

# Automated Classification of Skeletal Facial Asymmetry in CBCT Using a Reproducible 3D Slicer Workflow for Patient-Specific Decision Support

Angelina Sanna<sup>a</sup>, Janson Ng<sup>a</sup>, Antonio Ruellas<sup>b</sup>, Aron Aliaga<sup>a</sup>, Alban Gaydamour<sup>a</sup>, Enzo Tulissi<sup>a</sup>, Yanjie Huang<sup>c</sup>, Juan Carlos Prieto<sup>c</sup>, Najla Al Turkestani<sup>d</sup>, Marilla Yatabe<sup>a</sup>, Jonas Bianchi<sup>e</sup>, Steve Pieper<sup>f</sup>, Lucia Cividanes<sup>a,c</sup>, and Tengfei Li<sup>c</sup>

<sup>a</sup>University of Michigan, Ann Arbor, United States

<sup>b</sup>Federal University of Rio de Janeiro, Rio de Janeiro, Brazil

<sup>c</sup>University of North Carolina, Chapel Hill, United States

<sup>d</sup>King Abdulaziz University Faculty of Dentistry, Jeddah, Saudi Arabia

<sup>e</sup>University of the Pacific, San Francisco, United States.

<sup>f</sup>Isomics Incorporation, Cambridge, United States

## ABSTRACT

This study presents an automated pipeline for classifying skeletal facial asymmetry of 170 patients with Class III malocclusion using cone-beam computed tomography (CBCT). The methodology integrates standardized image preprocessing, mirroring, voxel-based registration, anatomical measurement extraction, dimensionality reduction, and clustering. Principal component analysis (PCA), t-distributed stochastic neighbor embedding (t-SNE), and uniform manifold approximation and projection (UMAP) were evaluated using Spectral Clustering with both k-nearest-neighbors (kNN) and radial basis function (RBF) affinities. PCA suggested four clusters, t-SNE found four with kNN and two with RBF, and UMAP identified six. UMAP with RBF affinity achieved the best overall clustering performance, as indicated by the silhouette score, C-index, Dunn index, and closed surface distance heat maps, slightly outperforming UMAP with kNN. The pipeline is implemented within open-source 3D Slicer Craniomaxillofacial automated modules to enhance reproducibility and clinical translation.

**Keywords:** Facial asymmetry, Cone-beam computed tomography (CBCT), 3D Slicer, Spectral clustering

## 1. INTRODUCTION

Facial asymmetry is a key clinical concern in orthodontics, affecting both appearance and functionality.<sup>1,2</sup> While minor asymmetries are common even in well-balanced faces,<sup>3</sup> marked skeletal asymmetry is often perceived as less attractive<sup>4,5</sup> and may motivate patients to seek orthodontic and/or surgical correction.<sup>6</sup> Facial asymmetry is particularly common in certain malocclusions, notably Class III cases characterized by a retrognathic maxilla, prognathic mandible, or both. In this group, chin asymmetry is detectable in up to 40% of patients,<sup>7</sup> with deviations often involving the mandible, maxilla, and related cranial structures. Through analyzing data from the Dentofacial Clinic at the University of North Carolina between 2006 and 2010, Proffit et al.<sup>8</sup> found that approximately 36% of all surgical patients underwent asymmetry correction. Accurate classification of asymmetry is therefore critical for diagnosis, treatment planning, and outcome assessment.<sup>1,9</sup>

Traditionally, postero-anterior cephalograms have been used to diagnose and plan treatment for facial asymmetry;<sup>10</sup> however, these two-dimensional methods can be limited in their ability to characterize the high-dimensional problem. Three-dimensional imaging, particularly cone-beam computed tomography (CBCT), now allows for volumetric analysis of craniofacial structures, allowing for more accurate identification of asymmetry and its regional distribution.<sup>11</sup> Recent advancements in artificial intelligence (AI) and machine learning, including unsupervised learning methods, allow for the detection of subgroups with similar phenotypic patterns that

---

Further author information: (Send correspondence to Angelina Sanna)  
: E-mail: asanna@andrew.cmu.edu

may not be apparent through conventional analysis.<sup>12</sup> While AI can offer valuable insights, models are not 100% accurate and cannot replace clinical expertise; therefore, professional interpretation by a qualified orthodontist/oral and maxilla facial surgeon is required to ensure high-quality care. The 3D Slicer<sup>13</sup> Craniomaxillofacial automated suite combines validated AI-based modules for head orientation, voxel-based registration at the cranial base, maxilla and mandible linear and angular quantification, surface distance mapping, and multi-method clustering approaches. Leveraging this standardized workflow enabled objective analysis of our dataset and the translation of results into precise, patient-specific facial asymmetry diagnoses.

The objectives of this study were as follows. First, to develop a reproducible, fully automated workflow for processing and analyzing CBCT images of Class III patients with and without facial asymmetry that is both clinician-facing and extensible across multimodal data. Second, to compare the performance of four clustering approaches in their ability in distinguishing patterns of facial asymmetry. Finally, to visualize regional asymmetry patterns using closed surface distance heat maps to facilitate clinical interpretation.

## 2. METHODS

### 2.1 Data Preprocessing

We analyzed CBCT scans from 170 patients diagnosed with Class III malocclusion. All images were anonymized and resampled to an isotropic voxel size of 0.3 mm. A standardized head orientation was established relative to the cranial base for assessment of cranial base and mandibular asymmetry, and relative to the occlusal plane for maxillary asymmetry assessment. Scans were mirrored to generate contralateral symmetry references and subsequently registered using automated voxel-based registration for the cranial base, mandible, and maxilla.

All image processing steps were performed in 3D Slicer<sup>13</sup> using automated extensions to ensure standardization and reproducibility. Cranial-base superimposition for T2–T1 (original–mirrored) comparisons were conducted with Automatrix and Areg’s voxel-based registration, using the cranial base as the stable reference to isolate maxillary and mandibular changes. AQ3DC extracted linear distances and pitch–yaw–roll angular metrics from predefined landmarks and planes, and Automated ModelToModelDistance computed dense surface maps visualized in ShapePopulationViewer. This workflow mirrors published CMF user-study protocols,<sup>14–16</sup> enabling transparent reuse and clinical deployment.

### 2.2 Feature Extraction and Dimensionality Reduction

We extracted 34 quantitative linear distances, volumetric indices, and angular deviations from the registered volumes, covering the cranial base, mandible, maxilla. Dimensionality reduction was applied to the measurements using three methods: UMAP, t-SNE, and PCA. UMAP and t-SNE were implemented as non-linear reduction techniques, while PCA served as a linear baseline. For PCA, the top ten principal components were retained to preserve the majority of the variance, and the explained variance ratio for each component was recorded. For visualization, three-dimensional scatter plots were generated from the reduced feature spaces: PC1–PC3 for PCA, and the first three embedding dimensions for UMAP and t-SNE. This allowed qualitative assessment of group separation alongside quantitative evaluation using silhouette score, C-index, and Dunn index.<sup>17,18</sup>

### 2.3 Clustering Approaches

Initially, we tested hierarchical and k-means approaches, then selected spectral clustering, which partitions data by analyzing the eigenstructure of an affinity matrix.<sup>19</sup> This method captures complex, non-linear relationships, making it well suited for accommodating irregular cluster shapes common in facial asymmetry phenotypes. Two affinity constructions of spectral clustering were evaluated: k-nearest-neighbors (kNN) graphs, which connect each sample to its closest neighbors to capture local structure,<sup>20</sup> and Radial Basis Function (RBF) kernels, which compute pairwise similarities based on distance, producing a fully connected graph with weights decaying over distance.<sup>21</sup> The choice of affinity impacts the clusters formed, with kNN graphs emphasizing local relationships and RBF kernels based on global similarity patterns.

For each combination, the number of clusters k was selected by maximizing the silhouette score computed in the same space used for clustering, and then evaluated with C-index and Dunn index along with qualitative inspection of 3D low-dimensional plots. Closed surface distance heat maps from representative cases were also generated to visualize and interpret regional asymmetry patterns.

### 3. RESULTS

#### 3.1 Model Selection Across Affinities

In the PCA space, both RBF and kNN produced consistent results, selecting  $k=4$  as the optimal cluster count (Table 1). The t-SNE embeddings produced different outcomes depending on the affinity: kNN favored  $k=4$ , while RBF selected  $k=2$ . In the UMAP space, both RBF and kNN agreed on  $k=6$  as optimal.

#### 3.2 Cluster Quality Metrics

Cluster quality was assessed using the silhouette score, C-index, and Dunn index for each manifold-affinity- $k$  combination (Table 1). In PCA space, silhouette scores were 0.071 for RBF and 0.070 for kNN, indicating minimal separation, with RBF performing marginally better. The t-SNE kNN configuration with  $k = 4$  achieved a silhouette score of 0.245, substantially higher than the t-SNE RBF score of -0.005 and both PCA results. UMAP with  $k = 6$  in both affinities consistently achieved the lowest C-index (0.047–0.048) and relatively high Dunn indices (0.133–0.135), suggesting superior compactness and separation compared to PCA and t-SNE. UMAP RBF yielded the highest silhouette score (0.411), slightly outperforming UMAP kNN (0.405), while also achieving the absolute minimum C-index of 0.047. The highest Dunn index overall was observed in PCA kNN (0.505), though its low silhouette score indicates that clusters were far apart but internally inconsistent.

Method	Number of Clusters	Silhouette Score	C-index	Dunn-index
t-SNE w/ kNN	4	.245	.169	.119
t-SNE w/ RBF	2	-.005	.506	.091
UMAP w/ kNN	6	.405	.048	.135
UMAP w/ RBF	6	.411	.047	.133
PCA w/ kNN	4	.070	.231	.505
PCA w/ RBF	4	.071	.227	.400

Table 1. Cluster quality metrics ranked by silhouette score, C-index, and Dunn index. Higher silhouette and Dunn indices indicate better separation, whereas lower C-index values indicate greater within-cluster homogeneity. The top-performing methods were t-SNE with kNN ( $k = 4$ ), UMAP with kNN ( $k = 6$ ), and UMAP with RBF ( $k = 6$ ).

Visual inspection of the embeddings confirmed the quantitative results (Figure 1). In PCA plots, clusters showed considerable overlap. The t-SNE kNN embedding produced visually compact, well-separated clusters, aligning with the improved metric scores and suggesting stronger subgroup definition. In contrast, the t-SNE RBF embedding collapsed data into broader partitions, likely obscuring clinically meaningful substructures.

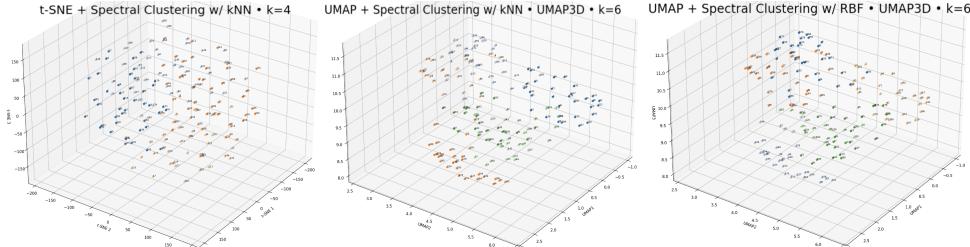


Figure 1. Top three clustering configurations plotted in three dimensions. Cluster assignments are color-coded, illustrating compactness and separation achieved by each method.

### 4. DISCUSSION

#### 4.1 Visual Assessment and Clinical Interpretability

UMAP embeddings exhibited the clearest separation (Figure 2), where Cluster 3 corresponded to no jaw asymmetry and Clusters 0-2 and 4-5 showed varying locations and degrees of asymmetry. From a clinical perspective, the greater separation in UMAP space may facilitate identification of meaningful patient categories, whereas PCA-derived clusters may require additional clinical validation.

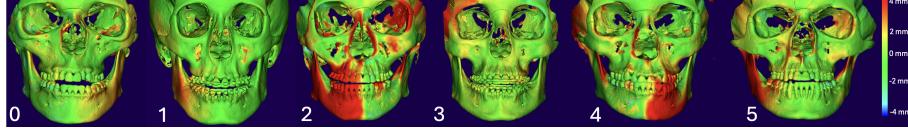


Figure 2. Heatmaps visualizing patients in each cluster obtained using UMAP with RBF affinity. Cluster 3 represents patients with no jaw asymmetry, Clusters 0-2 and 4-5 correspond to varying locations and degrees of asymmetry

#### 4.2 Methodological Insights

Our results demonstrate that the choice of embedding space had a greater impact on cluster quality than the affinity measure used within spectral clustering.

When retaining the top 10 principal components, PCA explained only  $\sim 60\%$  of the dataset's variance. The high Dunn index for PCA with kNN likely reflects large inter-cluster distances driven by a few widely separated clusters, but the low silhouette score indicates that these clusters were internally inconsistent. This limitation may be due to the inability of linear projections to reveal nonlinear relationships critical for subgroup differentiation in datasets with a large number of correlated measurements.<sup>22</sup> t-SNE with kNN produced more coherent clusters than PCA, as reflected by its higher silhouette score. However, its weaker performance with RBF affinity and lower Dunn index values compared to UMAP suggests that while t-SNE excels at identifying fine-grained local structures, it may not maintain global separability as effectively, which can reduce cluster compactness at larger scales. UMAP provided compact, well-separated clusters with high Dunn index values and low C-index scores, suggesting intra-cluster homogeneity and inter-cluster separation. This superior performance likely stems from its ability to balance local and global structure preservation, enabling it to capture clinically relevant subgroups in high-dimensional asymmetry data.

The RBF and kNN affinity matrices' relative performance varied depending on the embedding. In t-SNE, which preserves local structure strongly, kNN worked better. Both affinities work equally well in UMAP, as it balances local and global structure. In a linear embedding (PCA), affinity choice had little affect because the embedding itself limited performance. This highlights the interaction between manifold structure and affinity computation and underscores the need for systematic parameter tuning when designing clustering pipelines.

#### 4.3 Clinical Relevance and Future Directions

The clearer separation achieved with the nonlinearity of UMAP-based spectral clustering offers meaningful potential for clinical application. This approach could support the development of automated, phenotype-driven tools for diagnosis, treatment planning, and prognosis in cases of facial asymmetry. Well-defined clusters may represent clinically distinct subtypes with unique etiologies, therapeutic needs, and treatment outcomes, thereby enabling more personalized and effective clinical planning.

Despite these encouraging results, the analysis was conducted on a relatively small dataset, and the clustering outcomes showed sensitivity to parameters such as neighborhood size in manifold learning and scale selection in affinity functions. Future research should apply UMAP's superior capacity for global separation to larger, multi-center datasets. The integration of complementary data sources, such as 3D imaging, clinical records, and genetic information, could further strengthen the interpretability of the resulting clusters. Refining this approach to produce consistently well-separated and clinically meaningful subgroups has the potential to advance diagnosis precision, ultimately improving treatment outcomes for patients with craniofacial asymmetry.

### 5. CONCLUSION

Across all manifold-affinity combinations, UMAP consistently produced the highest-quality groupings, with the RBF affinity achieving the best overall performance. The t-SNE method with kNN formed compact, locally coherent clusters but showed less separation overall, while PCA-based clustering obtained low silhouette scores, despite achieving the highest Dunn index with kNN. These findings underscore the advantages of nonlinear dimensionality reduction, particularly UMAP, for distinguishing clinically relevant patterns in complex craniofacial shape data. Clinically, the optimized UMAP-RBF workflow offers a reproducible and objective framework for distinguishing asymmetry phenotypes in Class III patients, enabling more precise, personalized diagnosis and treatment planning to improve surgical outcomes.

## REFERENCES

- [1] Proffit, W. R., Fields, H. W., Larson, B. E., and Sarver, D. M., *Contemporary Orthodontics*, Sixth ed., Elsevier (2019).
- [2] Bishara, S. E., Burkey, P. S., and Kharouf, J. G., “Dental and facial asymmetries: a review,” *Angle Orthod.* **64**(2), 89–98 (1994).
- [3] Peck, S., Peck, L., and Kataja, M., “Skeletal asymmetry in esthetically pleasing faces,” *Angle Orthod.* **61**(1), 43–48 (1991).
- [4] Meyer-Marcotty, P., Stellzig-Eisenhauer, A., Bareis, U., Hartmann, J., and Kochel, J., “Three-dimensional perception of facial asymmetry,” *Eur. J. Orthod.* **33**(6), 647–653 (2011).
- [5] Jarosz, K. F., Bosio, J. A., Bloomstein, R., Jiang, S. S., Vakharia, N. S., and Cangialosi, T. J., “Perceptions of chin asymmetries among dental professionals and laypersons,” *Am. J. Orthod. Dentofacial Orthop.* **154**(2), 201–212 (2018).
- [6] Harun, N. A., Adam, K. B. C., Abdullah, N. A., and Rusli, N., “Is a symmetrical face really attractive?,” *Int. J. Oral Maxillofac. Surg.* **52**(6), 703–709 (2023).
- [7] Severt, T. R., and Proffit, W. R., “The prevalence of facial asymmetry in the dentofacial deformities population at the University of North Carolina,” *Int. J. Adult Orthodon. Orthognath. Surg.* **12**(3), 171–176 (1997).
- [8] Proffit, W. R., Phillips, C., and Dann, C., “Who seeks surgical-orthodontic treatment?,” *Int. J. Adult Orthodon. Orthognath. Surg.* **5**(3), 153–160 (1990).
- [9] Burstone, C. J., “Diagnosis and treatment planning of patients with asymmetries,” *Semin. Orthod.* **4**(3), 153–164 (1998).
- [10] Grummons, D. C., and Kappeyne van de Coppello, M. A., “A frontal asymmetry analysis,” *J. Clin. Orthod.* **21**(7), 448–465 (1987).
- [11] Liu, L. C., Chen, Y. A., Chen, R. F., Yao, C. F., Liao, Y. F., and Chen, Y. R., “Type of mandibular asymmetry affects changes and outcomes of bimaxillary surgery for class III asymmetry,” *Clin. Oral Investig.* **26**(1), 1077–1088 (2022).
- [12] Raza, K., and Singh, N. K., “A tour of unsupervised deep learning for medical image analysis,” *arXiv preprint arXiv:1812.07715* (2018).
- [13] Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.-C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F. M., Sonka, M., Buatti, J., Aylward, S. R., Miller, J. V., Pieper, S., and Kikinis, R., “3D Slicer as an image computing platform for the Quantitative Imaging Network,” *Magn. Reson. Imaging* **30**(9), 1323–1341 (2012).
- [14] Bianchi, J., Ruellas, A., Gonçalves, J. R., Nguyen, T., Paniagua, B., Prieto, J. C., Styner, M., and Cevidanes, L. H. S., “3D Slicer Craniomaxillofacial modules support patient-specific decision-making for personalized healthcare in dental research,” *Sci. Rep.* **10**, 123 (2020).
- [15] S. Bencherqui, S. Barone, L. Cevidanes, J.-P. Perrin, P. Corre, and H. Bertin, “3D analysis of condylar and mandibular remodeling one year after intra-oral ramus vertical lengthening osteotomy,” *Clin. Oral Investig.* **28**(1), 114 (2024).
- [16] Anchling, N. Hulin, Y. Huang, S. Barone, S. Roberts, F. Miranda, M. Gurgel, N. Al Turkestani, S. Tinawi, J. Bianchi, M. Yatabe, A. Ruellas, J. C. Prieto, and L. Cevidanes, “Automated orientation and registration of cone-beam computed tomography scans,” in *Clin. Image Based Proced., Fairness AI Med. Imaging, Ethical Philos. Issues Med. Imaging, Proc. SPIE 14242*, 43–58 (2023).
- [17] Renjith, S., “Evaluation of clustering algorithms using Silhouette Index and Dunn Index,” *Acta Sci. Technol.* **44**, e58653 (2022).
- [18] Hubert, L. and Levin, J. R., ”A general statistical framework for assessing categorical clustering in free recall,” *Psychol. Bull.* **83**, 1072–1080 (1976).
- [19] von Luxburg, U., “A tutorial on spectral clustering,” *Stat. Comput.* **17**(4), 395–416 (2007).
- [20] Altman, N. S., “An introduction to kernel and nearest-neighbor nonparametric regression,” *Am. Stat.* **46**(3), 175–185 (1992).
- [21] Schölkopf, B., Smola, A. J., and Müller, K. R., “Nonlinear component analysis as a kernel eigenvalue problem,” *Neural Comput.* **10**(5), 1299–1319 (1998).

- [22] Cunningham, J. P., and Ghahramani, Z., “Linear dimensionality reduction: Survey, insights, and generalizations,” *J. Mach. Learn. Res.* **16**, 2859–2900 (2015).