ARCHIVIO ISTITUZIONALE
DELLA RICERCA

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

Deep learning and defocus imaging for determination of three-dimensional position and orientation of microscopic objects

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:
Mehdizadeh Youshanlouei, M., Rossi, M. (2024). Deep learning and defocus imaging for determination of three-dimensional position and orientation of microscopic objects. PHYSICS OF FLUIDS, 36(8), 082006-1-082006-7 [10.1063/5.0219081].

Availability:
This version is available at: https://hdl.handle.net/11585/980403 since: 2024-09-02

Published:

DOI: http://doi.org/10.1063/5.0219081

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (https://cris.unibo.it/).
When citing, please refer to the published version.

(Article begins on next page)

09 November 2026

ACCEPTED MANUSCRIPT
Physics of Fluids
AIP Publishing
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0219081

Deep Learning and Defocus Imaging for Determination of Three-Dimensional Position and Orientation of Microscopic Objects

Mohammad Mehdizadeh Youshanlouei[1] and Massimiliano Rossi[1]
Department of Industrial Engineering, Alma Mater Studiorum - University of Bologna, Via Fontanelle 40, 47121 Forlì, Italy.

(*Electronic mail: mohammad.mehdizadeh2@unibo.it and massimiliano.rossi13@unibo.it)

(Dated: 15 July 2024)

We present a method to determine the three-dimensional (3D) position and orientation of microscopic, non-spherical objects in microfluidic and lab-on-a-chip systems observed through conventional optical microscopes. The method is based on the combination of the General Defocusing Particle Tracking technique [Barnkob *et al., Lab on a Chip*, 2015, **15**, 3556] and deep learning. It requires minimal input from the user, is suitable for real-time applications, and can be applied to any microscopic object with an approximately ellipsoidal shape, such as unicellular swimming organisms, red blood cells, or spheroidal colloids. The main challenge is linked to the construction of suitable training datasets for the neural network. We provide a procedure generally valid for active microswimmers and discuss possible strategies for other types of objects. An implementation using the Visual Geometry Group convolutional neural network (VGG-16) is presented and tested on synthetic images with different backgrounds and noise levels. The same implementation is used to track the position and orientation of different specimens of the heterotrophic ciliate *Euplotes Vannus* in free-swimming motion. The measurements were performed with a 10× objective over a depth of 800 μm with an average estimated uncertainty in the orientation angles of 9.0%.

I. INTRODUCTION

Tracking particles, cells, microorganisms, or other microscopic objects suspended in a fluid is a ubiquitous task in the experimental characterization of lab-on-a-chip and microfluidic systems. The most common situation is to observe the specimens through a conventional optical microscope connected to a digital camera. While tracking the position of individual objects in the images is an easy task for current image-processing packages, this is often insufficient to tackle more complex problems. In velocimetry applications, the displacement of a large number of passive tracer particles must be evaluated in image pairs to determine the instantaneous velocity fields. One popular technique in this situation is the microscopic particle image velocimetry (μPIV)[1,2], which statistically evaluates the particle displacement in small interrogation windows. However, µPIV provides a two-dimensional (2D) picture of the flow at the focal plane of the microscopic objective, while many applications require a full three-dimensional (3D) flow description.

Many 3D velocimetry techniques for microfluidics have been developed so far[3], however, their use is still limited among the lab-on-a-chip community, most likely due to their complexity or the requirement of special equipment such as a stereo[4] or holographic microscope[5]. An attempt to provide an accurate yet easy-to-use methodology for 3D particle tracking in microfluidics was the General Defocusing Particle Tracking technique (GDPT)[6]. GDPT is based on the defocusing principle[7,8], where tracer particles in a volume are observed with an objective with a small depth of field, and the out-of-plane particle coordinates are determined by the defocusing patterns of the respective particle images. GDPT works on conventional optical microscopes, uses a pattern recognition algorithm that requires minimum input from the user, and has been successfully used in several research investigations on acoustofluidics[9,10], evaporating droplets[11], blood and biological flows[12,13], small microorganisms[14].

The GDPT approach paved the way for artificial intelligence methods, in particular, algorithms based on deep convolutional neural networks (CNNs)[15–18]. Artificial intelligence is already revolutionizing microfluidics and lab-on-a-chip systems in many aspects[19], especially concerning data analysis and image processing approaches[20–22]. The flexibility of CNNs allows us to tackle tasks normally impossible with conventional single-camera defocusing approaches, such as the tracking of polydisperse particles, recently shown on particles with different diameters[23,24]. Another interesting application is the determination of the 3D position and orientation of non-spherical objects. This is relevant, for instance, in the characterization of non-spherical colloids in microfluidic systems[25], for the study of motion and propulsion mechanisms of active microswimmers[26], or in the steering of artificial microswimmers[27]. The measurement of position and orientation could also have a significant impact on the fundamental understanding of complex swimming behaviors of micro-organisms. The majority of recent works in this domain use conventional methods where only the projected orientation is considered, or the organisms are confined in 2D geometries[28].The 3D description of the motion could open new insight for instance in the understanding of the run-and-tumble motion[29] or the migration and anisotropy behavior of micro-swimmers[30]. While some results in this domain have been obtained with holographic microscopy[31], a more straightforward approach suitable for conventional microscopes is not yet available.

In this work, we present a CNN-based method to determine the 3D position and orientation of non-spherical objects observed through a small depth-of-field optical system. The method was first tested on synthetic images of spheroidal particles with different shapes, backgrounds, and noise lev-

ACCEPTED MANUSCRIPT
Physics of Fluids
AIP Publishing
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0219081

els. Then it was applied to experimental images of different specimens of the heterotrophic ciliate *Euplotes vannus* in free-swimming motion[32]. The main limitation of this approach, shared with methods based on supervised learning, is the requirement of a large set of labeled data to train the CNN effectively. This problem was solved here with an *ad hoc* procedure applicable for swimmers at low Reynolds numbers[33].

II. METHODOLOGY

The 3D position and orientation of a solid object in a measurement volume are fully described by the three coordinates of its center of mass (x, y, z) and three Euler angles (ψ, θ, ϕ), which following Tait-Bryan convention correspond to the angles of yaw, pitch, and roll, respectively (Figure 1a). The GDPT method determines the depth position z of a target particle image from the comparison with a reference set of defocus images obtained from an experimental calibration procedure (Figure 1b). The proposed methodology intends to use a similar approach to determine not only the z position but also the orientation angles ψ, θ, and ϕ of non-spherical objects (Figure 1c). However, relating a defocused image of a non-spherical object to the corresponding values of depth position and orientation angles cannot easily be achieved with a deterministic algorithm, as in the case of GDPT. CNN architectures, which are the standard AI tool for image processing, have already been proposed to solve similar types of regression problems with respect to particle depth position and size[23,24]. Following this approaches, we used CNNs trained on a large number of different defocus images to achieve

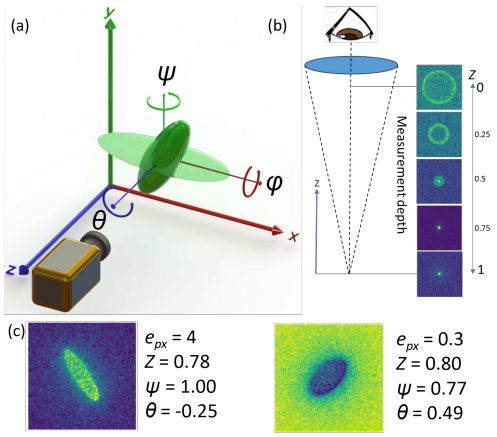

FIG. 1. (a) Illustration of 3D object's orientation angles yaw (ψ), pitch (θ), and roll (ϕ). (b) Schematic and synthetic images of non-spherical particles in different Z positions. (c) Synthetic images of non-spherical particles with different shapes in darkfield (e_{px} = 4) and brightfield (e_{px} = 0.3).

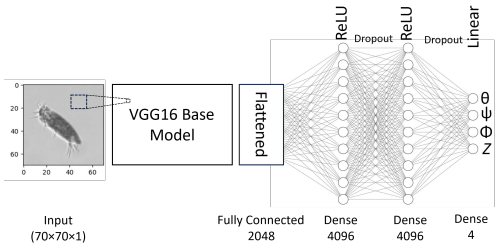

FIG. 2. Schematic of the VGG-16 deep learning convolutional neural network modified and utilized in presented study

this task. The determination of a suitable dataset of training data becomes then critical for the practical applicability of the method and it will be discussed in the following sections.

We tested different CNN architectures with different modifications and settings, namely ResNet18, ResNet50, ResNet101[34], and VGG-16[35]. The CNNs were implemented using the Pyhton-based library TensorFlow. The original versions of the mentioned architectures were designed for classification problems, while the present case is a regression problem, so we had to modify the output layers of each model. Concretely, each of the presented architectures was modified with a unique outlet box including a couple of Flatten, Dense, and Dropout layers (Figure 2). We used a custom input-image shape of $70 \times 70 \times 1$ pixels (gray-scale images), therefore the use of a pre-trained model was not possible. In our implementation, the architectures of the ResNet family provided significant over-fitting and non-acceptable training loss and validation loss trends during training and were discarded. On the other hand, the VGG-16 provided a satisfactory performance in all cases. Further CNN architectures were not investigated at this stage, but they shall be considered in follow-up works concerning the optimization of this approach.

Both mean squared error (MSE) and mean absolute error (MAE) were tested as loss functions, showing overall similar performances. MAE showed slightly lower uncertainty in the final results, most likely due to the presence of outliers, and it was was selected for the final regression using the Adam optimizer. The VGG-16 with a learning rate of 0.0001 and batch size of 64 gave the best results in our case. The implemented version of VGG-16 had one flatten layer followed by two Dense layers with the size of 4096, Relu activation, and Dropout layers, and a final Dense layer with a size of 4 and linear activation providing the output of Z, ψ, θ and ϕ. The depth position was given in normalized unit $Z = z/h$, where h is the height of the measurement volume.

The determination of the x-y coordinates was necessary only in the experimental case, and it was performed using the KCF[36] or CSRT[37] trackers from the Python-based OpenCV library.

ACCEPTED MANUSCRIPT
Physics of Fluids
AIP Publishing
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0219081

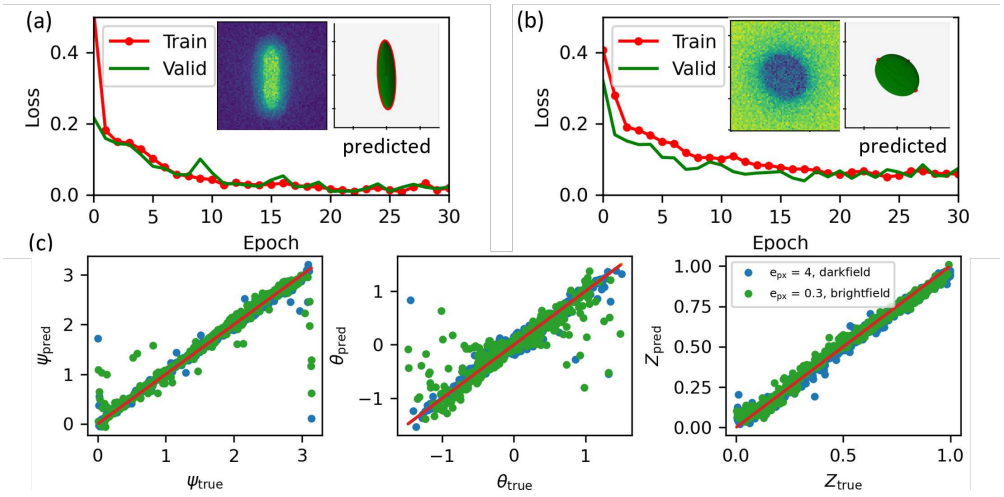

FIG. 3. Loss and validation loss of the training of the VGG-16 network with 5000 synthetic images for the (a) darkfield and (b) lightfield case. (c) True and predicted results of the orientation angles (ψ), (θ), and depth position Z of non-spherical particle for the two presented cases.

III. CASE I: SYNTHETIC IMAGES

A. Preparation of the dataset

A first validation and assessment of this approach was performed on synthetic images produced with the software MicroSIG[38] based on ray tracing. The software can create synthetic defocus images of spheroidal particles with a different elongation coefficient (e_{px}), backgrounds, and noise levels. In order to have a more general evaluation of the capabilities of the method, two different datasets were investigated. The first set included images with a darkfield background and an elongation coefficient of e_{px} = 4, and the second set images with a brightfield background and e_{px} = 0.3. Both sets contained a total of 5000 labeled images with the shape of $70 \times 70 \times 1$ pixels that were used to train the VGG-16 network. Moreover, 10% of the images were randomly selected for validation during the training process. Finally, 400 new images were generated and utilized in the testing step. The target outputs were the depth position z and two Euler angles ψ and θ. The roll angle was not considered here due to the rotational symmetry of the spheroidal particles, but it will be considered for the experimental images in Section IV. Figure 1c shows two examples of the used synthetic images and corresponding true data.

B. Results of CNN training and uncertainty assessment

The VGG-16 network presented in Section II was trained with 5000 images with a batch size of 64, for 30 epochs. Increasing the number of training images or epochs did not improve the final result significantly. The computations were run on an NVIDIA GeForce RTX 3070 Ti Laptop with GPU enabled for processing. With the mentioned system, the processing of each epoch takes approximately 5 seconds at 71 ms/step. The loss and validation loss of the training for both datasets are shown in Figure 3a-b. Overfitting was observed during training, which could be reduced by implementing the Dropout layers in the outlet box. The insets show examples of the original defocus images with the corresponding predicted orientation of the particle.

The predicted values of ψ, θ, and Z versus the true ones for the 400 test images are presented in 3c, for both cases. For the darkfield case with $e_{px} = 4$ (rod-shaped particles), a mean absolute error of 1.1%, 1.8% and 2.1% was obtained for ψ, θ and Z, respectively. For the lightfield case with $e_{px} = 0.3$ (disk-shaped particles), a mean absolute error of 1.9%, 2.8% and 2.5% was obtained for ψ, θ and Z, respectively. The results on synthetic images show that this approach can potentially predict the depth position and orientation of spheroidal particles from their defocused images with good accuracy. To validate the method on experimental images, we considered the case of a microswimmer with an ellipsoidal shape, the heterotrophic ciliate *Euplotes Vannus*.

ACCEPTED MANUSCRIPT
Physics of Fluids
AIP Publishing
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0219081

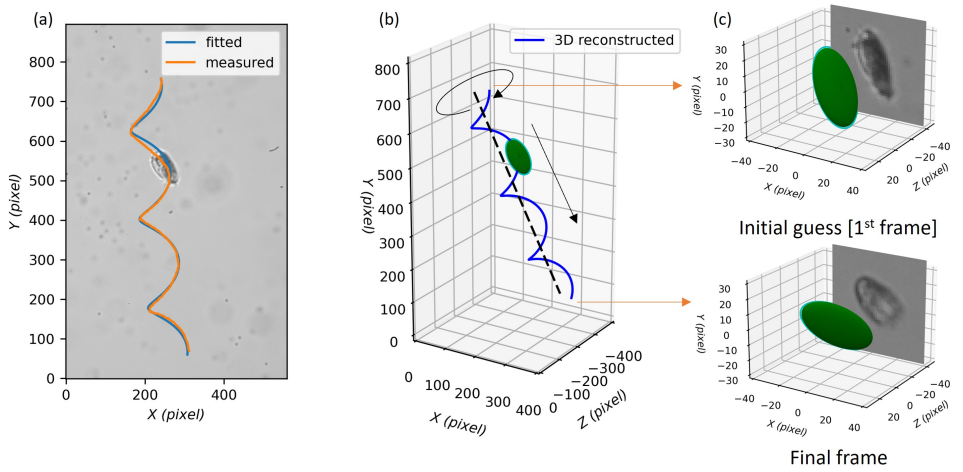

FIG. 4. (a) Measured 2D trajectory of animal with OpenCV and 3D fitted trajectory with the mathematical model of regular movement of the animal. (b) 3D illustration of the trajectory combined with 3D rotated objective animal. (c) Initial guess of the orientation of the swimmer and final resulted in the orientation of the animal in comparison with real images.

IV. CASE II: EXPERIMENTAL IMAGES OF *EUPLOTES VANNUS*

A. Determination of ground truth data for microswimmers

In experimental cases, labeled images with ground truth data are normally not readily available. This is a critical aspect of the method that will be discussed later. In the case of microswimmers, it is possible to exploit a general characteristic of the motion at low Reynolds numbers[33]. When the microswimmer proceeds on a quiescent fluid at a constant pace (i.e., with a regular beat of its flagellum or cilia), its trajectory necessarily follows a helical path, and its body rotates around the axis of the helix. Under these conditions, it is possible to obtain the microswimmer 3D position and orientation from subsequent 2D image recordings[11].

The first step is to obtain the specimen trajectory. We can write the equation of an arbitrary helical path using six free parameters

$$\begin{bmatrix} x(t) \\ y(t) \\ z(t) \end{bmatrix} = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} + R(\alpha,\beta,\gamma) \begin{bmatrix} ct \\ \rho\cos(\omega t) \\ \rho\sin(\omega t) \end{bmatrix}, \qquad (1)$$

where ρ, ω, and c describe a general helix along the x-axis, and the Euler angles α, β, γ orient the helix in the space. $R(\cdot,\cdot,\cdot)$ is a rotation matrix as a function of three Euler angles. We can estimate the six free parameters by fitting the x and y components of equation (1) to the experimental trajectory measured from the images $(x_m(t), y_m(t))$. The starting point can be chosen arbitrarily as $x_0 = x_m(0)$, $y_0 = y_m(0)$, $z_0 = 0$. After this step, the 3D position of the specimen over time is determined with good accuracy as shown in Figure 4a-b).

The next step is to obtain the orientation angles of the specimen as a function of time. We follow here a trial-and-error approach. Under the considered hypothesis of swimming motion, if the orientation angles are known for one time instant t_0, it is possible to obtain them for any time t just by applying an additional rotation around the helical axis u

$$R(\psi(t),\theta(t),\phi(t)) = R_u(\omega(t-t_0))R(\psi_0,\theta_0,\phi_0) \qquad (2)$$

where $R_u(\cdot)$ represents a rotation matrix as a function of the rotation axis u and a given angle. We begin by providing an initial guess of the orientation angles (ψ_0, θ_0, ϕ_0) that gives a result in agreement with the experimental image at time t_0 (Figure 4c). We then derive the orientations for other frames from equation (2) and check the agreement with the experimental images. The agreement is evaluated from a qualitative comparison between a projected ellipsoid and the experimental images, and repeated until the optimal match is obtained (Figure 4b-c). A technicality to consider is that the sign of ω cannot directly be identified by the fitting procedure. Therefore, it must be determined *a priori*, for instance looking at the rotation of patterns on the organism body.

B. Collection of experimental images and dataset preparation

The *E. vannus* culture (Culture Collection of Algae and Protozoa, SAMS Limited, Scottish Marine Institute, UK) was

ACCEPTED MANUSCRIPT
Physics of Fluids
AIP Publishing
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0219081

grown in artificial seawater at 18 °C and diluted 2–3 times per year in 65-mL flasks with artificial seawater and autoclaved rice grains to serve as a bacterial substrate. For the measurements, a group of different specimens was selected and placed in a custom-built, circular sample chamber and observed with an Olympus IX71 inverted microscope with a 10× objective. High-speed image recordings were obtained with a Phantom Miro LAB 320 at a frame rate of 150 fps.

The training data were obtained using selected frames from 5 videos of different specimens of *E. vannus* swimming at a regular pace, giving a total of approximately 500 frames. Since the frames belong to 5 different videos, overall we had 5 different background lights, noise, and contrast. For each video, the ground truth values of ψ, θ, ϕ, and Z were obtained from the procedure described in Section IV A. In order to increase the number of training data, data augmentation has been applied. In this regard Keras images preprocessing library called ImageDataGenerator has been used and mentioned augmentation includes three steps. First, frames were randomly rotated and the corresponding rotation was applied to ψ. In the next step, the brightness and contrast of all images have been adjusted randomly in a specific range. Finally, all images have been normalized between 0 and 1. As a result, a dataset with 13678 labeled images with an acceptable range of different background settings was prepared.

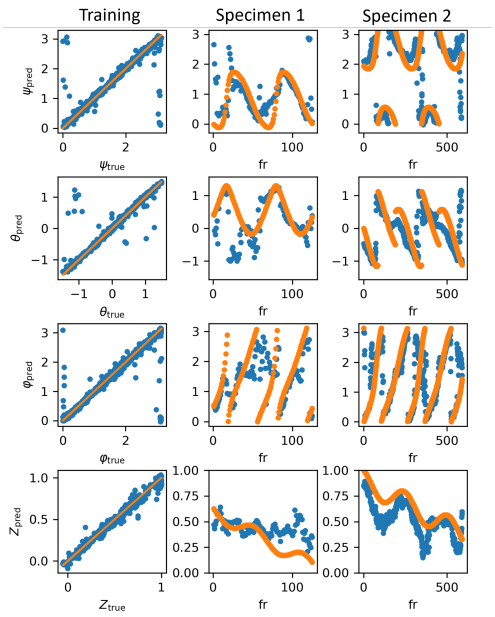

FIG. 5. (Training) True versus predicted results of the test set of microswimmer. (Specimen 1 and 2) True and predicted results of non-seen data of microswimmer, tested data are from videos that all frames are just used in testing part.

C. Results of CNN training and uncertainty assessment

From the whole dataset, 5% of the images were separated randomly for testing and 5% for validation. The VGG-16 network was trained with the remaining 12438 labeled images. The intensity of all images was rescaled to be between 0 and 1 . The network consisted of 39,903,940 trainable parameters that were trained for 37 epochs. Figure 5 (left column) shows the predicted versus true values of ψ, θ, ϕ, and Z, with a corresponding mean average error of 1.1%, 0.9%, 1.3%, and 1.7%, respectively. As can be seen from the plot, there is a remarkable error at $\psi \approx 0$ and $\psi \approx 3.14$ which is due to the similarity of related images in these points.

Finally, we tested the algorithm on non-seen cases. First, we applied the method on two specimens swimming at a regular pace, for which it was possible to determine the ground truth values using the procedure in Section IV A. The results are reported in Figure 5 (middle and right columns), with the true values in orange. The results indicate an average error of 9.0% and 9.0% in the orientation angles and 8.4% and 9.0% in the Z position for specimens 1 and 2, respectively. Then, we applied the method to a specimen swimming randomly. The results are presented in Figure 6 (Multimedia available online). Due to the random swimming pattern, true data are not available for this case and visualization is the only way of evaluating the data.

V. DISCUSSION

Overall, the presented methodology was able to successfully determine the 3D orientation and depth position of non-spherical objects both in synthetic and experimental images. Not surprisingly, a larger uncertainty was observed for non-seen experimental images. This is explained by the fact that, although the total number of training images was large enough, it was obtained only from five specimens. Expanding the training set to more specimens is expected to reduce the uncertainty. Another difficulty in the preparation of the experimental training data is that the procedure in Section IV A provides values of z relative to an arbitrary position z_0, which might be different for different movies. In this work, we compare the defocusing patterns to align the depths of the different movies. In future studies, more robust strategies can be implemented to improve this aspect, such as confined experimental setups where the specimen is constrained to swim within the measurement depth or an independent measurement of the z values through different approaches.

The implementation presented in this work can readily be applied to microswimmers, with practical application in the study of predator/prey behavior, feeding of microorganisms, and propulsion mechanisms. In particular, in the case of the *E.*

ACCEPTED MANUSCRIPT
Physics of Fluids
AIP Publishing
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0219081

Vannus specimens, one user needs to acquire images using the same optical settings described here and can then use the code right away. In case of different swimmers, the same CNN architecture can be used, but it must be trained on a different set of training images, prepared followed the instruction in Section IV A. Other possible fields of application could be the characterization of spheroidal colloids or biological samples, such as red blood cells. In these cases, it would be necessary to develop suitable microfluidic technologies or strategies to manipulate the sample in a controlled fashion, to obtain the labeled dataset with orientation and position, for instance using acoustic tweezers[39] or oscillating micro-bubbles[40,41]. In general, a collaborative effort among researchers working in this field will be necessary to build up a significantly large dataset of training images for different types of objects and optical systems to increase the applicability of this approach in more real-world applications.

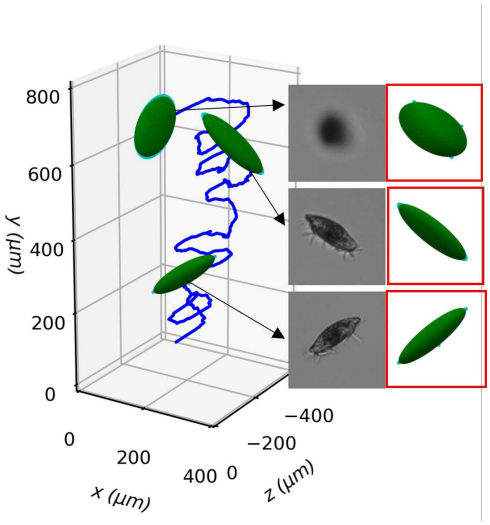

FIG. 6. Predicted trajectory and orientation of a non-seen case. Due to the non-regular movement of the microswimmer, finding true data was not possible for this case (Multimedia available online).

VI. CONCLUSIONS

In this work, we combined the GDPT method with deep learning to predict the 3D position and orientation of non-spherical objects observed through a conventional microscope. The method relied on a modified VGG-16 network and it was applied to synthetic images of single particles in different shapes and background light, and experimental images of different specimens of *E. vannus* swimming in a fluid volume. The results indicated that the method is capable of determining complex movements and rotating patterns, even in the more challenging case of non-seen data, with an average uncertainty for the estimation of the orientation angles ranging from 2% to 10%. The method can be extended to different applications including non-spherical colloids, red blood cells, or other biological samples. Furthermore, once the CNN has been trained, the processing time of new images is minimal making this approach suitable for real-time applications. Finally, it should be noted that the success and widespread of this method will strongly depend on the collaborative effort of research groups, especially concerning the preparation and sharing of training datasets. From this perspective, all the weights and codes of the presented study are available in a public repository.[42]

AUTHOR CONTRIBUTIONS

MMY: Conceptualization, Data Curation, Investigation, Methodology, Software, Visualization, Writing—Original Draft. MR: Conceptualization, Methodology, Software, Supervision, Writing—Review & Editing.

CONFLICTS OF INTEREST

There are no conflicts to declare.

ACKNOWLEDGMENTS

MR acknowledges financial support from the VILLUM Foundation under the Grant No. 00036098. The authors gratefully acknowledge Mads Rode and Federica Miano for providing the experimental test images of *E. Vannus*, and Anders Andersen and Thomas Kjørboe for abundant and helpful discussions.

[1]J. G. Santiago, S. T. Wereley, C. D. Meinhart, D. Beebe, and R. J. Adrian, "A particle image velocimetry system for microfluidics," Experiments in fluids **25**, 316–319 (1998).
[2]R. Lindken, M. Rossi, S. Große, and J. Westerweel, "Micro-particle image velocimetry (μpiv): recent developments, applications, and guidelines," Lab on a Chip **9**, 2551–2567 (2009).
[3]C. Cierpka and C. J. Kähler, "Particle imaging techniques for volumetric three-component (3D3C) velocity measurements in microfluidics," Journal of visualization **15**, 1–31 (2012).
[4]M. Bown, J. MacInnes, R. Allen, and W. Zimmerman, "Three-dimensional, three-component velocity measurements using stereoscopic micro-piv and ptv," Measurement science and technology **17**, 2175 (2006).
[5]T. Ooms, R. Lindken, and J. Westerweel, "Digital holographic microscopy applied to measurement of a flow in a t-shaped micromixer," Experiments in Fluids **47**, 941–955 (2009).
[6]R. Barnkob, C. J. Kähler, and M. Rossi, "General defocusing particle tracking," Lab Chip **15**, 3556–3560 (2015).
[7]C. Willert and M. Gharib, "Three-dimensional particle imaging with a single camera," Experiments in Fluids **12**, 353–358 (1992).
[8]R. Barnkob and M. Rossi, "General defocusing particle tracking: fundamentals and uncertainty assessment," Experiments in Fluids **61**, 110 (2020).
[9]J. T. Karlsen, W. Qiu, P. Augustsson, and H. Bruus, "Acoustic streaming and its suppression in inhomogeneous fluids," Physical review letters **120**, 054501 (2018).

AIP Publishing Physics of Fluids ACCEPTED MANUSCRIPT
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0219081

[10]W. Qiu, J. T. Karlsen, H. Bruus, and P. Augustsson, "Experimental characterization of acoustic streaming in gradients of density and compressibility," Physical Review Applied **11**, 024018 (2019).
[11]M. Rossi, G. Cicconofri, A. Beran, G. Noselli, and A. DeSimone, "Kinematics of flagellar swimming in Euglena gracilis: Helical trajectories and flagellar shapes," Proceedings of the National Academy of Sciences **114**, 13085–13090 (2017).
[12]G. Coutinho, A. S. Moita, M. Rossi, and A. L. Moreira, "Experimental perspective on the mechanisms for near-wall accumulation of platelet-size particles in pressure-driven red blood cell suspension flows," Physical Review Fluids **8**, 103101 (2023).
[13]D. Van Assche, E. Reithuber, W. Qiu, T. Laurell, B. Henriques-Normark, P. Mellroth, P. Ohlsson, and P. Augustsson, "Gradient acoustic focusing of sub-micron particles for separation of bacteria from blood lysate," Scientific reports **10**, 3670 (2020).
[14]K. Taute, S. Gude, S. Tans, and T. Shimizu, "High-throughput 3d tracking of bacteria on a standard phase contrast microscope," Nature communications **6**, 8776 (2015).
[15]J. König, M. Chen, W. Rösing, D. Boho, P. Mäder, and C. Cierpka, "On the use of a cascaded convolutional neural network for three-dimensional flow measurements using astigmatic ptv," Measurement Science and Technology **31**, 074015 (2020).
[16]R. Barnkob, C. Cierpka, M. Chen, S. Sachs, P. Mäder, and M. Rossi, "Defocus particle tracking: a comparison of methods based on model functions, cross-correlation, and neural networks," Measurement Science and Technology **32**, 094011 (2021).
[17]M. Dreisbach, R. Leister, M. Probst, P. Friederich, A. Stroh, and J. Kriegseis, "Particle detection by means of neural networks and synthetic training data refinement in defocusing particle tracking velocimetry," Measurement Science and Technology **33**, 124001 (2022).
[18]C. Sax, M. Dreisbach, R. Leister, and J. Kriegseis, "Deep learning and hybrid approach for particle detection in defocusing particle tracking velocimetry," Measurement Science and Technology (2023).
[19]K. Goda, H. Lu, P. Fei, and J. Guck, "Revolutionizing microfluidics with artificial intelligence: a new dawn for lab-on-a-chip technologies," Lab on a Chip **23**, 3737–3740 (2023).
[20]Y. Feng, Z. Cheng, H. Chai, W. He, L. Huang, and W. Wang, "Neural network-enhanced real-time impedance flow cytometry for single-cell intrinsic characterization," Lab on a Chip **22**, 240–249 (2022).
[21]K. Huang, H. Matsumura, Y. Zhao, M. Herbig, D. Yuan, Y. Mineharu, J. Harmon, J. Findinier, M. Yamagishi, S. Ohnuki, *et al.*, "Deep imaging flow cytometry," Lab on a Chip **22**, 876–889 (2022).
[22]Y. Wang, Z. Huang, X. Wang, F. Yang, X. Yao, T. Pan, B. Li, and J. Chu, "Real-time fluorescence imaging flow cytometry enabled by motion deblurring and deep learning algorithms," Lab on a Chip **23**, 3615–3627 (2023).
[23]S. Sachs, M. Ratz, P. Mäder, J. König, and C. Cierpka, "Particle detection and size recognition based on defocused particle images: a comparison of a deterministic algorithm and a deep neural network," Experiments in Fluids **64**, 21 (2023).
[24]M. Ratz, S. Sachs, J. König, and C. Cierpka, "A deep neural network architecture for reliable 3d position and size determination for lagrangian particle tracking using a single camera," Measurement Science and Technology **34**, 105203 (2023).
[25]S. Sachs, H. Schmidt, C. Cierpka, and J. König, "On the behavior of prolate spheroids in a standing surface acoustic wave field," Microfluidics and Nanofluidics **27**, 81 (2023).
[26]G. Gompper, R. G. Winkler, T. Speck, A. Solon, C. Nardini, F. Peruani, H. Löwen, R. Golestanian, U. B. Kaupp, L. Alvarez, *et al.*, "The 2020 motile active matter roadmap," Journal of Physics: Condensed Matter **32**, 193001 (2020).
[27]H.-W. Huang, F. E. Uslu, P. Katsamba, E. Lauga, M. S. Sakar, and B. J. Nelson, "Adaptive locomotion of artificial microswimmers," Science advances **5**, eaau1532 (2019).
[28]L. Zeng, W. Jiang, and T. J. Pedley, "Sharp turns and gyrotaxis modulate surface accumulation of microorganisms," Proceedings of the National Academy of Sciences **119**, e2206738119 (2022), https://www.pnas.org/doi/pdf/10.1073/pnas.2206738119.
[29]A. Patteson, A. Gopinath, M. Goulian, and P. Arratia, "Running and tumbling with e. coli in polymeric solutions," Scientific reports **5**, 15761 (2015).
[30]M. Guan, W. Jiang, L. Tao, G. Chen, and J. H. Lee, "Migration of confined micro-swimmers subject to anisotropic diffusion," Journal of Fluid Mechanics **985**, A44 (2024).
[31]A. Wang, R. F. Garmann, and V. N. Manoharan, "Tracking e. coli runs and tumbles with scattering solutions and digital holographic microscopy," Optics express **24**, 23719–23725 (2016).
[32]M. Rode, T. Kiørboe, and A. Andersen, "Feeding flow and membranelle filtration in ciliates," Physical Review Fluids **7**, 023102 (2022).
[33]E. M. Purcell, "Life at low Reynolds number," American journal of physics **45**, 3–11 (1977).
[34]K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2016) pp. 770–778.
[35]K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," arXiv preprint arXiv:1409.1556 (2014).
[36]J. F. Henriques, R. Caseiro, P. Martins, and J. Batista, "High-speed tracking with kernelized correlation filters," IEEE transactions on pattern analysis and machine intelligence **37**, 583–596 (2014).
[37]A. Lukezic, T. Vojir, L. ˇCehovin Zajc, J. Matas, and M. Kristan, "Discriminative correlation filter with channel and spatial reliability," in *Proceedings of the IEEE conference on computer vision and pattern recognition* (2017) pp. 6309–6318.
[38]M. Rossi, "Synthetic image generator for defocusing and astigmatic piv/ptv," Measurement Science and Technology **31**, 017003 (2019).
[39]A. Marzo and B. W. Drinkwater, "Holographic acoustic tweezers," Proceedings of the National Academy of Sciences **116**, 84–89 (2019).
[40]J. Zhang, S. Yang, C. Chen, J. H. Hartman, P.-H. Huang, L. Wang, Z. Tian, P. Zhang, D. Faulkenberry, J. N. Meyer, *et al.*, "Surface acoustic waves enable rotational manipulation of caenorhabditis elegans," Lab on a Chip **19**, 984–992 (2019).
[41]A. Volk, M. Rossi, B. Rallabandi, C. J. Kähler, S. Hilgenfeldt, and A. Marin, "Size-dependent particle migration and trapping in three-dimensional microbubble streaming flows," Physical review fluids **5**, 114201 (2020).
[42]Link to a public repository will be posted in the published version of the paper.