis. *Department of Radiology of Akershus University Hospital in Lørenskog, Norway, Departments of Radiology and Plastic Surgery of Erasmus Medical Center in Rotterdam and Department of Radiology of Alrijne Hospital in Leiderdorp, the Netherlands*. **2**
- [15] Simeon A Boyadjiev and International Craniosynostosis Consortium. Genetic analysis of non-syndromic craniosynostosis. *Orthodontics & craniofacial research*, 10(3):129–137, 2007. **1**
- [16] David J Brenner, Carl D Elliston, Eric J Hall, and Walter E Berdon. Estimated risks of radiation-induced fatal cancer from pediatric ct. *American journal of roentgenology*, 176(2):289–296, 2001. **1, 2**
- [17] Min-Jeong Cho, Rami R Hallac, Maleeh Effendi, James R Seaward, and Alex A Kane. Comparison of an unsupervised machine learning algorithm and surgeon diagnosis in the clinical differentiation of metopic craniosynostosis and benign metopic ridge. *Scientific reports*, 8(1):6312, 2018. **14**
- [18] M Michael Cohen Jr. Sutural biology and the correlates of craniosynostosis. *American journal of medical genetics*, 47(5):581–616, 1993. **1**
- [19] John P Connolly, Joseph Gruss, Marianne L Seto, Michael F Whelan, Richard Ellenbogen, Avery Weiss, Steven R Buchman, and Michael L Cunningham. Progressive postnatal craniosynostosis and increased intracranial pressure. *Plastic and reconstructive surgery*, 113(5):1313–1323, 2004. **1**
- [20] Martijn Cornelissen, Bianca den Ottelander, Dimitris Rizopoulos, René van der Hulst, Aebele Mink van der Molen, Chantal van der Horst, Hans Delye, Marie-Lise van Veelen, Gouke Bonsel, and Irene Mathijssen. Increase of prevalence of craniosynostosis. *Journal of Cranio-Maxillofacial Surgery*, 44(9):1273–1279, 2016. **1, 2**
- [21] Sebastian Daberdaku and Carlo Ferrari. Exploring the potential of 3d zernike descriptors and svm for protein–protein interface prediction. *BMC bioinformatics*, 19:1–23, 2018. **24**
- [22] K Ch Das. The laplacian spectrum of a graph. volume 48, pages 715–724. Elsevier, 2004. **13**
- [23] Guido de Jong, Elmar Bijlsma, Jene Meulstee, Myrte Wennen, Erik van Lindert, Thomas Maal, René Aquarius, and Hans Delye. Combining deep learning with 3d stereophotogrammetry for craniosynostosis diagnosis. *Scientific reports*, 10(1):15346, 2020. **14**
- [24] Guido de Jong, Elmar Bijlsma, Jene Meulstee, Myrte Wennen, Erik van Lindert, Thomas Maal, René Aquarius, and Hans Delye. Combining deep learning with 3d stereophotogrammetry for craniosynostosis diagnosis. *Scientific reports*, 10(1):15346, 2020. **14**
- [25] Linda Gaillard and ERN CRANIO-Craniosynostosis Workgroup. A photo score for aesthetic outcome in sagittal synostosis: an ern crania collaboration. *Journal of Craniofacial Surgery*, 34(8):2279–2283, 2023. **2**
- [26] Linda Gaillard and ERN CRANIO-Craniosynostosis Workgroup. A suture-specific photo score for metopic synostosis. *Journal of Craniofacial Surgery*, 35(1):75–79, 2024. **2**
- [27] Arturo Garrocho-Rangel, Lizeth Manríquez-Olmos, Joséln Flores-Velázquez, Miguel-Ángel Rosales-Berber, Ricardo Martínez-Rider, and Amaury Pozos-Guillén. Non-syndromic craniosynostosis in children: Scoping review. *Medicina oral, patología oral y cirugía bucal*, 23(4):e421, 2018. **1**
- [28] Thomas R Goodman, Adel Mustafa, and Erin Rowe. Pediatric ct radiation exposure: where we were, and where we are now. *Pediatric radiology*, 49:469–478, 2019. **1, 2**
- [29] Varun Jain and Hao Zhang. Shape-based retrieval of articulated 3d models using spectral embedding. In *International Conference on Geometric Modeling and Processing*, pages 299–312. Springer, 2006. **7**
- [30] Varun Jain and Hao Zhang. A spectral approach to shape-based retrieval of articulated 3d models. *Computer-Aided Design*, 39(5):398–407, 2007. **7**
- [31] Varun Jain and Hao Zhang. A spectral approach to shape-based retrieval of articulated 3d models. *Computer-Aided Design*, 39(5):398–407, 2007. **7**
- [32] Haidar Kabbani and Talkad S Raghuvver. Craniosynostosis. *American family physician*, 69(12):2863–2870, 2004. **1**
- [33] Nina Kajdic, Peter Spazzapan, and Tomaz Velnar. Craniosynostosis-recognition, clinical characteristics, and treatment. *Bosnian journal of basic medical sciences*, 18(2):110, 2018. **1**
- [34] Nina Kajdic, Peter Spazzapan, and Tomaz Velnar. Craniosynostosis-recognition, clinical characteristics, and treatment. *Bosnian journal of basic medical sciences*, 18(2):110, 2018. **1**
- [35] Nina Kajdic, Peter Spazzapan, and Tomaz Velnar. Craniosynostosis-recognition, clinical characteristics, and treatment. *Bosnian journal of basic medical sciences*, 18(2):110, 2018. **1**
- [36] Daisuke Kihara, Lee Sael, Rayan Chikhi, and Juan Esquivel-Rodriguez. Molecular surface representation using 3d zernike descriptors for protein shape comparison and docking. *Current Protein and Peptide Science*, 12(6):520–530, 2011. **11**
- [37] Daisuke Kihara, Lee Sael, Rayan Chikhi, and Juan Esquivel-Rodriguez. Molecular surface representation using 3d

- zernike descriptors for protein shape comparison and docking. *Current Protein and Peptide Science*, 12(6):520–530, 2011. [24](#)
- [38] Konstantin C Koban, Felix Härtnagl, Virginia Titze, Thilo L Schenck, and Riccardo E Giunta. Chances and limitations of a low-cost mobile 3d scanner for breast imaging in comparison to an established 3d photogrammetric system. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 71(10):1417–1423, 2018. [2](#)
- [39] Takeshi Kouga, Koji Tanoue, and Kiyoshi Matsui. Airway statuses and nasopharyngeal airway use for airway obstruction in syndromic craniosynostosis. *Journal of Craniofacial Surgery*, 25(3):762–765, 2014. [1](#)
- [40] Melissa SIC Kurniawan, Pauline AE Tio, Tareq Abdel Alim, Gennady Roshchupkin, Clemens MF Dirven, Mieke M Pleumeekers, Irene MJ Mathijssen, and Marie-Lise C van Veelen. 3d analysis of the cranial and facial shape in craniosynostosis patients: A systematic review. *Journal of Craniofacial Surgery*, 35(3):813–821, 2024. [1](#), [2](#)
- [41] Melissa SIC Kurniawan, Pauline AE Tio, Tareq Abdel Alim, Gennady Roshchupkin, Clemens MF Dirven, Mieke M Pleumeekers, Irene MJ Mathijssen, and Marie-Lise C van Veelen. 3d analysis of the cranial and facial shape in craniosynostosis patients: A systematic review. *Journal of Craniofacial Surgery*, 35(3):813–821, 2024. [2](#)
- [42] Hamid Laga, Yulan Guo, Hedi Tabia, Robert B Fisher, and Mohammed Bennamoun. *3D Shape analysis: fundamentals, theory, and applications*. John Wiley & Sons, 2018. [7](#)
- [43] Graciela Lara Lopez, Adriana Peña Pérez Negrón, Angelica De Antonio Jimenez, Jaime Ramirez Rodriguez, and Ricardo Imbert Paredes. Comparative analysis of shape descriptors for 3d objects. *Multimedia Tools and Applications*, 76:6993–7040, 2017. [24](#)
- [44] Shengjun Liu, Haibo Wang, Dong-Ming Yan, Qinsong Li, Feifan Luo, Zi Teng, and Xinru Liu. Spectral descriptors for 3d deformable shape matching: A comparative survey. *IEEE Transactions on Visualization and Computer Graphics*, 2024. [9](#)
- [45] Daniel Malacara. Interferogram evaluation and wavefront fitting. *Optical shop testing*, 13, 1992. [12](#)
- [46] Marlon Marcon, Riccardo Spezialetti, Samuele Salti, Luciano Silva, and Luigi Di Stefano. Unsupervised learning of local equivariant descriptors for point clouds. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12):9687–9702, 2021. [12](#)
- [47] Pouria Mashouri, Marta Skreta, John Phillips, Dianna McAllister, Melissa Roy, Senthujan Senkaiahliyan, Michael Brudno, and Devin Singh. 3d photography based neural network craniosynostosis triaging system. In *Machine Learning for Health*, pages 226–237. PMLR, 2020. [14](#)
- [48] Colm McAlinden, Mark McCartney, and Jonathan Moore. Mathematics of zernike polynomials: a review. *Clinical & experimental ophthalmology*, 39(8):820–827, 2011. [11](#)
- [49] Johanna M Meulepas, Cécile M Ronckers, Anne MJB Smets, Rutger AJ Nivelstein, Patrycja Gradowska, Choonsik Lee, Andreas Jahnen, Marcel van Straten, Marie-Claire Y de Wit, Bernard Zonnenberg, et al. Radiation exposure from pediatric ct scans and subsequent cancer risk in the netherlands. *JNCI: Journal of the National Cancer Institute*, 111(3):256–263, 2019. [1](#)
- [50] Stefano Antonio Mezzasalma, Marek Grzelczak, and Jordi Sancho-Parramon. Shape plasmonics and geometric eigenvalues: The crystal field plasmon splitting in a sphere-to-cube continuous transition. *arXiv preprint arXiv:1912.06454*, 2019. [2](#)
- [51] Katsuhiko Mizutani, Tomoru Miwa, Yoshiaki Sakamoto, and Masahiro Toda. Application of deep learning techniques for automated diagnosis of non-syndromic craniosynostosis using skull. *Journal of Craniofacial Surgery*, 33(6):1843–1846, 2022. [14](#)
- [52] Anthony M Musolf, Cristina M Justice, Zeynep Erdogan-Yildirim, Seppe Goovaerts, Araceli Cuellar, John R Shaffer, Mary L Marazita, Peter Claes, Seth M Weinberg, Jae Li, et al. Whole genome sequencing identifies associations for nonsyndromic sagittal craniosynostosis with the intergenic region of bmp2 and noncoding rna gene linc01428. *Scientific Reports*, 14(1):8533, 2024. [1](#)
- [53] Marc Niethammer, Martin Reuter, Franz-Erich Wolter, Sylvain Bouix, Niklas Peinecke, Min-Seong Koo, and Martha E Shenton. Global medical shape analysis using the laplace-beltrami spectrum. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2007: 10th International Conference, Brisbane, Australia, October 29–November 2, 2007, Proceedings, Part I 10*, pages 850–857. Springer, 2007. [8](#)
- [54] Kuo Niu and Chao Tian. Zernike polynomials and their applications. *Journal of Optics*, 24(12):123001, 2022. [12](#)
- [55] Kuo Niu and Chao Tian. Zernike polynomials and their applications. *Journal of Optics*, 24(12):123001, 2022. [12](#)
- [56] Marcin Novotni and Reinhard Klein. 3d zernike descriptors for content based shape retrieval. In *Proceedings of the eighth ACM symposium on Solid modeling and applications*, pages 216–225, 2003. [24](#)
- [57] Marcin Novotni and Reinhard Klein. Shape retrieval using 3d zernike descriptors. *Computer-Aided Design*, 36(11):1047–1062, 2004. [11](#)
- [58] Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Matching 3d models with shape distributions. In *Proceedings international conference on shape modeling and applications*, pages 154–166. IEEE, 2001. [6](#)
- [59] Robert Osada, Thomas Funkhouser, Bernard Chazelle, and David Dobkin. Shape distributions. *ACM Transactions on Graphics (TOG)*, 21(4):807–832, 2002. [6](#), [7](#)
- [60] Martin Reuter, Franz-Erich Wolter, and Niklas Peinecke. Laplace–beltrami spectra as ‘shape-dna’ of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006. [9](#)
- [61] Martin Reuter, Franz-Erich Wolter, Martha Shenton, and Marc Niethammer. Laplace–beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis. *Computer-Aided Design*, 41(10):739–755, 2009. [8](#)
- [62] Lee Sael, Bin Li, David La, Yi Fang, Karthik Ramani, Raif Rustamov, and Daisuke Kihara. Fast protein tertiary structure retrieval based on global surface shape similarity. *Proteins: Structure, Function, and Bioinformatics*, 72(4):1259–1273, 2008. [11](#)

- [63] Matthias Schaufelberger, Christian Kaiser, Reinald Kühle, Andreas Wachter, Frederic Weichel, Niclas Hagen, Friedemann Ringwald, Urs Eisenmann, Jürgen Hoffmann, Michael Engel, et al. 3d-2d distance maps conversion enhances classification of craniosynostosis. *IEEE Transactions on Biomedical Engineering*, 70(11):3156–3165, 2023. 14
- [64] Matthias Schaufelberger, Reinald Kühle, Andreas Wachter, Frederic Weichel, Niclas Hagen, Friedemann Ringwald, Urs Eisenmann, Jürgen Hoffmann, Michael Engel, Christian Freudlsperger, et al. A radiation-free classification pipeline for craniosynostosis using statistical shape modeling. *Diagnostics*, 12(7):1516, 2022. 3, 14
- [65] Spiros Sgouros, J Henry Goldin, Anthony D Hockley, Michael JC Wake, and Kalyan Natarajan. Intracranial volume change in childhood. *Journal of neurosurgery*, 91(4):610–616, 1999. 2
- [66] Neda Stefanovic, HAKAN El, DL Chenin, B Glisic, and Juan Martin Palomo. Three-dimensional pharyngeal airway changes in orthodontic patients treated with and without extractions. *Orthodontics & craniofacial research*, 16(2):87–96, 2013. 2, 5
- [67] Hang Su, Subhransu Maji, Evangelos Kalogerakis, and Erik Learned-Miller. Multi-view convolutional neural networks for 3d shape recognition. In *Proceedings of the IEEE international conference on computer vision*, pages 945–953, 2015. 13
- [68] Hang Su, Subhransu Maji, Evangelos Kalogerakis, and Erik Learned-Miller. Multi-view convolutional neural networks for 3d shape recognition. In *Proceedings of the IEEE international conference on computer vision*, pages 945–953, 2015. 14
- [69] Jian Sun, Maks Ovsjanikov, and Leonidas Guibas. A concise and provably informative multi-scale signature based on heat diffusion. In *Computer graphics forum*, volume 28, pages 1383–1392. Wiley Online Library, 2009. 10
- [70] A Tartaro, AR Larici, D Antonucci, B Merlino, C Colosimo, and L Bonomo. Optimization and diagnostic accuracy of computerized tomography with tridimensional spiral technique in the study of craniostenosis. *La Radiologia Medica*, 96(1-2):10–17, 1998. 1, 2
- [71] Rodolphe Vaillant. Definition of the laplacian matrix for triangle meshes. <https://rodolphe-vaillant.fr/entry/101/definition-laplacian-matrix-for-triangle-meshes>. Accessed: July 18, 2024. 10
- [72] Vishwesh Venkatraman, Lee Sael, and Daisuke Kihara. Potential for protein surface shape analysis using spherical harmonics and 3d zernike descriptors. *Cell biochemistry and biophysics*, 54:23–32, 2009. 11
- [73] Vishwesh Venkatraman, Lee Sael, and Daisuke Kihara. Potential for protein surface shape analysis using spherical harmonics and 3d zernike descriptors. *Cell biochemistry and biophysics*, 54:23–32, 2009. 24
- [74] Steven A Wall, Gregory PL Thomas, David Johnson, Jo C Byren, Jayaratnam Jayamohan, Shailendra A Magdum, David J McAuley, and Peter G Richards. The preoperative incidence of raised intracranial pressure in nonsyndromic sagittal craniosynostosis is underestimated in the literature. *Journal of Neurosurgery: Pediatrics*, 14(6):674–681, 2014. 2
- [75] Xin Wei, Ruixuan Yu, and Jian Sun. View-gcn: View-based graph convolutional network for 3d shape analysis. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1850–1859, 2020. 15
- [76] Ari M Wes, Sanjay Naran, James Sun, Daniel Mazzaferro, Wen Xu, Phuong Nguyen, Linton A Whitaker, Scott P Bartlett, and Jesse A Taylor. The whitaker classification of craniosynostosis outcomes: an assessment of interrater reliability. *Plastic and reconstructive surgery*, 140(4):579e–586e, 2017. 2
- [77] Ari M Wes, Sanjay Naran, James Sun, Daniel Mazzaferro, Wen Xu, Phuong Nguyen, Linton A Whitaker, Scott P Bartlett, and Jesse A Taylor. The whitaker classification of craniosynostosis outcomes: an assessment of interrater reliability. *Plastic and reconstructive surgery*, 140(4):579e–586e, 2017. 2
- [78] Honglin Yuan and Remco C Veltkamp. Free-viewpoint image based rendering with multi-layered depth maps. *Optics and Lasers in Engineering*, 147:106726, 2021. 4
- [79] Dan Zhang, Na Liu, Zhongke Wu, and Xingce Wang. 3d craniofacial similarity calculation and craniofacial relationships analysis based on spectral analysis method. *Multimedia Tools and Applications*, 83(5):14063–14084, 2024. 8, 10
- [80] Dan Zhang, Zhongke Wu, Xingce Wang, Chenlei Lv, and Na Liu. 3d skull and face similarity measurements based on a harmonic wave kernel signature. *The Visual Computer*, 37:749–764, 2021. 9
- [81] Hao Zhang. Discrete combinatorial laplacian operators for digital geometry processing. In <https://rodolphe-vaillant.fr/entry/101/definition-laplacian-matrix-for-triangle-meshes>, pages 575–592, 2004. 10