

Metaball-Trained 3D Bubble Reconstruction CNN: Cross-Validation via a Physics-Based Dataset (BubbleSH) and Encoder Structure Comparison

Ilhwan Jang, Sangoh Shin, Minseop Song*

Department of Nuclear Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea

*Corresponding author: hysms@hanyang.ac.kr

***Keywords:** 3D shape reconstruction; Convolutional Neural Network; Two-orthogonal-view imaging; Bubble morphology; Physics-based validation

1. Introduction

Bubble morphology, including volume and interfacial area, directly affects thermal-hydraulic parameters such as the critical heat flux (CHF) in nuclear reactor safety analysis [1]. Conventional measurement techniques are either invasive, such as wire-mesh sensors [2] or optical fiber probes, which offer high measurement accuracy but disturb the flow field being measured [3]. Non-intrusive optical approaches — most notably the 2-view technique, which orthogonally positions two cameras — have therefore been adopted as a standard approach due to their minimal spatial requirements [4].

A prior study [4] proposed a data-driven CNN framework reconstructing 3D bubble shapes from two orthogonal (front/side) 2D silhouettes, using an anisotropic Metaball algorithm to generate training data without geometric assumptions. The model was validated only on Metaball-generated test samples — the same geometric family used for training — leaving its reliability on physically realistic shapes unverified.

Two gaps remain, however. First, single- and dual-encoder architectures need to be trained under an identical data volume (200,000 samples), so that the effect of encoder structure can be isolated from the confound of unequal training data volumes present in prior architecture comparisons. Second, the trained models should be cross-validated against physically realistic geometry, rather than only against the same Metaball family used for training. BubbleSH, a physics-based dataset in which bubble surfaces are reconstructed via Spherical Harmonics (SH) representation fitted to Direct Numerical Simulation (DNS) results, was used here as a geometrically independent validation reference for this purpose.

In summary, this study provides a controlled comparison of the two encoder architectures under matched training conditions, the first cross-validation of the framework against a physics-based dataset rather than the self-referential Metaball-to-Metaball scheme used previously, and a camera-calibrated physical quantification of volume and surface area for both validation sets.

2. Methodology

2.1 Overall Framework

The overall framework follows three stages established in [4]: synthetic dataset generation via the anisotropic Metaball algorithm, CNN training on 2-view

silhouette/3D voxel pairs, and evaluation against independent test sets.

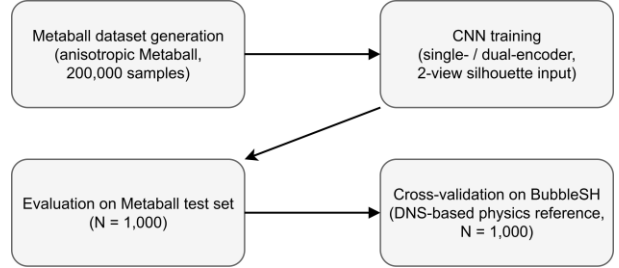


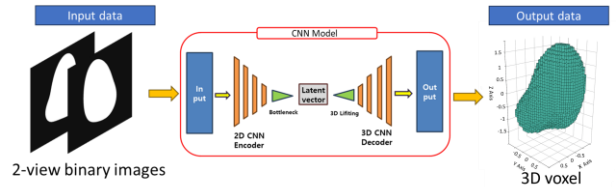
Fig. 1. Overall framework of the 3D bubble reconstruction and validation pipeline.

2.2 CNN Architecture: Single- and Dual-Encoder

Two encoder-decoder CNN architectures were compared to isolate the effect of view-fusion strategy on reconstruction accuracy. Both models take a pair of 128×128 binary silhouettes (front and side view) as input and output a $128 \times 128 \times 128$ binary voxel grid representing the reconstructed bubble volume.

The single-encoder model concatenates the front and side images into a single $2 \times 128 \times 128$ tensor processed by one shared 2D convolutional encoder, following [4]. The dual-encoder model instead processes each view through an independent convolutional branch, fusing the resulting latent vectors through a projection layer, relaxing the implicit assumption that the two orthogonal views are statistically interchangeable.

Unlike the comparison in [5], in which the two architectures were trained on datasets of different sizes, both models in this study were trained on an identical 200,000-sample dataset generated by the anisotropic Metaball algorithm, isolating the effect of the encoding strategy itself from that of training data volume.



(a)

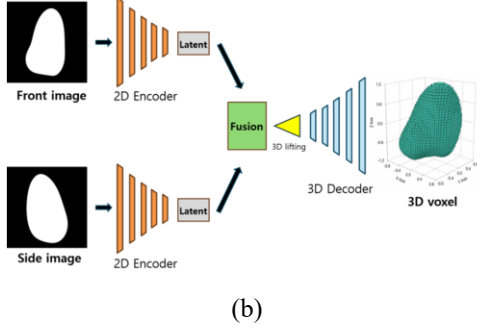


Fig. 2. Architectures of (a) the single-encoder and (b) the dual-encoder reconstruction models.

2.3 Physics-Based Validation Dataset (BubbleSH)

To evaluate the trained models beyond the Metaball geometric family, a physics-based validation dataset (BubbleSH) was constructed from Direct Numerical Simulation (DNS) results of rising bubble swarms[6]. Bubble surfaces at each time step were represented via a truncated Spherical Harmonics (SH) expansion up to degree $\ell = 14$ (225 coefficients per bubble), from which ground-truth volume and surface area were computed directly. The resulting mesh was voxelized on the same $128 \times 128 \times 128$ grid used for network input and output, and orthogonal front/side silhouettes were rendered using camera parameters matched to the physical imaging setup used for the Metaball dataset.

The validation set spans equivalent diameters of 4–6 mm and gas-fraction conditions of 5–40%. Each configuration simulates a swarm of 32 interacting bubbles, and samples were drawn from individual bubbles at randomly selected time steps across these swarm trajectories, spaced apart in time to avoid autocorrelation between samples. A total of $N = 1,000$ BubbleSH samples were used, matching the size of the independent Metaball test set. Fig. 3 shows a representative BubbleSH sample together with its voxelized ground truth and rendered input silhouettes.

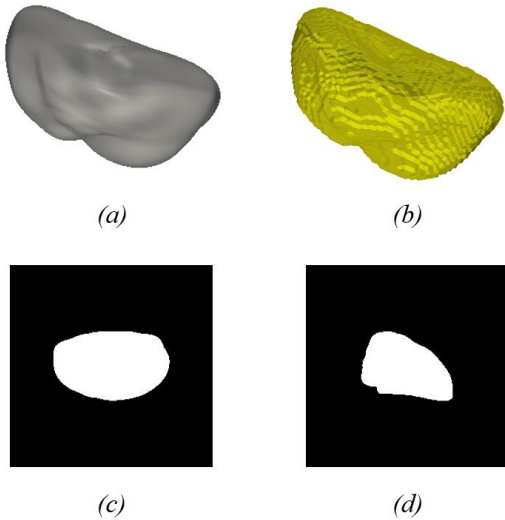


Fig. 3. Representative BubbleSH samples: (a) SH-fitted mesh, (b) voxelized ground truth, (c) rendered front silhouette, (d) rendered side silhouette.

2.4 Quantitative Evaluation Metrics

For samples with a known 3D ground truth, reconstruction accuracy was quantified via the Dice coefficient and Intersection-over-Union (IoU) between the predicted and ground-truth voxel grids, as defined in Eqs. (1) and (2):

$$\text{Dice} = \frac{2|V_{\text{pred}} \cap V_{\text{GT}}|}{(|V_{\text{pred}}| + |V_{\text{GT}}|)} \quad (1)$$

$$\text{IoU} = \frac{|V_{\text{pred}} \cap V_{\text{GT}}|}{|V_{\text{pred}} \cup V_{\text{GT}}|} \quad (2)$$

Bubble volume and surface area were computed from the reconstructed mesh in physical units (mm^3 , mm^2) using a camera-calibrated voxel spacing, with relative error defined as $|Q_{\text{pred}} - Q_{\text{GT}}| / Q_{\text{GT}} \times 100\%$. Because the Metaball and BubbleSH datasets share the same $128 \times 128 \times 128$ voxel grid and use matched camera parameters, as noted in Section 2.3, results from the two datasets are directly comparable without additional scale normalization.

3. Results and Discussion

3.1 Reconstruction Performance on the Independent Metaball Set

Both models, trained on an identical 200,000-sample Metaball dataset, were evaluated on an independent test set of $N = 1,000$ Metaball samples. Table 1 summarizes the mean $\pm$ standard deviation of all evaluation metrics.

Table 1. Reconstruction accuracy on the independent Metaball test set ($N = 1,000$).

Metric	Single-Encoder	Dual-Encoder
3D Dice	0.9314 ± 0.0519	0.9334 ± 0.0524
3D IoU	0.8757 ± 0.0851	0.8792 ± 0.0855
Volume Error (%)	6.14 ± 5.90	6.15 ± 6.26
Area Error (%)	3.31 ± 2.52	3.11 ± 2.42

The dual-encoder model outperformed the single-encoder model on three of the four metrics (3D Dice, 3D IoU, and area error), while the single-encoder model showed a marginally lower volume error. The difference between the two models is small relative to the sample standard deviation; under an identical training data volume, the effect of the view-fusion strategy on Metaball reconstruction accuracy is limited.

3.2 Cross-Validation on the Physics-Based Dataset (BubbleSH)

Both models were additionally evaluated on the $N = 1,000$ BubbleSH validation set (Section 2.3) to assess generalization beyond the Metaball geometric family. Table 2 summarizes the results.

Table 2. Reconstruction accuracy on the physics-based BubbleSH validation set ($N = 1,000$).

Metric	Single-Encoder	Dual-Encoder
3D Dice	0.8977 ± 0.0318	0.9061 ± 0.0323
3D IoU	0.8159 ± 0.0508	0.8298 ± 0.0520
Volume Error (%)	6.43 ± 4.56	5.05 ± 3.80
Area Error (%)	4.11 ± 3.00	3.27 ± 2.59

On the BubbleSH set, the dual-encoder model outperformed the single-encoder model on all four metrics, with a markedly larger margin than on the Metaball set (3D Dice difference of 0.0084 versus 0.0020; volume error difference of 1.38 versus 0.01 percentage points). This contrast suggests that the dual-encoder's independent per-view encoding matters comparatively little on in-distribution Metaball shapes, but becomes a meaningful source of robustness once the input geometry departs from the training distribution. Fig. 4 illustrates this range qualitatively, with samples selected to span from high-accuracy to lower-accuracy reconstructions rather than only the best-performing cases.

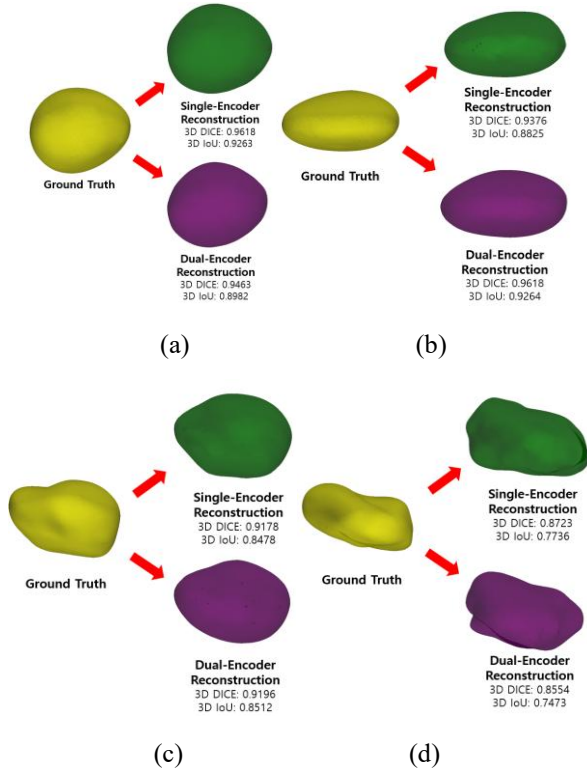


Fig. 4. Qualitative comparison of ground-truth and reconstructed shapes on representative BubbleSH samples (single- vs. dual-encoder), spanning the range of reconstruction accuracy observed in Table 2: (a), (b) high-accuracy cases; (c) a mid-range case, close to the BubbleSH-set mean; (d) a lower-accuracy case.

3.3 Comparison Between Metaball and BubbleSH Validation Sets

Table 3 compares the volume and area errors of each model between the two validation sets.

Table 3. Volume and surface-area errors compared between the Metaball and BubbleSH validation sets.

Model	Metric	Metaball	BubbleSH
Single	Volume Error (%)	6.14 ± 5.90	6.43 ± 4.56
Single	Area Error (%)	3.31 ± 2.52	4.11 ± 3.00
Dual	Volume Error (%)	6.15 ± 6.26	5.05 ± 3.80
Dual	Area Error (%)	3.11 ± 2.42	3.27 ± 2.59

For the single-encoder model, both errors increased slightly on BubbleSH relative to Metaball. For the dual-encoder model, the volume error on BubbleSH (5.05%) was lower than on Metaball (6.15%), while the area error remained comparable — a pattern that reinforces the interpretation above: the dual-encoder's advantage is not merely dataset noise, since it holds up (and even improves on one metric) when the evaluation geometry changes entirely. Fig. 5 shows predicted vs. ground-truth volume and surface area for both models across the two validation sets, and Fig. 6 compares the distribution of 3D IoU values across all four model-dataset combinations. The discrete clustering of ground-truth values visible in Fig. 5(c)–(d) reflects the three fixed bubble diameters (4, 5, and 6 mm) sampled in the BubbleSH dataset, rather than any limitation of the reconstruction.

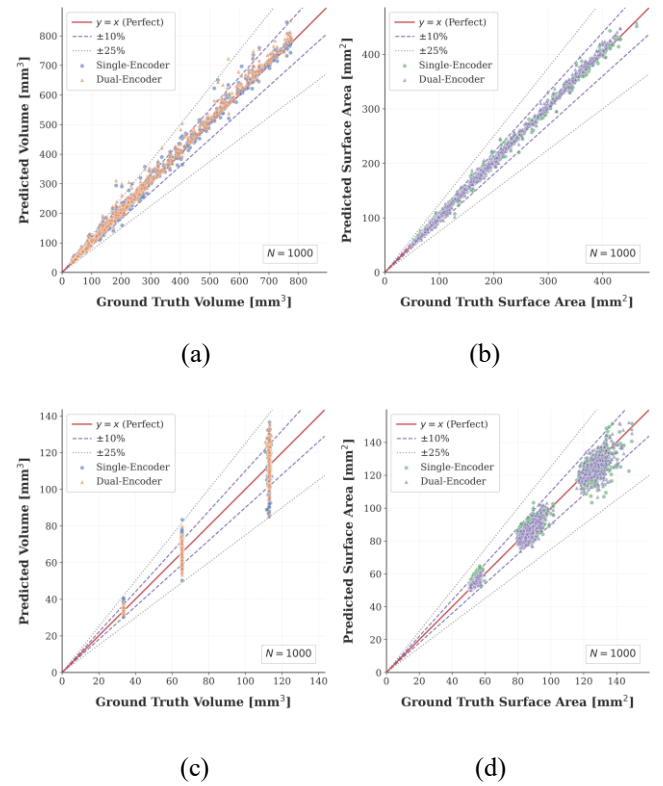
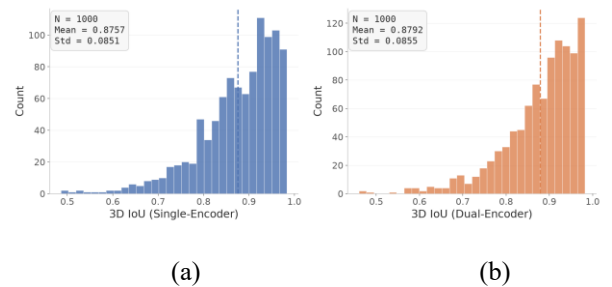


Fig. 5. Predicted vs. ground-truth volume and surface area for single- and dual-encoder models on the Metaball and BubbleSH test sets, with the $y = x$ reference line: (a) Metaball volume, (b) Metaball surface area, (c) BubbleSH volume, (d) BubbleSH surface area.



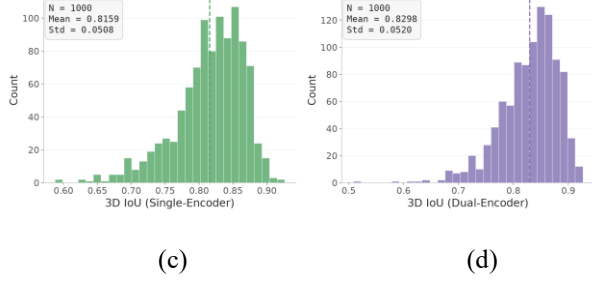


Fig. 6. Distribution of 3D IoU values for single- and dual-encoder models on the Metaball and BubbleSH test sets: (a) Metaball single-encoder, (b) Metaball dual-encoder, (c) BubbleSH single-encoder, (d) BubbleSH dual-encoder.

3.4 Limitations

Case-by-case inspection of low-scoring samples revealed a modest accuracy trend: bubbles with a relatively flattened (disk-like) geometry tended to score below the validation-set average on both Metaball and BubbleSH, although the underlying discrepancy differed between the two. Table 4 provides a representative example from each dataset. For the Metaball sample, the reconstructed shape was noticeably thicker than the ground truth along the axis perpendicular to both camera views, consistent with an overestimated volume. For the BubbleSH sample, the degree of flattening itself was reproduced well, but the extent of the flattened face did not match the ground truth; this discrepancy alone was sufficient to lower the 3D Dice and IoU despite the correct overall thickness. In both samples, 3D Dice and IoU fell well below the dataset means reported in Tables 1–2, with correspondingly larger volume and area errors. The likely cause is common to both cases: with no direct information available along the third, unobserved axis, two orthogonal cameras leave the reconstruction under-constrained in that direction, so that different 3D shapes can project to nearly identical two-view silhouettes (visual-hull ambiguity). The fact that this occurs, in different forms, on both datasets suggests a structural limitation of two-view reconstruction rather than a property specific to either dataset.

Table 4. Representative examples of relatively flattened bubble geometries showing below-average 3D reconstruction accuracy, observed in both the Metaball and BubbleSH validation sets (cf. dataset means in Tables 1–2).

Dataset	Model	3D Dice	3D IoU	Vol. Err (%)	Area Err (%)
Metaball	Single	0.7698	0.6257	43.93	11.73
	Dual	0.7560	0.6077	48.81	13.75
BubbleSH	Single	0.8358	0.7180	27.18	15.52
	Dual	0.8225	0.6985	21.97	12.48

This is a known limitation of two-view 3D reconstruction rather than an artifact specific to this framework; it is reported here for transparency, and points to additional view configurations or physics-informed regularization as directions for future work.

4. Conclusions

This study compared single- and dual-encoder CNN architectures under an identical training data volume (200,000 Metaball samples), cross-validated on an independent Metaball test set and a physics-based DNS-derived dataset (BubbleSH), each $N = 1,000$.

On the Metaball set, the dual-encoder model showed only a marginal advantage over most metrics: with training data volume held constant, encoder structure by itself has a limited effect within the training distribution. On the BubbleSH set, by contrast, the same dual-encoder model showed a more pronounced and consistent advantage across all four metrics — its benefit emerges specifically under distribution shift, not under matched, in-distribution conditions. A visual-hull ambiguity failure mode for flattened bubble geometries was also identified and is recommended for consideration in future model or data design. More broadly, whereas prior work validated the framework only against Metaball-generated samples from the same distribution used for training, this study additionally cross-validated it against a physics-based dataset derived from independent DNS simulations. The accuracy achieved on BubbleSH was of comparable magnitude to that on the Metaball set (Tables 1–2) rather than substantially worse; the framework therefore generalizes reasonably well to physically realistic bubble geometries, not only to the synthetic shapes it was trained on.

Overall, the dual-encoder architecture is the better default choice for two-view 3D bubble reconstruction, especially when the target geometry departs from the training distribution; on in-distribution (Metaball) data, its edge over the single-encoder model remains within sample variance. Building on this framework, future work will apply the same camera-calibrated pipeline to real experimental bubble images. The resulting volume and surface area measurements can be used to analyze heat transfer behavior and applied directly to thermal-hydraulic correlations such as the CHF correlation, contributing to nuclear reactor safety analysis.

ACKNOWLEDGEMENT

This work was supported by the Innovative Small Modular Reactor Development Agency grant funded by the Ministry of Science and ICT(No. RS-2024-00408520).

REFERENCES

- [1] Ishii, M. and Hibiki, T., *Thermo-Fluid Dynamics of Two-Phase Flow*, Springer, New York, NY, USA, 2006.
- [2] Prasser, H.-M., Böttger, A., and Zschau, J., "A New Electrode-Mesh Tomograph for Gas-Liquid Flows," *Flow Measurement and Instrumentation*, 9(2), pp. 111-119, 1998.
- [3] Monrós-Andreu, G., Peña-Monferrer, C., Martínez-Cuenca, R., Torró, S., Chiva, S., "Numerical

Evaluation and Assessment of Key Two-Phase Flow Parameters Using Four-Sensor Probes in Bubbly Flow," *Sensors*, 25(24), 7490, 2025.

[4] Shin, S. et al. (2026). "A Fully Data-Driven 3D-CNN Framework for Volumetric Reconstruction of a Deformed Bubble from Two Orthogonal Views," KNS Spring Meeting, 2026.

[5] Shin, S., Jang, I., Song, M., "Experimental Application of a CNN-based Framework for Three-Dimensional Reconstruction of a Single Rising Bubble from Two Orthogonal Views," submitted to NTHAS14, Yokohama, Japan, Nov. 2026.

[6] Ramesh, R., Bennema ten Brinke, K., Orij, D., Roghair, I., and Menkovski, V., "BubbleSH: A Dataset of Rising Bubbles with Deformable Interfaces," arXiv:2607.07275, 2026.