Volume estimation of liquid structures on shadowgraphy images using deep learning and high-fidelity simulations

Florian Granger, Magdeleine Airiau, Robin Devillers, Laure-Emilie Martin, Luc-Henry Dorey, Jean-Christophe Hoarau, Davide Zuzio, Jean-Luc Estivalezes

Volume estimation of liquid structures on shadowgraphy images using deep learning and high-fidelity simulations

Florian Granger[1,3], *Magdeleine Airiau*[1,3], *Robin Devillers*[1], *Laure-Emilie Martin*[1], *Luc-Henry Dorey*[1], *Jean-Christophe Hoarau*[1], *Davide Zuzio*[2], *Jean-Luc Estivalezes*[2]

[1]*DMPE, ONERA, Université Paris Saclay, F-91123 - Palaiseau, France*
[2]*DMPE, ONERA, Université Toulouse, F-31055 - Toulouse, France*
[3]*CNES - Transport Spatial - Etablissement de Daumesnil, F-75612, Paris Cedex, France*

florian.granger@onera.fr · magdeleine.airiau@onera.fr · robin.devillers@onera.fr

Abstract

This study presents a novel approach to accurately estimate the volume of atomized liquid structures in the combustion chamber of liquid rocket engines. Accurate volume estimation is challenging due to the 2D nature of shadowgraphy images. The approach aims to improve accuracy by leveraging deep learning and simulation advances, using computational fluid dynamics simulations as ground truth for training a neural network. Results are compared to classical image analysis algorithms.

1. Introduction

Liquid propellant rocket engines (LREs) are critical for space exploration and satellite deployment, serving as the primary propulsion system for various launch vehicles. Combustion within the chamber of a liquid rocket engine occurs under extreme conditions that demand precise control and optimization. The combustion chamber experiences high pressures, temperatures, and velocities, creating a harsh environment where efficient fuel atomization and mixing are vital to achieve optimal combustion efficiency. One common type of fuel injection system employed in liquid rocket engines is the coaxial injector. It features an inner oxidizer cylinder surrounded by an annular fuel passage, enabling an efficient mixing of the fuel and the oxidizer. Under the extreme conditions prevailing in the combustion chamber, the atomization process occurs through a combination of primary breakup mechanisms, secondary breakup, and subsequent dispersion of the liquid oxidizer. The primary breakup, the initial step in atomization, involves the disintegration of the liquid oxidizer into liquid structures of many sizes and shapes due to the interaction between the high-velocity gas flow and the oxidizer stream. The morphology of atomization regimes for coaxial jets has been classified based on non-dimensional parameters such as the Reynolds number (Re) and Weber number (We) [4]. Among these regimes, the fiber regime, encountered in liquid rocket engines, is the most violent. It is characterized by the formation of long, thin liquid threads or fibers that are stretched and elongated due to the aerodynamic forces exerted by the surrounding gas stream. These fibers exhibit high aspect ratios (length-to-diameter ratios), making them unstable and are prone to further break up into smaller droplets or detached ligaments. While empirical laws [16] exist for average atomized liquid structures, they provide limited information about the local instantaneous distribution or shape. However, accurately quantifying and understanding the distribution of atomized structures within the combustion chamber is crucial, as it directly influences combustion characteristics, including fuel-oxidizer mixing, flame stability, and combustion efficiency. A proper understanding of the atomized structures is essential to optimize fuel injection strategies, ensuring combustion stability, and improving overall engine performance.

Simulations, such as Direct Numerical Simulations (DNS), provide a powerful tool to study the atomization process and obtain detailed information about liquid structures [11]. However, the computational resources required for DNS are substantial, making it a costly option, particularly for parameter studies in an industrial context. On the experimental front, granulometry measurements using laser-based techniques [7] and shadowgraphy imaging [18] have proven to be effective in investigating atomized structures. Laser measurements enable the probing of a small portion of the spray, providing localized information about the size and spatial distribution of droplets. On the other hand, shadowgraphy imaging offers a comprehensive view of the entire spray length, allowing for direct visualization and measurement of the liquid structures within the combustion chamber. By capturing high-resolution images of the atomized structures, shadowgraphy imaging enables detailed analysis of their size, shape, and spatial distribution. These visual data serve as valuable inputs for quantitative analysis and can be coupled with image processing algorithms

to extract precise measurements of the liquid structure's sizes. However, the visualized droplets on the image are 2D projections of real 3D liquid structures: equivalent diameters or characteristic lengths estimated in 2D are hard to convert into 3D data such as accurate volume values, even more with the additional uncertainty due to blurred droplets. While stereo depth estimation is an old computer vision task [2], monocular depth estimation and volume reconstruction was performing poorly until recent deep learning advances [9].

In this study, we aim to enhance the volume estimation of liquid structures in monocular shadowgraphy images by leveraging the power of deep learning and high-fidelity simulations. The novel approach proposed combines experimental data from shadowgraphy imaging with projections of liquid structures obtained from DNS. These CFD simulations serve as ground truth data to train and test the neural network. The trained neural network will subsequently read the 2D shadowgraphy images and give back an accurate estimation of the actual liquid structures size. By exploiting the complementary strengths of experimental measurements and numerical simulations, the accuracy and reliability of volume estimation for liquid structures in shadowgraphy images can be improved. The subsequent sections of this article elaborate on the methodology for liquid structure volume estimation using deep learning and high-fidelity simulations. The details of the neural network architecture used in this study is presented and the creation of a learning and test database that incorporates high-fidelity simulations, ranging from baseline liquid/gas interface visualizations to realistic ray-tracing renderings is discussed. Furthermore, the results obtained from the deep learning model are presented and compared with those using classical image analysis algorithms.

2. Methodology

Traditionally, the size estimation of liquid structures involves applying a threshold to the pixel intensity and computing the projected diameter of the liquid structure, denoted as d_{proj}, by integrating its projected surface S_{proj}. This estimation method is based on the assumption that the liquid structures can be approximated as spherical drops:

$$d_{\text{proj}} = \sqrt{\frac{4S_{\text{proj}}}{\pi}} \, . \tag{1}$$

The projected surface can then be related to the volume of the liquid structure V, further referred as projected volume, as:

$$\frac{S_{\text{proj}}^{3/2}}{V} = \frac{3}{4}\sqrt{\pi} \approx 1.33 \, . \tag{2}$$

However, the fiber regime which has been extensively observed and described in the literature by [18, 13] is showing a wide range of liquid structure sizes and non-spherical shapes. In support of this, an experimental shadowgraphy image obtained from the water/air coaxial configuration by [8] is shown in Figure 1, illustrating the diverse range of liquid structure sizes and shapes present in the system. The volume of some of them, for example the 3$^{\text{rd}}$ one cannot be accurately estimated with a sphericity hypothesis.

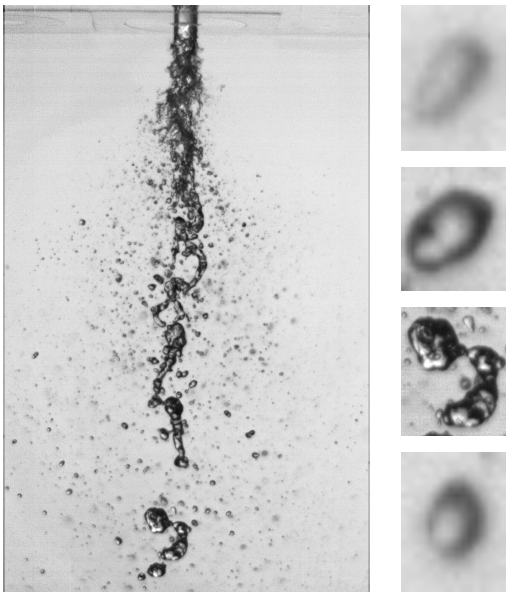


Figure 1: Shadowgraphy from [8] with some isolated liquid structures

The absence of stereo imaging results in the loss of depth information, further complicating the accurate recovery of liquid structure volumes. Deep learning can compensate for the loss of depth information by leveraging prior knowledge acquired during training. However, obtaining a reliable learning and test database with accurate liquid structure volume information poses a significant challenge. In this study, high-fidelity simulation from [11], based on the experimental configuration from [8], has yielded extensive insights into the characteristics of liquid structures atomized from a coaxial injector operating in the fiber regime. This simulation provides a comprehensive dataset of liquid structures with precise volume and location information. By combining these simulation results with rendering software, a valuable dataset of isolated liquid structures volume and the corresponding projection is obtained. This dataset serves as the training and testing data for the deep learning network as illustrated in Figure 2. A detailed explanation of each component will be presented in the subsequent sections.

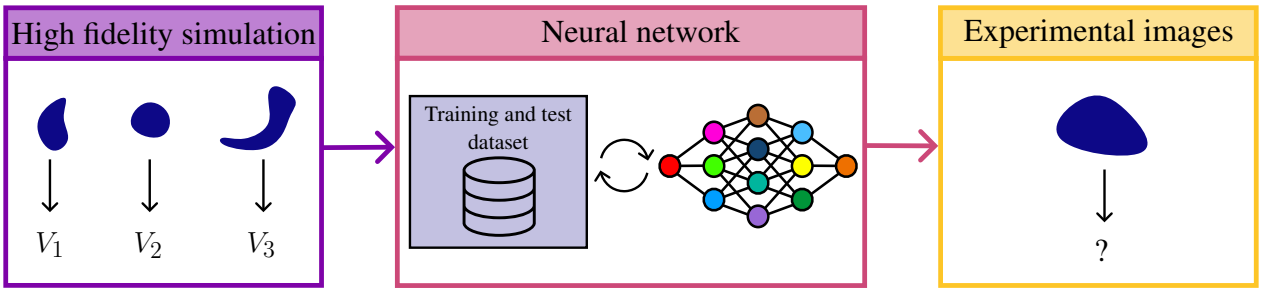


Figure 2: Methodology to combine deep learning and high fidelity simulations for liquid structure volume estimation

3. Neural network

3.1 Image Classification with Convolutional Neural Networks

Image classification is a fundamental task in computer vision that categorizes images into predefined classes or categories. The goal is to develop algorithms and models that automatically assign a label or class to an input image based on its visual features. The most effective method in this field uses Convolutional Neural Networks (CNNs), a learning architecture designed for computer vision [5]. Based on the convolution filters, they capture both low-level features (e.g., edges, textures) and high-level semantic information. Training a CNN model for image classification involves providing a large labeled dataset called Ground Truth (GT): the objective is to minimize the difference between the predicted class labels and the GT in the training data.

A neural network set-up involves two phases: the learning phase to build the model and the test phase to assess the model's ability to classify images correctly. The training process involves feeding the input images forward through the network, computing the predicted class probabilities, and comparing them with the true class labels using a loss function. It learns and optimizes its internal parameters (weights and biases) through a process called backpropagation. This iterative process continues for a defined number of iterations, called epochs, or until a convergence criterion is met, ensuring the model learns to classify images accurately. Once the CNN is trained, the test phase is performed to evaluate its performance. A separate dataset, called the test set, is used for this evaluation. This dataset must be distinct from the training dataset to provide an unbiased assessment. During testing, the CNN takes the unseen input images and predicts their class labels. The predicted labels are compared with the true labels to measure the model's performance using various evaluation metrics. The test phase helps assess the model's effectiveness and provides an indication of its performance in real-world scenarios.

3.2 Classification evaluation metric

Accuracy is a widely used metric for evaluating the performance of classification models. It measures the proportion of correctly classified instances from the total number of instances in a dataset by dividing the number of correctly classified instances by the total number of instances :

$$\text{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of predictions}} \tag{3}$$

While accuracy is a straightforward and intuitive metric, it does have limitations. It assumes that all classes are equally important and that misclassifications have the same cost. It is also recommended to consider a combination of

evaluation metrics to get a comprehensive understanding of the model's performance such as confusion matrices. These matrices allow to directly visualize the performance of an algorithm by comparing instances. Each row represents the real class associated with the object, and each column, the estimated class. For classification with n_{class}, let $i, j \in n_{\text{class}}$, two classes, the cell $c_{i,j}$ contains the number of elements of the real class i which have been estimated as belonging to the class j. If the classification is perfect, the confusion matrix is, therefore, a diagonal matrix with zeros on the other positions of the matrix. In this work, the accuracy and confusion matrix with metrics based on physical criteria, such as the Probability Density Function (PDFs) will be combined, in order to take into account the physics associated with the phenomena.

3.3 Volume classification

In the present study, the aim is to estimate the liquid structure's volume from images. However, volume reconstruction using 2D images is a particularly complex branch of computer vision, requiring large non-trivial neural networks [9]. For this reason, we have chosen to focus on classifying the structures into volume classes rather than trying to find the exact volume of the structure. The spectrum of volumes is discretized into different classes of variable sizes and bins, like a histogram. Every individual liquid structure is identified with an image and a label indicating which volume class it belongs to. The principle is, therefore, to classify the structures in the correct volume class and thus estimate where they lie in the volume histogram. Given the size of the database, which will be detailed in the next section, a large number of classes can be obtained for a refined classification.

To classify structures, a classic AlexNet-inspired classification network [12] with 9 layers, 6 convolution layers, and 3 fully connected layers is used. The loss function is calculated by cross-entropy. The training and test images are drawn randomly from the image database of structures.

4. Training database from high fidelity DNS

This section delves into the generation of projections for the training and test database, derived from high-fidelity Direct Numerical Simulations (DNS). With the deep learning architecture in place, the emphasis now shifts toward creating a robust dataset for training and testing purposes.

To isolate liquid structures within the simulation, a baseline image analysis technique is employed as the initial step. This process involves using a tool developed by [1], which facilitates the creation of statistical data for each individual structure present in the simulation. This tool provides valuable information such as the center of mass, the velocity, the interface surface, the volume, and an enclosing box for each structure. Leveraging this information, it becomes possible to isolate a single structure from the surrounding flow, thereby enabling focused analysis and characterization of individual liquid structures. Figure 3 plots an example from [11] of some isolated liquid structures after atomization with their enclosing boxes highlighting their accurate detection.

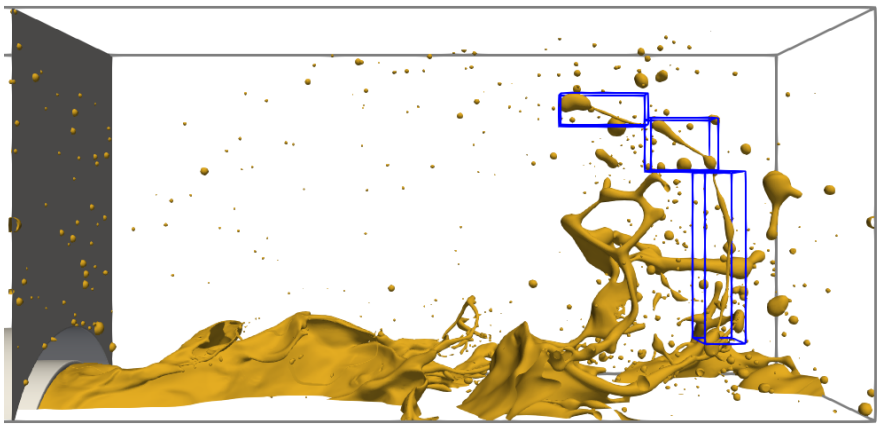


Figure 3: Enclosing boxes on detected liquid structures from ligament atomization, courtesy from [11]

4.1 Rendering

Rendering plays a crucial role in enabling a deeper understanding of the physical phenomena occurring in two-phase flow simulations. In this study, ParaView [10], a widely adopted and powerful visualization tool, is used to render and analyze simulation data. It enables visualizing simulation results applying advanced rendering techniques.

4.1.1 Baseline rendering

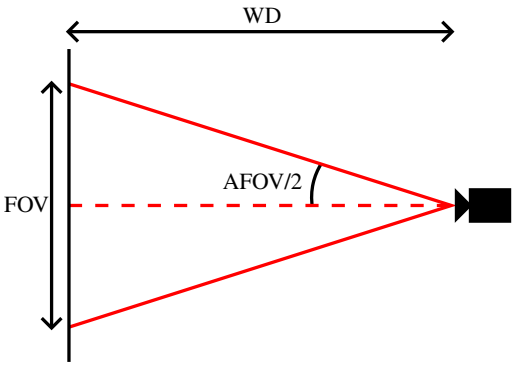
The baseline rendering technique is commonly employed in the visualization of two-phase flow phenomena, as demonstrated in previous works from [14], [3], [6], and [19]. This approach involves coloring the surface of the interface with an arbitrary color, without considering the refraction of light within the liquid structures. It primarily focuses on the reflection of light, resulting in the rendering of volumes with shadows, as illustrated in Figure 3. This rendering technique serves as a foundation for analyzing and understanding the behavior of two-phase flow systems.

In the context of this study, the objective is to generate images that capture the liquid structure with a high enough resolution to discern its intricate details, such as curvatures. To ensure precise visualization, the camera setup is carefully adjusted to center the liquid structure within the image. The angular field of view AFOV, the working distance WD and the number of pixels in both directions are given in Table 1.

WD (m)	AFOV (°)	nb_{pix}
7×10^{-3}	30	512

Table 1: Baseline rendering parameters

Then the spatial resolution can be deduced from the Field Of View (FOV) of the projection as illustrated in Figure 4 in order to measure quantitative data such as volumes. By dividing the FOV into the desired number of pixels, denoted as nb_{pix}, the spatial resolution of the resulting liquid structure projection can be determined as in equation 4.



$$\begin{cases} \text{FOV} & = 2 \times \text{WD} \times \tan\left(\frac{\text{AFOV}}{2}\right) \\ \text{Resolution} & = \text{FOV}/\text{nb}_{\text{pix}} \approx 7.16 \mu m/\text{px} \end{cases} \tag{4}$$

Figure 4: Illustration of camera setup quantities

Figure 5 showcases several liquid structures that have been isolated using the presented method. By considering liquid structures with varying volumes and fixing the WD, it is necessary for the apparent size to be proportional to their respective volumes. This choice enables the neural network to establish a correlation between the apparent size and the volume of the liquid structures.

Figure 5: Baseline rendering for several liquid structures

The baseline rendering technique described above will serve as the initial methodology test for the neural network in the subsequent sections.

4.1.2 Realistic rendering

Nevertheless, to accurately estimate the volume of liquid structures in experimental shadowgraphy images, it is crucial to ensure that the rendered simulation projections in the learning and testing databases closely resemble the experimental liquid structures. Previous works by authors such as [18], [17], [15], and [20] have provided shadowgraphy images of inert two-phase flow. Shadowgraphy images from [8] are used as a qualitative reference, as exemplified in Figure 1. The luminosity gradient along the edges of a liquid structure is a distinguishing feature that separates it from the background. It provides important visual information that a realistic rendering of a simulation should be able to capture and reproduce accurately.

Raytracing is a sophisticated rendering technique used in computer graphics to achieve highly realistic images. It simulates the behavior of light by tracing the paths of individual rays as they interact with objects in a scene. By considering optical properties like reflection and refraction, raytracing calculates the color and intensity of each pixel, resulting in visually stunning renderings with realistic lighting, shadows, and materials. Paraview uses the Ospray raytracing library to produce highly realistic rendered images. In this study, specific material parameters are defined for the liquid, creating a glass-like material (as defined by the Ospray library whose values are given in Table 2) that supports refraction and volumetric attenuation. The transparency color varies based on the geometric thickness, and the refraction coefficient is set to the water/air interface value. It is important to note that the attenuation coefficients used in the rendering are not based on physical parameters, but rather chosen arbitrarily to achieve the desired visual effect. Finally, the scene is designed to mimic a backlit experimental shadowgraphy setup, enhancing the realism of the rendered images.

Refraction coefficient	Attenuation distance (m)	Attenuation color
1.33	0.01	[0.01, 0.01, 0.01]

Table 2: Ray tracing, Paraview material parameters for water shadowgraphy rendering

Figure 6 presents the rendered images of the liquid structures that were selected from Figure 5. They exhibit a distinct luminosity gradient along the edges of the liquid structures and a brighter center, similar to what is observed in experimental images shown in Figure 1. These characteristics contribute to the visual realism and accuracy of the rendered images, making them more comparable to the experimental counterparts.

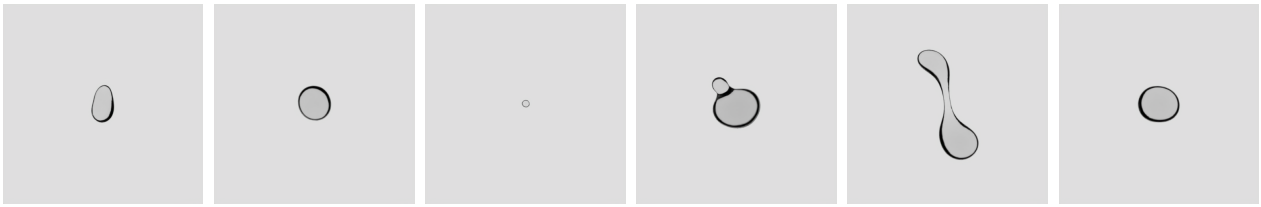


Figure 6: Ray tracing rendering with infinite depth of field

4.1.3 Training database specifications

The generation of projections is systematically applied to all liquid structures from the DNS field across multiple snapshots. To increase the number of images available for a given number of liquid structures, the process is performed in three directions: X, Y, and Z. Each resulting image is annotated with the corresponding volume V and saved in a separate file, serving as ground truth data. Two datasets are generated for the baseline and realistic renderings, one for training and a separate one for testing. The number of images for each dataset is detailed in Table 3.

	Direction	Training	Testing
Baseline	XYZ	425k	6k
Realistic	XYZ	278k	6k

Table 3: Number of simulation projections for the training and testing datasets

It is important to note that the simulation does not equally represent all volume classes, as the distribution of liquid structure volumes follows a log-normal pattern [11]. In order to ensure proper learning across all volume

classes, the dataset is sorted into an equal number of liquid structure projections, denoted as n_{proj}, for each class. The n_{proj} for each scenario will be presented and discussed in the subsequent sections.

5. Limitations with classical image analysis on the projected-volume correlation

Using both baseline and realistic rendering test datasets, a projected volume is estimated through projected surface estimation. Then a correlation map depicting the ratio between the volume estimated from the projected surface $S_{\text{proj}}^{3/2}$ and the exact volume of the liquid structures V is presented in Figure 7. It is important to note that all results are scaled by a characteristic size of the flow, namely the thickness of the coaxial injector lip denoted as δ_{lip}, as established in [11]. The maps reveal a significant linear correlation, with a coefficient 1.60. This value observed on a DNS dataset is not far from the 1.33 correlation coefficient for perfect spherical structures determined in section 2. However, it should be noted that the liquid structures exhibit some scatter around the $S_{\text{proj}}^{3/2}/1.6V$ correlation, especially for smaller diameters on the realistic rendering.

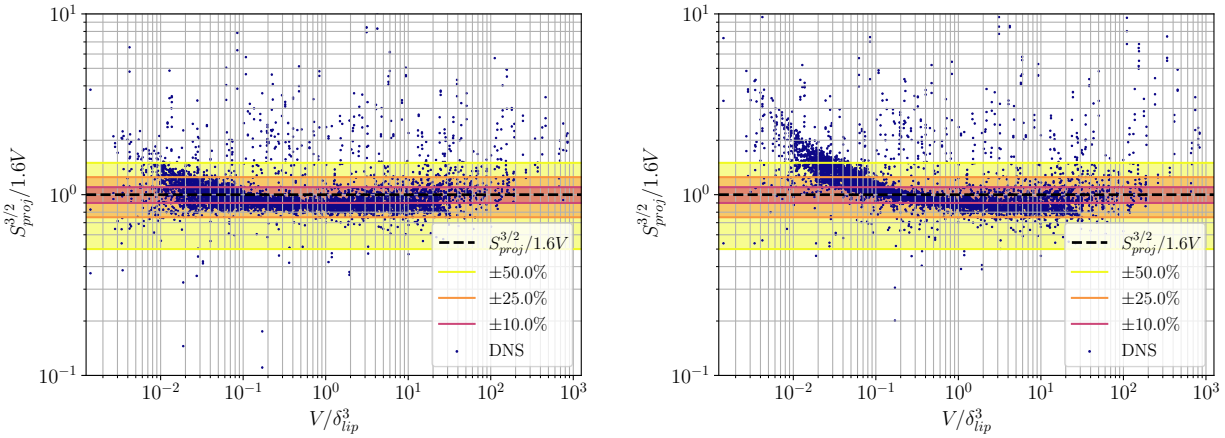


Figure 7: Correlation between the projected and exact volume from baseline (left) and realistic (right) rendering

Figure 8 provides a more detailed analysis of the scattering range observed in Figure 7, represented by the yellow and orange colored stripes. To better understand the correlation patterns, the volume spectrum is divided into three diameter classes: $d_{\text{proj}}/\delta_{\text{lip}} < 0.5$, $0.5 < d_{\text{proj}}/\delta_{\text{lip}} < 1.0$, and $1.0 < d_{\text{proj}}/\delta_{\text{lip}}$. On the left side of Figure 8, the cumulative probability is plotted, which represents the integration of the liquid structures number with increasing projected volume. The curves show a larger dispersion for the realistic rendering compared to the baseline rendering, indicating a volume-dependent distribution. Notably, the smallest liquid structures where $d_{\text{proj}}/\delta_{\text{lip}} < 0.5$, show a very different correlation compared to other volume classes as observed in Figure 7 with values very different from 1 for the $S_{proj}^{3/2}/1.6V$ ratio. On the right side of Figure 8, the tolerance according to the 1.6 correlation is plotted. This reveals a wider spread of results for the realistic rendering (only 70% of the overall population is found within a 25% tolerance for the realistic rendering, against 85% for the baseline rendering), particularly for the smallest liquid structures. This suggests that the accuracy of volume estimation through classical image analysis is reduced in these cases. It highlights the importance of alternative approaches such as deep learning techniques and the potential for improvements when dealing with smaller liquid structures in realistic rendering scenarios.

6. Neural network results

In this section, we present the results achieved through the methodology outlined in the preceding sections, with a specific emphasis on the learning convergence of the neural network and the precision of the obtained outcomes for both the baseline and realistic rendering scenarios.

6.1 Classification and learning parameters

To investigate the learning convergence of the neural network, experiments were conducted by varying the number of volume classes and projections per class during the training process. The objective is to determine the impact of these parameters on the performance of the network in estimating the volume of liquid structures. In this study, a

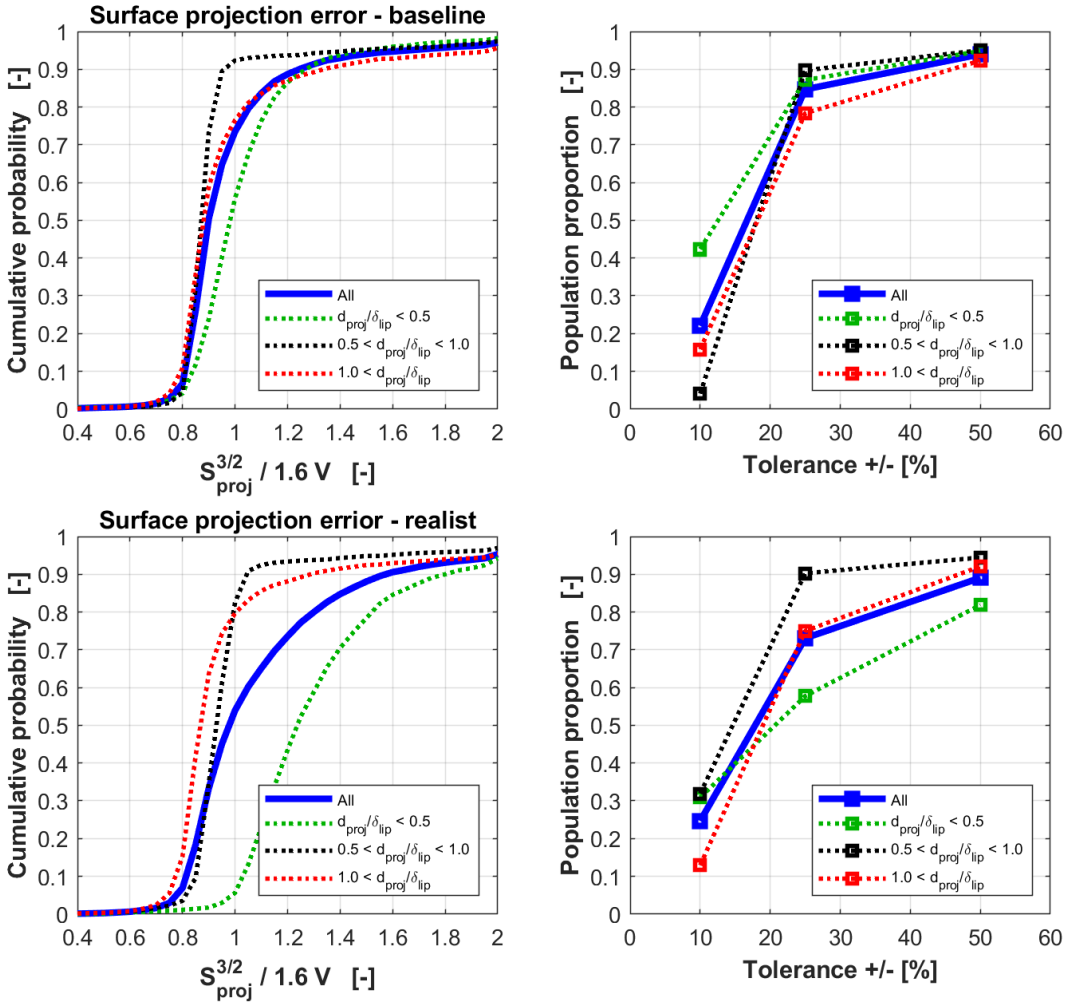


Figure 8: Liquid structures margins error count from baseline (top) and realistic rendering (bottom) for baseline

range from 4 to 10 volume classes is considered with an adjusted number of projections per class from 700 to 1500. Also, an additional classification in 5 classes tagged as specialized (sp) is detailed and will be used in the last section with narrower classes. It is worth noting that the configuration of the liquid structures exhibits a log-normal size distribution, as observed in previous studies [11]. This distribution implies that smaller liquid structures are more abundant compared to larger ones. To ensure a sufficient number of projections available per class and to maintain a balanced representation of the different volume classes, the distribution of classes was adapted according to the PDF from [11] as detailed in Table 4. By systematically analyzing the learning convergence of the neural network under various training configurations, the aim is to identify the optimal settings that would result in accurate and reliable volume estimations. This investigation provides valuable insights into the interplay between the number of volume classes, projections per class, and the learning convergence behavior of the network, allowing the optimization of the training process for this specific application.

	Class range ($d_{30}/\delta_{\text{lip}}$)									
4 cl.	[0-0.35]	[0.35-0.5]	[0.5-0.75]	[0.75-2.5]						
6 cl.	[0-0.35]	[0.35-0.45]	[0.45-0.6]	[0.6-0.75]	[0.75-1.05]	[1.05-2.5]				
8 cl.	[0-0.35]	[0.35-0.4]	[0.4-0.5]	[0.5-0.6]	[0.6-0.75]	[0.75-1.0]	[1.0-1.25]	[1.25-2.5]		
10 cl.	[0-0.35]	[0.35-0.4]	[0.4-0.5]	[0.5-0.6]	[0.6-0.7]	[0.7-0.85]	[0.85-1.05]	[1.05-1.25]	[1.25-1.5]	[1.5-2.5]
sp 5 cl.	[0-0.1]	[0.1-0.2]	[0.2-0.3]	[0.3-0.4]	[0.4-0.5]					

Table 4: Range of $d_{30}/\delta_{\text{lip}}$ sizes for classification into 4, 5, 6, 8 and 10 classes

6.1.1 Training convergence

The convergence of the learning process for the neural network is further examined using the baseline database. First, the number of classes is fixed to 6, with a corresponding range of volume classes as outlined in Table 4. The study focuses on evaluating the network convergence by varying the number of projections per class, ranging from 700 to 1500. Multiple training iterations are conducted for each n_{proj} value, resulting in varying convergence behavior due to the random selection of projections during the learning process. To provide a comprehensive analysis, both the actual convergence data and a fitted polynomial curve are plotted in Figure 9. The results demonstrate that, regardless of the number of projections per class, the neural network loss converges after approximately 5-6 epochs. However, when examining the convergence after 2 epochs, it is observed that a higher number of projections per class results in faster convergence. Overall, these findings highlight the importance of carefully selecting the number of projections per class to optimize the convergence of the neural network.

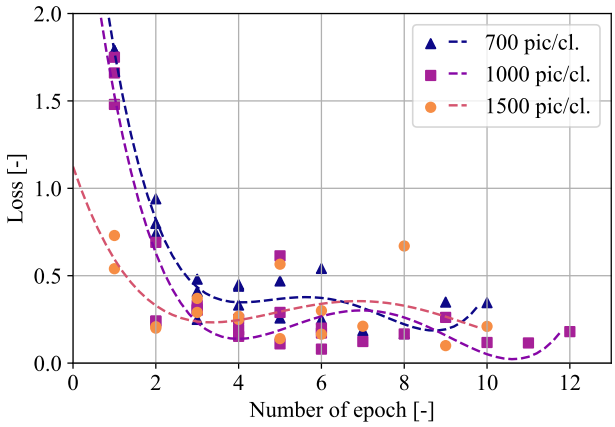


Figure 9: Loss convergence for 6-class training

Subsequently, the number of projections per class, denoted as n_{proj}, is fixed at 1000, while the number of classes is varied from 4 to 10, following the volume class ranges specified in Table 4. The results plotted in Figure 10 reveal that the networks with 4 and 6 classes exhibit similar loss convergence patterns, reaching a stable state after approximately 4 epochs. This suggests that a moderate number of classes can achieve satisfactory convergence with 1000 projections. However, for the network with 8 and 10 classes, a slightly slower convergence is observed, indicating the need for a higher number of projections to ensure optimal convergence of the training.

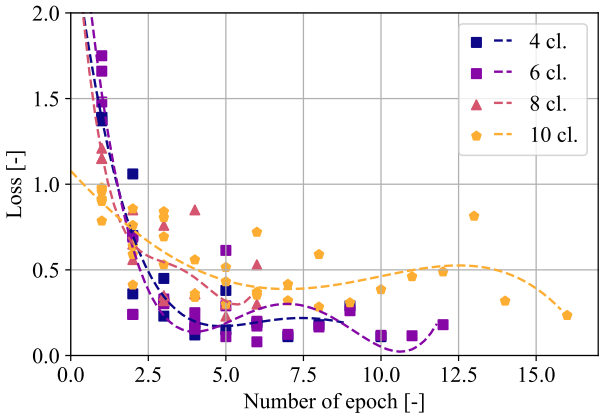


Figure 10: Loss convergence for 1000 projections by class

In order to compare the convergence behavior of the neural network using the baseline and realistic rendering, experiments were conducted with a dataset consisting of 1000 images per class and 10 volume classes. The inclusion of a significant number of classes is necessary to achieve reasonable accuracy in volume prediction. The results of this comparison are presented in Figure 11. It can be observed that the network convergence is slower when using more complex realistic images compared to the baseline rendering. The learning process also seems more dispersed

for realistic conditions from one epoch to the following and from one run to another. This can be attributed to the additional visual complexity and intricacies present in the realistic rendering, which may require a more complete training process for the network to learn and accurately estimate the volume of the liquid structures (larger number of projections for each class, training over a larger number of epochs...).

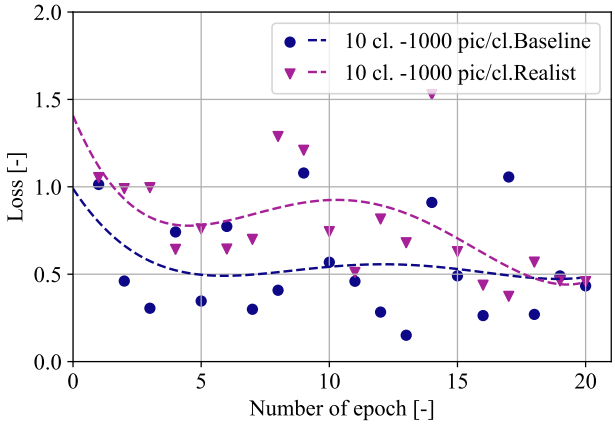


Figure 11: Loss convergence comparison between baseline and realistic dataset

This finding highlights the importance of considering the rendering technique and the level of realism in the training data when designing and training neural networks for volume estimation tasks. While the realistic rendering provides a more visually accurate representation of the liquid structures, it requires a higher number of training epochs to achieve convergence comparable to the baseline rendering.

6.1.2 Performance for volume classification

The performance of the trained neural network for estimating the volume of liquid structures is evaluated in this section. To assess the classification accuracy (i.e. classification of projections in the proper volume class), the trained neural network is applied to the test dataset detailed in Table 3. The evolution of accuracy is analyzed by varying the number of training images per class for a range of volume classes, from 4 to 10. Figure 12 illustrates the relationship between accuracy and the number of training images per class. The results demonstrate a consistent trend where increasing the number of training images per class leads to higher accuracy across all volume classes. This observation holds true for all evaluated scenarios, as indicated by the convergence-like behavior of the accuracy curve. It suggests that the neural network benefits from a larger number of training images, allowing it to capture the underlying patterns and improve its estimation precision. In this context, the use of simulated data for training is very helpful since it is able to generate large datasets with much less practical limitations compared to manual annotations.

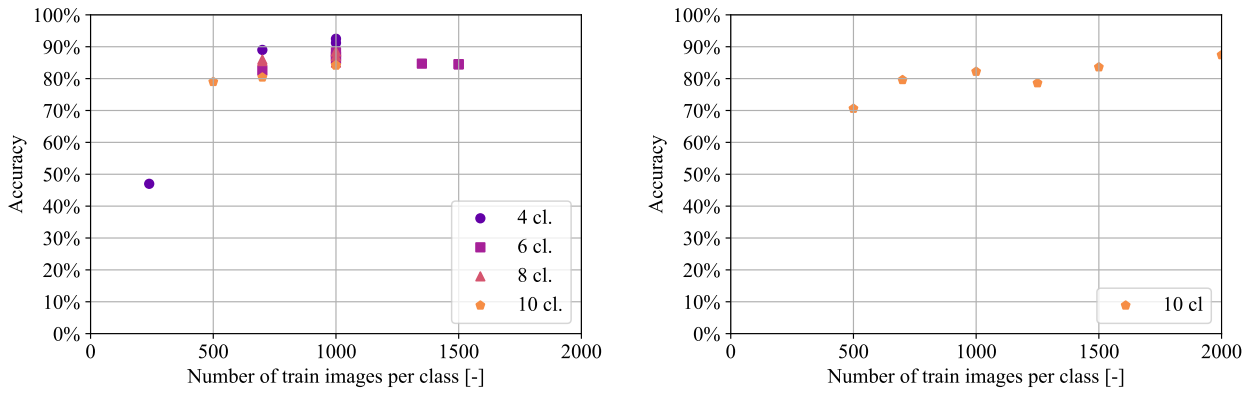


Figure 12: Model accuracy as a function of the number of pictures for each class in the training for the baseline (left) and realistic rendering (right)

The volume error made by the neural network is plotted in Figure 13 to assess the impact of classification errors on physical parameters in the spray. The total volume estimation is obtained by summing the volume estimates for

each projection classified in the testing stage. Typically, there are 50 projections per class, resulting in a total of 200 to 500 projections, depending on the number of classes. The total volume estimated by the neural network classification is compared to the total volume obtained from the ground truth (GT). The maximum error observed is up to 15% for the baseline rendering and 5% for the realistic rendering. These results demonstrate the overall good performance of the network in predicting the overall volume of liquid structures.

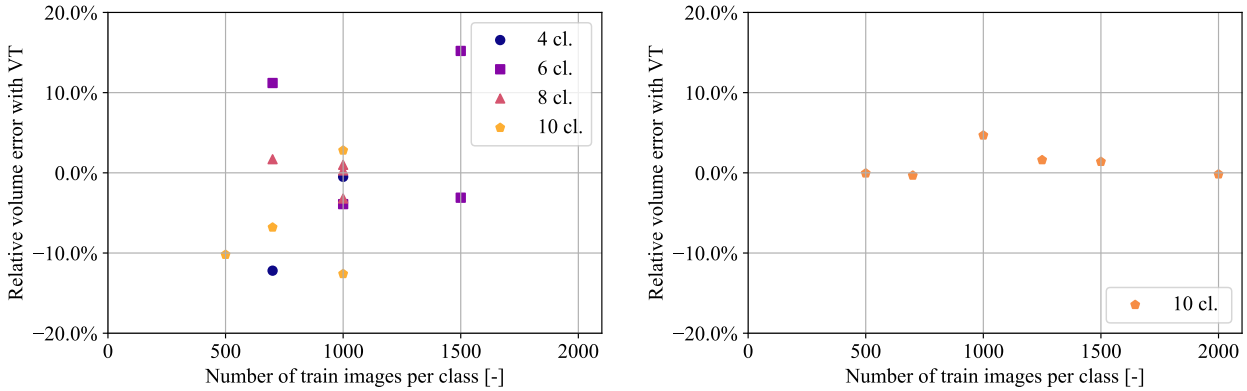


Figure 13: Relative error on the total estimated volume as a function of the number of pictures for each class in training

A more detailed analysis of the neural network accuracy on volume estimation can be achieved by examining the confusion matrices, i.e. comparing the classification for GT and volume estimation via a matrix representation: x-axis for GT volume and y-axis for estimation (correlation or CNN). The dotted line along the diagonal corresponds to an accurate classification (for instance, GT projections from class 4 are estimated in class 4). Figure 14 displays the confusion matrices for the classical image analysis projected volume correlation $S_{\text{proj}}^{3/2}/1.6$, and the classifications made by the neural network for both the baseline and realistic rendering scenarios. The results reveal that the projected volumes from baseline and realistic rendering exhibit similar performance for larger volume classes. However, when it comes to smaller liquid structures, the projections from the baseline rendering outperform those from the realistic rendering, as observed in Figure 7: the confusion matrix shows more dispersion around the diagonal for the small volume range (classes 1 to 4) in the case of the realistic rendering. On the other hand, the neural network classifications show reduced scattering compared to the projected volumes, particularly for large volumes. This highlights the effectiveness of the neural network approach in estimating liquid structure volumes.

In order to estimate global statistics on the flow, various PDFs are constructed and displayed in Figure 15, using the ground truth data (GT), the classified ground truth (GT n cl) for the same number of classes used for CNN, the projected volume correlation $S_{\text{proj}}^{3/2}/1.6$ with its classification, and the trained neural network outputs for both the baseline and realistic rendering scenarios. The results show that the trained neural network enables a precise reconstruction, despite large- range classes, of the PDFs and yields results that are close to those estimated using the projected volume correlation $S_{\text{proj}}^{3/2}/1.6$. This demonstrates the capability of the neural network to accurately estimate the volume of liquid structures. However, it is important to note that achieving precise estimations of volume requires a larger number of volume classes which comes with a higher computational cost that increases exponentially. To mitigate this computational burden, one possible approach is to train the neural network for a specific range of sizes, focusing on a narrower size distribution of liquid structures (for instance small structures only). This can help reduce the training time and computational resources required while still achieving accurate volume estimations within the targeted size range.

To address the challenge of estimating liquid structure sizes within a narrow range, a specialized neural network example was developed. This neural network is specifically designed to handle liquid structures with sizes ranging from 0 to 0.5 times the characteristic size δ_{lip}. The size range is discretized into five classes, as defined in Table 4. The training dataset consists of 500 images per class. By focusing on this fine discretization of the size scale, the neural network is able to provide a better discretization of small volume, as demonstrated by the probability density function plotted in Figure 16.

This approach highlights the possibility of training a neural network for precise volume estimation using a relatively low number of training images, provided that the network is specialized for a narrow spectrum of sizes. To cover the entire range of liquid structure sizes, multiple neural networks can be trained, each targeting a specific range of sizes.

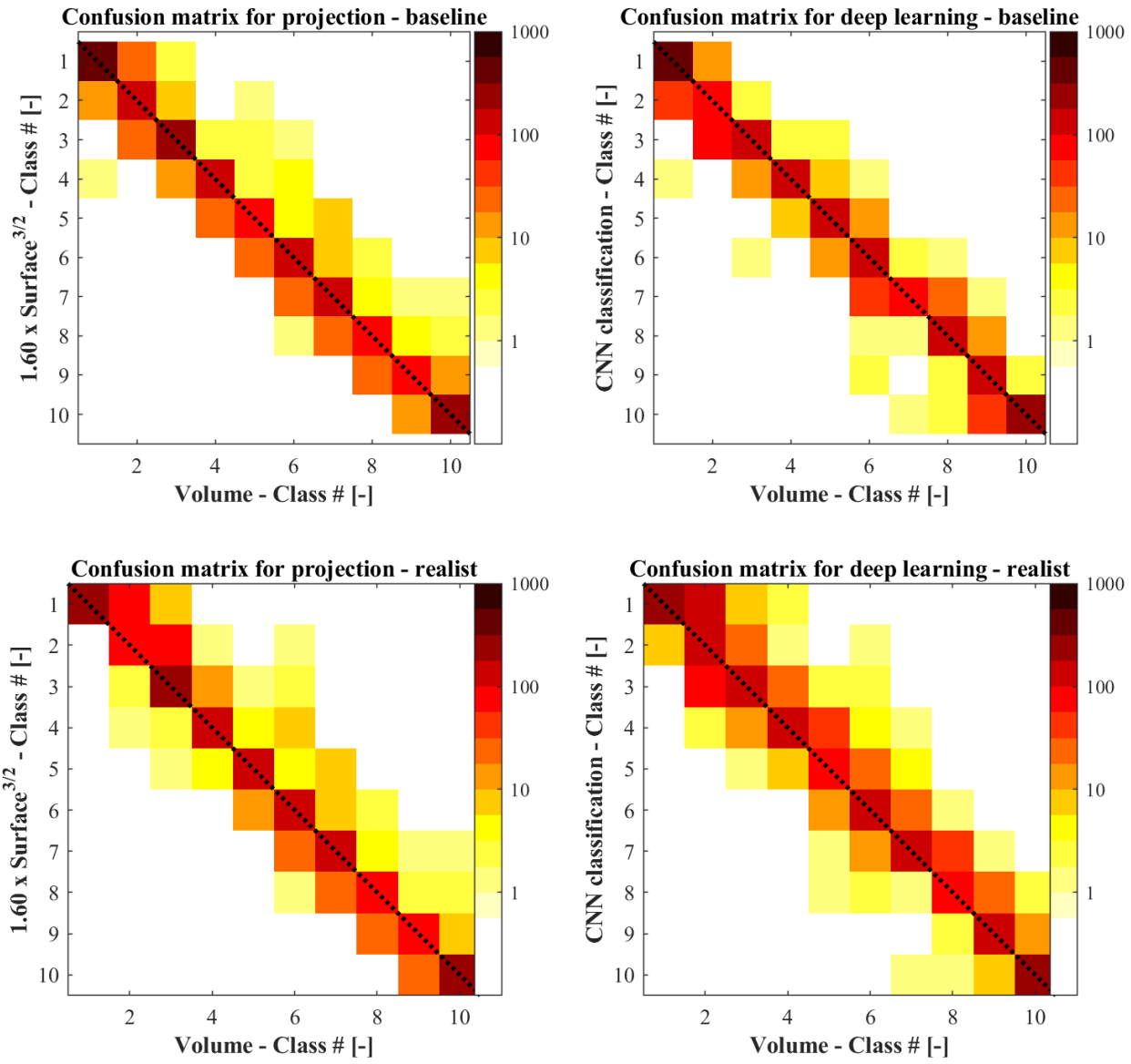


Figure 14: Confusion matrix for the $S_{\text{proj}}^{3/2}$ (left) and neural network (right) estimation on the baseline (top) and realistic rendering (bottom)

7. Conclusion

In conclusion, this paper designs a methodology to estimate the volume of a liquid structure in sprays from a single projection using deep learning. A neural network proved to be a viable solution for this task, leveraging its ability to learn from prior information and compensate for the lack of depth information in the images. To generate the training and test dataset, high-fidelity simulation results were employed. The results provided a significant advantage in terms of dataset size and annotation, as they offered a large number of accurately annotated liquid structures, which is usually a bottleneck. The results obtained with the baseline rendering were excellent, demonstrating fast convergence of the neural network with a limited number of training epochs. The training process could be completed within a few hours, highlighting the efficiency of the approach and its applicability in an industrial context. While the results on realistic rendering were also promising, achieving comparable performance to baseline rendering requires a higher number of training epochs and training images per class. This limitation could be addressed by increasing the dataset per volume class with additional images or training specialized neural networks for narrower size bands.

Also, the training configurations for the neural network were carried out on a single Nvidia RTX 2080 Ti GPU and completed within a few hours. This relatively short training time and the use of a single GPU highlight the

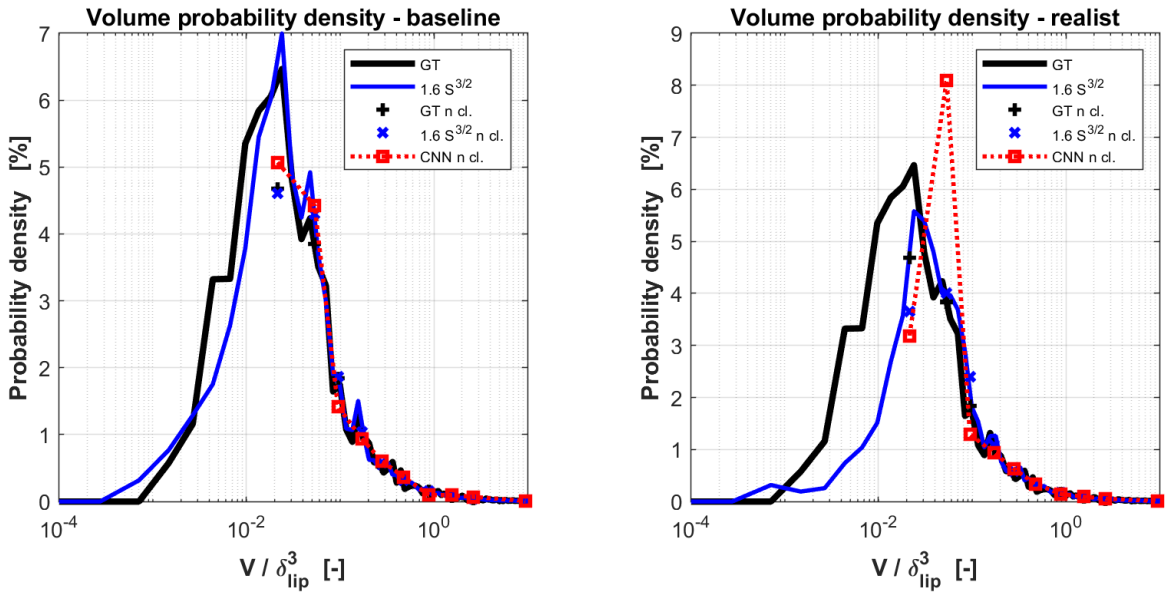


Figure 15: Probability density function on the testing dataset with ground truth, classed ground truth, projected volume correlation $S_{proj^{3/2}}/1.6V$ and prediction of the neural network for both baseline and realistic rendering

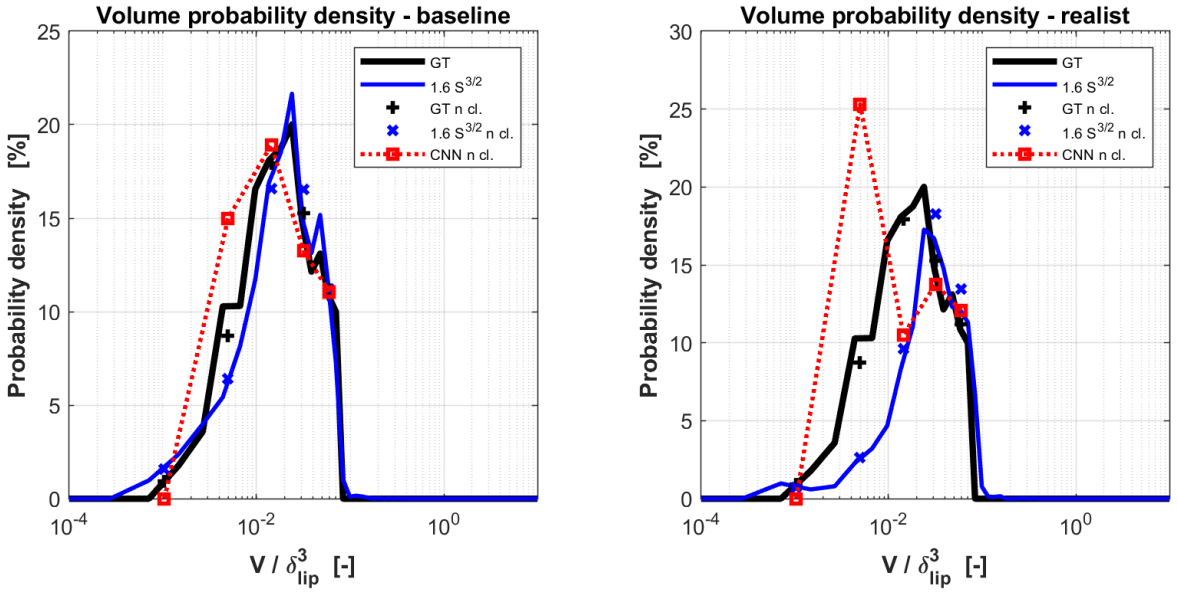


Figure 16: Probability density function of the specialized network for $0 < d_{30}/\delta_{\text{lip}}$

relatively low computational cost associated with the training process. It also suggests that there is potential for further improvements and optimizations in terms of training efficiency and performance. With advancements in hardware and training techniques, it is possible to achieve even faster training times and potentially enhance the accuracy and capabilities of the neural network model.

Looking ahead, future work should consider the inclusion of a limited depth of field during the generation of realistic renderings, which introduces blur in the images and provide more realistic results comparable to experimental conditions. This can be achieved using the same tools and techniques as those presented in this study. More comprehensive metrics such as precision or recall could be used to take account of these dissimilarities. The developed methodology can finally be applied to real experimental images, bringing a new way of quantifying liquid structures of any shape and improving the analysis of experimental data. This would contribute to a better comparison between numerical results and experimental data, which is crucial for model validations and advancing our understanding of the underlying phenomena.

8. Acknowledgments

This work was made possible through the support of a PhD grant from CNES (Centre National d'Etudes Spatiales) and ONERA (Office National d'Etudes et de Recherches Aéronautiques). The authors would also like to express their gratitude to the CORIA laboratory for providing the experimental shadowgraphy images from the MARACA test bench.

References

[1] M. Averseng. *Contribution à la modélisation de l'atomisation assistée par analyse de simulations haute fidélité*. PhD thesis, Toulouse University, 2022.

[2] S. T. Barnard and M. A. Fischler. Computational Stereo. *ACM Computing Surveys (CSUR)*, 14(4):553–572, 1982.

[3] P. A. Beau, T. Ménard, R. Lebas, A. Berlemont, S. Tanguy, and F. X. Demoulin. Numerical jet atomization. Part II: modeling information and comparison with DNS results. In *7th International Symposium on numerical methods for multiphase flows*, 2006.

[4] N. Chigier and Z. Farago. Morphological Classification of Disintegration of Round Liquid Jets in a Coaxial Air Stream. *Atomization and Sprays*, 2(2):137–153, 1992.

[5] Dan C. Cireşan, Ueli Meier, Jonathan Masci, Luca M. Gambardella, and Jürgen Schmidhuber. Flexible, high performance convolutional neural networks for image classification. In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence - Volume Volume Two*, IJCAI'11, page 1237â1242. AAAI Press, 2011.

[6] B. Duret, J. Reveillon, T. Menard, and F. X. Demoulin. Improving primary atomization modeling through DNS of two-phase flows. *International Journal of Multiphase Flow*, 55:130–137, 2013.

[7] N. Fdida, L. Vingert, Y. Mauriot, L. H. Dorey, and M. Théron. Comparison of LOX / Methane and LOX / Hydrogen cryogenic spray combustion with simultaneous optical diagnostics. *Eucass 2019*, pages 2–11, 2019.

[8] A. Ficuciello, J. B. Blaisot, C. Richard, and F. Baillot. Investigation of air-assisted sprays submitted to high frequency transverse acoustic fields: Droplet clustering. *Physics of Fluids*, 29(6):1–17, 2017.

[9] H. Fu, M. Gong, C. Wang, K. Batmanghelich, D. Tao, Ligm Umr, and Esiee Paris. Deep ordinal regression network for monocular depth estimation. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2002–2011, 2018.

[10] A. Henderson. *The ParaView guide : a parallel visualization application*. Kitware, Inc, 2007.

[11] J. C. Hoarau, L. H. Dorey, D. Zuzio, F. Granger, and J. L. Estivalèzes. Direct numerical simulation of a subcritical coaxial injection in fiber regime using sharp interface reconstruction. *International Journal of Multiphase Flow*, (to be submitted).

[12] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. Red Hook, NY, USA, 2012. Curran Associates Inc.

[13] A. Kumar and S. Sahu. Liquid jet breakup unsteadiness in a coaxial air-blast atomizer. *International Journal of Spray and Combustion Dynamics*, 10(3):211–230, 2018.

[14] C. Le Touze, L. H. Dorey, N. Rutard, and A. Murrone. A compressible two-phase flow framework for Large Eddy Simulations of liquid-propellant rocket engines. *Applied Mathematical Modelling*, 84:265–286, 2020.

[15] H. Malot and J. B. Blaisot. Droplet size distribution and sphericity measurements of low-density sprays through image analysis. *Particle and Particle Systems Characterization*, 17(4):146–158, 2000.

[16] P. Marmottant. *Atomisation d'un jet liquide par un courant gazeux*. PhD thesis, Polytechnique Institute of Grenoble, 2001.

[17] P. Marmottant and E. Villermaux. Fragmentation of stretched liquid ligaments. *Physics of Fluids*, 16(8):2732–2741, 2004.

[18] P. Marmottant and E. Villermaux. On spray formation. *Journal of Fluid Mechanics*, 498(498):73–111, 2004.

[19] H. Pitsch, O. Desjardins, G. Balarac, and M. Ihme. Large-eddy simulation of turbulent reacting flows. *Progress in Aerospace Sciences*, 44(6):466–478, 2008.

[20] K. A. Sallam, C. Aalburg, and G. M. Faeth. Breakup of round nonturbulent liquid jets in gaseous crossflow. *AIAA Journal*, 42(12):2529–2540, 2004.