Neural networks for classification and segmentation of thermally-induced droplet breakup in spray-flame synthesis

Jüngst, Niklas; Ersoy, Veysel; Smallwood, Gregory J.; Kaiser, Sebastian A.

Publisher's version / Version de l'éditeur:

https://doi.org/10.1016/j.jaerosci.2023.106314

Journal of Aerosol Science, 176, C, pp. 1-15, 2023-12-10

Journal of Aerosol Science 176 (2024) 106314

Contents lists available at ScienceDirect

Journal of Aerosol Science

journal homepage: www.elsevier.com/locate/jaerosci

Neural networks for classification and segmentation of thermally-induced droplet breakup in spray-flame synthesis

Niklas Jüngst [a,c,*], Veysel Ersoy [a,c], Gregory J. Smallwood [b], Sebastian A. Kaiser [a,c]

[a] Institute for Energy and Materials Processes – Reactive Fluids University of Duisburg-Essen Duisburg, Germany
[b] Metrology Research Centre National Research Council Canada, Ottawa, ON, Canada
[c] CENIDE, Center for Nanointegration Duisburg-Essen University of Duisburg-Essen Duisburg, Germany

ARTICLE INFO

Keywords:
Image analysis
Mask R–CNN
Faster R–CNN
Spray-flame synthesis
Droplet puffing and micro-explosion
High-speed microscopy

ABSTRACT

Spray-flame synthesis (SFS) enables one-step production of functional metal-oxide nanoparticles from inexpensive precursors such as metal salts. Precursor-laden droplets show thermally-induced breakup, i.e., puffing and micro-explosion, considered a key step in SFS. In this work, shadowgraphy images of droplets are investigated with neural networks with the aim of extracting quantitative data on this breakup process. Faster and Mask region-based convolutional neural networks (R-CNNs) were applied to segment in-focus droplet shadowgraphs (Mask R–CNN) and to distinguish regular from disrupting droplets and detect sequences of consecutive droplet shadowgraphs (Faster R–CNN). The networks were trained with about 400 manually annotated images containing a total of about 1200 objects. Precision and recall of the trained networks reach 80% and 90%, respectively. The results from the two networks were combined and compared with those from conventional particle/droplet image analysis (PDIA). Mask R–CNN and PDIA identify similar droplets as being in-focus and segment similar regions of the droplets. Puffing or micro-explosion sometimes leave behind secondary droplets that are too small or defocused for reliable segmentation. Such shadowgraphs are classified erroneously by PDIA. In contrast to this, Faster R–CNN is more sensitive as it does not rely on the segmentation and thus classifies such low-contrast shadowgraphs correctly. Therefore, the CNNs find on average 30% higher droplet-disruption probabilities than PDIA.

1. Introduction

1.1. Background

Spray-flame synthesis (SFS) produces various functional metal-oxide nanomaterials (Mädler et al., 2002). Solid or liquid precursors are dissolved in a combustible liquid and atomized into a pilot flame where the solution ignites into a spray flame. Such precursor-laden droplets are known to undergo thermally-induced secondary atomization, i.e., droplet disruption by “puffing” and “micro-explosion” (Jüngst et al., 2022; Rosebrock et al., 2013; Stodt et al., 2022; Witte & Mädler, 2023). In puffing, the droplet locally ruptures and ejects secondary droplets and vapor, while in a micro-explosion the droplet entirely disintegrates into multiple fragments (Shinjo et al., 2014). Both types of events increase the droplet’s specific surface area and thus help bringing the precursor into the gas

* Corresponding author. Institute for Energy and Materials Processes – Reactive Fluids University of Duisburg-Essen Duisburg, Germany.
E-mail address: niklas.juengst@uni-due.de (N. Jüngst).

https://doi.org/10.1016/j.jaerosci.2023.106314
Received 19 September 2023; Received in revised form 29 November 2023; Accepted 4 December 2023
Available online 10 December 2023

phase from where it eventually solidifies into nanoparticles. Therefore, puffing and micro-explosion are key steps in SFS and need to be included in process models and simulations. However, little data is available on probabilities of these events or the sizes, velocities, and concentrations of the participating droplets. Acquisition of such data requires high-speed high-magnification imaging of the droplets and subsequent image analysis.

Conventional particle/droplet image analysis (PDIA) uses intensity or intensity-gradient thresholding (or a combination of both) for the identification and segmentation of in-focus objects (Jiang et al., 2010; Kashdan et al., 2007; Kim & Kim, 1994; Koh et al., 2001; Kumar & Sahu, 2020; Lecuona et al., 2000; Lee & Kim, 2004; Malot & Blaisot, 2000). Particularly, if the evolution of a droplet's morphology, e.g., its disintegration, is of interest, complex programming is required to track a droplet and its secondary droplets through the spatial domain. Also, varying background intensity, overlapping objects, complex object shapes, objects of different classes, and varying experimental boundary conditions may require specific rules and exceptions that are time-consuming to program (Li et al., 2021).

Convolutional neural networks (CNNs) are a promising approach to overcome these difficulties. CNNs represent a class of deep learning models especially suited for image analysis tasks due to their ability to capture local features and build hierarchical feature representations (Lecun et al., 1998). They implicitly learn rules and conditions from a set of training data and operate via repeated convolution, activation, and sub-sampling procedures. Fig. 1 shows shadowgraphy and mask images from droplets in SFS and summarizes typical image-analysis tasks that CNNs may perform. CNNs are either applied for the classification of an entire image, the localization and classification of objects (i.e., object detection), semantic segmentation, or instance segmentation. These are discussed in more detail below.

Region-based convolutional neural networks (R-CNNs) were originally developed to perform object detection based on region proposals in images (Girshick et al., 2014). Object detection refers to predicting a bounding box around each object and classifying each object. Several improvements of R–CNN, such as the open-source Fast R–CNN and Faster R–CNN, have yielded higher detection quality and less computation time than R–CNN (Girshick, 2015; Ren et al., 2017). Today, R-CNNs are also applied for segmentation tasks. Semantic segmentation associates each pixel with a certain class, e.g., droplet or background. Instance segmentation detects all individual objects of a certain class, e.g., droplet, and demarcates them from each other. A state of the art technique for instance segmentation is Mask R–CNN (He et al., 2017).

In multiphase-flow applications, CNNs typically classify images showing bubbles (Fu & Liu, 2019; Haas et al., 2020; Poletaev et al., 2016; Poletaev et al., 2020), droplets (Hadikhani et al., 2019; Narayanan et al., 2021; Rutkowski et al., 2022; Yee et al., 2022), particles (Li et al., 2021), or flame luminosity (Li et al., 2023; Pan et al., 2021) In droplet-based microfluidic devices, CNNs were found to be suitable for real-time monitoring with image rates of more than 200 frames per second (fps) (Rutkowski et al., 2022). Single coal particles were ignited in premixed lean methane flames and the laser-induced fluorescence of OH was detected at high frame rates. Two CNNs classified images showing ignited particles to determine the ignition delay time (Li et al., 2023). To assess flame instabilities during SFS, images of visible chemiluminescence from the spray flame were acquired. Faster R–CNN was used to detect stable and unstable spray flames with the aim of monitoring the flame stability and eventually real-time controlling of SFS parameters (Pan et al., 2021; Paulson et al., 2020).

The U-Net architecture originally evolved from the semantic segmentation of biomedical images (Ronneberger et al., 2015). Today it is also used in various multiphase systems (Chaussonnet et al., 2019; Devillers et al., 2019; Li et al., 2021; Liang, et al., 2019; Schäfer et al., 2019; Wang et al., 2022a, 2022b). A modified U-Net model segmented bubbles (Li et al., 2021) and droplets (Schäfer et al., 2019) in shadowgraphy images as well as solid particles (Liang et al., 2019) in scattered-light images. A generative adversarial network (GAN) identified and segmented in-focus droplet shadowgraphs in spray images. In that work, synthetic training images containing focused and defocused droplets were generated, the latter applying Gaussian blur (Wang et al., 2022a, 2022b). Also in spray images, U-Net was combined with you only look once (YOLO) to segment droplets and classify characteristic liquid structures in primary atomization, such as attached and detached ligaments, lobes, bags, and rims (Chaussonnet et al., 2019). Complementing this, an artificial neural network (ANN) was trained to also extract the droplet Sauter mean diameter and velocity as well as the ligament length and deformation velocity (Chaussonnet et al., 2021).

The segmentation of individual objects of a certain class, i.e., instances, requires more complex networks than semantic

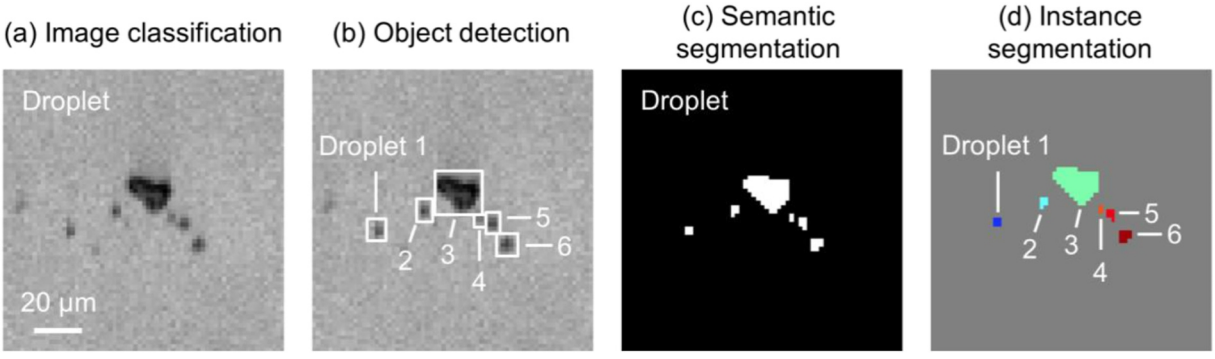

Fig. 1. Image analysis with R-CNNs: (a) image classification, (b) object detection, (c) semantic segmentation, (d) instance segmentation.

segmentation (Frei & Kruis, 2018; Frei & Kruis, 2020; Frei & Kruis, 2022; Kim & Park, 2021; Pathan et al., 2022; Petrucci et al., 2022; Sipkens et al., 2021; Wu et al., 2020; Zhang et al., 2019). In transmission and scanning electron microscope (SEM/TEM) images, Mask R–CNN segmented nanometer-sized particles (Zhang et al., 2019), sometimes sintered in aggregates (Frei & Kruis, 2018; Frei & Kruis, 2020; Frei & Kruis, 2022; Sipkens et al., 2021). The performance of U-Net and Mask R–CNN was compared for visualizations of a disintegrating liquid jet with very high object density, generated with direct numerical simulations (DNS). It was found that U-Net outperformed Mask R–CNN after few training epochs. However, after longer training, Mask R–CNN reached higher precision in detecting small objects and an otherwise similar performance as U-Net (Pathan et al., 2022).

1.2. Scope of this work

In our previous work, we developed high-speed high-magnification shadowgraphy to image the thermally-induced droplet breakup in SFS (Jüngst et al., 2022). Explicitly programmed PDIA was developed to extract droplet sizes, velocities, concentrations, and breakup probabilities from the images (Jüngst et al., 2023). However, advanced burner versions with smaller and faster droplets decrease the contrast in such shadowgraphy images and pose challenges for the image analysis. This was further incentive to explore the use of data-driven image analysis. Particularly, small fragments generated during droplet breakup tend to be found less frequently by PDIA. In this study, we trained object detection and segmentation CNNs on manually annotated images from a former burner version and then applied the networks to images from a more advanced burner. This way we want to establish a robust tool for assessing thermally-induced droplet breakup under versatile conditions. Mask R–CNN predicts a segmentation mask for each in-focus droplet shadowgraph, while Faster R–CNN localizes and classifies droplet shadowgraphs. Faster R–CNN also localizes sequences of temporally consecutive droplet shadowgraphs. In contrast to PDIA, classification and segmentation are independent from each other in this CNN framework. The results from both networks are combined, yielding droplet sizes and classes, and eventually compared to those from PDIA.

2. Methods

2.1. Spray-flame synthesis and image acquisition

The SFS burner in this work, the so called SpraySyn burner, is the result of the collaborative work from 27 research groups, located at multiple German universities. Throughout the last six years, this burner has gone through many different versions. The original SpraySyn1 (Schneider et al., 2019) has global spray and flame fluctuations, with potential carbon contamination of the product (Bieber et al., 2021). If the high-velocity flow of the dispersion gas is made to exit inwardly towards the central liquid precursor flow, as schematically shown in Fig. 2a, the fluctuations are much reduced and the spray atomization is finer (Bieber et al., 2021). This also yields smaller particles with a narrower size distribution. These insights resulted in the creation of SpraySyn2 which was used in this work. However, the smaller and faster droplets also make image acquisition and analysis more challenging – which is why the data from SpraySyn2 were used as test cases for the trained networks here.

Iron (III) nitrate nonahydrate (INN) was used as a precursor for iron-oxide nanoparticles. INN was dissolved at various concentrations (specific values are indicated at the end of this Section) in a mixture of 35 vol% ethanol (EtOH) and 65 vol% 2-ethylhexanoate acid (EHA). The liquid was supplied at a flow rate of 2 ml/min (exit velocity 0.26 m/s). An oxygen flow of 6 slm (exit velocity >140 m/

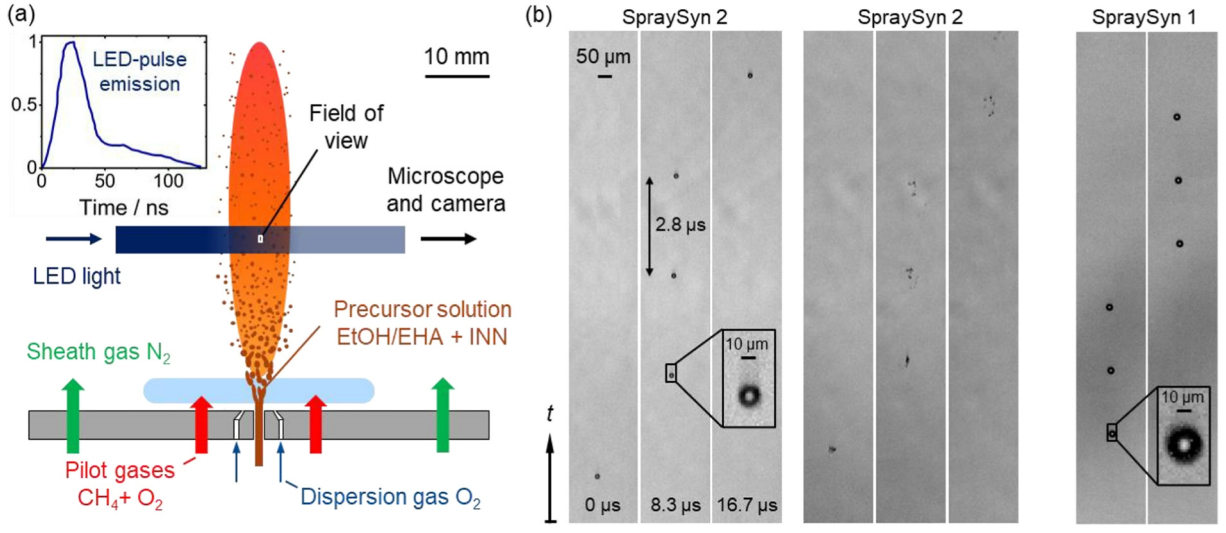

Fig. 2. (a) Sketch of SpraySyn2 burner, optics, probe volume, and time-resolved LED emission. (b) Shadowgraphy images of regular and disrupting droplets from SpraySyn1 and 2.

s) atomized the liquid into small and fast droplets which ignite as they pass through a pilot flame. The pilot flame burned a mixture of 16 slm oxygen and 2 slm methane (exit velocity 2.3 m/s). A nitrogen co-flow of 120 slm (exit velocity 0.6 m/s) shielded the near-nozzle region from the environment.

To visualize micrometer-sized regular and disrupting droplets, a high-speed high-magnification shadowgraphy technique was developed, as summarized below. Details on the method can be found in (Jüngst et al., 2022, 2023). A blue LED (Cree XP-E2 SMD-LED, 550 mW, 450 nm) was pulsed by a commercial driver (Picolas LDP-V 50–100 V3.3) to illuminate the spray at selected heights above the burner (HAB). Fig. 2a shows the time-resolved emission of the LED pulse, which is about 120 ns in duration. Such short light pulses were necessary to avoid motion blurring of the shadowgraphs. A microscope objective (Mitutoyo M Plan Apo 10×) imaged the droplet shadowgraphs onto the sensor of a high-speed camera (Photron SA-Z), resulting in a projected pixel size of 1.78 µm/pixel. The diffraction limit of the microscope being about 1 µm, a single pixel corresponds to the smallest detected droplet size, yielding an area-equivalent diameter of 2 µm. The sensor was cropped to 1024 rows × 144 columns and images were acquired at 120 thousand frames per second (kfps). To increase the temporal resolution, three LED pulses were stacked in a single exposure (LED pulsed at 360 kHz), as shown in Fig. 2b, showing round regular droplets and a micro-exploding one. The images show that in the spray flame of SpraySyn2, droplets are significantly smaller and faster (the latter indicated by the greater vertical spacing between the shadowgraphs in an image) than in the flame of SpraySyn1. Images acquired in SpraySyn2 show a weak shadow "tail" in the downstream direction of the droplets. This corresponds to the tail in the LED emission between 50 and 120 ns. Images acquired in SpraySyn1 suffer less from such motion blurring as mean droplet velocities are about 30% lower than in SpraySyn2.

For a comprehensive description of thermally-induced droplet breakup in the SpraySyn2 burner, measurements were performed at various INN concentrations and HAB. At a concentration of 0.2 mol/l INN, measurements were performed at positions of 20, 30, 40, 50, 60, and 70 mm HAB. At 40 mm HAB, seven INN concentrations were investigated, at 0, 0.05, 0.1, 0.2, 0.3, 0.4, and 0.5 M. Each measurement was done on the centerline of the burner and comprises 150 000 images. An averaged bright field image – image without spray – was divided by each raw image.

2.2. Convolutional neural networks (CNNs)

This section describes the structure of Faster and Mask R–CNN as well as training, validation, and inference parameters and procedures. Fig. 3 shows the structure and data flow-chart of Mask and Faster R–CNN (enclosed within the red dashed line) as well as the training workflow (outside of the dashed line). Both networks have a backbone, which is a CNN primarily used for feature extraction from the supplied image. The feature maps are then fed into the region proposal network (RPN), a smaller CNN creating candidate regions of predefined sizes and aspect ratios (anchor boxes), grouping them either as region of interest (ROI) or as background. ROI-pool (in Faster R–CNN) and ROI-align (in Mask R–CNN) scale ROIs to a fixed size and map them onto the output feature maps of the backbone for further processing by the final component of the architecture, the head. Faster R–CNN uses two branches of fully connected layers to perform object classification and bounding box regression, yielding class labels and bounding box coordinates (Ren et al., 2017). Mask R–CNN uses an additional (two-layered) CNN to also generate the mask, i.e., the pixel-level segmentation, of an object (He et al., 2017). Masks, in contrast to bounding boxes, enable a more detailed extraction of object position and contour. Training, validation, and inference were performed on a GPU (NVIDIA GeForce 3070) and the image throughput (in frames per second) is indicated in Table 1.

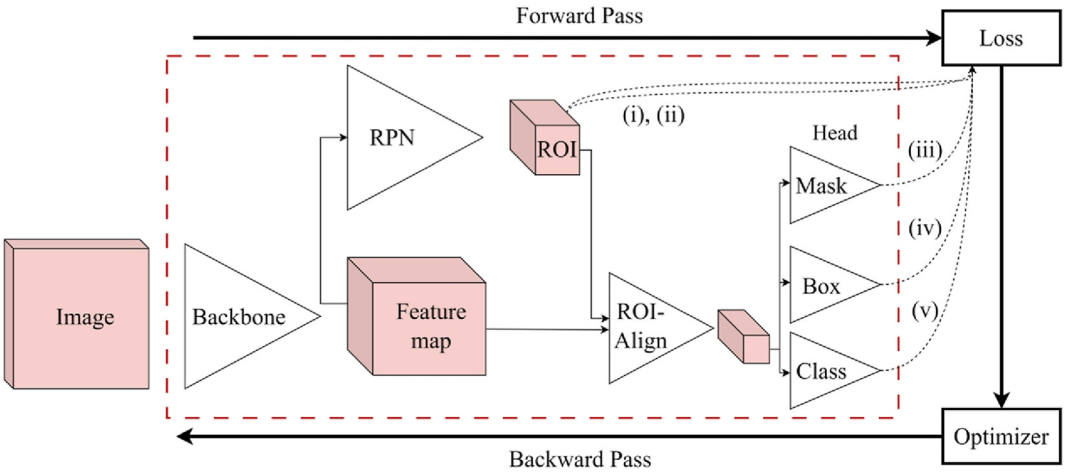

Fig. 3. Schematic overview of the architecture of Faster and Mask R–CNN (enclosed in red) and training procedure (surrounding). Components of the model are white, while data is pink. Note that Faster R–CNN does not have the mask branch in the head and uses ROI-pool instead of ROI-align. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2.2.1. Training

The main objective of training is to find a set of weighting values (weights) for which model predictions accurately align with human-specified target values that are here considered the ground truth. In inference, these weights can then be used to predict unlabeled data.

The outer part of Fig. 3 shows the workflow in training. Input images are used to calculate model predictions in a forward pass as described above. Discrepancies between predictions and ground truths are calculated in a loss function. In Mask R–CNN, it consists of five main components (four in Faster R–CNN as masks are not predicted). Each branch of the head, as well as the outputs of the RPN, contribute to the loss (Fig. 3(i)–(v)) (Ren et al., 2017). The RPN loss describes whether anchor boxes are predicted correctly as object containing or not by using a predicted probability (i) and bounding box' positions (ii). The loss of the head quantifies deviations amongst final (iii) mask, (iv) bounding box, and (v) class predictions and their respective target values (He et al., 2017; Ren et al., 2017).

Loss values are aggregated in (mini-)batches of multiple images before an optimizer is used to first compute the gradient of the loss with respect to the weights, and second update the weights using the gradient (backward pass). Some of the optimizer's user-defined configuration parameters (hyperparameters) are described briefly in the following. To appropriately scale each update step of the optimizer, a scaling factor – the learning rate (LR) – is introduced. To avoid overfitting and divergence of the network, the learning rate is scheduled to be reduced every fixed number of epochs. Proper tuning of momentum and weight decay further improve the networks' performance. The values for the hyperparameters were chosen based on previous works (Girshick, 2015; Lin et al., 2017; Ren et al., 2017) and manual testing iterations. They are listed in Table 1. The sensitivity of the network's performance (Mask R–CNN) to the learning rate is shown in the supplementary material S1.

Fig. 4 shows cropped regions from three example images. All training and validation images for Faster and Mask R–CNN were acquired in SpraySyn1. The left image of each pair shows a raw image with the corresponding bounding boxes of sequences (white), regular droplets (orange), erupted droplets (green), and fragments (magenta). Fragments are the result of a micro-explosion. An erupted droplet is the result of a puffing event, i.e., a weak disruption ejecting few secondary droplets where the parent droplet remains largely intact. The bounding boxes were manually generated with MATLAB's Image Labeler. The raw images, bounding box coordinates, and classes were used for training and validation of Faster R–CNN. The right image of each pair shows the corresponding mask image of in-focus droplets for training and validating the Mask R–CNN. Here, 'in-focus' refers to which droplets the human annotator considered as 'sufficiently' focused. Mask images were generated with ImageJ. This was done by thresholding the intensity in each region of the image where an in-focus droplet was identified, until the annotator considered the result appropriate. In total, 307 and 432 images, selected by a second, independent, annotator from a larger set of images, were used for training and validating Mask and Faster R–CNN, respectively. This corresponds to a total of 1125 masks and 1228 localized and classified illuminations and sequences. The annotation time per image was around 0.5 and 1.5 min for Faster and Mask R–CNN, respectively. In each network, the number of images used for validation was 20% of the number of total images, as indicated in Table 1. Both networks were loaded pretrained on the public-domain common objects in context (COCO) dataset, consisting of 2.5 million labeled, annotated, and segmented objects from more than 320 000 images and 91 categories (Lin et al., 2014).

2.2.2. Validation

The validation of the networks is done in every epoch after the training. To determine the performance of the model, the metrics precision and recall are used. These metrics are based on different types of errors and are commonly used for object detectors. The predictions of the model can be classified as correct or incorrect based on the intersection over union (IoU). This is the ratio of the area of intersection and the area of union between the prediction and the ground truth. Fig. 5 illustrates the IoU of objects being composed of 9, 25, and 81 pixels. If the IoU exceeds a certain threshold, typically between 0.5 and 0.95 (Lin et al., 2014), the mask or bounding box is classified as true positive (*TP*). If the model makes a prediction but a corresponding ground truth is missing, i.e., the IoU is below the threshold, the prediction is false positive (*FP*). If a ground truth exists but no prediction was made, the prediction is false negative (*FN*). Based on this, precision P and recall R become:

$$P = \frac{TP}{TP + FP} \qquad R = \frac{TP}{TP + FN} \tag{1}$$

Table 1
Hyperparameters for training of Mask and Faster R–CNN.

Parameter	Mask/Faster R–CNN
Learning rate	0.01
Weight decay	0.001
Momentum	0.9
LR scheduler: step size	10 epochs
LR scheduler: interval	0.1
Training images	246/346
Validation images	61/86
Batch size (training)	2
Epochs	50
Training throughput (fps)	11/16
Inference throughput (fps)	20/22

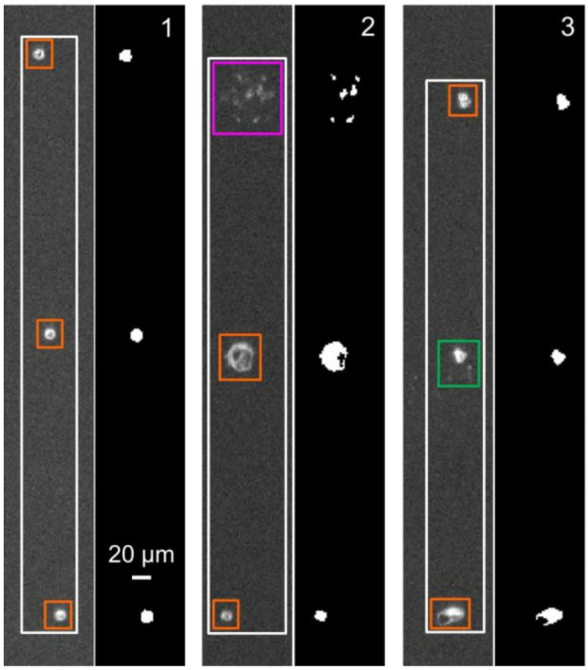

Fig. 4. Three pairs of training/validation images: Left: shadowgraphy images with bounding boxes of sequences (white), regular droplets (orange), erupted droplets (green), and fragments (magenta). Right: Corresponding mask images. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

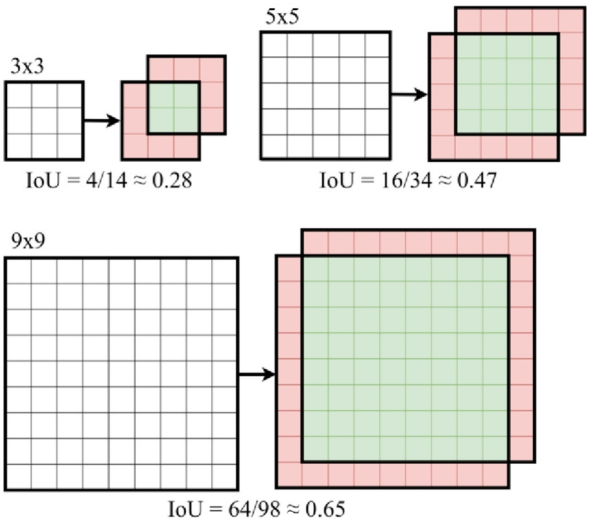

Fig. 5. Intersection (green) over union (red + green) of objects of different sizes. Note that in this illustration the objects are square, while in the real data they typically have irregular shapes. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

The precision is the fraction of true positives out of all positive predictions. Recall is the fraction of true positives out of all real (ground truth) objects and indicates the ability to recover all existing objects. After each epoch, thus after the evaluation of all validation images, precision and recall are calculated over all objects, regardless of class and size (Lin et al., 2014), and with an IoU threshold of 0.5. The comparatively low IoU threshold of 0.5 is chosen because most objects to be detected are small. For example, a 10 µm droplet, a typical droplet size in SFS, has an equivalent area of 5·5 pixels (compare to (Lin et al., 2014), where only objects with an area of less than 32·32 pixels are classified as small). This small size means the IoU is very sensitive to the absolute misalignment of ground truth and prediction, as illustrated in Fig. 5.

Fig. 6 shows the total loss during training, precision, and recall of both CNNs versus the number of epochs. Faster R–CNN and Mask

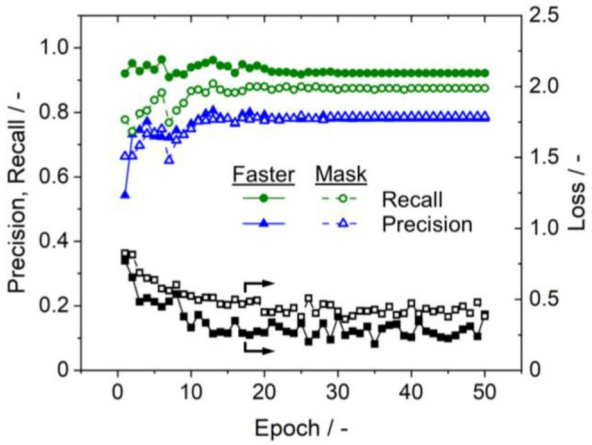

Fig. 6. Training loss, recall, and precision of Faster and Mask R–CNN.

R–CNN reach recall values of up 92% and 88%, respectively. That means that about 90% of the labeled (ground truth) objects are localized and segmented correctly while 10% are not found. The precision reaches about 80% for both networks, which means that 20% of the predicted objects do not have a corresponding ground truth. However, this predominantly concerns objects that were considered as slightly too defocused by the annotator which may have been counted by a different annotator. After 50 epochs, the losses – the summation of errors in a training epoch – converged to about 0.28 and 0.42 for Faster R–CNN and Mask R–CNN, respectively. The trained networks are now used to make predictions in shadowgraphy images from the SpraySyn2 burner.

2.2.3. Inference and post-processing

Fig. 7 schematically shows the inference procedure and the post-processing of the predictions from both networks. Mask R–CNN localizes and segments all instances of the object class droplet. Faster R–CNN localizes objects and classifies them into regular droplet, erupted droplet, fragments, and sequence. As multiple masks or bounding boxes might be found for a single prediction, non-maximum suppression (NMS) is applied to extract the one with the highest score (Felzenszwalb et al., 2010). In a first step, all predicted bounding boxes of an object are grouped that share an IoU above a certain threshold. Those are considered 'competing'. From that group, the bounding box with the highest score is kept while the remaining predictions are discarded. In our data set, the predicted masks' bounding boxes only intersect rarely, and if so, only to a small extent. An IoU of 0.05 was found suitable for Mask R–CNN NMS to not discard neighboring droplet instances. For Faster R–CNN, an intra-class NMS with an IoU of 0.5 and an inter-class NMS with an IoU of 0 were used. The latter means that if, for example, the bounding box of a regular droplet and an erupted droplet intersect, the one with the higher score is kept. Finally, mask pixels and bounding boxes with a score below 0.8 are also discarded from the results of Mask and Faster R–CNN, respectively. Visual inspection of predictions indicates that 98%, 96%, and 96% of the objects classified as regular droplet, erupted droplet, and fragments, respectively, are correct.

The final predictions from Faster and Mask R–CNN are combined as described in the following. For each predicted mask from Mask R–CNN, it is checked whether an intersecting bounding box from Faster R–CNN exists. If a bounding box is not available, the mask is discarded. Conversely, if a bounding box exists, but a mask is not available, the bounding box is discarded. From the masks, area-equivalent diameters are calculated and assigned to the class depending on the illumination's bounding box, i.e., regular, erupted, or fragments.

Fig. 8 shows three pairs of temporally consecutive shadowgraphy images and bounding box predictions from Faster R–CNN. In images where Faster R–CNN finds a sequence, all illuminations in that sequence are considered consecutive. For illuminations not enclosed by a sequence's bounding box, it is checked whether an illumination in certain horizontal (25 columns) and vertical distance (between 200 and 250 rows, depending on the mean droplet velocity at the HAB, which was taken from phase Doppler interferometry (PDI) measurements (Kirchmann et al., 2023; Prenting, 2023)) exists. This is also done with the topmost illuminations of the current image and the bottommost illuminations of the next image. Illuminations fulfilling this condition are also considered consecutive. A sequence either shows a regular droplet, an already disrupted droplet (erupted or fragments), puffing, or micro-explosion. A sequence showing puffing or micro-explosion indicates the evolution of a regular droplet into an erupted droplet or fragments.

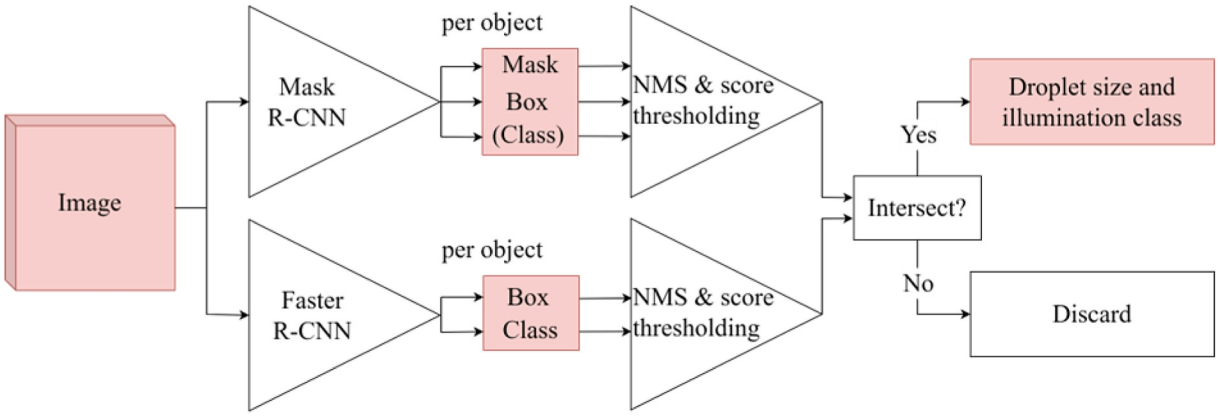

Fig. 7. Inference procedure and post-processing of results from Faster and Mask R–CNN.

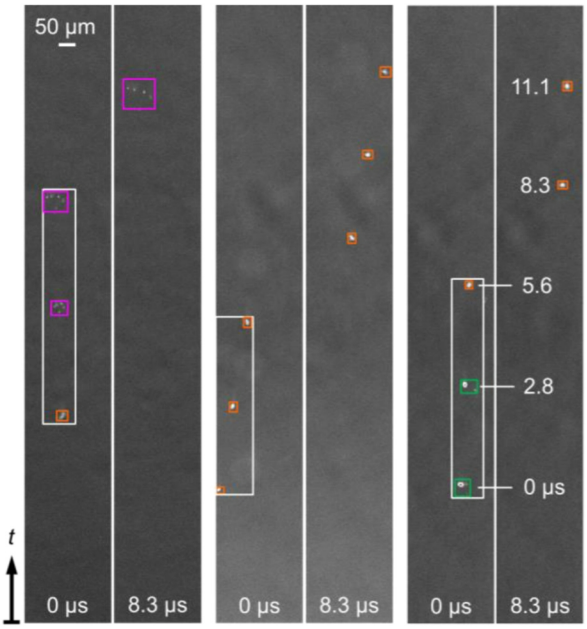

Fig. 8. Inference: Sequences of shadowgraphy images with overlays of bounding box predictions from Faster R–CNN showing (from left to right) a micro-explosion, a regular droplet, and an erupted droplet. Bounding box colours: white = sequence; orange = regular; green = erupted; magenta = fragments. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

2.3. Particle/droplet image analysis (PDIA)

Previously, we developed and used PDIA implemented in MATLAB R2022b to find in-focus droplets, infer their size and velocity, and classify shadowgraphs and their sequences. This is presented in detail in (Jüngst et al., 2022, 2023) and summarized in Fig. 9. First, images without spray – flat-field images – are divided by the raw spray images. This yields a brightness-inversion in the processed images. Then, thresholding the intensity gradient finds edges and thus identifies in-focus droplet instances. The intensity-gradient threshold was set to 0.0325, based on visual perception of the results. This results in accepting all droplets within ± 42.5 µm from the focal plane as 'in focus' (Jüngst et al., 2022). Subsequent local intensity-based thresholding finds the area and the area-equivalent diameter of each instance. Instances that are in close horizontal proximity and at a vertical spacing corresponding to their convection velocity are likely to stem from the three illuminations of the same droplet. Those instances are grouped in a sequence of consecutive illuminations and each instance is assigned to an illumination. An illumination contains either a single instance, a regular droplet, exactly two instances, i.e., an erupted droplet, or more than two instances, i.e., fragments. Analogous to the CNNs, a sequence then shows either a regular, already disrupted, or disrupting droplet. Visual inspection revealed that more than 95% of the images were classified correctly (Jüngst et al., 2023). The throughput of PDIA on a CPU was 35 fps.

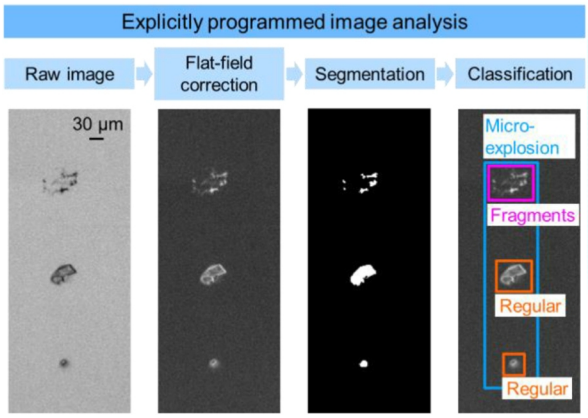

Fig. 9. Explicitly programmed particle/droplet image analysis (PDIA): Flat-fielding, gradient- and intensity-based thresholding, and sequence analysis to identify in-focus instances, estimate their morphological metrics, and analyze their spatiotemporal evolution.

3. Results

3.1. Illumination classifications and droplet masks

Fig. 10 shows raw images and masks of temporally consecutive droplet shadowgraphs, acquired at different HAB and INN concentrations. Bounding boxes localize and classify the illuminations in the raw images. The corresponding masks are shown from Mask R–CNN and PDIA. Sequence 1 shows an initially regular droplet that then stretches along the flow and ejects a secondary droplet downwards. Faster R–CNN and PDIA classify the second illumination as an erupted droplet. As a third droplet appears in the third illumination, PDIA classifies this illumination as fragments and thus the sequence as a micro-explosion. Even though Mask R–CNN also segments the third droplet, Faster R–CNN classifies the illumination as an erupted droplet, such that the CNNs find a puffing event for this sequence. This appears reasonable as the parent droplet largely remains and the disruption is relatively 'weak'. Sequence 2 shows a regular droplet, correctly identified by CNNs and PDIA, throughout all three illuminations. In the first illumination, Mask R–CNN finds a larger mask than PDIA. In the second and third illumination, PDIA and Mask R–CNN find masks of similar size and segment similar regions of the droplet. Sequence 3 shows a regular, deformed droplet in the first illumination. PDIA and Faster R–CNN find the droplet erupted in the second illumination. Faster R–CNN correctly recognizes fragments in the third illumination such that the sequence becomes a micro-explosion. Since most fragments are defocused or too small for reliable detection by either Mask R–CNN or PDIA, both find only two of them. This results in PDIA classifying the sequence erroneously as puffing.

Sequence 4 shows a regular droplet that deforms in the third illumination and appears to eject droplets and/or vapor to the right. The ejections are too weak in contrast for detection by PDIA or Mask R–CNN. However, Faster R–CNN classifies this illumination as an erupted droplet and thus the sequence as puffing. PDIA in contrast classifies the sequence as a regular droplet. CNNs and PDIA find puffing in sequence 5. The small droplet in the top of the third illumination, detected by Mask R–CNN, is not found by Faster R–CNN. Both methods identify sequence 6 correctly as a micro-explosion. However, both apparently do not segment all fragments (probably because they are too defocused or small). In general, Mask R–CNN appears to find smoother edges of the masks than PDIA.

3.2. Droplet diameters and disruption probabilities

Fig. 11 shows frequency histograms of the area-equivalent diameter of regular droplets, erupted droplets, and fragments at 30 mm HAB and 0.2 mol/l INN from PDIA and CNNs. In the CNN data, the diameters are from Mask R–CNN and their classes from Faster R–CNN. Also indicated are log-normal fits to the data and arithmetic mean diameters. The CNN data comprises a total of 42 493, 6753, and 357 regular droplets, erupted droplets, and fragments, respectively. The PDIA data comprises 42 717, 4416, and 777 regular droplets, erupted droplets, and fragments, respectively. The diameter of regular droplets determined by the CNNs is 17% larger than that from PDIA. However, both methods find very similar diameters for erupted droplets and fragments. These findings qualitatively also hold at different INN concentrations and HABs. Mean droplet diameters, velocities, and concentrations versus the INN concentration and HAB are given in appendix A1. PDIA results indicate that erupted droplets, and fragments are as large in SpraySyn2 as in SpraySyn1 (7.6 and 6.4 µm), while regular droplets are about 20% smaller than in SpraySyn1 (11.4 µm) (Jüngst et al., 2023).

The aggregation of consecutive illuminations and the analysis of the sequences, as introduced in Sections 2.2 and 2.3, yield the fractions of regular + disrupted and disrupting droplets. These fractions can be understood as relative frequencies or probabilities. On average, a coherent sequence of a droplet is 3–3.5 illuminations (depending on the HAB), corresponding to 5.6–6.9 µs in duration (4–5 illuminations in SpraySyn1 (Jüngst et al., 2023)). For comparability with the previous results, the probabilities in Fig. 12 are given per 10 µs lifetime by assuming the disruption probability follows a binomial distribution. Fig. 12a shows the probabilities from CNNs and PDIA as a function of the INN concentration at 40 mm HAB. Qualitatively consistent with data from SpraySyn1, the probability for disruption increases with the INN concentration to 14% at 0.5 mol/l INN (11.5% in SpraySyn1 (Jüngst et al., 2023)). The difference

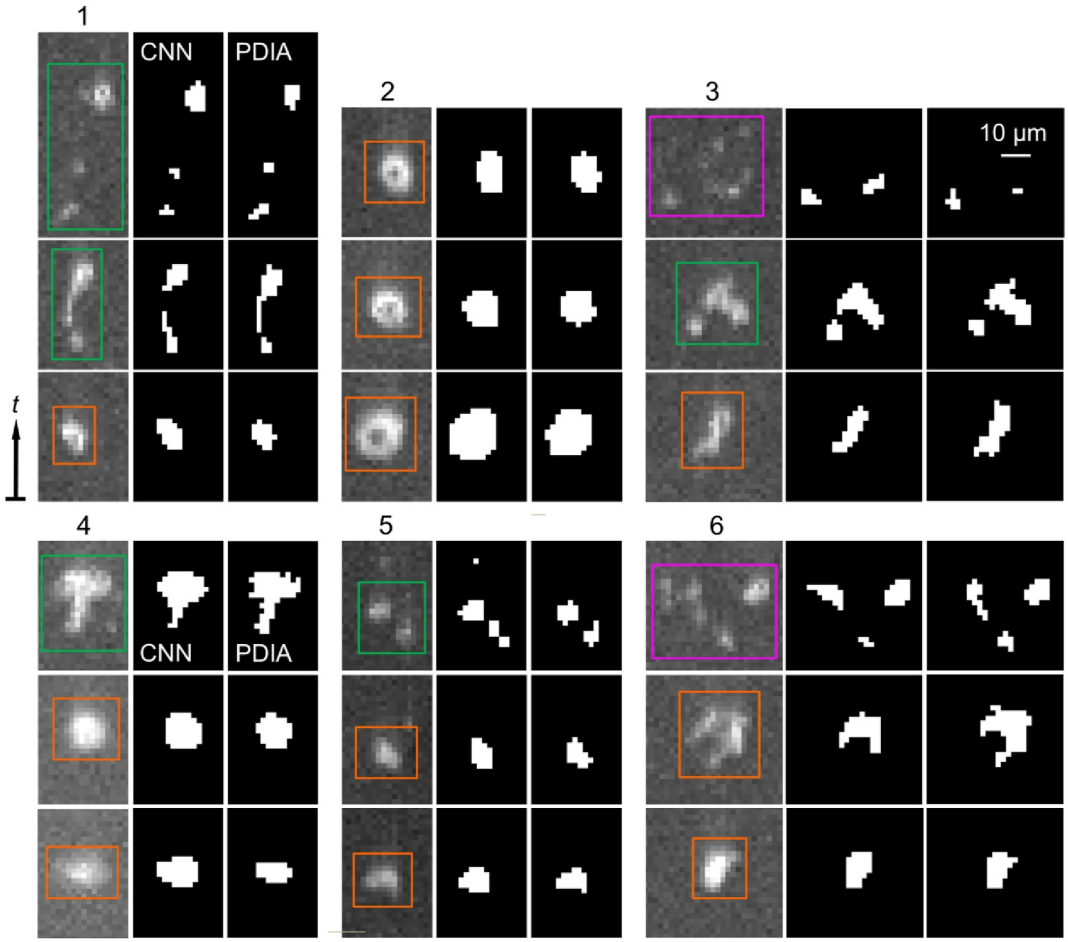

Fig. 10. Sequences of consecutive droplet shadowgraphs and masks predicted by the CNNs and by PDIA. Colored bounding boxes indicate the classification by Faster R–CNN: orange = regular; green = erupted; magenta = fragments. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

between puffing probability and micro-explosion probability is much larger in SpraySyn2. Throughout all INN concentrations, CNNs and PDIA find the probability for puffing 8 and 4 times higher, respectively, than for micro-explosion (1.5 times in SpraySyn1 (Jüngst et al., 2023)). Between 0.1 and 0.4 mol/l INN, CNNs find the probability for puffing higher than PDIA and constantly increasing to 16% at 0.4 mol/l INN before it decreases again towards 0.5 mol/l INN. The fact that the neural networks mostly predict a higher probability for puffing than PDIA is consistent with Fig. 10. Weak disruptions, where ejected secondary droplets are barely visible and thus not segmented, are classified as a regular droplet by PDIA. In contrast, Faster R–CNN correctly identifies such illuminations as an erupted droplet and thus the corresponding sequence as puffing.

Fig. 12b shows the probabilities versus the HAB with 0.2 mol/l INN. The top axis of Fig. 12b shows the mean residence time of a droplet in the spray flame, determined from the mean droplet velocity (Kirchmann et al., 2023; Prenting, 2023), at each HAB where measurements were recorded. Note that the residence time scales non-linearly with the HAB due to deceleration of the flow field, as highlighted by the breaks in the top axis in Fig. 12b (The disruption probabilities as well as mean droplet diameters, mean droplet velocities, and droplet concentrations as a function of the linearly-scaled residence time are shown in the supplementary material S2). Even though Faster R–CNN can detect illuminations and sequences at 20 mm HAB, the high droplet density causes erroneous results. Neighboring individual droplets were sometimes classified as erupted droplets or fragments. Therefore, the probabilities at 20 mm HAB for PDIA and CNNs (greyed out in Fig. 12b) were determined by manual inspection of 5000 images, counting sequences of regular droplets, disrupted droplets, puffing, and micro-explosion (total of 1600 sequences found). CNNs and PDIA find the probability of disruption generally increasing towards higher HAB, i.e., longer droplet residence times, whereas it was less dependent on the HAB in SpraySyn1. Throughout all HAB, the probability of puffing exceeds that of micro-explosion. Again, CNNs find higher probabilities for puffing than PDIA. In contrast to SpraySyn1, where droplets remain up to 70 mm HAB, here almost all droplets have evaporated at already 50 mm HAB. Again, PDIA and CNNs predict similar probabilities of micro-explosion.

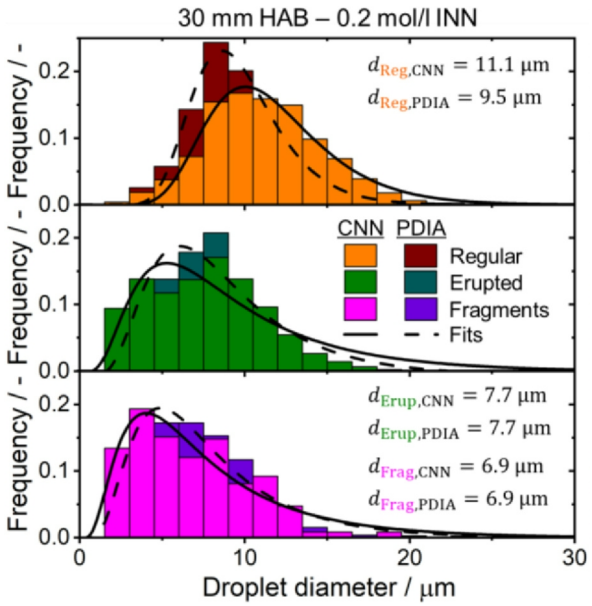

Fig. 11. Frequency histograms of the area-equivalent diameter of regular droplets, erupted droplets, and fragments, derived from PDIA and CNNs. Images acquired at 30 mm HAB and with 0.2 mol/l INN.

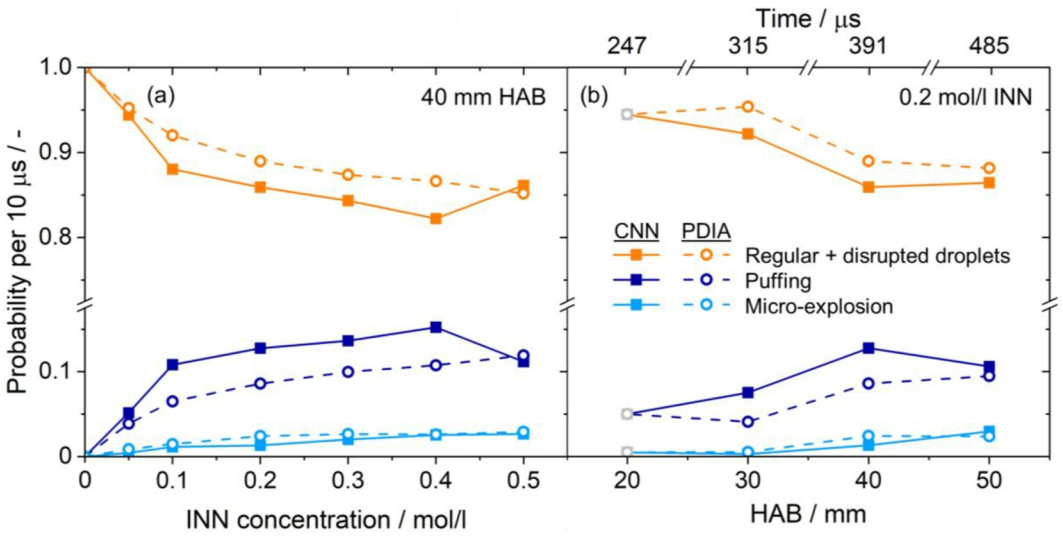

Fig. 12. (a, b) Probability per 10 µs for finding regular + disrupted droplets, puffing, or micro-explosion versus the INN concentration and HAB. Data at 20 mm HAB (greyed out) were obtained by manual inspection because the high droplet concentration made automated analysis too unreliable.

4. Conclusions

Mask and Faster region-based convolutional neural networks (Mask R–CNN and Faster R–CNN) were trained and applied on high-speed shadowgraphy images of droplets in the spray flame of a burner for nanoparticle synthesis. The burner design of SpraySyn2 creates a fine spray of micrometer-sized droplets that travel with up to 140 m/s through the measurement volume, posing challenges for the imaging and subsequent image analysis. As precursor-laden droplets are subject to thermally-induced droplet breakup, i.e., droplet disruption by puffing and micro-explosion, the networks were used to detect such events and extract statistical quantities. The results were compared with those from traditional, explicitly programmed particle/droplet image analysis (PDIA).

In the experiments, to increase the temporal resolution in imaging of droplet shadowgraphs, three light pulses from an LED were stacked onto a single frame of a high-speed camera. Faster R–CNN was used to localize and classify shadowgraphs into regular droplets, erupted droplets (as a result of puffing), or fragments (as a result of a micro-explosion). Also, sequences of consecutive droplet

shadowgraphs/LED-illuminations were localized. Mask R–CNN created masks of in-focus droplet shadowgraphs which were combined with the results from Faster R–CNN, yielding size and class of droplet shadowgraphs. The precision of the trained networks reaches 80% after 50 epochs. The recall reaches 92% and 88% for Faster R–CNN and Mask R–CNN, respectively.

Mask R–CNN and PDIA identify mostly the same droplets as in-focus and mask similar regions of the instances. Discarded droplets are too small or defocused, thus of very low contrast in the images, for being reliably found by PDIA or Mask R–CNN. Particularly in shadowgraphs of erupted droplets and fragments, small secondary droplets are often discarded by Mask R–CNN and PDIA. However, as in the CNN framework classification and segmentation are independent from each other, Faster R–CNN finds and correctly classifies such illuminations.

Due to the above, the CNNs find an on average 30% higher probability for disruption than PDIA. Among the INN concentration and the HAB, both methods find probabilities per 10 μs droplet lifetime increasing to 14% at 0.5 mol/l INN and 13% at 50 mm HAB. On average, Mask R–CNN finds regular droplets being 22% larger than PDIA while erupted droplets and fragments are of similar size in both methods. The probability for droplet disruption is generally higher in SpraySyn2 than in SpraySyn1. At the same time, disrupted droplets are 31% smaller than regular droplets in SpraySyn2 while they are 45% smaller in SpraySyn1. This results in a larger relative increase of the specific surface area in SpraySyn1 than in SpraySyn2.

As the networks performed well on data from SpraySyn2 (droplets smaller and faster than in SpraySyn1), we assume they will perform well also in different burners, under very different operating conditions, or with a different solution and precursor chemistry. However, the current Faster R–CNN fails in classifying droplet shadowgraphs and sequences under very high droplet concentration, i.e., low HAB. Training the networks on images from SpraySyn2 in addition to those from SpraySyn1, or a mixture of images from both burner versions, could enhance the networks' performances.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This work was supported by the German Research Foundation (DFG) within the Priority Program SPP 1980 SpraySyn under project number 447500013. The DFG Mercator Fellowship held by Greg Smallwood is supported by project number 374957702. We thank Alper Emre Sahin for image annotation.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jaerosci.2023.106314.

Appendix

A1. Droplet diameter, velocity, and concentration

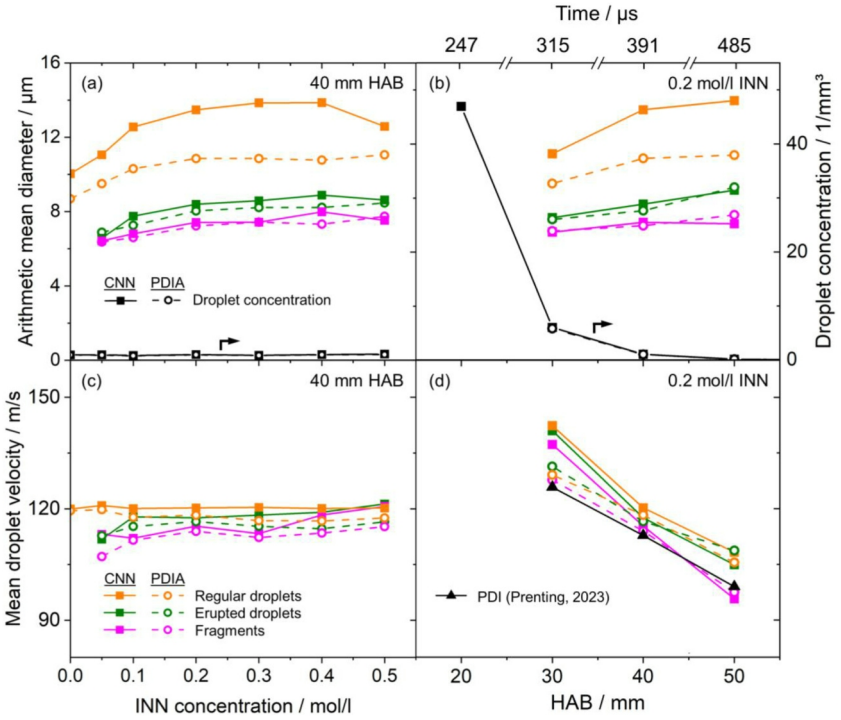

Fig. 13. (a, b) Arithmetic mean diameter of regular droplets, erupted droplets, and fragments as well as total droplet concentration versus the INN concentration and HAB. (c, d) Mean droplet velocity of regular droplets, erupted droplets, and fragments versus the INN concentration and HAB. Note that PDI measurements show the mean droplet velocity in a spray flame of EtOH/EHA.

For PDIA and Mask R–CNN, droplet concentrations were calculated by summing up all in-focus instances in a measurement and dividing the sum by the number of images and the measurement volume. Since the depth of the CNN's measurement volume is not known, the same value as in PDIA, ± 42.5 µm from the focal plane, was chosen. CNNs and PDIA find almost same droplet concentrations from 0 to 0.5 mol/l INN and 30–50 mm HAB. Most likely, this results from similar visual perceptions of what is 'in-focus' during PDIA threshold-setting and image annotation in CNNs. In other words, the CNNs were trained to also accept instances within ± 42.5 µm from the focal plane. Even though detections and classifications of illuminations and sequences at 20 mm HAB are erroneous, predictions from Mask R–CNN can be used to derive the overall droplet concentration at that height. Droplet velocities were calculated from the distance between two consecutive illuminations and the time between two LED pulses, as described in (Jüngst et al., 2023). Fig. 13d also shows mean droplet velocities from PDI measurements in a spray flame of EtOH/EHA. Velocities from PDI, PDIA, and CNN generally agree well (maximum deviation of 9% between CNN and PDI at 30 mm HAB).

References

Bieber, M., Al-Khatib, M., Fröde, F., Pitsch, H., Reddemann, M., Schmid, H. J., et al. (2021). Influence of angled dispersion gas on coaxial atomization, spray and flame formation in the context of spray-flame synthesis of nanoparticles. *Experiments in Fluids, 62*(5), 98. https://doi.org/10.1007/s00348-021-03196-6

Chaussonnet, G., Gepperth, S., Holz, S., Koch, R., & Bauer, H.-J. (2021). Best architecture of an artificial neural network to model prefilming airblast atomization: Not so deep learning. *Journal of Engineering for Gas Turbines & Power, 143*(7), Article 071006. https://doi.org/10.1115/1.4049911

Chaussonnet, G., Lieber, C., Yikang, Y., Gu, W., Bartschat, A., Reischl, M., et al. (2019). *Towards DeepSpray: Using convolutional neural network to post-process shadowgraphy images of liquid atomization*. Karlsruhe Institute of Technology. https://doi.org/10.5445/IR/1000097897/v3. Technical report.

Devillers, R., Nugue, M., Chan Hon Tong, A., Le Besnerais, G., & Pichillou, J. (2019). *Experimental analysis of aluminum-droplet combustion in solid- propellant conditions using deep learning*. Madrid, Spain: European Conference For Aeronautics And Space Sciences. https://doi.org/10.13009/eucass2019-593

Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2010). Object detection with discriminatively trained part-based models. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(9), 1627–1645. https://doi.org/10.1109/TPAMI.2009.167

Frei, M., & Kruis, F. E. (2018). Fully automated primary particle size analysis of agglomerates on transmission electron microscopy images via artificial neural networks. *Powder Technology, 332*, 120–130. https://doi.org/10.1016/j.powtec.2018.03.032

Frei, M., & Kruis, F. E. (2020). Image-based size analysis of agglomerated and partially sintered particles via convolutional neural networks. *Powder Technology, 360*, 324–336. https://doi.org/10.1016/j.powtec.2019.10.020

Frei, M., & Kruis, F. E. (2022). Image-based analysis of dense particle mixtures via mask R-CNN. *Eng, 3*(1), 78–98. https://doi.org/10.3390/eng3010007

Fu, Y., & Liu, Y. (2019). BubGAN: Bubble generative adversarial networks for synthesizing realistic bubbly flow images. *Chemical Engineering Science, 204*, 35–47. https://doi.org/10.1016/j.ces.2019.04.004

Girshick, R. (2015). Fast R-CNN. In *IEEE international conference on computer vision, santiago, Chile*. https://doi.org/10.1109/ICCV.2015.169

Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *IEEE conference on computer vision and pattern recognition, columbus, USA*. https://doi.org/10.1109/ICCV.2015.169

Haas, T., Schubert, C., Eickhoff, M., & Pfeifer, H. (2020). BubCNN: Bubble detection using Faster RCNN and shape regression network. *Chemical Engineering Science, 216*, Article 115467. https://doi.org/10.1016/j.ces.2019.115467

Hadikhani, P., Borhani, N., Hashemi, H., S, M., & Psaltis, D. (2019). Learning from droplet flows in microfluidic channels using deep neural networks. *Scientific Reports, 9*(1), 8114. https://doi.org/10.1038/s41598-019-44556-x

He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2017). Mask R-CNN. In *IEEE international conference on computer vision, venice, Italy*. https://doi.org/10.1109/ICCV.2017.322

Jiang, Y., Jeon, H. Y., Tian, L., & Bode, L. E. (2010). Measuring particle size distribution using LED-illumination. *International Journal of Multiphase Flow, 36*(3), 193–201. https://doi.org/10.1016/j.ijmultiphaseflow.2009.11.004

Jüngst, N., Skenderović, I., Suedholt, B., Smallwood, G. J., Kruis, F. E., & Kaiser, S. A. (2023). High-speed imaging and statistical analysis of droplet puffing and micro-explosion in spray-flame synthesis. *Applications in Energy and Combustion Science, 15*, Article 100167. https://doi.org/10.1016/j.jaecs.2023.100167

Jüngst, N., Smallwood, G. J., & Kaiser, S. A. (2022). Visualization and image analysis of droplet puffing and micro-explosion in spray-flame synthesis of iron oxide nanoparticles. *Experiments in Fluids, 63*(3), 60. https://doi.org/10.1007/s00348-022-03411-y

Kashdan, J. T., Shrimpton, J. S., & Whybrew, A. (2007). A digital image analysis technique for quantitative characterisation of high-speed sprays. *Optics and Lasers in Engineering, 45*(1), 106–115. https://doi.org/10.1016/j.optlaseng.2006.03.006

Kim, K. S., & Kim, S.-S. (1994). Drop sizing and depth-of-field correction in TV imaging. *Atomization and Sprays, 4*(1), 65–78. https://doi.org/10.1615/AtomizSpr.v4.i1.30

Kim, Y., & Park, H. (2021). Deep learning-based automated and universal bubble detection and mask extraction in complex two-phase flows. *Scientific Reports, 11*(1), Article 8940. https://doi.org/10.1038/s41598-021-88334-0

Kirchmann, J., Kronenburg, A., Prenting, M. M., Karaminejad, S., Dreier, T., Endres, T., et al. (2023). Characterizing the SpraySyn burners with MMC-LES. *Applications in Energy and Combustion Science, 15*, Article 100182. https://doi.org/10.1016/j.jaecs.2023.100167

Koh, K., Kim, J., & Lee, S. (2001). Determination of in-focus criteria and depth of field in image processing of spray particles. *Atomization and Sprays, 11*(4), 317–333. https://doi.org/10.1615/AtomizSpr.v11.i4.20

Kumar, A., & Sahu, S. (2020). Liquid jet disintegration memory effect on downstream spray fluctuations in a coaxial twin-fluid injector. *Physics of Fluids, 32*, Article 73302. https://doi.org/10.1063/5.0009188

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278–2324. https://doi.org/10.1109/5.726791

Lecuona, A., Sosa, P. A., Rodriguez, P. A., & Zequeira, R. (2000). Volumetric characterization of dispersed two phase flows by digital image analysis. *Measurement Science and Technology, 11*(8), Article 1152. https://doi.org/10.1088/0957-0233/11/8/309

Lee, S., & Kim, Y.-D. (2004). Sizing of spray particles using image processing techniques. *Journal of Mechanical Science and Technology, 18*(6), 879–894. https://doi.org/10.1007/BF02990860

Liang, Z., Nie, Z., An, A., Gong, J., & Wang, X. (2019). A particle shape extraction and evaluation method using a deep convolutional neural network and digital image processing. *Powder Technology, 353*, 156–170. https://doi.org/10.1016/j.powtec.2019.05.025

Li, T., Liang, Z., Dreizler, A., & Boehm, B. (2023). Accurate determination of homogeneous ignition of single solid fuel particles enabled by machine learning. *Fuel, 338*, Article 127171. https://doi.org/10.1016/j.fuel.2022.127171

Lin, T.-Y., Dollár, P., Girshick, R., He, K., Hariharan, B., & Belongie, S. (2017). *Feature pyramid networks for object detection*. Hawaii, USA: IEEE Conference on Computer Vision and Pattern Recognition. https://doi.org/10.1109/TPAMI.2016.2577031

Li, J., Shao, S., & Hong, J. (2021). Machine learning shadowgraph for particle size and shape characterization. *Measurement Science and Technology, 32*(1), Article 015406. https://doi.org/10.1088/1361-6501/abae90

Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., et al. (2014). *Microsoft COCO: Common objects in context*. Zurich, Switzerland: European Conference on Computer Vision. https://doi.org/10.1007/978-3-319-10602-1_48

Mädler, L., Kammler, H. K., Mueller, R., & Pratsinis, S. E. (2002). Controlled synthesis of nanostructured particles by flame spray pyrolysis. *Journal of Aerosol Science, 33*(2), 369–389. https://doi.org/10.1016/S0021-8502(01)00159-8

Malot, H., & Blaisot, J.-B. (2000). Droplet size distribution and sphericity measurements of low-density sprays through image analysis. *Particle & Particle Systems Characterization, 17*(4), 146–158. https://doi.org/10.1088/0957-0233/21/2/025501

Narayanan, B. N., Gunasekaran, S., Ivarson, J., & Maneck, L. (2021). *Deep learning algorithm for atomization characterization using shadowgraph images*. Dayton, USA: IEEE National Aerospace and Electronics Conference. https://doi.org/10.1109/NAECON49338.2021.9696443

Pan, J., Libera, J. A., Paulson, N. H., & Stan, M. (2021). Flame stability analysis of flame spray pyrolysis by artificial intelligence. *International Journal of Advanced Manufacturing Technology, 114*(7), 2215–2228. https://doi.org/10.1007/s00170-021-06884-z

Pathan, R., Lim, W. L., Lau, S. L., Ho, C. C., & Kare, P. (2022). Experimental analysis of U-net and mask R-CNN for segmentation of synthetic liquid spray. In *IEEE international conference on computing, Kota Kinabalu, Malaysia*. https://doi.org/10.1109/ICOCO56118.2022.10031951

Paulson, N. H., Libera, J. A., & Stan, M. (2020). Flame spray pyrolysis optimization via statistics and machine learning. *Materials and Design, 196*, Article 108972. https://doi.org/10.1016/j.matdes.2020.108972

Petrucci, L., Ricci, F., Martinelli, R., & Mariani, F. (2022). Detecting the flame front evolution in spark-ignition engine under lean condition using the mask R-CNN approach. *Vehicles, 4*(4), 978–995. https://doi.org/10.3390/vehicles4040053

Poletaev, I., Pervunin, K., & Tokarev, M. (2016). Artificial neural network for bubbles pattern recognition on the images. *Journal of Physics: Conference Series, 754*, Article 072002. https://doi.org/10.1088/1742-6596/754/7/072002. Article.

Poletaev, I., Tokarev, M. P., & Pervunin, K. S. (2020). Bubble patterns recognition using neural networks: Application to the analysis of a two-phase bubbly jet. *International Journal of Multiphase Flow, 126*, Article 103194. https://doi.org/10.1016/j.ijmultiphaseflow.2019.103194

Prenting, M. M. (2023). *Optical and laser-based in situ diagnostics in lab-scale and pilot-scale spray flames for nanoparticle synthesis*. [Unpublished doctoral dissertation, University of Duisburg-Essen.

Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 39*(6), 1137–1149. https://doi.org/10.1109/TPAMI.2016.2577031

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In *International conference on medical image computing and computer-assisted intervention, Munich, Germany*. https://doi.org/10.1007/978-3-319-24574-4_28

Rosebrock, C. D., Riefler, N., Wriedt, T., Mädler, L., & Tse, S. D. (2013). Disruptive burning of precursor/solvent droplets in flame-spray synthesis of nanoparticles. *AIChE Journal, 59*(12), 4553–4566. https://doi.org/10.1002/aic.14234

Rutkowski, G. P., Azizov, I., Unmann, E., Dudek, M., & Grimes, B. A. (2022). Microfluidic droplet detection via region-based and single-pass convolutional neural networks with comparison to conventional image analysis methodologies. *Machine Learning with Applications, 7*, Article 100222. https://doi.org/10.1016/j.mlwa.2021.100222

Schäfer, J., Schmitt, P., Hlawitschka, M. W., & Bart, H.-J. (2019). Measuring particle size distributions in multiphase flows using a convolutional neural network. *Chemie Ingenieur Technik, 91*(11), 1688–1695. https://doi.org/10.1002/cite.201900099

Schneider, F., Suleiman, S., Menser, J., Borukhovich, E., Wlokas, I., Kempf, A., et al. (2019). SpraySyn—a standardized burner configuration for nanoparticle synthesis in spray flames. *Review of Scientific Instruments, 90*(8), Article 085108. https://doi.org/10.1063/1.5090232

Shinjo, J., Xia, J., Ganippa, L. C., & Megaritis, A. (2014). Physics of puffing and microexplosion of emulsion fuel droplets. *Physics of Fluids, 26*(10), Article 103302. https://doi.org/10.1063/1.4897918

Sipkens, T. A., Frei, M., Baldelli, A., Kirchen, P., Kruis, F. E., & Rogak, S. N. (2021). Characterizing soot in TEM images using a convolutional neural network. *Powder Technology, 387*, 313–324. https://doi.org/10.1016/j.powtec.2021.04.026

Stodt, M., Groeneveld, J., Mädler, L., Kiefer, J., & Fritsching, U. (2022). Microexplosions of multicomponent drops in spray flames. *Combustion and Flame, 240*(6), Article 112043. https://doi.org/10.1016/j.combustflame.2022.112043

Wang, Z., He, F., Zhang, H., Hao, P., Zhang, X., & Li, X. (2022a). Three-dimensional measurement of the droplets out of focus in shadowgraphy systems via deep learning-based image-processing method. *Physics of Fluids, 34*(7), Article 073301. https://doi.org/10.1063/5.0097375

Wang, Z., He, F., Zhang, H., Hao, P., Zhang, X., & Li, X. (2022b). Characterization of the in-focus droplets in shadowgraphy systems via deep learning-based image processing method. *Physics of Fluids, 34*(11), Article 113316. https://doi.org/10.1063/5.0121174

Witte, A., & Mädler, L. (2023). Single droplet combustion of iron nitrate-based precursor solutions: Investigation of time- and size scales of isolated burning FSP-droplets. *Applications in Energy and Combustion Science, 14*, Article 100147. https://doi.org/10.1016/j.jaecs.2023.100147

Wu, Y., Lin, M., & Rohani, S. (2020). Particle characterization with on-line imaging and neural network image analysis. *Chemical Engineering Research and Design, 157*, 114–125. https://doi.org/10.1016/j.cherd.2020.03.004

Yee, J., Yamanaka, A., & Tagawa, Y. (2022). Image features of a splashing drop on a solid surface extracted using a feedforward neural network. *Physics of Fluids, 34* (1), Article 013317. https://doi.org/10.1063/5.0077050

Zhang, F., Zhang, Q., Xiao, Z., Wu, J., & Liu, Y. (2019). Spherical nanoparticle parameter measurement method based on Mask R-CNN segmentation and edge fitting. In *International conference on computing and pattern recognition, New York, USA*. https://doi.org/10.1145/3373509.3373590