A Fully Data-Driven 3D CNN Framework for Volumetric Reconstruction of a Deformed Bubble from Two Orthogonal Views

Sangoh Shin, Janghun Han, 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 : Convolutional Neural Network (CNN), 3D reconstruction, Single bubble, Machine learning

1. Introduction

Gas-liquid two-phase flows occurring in the core and steam generators of nuclear power plants are critical factors determining heat transfer efficiency and system safety. In particular, the behavior of individual bubbles under subcooled or saturated boiling conditions is directly linked to the void fraction and interfacial area concentration, which serve as essential parameters for predicting the Critical Heat Flux (CHF) and evaluating the thermal-hydraulic safety margin of nuclear reactors [1]. Recently, with the advent of Small Modular Reactors (SMRs) and advanced nuclear systems, reactor internals and heat exchangers have become increasingly compact and geometrically complex. Under these intricate operational environments, bubbles are subjected to highly complex flow patterns and strong shear forces, resulting in severe morphological deformations and irregular topological changes. To accurately predict the macroscopic behavior of such complex two-phase flow systems, a fundamental understanding of the microscopic deformation mechanisms and interfacial characteristics at the individual bubble level is strictly required as a prerequisite. Therefore, measuring and simulating the three-dimensional (3D) shape, volume, and topological changes of a highly deformed single rising bubble in high-speed flow fields with high precision has been a continuous research focus in nuclear thermal hydraulics.

Traditionally, intrusive measurement techniques, such as wire-mesh sensors and local electrical probes, have been widely utilized to measure the phase distribution and bubble morphology [2]. While these methods provide valuable high-resolution cross-sectional data, they inevitably cause physical disturbances to the flow field and may alter the natural dynamics of the bubbles due to their intrusive nature. To fundamentally eliminate these physical disturbances, non-intrusive optical measurement techniques have emerged as a compelling alternative. Among various optical methods, the 2-View visualization technique, which orthogonally positions two cameras, has been adopted as a standard approach due to its minimal spatial requirements for optical windows [3]. However, reconstructing a complete 3D volume from only two 2D orthogonal projection images presents a challenging mathematical problem due to the lack of depth information. To address this, previous studies have introduced geometrical assumptions, such as approximating the entire bubble as a combination of semi-ellipsoids [4], stacking horizontally sliced elliptical cross-sections [5], or merging 2D contours using active contour algorithms [3]. While these traditional approaches yield reasonable results for spherical or mildly oblate bubbles, they exhibit limitations when applied to highly asymmetric or concave bubbles deformed by strong shear forces, where reconstruction errors tend to increase. Furthermore, utilizing three or more cameras to improve accuracy [5] poses practical installation difficulties in complex nuclear thermal-hydraulic experimental facilities.

Recently, with the advancement of computer vision, deep learning technologies, particularly Convolutional Neural Networks (CNNs), have been actively introduced into fluid dynamics and two-phase flow analysis [6]. Although artificial intelligence has demonstrated significant potential in tracking individual bubbles [7] and predicting flow fields using Physics-Informed Neural Networks (PINNs) [8], applying 3D CNNs directly to the 2-View based 3D volumetric bubble reconstruction problem has been relatively scarce. This is largely due to the "data bottleneck" in building large-scale paired datasets—i.e., multi-view 2D images and their corresponding 3D ground-truth dataset—which is essential for training 3D neural networks. Generating a massive paired dataset through actual multiphase flow experiments or high-fidelity Computational Fluid Dynamics (CFD) simulations require prohibitive amounts of time and computational resources. Furthermore, previous data-driven studies have tended to generate training data tailored to specific physical phenomena or flow conditions. While this approach may yield excellent reconstruction performance under similar conditions, it inherently suffers from degraded generalization capabilities, making it difficult to reconstruct unseen, highly irregular bubble topologies that deviate from the training distribution.

To overcome these challenges, this study proposes a 3D CNN-based volumetric reconstruction framework that precisely rebuilds the 3D shape of a highly deformed object at the voxel level using only two orthogonal projection images, without relying on strict geometric assumptions. To bypass the computationally demanding data generation based on CFD or experiments, and to maximize the morphological versatility of the model, we

developed an efficient pipeline that rapidly generates a massive synthetic dataset comprising diverse, completely random 3D free-form geometries and their corresponding 2D projection silhouettes using the anisotropic Metaball algorithm. Rather than explicitly simulating specific physical bubble shapes governed by fluid dynamics, this approach deliberately focuses on creating arbitrary, highly irregular 3D topological structures to train and rigorously test the model's fundamental geometric inference capabilities. The encoder-decoder 3D CNN is trained exclusively on this geometric dataset to probabilistically infer hidden curvatures and output a $2 \times 128 \times 128$ binary voxel grid.

As an intermediate and foundational stage toward the ultimate goal of real-world single bubble reconstruction, this paper concentrates on validating the purely geometrical reconstruction performance of the proposed model. The trained model undergoes rigorous quantitative and qualitative evaluations using a newly generated test set of completely random 3D free-form topologies that are strictly independent of the training data. To ensure a rigorous evaluation of the reconstructed shapes, we utilized the 3D Intersection over Union (IoU) to accurately quantify the voxel-wise morphological agreement, complementing the macroscopic volumetric error metrics commonly used in thermal-hydraulic studies. Validations of the proposed framework against physical CFD simulations and actual experimental measurements of multiphase flows are reserved for future work. This study provides a highly efficient, data-driven methodological foundation capable of precisely reconstructing complex 3D free-form volumes from only two orthogonal 2D projections, effectively overcoming severe optical window constraints in advanced nuclear and thermal-hydraulic experimental facilities.

2. Methodology

2.1. Overall Framework

To overcome the inherent data bottleneck in training 3D CNNs for volumetric bubble reconstruction, this study proposes a fully data-driven, geometry-free framework. As illustrated in Fig. 1, the overall research pipeline consists of three sequential stages: synthetic dataset generation, model training, and geometric cross-validation.

First, an extensive synthetic dataset of arbitrary 3D free-form topologies and their corresponding 2D orthogonal binary images (front and side views) is rapidly generated computationally using a stochastically randomized Metaball algorithm in Python. Second, a hybrid 2D-to-3D CNN architecture is trained exclusively on this geometric dataset to learn the hidden mapping between 2D silhouettes and 3D volumes. Finally, instead of immediate physical validation, the trained model undergoes rigorous geometric cross-validation using an independent, unseen test set of highly complex synthetic topologies. Through this end-to-end pipeline, the fundamental geometrical capabilities limits and the 3D volumetric reconstruction accuracy of the proposed model can be explicitly evaluated prior to its application in actual fluid dynamics.

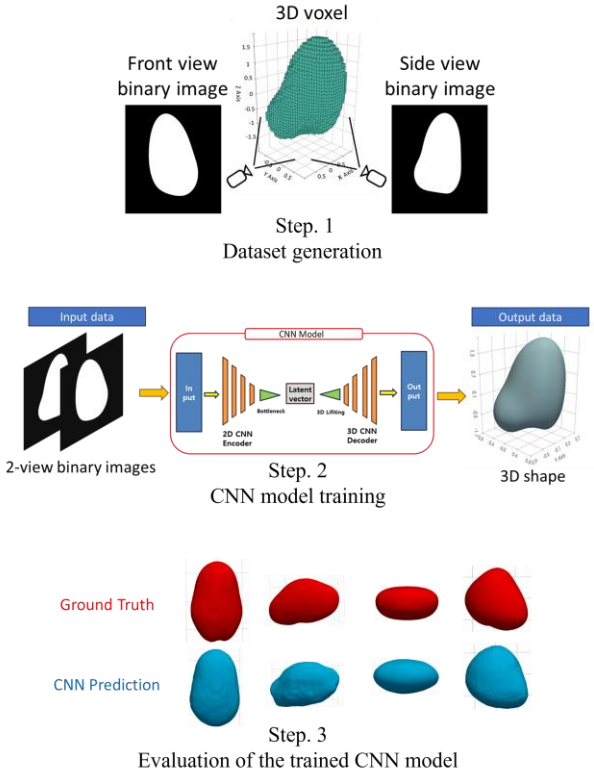


Fig. 1. The overall framework of the proposed data-driven 3D bubble reconstruction method.

2.2. Data Generation via Anisotropic Metaball Algorithm

Generating a massive paired dataset through high-fidelity CFD simulations or actual experiments is computationally prohibitive and risks model overfitting to specific flow conditions. To construct a comprehensive paired dataset of 2D binary projection images and their corresponding 3D volumetric ground truths while fundamentally bypassing these limitations, an advanced, data-driven geometric pipeline was developed using Python. To realistically emulate the highly irregular, completely free-form topologies of a single rising bubble—including arbitrary concave deformations—a stochastically randomized, anisotropic Metaball algorithm was employed. The fundamental mechanism of evaluating the scalar density field from control points and extracting the 3D isosurface is illustrated in Fig. 2.

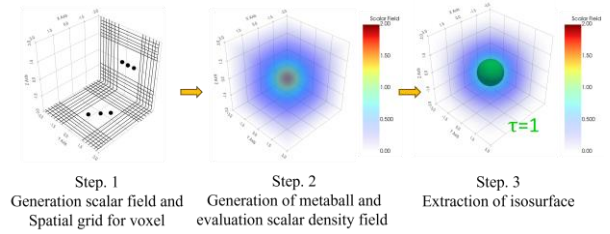


Fig. 2. Fundamental process of evaluating the scalar density field from metaball and extracting the corresponding 3D isosurface.

Unlike standard spherical metaballs, this study utilizes anisotropic ellipsoids as the base control units. By applying a completely random 3D rotation matrix ($\mathbf{R}$) generated via QR decomposition, the global spatial coordinate $\mathbf{x} = [\mathrm{x}, \mathrm{y}, \mathrm{z}]^{\mathrm{T}}$ is transformed into a local coordinate system $\mathbf{q}_{\mathrm{i}} = \left[\mathrm{q}_{\{\mathrm{x,i}\}}, \mathrm{q}_{\{\mathrm{y,i}\}}, \mathrm{q}_{\{\mathrm{z,i}\}}\right]^{\mathrm{T}} = R^{\mathrm{T}}(\mathbf{x} - \mathbf{c}_{\mathrm{i}})$ relative to the center $\mathbf{c}_{\mathrm{i}}$ of the i-th metaball. The cumulative spatial density field $f(x)$ is defined as in Eq. (1).

$$f(x) = \sum_{i=1}^{N} \frac{s_i}{\left(\frac{q_{x,i}}{r_{x,i}}\right)^2 + \left(\frac{q_{y,i}}{r_{y,i}}\right)^2 + \left(\frac{q_{z,i}}{r_{z,i}}\right)^2 + \epsilon} \qquad (1)$$

N is the number of metaballs (randomly selected between 5 and 13), $r_{x,i}, r_{y,i}, r_{z,i}$ are the distinct directional radii representing the anisotropic nature of the ellipsoids, s_i is the strength parameter of the i-th metaball, and ϵ is a small constant to prevent singularity. Furthermore, to explicitly model the realistic concave interfaces (dimples) frequently observed in actual fluid dynamics, negative strength values ($s_i < 0$) were probabilistically introduced, effectively carving out local volumes from the primary shapes. As depicted in Fig. 3, the fusion of these multiple interacting control points effectively generates a wide variety of highly irregular and complex 3D topologies.

Based on a predefined threshold (τ), the isosurface satisfying $f(\boldsymbol{x}) = \tau$ was extracted using the Marching Cubes algorithm. After rigorous filtering processes—verifying bounding box inclusion, single-component connectivity, and surface curvature reasonability—the verified 3D mesh was perfectly centered at the origin. To evaluate the purely morphological reconstruction capability independent of the physical bubble scale, the entire 3D rendering and subsequent voxelization processes were performed within a normalized dimensionless coordinate system ($[-3,3]^3$). The enclosed space was then discretized into a 128×128×128 binary voxel grid, establishing the exact 3D ground truth.

Finally, to generate the 2D model input, a virtual perspective projection was implemented using the PyVista rendering engine. Virtual cameras with a 35-degree field of view (FOV) were orthogonally positioned at a fixed distance to capture the front and side view silhouettes of the centered 3D meshes. These projections were thresholded to yield two perfectly aligned 600 × 600 pixel binary images per 3D volume. Through this geometry-free computational pipeline, a vast synthetic dataset pairing 2D images with 3D voxels was rapidly and robustly constructed.

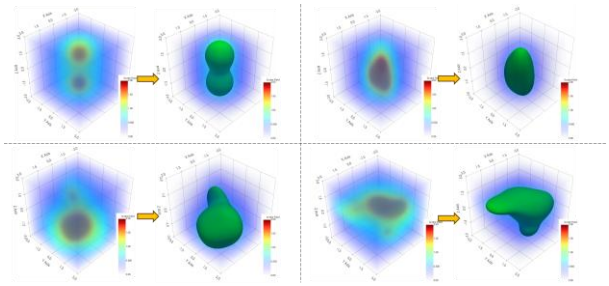


Fig. 3. Examples of complex 3D free-form topologies generated by the fusion of multiple interacting metaballs.

2.3. 2D-3D CNN Architecture and Training Strategy

To reconstruct a fully 3D volumetric bubble from limited 2D orthogonal projections, an Encoder-Decoder-based CNN architecture was designed. As illustrated in Fig. 4, the model accepts two perfectly aligned 2D binary images concatenated along the channel dimension, forming an input tensor of size $2 \times 128 \times 128$. The feature extraction begins with a 2D CNN encoder comprising four sequential convolutional blocks with 4×4 kernels and Rectified Linear Unit (ReLU) activations. These layers systematically downsample the spatial dimensions while progressively increasing the feature channels ($32 \rightarrow 64 \rightarrow 128 \rightarrow 256$). The extracted high-level 2D spatial features are then flattened and passed through a fully connected layer to compress the information into a dense, 1-dimensional latent vector of size 128, effectively capturing the core morphological representation of the free-form geometry.

To bridge the dimensional gap between 2D and 3D spaces, a 3D lifting module expands the latent vector via a fully connected layer and unflattens it into a 3D tensor of size $128 \times 16 \times 16 \times 16$. This tensor serves as the foundational structural grid for the 3D CNN decoder, which consists of three sequential 3D transposed convolutional layers with $4 \times 4 \times 4$ kernels. The decoder symmetrically upsamples the spatial resolution while reducing the channel depth ($128 \rightarrow 64 \rightarrow 32 \rightarrow 1$), ultimately yielding a 3D logit tensor of size $1 \times 128 \times 128 \times 128$ that represents the unnormalized probability of each voxel belonging to the internal bubble volume.

A critical challenge in 3D voxel generation is the severe class imbalance, as the vast majority of the 128^3 bounding box consists of empty background space compared to the actual bubble volume. To prevent the model from converging to a trivial all-zero prediction, a dynamically weighted Binary Cross-Entropy (BCE) loss was employed. The loss function $\mathcal{L}_{\mathcal{BCE}}$ is defined as follows in Eq. (2).

$$\mathcal{L}_{BCE} = -\frac{1}{V}\sum_{j=1}^{V}\begin{bmatrix} w_p \cdot y_j \log\left(\sigma(\hat{y}_j)\right) \\ +(1-y_j)\log\left(1-\sigma(\hat{y}_j)\right)\end{bmatrix} \quad (2)$$

V is the total number of voxels in the 3D domain, $y_j \in \{0,1\}$ is the ground truth binary label of the j-th voxel, $\hat{y}_j$ is the predicted logit from the model, $\sigma(\cdot)$ represents the sigmoid activation function, and w_p is the positive weight assigned to mitigate the severe class imbalance.

A total of 10,000 synthetic paired datasets were generated and randomly partitioned into training (80%), validation (10%), and testing (10%) sets. The model was optimized using the Adam optimizer with an initial learning rate of 1×10^{-4}, coupled with a learning rate scheduler that decays the rate by a factor of 0.5 upon validation stagnation. For the quantitative evaluation of the model's reconstruction performance, the IoU was primarily used as an evaluation parameter.

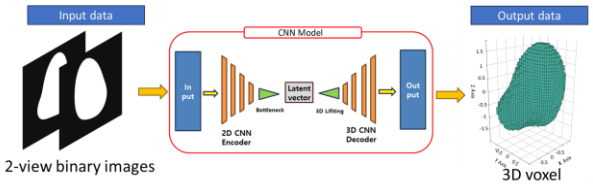


Fig. 4. 2D-to-3D CNN model architecture.

2.4. Quantitative Evaluation of Trained CNN Model

To comprehensively evaluate the 3D volumetric reconstruction performance of the proposed CNN model, a dual-scale assessment strategy was adopted. The evaluation bridges traditional global geometric parameters with a voxel-wise spatial overlap parameter widely utilized in computer vision and medical image analysis. All evaluation parameters were calculated using the $128 \times 128 \times 128$ binary voxel grid of the ground truth (V_{GT}) and the model's binarized prediction (V_{pred}).

2.4.1. Global Geometric Parameters: Volumetric and Surface Area Errors.

In fluid dynamics, accurately predicting the total volume and interfacial area of a bubble is crucial, as they directly govern buoyancy and mass/heat transfer rates. The Volumetric Error (E_V) evaluates global volume preservation and is defined as the relative difference between the total number of positive voxels in the predicted and ground truth grids as Eq. (3).

$$E_V = \frac{|V_{pred} - V_{GT}|}{V_{GT}} \quad (3)$$

Similarly, the Surface Area Error (E_S) assesses the model's ability to reconstruct the global interfacial area. To compute this, the discrete 3D voxel grids of both the ground truth and the prediction were first converted into continuous 3D polygon meshes using the Marching Cubes algorithm. The total surface area was then calculated by summing the areas of the extracted triangular faces as Eq. (4).

$$E_S = \frac{|S_{pred} - S_{GT}|}{S_{GT}} \quad (4)$$

S_{GT} and S_{pred} denote the extracted surface areas of the ground truth and predicted meshes, respectively.

2.4.2. Spatial Overlap Parameter: Voxel-wise Shape Fidelity.

While global errors can evaluate the overall scale, they cannot penalize spatial distortions or localized shape mismatches if the total volume remains identical. To strictly quantify the 3D morphological fidelity and localized deformations, the 3D IoU was employed. The IoU yields a value ranging from 0 to 1, where a value closer to 1 indicates a mathematically identical morphological match between the predicted and actual geometries, while 0 signifies entirely disjoint volumes. This spatial overlap parameter is an established gold standard in highly critical fields such as medical image segmentation (e.g., 3D MRI and CT analysis), providing strong reliability for evaluating complex topological matching.

The IoU measures the ratio of the spatially overlapping volume to the combined total volume of both grids as described in Eq. (5).

$$\text{IoU} = \frac{|V_{pred} \cap V_{GT}|}{|V_{pred} \cup V_{GT}|} = \frac{TP}{TP+FP+FN} \quad (5)$$

TP (True Positive), FP (False Positive), and FN (False Negative) represent the voxel-wise classification outcomes. By combining these global and spatial parameters, the model's purely geometrical inference capability was robustly validated.

3. Results and Discussion.

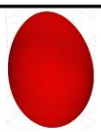
3.1. Qualitative Visual Comparison

To provide an intuitive understanding of the CNN's predictive performance, representative visual comparisons of the ground truth and reconstructed 3D shapes are presented in Table I. The examples are categorized into best (a), average (b), and worst cases (c) based on their IoU scores.

In the best and average cases (IoU > 0.90), the model accurately reconstructed highly asymmetric geometries and mildly concave dimples. Although the input consists solely of two orthogonal 2D silhouettes lacking depth information, the CNN effectively learned the geometric mapping to interpolate the hidden curvatures.

Conversely, the worst-case examples (IoU < 0.70) reveal the inherent optical limitations of the 2-View setup rather than a failure of the neural network. Severe discrepancies primarily occurred when deep concavities or internal voids were oriented perfectly parallel to the viewing axes of both cameras. Under such specific orientations, the carved features are completely obscured within the external silhouettes, making deterministic 3D reconstruction mathematically impossible. Nevertheless, the CNN probabilistically inferred physically plausible solid structures to encompass the hidden concavities, maintaining the overall structural integrity without generating scattered voxels.

Table I: Examples of the 3D volumetric reconstructions

		Volume	Surface Area	IoU
(a)	Groud Truth 	41.9	59.5	0.97
	CNN prediction 	42.8	60.4	
(b)	Groud Truth 	39.1	58.5	0.88
	CNN prediction 	41.4	59.7	
(c)	Groud Truth 	13.2	29.9	0.66
	CNN prediction 	9.4	26.2	

3.2. Quantitative Performance Metrics

The overall quantitative evaluation of the 1,000 independent test samples is summarized in Table II. The model achieved a high mean IoU of 0.91, demonstrating precise voxel-wise agreement between the predicted grids and the ground truth. Furthermore, the global geometric parameters showed robust performance, yielding low mean error rates of 7.15% and 3.76% for the volume and surface area, respectively.

Table II: Quantitative performance metrics of the 3D volumetric reconstruction.

	IoU	Volume Error, E_V [%]	Surface Area Error, E_S [%]
Mean	0.915	7.15	3.76
Min	0.659	0.15	0.02
Max	0.978	34.56	19.87

The detailed distribution of the IoU scores is illustrated in Fig. 5. The histogram is strongly left-skewed, with most of the predictions densely concentrated in the high-accuracy region (IoU > 0.90). The minimum recorded IoU remained above 0.65, confirming that the framework avoids catastrophic reconstruction failures even for highly irregular geometries.

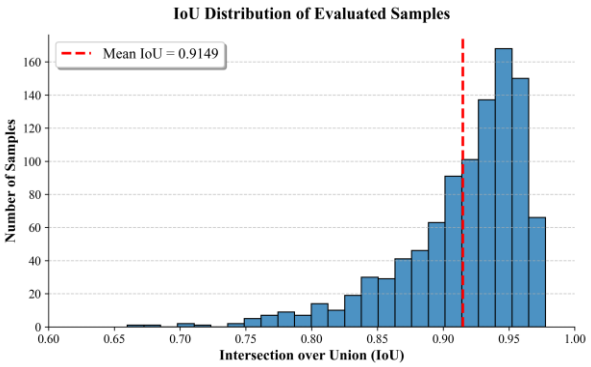


Fig. 5. Distribution of IoU scores for the evaluated samples (N=1000).

To further validate the volume preservation capability, a parity plot comparing the normalized ground truth and predicted volumes is presented in Fig. 6. The predicted volumes closely align with the perfect prediction line (y = x). Notably, regardless of the absolute scale of the generated structures, almost all data points fall within the ±10% error margin. Since a ±10% error bound is widely considered an acceptable standard for two-phase flow parameter predictions in nuclear thermal hydraulics, these results confirm that the proposed purely data-driven model can provide highly reliable volumetric data for evaluating the void fraction and interfacial area.

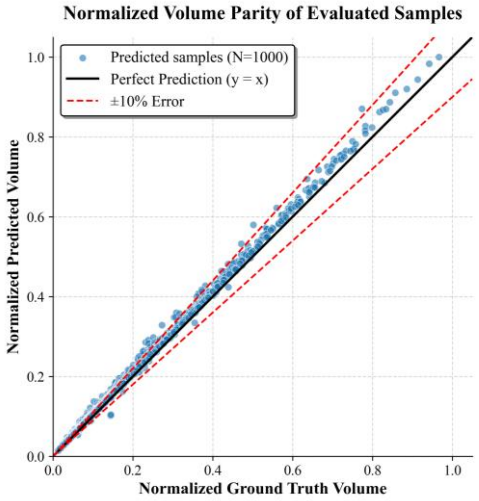


Fig. 6. Parity plot comparing the normalized ground truth and predicted volumes of the evaluated samples (N=1000).

4. Conclusions

This study proposed a fully data-driven, 3D CNN framework designed to reconstruct the highly deformed, complete 3D volumetric shapes of single bubbles using only two orthogonal 2D projection images. To fundamentally overcome the critical "data bottleneck" associated with training 3D neural networks in fluid dynamics, an efficient and novel computational pipeline utilizing a stochastically randomized, anisotropic Metaball algorithm was developed. This geometry-free approach successfully generated a massive synthetic dataset of arbitrary 3D free-form topologies, eliminating the computationally prohibitive reliance on physical high-fidelity simulations or experimental measurements for data generation.

Rigorous geometric cross-validation utilizing 1,000 independent synthetic test samples demonstrated the exceptional predictive performance of the proposed model. The framework achieved a high mean IoU of 0.91 and strictly bounded the overall volumetric predictions within a $\pm$ 10% error margin. Qualitatively, the model successfully generalized the hidden morphological mappings, precisely interpolating highly asymmetric curvatures and concave dimples without relying on idealized mathematical assumptions. Even under severe optical depth-ambiguity inherent to the 2-View setup, the CNN demonstrated remarkable robustness by probabilistically inferring physically plausible internal structures rather than experiencing catastrophic reconstruction failures.

Ultimately, this study establishes a highly reliable and efficient methodological foundation for non-intrusive, voxel-level 3D interfacial measurements using only two orthogonal 2-View images in advanced nuclear experimental facilities subject to severe optical constraints. By demonstrating the fundamental geometrical inference capability using the proposed CNN model, this framework offers a promising tool for acquiring highly accurate void fraction and interfacial area data critical for reactor safety analyses. Building upon this validated geometrical foundation, future work will focus on deploying the pre-trained model to physical CFD datasets and actual multiphase flow experimental measurements to thoroughly evaluate its practical applicability in real-world fluid dynamics.

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) and the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (RS-2025-02310831)

REFERENCES

[1] M. Ishii and T. Hibiki, Thermo-fluid Dynamics of Two-Phase Flow, Springer, New York, 2006.
[2] H. M. Prasser, A. Böttger, and J. Zschau, A New Wire-Mesh Sensor for Fast Measurement of Phase Boundaries in Pipes, Flow Measurement and Instrumentation, Vol.9, p.111, 1998.
[3] S. J. Kim and G. C. Park, Development of an Orthogonal Double-Image Processing Algorithm to Measure Bubble Volume in a Two-Phase Flow, Nuclear Engineering and Technology, Vol.39, p.623, 2007.
[4] Y. Bian, H. Wang, Z. Zhang, and X. Wang, 3D Reconstruction of Single Rising Bubble in Water Using Digital Image Processing and Characteristic Matrix, Particuology, Vol.11, p.784, 2013.
[5] Y. Zhang, X. Que, M. Hu, and Y. Zhou, 3D Reconstruction of a Single Bubble in Transparent Media Using Three Orthographic Digital Images, Applied Sciences, Vol.10, p.5936, 2020.
[6] S. L. Brunton, B. R. Noack, and P. Koumoutsakos, Machine Learning for Fluid Mechanics, Annual Review of Fluid Mechanics, Vol.52, p.477, 2020.
[7] R. F. L. Cerqueira and E. E. Paladino, Development of a Deep Learning-Based Image Processing Technique for Bubble Pattern Recognition and Shape Reconstruction in Dense Bubbly Flows, Chemical Engineering Science, Vol.230, p.116182, 2021.
[8] M. Dreisbach, E. Kiyani, J. Kriegseis, G. Karniadakis, and A. Stroh, PINNs4Drops: Convolutional Feature-Enhanced Physics-Informed Neural Networks for Reconstructing Two-Phase Flows, arXiv preprint arXiv:2411.15949, 2024.