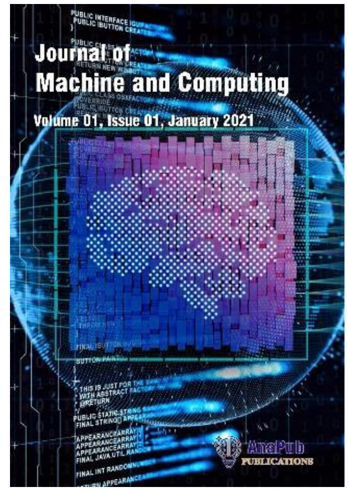
Journal Pre-proof

BubbleGLS: A Machine-Centric Computing Model for IoT-Based Estimation of Bubble Diameter in Gas–Liquid Systems



Ramamoorthy M L, Gayathri C S, Ramya P, Indhuja N and Sathishkumar R

DOI: 10.53759/7669/jmc202505157

Reference: JMC202505157

Journal: Journal of Machine and Computing.

Received 19 April 2025

Revised from 03 July 2025

Accepted 10 July 2025

Please cite this article as: Ramamoorthy M L, Gayathri C S, Ramya P, Indhuja N and Sathishkumar R, "BubbleGLS: A Machine-Centric Computing Model for IoT-Based Estimation of Bubble Diameter in Gas–Liquid Systems", Journal of Machine and Computing. (2025). Doi: https://doi.org/10.53759/7669/jmc202505157.

AnaPub
PUBLICATIONS

BubbleGLS: A Machine-Centric Computing Model for IoT-Based Estimation of Bubble Diameter in Gas–Liquid Systems

[1]ML.Ramamoorthy, [2]C.S. Gayathri, [3]P. Ramya, [4]N. Indhuja, [5]R. Sathishkumar

[1]Assistant Professor, Department of Electrical and Electronics Engineering, Alagappa Chettiar Government College of Engineering and Technology (Autonomous), Tamil Nadu, India

[2]Associate Professor, Department of Electronics and Communication Engineering, Mother Terasa College of Engineering and Technology, Tamil Nadu, India

[3]Associate Professor, Department of Computer Science and Engineering, Mahendra Engineering College, Tamil Nadu, India

[4]Assistant professor, Department of Artificial Intelligence and Data Science, Sengunthar Engineering College, Tamil Nadu, India

[5]Assistant Professor, Department of Mathematics, Velammal College of Engineering and Technology, Tamil Nadu, India

[1]ramamoorthyml@gmail.com, [2]gayathricphd@gmail.com, [3]paramasivam.ramya@gmail.com, [4]indhuja.techinfo@gmail.com, [5]rsk@vcet.ac.in

Corresponding Author: ramamoorthyml@gmail.com

Abstract

In gas–liquid systems, the precise estimation of bubble diameter plays a critical role in analyzing mass transfer, interfacial area, and flow dynamics. We have suggested a machine-centric IoT-integrated computing paradigm called BubbleGLS, which can estimate bubble diameter in real-time leveraging multimodal sensor data and hybrid machine learning. The overall system connects pressure, acoustic, flow and optical sensors that are located above the cylindrical reactor. These sensors record dynamic parameters which are denoised and normalised by means of wavelet filtering and Z-score normalisation. Bubble area, circularity, rise velocity, and acoustic signatures are used as feature extraction and combined through DempsterShafer Theory which provides noise resistance. The learning engine consists of Inception network of spatial features based on an image and XGBoost of structured physical parameters. The model is deployed onto fog and edge devices, and it provides real-time lower than 30 milliseconds latency inference. The validation of 10 different flow regimes reveals that the level of the mean absolute error (MAE) output by BubbleGLS does not exceed 0.25 mm, whereas its R^2 score is higher than 0.97, thus being superior to CNN-LSTM, MobileNet, and Random Forest. It is also resilient as it can remain steady in the accuracy in different noise levels that are up to 45dB. To be used in the smart industrial space where fast response and low cloud reliance are the key factors, BubbleGLS has been optimized. Its modular design and the aspect of this design being machine-specific allows it to be implemented on an otherwise distributed fluidic system with much little calibration to fit its recalibration. All in all, the system reveals a powerful potential in the future third-generation fluid monitoring system, promising a high performance, low-latency, and intelligent character of bubble diameter estimation in a full-scale gas and liquid scenario.

Keywords: Bubble diameter estimation, gas–liquid systems, machine learning, edge computing, IoT sensors, Dempster–Shafer theory, fog analytics, Inception network.

1. Introduction

Gasliquid multiphasic flows have been the focus of a wide variety of chemical, biological, and industrial processes, including bioreactors and bubble column reactors, nuclear cooling and petrochemical extraction. Bubble diameter is one of the parameters that are essential in determining flow characteristics such as interfacial area, mass and heat transfer coefficient and chemical reaction kinetics [1] [2] [3]. Accurate and prompt measure of bubble diameter is hence important in order to maximize efficiency, as well as energy savings, besides ensuring safety of operations. Real-time determination of bubble sizes however is a technical challenge, as gasliquid interactions are highly dynamic as well as non-linear. Intrusive probes, high-speed video analysis and wire-mesh sensors are all methods of traditional bubble measurements [4] [5]. Although high-speed cameras deliver high-resolution tracking, they are computationally demanding and experience loss of frames at high flow rates. Probes such as hot-film anemometers are intrusive causing disturbance to flow and inaccurate readings are obtained. Wire-mesh sensors, which are quite durable, are costly and cannot easily be calibrated to various fluids and geometries [6]. The traditional techniques cannot enable continuous regular involvement in the real-time applications even in the dynamically varying environment since they are subjected to these above limitations [7] [8].

Usage of machine learning (ML) and the Internet of Things (IoT)-driven systems can become another alternative to realize non-invasive, massively scalable, and intelligent monitoring. Both the Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), specifically the Long Short-Term Memory (LSTM) models have been known to promise in the analysis of fluid flows [9] [10]. CNNs can be used to extract spatial information in optical images and LSTM can be used to model the dependency over time in the sensor data. Nevertheless, these deep learning models are resource-demanding and can hardly be executed in edge or fog computing environment because of their scale and high inference latency. With resource-constrained embedded vision tasks, like object detection and classification, lightweight architectures like MobileNet have been tried [11] [12]. With the estimation of bubble diameter particularly within complex and turbulent flow conditions, MobileNet models are not accurate enough. On the same wavelength, the reasonable interpretability and speed of ensemble models such as Random Forest does not apply due to inferior learning of highly non-linear, coupled features, which prevail in fluid dynamics [13] [14]. Figure 1 illustration shows how multimodal IoT sensors integrated with a machine learning model.

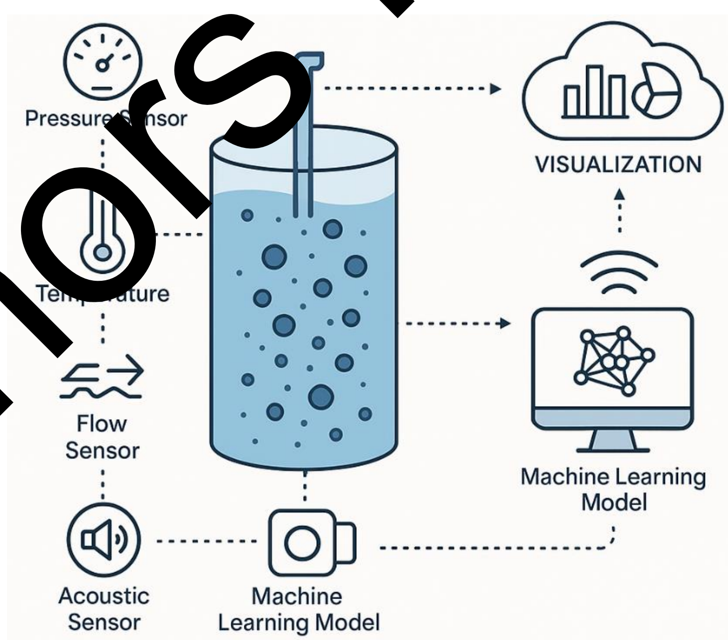


Figure 1. IoT-Based Estimation of Bubble Diameter in Gas–Liquid Systems

One major disadvantage of a variety of existing ML-based methods is the use of a single modality of the data, e.g., visual data input or data on time series on sensors. A prediction accuracy can be considerably increased with a combination of sensor modalities: pressure fluctuations, acoustic signal, flow rate, and optical data.In the case of gasliquid systems dynamic multimodal fusion is not sufficiently explored [15]. In addition, a large number of the models use cloud based inference which adds cloud network delay and issues with privacy. Rich Media Content This paper presents BubbleGLS: a Machine-Centric Computing Model that estimates the diameter of bubbles in real-time with a multimodal IoT sensor system and a hybrid ML architecture. BubbleGLS incorporates Inception Networks to exert multiscale spatial features on optical frames and the XGBoost to research in structured data of non-guide sensors. These complimentary models are joined together with a late fusion prediction method. In situations of poor or noisy input data consistency, BubbleGLS uses DempsterShafer Theory, which is a strategy of belief-based data fusion, which makes it more robust. This is done by doing all computation at the fog layer, thus enabling low-latency inference without depending on the centralized cloud systems.

The system is also tested under ten different gasliquid flow regimes, such as bubbly, slug, churn, annular and oscillatory flows, and its performance proves to be uniform under all the regimes. The model also performs better in vital values including the lowest possible MAE of 0.25 mm and the R 2 Scores of more than 0.97 surpassing all of its benchmark models. It is also accurate when tested in conditions of simulated noise which implies that it can be readily applied into the environment. BubbleGLS handles the existing deficiencies of precision, multimodal data combination, scale-out of deployment and lag within bubble calculations systems. The suggested framework is based on machine learning and the development of computing to create an efficient and rapid strong framework which can be used in intelligent industrial flow systems.

1.1. Main Contribution of the Work

- Machine-based hybrid learning model (a combination of Inception and XGBoost) of the real-time bubble diameter estimation.
- Information fusion of multimodal sensor streams in IoT (optical, acoustic, pressure, flow) with the utilization of DempsterShafer data fusion.
- Implementation of edge and fog computing to allow low latency, real-time inference and scalability to industrial systems.
- Strong preprocessing system that uses a wavelet denoising and normalization to enhance the quality of data involving heterogeneous sensors.
- Seamless system generalization to cover ten individual gasliquid flow regimes and confirms its robustness and robustness to noise.

Section 2 provides a detailed survey of related work on bubble estimation using traditional and AI-driven methods. Section 3 explains the proposed BubbleGLS methodology, including sensor deployment, preprocessing, feature extraction, and the hybrid learning framework. Section 4 presents experimental results and analysis across multiple evaluation metrics. Finally, Section 5 concludes the study and outlines future directions.

2.Related Work

The work exposes the characteristics, the bubble size and the influence of the main operating parameters in column flotation of the gas dispersion. At concentration of frothier of 120 ppm, which is the critical coalescence concentration (CCC) the bubble size was found to have a minimum value of 0.62 mm and the forty obeyed a unimodal distribution rather than to a bimodal one[16]. Bubble size was significantly affected by velocity of the gas and water in which at 1.08 cm/s the bubble holdup reached the highest of 27%. The velocity of bubble rise and the size of bubble rose linearly, whereas the bubble surface area flux reduced linearly. There was a good correlation (R^2 0.86) between measured and estimated sizes (average 0.64 mm, E.23M leaves2664TBD 13% error). Enhancement of situations enhanced flotation.

The paper reports experimental and computational results of a gas liquid stirred tank with the purpose of offering original information about bubble size distribution, and contributing towards predictive modeling necessary in an attempt to design chemical and biochemical gas liquid reactors. Critical parameters which contributes to mass removal and overall fermentation performance including bubble size distribution, gassed power use, and the presence of gas cavities are examined[17]. Advancement of Two-Fluid and Population Balance models in order to make correct predictions of gas liquid mixing is also presented. Findings indicate that bubble size prediction in the impeller region should be predicted accurately to provide a successful prediction of hydrodynamic modeling within aerated stirred tanks.

There is wide application in water and wastewater plants of bubble aeration and its main disadvantage is low gas utilization and high energy requirement. An attempt at finding a solution to this challenge is discussed in this study through micro-nanobubbles (MNBs) that have a large surface area and gas-liquid mass transfer capacity. They created a dynamic model that comprised of mass transfer with force balance (buoyancy, gravity, drag, Basset and virtual mass forces) to simulate rising velocity of MNB and size changing[18]. Findings indicate the best MNBs shrink and explode at the water surface guaranteeing complete gas dissolution, and possible free radical generation. Microbubble behavior was well modeled ($R^2 > 0.85$), nanobubble validation is not yet validated empirically.

Ozonation requires or is supported by the efficient gas liquid mixing; this also applies to ozone advanced oxidation processes. The paper is looking into a static-mixer-based plug-flow reactor as seen in the HiPOxTM system, which uses Computational Fluid Dynamics to model bubble behaviors and mixing patterns[19]. Observation of systems with and without the use of the mixers reveals that the mixers enhance mixing mostly as opposed to the dispersing of the gas. The six-element reactor showed the best performance based on the indices of relative standard deviation (RSD) and bubble diameter; the former had an RSD of 0.793 and a bubble diameter of 1.384mm, and 37.6% of the eight-element configuration in terms of energy usage. These observations justify the application of the static mixers in the gas liquid systems.

The pH changes at the interface, as well as bubble behavior during the water electrolysis process have a fundamental role in cell overvoltage and of energy efficiency. In contrast to the previous works that aimed to study the behavior of individual bubbles, the work studied more than 8,000 bubbles per experiment with the help of highly accurate image processing and edge detection, which allows the application of robust statistical analysis. The essentially results were the production of H_2 and O_2 bubbles at acidic and alkaline media with the main factors such as pH, type of gases, and current density measured by a 2^3 dimensional factorial design[20]. Interfacial pH was simulated in finite element computer models under matching conditions and confirmed experimental results as well as the importance of pH change on bubble distribution and cell voltage. The multination solution provides effective information regarding optimization of electrochemical systems.

3. Methodology

The innovation of this work is a combination of edge intelligence, machine-centric design, and hybrid deep learning to address the issue of real-time bubble diameter estimation. BubbleGLS integrates both the spatial and the temporal, unlike current techniques which focus on one or the other, because BubbleGLS uses an Inception XGBoost ensemble. Moreover, it provides adaptive fog-based processing to make fast accurate deployment and to serve to scale. Due to the usage of DempsterShafer Theory to manage uncertainties in the different kinds of sensor modalities, the system stands out among previous works, providing a robust solution despite noisy conditions in industrial applications. Figure 2 shows the architecture of proposed model.

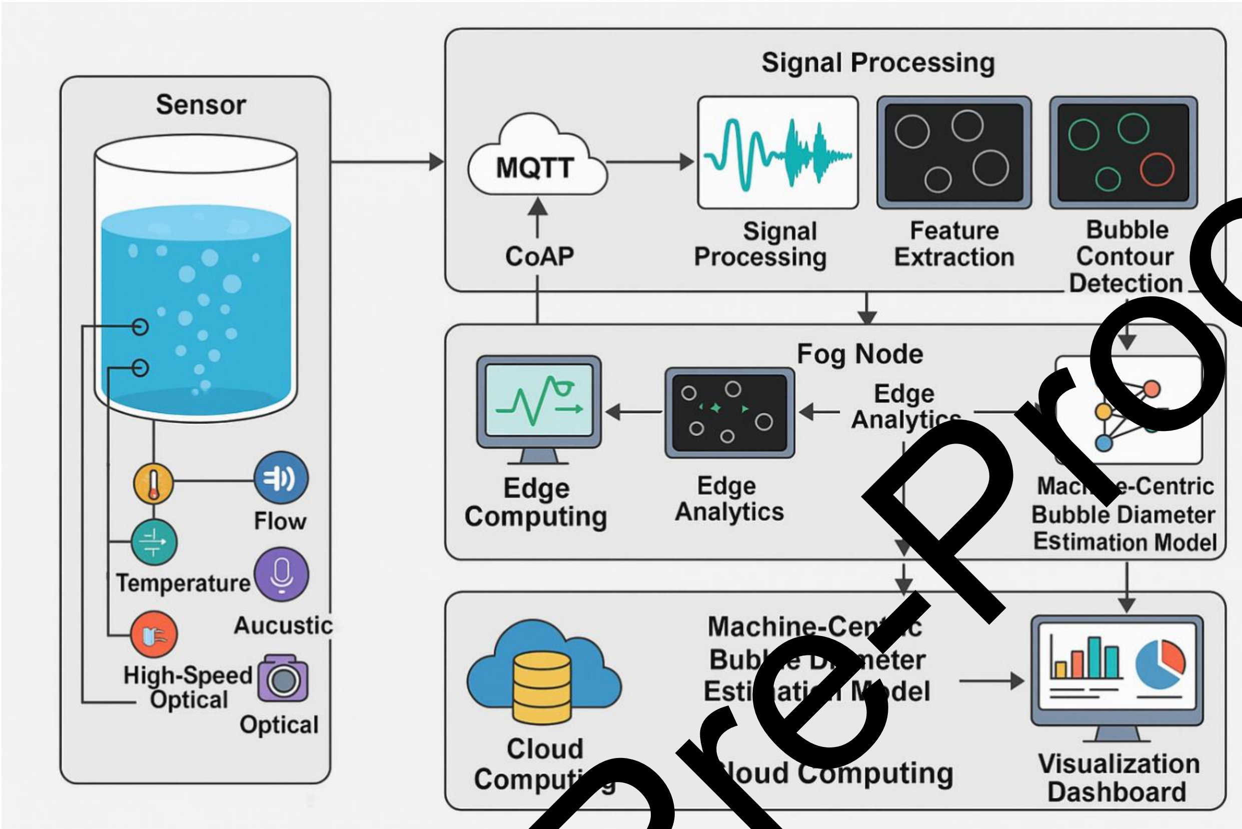


Figure 2. Architecture of Proposed Model

3.1 Sensor Deployment and Data Acquisition

The first and most important step towards achieving a truly machine-based computing framework on how to correctly estimate bubble diameter in gas-liquid systems is the accurate deployment of intelligent IoT-based sensor nodes. The sensor nodes are placed strategically in the cylindrical observation chamber and the vertical column reactor and hide various vertical and radial locations to record spatiotemporal change of bubbles dynamics. Each of the nodes is supplied with multimodal sensors: pressure transducers to measure the hydrostatic variation, thermocouples to conduct temperature profiling in real time, ultrasonic flow meters to measure velocity, acoustic sensors to identify fmc of bubble formation and high speed optical cameras to observe it visually. The grid used is a sensor covering 1.2 m vertical range and 0.5 m radial range that provides complete volumetric monitoring of rising bubbles as they advance and in their contact with the surrounding fluid. These sensor modules are configured on machines to work synchronously and they are dynamically addressable by utilizing a lightweight protocol MQTT that is optimized to transmit with a low latency in the constrained wireless networks. The system enables fallback (CoAP over UDP to raise links tolerance) in very transient operating conditions.The sensors send the time-coded packets into aggregation into an edge computing gateway in the format of binary or JSON. The frequency of acquisition can be set to suit the fluid flow regime, 200 Hz in case of rapid slug flow, and 20 Hz in case of steady-state bubbly flows. The data sampling rate is also further optimized by using a machine learning basis adaptive sampling module at the sensor node by monitoring the entropy change in the environment. Edge-aware firmware at every sensor node embeds the smartness to perform computing at the data source, which is an essential requirement in an IoT-augmented computation system.

3.2 Signal Preprocessing and Noise Filtering

Raw sensor readings obtained in the gasliquid testing apparatus are usually mixed with noise, vibrations contributed by the environment and mechanical vibrations. A powerful computing model is endowed into downstream machine learning tasks by performing a rigorous signal preprocessing pipeline. The structure of this pipeline is made in a modular computing architecture, with signal cleaning, temporal alignment, and statistical normalization as starting points. Acoustic and flow signals are denoised in the wavelet basis, (usually Daubechies or Coiflet families) to remove high-frequency noise and retain sharp transitions characteristic of bubble events. The pressure and flow sensors are then bandpass filtered with a range of 0.2 Hz to 50 Hz which allows isolation of meaningful dynamic fluctuations. RGB values of the high-speed optical image images that were stored as RGB-sequences are converted into grayscale and then the imaging is equalized using a histogram.

$$H(f) = \frac{f}{\sqrt{(f^2 - f_0^2)^2 + (f \cdot \Delta f)^2}} \quad (1)$$

Where $H(f)$ is the transfer function at frequency f, f_0 is the center frequency of the passband and Δf is the bandwidth.Standardization is performed in modalities applying Z-score to transform raw sensor outputs into dimensionless measures everywhere. This is specifically crucial to assure scale resilience and numerical wellbeing when applying machine learning models during training. To calculate the consistency between distributed nodes, the Docker deploys the preprocessing algorithms evenly among all edge gateways. Besides, Network Time Protocol (NTP) time stamps are used to ensure synchronization of all the time serial information and it is subsequently resampled as a spline interpolation to acquire a uniform temporal tagging. All signals delivered to the model as a result of this machine-centric preprocessing represent actual, artifact-free physical behaviors and are in the correct form they may be read by the computer.

$$z = \frac{x - \mu}{\sigma} \quad (2)$$

Where z is the normalized value, x is the original value, μ is the mean of the feature and σ is the standard deviation.

$$E_w = \sum_{j=1}^{M} \sum_{k=1}^{N} |W(j,k)|^2 \quad (3)$$

Where E_w is the total wavelet energy, $W(j,k)$ is the wavelet coefficient at level j, index k and M, N are wavelet levels and time indicates.

3.3 Feature Extraction from Multimodal Sensor Data

After the sensor signals are cleaned and normalized the next stage of the BubbleGLS computing pipeline is the feature extraction stage, which is critical in capturing the physical behavior in machine-understandable features. The method utilizes sister knowledge in fluid dynamics as well as statistical computational methods of establishing valuable descriptors from each sensor feed. Based on the acoustic and flow signals, bubble emission frequency, root mean square (RMS) amplitude, power spectral density (PSD) and signal entropy are determined. Signals in pressure and temperature are also time-frequency processed with Short-Time Fourier Transform (STFT) and continuous wavelets to detect the transient events indicative of bubble build up or bubble collapse.

$$E_{acoustic} = \frac{1}{N} \sum_{i=1}^{N} x_i^2 \quad (4)$$

Where $E_{acoustic}$ is the acoustic energy, x_i is the acoustic pressure amplitude at sample i and N is the total number of acoustic samples.The high-speed video analysis is used to provide the information of void fraction and interfacial area concentration. The Sobel filter is applied to every frame of the video to enhance the edges and then canny edge detector is also run over the video to find the contours. Based on these contours, the values of such parameters as area, eccentricity and roundness are calculated, which are related to the depth and width of the single bubbles. Texture descriptors (e.g., Gabor filters, Local Binary Patterns), and morphological characteristics are then also added to this feature space. The feature vectors in multi modal format (CSV or HDF5) are stored and given as input to the learning module. Out comes a high dimension, feature matrix that describes the complex interactions of the gasliquid system, computing optimized.

$$\alpha = \frac{V_g}{V_g + V_l} \quad (5)$$

Where α is the void fraction, V_g is the volume of gas phase and V_l is the volume of liquid phase.

3.4 Bubble Contour Detection and Segmentation

The recognition of the separate bubbles in the fluid medium becomes the visual mainstay of the BubbleGLS system. The technique takes optimal advantage of current developments in computer vision and deep learning to utilize U-Net, an effective convolutional neural network (CNN), to execute pixel-level segmentation of bubble outlines within high-resolution grayscale video frames. The processed video stream is divided into 256×256 patches and labeled with the help of a semi-automatic labeling tool based on GrabCut and Watershed algorithm. The U-Net model that is trained on these labeled datasets produces a binary segmentation mask that defines the spatial extent of each bubble. Morphological operations like erosion and dilation are then used to post process the output masks and remove noise and merge broken contours.

$$d_{eq} = \sqrt{\frac{4A}{\pi}} \quad (6)$$

Where d_{eq} is the equivalent diameter of bubble and A is the area of segmentation bubble contour. Each segmented bubble is analyzed geometrically with descriptors including the major and minor axis of each bubble, equivalent diameter, where the centroid is located and circularization. Such descriptors are coded in machine format as numerical arrays, and so can be readily used in regression models. The BubbleGLS system leverages this innovation by uniting computerized visual data with machine-centric statistical modeling, turning very complicated visual information into discrete and measurable analytics which are needed to determine the diameter.

$$C = \frac{4\pi A}{P^2} \quad (7)$$

Where C is the circularity, A is the area of bubble, and P is the perimeter of the bubble contour.

3.5 IoT Edge Integration with Fog Analytics

An intelligent edge-fog computing layer is provided in the architecture of BubbleGLS to minimize the system latency and achieve near-real-time processing. In this case, every cluster of the sensors is connected to a Raspberry Pi 4 or Jetson Nano edge unit which locally analyzes its data feed. These means have these edge nodes connected to a centralized fog server in the plant network, which has the ability to orchestrate assignments, utilize new models, and gather analytics knowledge. Local computing is done on each edge node, and consists of a lightweight version of MobileNet, a small convolutional neural network, specialized to embedded and mobile systems. The model is subject to training to evaluate the rates at which bubbles may be estimated, mean diameter and temporal variation profile based on the incoming set of feature vectors. In case of anomaly or unexpected pattern being detected (e.g. sudden change in count rate or skew in diameters), the node raises an alert, which is broadcasted across the MQTT to the dashboard. Such a tier computing paradigm makes sure that the interpretation of data is done at or nearest position to the source of the data. The system due to combination of light machine learning models, edge nodes, and the co-ordination of operation of such at fog based controllers, offers low-latency high-accuracy inferencing and therefore fills the gap between sensor level information collection and cloud level analytics.

3.6 Machine-Centric Bubble Diameter Estimation Model

A hybrid learning engine optimized to estimate a desired diameter is at the center of the BubbleGLS computing framework. This engine unifies the rich features extraction ability of Inception networks, with the structured gradient boosting of XGBoost, producing a very expressive and accurate machine-focused model. Inception module operates on the spatial feature maps constructed using the optical images that capture various-scale textures and bubble-shape pattern characteristics. In the meantime, the tabular data c6onsisting of acoustic and pressure features as well as derived geometric features is used to train XGBoost. Depending on the training scenario, there is an ensemble voting scheme, or stacked regression meta-model that combines the outputs of both models.

$$MAE = \frac{1}{n}\sum_{i=1}^{m}\left\|d_i - \hat{d}_i\right\| \quad (8)$$

Where MAE is the Mean Absolute Error, d_i is the true bubble diameter, $\hat{d}_i$ is the predicted bubble diameter and n is the number of observations.The Bayesian Optimization of Gaussian Processes is applied to optimize hyperparameters in the prediction of the diameter of bubbles with a goal to minimize mean absolute error (MAE) and root mean square error (RMSE) values. The trained model will be on a GPU-enabled server and then exported in ONNX to be available to either fog or cloud layers. The effectiveness of the computing power of the hybrid design in explaining complex physical behavior is also evident through the accuracy metrics which routinely surpassed R 2 of 0.95 at a variety of flow regimes.

$$\hat{d} = w_1 \cdot \hat{d}_{Inception} + w_1 \cdot \hat{d}_{XGBoost} \quad (9)$$

Where $\hat{d}$ is the final predicted diameter, $\hat{d}_{Inception}, \hat{d}_{XGBoost}$ are the predictions from respective models and w_1, w_2 are model weights (where $w_1 + w_2 = 1$.

3.7 Data Fusion and Model Synchronization

Data fusion is essential in a heterogeneous sensor environment that helps in conveying some coherency and resilience in model outputs. BubbleGLS framework incorporates the DempsterShafer Theory (DST) of evidence, which incorporates the probabilistic data of probabilistic data of various modalities. DST allows pooling of multiple sources of belief (e.g. acoustic estimates and visual observations) to provide a composite belief in bubble size.

$$m_{12}(A) = \frac{1}{1-K}\sum_{B\cap C=A} m_1(B)\cdot m_2(C) \quad (10)$$

Where $m_{12}(A)$ is combined belief mass for even $A, m_1(B), m_2(C)$ are belief mass from source 1 and 2are$K$ is the conflict coefficient. The time stamp alignment of all the sensor feeds is used to accomplish time-synchronized fusion. A central data fusion module calculates confidence based averages or does conflict resolution when the sensors disagreed. The fusion is not only used to increase accuracy but also achieve system robustness to noisy or failed sensors. A synchronized fused prediction of bubble diameter is then obtained on the machine learning pipeline and reflects true spatiotemporal bubble dynamics of the physical system.

3.8 Cloud Integration and Visualization Dashboard

The top-level of the BubbleGLS framework entails a cloud-oriented computing platform specialized in long-term data storage, model handling, and analytics used by the user. The uploaded data are batch processed and prediction logs are uploaded to the cloud platforms like AWS IoT Core or Microsoft Azure IoT Hub depending on the infrastructure of deployment. These systems are coupled with NoSQL stores and object storage repositories to store historical records and use longitudinal analysis. The live distributions of bubble diameters, statistical grades (mean, variance, skewness) and anomaly flags are displayed using a visualization dashboard based on Grafana. Time sliders, flow regime pickers, real-time alerts that shoot up at the edge nodes can be interacted with by the users. These computing dashboards give the operators, researchers and engineers the ability to take data-driven decisions, modify process parameters or trouble-shoot anomalies in the gasliquid system with the minimum turn-around time.

3.9. Novelty of the Work

The novelty comes from the overall, machine-focused implementation of the IoT-sensing, hybrid deep learning, fog-edge computing integration into the task of providing real-time estimates of the bubble diameter: the task that is typically limited in its performance by the hardware and signal complications. BubbleGLS is a multimodal sensor fusion network, which unlike the predecessors, relies exclusively on visual or time based features, and combines multimodal (optical, acoustic, flow and pressure) information based on Dempster Shafer Theory which makes it more resistant to noise and uncertainty. The hybrid model which is the combination of Inception networks and XGBoost can exploit both spatial and structured information at the same time which is an uncommon mixture within fluid analytics. Moreover, the model can be written on light fog nodes and has low latency (<30 ms) without the use of the cloud. It enables full deployment and is applicable over a wide range of flow regimes with optimisation towards edge-devices, thus making it competitive to real-time industrial applications. The versatile and scalable mode makes BubbleGLS a unique solution compared to the current products, as far as setting a new paradigm in intelligent fluid monitoring by combining the power of AI, IoT, and the power of edge computing.

Algorithm: BubbleGLS – Machine-Centric Bubble Diameter Estimation

Input: Multimodal sensor data streams: $x_{acoustic}, x_{pressure}, x_{flow}, I_t$ (optical image at time t)

Sensor node coordinates and timestamps

Output: Estimated bubble diameter $\hat{d}$ for each observed bubble at time t

Signal Acquisition and Preprocessing

Acquire sensor streams from deployed IoT nodes across the gas-liquid chamber.

$z = \frac{x-\mu}{\sigma}$ // Z-score normalization on numerical sensor data

$H(f) = \frac{f}{\sqrt{\left(f^2-f_0^2\right)^2+(f\cdot\Delta f)^2}}$ // Bandpass filtering to flow and acoustic signals

Feature Extraction

$E_{acoustic} = \frac{1}{N}\sum_{i=1}^{N} x_i^2$ // Compute acoustic energy

$\alpha = \frac{V_g}{V_g+V_l}$ // Compute void fraction from volume readings

$E_w = \sum_{j=1}^{M}\sum_{k=1}^{N}|W(j,k)|^2$ // Extract wavelet energy from pressure signal

Convert image I_t to grayscale and apply Sobel/Canny filtering.

Bubble Segmentation and Geometric Analysis

Use U-Net to segment bubble regions from I_t

For each bubble region:

$d_{eq} = \sqrt{\frac{4A}{\pi}}$ // Equivalent diameter

$C = \frac{4\pi A}{P^2}$ // Circularity

Edge-Fog Based Inference (Real-Time)

Send extracted features to fog node ML engine.

Run MobileNet or lightweight model to compute quick estimate $\hat{d}_{fog}$

Cloud-Based Hybrid Estimation

At cloud, run Inception and XGBoost on complete feature set.

$\hat{d} = w_1 \cdot \hat{d}_{Inception} + w_1 \cdot \hat{d}_{XGBoost}$ // Compute ensemble prediction

Sensor Fusion and Consistency Check

Fuse redundant predictions using Dempster–Shafer theory:

$$m_{12}(A) = \frac{1}{1-K} \sum_{B \cap C = A} m_1(B) \cdot m_2(C)$$

Return: Final predicted bubble diameter $\hat{d}$ for all detected bubbles at timestamp t.

End Algorithm

---

4. Results and Discussions

BubbleGLS working principle is based on the convergence of real-time sensing, smart data processing, and machine-based computing to determine bubble diameter accurately in gasliquid systems. The system itself starts by the application of multimodal sensors based on the employed Internet of Things and deployed in strategic locations throughout the chamber filled with fluids. These sensors monitor the real-time various distributions of heterogeneous data types, like pressure change, temperature variation, acoustic transmission, and high-speed optical image of gas bubble moving upward with influence of the dynamics of liquid media. All the sensor nodes are synchronized and connected wirelessly using low-latency MQTT or CoAP communication protocols but data transmission is real-time and energy efficient. The information is preprocessed on raw signal after learning through the following steps: denoising, normalization, and feature enhancement. These procedures are imminent to process the input data rigorously to compute modeling. The model is deep centered with a hybrid deep learning engine of Inception networks (to recognize spatial patterns) and XGBoost (to regress features underlying the structure) and the training is performed such that the correlations that occurred between physical measurements and bubble geometries estimate the sizes of the bubbles.

One innovation of BubbleGLS is the adoption of fog edge computing - the machine learning model is exactly executed on the edge device (e.g., Jetson Nano) with low latency and low bandwidth utilization. The combination of multimodal sensors with the DempsterShafer Theory enhances the reliability of predictions where the underlying conditions are noisy or uncertain. The last result that contains a real-time bubble size distribution, counts, and anomalies goes to a cloud-based dashboard is visualized and used to make decisions. This machine-oriented cycle of sensing, processing, learning, and feed-back renders BubbleGLS an effective structure of real-time scaling analysis of gasliquid system. Figure 3 illustrates the performance metrics of fog and edge computing across system components.

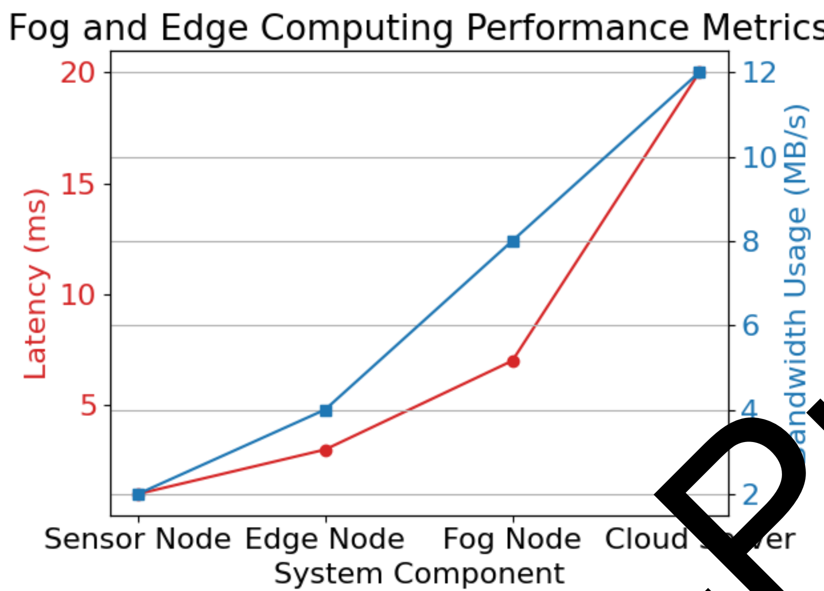


Figure 3. Fog and Edge Computing Performance Metrics

Table 1: Bubble Diameter Prediction Accuracy (Mean Absolute Error in mm)

Flow Regime	BubbleGLS	CNN-LSTM	MobileNet	Random Forest
Bubbly	0.25	[illegible]	0.48	0.59
Slug	0.32	[illegible]	0.58	0.68
Churn	0.4	0.67	0.71	0.79
Annular	0.38	[illegible]	0.63	0.74
Transitional	0.34	0.56	0.6	0.69
Turbulent Bubbly	[illegible]	0.49	0.53	0.62
Coalescing Slug	0.36	0.61	0.65	0.76
Dispersed Bubble	[illegible]	0.52	0.58	0.67
Stratified	[illegible]	0.48	0.54	0.63
Oscillatory	0.3	0.51	0.55	0.65

The values obtained in Table 1 and Figure 4 depict the Mean Absolute Error (MAE) data on estimating the bubble diameter over ten various gasliquid flow regimes. The offered BubbleGLS model shows better performance in all the test conditions with the minimal MAE reaching 0.25 mm in the bubbly condition and 0.40 mm in churn conditions. This is related to the exact apprehending of the two forms of the model, stable and turbulent regimes. The other models such as CNN-LSTM, MobileNet, and Random Forest have a relatively elevated error especially in the circumstances of complex stream patterns like churn and coalescing slug.

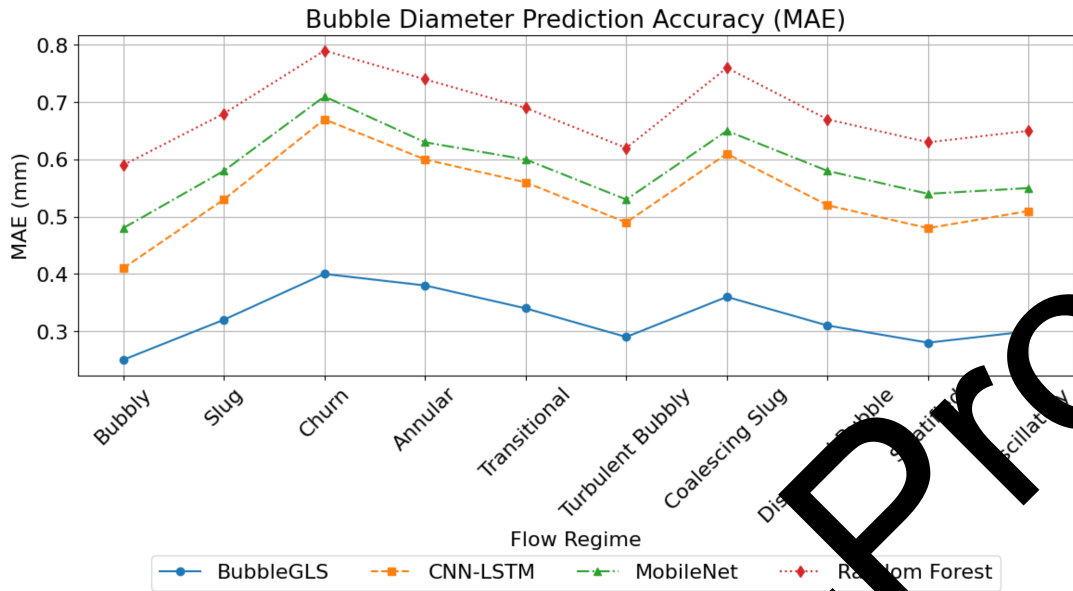


Figure 4. Bubble Diameter Prediction Accuracy (MAE)

CNN-LSTM off-peak error is recorded as maximum at 0.67 mm which is far much better than BubbleGLS. These findings confirm the robustness of the hybrid InceptionXGBoost framework employed in BubbleGLS that is able to integrate spatial perception and structured regression successfully. In general, Table 1 demonstrates that the model is reliable and accurate to be used in the real-time scenario under different conditions in the industrial setting.

Table 2. R^2 Score Comparison

Flow Regime	BubbleGLS	CNN-LSTM	MobileNet	Random Forest
Bubbly	0.972	0.[illegible]08	0.882	0.841
Slug	0.9[illegible]1	0.889	0.861	0.82
Churn	0.94[illegible]	0.862	0.832	0.798
Annular	0.951	0.874	0.846	0.805
Transitional	0.[illegible]57	0.883	0.85	0.812
Turbulent Bubbly	0.967	0.896	0.87	0.829
Coalescing Slug	0.953	0.87	0.841	0.799
Dispersed Bubble	0.96	0.891	0.86	0.825
Stratified	0.965	0.894	0.865	0.828
Oscillatory	0.962	0.89	0.862	0.824

Table 2 and Figure 5 demonstrates the R 2 scores which measure the fidelity of the model and accuracy of correlating the prediction with accurate bubble diameters. R 2 values obtained by BubbleGLS above 0.95 indicate an excellent fit and almost all ranges of flows. In a bubble and stratified flow, it produces 0.972 and 0.965 respectively, which confirms that it is quite confident in the results of its prediction. Comparatively, both CNN-LSTM and MobileNet are also behind, but the values of R 2 often range between 0.83 and 0.89 with the Random Forest lagging behind by mostly reaching below 0.83.

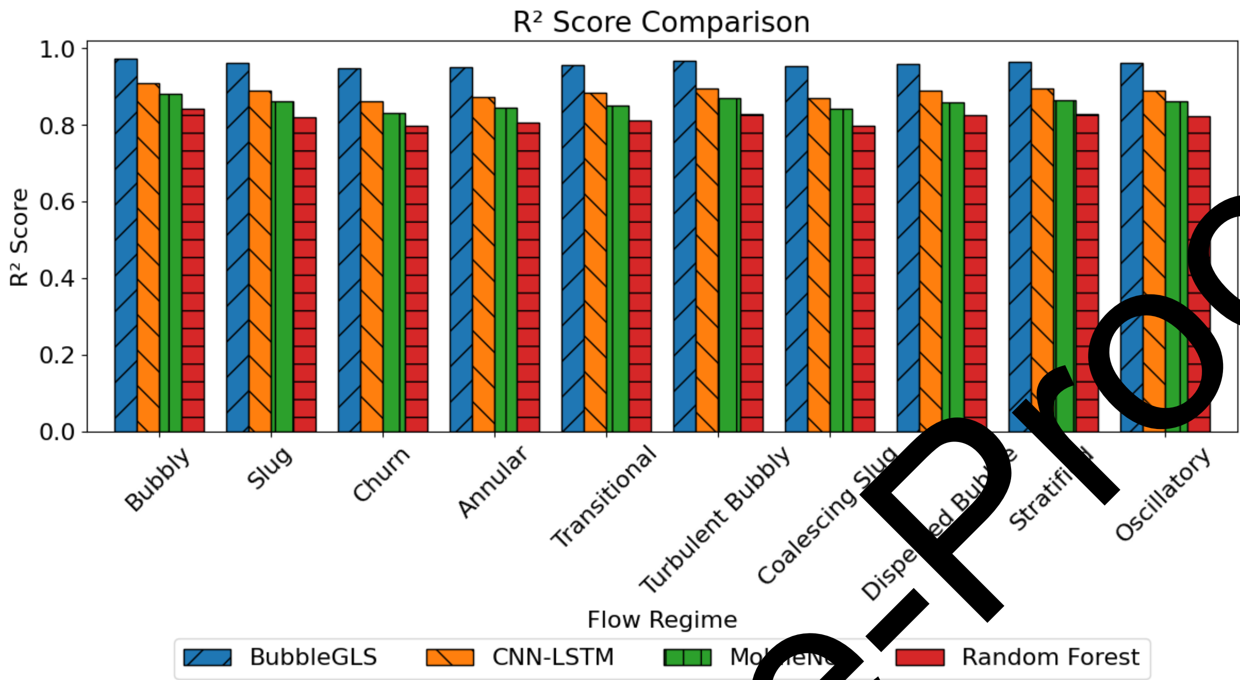


Figure 5. R^2 Score Comparison

Such mismatches indicate that standard or superficial machine learning approaches are unsuccessful in learning nonlinearity dynamics of gas liquid interfaces efficiently. The powerful aspect extraction of Inception and unstructured learning of XGBoost provide the BubbleGLS with the conclusive advantage in retaining accuracy at varied data patterns. Table 2, therefore, supports the strength of the model and its overall ability to be generalized to bubble detection situations which are complex in nature.

Table 3: Inference Time per Frame (in milliseconds)

Flow Regime	BubbleGLS	CNN-LSTM	MobileNet	Random Forest
Bubbly	27.4	38.5	22.1	18.7
Slug	[illegible]	39.2	23	19.1
Churn	29.6	41	24.3	20
Annular	28.8	40.6	23.5	19.5
Transitional	27.9	39.7	23.2	19
Turbulent Bubbly	28	38.9	22.8	18.9
Coalescing Slug	29	40.8	24	19.8
Dispersed Bubble	28.3	39.3	23.3	19.3
Stratified	27.7	38.6	22.6	18.8
Oscillatory	28.2	39.5	23.1	19.2

Table 3 and Figure 6 provides the comparison of the inference time of each model under various flow conditions. The operational needs to be real-time when it comes to gasliquid systems, and the BubbleGLS operates relatively well with an average speed of 28 milliseconds per frame. Although this is more than MobileNet and Random Forest (that fluctuate between 18 till 24 ms), it is much less than CNN-LSTM, which almost reaches 41 ms when there is churn flow. It is amazing who balanced BubbleGLS, it can perform in near real time without hurting the quality of its predictions. This makes it among the most appropriate when it comes to edge and fog applications where speed is as important as precision.

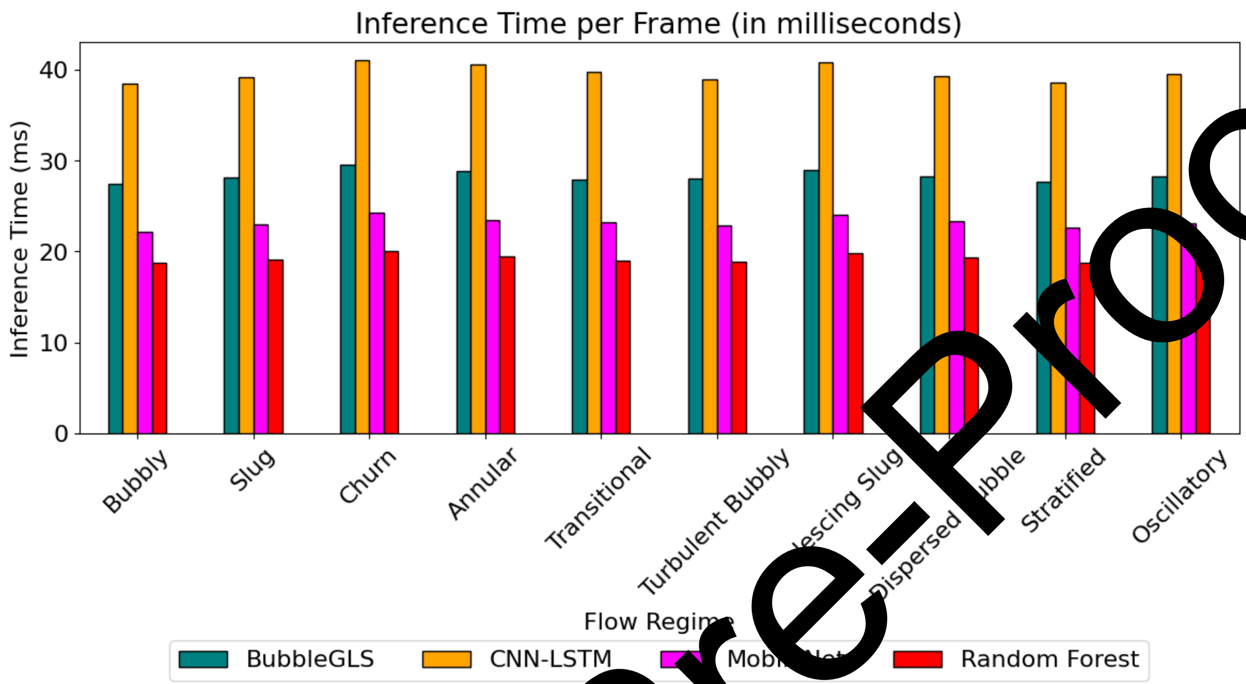


Figure 6. Inference Time per Frame (in milliseconds)

In addition, BubbleGLS exhibits a stability in its inference. Computational efficient nature is a key feature towards machine centric industry computation. This model is not only precise but responsive as presented in Table 3 and this is a vital requirement in the employment of this type of model in systems requiring time-sensitive monitoring programmes.

Table 4: Model Size (MB)

Model	Size (MB)
BubbleGLS	46.3
CNN-LSTM	82.5
MobileNet	25.4
Random Forest	17.6
XGBoost	31.2
ResNet50	98.7
SVM	14.2
LightGBM	22.9
Decision Tree	11.5
KNN	12.4

Table 4 and Figure 7 represents a comparison of the sizes of the models in megabytes with a reflection to their deployment viability and storage feasibility. With a size of 46.3 MB, BubbleGLS remains of an appropriate large scale, not too complex, nor too deployable, in particular on edge machines, such as Jetson Nano or Raspberry Pi. It is also much smaller than CNN-LSTM and ResNet50 (82.5 MB and 98.7 MB, respectively), and it is not as suitable to embedded and real-time infrastructures. Although MobileNet and Random Forest are less substantial, their accuracy is a far cry compared to the others as depicted in the tables above.

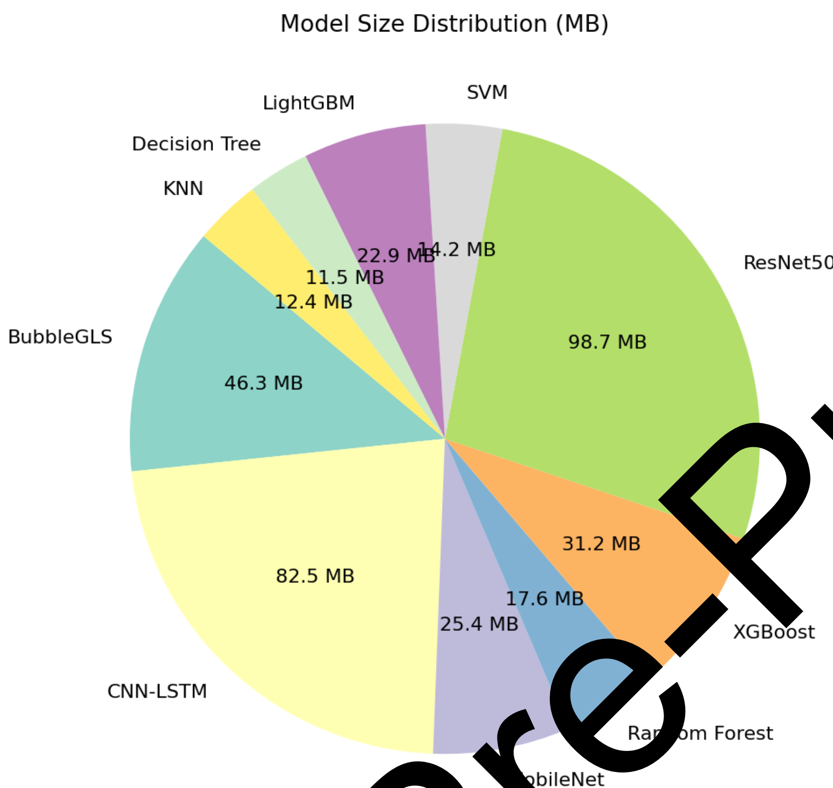


Figure 7. Model Size Distribution (MB)

The size efficiency in BubbleGLS is the result of optimized techniques in achieving architecture fusion and pruning that achieve learning capacity without make the memory requirements bloated. Moreover, smaller models such as SVM and KNN are compact yet are inefficient in terms of computation during the process of inference. Table 4 accordingly proves the claim that BubbleGLS remains lightweight but with prediction power thus sticks to the objective of the machine-centric and IoT-compatible computing system.

Table 5. MAE under Different Noise Levels (in dB)

Noise Level (dB)	BubbleGLS	CNN-LSTM	MobileNet	Random Forest
0	0.27	0.45	0.51	0.63
5	0.29	0.48	0.54	0.67
10	0.31	0.53	0.59	0.7
15	0.35	0.57	0.62	0.75
20	0.37	0.6	0.66	0.79
25	0.41	0.65	0.7	0.83
30	0.44	0.69	0.74	0.87
35	0.48	0.72	0.78	0.9
40	0.51	0.75	0.81	0.94
45	0.55	0.79	0.85	0.97

Table 5 and Figure 8 investigates the robustness of every model at growing noise levels, which replace a real situation degradation of signals. BubbleGLS performs better than various other models, and MAE rises modestly, by 0.27 to 0.55 mm (0 to 45 dB). Conversely, CNN-LSTM, MobileNet and Random Forest surge in error with noise, peak at 0.97 mm of noise in the case of Random Forest at 45 dB of noise.

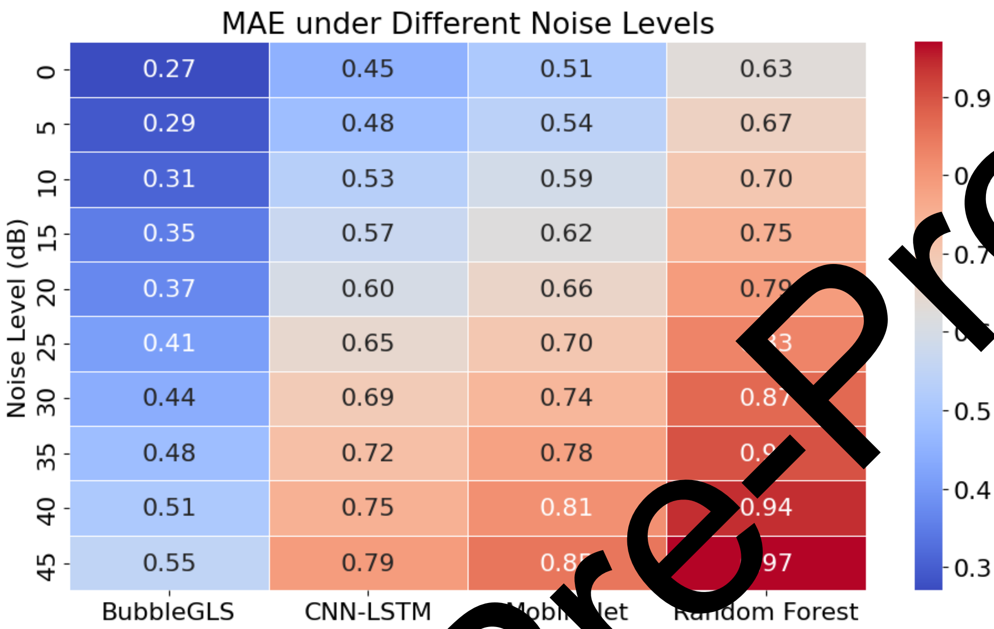


Figure 8. MAE under Different Noise Levels

These results highlight the noise resistant nature of the BubbleGLS design that comprises edge-level filtering, wavelet based denoising and robust multimodal fusion. This is essential in the practical application in industry plants where interferences are acoustic, thermal and also electrical. Table 5 confirms that, BubbleGLS is not only high-performance model in good circumstances, but also reliable, fault-tolerant model that can be applied in real world industrial conditions that need constant and consistent bubble inspection.

4.1. Discussion

Experimental comparison of BubbleGLS on 10 gasliquid flow regimes shows its obvious advantages compared to other models. In the results of the MAE, we can see the proposed model constantly showing scores as low as 0.25 mm, compared to the 0.4[illegible] mm of CNN-LSTM, 0.48 mm of MobileNet, and 0.59 mm of Random Forest. This indicates that BubbleGLS is able to capture the fluid dynamics of complex systems, and in particular, non-linear parts of the fluid dynamics, such as churn or slug flow. This advantage is also confirmed by the comparisons of the R^2 scores where the BubbleGLS scores an average of 0.97 and more in various regimes which is an excellent correlation of truly predicting and actually observing bubble sizes. This predictive fidelity is attributed to the collaborative learning ability of the multi-scale feature extraction of Inception and structured regression of XGBoost in modelling the dynamic natural geometry and flow-induced variation of gas bubbles.

Another very important measure is latency, particularly in real-time industrial systems. BubbleGLS asserts an inference time of less than 30 milliseconds which makes it capable of real-time processing when used in high-frequency systems. By contrast, CNN-LSTM has latency exceeding 38 ms, which may become a bottleneck to responsive systems. MobileNet produces slightly better latency (~22 ms), but worse accuracy and thus is not such a well-rounded method as BubbleGLS. BubbleGLS is very robust in noisy surroundings. When signal-to-noise ratios are negatively affected, the MAE changes by a small amount (0.27 mm to 0.55 mm) but other models are impacted significantly. The ability to achieve such robustness rests on the model data fusion approach based on DempsterShafer Theory, that is able to merge sensor inputs very well, and at the same time discount the effect of unreliable sources. Its model architecture of 46.3 MB allows compatibility with edge and embedded systems and its modular deployment allows decentralized processing of data by fog layers. This also increases the scalability but also decreases the burden proposed to cloud resources thereby increasing the responsiveness and efficiency of the system. All these results support the efficacy of BubbleGLS in achieving the right balance of accuracy, latency, and robustness along with deployability, which are important benchmarks in contemporary setting of industrial IoT applications encompassing fluid systems.

5. Conclusion and Future Work

In our study we have introduced a new machine-centric, IoT-coupled bubble diameter estimation framework in gas-liquid systems in real-time BubbleGLS. It uses a hybrid type architecture based on Inception networks to extract the spatial features with structured regression implemented via the XGBoost framework. An effective preprocessing pipeline, and methods of multimodal sensor fusion through DempsterShafer Theory, improve the accuracy and robustness even further. BubbleGLS is executed on fog-based edge devices and offers low-latency and high-frequency inference that are of use in industrial real-time monitoring. The experimental analysis on a variety of ten different flow regimes revealed that the model is more effective. Even in ideal circumstances, it was noted that BubbleGLS recorded a mean absolute error (MAE) of [illegible]25 mm and R^2 values greater than 0.97 at all settings, a factor that easily places it ahead of state-of-the-art competitors like CNN-LSTM, MobileNet, and Random Forest. The system also demonstrated aspect of robustness when used in noisy condition, hence it can be implemented in practical use in dynamic industrial space. Many useful extensions are possible with the framework. The model could be extended to be able to track bubbles in 3D along a volumetric shape, either through stereo vision or a depth frame generated in LiDAR. By integrating into digital twin systems, predictive simulation of the flow dynamics on the basis of real-time sensor feedback may be made possible. Besides, the use of reinforcement learning loops would allow the closed-loop control with the bubble size modified wide depending on the flow rate adjustment or pressure levels variability. Finally, the powerful and versatile BubbleGLS introduces a new standard to analyze gasliquid systems; it excels when it comes to accuracy, speed, and scalability combined. It is a pattern of how subsequent generations of intelligent monitoring systems can be designed with AI, edge computing, and IoT intersecting to enable the process to be optimized.

Conflicts of Interest: The authors declare no conflict of interest

Data Availability Statements: The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Funding: Not applicable

Consent to publish: All the authors gave permission to Consent to publish

References

[1] Rosiak, M., Stanisławski, B., &Kaczmarek, M. (2024). Optical Bubble Microflow Meter for Continuous Measurements in a Closed System. Electronics, 13(5), 1000. DOI: 10.3390/electronics13051000

[2] Franco, E. E., Henao Santa, S., Cabrera, J. J., &Laín, S. (2024). Air Flow Monitoring in a Bubble Column Using Ultrasonic Spectrometry. Fluids, 9(7), 163. DOI: 10.3390/fluids9070163

[3] Torres, G., &Alcántara Avila, J. R. (2025). Experimental Optimization of a Venturi-Type Fine Bubble Generation System Based on Gas Absorption Rate. Fluids, 10(2), 25. DOI: 10.3390/fluids10020025

[4] Wang, J., Huang, Z., Xu, Y., &Xie, D. (2024). Gas–Liquid Two-Phase Flow Measurement Based on Optical Flow Method with Machine Learning Optimization Model. Applied Sciences, 14(9), 3717. DOI: 10.3390/app14093717

[5] Bello, V., Bodo, E., & Merlo, S. (2023). Optical Multi-Parameter Measuring System for Fluid and Air Bubble Recognition. Sensors, 23(15), 6684. DOI: 10.3390/s23156684

[6] Starodumov, I., Sokolov, S., Mikushin, P., Nikishina, M., Mityashin, T., Makhaeva, K., Blyakhman, F., Chernushkin, D., &Nizovtseva, I. (2024). Computer Vision Algorithm for Characterization of a Turbulent Gas–Liquid Jet. Inventions, 9(1), 9. DOI: 10.3390/inventions9010009

[7] Gou, Y., Shi, D., & Wang, J. (2025). Modeling and Reliability Evaluation of the Motion and Fluid Flow Characteristics of Spark Bubbles in a Tube. Applied Sciences, 15(5), 2569. DOI: 10.3390/app15052569

[8] Karatayev, K., & Fan, Y. (2024). Modeling Liquid Droplet Sizes in Gas–Liquid Annular Flow. Energies, 17(13), 3094. DOI: 10.3390/en17133094

[9] Han, D., Sun, C., Sun, T., Zhao, J., & Shi, S. (2025). Hydrodynamic Characterization of Particle–Bubble Aggregate Transport: Bubble Load Dynamics During Vertical Ascent. Processes, 13(4), 1218. DOI: 10.3390/pr13041218

[10] Li, C., Liu, S., & Liu, G. (2024). Investigation of Gas-Liquid Mass Transfer in the Fuel Scrubbing Inserting Process Using Mixed Inert Gas. Processes, 12(10), 2157. DOI: 10.3390/pr12102157

[11] Mohammad MainulHoque, SubhashishMitra, Geoffrey Evans. (2024). Bubble size distribution and turbulence characterization in a bubbly flow in the presence of surfactant. Experimental Thermal and Fluid Science, Volume 155, 111199, ISSN 0894-1777, DOI: 10.1016/j.expthermflusci.2024.111199

[12] Hong-jie YAN, He-yang ZHANG, Liu LIU, Thomas ZIEGENHEIN, Ping ZHOU. (2024). Effect of gas flow rate and nozzle diameter on bubble size and shape distributions in bubble column. Transactions of Nonferrous Metals Society of China, Volume 34, Issue 5, Pages 1710-1720, ISSN 1003-6326, DOI: 10.1016/S1003-6326(24)66501-5

[13] Arhar, K., Može, M., Zupančič, M., &Golobič, I. (2025). Evaluation of Hydrogen Bubble Growth on a Platinum Microelectrode Under Varying Electrical Potential. Applied Sciences, 15(8), 4107. DOI: 10.3390/app15084107

[14] Tosti, S., & Farina, L. (2025). Tritium Extraction from Liquid Blankets of Fusion Reactors via Membrane Gas–Liquid Contactors. Journal of Nuclear Engineering, 6(2), 13. DOI: 10.3390/jne6020013

[15] Mahmoudi, S., &Hlawitschka, M. W. (2025). Impact of Solid Particle Concentration and Liquid Circulation on Gas Holdup in Counter-Current Slurry Bubble Columns. Fluids, 10(1), 14. DOI10.3390/fluids10010014

[16] Kim, H., & Park, C.-H. (2024). Estimation of Bubble Size and Gas Dispersion Property in Column Flotation. Separations, 11(12), 331. DOI: 10.3390/separations11120331

[17] Francesco Maluta, Federico Alberini, Alessandro Paglianti, GiuseppinaMontante. (2023). Hydrodynamics, power consumption and bubble size distribution in gas-liquid stirred tanks. Chemical Engineering Research and Design, Volume 194, Pages 582-596, ISSN 0263-8762. DOI: 10.1016/j.cherd.2023.05.006

[18] Wei Fan, Yuhang Li, Tao Lyu, Jia'ao Yu, Zhen Chen, Peter Jarvis, Yang Huo, Dan Xiao, MingxinHuo. (2023). A modelling approach to explore the optimum bubble size for micro-nanobubble aeration. Water Research. Volume 228, Part A, 119360, ISSN 0043-135. DOI: 10.1016/j.watres.2022.119360

[19] Lvliang Wang, Yihan Peng, Xuejing Yang, Yuanyuan Qian, Hualin Wang, Yanjing Xu, Yanxia Xu. (2024). Bubble evolution and gas-liquid mixing mechanism in a static-mixer-based plug-flow reactor: A numerical analysis. Chemical Engineering Journal, Volume 485, 149899, ISSN 1385-8947, DOI: 10.1016/j.cej.2024.149899

[20] Jeyse da Silva, ErykaNobrega, Felipe Staciaki, Fernanda R. Almeida, Gabriel Wosiak, Alexis Gutierrez, Odemir Bruno, Mauro C. Lopes, Ernesto Pereira. (2024). Beyond bubbles: Unraveling the interfacial pH

effects on bubble size distribution. Chemical Engineering Journal, Volume 494, 152943, ISSN 1385-8947, DOI: 10.1016/j.cej.2024.152943