MBRCNet: Multi-View Breast Reconstruction and Classification Network

EUROGRAPHICS AACHEN 2026

Yan Pang, Rubi Quiñones
Computer Science Department, School of Engineering
Southern Illinois University Edwardsville, USA

Introduction

Accurate 3D reconstruction of breast morphology remains challenging in clinical settings. Clinical imaging methods such as CT and MRI provide detailed anatomy but are costly and not verified for a variety of breast shapes. Breast shape variation is clinically relevant in applications such as surgical planning and lactation-related morphology assessment. Our framework addresses these challenges with dual reconstruction paths for different supervision settings and uses reconstructed 3D anatomy as a basis for downstream morphology analysis. Reconstruction therefore serves both as the primary task and as a shape representation for downstream grouping.

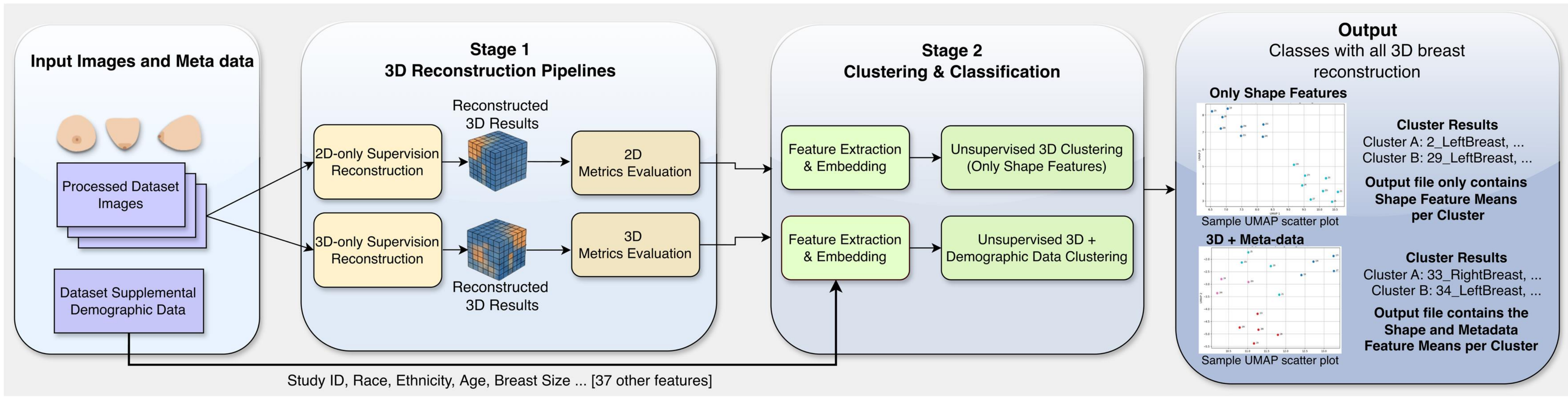

Figure1: *Block diagram of the proposed MBRCNet framework.*

METHODOLOGY

Input and Preprocessing: Dataset. The dataset contains 50 human participants and 100 breast instances, each with three calibrated orthographic views (front, side, top), split into 70/14/16 train/validation/test samples. Ground truth 3D occupancy is provided as 64^3 binary voxel grids, and 2D supervision is provided as binary segmentation masks derived from the same views. Segmented multi-view images are used as input to isolate the target anatomy. For morphology analysis, each reconstructed instance is additionally represented by geometric shape descriptors and, where available, participant metadata.

As shown in Figure1, the model contains two reconstruction pipelines followed by a morphology classification stage operating on reconstructed 3D shape features.

Stage 1: 3D Reconstruction Pipelines. The 2D pipeline uses a pretrained VGG-16 backbone to extract per-view feature maps. A Merger module fuses multi-view features via 3D convolutions and softmax attention, and a Refiner module sharpens the result through a 3D encoder-decoder with skip connections[XMZ*19]. Training uses Binary Cross-Entropy, Dice, and L1 projection losses combined with volumetric regularizers for smoothness, connectivity, and solid interior, requiring no 3D annotations.

The 3D-supervised pipeline encodes each view with ResNet-50, fuses multi-view features with a lightweight Transformer encoder into a latent combining instance and attribute embeddings, and decodes it into a 64^3 voxel volume refined by a geometry block. Auxiliary supervision is provided through attribute regression, voxel statistics, and implicit surface heads. The resulting reconstructed volume and latent shape representation are then used as inputs for downstream morphology classification.

Stage 2: Clustering and Classification. Following reconstruction, the framework performs morphology classification through two routes. First, an unsupervised path groups reconstructed breasts into shape-consistent clusters for exploratory morphology discovery. Second, a clustering path uses the same reconstructed shape descriptors combined with demographic features to discover morphology groupings through unsupervised cluster assignments. The 3D-only path extracts 16 hand-crafted geometric features from each reconstructed volume, including volume, surface area, sphericity, moments of inertia, projections, symmetry scores, and aspect ratios, combined with Principal Component Analysis components retaining 95% of raw voxel variance. The 3D + demographic path fuses the same 3D features with participant metadata using a 50/50 energy normalization. Both paths reduce features via UMAP and cluster with HDBSCAN, with KMeans as fallback, selecting the optimal cluster size by silhouette score. Both routes output per-cluster mean voxel shapes, medoid representatives, and feature statistics.

RESULTS

2D Pipeline(Shown in Figure2). Our predicted masks closely resemble the ground truth silhouettes across all three views in both shape and boundary smoothness. The 2D pipeline achieves a mean IoU of 0.776 and HD95 of 5.193, substantially outperforming CapNet [NPAB19]. These silhouettes provide stable view-consistent constraints for the downstream 3D reconstruction.

3D Pipeline(Shown in Figure3). Our method produces compact, smooth volumes that generally align with ground truth morphology, outperforming Pix2Vox[XYS*19] in volumetric IoU (0.383 vs. 0.292) and Chamfer Distance (0.001742 vs. 0.00241). Pix2Vox recovers the approximate shape but exhibits block artefacts from its coarser 64^3 resolution. Point-E[NJD*22] generates unstructured, scattered point clouds with no coherent anatomical boundary. These results indicate that MBRCNet recovers more faithful instance-specific geometry, which is important not only for reconstruction accuracy but also for subsequent morphology classification.

Classification(Shown in Figure4). Morphology analysis was conducted on reconstructions from the 3D-supervised pipeline. In the shape-only setting, the test set separated into two clusters (8 samples each; silhouette = 0.64), primarily differing in sphericity and surface symmetry. Adding demographic features produced four smaller clusters (3 to 5 samples each) with a lower silhouette score (0.46). Demographic features were assessed using ANOVA and chi-square tests ($p < 0.05$), with height ($p = 0.004$) and delivery method ($p = 0.005$) showing the strongest association with cluster membership.

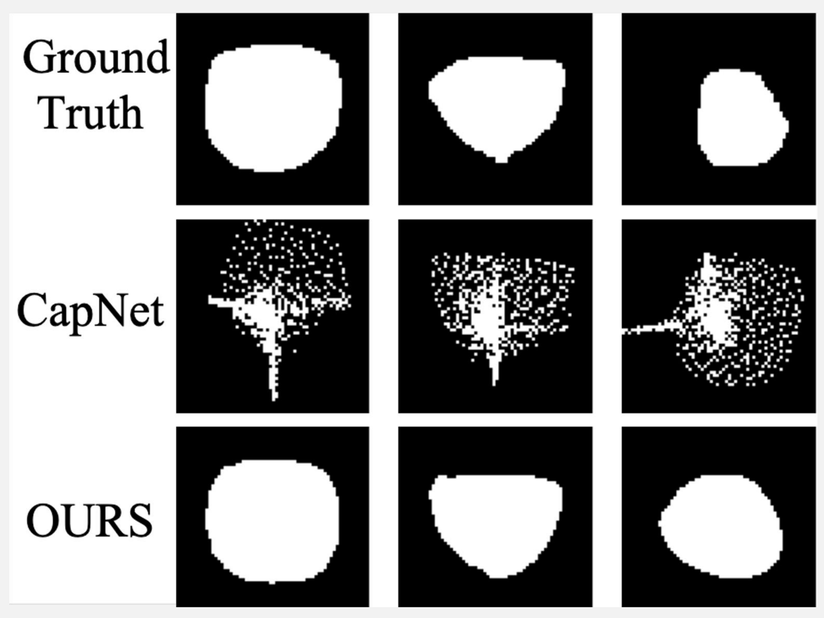

Figure2: *Qualitative comparison of 2D projection masks. Rows: Ground Truth, CapNet, Ours; columns: Front, Top, Side views.*

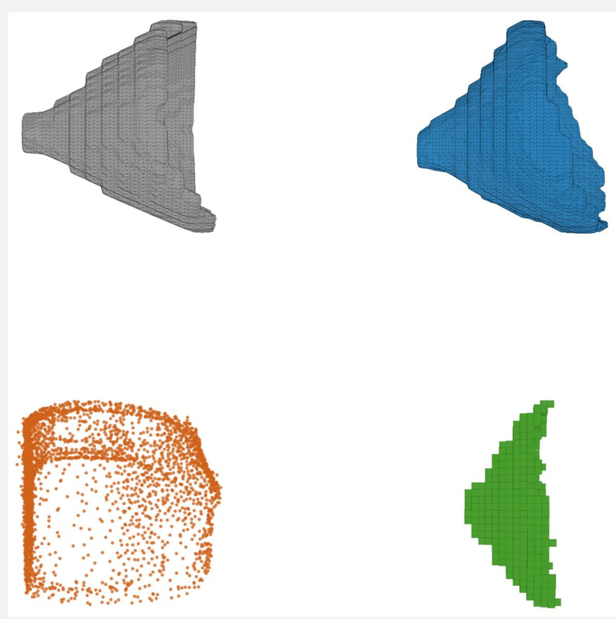

Figure3: *Qualitative 3D reconstruction results across three test samples. Top-left: Ground Truth; Top-right: Ours; Bottom-left: Point-E; Bottom-right: Pix2Vox.*

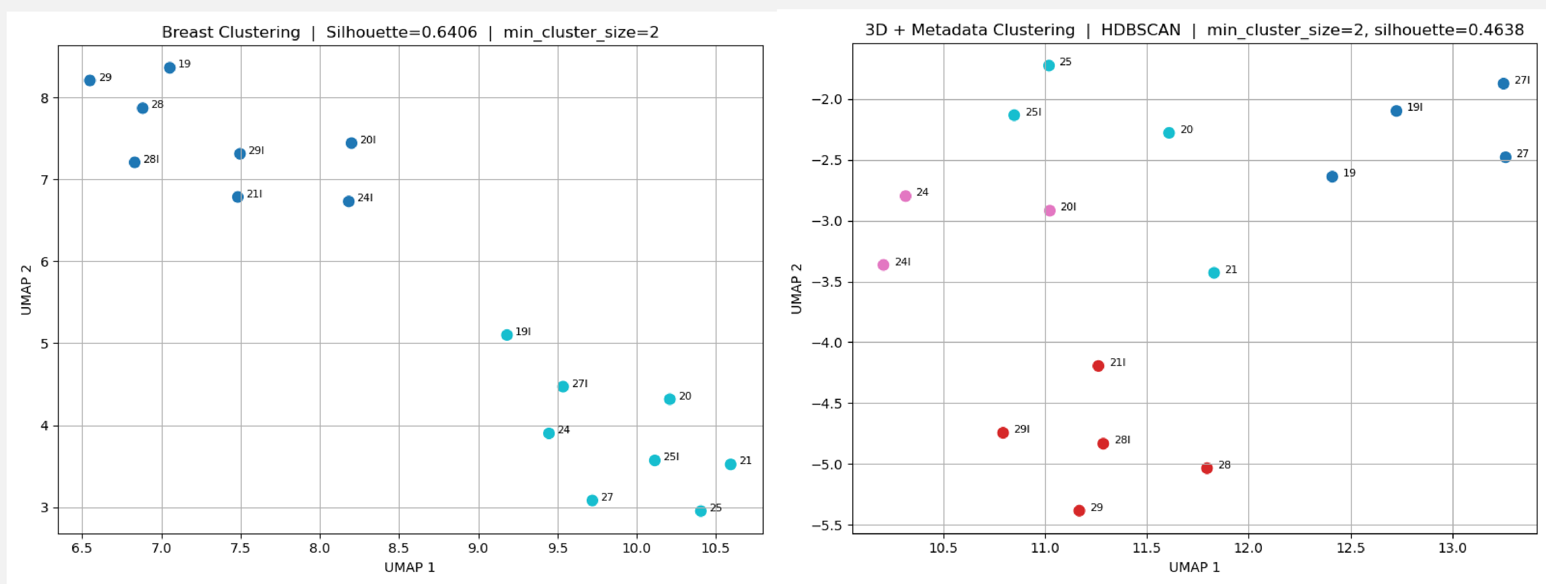

Figure4: UMAP visualisation of clustering results using 3D shape descriptors only (left, 2 clusters, silhouette = 0.64) and combined with demographic features (right, 4 clusters, silhouette = 0.46).

Conclusion

We presented a dual-pipeline framework for 3D breast reconstruction from multi-view RGB images. Both pipelines shape fidelity over relevant baselines, supporting RGB-based reconstruction as a promising direction for morphological analysis. Preliminary clustering results further suggest that reconstructed 3D shape representations may support downstream morphology grouping. These exploratory finding should be interpreted with caution given the small sample size potential confounding. Future work will focus on robustness under limited data and validation on larger cohorts.

AFFILIATIONS

SIUE NIH USDA

This work was supported by the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number R41HD115477. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

[NJD*22] NICHOL A., JUN H., DHARIWAL P., MISHKIN P., CHEN M., RADFORD A., SUTSKEVER I., JAIN P.: Point-e: A system for generating 3d point clouds from complex prompts. arXiv preprint arXiv:2212.08751 (2022).

[NPAB19] NAVANEET L. K., PRIYANKA M., AGARWAL M., BABU R. V.: Capnet: Continuous approximation projection for 3d point cloud reconstruction using 2d supervision. In Proceedings of the AAAI conference on artificial intelligence (2019), vol. 33, pp. 8819–8826.

[XYS*19] XIE H., YAO H., SUN X., ZHOU S., ZHANG S.: Pix2vox: Context-aware 3d reconstruction from single and multi-view images. In Proceedings of the IEEE/CVF International Conference on Computer Vision (2019), pp. 2690–2698.