

Developmental Geometry in Facial Embeddings: A Boundary-Layer Transition

Supplementary Material

8. Appendix. Additional Results

This appendix provides full supporting statistics referenced in Section 4. These tables expand on local separability, centroid geometry, mixing structure, and scale sensitivity across encoders.

8.1. Full kNN Purity Results

Table 7 reports per-bin kNN purity ($k = 10$) across developmental bins for UTKFace and FG-NET. Across both datasets, early childhood (0–5) consistently forms the most compact region of the embedding space, while adolescence (13–17) exhibits the lowest local separability. Purity statistics on FG-NET follow exhibit higher variance due to smaller sample size. We include them for completeness.

8.2. Centroid Distance Structure

Table 8 reports pairwise centroid distances for CLIP ViT-L/14 on UTKFace. Distances exhibit ordered progression across developmental bins, with adolescence positioned between late childhood and adulthood.

8.3. kNN Mixing Structure

Table 9 reports the kNN mixing matrix ($k=10$) for CLIP ViT-L/14. Rows sum to 1. Adolescence mixes substantially with both 6–12 and 18–40 bins, while early childhood remains highly self-contained.

8.4. Image-Scale Drift

Table 10 reports mean embedding drift under scale perturbations ($\alpha \in \{0.8, 1.2\}$). Drift increases modestly with age but does not exhibit a transition spike.

8.5. Cross-Encoder Replication Summary

Across all encoders and datasets:

- Early childhood consistently exhibits highest local separability and largest margins.
- Adolescence consistently exhibits lowest margins and highest perturbation fragility.
- Local intrinsic dimensionality increases toward adulthood.
- Scale sensitivity increases gradually but does not display a transition spike.

These consistent patterns across architectures and corpora support the interpretation of adolescence as a boundary-adjacent transition regime in facial embedding space.

Table 7. kNN purity ($k = 10$) across developmental bins on UTKFace and FG-NET.

Model	UTKFace					FG-NET				
	0-5	6-12	13-17	18-40	40+	0-5	6-12	13-17	18-40	40+
CLIP ViT-L/14	0.926	0.551	0.315	0.873	0.754	0.749	0.551	0.393	0.708	0.481
OpenCLIP ViT-H/14	0.926	0.582	0.332	0.871	0.775	0.777	0.574	0.403	0.703	0.543
OpenCLIP ViT-B/16	0.922	0.545	0.315	0.871	0.754	0.752	0.574	0.348	0.676	0.495
OpenCLIP ViT-L/14 (DataComp)	0.912	0.508	0.271	0.869	0.731	0.718	0.532	0.351	0.694	0.502

Table 8. Centroid distance matrix (CLIP ViT-L/14, UTKFace).

	0-5	6-12	13-17	18-40	40+
0-5	0.000	0.283	0.371	0.436	0.496
6-12	0.283	0.000	0.184	0.372	0.433
13-17	0.371	0.184	0.000	0.283	0.372
18-40	0.436	0.372	0.283	0.000	0.242
40+	0.496	0.433	0.372	0.242	0.000

Table 9. kNN mixing matrix (CLIP ViT-L/14, UTKFace).

Source \rightarrow Target	0-5	6-12	13-17	18-40	40+
0-5	0.926	0.062	0.004	0.006	0.002
6-12	0.246	0.551	0.151	0.050	0.002
13-17	0.031	0.312	0.315	0.334	0.008
18-40	0.002	0.005	0.019	0.873	0.101
40+	0.001	0.000	0.001	0.244	0.754

Table 10. Mean image-scale drift (UTKFace).

Model	Child (<18)	Adult (≥ 18)
CLIP ViT-L/14	0.399	0.412
OpenCLIP ViT-H/14	0.394	0.396
OpenCLIP ViT-B/16	0.373	0.383
OpenCLIP ViT-L/14 (DataComp)	0.386	0.400