



# Particle detection and size recognition based on defocused particle images: a comparison of a deterministic algorithm and a deep neural network

Sebastian Sachs<sup>1</sup> · Manuel Ratz<sup>1</sup> · Patrick Mäder<sup>2</sup> · Jörg König<sup>1</sup> · Christian Cierpka<sup>1</sup>

Received: 15 September 2022 / Revised: 14 October 2022 / Accepted: 6 January 2023  
© The Author(s) 2023

## Abstract

The systematic manipulation of components of multimodal particle solutions is a key for the design of modern industrial products and pharmaceuticals with highly customized properties. In order to optimize innovative particle separation devices on microfluidic scales, a particle size recognition with simultaneous volumetric position determination is essential. In the present study, the astigmatism particle tracking velocimetry is extended by a deterministic algorithm and a deep neural network (DNN) to include size classification of particles of multimodal size distribution. Without any adaptation of the existing measurement setup, a reliable classification of bimodal particle solutions in the size range of 1.14  $\mu\text{m}$ –5.03  $\mu\text{m}$  is demonstrated with a precision of up to 99.9 %. Concurrently, the high detection rate of the particles, suspended in a laminar fluid flow, is quantified by a recall of 99.0 %. By extracting particle images from the experimentally acquired images and placing them on a synthetic background, semi-synthetic images with consistent ground truth are generated. These contain labeled overlapping particle images that are correctly detected and classified by the DNN. The study is complemented by employing the presented algorithms for simultaneous size recognition of up to four particle species with a particle diameter in between 1.14  $\mu\text{m}$  and 5.03  $\mu\text{m}$ . With the very high precision of up to 99.3 % at a recall of 94.8 %, the applicability to classify multimodal particle mixtures even in dense solutions is confirmed. The present contribution thus paves the way for quantitative evaluation of microfluidic separation and mixing processes.

## 1 Introduction

The mixing or separation of microparticles with various sizes is of crucial importance for biological or chemical analysis as well as for customizing pharmaceutical products, paints and coatings (Sajeesh and Sen 2014; Hessel et al. 2005). Innovative processes based on microfluidic lab-on-a-chip devices have achieved major advances for the systematic manipulation of the composition of particle solutions (Ahmad et al. 2019; Gossett et al. 2010; Miyagawa and Okada 2021; Zhang et al. 2020; Zoupanou et al. 2021). In order to evaluate the performance and further optimize these processes, a particle size recognition is essential.

Commercially available particle size analysis devices are typically designed for the separate examination of samples obtained from outlets (Shekunov et al. 2007; Gross-Rother et al. 2020). Common devices mostly utilize particle-light interactions (Stetefeld et al. 2016; Dannhauser et al. 2014; Xu 2015), resistive pulse sensing (Zhang et al. 2019; Weatherall and Willmott 2015; Grabarek et al. 2019; Caputo et al. 2019), suspended microchannel resonators (Gross-Rother et al. 2020; Stockslager et al. 2019; de Pastina and Villanueva 2020), or microscopic imaging (Sharma et al. 2010; Rice et al. 2013). Among the available techniques, the microfluidic flow cytometry emerged as a widely used tool for the analysis of suspended particle mixtures (Mach et al. 2011; Picot et al. 2012; Gong et al. 2018; McKinnon 2018). By means of hydrodynamic focusing, the particles are initially guided to a narrow pathline within a capillary and get spatially separated in flow direction. Following the typically laser-based illumination, a single-particle analysis based on an emitted fluorescence signal or scattered light is achieved. By analyzing the samples in a well-defined microfluidic environment, a high throughput and automated

✉ Sebastian Sachs  
sebastian.sachs@tu-ilmenau.de

<sup>1</sup> Institute of Thermodynamics and Fluid Mechanics,  
Technische Universität Ilmenau, 98693 Ilmenau, Germany

<sup>2</sup> Data-Intensive Systems and Visualization Group (dAISY),  
Technische Universität Ilmenau, 98693 Ilmenau, Germany

measurements are realized. The gathered data allow a statistical process evaluation; however, it does not provide a deep insight into the particle behavior during the process. In addition, a characterization and optimization of the process can hardly be done. A label-free and in-process size recognition of particles is further realized by electrical impedance flow cytometry (de Bruijn et al. 2021; Zhong et al. 2021). However, the required integration of electrode arrays in a microchannel increases the fabrication effort and still provides no insight into the spatial distribution of the involved particle species. Non-invasive imaging techniques enable in-process size recognition and detection of particles without significantly affecting the process, thus manifesting a solid database for optimization in early stages (Farkas et al. 2021; Gao et al. 2018; Chen et al. 2019). In this context, a classification of the suspended particle species is usually achieved by image evaluation while using various fluorescent dyes in conjunction with multiple long-pass filters or a color camera (Ahmed et al. 2018; Sehgal and Kirby 2017). The extended demands on the measurement setup represent a major drawback of these techniques. To overcome these limitations, the present work aims at the simultaneous size recognition and three-dimensional position determination of suspended particles by analyzing defocused particle images obtained from astigmatism particle tracking velocimetry (APTV) measurements. In this way, no adaptation of the existing measurement setup is necessary, while the spatial distribution and classification of the involved particle species are obtained in the process.

The APTV evolved as a reliable tool for three-dimensional, three-component (3D3C) velocity measurements in microfluidic environments with limited optical access by integrating a cylindrical lens in the optical path to the camera (Cierpka et al. 2010). As this lens reduces the focal length along one axis of the camera sensor while keeping the focal length constant along the other axis, captured particle images appear elliptically shaped depending on their depth position. Based on the lengths of the semi-axes of the ellipses ( $a_x, a_y$ ), the depth position ( $z$ ) of the particles is mapped by a calibration function that is obtained from previously captured particles with known  $z$ -position (Cierpka et al. 2011). Due to the geometric dependency between the size of a particle and the shape of its particle image (Rossi et al. 2012), a monomodal size distribution is commonly used for velocity measurements (Rossi and Kähler 2014). However, by including particle size recognition algorithms in the evaluation of particle images, the depth position of particles with multimodal size distribution can be determined based on the corresponding calibration function. In this way, precise APTV measurements are achievable throughout the entire measurement volume with only one fluorescent dye and one camera, even when a multimodal particle solution is used (König et al. 2020).

The size recognition is accomplished by two algorithms. First, a deterministic algorithm was developed to achieve the discrimination between various particle sizes by the intensity of the particle images (see Sect. 3.2). Second, a deep neural network (DNN) was trained for object detection of particles images. In recent years, DNNs gained increased attention for the fast segmentation and object classification, which make them suitable for image-based in-process evaluation of particle images (Galata et al. 2021; Gao et al. 2018; Chen et al. 2019). By using a Mask R-CNN (Regional Convolutional Neural Network, He et al. 2017), real-time image processing was demonstrated to monitor the crystal size and shape during a L-glutamic acid crystallization process (Gao et al. 2018; Chen et al. 2019). However, since the size of the crystals was determined based on the size of the detected objects in the images, this approach is not applicable for astigmatically distorted particle images. Furthermore, the annotation of the training data poses a challenge, which is often performed manually at great expense. Advanced strategies to provide rich semi-synthetic and fully labeled training data enable DNNs to correctly detect even overlapping particle images, thus exceeding the limitations of classical image processing (Franchini and Krevor 2020; Dreisbach et al. 2022). In the work of König et al. (2020), a cascaded DNN performed robustly even when using bimodal particle mixtures for velocity measurements, with deviations of an order of magnitude smaller than those of the classical APTV evaluation. However, in their study, no classification regarding the particle size was done. In this study, a DNN with a Faster R-CNN architecture (Ren et al. 2015) is used, which extends the object detection in the first stage of the cascaded DNN of König et al. (2020) to include a size classification of the found particles. Furthermore, small-scale features of the particle images are utilized by implementing a feature pyramid network (FPN, Lin et al. 2017) architecture. The simultaneous determination of the in-plane coordinates ( $x, y$ ) of suspended particles and assignment of these to different classes of objects according to their size provides a powerful tool for optimizing microfluidic particle separation and mixing processes.

The performance of the proposed algorithms is initially evaluated on two types of data sets using bimodal particle mixtures in Sect. 4.1. On the one hand, experimental images were captured in a laminar fluid flow with spatially separated particle species. On the other hand, detected particle images were cut out of the experimental images and resampled to create semi-synthetic images. In this way, the ground truth is well defined, even for overlapping particle images (Franchini and Krevor 2020). To investigate the potential of the DNN trained on semi-synthetic images for real-world applications, the inference is carried out on experimental images in Sect. 4.2. By combining the images of different bimodal particle mixtures, a size classification into multiple classes is

investigated in Sect. 4.3. Finally, the results are summarized and compared in Sect. 5.

## 2 Experimental setup

To compare the performance of the different algorithms, images of spherical particles with monodisperse diameter distribution of  $(1.14 \pm 0.03) \mu\text{m}$ ,  $(2.47 \pm 0.08) \mu\text{m}$ ,  $(3.16 \pm 0.07) \mu\text{m}$  and  $(5.03 \pm 0.07) \mu\text{m}$  were acquired. The fluorescent polystyrene particles (PS-FluoRed, Ex/Em 530 nm/607 nm, MicroParticles GmbH) were suspended in an aqueous glycerol solution (80/20 w/w deionized water/glycerol) to reduce sedimentation. The fluid was pumped through a microchannel made of polydimethylsiloxane (PDMS) with a rectangular cross section (approx.  $500 \times 185 \mu\text{m}^2$ ) at a constant flow rate of  $5 \mu\text{L}/\text{min}$  using a syringe pump (neMESYS, cetoni GmbH). In order to calculate regression planes of the intensity distribution (see Sect. 3.2) and to provide training data for the DNN (see Sect. 3.3), APTV measurements were first performed with a monomodal particle solution (see Fig. S1 (ESI)). Subsequently, the size recognition was based on measurements using particles of bimodal size distribution in a microchannel with three inlets. The two particle types of different sizes were fed individually through the outer inlets, while pure fluid was fed in the middle inlet to spatially separate both particle types across the channel width. In this way, the ground truth for particle size recognition was determined based on the particle positions along the channel width, while particle images of two different classes were contained in one image capture (see Fig. S2 (ESI)). The ratio between the outer and inner volume flow rates was set to 1.5:2:1.5.

The microchannel was positioned on top of an inverted microscope (Axio Observer 7, Zeiss GmbH) with a plan neofluar objective (M20x, NA = 0.4, Zeiss GmbH). A modulatable OPSL laser (tarm laser technologies tlt GmbH & Co.KG) was used to illuminate the particles through the channel bottom composed of  $128^\circ$  YX LiNbO<sub>3</sub>. This piezoelectric substrate is widely used for particle separation devices based on surface acoustic waves (Wu et al. 2019; Sachs et al. 2022). Since particle size recognition provides essential advantages for the optimization of those devices, the described experiment was designed close to the intended application. Due to the birefringence of the fluorescent light beam caused by the LiNbO<sub>3</sub> crystal, a linear polarization filter was inserted into the optical path of the camera (Kiebert et al. 2017; Sachs et al. 2022). For detection, the reflected laser light was removed from the fluorescent light of the particles by a dichroic mirror (DMLP567T, Thorlabs Inc) and a long-pass filter (FELH0550, 550 nm, Thorlabs Inc). In addition, a cylindrical lens with a focal length of 250 mm was positioned approximately 40 mm in front

of the camera (imager sCMOS, LaVision GmbH, 16 bit,  $2560 \times 2160$  pixel) to introduce astigmatism into the optical system. To cover the entire channel height, the measurement volume was traversed in 8 steps in depth direction with a step size of  $\Delta z = 20 \mu\text{m}$ . At each measurement position, 1000 double frame images were acquired with a frame rate of 10 Hz. For further evaluation, the relative depth position  $z'$  of the particles in the respective measurement volume was used and corrected by the refractive index of the fluid. Additional details about the measurement setup can be found in Sachs et al. (2022).

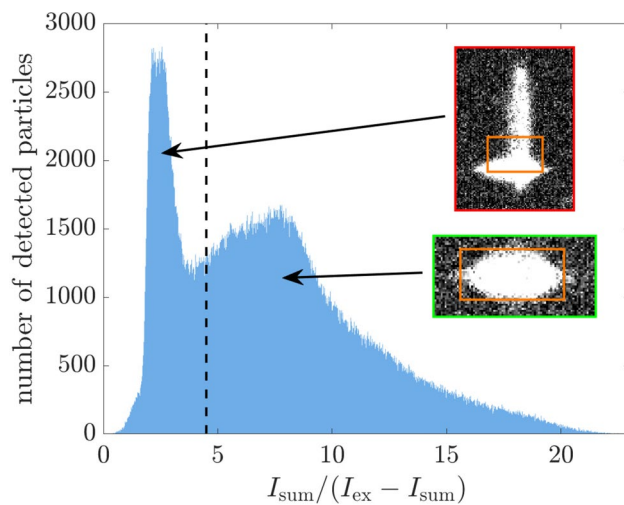
## 3 Methods

### 3.1 Image processing

#### 3.1.1 Classical image evaluation

The acquired images were preprocessed using a background subtraction and smoothed by a Gaussian filter with a kernel size of  $5 \times 5$  pixel. Particle images were then segmented by a global threshold in the intensity profile. The in-plane positions and the lengths of the semi-axes of the elliptical particle images ( $a_x, a_y$ ) were determined by one-dimensional Gaussian fits in the  $x$ - and  $y$ -direction with subpixel accuracy (Cierpka et al. 2012). The depth position of the particles was assigned based on ( $a_x, a_y$ ) using a calibration function determined by calibration measurements with well-known  $z$ -positions (Cierpka et al. 2011).

In order to define the ground truth for the particle detection and size recognition in the experimental images, an outlier detection based on the Euclidean distance  $Z_{\text{err}} = ((a_{x,i} - a_{x,c}(z'_i))^2 + (a_{y,i} - a_{y,c}(z'_i))^2)^{\frac{1}{2}}$  of the found particle images ( $a_{x,i}, a_{y,i}$ ) to the respective calibration function ( $a_{x,c}(z'_i), a_{y,c}(z'_i)$ ) at the assigned  $z'_i$  was used (Barnkob et al. 2021). Here,  $i = 1, 2, \dots, N$  and  $N$  is the amount of detected particles. Particles with  $Z_{\text{err}}$  below a global threshold of 7.5 pixel were initially considered as valid. In a second step, the acceptable limit  $Z_{\text{err,lim}}(z')$  of the Euclidean distance was adaptively determined as a function of the relative depth position  $z'$ , leaving 90 % of the remaining particles images as valid. However, if  $Z_{\text{err,lim}}(z')$  was locally exceeding the arithmetic mean  $\bar{Z}_{\text{err,lim}}$  across all relative depth positions,  $\bar{Z}_{\text{err,lim}}$  was set as the local threshold instead. To further remove incorrectly detected overlapping particle images, the ratio of the summed intensities  $I_{\text{sum}}/(I_{\text{ex}} - I_{\text{sum}})$  was used (see Fig. 1). Here,  $I_{\text{sum}}$  refers to the region within the analyzed ellipsoidal particle image determined by ( $a_x, a_y$ ) and  $I_{\text{ex}}$  to the ellipsoidal region  $\lambda(a_x, a_y)$  extended by a factor of  $\lambda = 1.25$  (Franchini and Krevor 2020). If this ratio exceeded a value of 4.5, the particle image was declared valid.



**Fig. 1** Histogram of detected particles with sizes of 1.14  $\mu\text{m}$  and 2.47  $\mu\text{m}$  as a function of intensity ratio  $I_{\text{sum}}/(I_{\text{ex}} - I_{\text{sum}})$ . The first peak corresponds to incorrectly detected clusters or fragments of particle images, as exemplified in the image framed in red. The orange box indicates the prediction from the classical image evaluation. Particle images with an intensity ratio greater than 4.5 (dashed line) were considered valid, as illustrated by the example in the green framed image

### 3.1.2 Preparation of data sets

The performance of the algorithms was evaluated on two types of data sets. These were either based on experimental images from APTV measurements or semi-synthetic images. The former each contains 16,000 raw images of suspended particles with bimodal size distribution (see Fig. S2 (ESI)), which were further merged to enable a multi-class size recognition (see Sect. 4.3). By adjusting the seeding concentration with respect to the particle size, agglomerations and particle image overlaps were reduced. As a result, the number of particle images contained within a data set decreased with increasing particle size. Since the position and shape of the particle images were unknown a priori, the outcome of the classical image evaluation (see Sect. 3.1.1) was used as ground truth. As a consequence, incorrect labels could not be excluded entirely. Furthermore, overlapping and highly defocused particle images were not considered in the reference labels, since they cannot be detected by the classical segmentation algorithm.

The validated particle images (see Sect. 3.1.1) provided the basis for the preparation of the semi-synthetic images. After a background subtraction, the detected particle images were extracted from the experimentally acquired images along elliptical contours, which were determined by the lengths of the extended semi-axes  $\lambda(a_x, a_y)$ . Here, the expansion factor was constant at  $\lambda = 1.25$ . A total of 100,000 individual particle images were selected for each particle species to generate a balanced data set. This data set was divided

into training, validation and test data in the ratio 70/10/20 %. Randomly selected particle images were then placed on a new synthetic image ( $2160 \times 2560$  pixel) at their original position  $(x_i, y_i)$ . In this way, systematic distortion features in the intensity profile (König et al. 2020) that depend on the in-plane position were retained. For the same reason, conventional image augmentations such as rotation and flipping were avoided. The intensity of overlapping particle images was accumulated. A schematic representation of the image preparation is given in Fig. S3 (ESI). In order to generate a sufficient amount of training data, the individual particle images were arranged multiple times within the respective data set. Overall, the training data comprise at least 558,294 particles in each set.

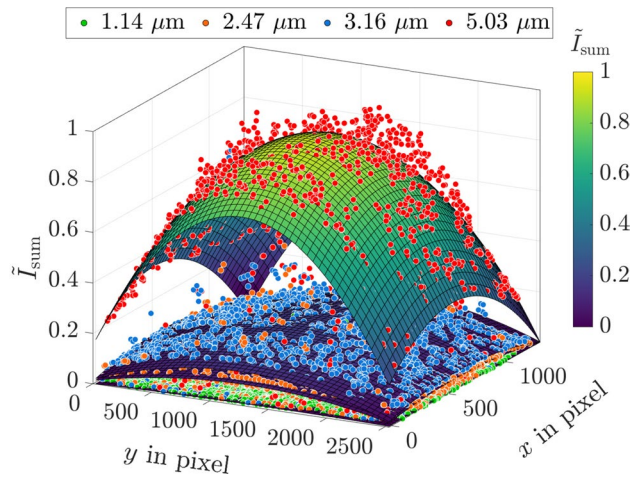
The arbitrary composition of the particle images on each semi-synthetic image can be considered as augmentation (Ghiasi et al. 2021), which reduces the risk of overfitting when using the DNN. The particle image density  $N_s$ , which is described as the ratio of the sum of the particle image area to the full image area (Barnkob and Rossi 2020), was set to 0.02 in accordance with the experimental images. Particle images are overlapping by a mean fraction of 9.9 %. These were labeled within the semi-synthetic images and can potentially be found by the DNN. Finally, Gaussian noise with a mean of 0.5 counts and a standard deviation of 4.5 counts was added to the background of the images. The given ground truth represents a major advantage compared to the experimental images, which contain unlabeled particle images that were not detected in the classical image evaluation.

### 3.2 Deterministic algorithm for particle size recognition

In the proposed algorithm, the influence of particle size on particle image intensity (PII, Hill et al. 2001) is used to implement a deterministic particle size recognition based on APTV measurements. However, the particle image intensity remains a complex quantity, which further depends on the intensity and duration of the illumination, the sensitivity of the camera sensor, the optical system, the experimental setup, the grade of fluorescence and type of dye as well as the position of the particle within the measurement volume. Since the experimental conditions and the measurement system remained unchanged in all experiments, systematic influences on the intensity of the particle images as well as random fluctuations in the fluorescence grade of a particle species are neglected.

However, the spatial position of a particle  $(x_i, y_i, z_i)$  has a significant effect on the intensity of its particle image. Due to aberrations in the optical system and the laser light intensity profile, characteristic intensity distributions are formed in the  $xy$ -plane, as shown in Fig. 2 for different

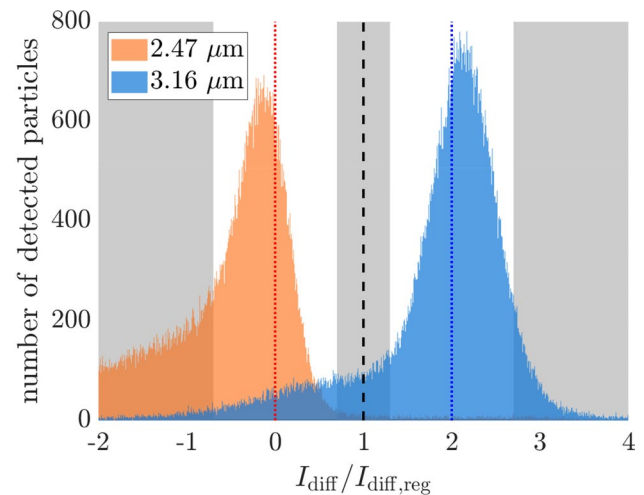




**Fig. 2** Summed intensity from images of spherical polystyrene particles with various sizes along with regression planes at a relative depth position  $z' = 0 \mu\text{m}$ . The intensity was normalized by the maximum intensity  $I_{\text{sum,max}}$  of the  $5.03 \mu\text{m}$  particles to  $\tilde{I}_{\text{sum}} = I_{\text{sum}}/I_{\text{sum,max}}$ . All data points were obtained from APTV measurement of a laminar channel flow

particle species at a relative depth position  $z'_0 = 0 \mu\text{m}$ . Furthermore, the stronger defocusing leads to a monotonically decreasing intensity as the distance from the focal planes ( $z_1, z_2$ ) increases (Brockmann et al. 2020). To account for these influences, 2D regressions of the intensity profiles are approximated by second-order polynomials in the  $xy$ -plane for each particle species at a constant depth position  $z'$  (see Fig. 2). For this, the approximations are based on APTV measurements of a laminar fluid flow using particles with a monomodal size distribution. The step size in  $z'$ -direction for calculating the regression planes amounts to  $\Delta z' = 1 \mu\text{m}$ . The full training pipeline is illustrated schematically in Fig. S3 (ESI). The size recognition of an unknown particle is performed according to the minimum absolute difference between its intensity and the value of the respective regression plane at the particle position  $(x_i, y_i, z'_i)$ . The histogram of the difference in the summed intensity level  $I_{\text{diff}}$  of the detected particle images and one regression plane, normalized by half the intensity difference between both regression planes  $I_{\text{diff,reg}}$ , is shown in Fig. 3 for a bimodal mixture of  $2.47 \mu\text{m}$  and  $3.16 \mu\text{m}$  particles. Particles with  $I_{\text{diff}}/I_{\text{diff,reg}} < 1$  are assigned to the  $2.47 \mu\text{m}$  class, while particles with  $I_{\text{diff}}/I_{\text{diff,reg}} \geq 1$  are classified as particles with a diameter of  $3.16 \mu\text{m}$ .

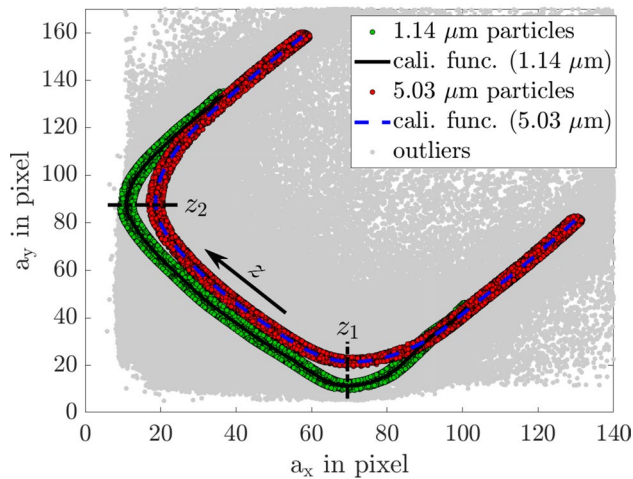
In order to determine the particle image intensity, the intensity values of all pixels within the ellipsoidal area of the particle image are summed up. This approach is more robust against random noise in the intensity level at individual pixel than taking the intensity value at the center of the particle image (Massing et al. 2016). Furthermore, not



**Fig. 3** Histogram of detected  $2.47 \mu\text{m}$  and  $3.16 \mu\text{m}$  sized particles against the normalized difference between the intensity of the particles and the regression plane in the intensity level of the  $2.47 \mu\text{m}$  particles. The regression planes of the particles with sizes of  $2.47 \mu\text{m}$  and  $3.16 \mu\text{m}$  are located at  $I_{\text{diff}}/I_{\text{diff,reg}} = 0$  and  $2$ , respectively. Particles in the gray colored domains are discarded

only the intensity peak but also the particle image size is considered, which proved to be beneficial for the size recognition in comparison to using the peak of a 2D Gaussian fit of the particle image intensity.

The performance of the deterministic algorithm is further improved by applying different outlier criteria. A schematic representation of the inference pipeline is given in Fig. S4 (ESI). Initially, incorrectly detected fragments of defocused or clustered particle images are rejected by a threshold in the intensity ratio  $I_{\text{sum}}/(I_{\text{ex}} - I_{\text{sum}})$  (see Sect. 3.1.1). Moreover, there is a link between the physical particle diameter and the particle diameter in the image (Rossi et al. 2012), causing misclassified particle images to potentially diverge to a greater extent from the respective calibration function in the  $a_x a_y$ -plane. The different calibration functions obtained for particles of  $1.14 \mu\text{m}$  and  $5.03 \mu\text{m}$  in size are illustrated in Fig. 4. The Euclidean distance to the respective calibration function can be used as an independent discriminator for the particle size recognition (Brockmann et al. 2020) yielding decent results in the vicinity of the focal planes, as shown by Sachs et al. (2022). However, the performance drops significantly for particles located outside the focal planes ( $z' < z_1 \vee z' > z_2$ ). To utilize this feature in the proposed algorithm, only particle images with an Euclidean distance  $Z_{\text{err}}$  smaller than the threshold  $\hat{Z}_{\text{err,lim}}$  to the assigned calibration function are considered valid after size recognition based on the particle image intensity (Cierpka et al. 2011). The threshold  $\hat{Z}_{\text{err,lim}}$  was varied in the range of  $[1, 10]$  pixel during optimization (see Sect. 4.1). Furthermore, the distance to the nearest regression plane in the intensity level of



**Fig. 4** Valid 1.14  $\mu\text{m}$  (green, 60.3 %) and 5.03  $\mu\text{m}$  (red, 45.7 %) polystyrene particles with the corresponding calibration functions in the  $a_x a_y$ -plane. The particles are detected in a laminar fluid flow with only one particle species present. Outliers with an Euclidean distance of  $Z_{err} > 2$  pixel to the respective calibration function are depicted in gray for visualization. The positions of the two focal planes are marked as  $z_1$  and  $z_2$ .

the particle images is considered. The latter is investigated by the parameter  $\text{err}_{\max}$  in the range of  $[0.1, 1.1]$  times half the distance between the neighboring regression planes at the respective particle position  $(x_i, y_i, z'_i)$ . Adjusting  $\text{err}_{\max}$  remains a trade-off, since both the number of incorrectly and correctly classified particles decrease with  $\text{err}_{\max}$ . By setting  $\text{err}_{\max}$  for instance to 0.7, particles within the gray domains in Fig. 3 are rejected.

### 3.3 Deep neural network (DNN)

The proposed particle size recognition is a typical classification task for which great progress has been made in recent years by applying DNNs (Russakovsky et al. 2015). These networks use multiple layers of connected artificial neurons to detect image regions in which a desired object is suspected, extract associated features and assign the detected objects based on a given set of classes. The application of DNNs extend into fields of medicine, where degenerative brain diseases can be detected on the basis of neuroimaging data (Zeng et al. 2021), up to robotics, where the fast segmentation of images enables the use of mobile robots (Lewandowski et al. 2019; Seichter et al. 2021). Advanced techniques for using image features at multiple scales allow even small objects to be reliably detected (Lin et al. 2017; Hu et al. 2018; Li et al. 2019; Zeng et al. 2022). This opens the way for the use in industrial applications, such as surface defect detection of printed circuit board (Zeng et al. 2022). The use of DNNs to detect the three-dimensional particle positions from raw images obtained in APTV measurements

has been successfully demonstrated by König et al. (2020). Although the achievable accuracy in position determination remained lower for that specific DNN compared to classical image processing algorithms for particle detection using monomodal particle solutions (Barnkob et al. 2021), the application of DNNs showed promising results when using bimodal particle mixtures (König et al. 2020).

The DNN used in this work combines the in-plane position determination of suspended particles with the classification of their size in multimodal particle solutions. For this, a network with a Faster R-CNN architecture (Ren et al. 2015) and a ResNet-50 (He et al. 2016) as backbone for object detection is trained in PyTorch version 1.11.0. The DNN incorporates a feature pyramid network (FPN) architecture, which was first introduced by Lin et al. (2017) to enable an inverse information transfer between later and earlier layers of the DNN. By using the FPN, the strong semantic information in final layers with low resolution is combined with the high resolution in initial layers, which lead to rich semantics at all levels. In this way, small-scale features in the intensity distribution of the particle images are utilized to improve the precision of predictions (Lin et al. 2017; Dreisbach et al. 2022). When using the experimental data set, raw images from flow measurements in the microchannel with only one particle species are employed during training (see Fig. S3 (ESI)). The results presented are based on the inference on experimental images of bimodal particle solutions. In contrast, the semi-synthetic images are split into 70 % training, 10 % validation and 20 % test data. The learning rate of the pre-trained DNN is scheduled to linearly increase over one epoch from  $10^{-5}$  to a maximum of  $10^{-4}$  and successively decrease by a factor of 10 at the beginning of each of the last two epochs. The network is trained for 5 epochs with a batch size of 16, and the AdamW optimizer (Loshchilov and Hutter 2019) with a weight decay of  $10^{-4}$  as regularization. Due to a large training set and short training duration, no additional data augmentation is used. The images are transmitted to the DNN with a bit depth of 16 bit to accurately resolve gradients in the intensity distribution for all particle species involved while avoiding clipping of high intensities due to large particles or overlapping particle images. The different particle species are labeled as classes of objects according to their size. During the training process, regions which contain suspected particles are identified by bounding boxes enclosing the width and height of the particle images. The in-plane positions of the particles are determined based on the center of the bounding boxes. Correctly matched particles are required to have an Euclidean distance of less than 5 pixel to the corresponding label in the respective ground truth. Once the DNN has been trained on an NVIDIA A100 tensor-core GPU at the high-performance computing (HPC) cluster of TU Ilmenau, combinations of the acquisitions of bimodal particle flows serve to compare the size classification with

the presented deterministic approaches during the inference (see Fig. S4 (ESI)). Here, particle labels predicted by the DNN are validated by a threshold in the classification score, which gives a measure for the confidence of the prediction.

## 4 Results and discussion

### 4.1 Particle size recognition in bimodal particle mixtures

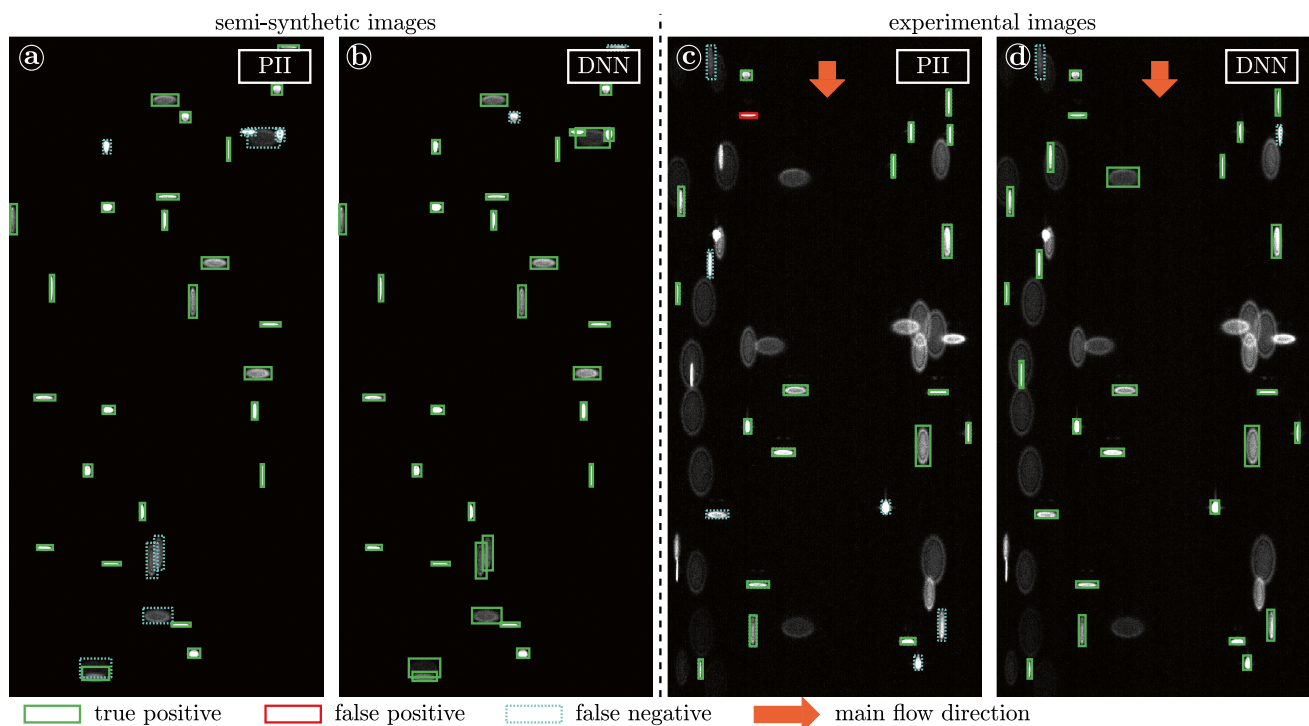
The performance of the proposed algorithms for particle size recognition is initially investigated on the basis of semi-synthetic and experimental images from APTV measurements with particle solutions of bimodal size distribution. For this, spherical polystyrene particles with a diameter of  $2.47\text{ }\mu\text{m}$  are to be distinguished from  $1.14\text{ }\mu\text{m}$ ,  $3.16\text{ }\mu\text{m}$  or  $5.03\text{ }\mu\text{m}$  particles. Representative samples of semi-synthetic images containing  $2.47\text{ }\mu\text{m}$  and  $3.16\text{ }\mu\text{m}$  particles are given in Fig. 5a,b, while related experimental images are shown in Fig. 5c,d. In contrast to the experimental images, Fig. 5a,b do not include strongly defocused particle images in the background. These reduce the signal-to-noise ratio (SNR) of brighter particle images in the foreground, and thus complicate their detection. Predictions of classified particles given by the deterministic algorithm are

depicted in Fig. 5a,c, while particles suggested by the DNN are labeled in Fig. 5b,d. Particles correctly classified by the algorithms are declared as true positives (TP, green), incorrectly classified particles as false positives (FP, red), and non-detected particles are indicated as false negatives (FN, cyan), if they are considered as valid in the ground truth. Based on this sampling evaluation, a reliable size classification by the DNN is evident, while some undetected or misclassified particles are present in the predictions of the deterministic algorithm. The DNN is able to correctly detect and classify overlapping particle images in the semi-synthetic images, which is a significant improvement compared to the deterministic algorithm. In the experimental images, overlaps are not labeled within the ground truth, which causes them to remain unlabeled by the DNN during the training process.

To enable a quantitative comparison of the algorithms, the following metrics are determined.

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (1)$$

$$\text{recall} = \frac{\text{TP}_{\text{ref}}}{\text{TP}_{\text{ref}} + \text{FN}} \quad (2)$$



**Fig. 5** Elliptically distorted particle images with labels given by the particle image intensity (PII, **a**, **c**) and the deep neural network (DNN, **b**, **d**). Semi-synthetic images are shown in (**a**, **b**), while experimental images are present in (**c**, **d**). Spherical particles with  $2.47\text{ }\mu\text{m}$  and

$3.16\text{ }\mu\text{m}$  in size are used in a bimodal mixture, which are hydrodynamically pre-positioned to the left and right side of the microchannel in (**c**, **d**), respectively. Unlabeled particle images are not part of the ground truth



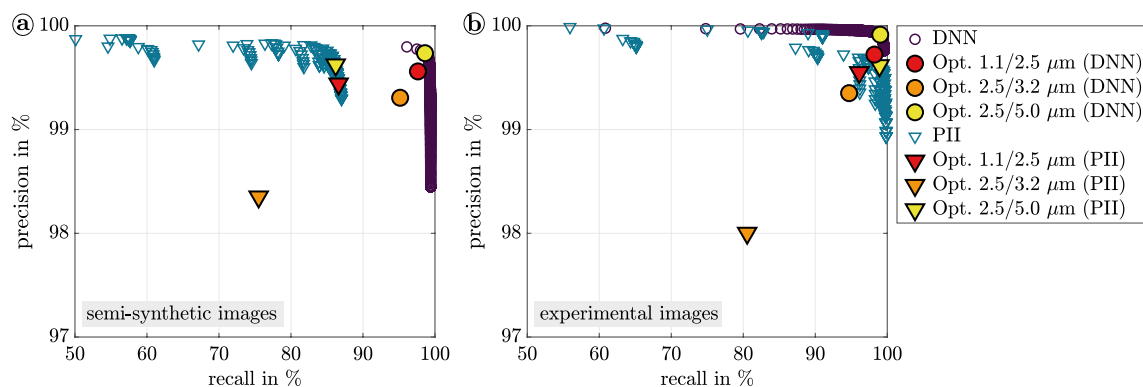
The precision serves as a measure for the quality of the classification and considers all predictions given by the respective algorithm. In contrast, the recall is based on the reference labels, which are validated with knowledge of the ground truth (see Sect. 3.1.1), and quantifies the proportion of particles that are correctly assigned ( $TP_{ref}$ ) to the total number of valid particles.

In Fig. 6, the achieved precision is plotted against the recall for the utilized algorithms under variation of the accepted threshold values in the outlier criteria (see Sect. 3) for a particle mixture of 2.47  $\mu\text{m}$  and 5.03  $\mu\text{m}$  particles. For both algorithms, a characteristically increasing precision is revealed as the recall decreases. This trend is caused by a more rigorous validation, which leads to less particles being declared as valid and to a decrease in the recall. Simultaneously, more reliable predictions are kept and the precision increases. To identify the best combination of thresholds in the outlier criteria within the grid search, the minimum of the weighted Euclidean distance  $d = [(1 - \text{recall})^2 + \xi(1 - \text{precision})^2]^{\frac{1}{2}}$  to the optimal case with precision and recall equal one is calculated. In order to prioritize a reliable size classification higher than the yield, the precision is weighted higher by the coefficient  $\xi = 100$ .

The achievable precision following this optimization is plotted in Fig. 6a for all combinations of particle species using semi-synthetic images. Two trends emerge from the diagram: First, both the recall and the precision for a given particle combination are consistently highest using the DNN. Second, a distinction between 2.47  $\mu\text{m}$  and 5.03  $\mu\text{m}$  particles is more reliable than between 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  particles, as indicated by the precision. This is caused by less distinctive features regarding the intensity distribution of the particle images corresponding to the latter combination of particle sizes. However, a precision above 98 % is

reached consistently. Greater limitations exist with regard to the recall for size detection according to PII, while values above 95 % were always attained with the DNN. One major reason for this is found in overlapping particle images, which are captured in the ground truth but remain undetected by the deterministic algorithm. Since the labels in the experimental images are based on the classical image evaluation, the achievable recall based on PII increases for all combinations of particle species (see Fig. 6b). It applies that the higher the precision obtained for semi-synthetic images, the more the recall increases when using experimental images. This is determined by the more rigorous validation of hard-to-distinguish particle images based on the experimental images, which becomes necessary as a result of the strongly defocused particle images in the background. Concurrently, the precision achieved with the DNN increases slightly, while a drop of about 0.35 % is observed in the distinction between 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  sized particles using the PII. The results prove that the DNN performs more robust than the deterministic algorithm against defocused particle images in the background.

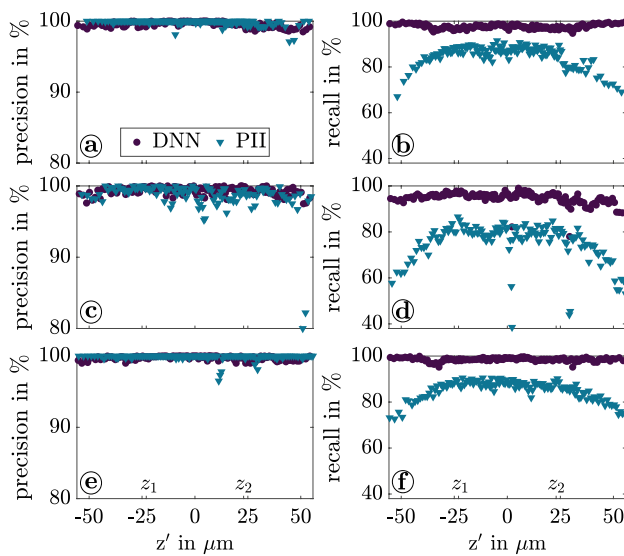
The achievable precision and recall are evaluated as a function of the relative depth position  $z'$  in Fig. 7 using semi-synthetic images. Both algorithms yield a very uniform precision with minute variations when applied with optimized settings. A reliable size classification of bimodal particle solutions is thus demonstrated for the entire depth range studied. The proportion of correctly classified particles among the ensemble of valid particles is reflected by the recall, which is in contrast always lower when using the deterministic algorithm. For the latter, a decrease is especially noticeable for particle images obtained from particles outside the focal planes, e.g.  $z' < z_1 \vee z' > z_2$ . In these regions, a greater portion of particle images is filtered by the outlier criteria in order to maintain accurate



**Fig. 6** Precision as a function of recall achieved for size recognition of 2.47  $\mu\text{m}$  and 5.03  $\mu\text{m}$  particles with DNN (dots) and PII (triangle) under variation of the threshold in the outlier criteria. The results are obtained by using either semi-synthetic images (a) or experimental

images (b). Optimized results are given for different bimodal combinations of particle species. The legend on the right hand side applies to both subplots

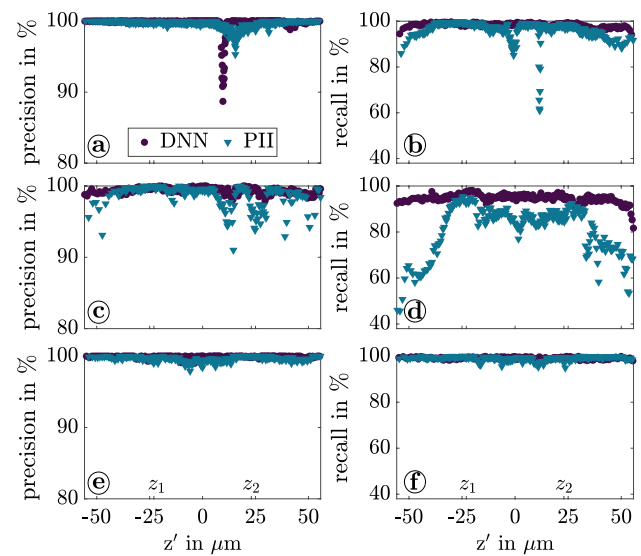




**Fig. 7** Results obtained using semi-synthetic images. Precision and recall as a function of the relative depth position  $z'$  for size classification of bimodal particle solutions of sizes 1.14  $\mu\text{m}$  and 2.47  $\mu\text{m}$  (a, b), 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  (c, d), as well as 2.47  $\mu\text{m}$  and 5.03  $\mu\text{m}$  (e, f). Please note that the legend in (a) applies to all subplots

size recognition. Overall, a significantly enhanced performance was achieved using the DNN with respect to recall. Across all combinations of particle species examined, a precision above 99 % is obtained for 85.68 % of the depth range, while maintaining a recall above 98 % for 84.55 % of the depth range when applying the DNN. Thus, only 2156 of the 465814 predicted particles are incorrectly classified in total with this algorithm.

In Fig. 8, precision and recall for the size recognition of different bimodal mixtures based on experimental images are depicted. The strongly defocused particle images in the background of the experimental images significantly affect the size classification and can lead to local drops in precision. These drops are particularly evident in the distinction between 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  sized particles using PII, since differences in the intensity of the particle images of both species are smallest. Considerable variations in the precision are evident across  $z'$  in Fig. 8c. Due to the close intensity levels, a more rigorous validation with respect to  $Z_{\text{err}}$  was chosen, which is reflected in a decreasing recall with increasing distance to the focal planes ( $z_1, z_2$ ) (see Fig. 8d). Furthermore, a drop in precision is apparent within a narrow band when discriminating between 1.14  $\mu\text{m}$  and 2.47  $\mu\text{m}$  sized particles using the DNN (see Fig. 8a). This reduced precision can be attributed to a single particle sticking to the channel floor, which is consistently misclassified by the DNN. Since the associated ground truth on the particle species is determined based on the position along the channel width, the sticking



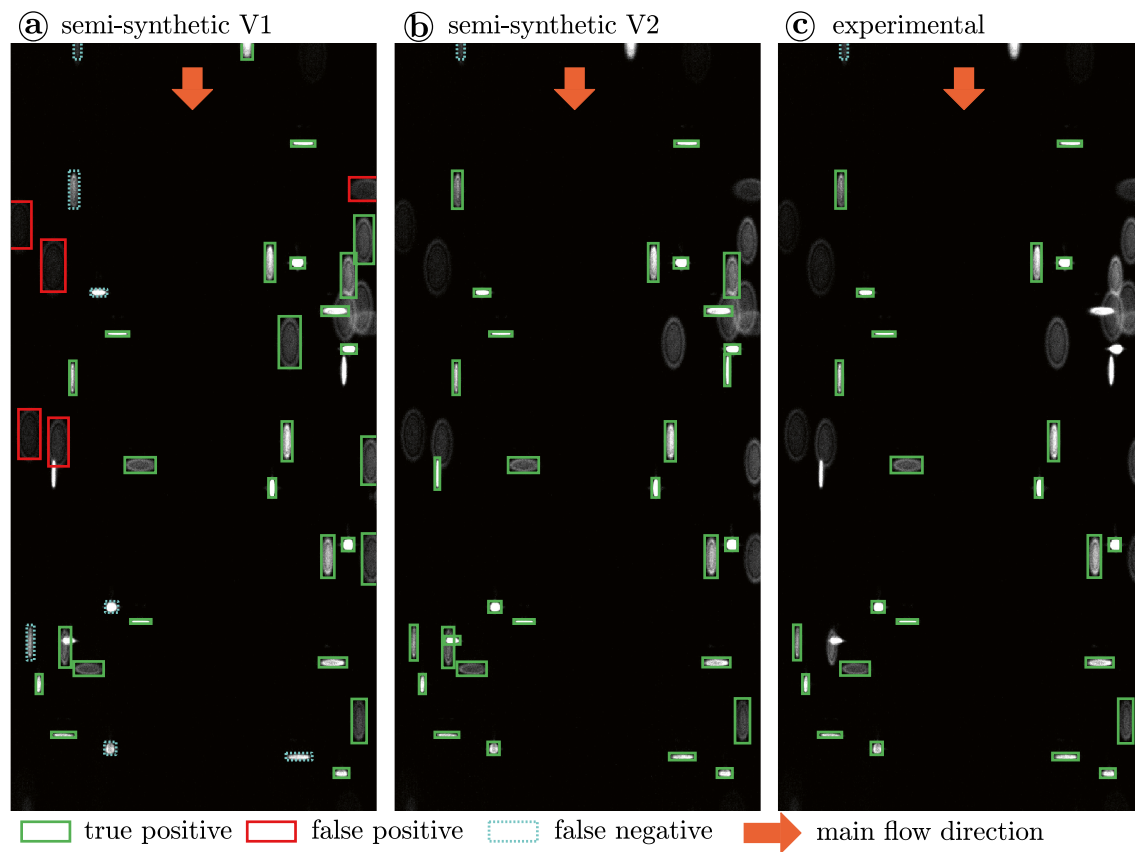
**Fig. 8** Particle size recognition based on experimental images. Precision and recall against the relative depth position  $z'$  obtained in the distinction between 1.14  $\mu\text{m}$  and 2.47  $\mu\text{m}$  (a, b), 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  (c, d), as well as 2.47  $\mu\text{m}$  and 5.03  $\mu\text{m}$  (e, f) sized particles. Please note that the legend in (a) applies to all subplots

particle is labeled to have a diameter of 2.47  $\mu\text{m}$ . However, an incorrect label cannot be discounted confidently when considering the intensity of the particle image.

## 4.2 Application to different data sets

Training the DNN on fully annotated semi-synthetic images enables the detection and classification of overlapping particle images, as shown in Sect. 4.1. Applied experimentally, this provides the potential to use a higher particle seeding concentration, which could yield more Lagrangian particle tracks and a finer resolution of the velocity field at constant measurement effort and time (Kähler et al. 2012). To utilize this improvement in the particle size recognition, the DNN needs to be trained on semi-synthetic images and tested on experimental images during the inference phase. Since both types of data sets were sampled from similar distributions, a practical conversion is conceivable and will be explored below. The particle size recognition of 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  sized particles is used to evaluate the performance of the DNN, which is found to be the most difficult bimodal size distinction in Sect. 4.1. The inference is consistently based on background subtracted experimental images, while the training is conducted on two modifications of semi-synthetic images. Examples of the modified semi-synthetic images can be found in Fig. S5 (ESI).

The first modification (V1) is similar to the semi-synthetic images introduced in Sect. 3.1.2. These images contain only annotated particle images extracted from experimental data sets. Exemplary predictions by the DNN are shown in Fig. 9



**Fig. 9** Exemplary predictions given by the DNN in the size recognition of 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  particles using experimental images during inference. The training of the DNN was performed on semi-

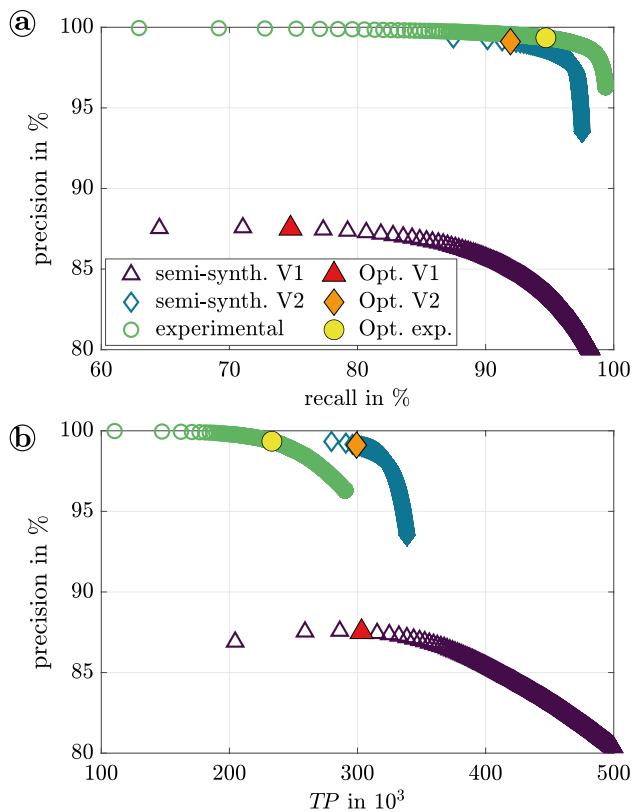
synthetic images of modifications V1 (a) or V2 (b) as well as on experimental images (c). Unlabeled particle images are not present in the ground truth and correspond to the background

for the inference on experimental images. After successfully training the DNN on semi-synthetic images of modification V1 (Fig. 9a), a higher amount of particle images are correctly detected and classified (TP) as compared to training on experimental images (Fig. 9c). However, the number of incorrectly classified particles (FP) increases as well, which are mainly related to strongly defocused particle images and particles at the edge of the microchannel or image. These are rarely detected by the classical image evaluation and thus underrepresented in the semi-synthetic training images. In order to optimize the performance of the DNN according to the weighted Euclidean distance  $d$  (see Sect. 4.1), the threshold in the classification score is set to be higher than when trained on experimental images. Consequently, a larger number of correct predictions is rejected as well, which increases the amount of false negatives (FN). The reduced performance of the DNN after training on semi-synthetic images of modification V1 is thus caused by a lack of similarity between the training and test images.

In order to overcome this lack of similarity, the semi-synthetic images were further refined, creating a data set referred to as version V2. The modifications comprise the

implementation of strongly defocused and near-edge particle images without annotation in the background of the semi-synthetic images. Since these particle images were not labeled by the classical image evaluation, they are extracted manually from the experimental training images. Providing accurate annotations could further enhance the performance of the DNN and the detectable depth range; however, this is not feasible in the manual extraction of the particle images. The complexity of object detection increased significantly due to the superposition of labeled overlapping particle images with strongly defocused particle images in the background. As a result, an extended training duration of ten epochs proved to be beneficial. With respect to Fig. 9b, the number of FPs could be reduced. Furthermore, fewer strongly defocused particle images are detected, since these are now included unannotated in the training. Simultaneously, predictions for particles closer to the focal planes are given with higher confidence and are not rejected as outliers.

A quantitative comparison of the results obtained with both versions of semi-synthetic images is given in Fig. 10a and (b) based on the precision as a function of recall and number of TPs, respectively. For an optimized threshold in



**Fig. 10** Results obtained for the inference on experimental images and training on either semi-synthetic images in different modifications (V1, V2) or experimental images. The precision is depicted as a function of the recall (**a**) and the number of correctly classified particles (TP, **b**). The results with an optimized threshold in the classification score are highlighted by filled markers. Here, 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  particles are distinguished in a bimodal particle solution. Please note that the legend in (**a**) applies to both subplots

the classification score, both precision and recall increase due to the modifications by about 11.59 % and 17.17 %, respectively. Concurrently, the number of TPs decreases by only 1.29 %, indicating a reliable detection and classification of overlapping particle images. This is further reflected by the steeper slope of the precision when using version V2 in Fig. 10b, which indicates that most of the incorrect predictions show a lower classification score.

In comparison to training on experimental images, a marginal decrease of 0.24 % in precision is observed, which can be attributed to the much more complex object detection of overlapping particle images. However, the increase of 0.92 % in precision among the detected particle images that are not annotated in the ground truth underlines the benefit of training on semi-synthetic images. A larger discrepancy is seen in recall, which dropped by 2.76 %. The calculation of the recall according to equation 2 is based on the ground truth and disregards the 85,389 additionally found TPs. Overall, the number of correctly classified particle images increased

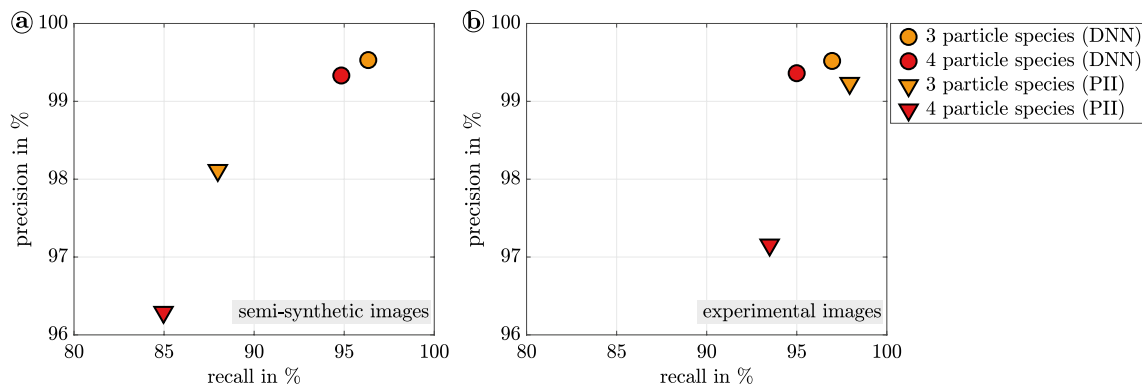
by 28.34 % compared to training on experimental images. This is a significant improvement in performance, while less individual particle images are required in training and the experimental effort can be reduced. The results highlight the relevance of a high similarity between the training and test data sets in order to utilize the potential of a improved detection rate when training the DNN on semi-synthetic images.

### 4.3 Multi-class size recognition

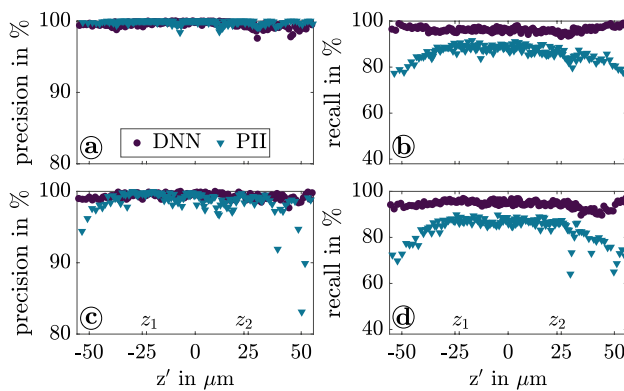
After applying and optimizing the particle size recognition on experimental and semi-synthetic images of bimodal particle solutions, the distinction between multiple particle species is analyzed in the following. For this, the existing experimental images each containing two different particle species are concatenated into new experimental data sets without manipulating the individual images. In this way, experimental data sets are generated that include multiple particle species, while only two particle species are present on each image. Furthermore, semi-synthetic images of modification V1 are used in which all involved particle species are present on each image. In the different data sets either particles with 1.14  $\mu\text{m}$ , 2.47  $\mu\text{m}$  and 5.03  $\mu\text{m}$  or 1.14  $\mu\text{m}$ , 2.47  $\mu\text{m}$ , 3.16  $\mu\text{m}$  and 5.03  $\mu\text{m}$  in diameter are investigated as three- and four-class distinction problems, respectively. The outlier criteria optimized in Sect. 4.1 are adaptively applied within the deterministic algorithms depending on the predicted particle class. For size recognition by the DNN, the outlier detection is performed using a constant threshold in the classification score, which is optimized for each combination of particle species according to the minimum of the criterion  $d$  (see Sect. 4.1).

The achieved precision in particle size recognition is plotted in Fig. 11 for the proposed algorithms as a function of recall. With the extension to three and four classes of particles, the precision does not significantly decrease with values above 96 % or 97 % using semi-synthetic or experimental images, respectively. For both algorithms, the achievable precision and recall in the classification of three particle species are consistently higher than the results of the four-class discrimination. This trend is expected to persist for an increasing number of classes within the similar size range.

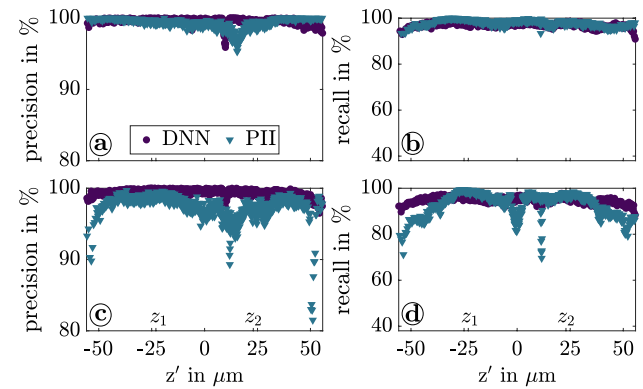
Furthermore, only a marginal decrease in precision is revealed when using PII in the three-class problem compared to the distinction between two of the involved particle species. The reason for this is found in the significant deviation between the intensity levels of the different particle images, which facilitates the size classification. However, a more challenging particle size recognition is constituted by the four-class discrimination, in which 3.16  $\mu\text{m}$  sized particles were complemented to the data set. Herein, the minimum mean deviation between adjacent regression



**Fig. 11** Precision as a function of recall obtained in the classification of multiple particle species by using DNN (dots) and PII (triangles). The results are achieved based on semi-synthetic (a) or experimental images (b)



**Fig. 12** Precision and recall as a function of the relative depth position  $z'$  for the classification of 1.14  $\mu\text{m}$ , 2.47  $\mu\text{m}$  and 5.03  $\mu\text{m}$  (a, b) as well as 1.14  $\mu\text{m}$ , 2.47  $\mu\text{m}$ , 3.16  $\mu\text{m}$  and 5.03  $\mu\text{m}$  sized particles (c, d) by using DNN and PII. Here, semi-synthetic images are used. The legend in (a) applies to all subplots



**Fig. 13** Precision and recall over the relative depth position  $z'$  for the classification of particles with sizes of 1.14  $\mu\text{m}$ , 2.47  $\mu\text{m}$  and 5.03  $\mu\text{m}$  (a, b) as well as 1.14  $\mu\text{m}$ , 2.47  $\mu\text{m}$ , 3.16  $\mu\text{m}$  and 5.03  $\mu\text{m}$  (c, d). The results are obtained by using DNN and PII on experimental images. Please note the legend in (a), which applies to all subplots

planes in the intensity level is reduced by about 30 %, which leads to a reduced precision consistent with the results of the bimodal size classification. Compared to the application to three particle species, the precision decreases by approx. 1.82 % or 2.07 % and the recall by about 3.01 % or 4.45 % when using semi-synthetic or experimental images, respectively. Employing the DNN, the drop in precision and recall observed between the three- and four-class problem is considerably reduced to about 0.2 % and 1.75 %, respectively. In conjunction with the excellent precision clustering around 99.44 %, the results indicate a robust size classification even with four particle species.

The analysis of precision and recall as a function of relative depth position  $z'$  is shown in Figs. 12 and 13 using semi-synthetic and experimental images, respectively. The obtained results for the particle size recognition of multiple species are somewhat similar to those achieved for bimodal particle mixtures. Using either algorithm, very uniform

profiles of precision with values higher than 98 % in about 89 % of the depth range were reached consistently in the semi-synthetic images. However, local drops are revealed in the discrimination of four particle species by the PII using experimental images (see Fig. 13c). These are caused by the superposition of the inaccuracies in the classification of the bimodal size-distributed particles, since this data set is based on combinations of the images used in Sect. 4.1. For the same reason, the local drop at  $z' \approx 7 \mu\text{m}$  in the precision obtained by the DNN in Fig. 13a reappears, which is attributed to an adherent particle in the images of 1.14  $\mu\text{m}$  and 2.47  $\mu\text{m}$  sized particles (see Sect. 4.1).

In Fig. 12, a decrease in recall for defocused particle images ( $z' < z_1 \vee z' > z_2$ ) is evident when employing the deterministic algorithm. A similar trend is further present in Fig. 13d due to the more rigorous validation of the involved 3.16  $\mu\text{m}$  particles according to  $Z_{\text{err}}$ . In these domains, a higher amount of particles are correctly



classified by the DNN, circumventing limitations regarding the defocusing of particle images in the observed depth range. Furthermore, detected overlapping particle images lead to a consistently higher recall by the DNN in the semi-synthetic images.

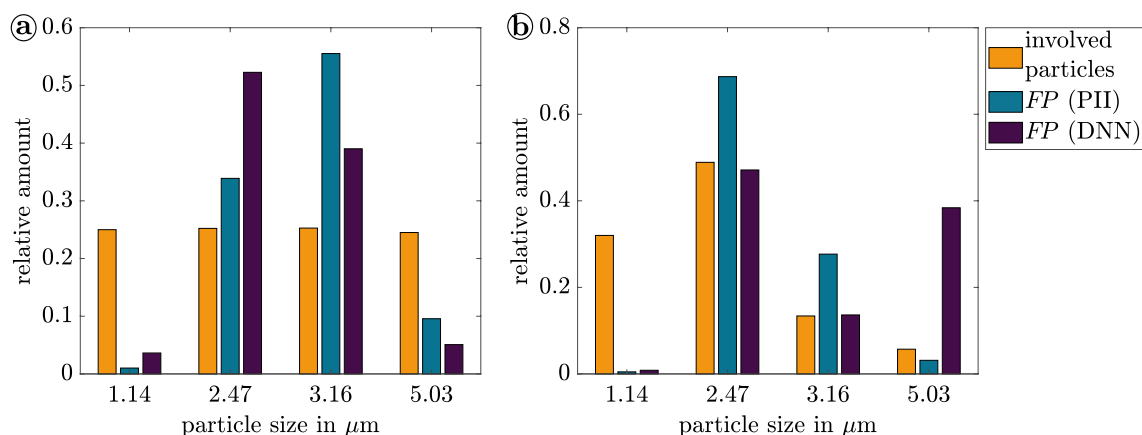
The results demonstrate the applicability of the proposed algorithms for multi-class size recognition with a very stable performance of the DNN. For a deeper analysis of the misclassifications in the discrimination between four particle species, the relative number of involved and incorrectly classified particles is illustrated in Fig. 14. In the semi-synthetic images (Fig. 14a), an even distribution across all particle species was set. The highest frequency of misclassifications occur due to 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  sized particles using either algorithm. This is to be expected since both the intensity of the particle images and features in the intensity distribution are closest in similarity between these particle species. In contrast, distinctive features are more pronounced in the images of particles with sizes of 1.14  $\mu\text{m}$  and 5.03  $\mu\text{m}$ , which facilitates their classification. The experimental data set contains a constant number of images for each bimodal particle combination investigated, thus resulting in a varying relative number of valid particles for each particle species (see Fig. 14b). The smaller the particles, the higher the chosen seeding concentration, which means that large particles are underrepresented in this data set. An exception to this trend is the proportion of 2.47  $\mu\text{m}$  particles of about 49 % of the total number of valid particles, since these occur in every bimodal particle mix. The underrepresentation of 5.03  $\mu\text{m}$  particles applies also to the training data, which is reflected in a significantly increased proportion of misclassifications by the DNN. Simultaneously, the deterministic algorithm behaves robust against the imbalance of the data set. Resembling the results with semi-synthetic images,

the highest relative amount of misclassifications occurs for 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  sized particles.

## 5 Conclusion

In this study, the particle size recognition of suspended tracer particles is investigated by implementing a deterministic algorithm and a deep neural network (DNN). The performance of both algorithms is compared based on experimental and semi-synthetic images from astigmatism particle tracking velocimetry (APTV) measurements in a microchannel. By utilizing differences in the particle image intensity (PII), a discrimination of spherical 2.47  $\mu\text{m}$  and 3.16  $\mu\text{m}$  sized particles is achieved with a precision above 98.0 % by the deterministic algorithm. After training the DNN for object detection and particle size classification, an even higher precision beyond 99.3 % is achieved in all examined bimodal particle solutions. Concurrently, the recall remains consistently above 92 % across approx. 90.1 % of the investigated depth range, which further demonstrates the applicability of the DNN for particle size recognition of highly defocused particle images. By training the DNN on modified semi-synthetic images, a reliable detection and size classification of overlapping particle images is achieved even in experimental images. The increased detection rate represents a significant improvement in the particle size recognition for real-world applications.

Once various outlier criteria are optimized based on the discrimination of bimodal particle mixtures, a size recognition of multimodal distributed particles is performed. Fluorescent polystyrene particles in the range of 1.14  $\mu\text{m}$  to 5.03  $\mu\text{m}$  are distinguished into four classes with a precision of up to 99.3 %, while a recall of 94.8 % is maintained. The



**Fig. 14** Relative distribution of the number of valid particles (involved particles) and relative frequency of incorrectly classified particles (FP) by the proposed algorithms (PII, DNN) versus particle

size. Illustrated are the results of the particle size recognition of four particle species, which is based on semi-synthetic (a) or experimental images (b)

obtained results demonstrate the capability of a reliable particle size recognition using a standard APTV setup for microfluidics. The introduced algorithms allow a low cost extension of APTV to particle solutions with multimodal size distribution, which does not require any adaptation of the existing measurement equipment or the use of different fluorescent dyes for each particle species. Furthermore, the training of the DNN is based on flow measurements, which can be integrated into the measurement procedure with reasonable effort and does not require a rather time consuming acquisition of training data using sedimented particles as usual. Variations in the optical setup, such as the magnification, might affect the appearance of features in the particle images and the performance of the algorithms, which requires further investigation in upcoming studies. Meeting the requirement of high similarity between training and test data sets, by using e.g. modified semi-synthetic images, is an important aspect of ongoing research that opens new prospects to improve the applied measurement technique. For this, an assessment of the positional uncertainty of the predictions by the introduced DNN based on synthetic or semi-synthetic data sets with rich ground truth is further needed. The application of the algorithms for label-free identification of targeted particle species paves the way for future usage as triggers for online on-demand particle separation devices (Ma et al. 2017; Zhang et al. 2020; Bakhtiari and Kähler 2022).

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00348-023-03574-2>.

**Acknowledgements** The authors thank the German Research Foundation (DFG) for financial support within the priority program PP2045 “MehrDimPart” (CI 185/8-1). Support by the Center of Micro- and Nanotechnologies (ZMN) (DFG RI-sources reference: RI\_00009), a DFG-funded core facility of TU Ilmenau, and by the project no. P 2018-02-001 “DeepTurb - Deep Learning in and of Turbulence”, which is funded by the Carl Zeiss Foundation, is gratefully acknowledged. Furthermore, the authors are grateful to Henning Schwanbeck for his technical support using the GPU computing cluster at the Universitätsrechenzentrum of TU Ilmenau.

**Author contributions** SS contributed to conceptualization, formal analysis, investigation, methodology, software, validation, visualization, writing—original draft, and writing—review and editing. MR contributed to methodology, formal analysis, investigation, software, validation, and writing—review and editing. PM contributed to supervision, resources, and writing—review and editing. JK contributed to conceptualization, project administration, resources, supervision, and writing—review and editing. CC contributed to conceptualization, project administration, resources, supervision, and writing—review and editing.

**Funding** Open Access funding enabled and organized by Projekt DEAL. This study was funded by the German Research Foundation (DFG) within the priority program PP2045 “MehrDimPart” under Grant no. CI 185/8-1 and by the Carl Zeiss Foundation under project no. P 2018-02-001 “DeepTurb - Deep Learning in and of Turbulence”.

**Data availability** The full data sets that support the findings of this study are available from the corresponding author, S.S., upon reasonable request.

## Declarations

**Conflict of interest** There are no conflicts of interest to declare.

**Ethical approval** Not applicable.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Ahmad M, Bozkurt A, Farhanieh O (2019) Evaluation of acoustic-based particle separation methods. *World J Eng* 16(6):823–838. <https://doi.org/10.1108/WJE-06-2019-0167>
- Ahmed H, Destgeer G, Park J, Afzal M, Sung HJ (2018) Sheathless Focusing and Separation of Microparticles Using Tilted-Angle Traveling Surface Acoustic Waves. *Analytical chemistry* 90(14):8546–8552. <https://doi.org/10.1021/acs.analchem.8b01593>
- Bakhtiari A, Kähler CJ (2022) Automated monitoring and positioning of single microparticle via ultrasound-driven microbubble streaming. *Microfluid Nanofluid* 26(8):59. <https://doi.org/10.1007/s10404-022-02566-8>
- Barnkob R, Rossi M (2020) General defocusing particle tracking: fundamentals and uncertainty assessment. *Exp Fluids* 61(4):1690. <https://doi.org/10.1007/s00348-020-2937-5>
- Barnkob R, Cierpka C, Chen M, Sachs S, Mäder P, Rossi M (2021) Defocus particle tracking: a comparison of methods based on model functions, cross-correlation, and neural networks. *Meas Sci Technol* 32(9):094011. <https://doi.org/10.1088/1361-6501/abfeb6>
- Brockmann P, Kazerooni HT, Brandt L, Hussong J (2020) Utilizing the ball lens effect for astigmatism particle tracking velocimetry. *Exp Fluids* 61(2):67. <https://doi.org/10.1007/s00348-020-2900-5>
- Caputo F, Clogston J, Calzolari L, Rösslein M, Prina-Mello A (2019) Measuring particle size distribution of nanoparticle enabled medicinal products, the joint view of EUNCL and NCI-NCL. A step by step approach combining orthogonal measurements with increasing complexity. *J Control Release* 299:31–43. <https://doi.org/10.1016/j.jconrel.2019.02.030>
- Chen S, Liu T, Xu D, Huo Y, Yang Y (2019) Image based Measurement of Population Growth Rate for L-glutamic acid crystallization, In: *Proceedings of the 38th Chinese Control Conference* <https://doi.org/10.23919/ChiCC.2019.8866441>
- Cierpka C, Rossi M, Segura R, Mastrangelo F, Kähler CJ (2012) A comparative analysis of the uncertainty of astigmatism- $\mu$ PTV, stereo- $\mu$ PIV, and  $\mu$ PIV. *Exp Fluids* 52(3):605–615. <https://doi.org/10.1007/s00348-011-1075-5>
- Cierpka C, Rossi M, Segura R, Kähler CJ (2011) On the calibration of astigmatism particle tracking velocimetry for microflows.

- Measurement Science and Technology 22(1):015–401. <https://doi.org/10.1088/0957-0233/22/1/015401>
- Cierpka C, Segura R, Hain R, Kähler CJ (2010) A simple single camera 3C3D velocity measurement technique without errors due to depth of correlation and spatial averaging for microfluidics. *Measurement Science and Technology* 21(4):045–401. <https://doi.org/10.1088/0957-0233/21/4/045401>
- Dannhauser D, Romeo G, Causa F, de Santo I, Netti PA (2014) Multiplex single particle analysis in microfluidics. *Analyst* 139(20):5239–5246. <https://doi.org/10.1039/c4an01033g>
- de Bruijn DS, Jorissen KFA, Olthuis W, van den Berg A (2021) Determining particle size and position in a coplanar electrode setup using measured opacity for microfluidic cytometry. *Biosensors* 11(10):353. <https://doi.org/10.3390/bios11100353>
- de Pastina A, Villanueva LG (2020) Suspended micro/nano channel resonators: a review. *J Micromech Microeng* 30(4):043. <https://doi.org/10.1088/1361-6439/ab6df1>
- Dreisbach M, Leister R, Probst M, Friederich P, Stroh A, Kriegseis J (2022) Particle detection by means of neural networks and synthetic training data refinement in defocusing particle tracking velocimetry. *Meas Sci Technol* 33(12):124001. <https://doi.org/10.1088/1361-6501/ac8a09>
- Farkas D, Madarász L, Nagy ZK, Antal I, Kállai-Szabó N (2021) Image analysis: a versatile tool in the manufacturing and quality control of pharmaceutical dosage forms. *Pharmaceutics* 13(5):685. <https://doi.org/10.3390/pharmaceutics13050685>
- Franchini S, Krevor S (2020) Cut, overlap and locate: a deep learning approach for the 3D localization of particles in astigmatic optical setups. *Exp Fluids* 61(6):140. <https://doi.org/10.1007/s00348-020-02968-w>
- Galata DL, Mészáros LA, Kállai-Szabó N, Szabó E, Pataki H, Marosi G, Nagy ZK (2021) Applications of machine vision in pharmaceutical technology: A review. *Eur J Pharm Sci* 159(105):717. <https://doi.org/10.1016/j.ejps.2021.105717>
- Gao Z, Wu Y, Bao Y, Gong J, Wang J, Rohani S (2018) Image analysis for in-line measurement of multidimensional size, shape, and polymorphic transformation of L-glutamic acid using deep learning-based image segmentation and classification. *Cryst Growth Des* 18(8):4275–4281. <https://doi.org/10.1021/acs.cgd.8b00883>
- Ghiasi G, Cui Y, Srinivas A, Qian R, Lin TY, Cubuk ED, Le QV, Zoph B (2021) Simple copy-paste is a strong data augmentation method for instance segmentation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
- Gong Y, Fan N, Yang X, Peng B, Jiang H (2018) New advances in microfluidic flow cytometry. *Electrophoresis* 40:1212–1229. <https://doi.org/10.1002/elps.201800298>
- Gossett DR, Weaver WM, Mach AJ, Hur SC, Tse HTK, Lee W, Amini H, Di Carlo D (2010) Label-free cell separation and sorting in microfluidic systems. *Anal Bioanal Chem* 397(8):3249–3267. <https://doi.org/10.1007/s00216-010-3721-9>
- Grabarek AD, Weinbuch D, Jiskoot W, Hawe A (2019) Critical evaluation of microfluidic resistive pulse sensing for quantification and sizing of nanometer- and micrometer-sized particles in biopharmaceutical products. *J Pharm Sci* 108(1):563–573. <https://doi.org/10.1016/j.xphs.2018.08.020>
- Gross-Rother J, Blech M, Preis E, Bakowsky U, Garidel P (2020) Particle detection and characterization for biopharmaceutical applications: current principles of established and alternative techniques. *Pharmaceutics* 12(11):1112. <https://doi.org/10.3390/pharmaceutics12111112>
- He K, Gkioxari G, Dollár P, Girshick R (2017) Mask R-CNN. In: *2017 IEEE International Conference on Computer Vision (ICCV)*
- Hessel V, Löwe H, Schönfeld F (2005) Micromixers—a review on passive and active mixing principles. *Chem Eng Sci* 60(8–9):2479–2501. <https://doi.org/10.1016/j.ces.2004.11.033>
- He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
- Hill SC, Pinnick RG, Niles S, Fell NF, Pan Y-L, Bottiger J, Bronk BV, Holler S, Chang RK (2001) Fluorescence from airborne micro-particles: dependence on size, concentration of fluorophores, and illumination intensity. *Appl Opt* 40(18):3005–3013. <https://doi.org/10.1364/ao.40.003005>
- Hu H, Gu J, Zhang Z, Dai J, Wei Y (2018) Relation networks for object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
- Kähler CJ, Scharnowski S, Cierpka C (2012) On the resolution limit of digital particle image velocimetry. *Exp Fluids* 52(6):1629–1639. <https://doi.org/10.1007/s00348-012-1280-x>
- Kiebert F, Wege S, Massing J, König J, Cierpka C, Weser R, Schmidt H (2017) 3D measurement and simulation of surface acoustic wave driven fluid motion: a comparison. *Lab Chip* 17(12):2104–2114. <https://doi.org/10.1039/c7lc00184c>
- König J, Chen M, Rösing W, Boho D, Mäder P, Cierpka C (2020) On the use of a cascaded convolutional neural network for three-dimensional flow measurements using astigmatic PTV. *Meas Sci Technol* 31(7):074015. <https://doi.org/10.1088/1361-6501/ab7bfd>
- Lewandowski B, Seichter D, Wengelfeld T, Pfennig L, Drumm H, Gross HM (2019) Deep orientation: fast and robust upper body orientation estimation for mobile robotic applications. In: *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*
- Li Y, Chen Y, Wang N, Zhang Z (2019) Scale-aware trident networks for object detection. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*
- Lin TY, Dollar P, Girshick R, He K, Hariharan B, Belongie S (2017) Feature pyramid networks for object detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*
- Loshchilov I, Hutter F (2019) Decoupled weight decay regularization. In: *International Conference on Learning Representations (ICLR)*
- Ma Z, Zhou Y, Collins DJ, Ai Y (2017) Fluorescence activated cell sorting via a focused traveling surface acoustic beam. *Lab Chip* 17(18):3176–3185. <https://doi.org/10.1039/c7lc00678k>
- Mach H, Bhambhani A, Meyer BK, Burek S, Davis H, Blue JT, Evans RK (2011) The use of flow cytometry for the detection of subvisible particles in therapeutic protein formulations. *J Pharm Sci* 100(5):1671–1678. <https://doi.org/10.1002/jps.22414>
- Massing J, Kaden D, Kähler CJ, Cierpka C (2016) Luminescent two-color tracer particles for simultaneous velocity and temperature measurements in microfluidics. *Meas Sci Technol* 27(11):115301. <https://doi.org/10.1088/0957-0233/27/11/115301>
- McKinnon KM (2018) Flow cytometry: an overview. *Curr Protoc Immunol* 120:5.1.1–5.1.11. <https://doi.org/10.1002/cpim.40>
- Miyagawa A, Okada T (2021) Particle manipulation with external field; from recent advancement to perspectives. *Anal Sci Int J Jpn Soc Anal Chem* 37(1):69–78. <https://doi.org/10.2116/analsci.20SAR03>
- Picot J, Guerin CL, van Kim C, Boulanger CM (2012) Flow cytometry: retrospective, fundamentals and recent instrumentation. *Cytotechnology* 64(2):109–130. <https://doi.org/10.1007/s10616-011-9415-0>
- Ren S, He K, Girshick R, Sun J (2015) Faster R-CNN: towards real-time object detection with region proposal networks. In: *Proceedings of the 28th International Conference on Neural Information Processing Systems (NIPS) - Volume 1*
- Rice SB, Chan C, Brown SC, Eschbach P, Li Han, Ensor DS, Stefaniak AB, Bonevich J, Vladar AE, Hight WAR, Zheng J, Starnes C, Stromberg A, Ye J, Grulke EA (2013) Particle size distributions by transmission electron microscopy: an interlaboratory

- comparison case study. *Metrologia* 50(6):663–678. <https://doi.org/10.1088/0026-1394/50/6/663>
- Rossi M, Kähler CJ (2014) Optimization of astigmatic particle tracking velocimeters. *Exp Fluids* 55(9):1809. <https://doi.org/10.1007/s00348-014-1809-2>
- Rossi M, Segura R, Cierpka C, Kähler CJ (2012) On the effect of particle image intensity and image preprocessing on the depth of correlation in micro-PIV. *Exp Fluids* 52(4):1063–1075. <https://doi.org/10.1007/s00348-011-1194-z>
- Russakovsky O, Deng J, Su H, Krause J, Satheesh S, Ma S, Huang Z, Karpathy A, Khosla A, Bernstein M, Berg AC, Fei-Fei L (2015) ImageNet large scale visual recognition challenge. *Int J Comput Vis* 115(3):211–252. <https://doi.org/10.1007/s11263-015-0816-y>
- Sachs S, Baloochi M, Cierpka C, König J (2022) On the acoustically induced fluid flow in particle separation systems employing standing surface acoustic waves - Part I. *Lab Chip* 22:2011–2027. <https://doi.org/10.1039/d1lc01113h>
- Sachs S, Cierpka C, König J (2022) On the acoustically induced fluid flow in particle separation systems employing standing surface acoustic waves - Part II. *Lab Chip* 22:2028–2040. <https://doi.org/10.1039/d2lc00106c>
- Sachs S, Ratz M, Mäder P, König J, Cierpka C (2022) Particle size recognition by deterministic approaches and deep neural networks using astigmatism particle tracking velocimetry (APTV). In: *Proceedings of the 20th International Symposium on the Application of Laser and Imaging Techniques to Fluid Mechanics*
- Sajeesh P, Sen AK (2014) Particle separation and sorting in microfluidic devices: a review. *Microfluid Nanofluid* 17(1):1–52. <https://doi.org/10.1007/s10404-013-1291-9>
- Sehgal P, Kirby BJ (2017) Separation of 300 and 100 nm Particles in Fabry-Perot Acoustofluidic Resonators. *Analytical chemistry* 89(22):12192–12200. <https://doi.org/10.1021/acs.analchem.7b02858>
- Seichter D, Kohler M, Lewandowski B, Wengelfeld T, Gross HM (2021) Efficient RGB-D semantic segmentation for indoor scene analysis. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*
- Sharma DK, King D, Oma P, Merchant C (2010) Micro-flow imaging: flow microscopy applied to sub-visible particulate analysis in protein formulations. *AAPS J* 12(3):455–464. <https://doi.org/10.1208/s12248-010-9205-1>
- Shekunov BY, Chattopadhyay P, Tong HHY, Chow AHL (2007) Particle size analysis in pharmaceuticals: principles, methods and applications. *Pharm Res* 24(2):203–227. <https://doi.org/10.1007/s11095-006-9146-7>
- Stetefeld J, McKenna SA, Patel TR (2016) Dynamic light scattering: a practical guide and applications in biomedical sciences. *Biophys Rev* 8(4):409–427. <https://doi.org/10.1007/s12551-016-0218-6>
- Stockslager MA, Olcum S, Knudsen SM, Kimmerling RJ, Cermak N, Payer KR, Agache V, Manalis SR (2019) Rapid and high-precision sizing of single particles using parallel suspended microchannel resonator arrays and deconvolution. *Rev Sci Instrum* 90(8):085. <https://doi.org/10.1063/1.5100861>
- Weatherall E, Willmott GR (2015) Applications of tunable resistive pulse sensing. *Analyst* 140(10):3318–3334. <https://doi.org/10.1039/c4an02270j>
- Wu M, Ozcelik A, Rufo J, Wang Z, Fang R, Jun Huang T (2019) Acoustofluidic separation of cells and particles. *Microsyst Nanoeng* 5:32. <https://doi.org/10.1038/s41378-019-0064-3>
- Xu R (2015) Light scattering: a review of particle characterization applications. *Particuology* 18:11–21. <https://doi.org/10.1016/j.partic.2014.05.002>
- Zeng N, Li H, Peng Y (2021) A new deep belief network-based multi-task learning for diagnosis of Alzheimer's disease. *Neural Comput Appl*. <https://doi.org/10.1007/s00521-021-06149-6>
- Zeng N, Wu P, Wang Z, Li H, Liu W, Liu X (2022) A small-sized object detection oriented multi-scale feature fusion approach with application to defect detection. *IEEE Trans Instrum Meas* 71:1–14. <https://doi.org/10.1109/TIM.2022.3153997>
- Zhang W, Hu Y, Choi G, Liang S, Liu M, Guan W (2019) Microfluidic multiple cross-correlated Coulter counter for improved particle size analysis. *Sensors Actuators B Chem* 296:126615. <https://doi.org/10.1016/j.snb.2019.05.092>
- Zhang S, Wang Y, Onck P, den Toonder J (2020) A concise review of microfluidic particle manipulation methods. *Microfluid Nanofluid* 24(4):24. <https://doi.org/10.1007/s10404-020-2328-5>
- Zhang J, Hartman JH, Chen C, Yang S, Li Q, Tian Z, Huang PH, Wang L, Meyer JN, Huang TJ (2020) Fluorescence-based sorting of *Caenorhabditis elegans* via acoustofluidics. *Lab Chip* 20(10):1729–1739. <https://doi.org/10.1039/d0lc00051e>
- Zhong J, Liang M, Ai Y (2021) Submicron-precision particle characterization in microfluidic impedance cytometry with double differential electrodes. *Lab Chip* 21(15):2869–2880. <https://doi.org/10.1039/d1lc00481f>
- Zoupanou S, Chiriaco MS, Tarantini I, Ferrara F (2021) Innovative 3D microfluidic tools for on-chip fluids and particles manipulation: from design to experimental validation. *Micromachines* 12(2):104. <https://doi.org/10.3390/mi12020104>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.