Aviral Verma Paper check

By Aviral Verma

WORD COUNT	2691	TIME SUBMITTED	20-APR-2024 05:00PM
		PAPER ID	108398582

Machine Learning Based Digital Holographic Microscopy

Aviral Verma
National Insititute of Science Education and Research
Bhubaneshwar, India -752050
aviral.verma@niser.ac.in

Abstract

In this report, we briefly described the experimental details used for recording the digital hologram. We had reproduced the previous works and provided their results along with the bit of explanation about the LodeSTAR model. We also showed the produced synthetic data in comparison with the experimental image with their pixel intensity hologram. Future work has been provided along with the difficulties that are currently being faced.

1 Introduction

Digital Holographic Microscopy (DHM) is a technique that can be used to store information about the particles in 3 dimensions along with various other properties. The interference pattern, which is called the hologram, is created using a coherent illumination source like a laser. Using a beam splitter, the laser is split into two parts, one of which goes and interacts with the specimen called the object beam, and the other is the reference beam, which interferes with the object beam to create holograms that carry the phase information. In this project, we aim to track the particles' position in 3D and size in the range of micrometres in a microfluidic device using deep learning algorithms on the DHM data of the particles.

The particles, in general, follow Brownian motion in a fluid. But when the particle comes under a local thermal perturbation, it starts to move due to effects like convective flow, thermo-osmotic flow, thermoviscous flow, and thermophoresis in the fluid. when the fluid transport is dominant, the particle trajectories directly report on the fluid dynamics. Achieving this local thermal perturbation will be explained in the experimental setup section. This local thermal perturbation at the microscale can be used to achieve long-range transport, which can lead to changes in mm scale for both fluid and particle dynamics. The motivation behind the work is that this long-range movement of particles and fluids can be used in lab-on-a-chip technology for the mixing of samples and fluids, which is still a big hurdle. It is very difficult to control the fluid dynamics on such a small scale as the available external fluid controllers are pretty bulky and cannot be used on such small devices. There is an apparent mismatch in the size.

2 Experimental Setup in Brief

The experimental setup consists of a flow cell system consisting of two coverslips spaced by a silica space of known length, and the cell contains water (n = 1.333) and silica particles having a diameter of 1.5μm with refractive index (n) = 1.4645. The lower slip contains the uniform distribution of gold nanorods (AuNRs) on which we shine a light to produce local thermal perturbation. The flow cell is illuminated using the pump-probe technique. The AuNRs are illuminated by a pump beam with a wavelength of 780 nm to create heat, and the DHM is done using a probe beam of 465 nm. CMOS

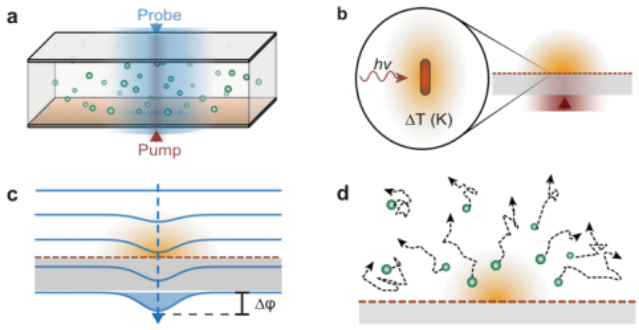

Figure 1: Schematics of Experimental Setup

camera with a resolution of 0.106μm/pixel with a magnification of 55x and a numerical aperture (NA) of 0.65 was used to record the hologram. The experiment utilizes a label-free thermometry method that identifies small phase changes caused by temperature dependent variations in the refractive index of the imaging medium. These changes accumulate as the incident wavefront traverses the sample, resulting in an overall difference in optical path length. The fluid dynamics are determined by tracking the movement of tracer particles in three dimensions. B. Ciraulo et al. developed a custom off-axis digital holographic microscope that integrates wavefront sensing and 3D particle-tracking velocimetry, operating in a pump/probe setup. Detailed information about the experimental setup and the achievement of long-range transport can be found in the reference [1].

3 Reproduced work:

Similar work has been done by Midtvedt et al. [2,] using two different models: one is U-Net, which is a supervised learning model, and the other is Localization and detection from Symmetries, Translations And Rotations (LodeSTAR), which is a self-supervised model.

The LodeSTAR model has achieved better results using the same dataset than the U-Net for the same task. The results they have achieved compared to the existing numerical algorithms have been shown in fig.(4,5) and their loss in fig.(2,3).

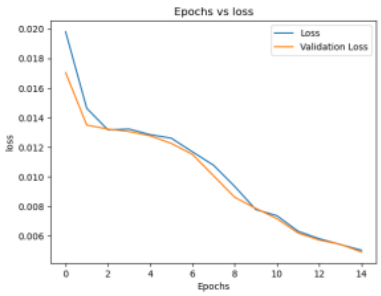

Figure 2: For U-Net

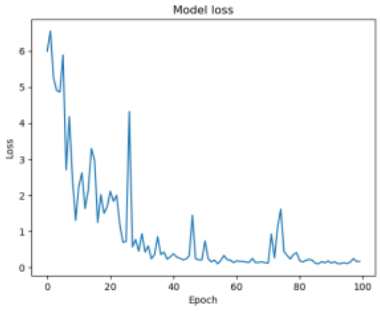

Figure 3: For LodeSTAR

LodeSTAR is a system designed to detect microscopic objects by leveraging inherent symmetries. It operates on the principle of equivariance, meaning that transformations applied to the object image, such as translations, rotations, and reflections, result in corresponding transformations in the object's predicted position. This allows LodeSTAR to accurately locate the centre of an object, even if its absolute position is unknown. The system achieves this by training a neural network to establish an exact correspondence between the transformations applied to the input image and their

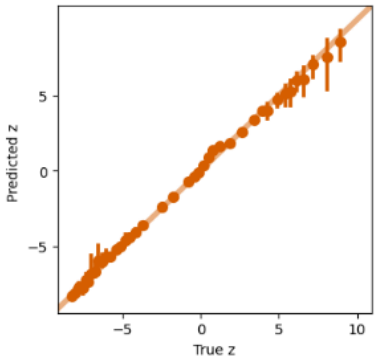

Figure 4: LodeSTAR output vs Numerical Algorithm predicted z

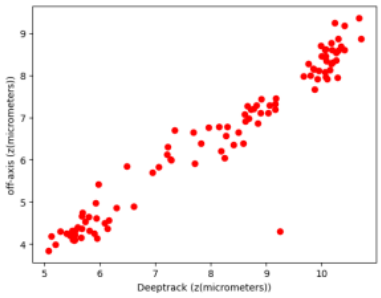

Figure 5: U-Net output vs Numerical Algorithm predicted z

effects on the output prediction. In addition to the aforementioned capabilities, LodeSTAR can also process holographic images. These images can be moved to various planes or different axial positions from the focal plane using Fourier transforms. This introduces another form of equivariance that LodeSTAR can learn, similar to the equivariances on the plane. This means that LodeSTAR can adapt to changes in the axial position of the object image, further enhancing its ability to locate microscopic objects accurately. The precise mathematical explanation of the LodeSTAR model can be found in the method section of ref.[3].

3.1 Experimental details, Data representation and preprocessing:

The experimental data [4] used in the models above have very different optical details than ours, on which the above models and results have been shown. In their dataset, the particles have a size of 190nm with a refractive index of 1.45. Their hologram has a resolution of 0.345 μ/m, and the Δ z = 28μm. In the U-Net model, the no. of features, which are the no. of the output of the model, is defined as Δz/resolution.

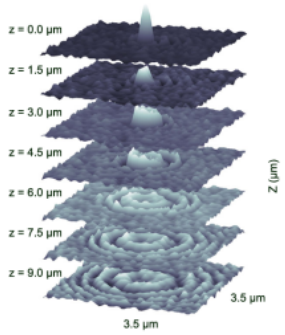

Figure 6: Propagation of focal plane in z

To find the 3D position from the 2D images, the interference pattern was reconstructed into the Fourier space, which has been used for analysis. Their data representation was done as a Matlab file containing the reconstructed field. These traces contain the Region of Interest (ROI) and position in x,y, and z, which were evaluated using numerical algorithms and the mapping matrix. For U-Net,

Each particle's ROI was selected and cropped into a 64x64 image size before giving it to the model. The holographic image can be moved to various planes, meaning different axial positions from the focal plane, by using Fourier transforms. This introduces an equivariance that LodeSTAR can learn, similar to the equivariances on the plane. By training on the image in the top slice, i.e., z = 0, as shown in Fig. (6), LodeSTAR learns to identify the location of the polystyrene spheres in 3D space. Here, the measured vertical position is represented as the distance in the image.

4 Synethetic Images

This section will show the images generated to train the U-Net. Before giving images to the model, we must ensure that the pixel-intensity histograms of simulated and experimental images are the same. DeepTrack[2] framework produced the simulated images, brightfield microscopy, and other optical details. The results are shown in fig.(7)

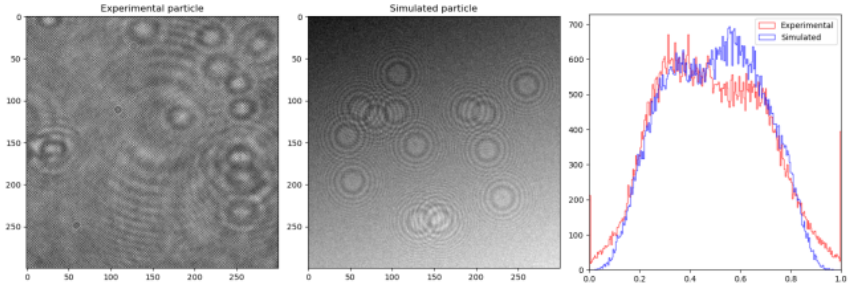

Figure 7: Pixel intensity of Histogram

5 Update on LodeStar Model

I cannot find any relevant information on how to do the data representation, which was done in the previous work. I tried to take the Fourier transform of the data and then apply it to the LodeStar model. Still, it could not provide any significant results even after following the procedure given in the LodeStar paper.

6 Using U-Net to predict the particles labels

First, we create the spherical labels for the particles to validate the model.

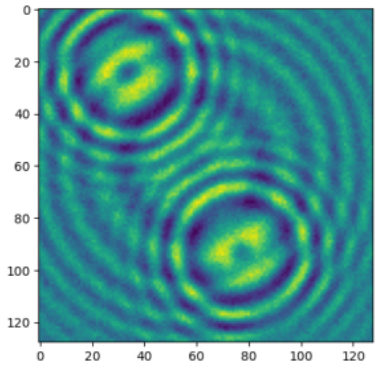

Figure 8: For U-Net

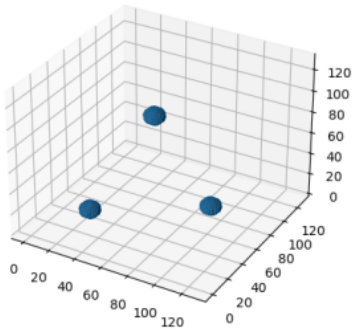

Figure 9: For LodeSTAR

When an input image is given to the Unet model, it results in a volume with particles, as shown below. Where the height of the volume is cut into pieces, which we call bins. Currently, for our data, we have 122 bins.

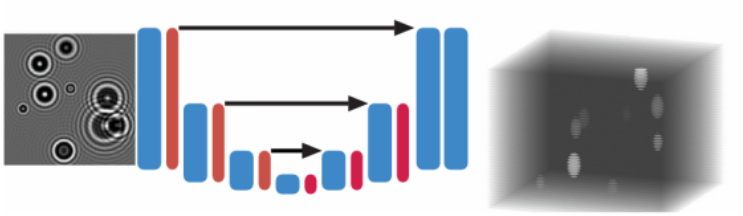

Figure 10: Schematics of Unet

Since the Unet used in the previous work was very basic, with just 0.9M parameters, We tried to increase the complexity of the model so the model could learn from highly noisy data and extract features from the images. We have increased the trainable parameters for the model from 0.9 M to 57.6 M. We have used the Adam optimiser to have a slow learning rate and have used weighted cross-entropy loss with the weight of (20,1). The loss is weighted so that the pixel value of 1 to 0 is penalised more than the 0 to 1.

$$L_{WCE} = -(\beta y log(\hat{y}) + (1 - y) log(1 - \hat{y})) \tag{1}$$

The parameter β, representing adjustment weights, can either be a scalar or a vector, depending on the application of the cross-entropy function. It takes on a scalar value when the function is applied to binary scenarios and a vector when dealing with categorical situations. By adjusting these weights, it's possible to prioritise either false negative or false optimistic predictions. A β value less than 1 penalises errors in false optimistic predictions, while a β value more excellent than 1 penalises errors in false negatives.

Using the same parameters in the code as used in the experiment was causing memory issues as the output tensor for the model was becoming too big. So, we decreased the expected height of the chamber (height in z) in the model while keeping the pixel size the same, as the height was directly proportional to the size of the chamber along with the pixel size, but if we change the pixel size, the synthetically produced image optics will change which we do not want. Hence, we limit the height so the output tensor does not cause memory issues.

Finally, we trained the model on 3000 synthetic images for 20 epochs in a batch of 8 (We tried to increase the batch size beyond 8, hoping the model could generalise well, but this again caused memory errors). The plot of the loss given below suggests that the training went well.

Since I have the Trace, X, Y and Z for the experimental data, we use the trace of a single particle, which consists of the X, Y and Z positions of that particle frame by frame. We crop the particle in each frame using the X and Y location into a 64x64 image and give it to the model as an input of (1,64,64,1) tensor. Since the model returns us a volume with the label of the particles, we found that the model provides good labelling results, as shown below, along with the image provided as the input to the model. (Here, I have used a 128x128 image to show that the model predicts pretty good labels)

We can see in the above pictures that the predicted labels match the particle position in the input image to a very good extent. If we stack up all the predicted images, we will get the volume with particles labelled as spheres inside it.

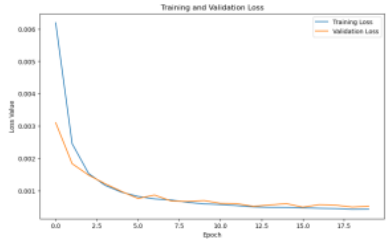

Figure 11: Loss vs Epochs

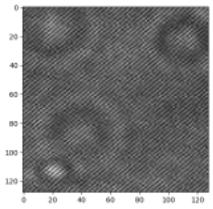

Figure 12: Image given to the model

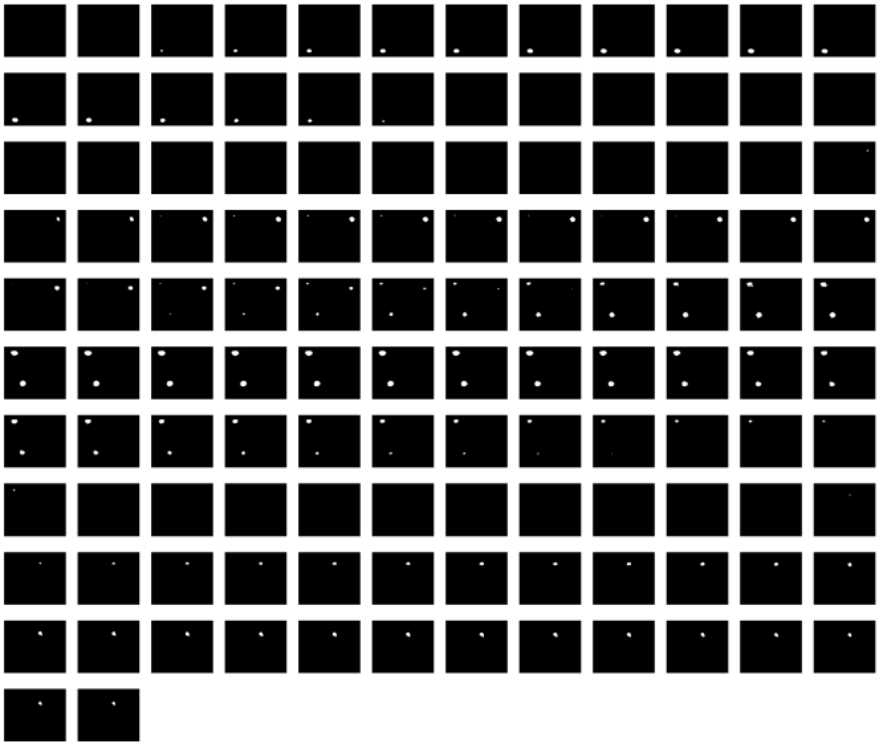

Figure 13: Predicted label of particles from the model

7 Problems with Post Processing the model's output

- We used the centroid algorithm to get the z position from the labels to post-process the model's output. In the previous work, the data being used had a very low particle density; hence, they could crop a single particle, and in the predicted label, they get a volume with the label of the single particle inside. In that case, using the centroid algorithm will provide you with the correct z position of the particle. But in our case, due to high particle density, even if you want to crop a single particle, one will get the other particle in the frame and then in the predicted volume. Since we have multiple particles, using the centroid algorithm will consider all the particle labels while giving the z position, which will obviously be wrong. On discussing this problem with my mentor for the project, he suggested that I try to take an even smaller crop of the frame such that one particle is in the frame, but again, doing this will reduce the features of the image. I tried with the 8x8 frame, 16x16 frame, and 32x32 frame, but the model could not predict the labels and gave errors.

- Another problem we face is the defocusing of the particle. Since the camera used in the experiment is focused on a particular z, sometimes the particle's interference patterns are not even visible when in different z positions. So, while giving a frame to the model, the model does not predict any labels if this is the case.

8 Possible Solutions to above-discussed problems

To solve the defocus problem, the only way is to numerically propagate the focus to different z positions and get the focus frames. For this, I have the Matlab code, which I am converting to Python. However, the code had some user-defined functions that were not there in the code, which I will probably receive next week. Then, I will try to propagate the fields numerically, get the focused particle, and get better outputs.

On discussing the multiple particle problem with the TA, she suggested that we use the DETR(Detection Transformer) bipartite matching loss function for this purpose and translate the Object detection problem to our problem statement.

In dETR, many classes are defined as containing information about the object along with a class of no object along with the information of bounding boxes that tell the object's position based on which model is being trained. Hence, when we give an image as an input, the model tries to predict the classes and bounding boxes of the objects in the image. During the training, bipartite matching matched the object with a bounding box and class.

Since we use synthetic data, we have all the information about the particle's location in x,y and z, and as we divide the height in z into bins, we will define these bins as classes in the model. As DETR is made for the COCO dataset, we need to prepare the custom dataset with the appropriate information, which we can give to the model. We can use DETECTRON 2, as the information is available on creating the custom dataset and how to train the model using it. The model will try given particles as the object and learn to map the particles into the bins, which will directly give us the z position. As we cut the z-axis into discrete bins, there will be some errors with the particle's position.

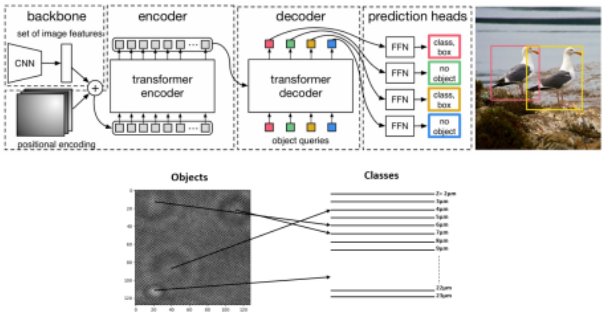

Figure 14: DETR analogy with our current problem

9 Future Work:

I will implement the DETR with our custom dataset during the summer (as it is time-consuming with my current programming skills). If I receive the code for numerical propagation, I will try to use it to get better-focused images and better predictions, and I will try to update the report before the May 10 deadline.

References

[1] Ciraulo, B., Garcia-Guirado, J., de Miguel, I. et al. Long-range optofluidic control with plasmon heating. Nat. Commun. 12, 2001 (2021). [Link]

[2] Benjamin Midtvedt, Saga Helgadottir, Aykut Argun, Jesús Pineda, Daniel Midtvedt, Giovanni Volpe. "Quantitative Digital Microscopy with Deep Learning.“ Applied Physics Reviews 8 (2021), 011310 [Link]

[3] Midtvedt, B., Pineda, J., Skärberg, F. et al. Single-shot self-supervised object detection in microscopy. Nat Commun 13, 7492 (2022). [Link]

[4] D. Midtvedt, F. Eklund, E. Olsen, B. Midtvedt, J. Swenson, and F. H€ o€ ok, “Size and refractive index determination of subwavelength particles and air bubbles by holographic nanoparticle tracking analysis,” Anal. Chem. 92, 1908–1915 (2020d).

[5] Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., Zagoruyko, S. (2020). End-to-End Object Detection with Transformers. ArXiv, abs/2005.12872. [Link]

Aviral Verma Paper check

ORIGINALITY REPORT

8%
SIMILARITY INDEX

PRIMARY SOURCES

1	www.nature.com Internet	93 words —	4%
2	www.ncbi.nlm.nih.gov Internet	77 words —	3%
3	www.arxiv-vanity.com Internet	14 words —	1%
4	Zhihan Chen, Jingang Li, Yuebing Zheng. "Heat-Mediated Optical Manipulation", Chemical Reviews, 2021 Crossref	8 words —	< 1%
5	B. Ciraulo, J. Garcia-Guirado, I. de Miguel, J. Ortega Arroyo, R. Quidant. "Long-range optofluidic control with plasmon heating", Nature Communications, 2021 Crossref	6 words —	< 1%
6	Benjamin Midtvedt, Jesús Pineda, Fredrik Skärberg, Erik Olsén et al. "Single-shot self-supervised object detection in microscopy", Nature Communications, 2022 Crossref	6 words —	< 1%

EXCLUDE QUOTES	ON	EXCLUDE SOURCES	OFF
EXCLUDE BIBLIOGRAPHY	ON	EXCLUDE MATCHES	OFF