EP-297
Exploratory Project Report
2020-2021, Semester-4

Machine learning holography for 3D particle field imaging

Presented By:

Pushker Naresh

Engineering Physics, IDD, Part II

Roll No. 19174018

Supervised by:

Dr. Rakesh Kr. Singh

Department of Physics,
Indian Institute of Technology (BHU),
Varanasi - 221005

Contents

Machine learning holography for 3D particle field imaging

- **Objective**

As per the challenges arrising in particle holograms where accurate measurement of individual particles is crucial.

There were some methods that were developed earlier for the computational imaging tasks as for the significant improvement in the particle extraction rate, localization accuracy and speed compared to prior methods over a wide range of the particle concentrations, including highly dense concentrations but those methods were found to be unsuitable.

So in order to cope up those challenges discover a new learning based approach that can potentially improve our computational imaging tasks or it could be extended to the other types of computational imaging tasks with similar features. For the 3D particle field imaging using holography and demonstrate them with synthetic and experimental holograms.

Introduction

In the past few decades, three-dimensional (3D) imaging has grown in popularity for measurements of particle size, shape, position, and motion in fields such as fluid dynamics, environmental science etc.

Digital holography (DH) has recently emerged as a powerful tool for such imaging tasks and is particularly useful for many in situ applications owing to its simple and compact setup. DH encodes complex information from the particles (e.g., 3D position and size) onto a 2D image called a hologram by recording the interference between a reference light wave and light scattered from the particles.The information can subsequently be recovered from the hologram through a digital reconstruction process.

Many approaches to overcome these issues focus on hardware design to improve the hologram quality or encode more information during the recording step of holography. However, the implementation of these approaches requires sophisticated mechanical and optical components.

Numerical approaches replace the mechanical complexity with computational complexity. However, these advanced methods are computationally intensive and require fine tuning parameters to get optimal results.

Recently, machine learning using deep neural networks (DNNs) has emerged as a prevailing tool for various image analysis tasks. Adoption of DNNs has drastically enhanced processing speed and yielded more accurate results than conventional inverse approaches for some applications.Machine learning in DH has been adopted for transforming hologram reconstructions to microscopic images similar to those commonly used in biological and medical examination, and classification of the particle species captured in the hologram.

Based on the literature review and compared with other learning-based image processing tasks, we have identified three unique challenges associated with 3D particle imaging using DH.

1.**First,unstabilty in the training process:**
while the signal of an individual object can spread over a large region of the hologram the reconstructed particle field usually consists of a group of sparse objects. When a learning-based approach is used to replace the reconstruction, this sparsity causes the training process to be highly unstable and produce incorrect results.

2.**Second,requirement for more accuracy:**
3D particle field reconstruction requires very accurate measurements for each particle which differs from many conventional learning-based imaging tasks such as classification or global regression of the image .

3.**Finally,limited generalisability of model:** the desired metrics, recording parameters, and hologram appearance are coupled, limiting the generalizability of a model trained on a specific set of data. It is worth noting that these challenges can also appear in light field imaging , imaging through diffusive media, defocus imaging and other methods.

To address the above-mentioned issues, we present in this paper a specially-designed machine learning approach for 3D particle field reconstruction in DH, which can also be employed in other computational imaging tasks sharing similar traits.

Our approach uses a U-net architecture incorporating residual connections, Swish activation, hologram preprocessing, and transfer learning to cope with challenges arising in particle holograms where accurate measurement of individual particles is crucial.

Methodolgy

Our machine learning approach for particle field reconstruction from a hologram uses a specially_x0002_designed U-net architecture, which takes the holograms as input and computes the 3D location of particles as output.U-net is a type of CNN developed for medical and biological image segmentation that has also been used in learning-based image-to-image transformations and multi-object classification from single images.

U-net consists of a series of encoder and decoder blocks, corresponding to solid and dashed black boxes.In the encoder block, two consecutive sets of convolution layers and activation functions are used to encode local features of input images into channels. Two encoder blocks are connected by a maximum pooling layer which downsamples the feature maps in order to extract global features. The decoder block is similar but in reverse. Two consecutive convolution layers are used to decode the channels to form an image and two decoder blocks are connected by up-convolution layers to resize feature maps.

U-net also includes skip connections (black arrows in Fig. 1) whereby the encoder output is concatenated to the same size decoder block which combines the local and global features of images for training in the deeper stage of the network.In comparison to a simple CNN architecture without skip connections, we suggest that U-net is more suitable or particle field reconstruction from a hologram because the skip connections make use of the spread of individual particle information over the a large portion of the image.

In our U-net architecture, we use a U-net with 4 encoders and 3 decoders, and the number of output encoder channels are 64 and 512 for the first and last encoder, respectively.

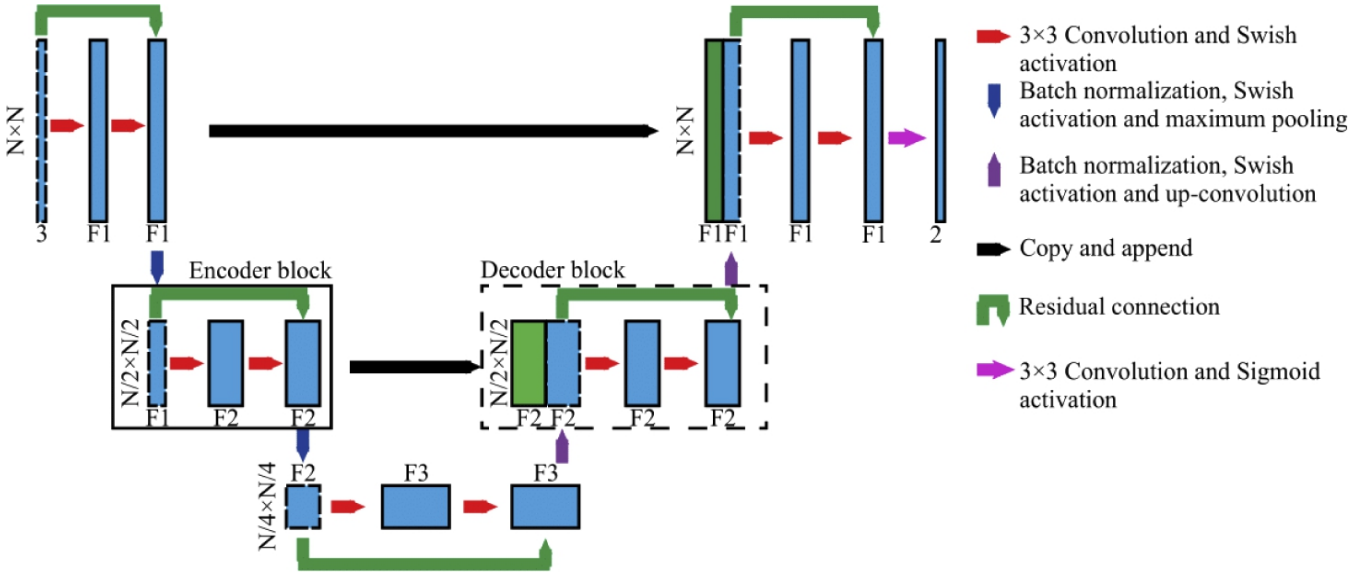


Fig. 1. The specially-designed U-net architecture for holographic reconstruction of 3D particle field

Compared with the conventional U-net architecture, our U-net has a residual connection within each encoder and decoder block (green arrows in Fig. 1) and uses a Swish (Sigmoid-weighed Linear Unit) activation function (Eq. (1)) for all except the last layer. The residual connection increases the training speed and reduces the likelihood of the training becoming trapped at a local minimum.

In the decoder the residual connection uses the channels from the previous decoder block connected by an up-convolution layer. Using activation function with the recently proposed Swish activation function (f(x), Eq. (1) in our architecture. Note that the x in Eq. (1) corresponds to the outputs from previous layer and f(x) is the input to the next layer.

$$f(x) = \frac{x}{1 + e^{-x}} \quad - \quad (1)$$

For particle holograms, the target images (Fig. 2) are usually sparse due to the small size of particle centroids which leads to the majority of values in the feature layers being equal to 0.Therefore, during training the parameters within the model have a higher tendency to be 0, we use a Sigmoid activation in the final decoder block (magenta arrow in Fig. 1) to produce results within the range from 0 to 1.

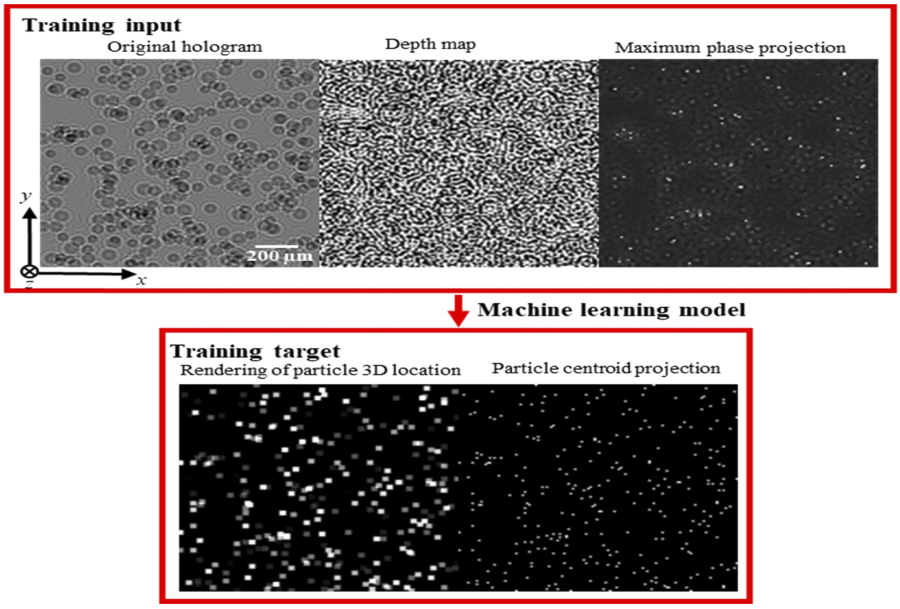


Fig. 2. A sample training input and training target consisting of 300 particles (i.e.,concentration at 0.018 ppp) with a hologram size of 128 ×128 pixels.

The training input consists of three channels: an original hologram, the corresponding images of pixel depth projection (i.e., depth map) and maximum phase projection (Fig. 2) The particles within the holograms are randomly distributed with a distance between 1 mm and 2.28 mm to the sensor. The resolution in z direction is also 10 μm with 128 discrete depth levels.

Compared with the depth map and phase projections are additional information obtained from preprocessing the holograms.Preprocessing is employed to incorporate existing hologram formation knowledge into the model and reduce the need of the model to fully learn the required physics during training. A noticeable improvement was observed for particle extraction rate especially for high concentration cases with these preprocessing steps in comparison to training directly on the raw holograms.

Using the angular spectrum method, a 3D complex optical field, up (x, y, z), is generated from the original hologram I(x,y) (Eq. (2)), where λ is the wavelength, k is the wave number and F is the Fourier transform operator. The depth map is generated by projecting the z locations where the pixels have the maximum intensity from up (x, y, z) to xy plane (Eq. (3)), and the maximum phase projection is calculated from Eq. (4).

$$u_{\mathrm{p}}(x,y,z) = \mathcal{F}^{-1}[\mathcal{F}(I(x,y)) \times \mathcal{F}(\frac{\exp(jkz)}{j\lambda z} \exp\left\{j\frac{k}{2z}[(x^2+y^2)]\right\})]$$

$$z_{\mathrm{approx}} = \arg\max_{z} \{u_{\mathrm{p}}(x,y,z) \times \mathrm{conj}[u_{\mathrm{p}}(x,y,z)]\}$$

$$P(x,y) = \max_{z}\{\mathrm{angle}[u_{\mathrm{p}}(x,y,z)]\}$$

The training target consists of two output channels. The first is a grayscale channel in which the pixel intensity corresponds to the relative depth of each particle and the second is a binary image of the particle xy centroids (Fig. 2).

Huber loss improves the training robustness and prediction accuracy by using MAE once the averaged pixel intensities are biased by the outliers. the parameter δ in Eq. (5) can be determined based on the measurement accuracy requirements, with a smaller δ resulting in an improved particle depth resolution.

$$L = \begin{cases} \frac{1}{2}||Y - X||_2^2 & ||Y - X||_1 \leq \delta, \\ \delta||Y - X||_1 - \frac{1}{2}\delta^2 & \text{otherwise.} \end{cases}$$

An MSE loss regularized by the total variation (TV) of the prediction is used for the xy centroid channel (Eq. (6)). As shown in Eq. (7), TV is the sum of first-order gradients over the image of size Nx×Ny.

$$L = (1 - \alpha)(||Y - X||_2^2) + \alpha||Y||_{TV}^2$$

$$||Y||_{TV} = \sum_{i=1}^{N_x}\sum_{j=1}^{N_y}\sqrt{(Y_{i,j} - Y_{i-i,j})^2 + (Y_{i,j} - Y_{i,j-1})^2}$$

The U-net architecture is implemented using Keras with the TensorFlow backend. The training is conducted on a Nvidia RTX 2080Ti GPU. The Adam optimizer is used with the default learning rate of 0.001. Thirteen datasets, with particle concentration between 1.9×10−4 and 6.1×10−2 ppp, are generated to train and test the models.

A base model is first trained on 2500 holograms with a particle concentration of 1.8×10−2 ppp for 480 epochs (in total 13.5 hours). For each subsequent particle concentration, a dataset of 2000 holograms is trained for 120 epochs with the training initialized by the base model (2.7 hours for each case). This transfer learning approach substantially decreases the training requirement (i.e., dataset size and training time) for new hologram datasets.

To extract the particles from the model output, the predicted particle xy centroid map is first binarized with a threshold of 0.5 to extract the xy centroids of the particles.

Subsequently, from the depth-encoded grayscale output, we use the intensity value of the corresponding pixels in the depth map as the particle depth.

Results

1. Assessment using synthetic holograms with constant particle concentration

To analyze the impacts of the unique features in our U-net architecture on the training process, we compare the training loss decay for the first 200 epochs of different model variants on the 1.8×10−2 ppp dataset. As shown in Fig. 4, compared to the proposed method, removing the residual connections or using a conventional loss function (MSE) both destabilize the training process. The removal of residual connections (Fig. 4(b)) leads the training process to be susceptible to local minima during training as discussed by He et al.

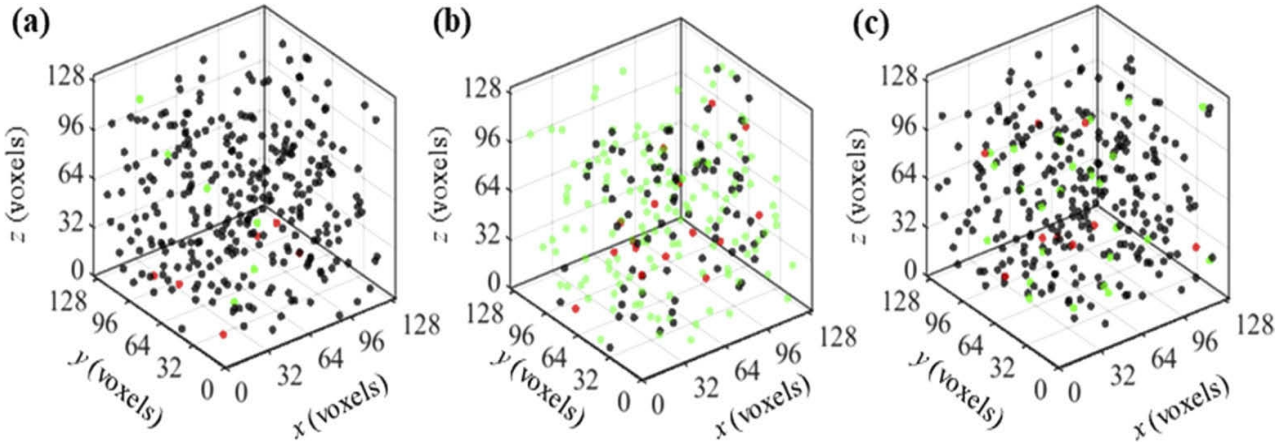


Fig. 3. Prediction results from the trained model using (a) our U-net architecture and (b) the method presented in Shimobaba et al. [44] (c) and Mallery and Hong [31] for the case of
0.018 ppp (300-particle holograms). The black dots are extracted true particles, red dots are false positives (i.e., unpaired particles from ground truth)

And green dots are the false negatives (unpaired particles from the ground truth).ss fluctuation of the loss curve even early in the training (Fig. 4(c)). The result is that the model does not converge to an optimal solution and produces very inaccurate pixel intensity predictions or white noise outputs for the worst scenarios.

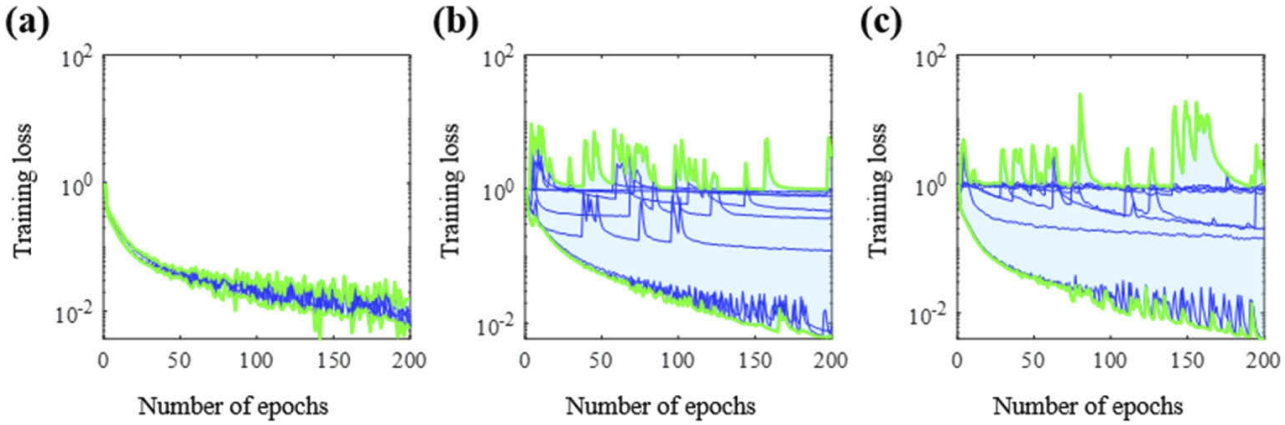


Fig. 4. Demonstration of the impact of the proposed model improvements on the training process over the first 200 epochs. (a) Proposed approach, (b) using U-net architecture without residual connection and (c) using mean squared error as loss function. The loss is normalized by its initial value, and each case is randomly initialized 10 times to show the resultant instability of the training for cases (b) and (c). In the image, the green curves correspond to the maximum and minimum normalized loss at each epoch, the blue curves corresponding to each initialization, and the shaded region is the range of loss.

Finally, replacement of ReLU activation functions with Swish optimizes the training process by avoiding untrainable parameters (i.e., dead neurons). From our tests, the cases without Swish can produce >6000 dead neurons in the last convolution layer at the end of the first epoch which substantially degrades the model. As a result, the training is likely to reach a plateau in the first few epochs and the resulting models generate 2D white noise images.

2. Assessment using synthetic holograms with variable particle concentration

The particle extraction rate and positioning accuracy using the proposed transfer learning approach (see Section 2) are assessed for variable particle concentrations from 1.9×10−4 ppp to 6.1×10−2 ppp. In Fig. 5, we present a case of a 100-particle hologram (6.1×10−3 ppp) and a case of a 1000-particle hologram (6.1×10−2 ppp). The lowest concentration instance shown here has an

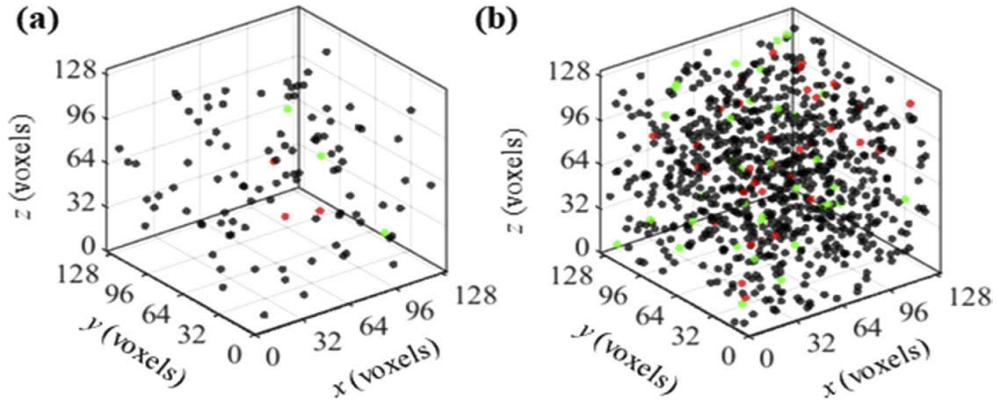


Fig. 5. Comparison of prediction results with a 100-particle h extraction rate of 97.0% (both false positive and false negative rates of 3.0%) while the highest concentration case has a 95.0% extraction rate (false positive rate of 1.6% and false negative rate
of 5.0%).

From an assessment of 100 holograms for each particle concentration, the extraction rate is above 94.4% (false negative less than 5.6%) over the range for 1.9×10−4 ppp to 6.1×10−2ppp (Fig. 6(a)) with a median particle positioning accuracy less than one voxel for x and y and less than 3.2 voxels for z for all concentrations (Fig. 6(b)). For all cases, the false positive of the model prediction is less than 10%.

Our machine learning approach shows no such drop within the studied range (Fig. 6(a)). We suggest that our pre-processing and transfer learning approaches significantly reduce the training requirements to yield a high extraction rate and positioning accuracy for new holograms especially for high concentration cases.

The increased extraction rate for high particle concentration holograms potentially enables improved spatial resolution tracer-based flow diagnostic techniques such as particle image velocimetry (PIV) and particle tracking velocimetry (PTV).

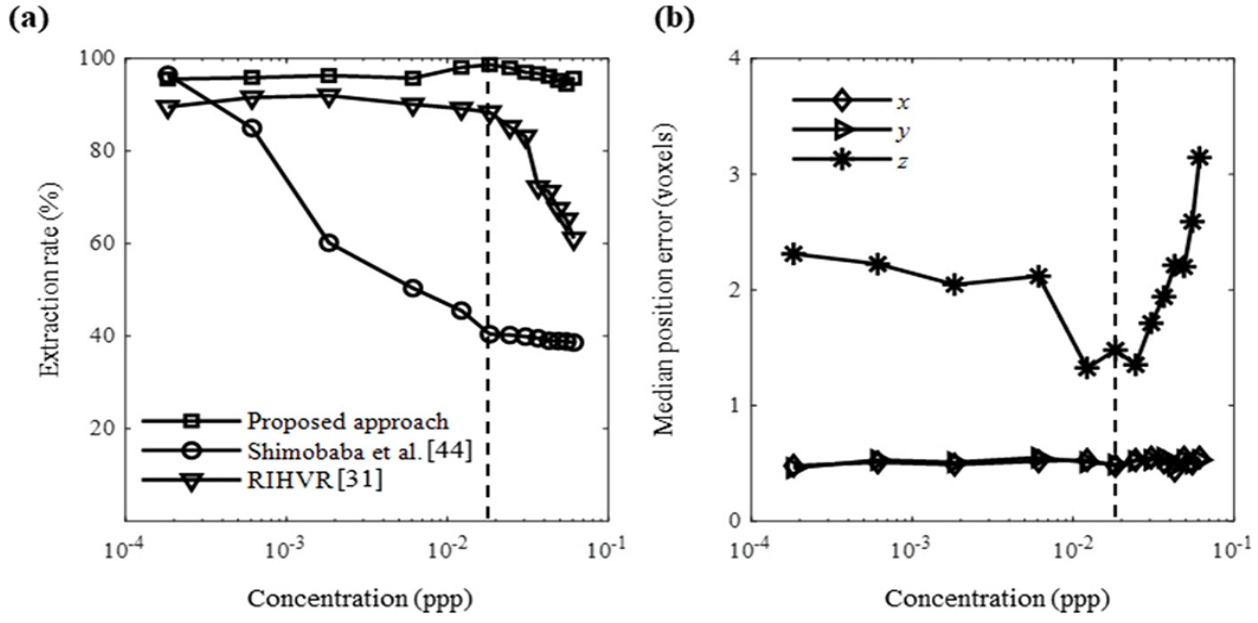


Fig. 6. (a) Extraction rate under different particle concentrations of the proposed method and compared with the case

3. Assessment using experimental data

The training dataset consists of 7500 randomly cropped 128×128 pixel tiles from the enhanced hologram and their corresponding particle locations from the 3D fluorescence scan volume. Here the target images are saved with 16-bit precision to encode a large number

of reconstruction planes. The Huber loss δ in Eq. (4) is set as 0.001 which is lower than the synthetic cases since the pixel intensity in the labels encode higher resolution in the depth (z direction in the experimental case).

We use the same method as the synthetic case to pair the predicted particles to the ground truth. As shown in Fig. 7(c), despite the noisy input (Fig. 7(a)), the predicted results from the trained model match well to the ground truth. The test of 100 randomly cropped 128×128 pixel tiles from a validation hologram imaging a different region of this sample yields a 90% extraction rate with a positioning error less than 1 voxel (0.65 μm) in the x and y directions and 5.24 voxels (13.1 μm) in the z direction.

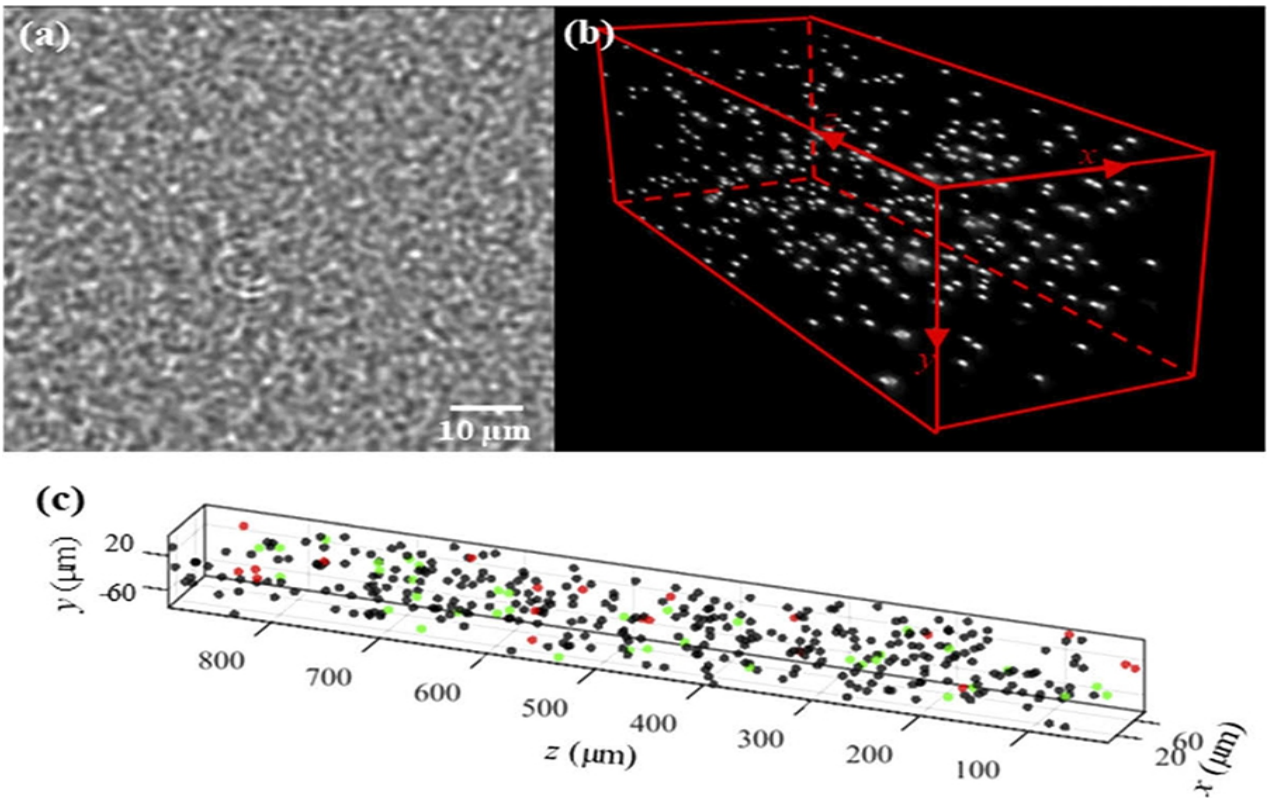


Fig. 7. (a) A 128×128-pixel enhanced hologram from the experimental data and corre_x0002_sponding volumetric image through the stacking of

Summary and discussion

In the present paper, we introduce a new learning-based approach for 3D particle field reconstruction from holograms. For holograms, the information in the training input (i.e. the diffraction fringes of an individual particle) can significantly spread beyond the in-focus location of the particle, while the training target consists of a sparse particle field where accurate measurement of each particle is crucial. To handle these traits, our specially-designed U-net architecture has three input channels (original holograms, depth, and maximum phase projection maps) and two output channels (depth and centroid maps of reconstructed particles).

Compared to simple feed-forward networks, U-net combines both local and global features which is particularly suitable for particle hologram reconstruction where the reconstructed particle target is small, but its information can spread significantly beyond its in-focus position in the hologram. The 2D depth and maximum phase projection map channels use the angular spectrum method from conventional holographic reconstruction to incorporate hologram formation knowledge into the U-net training and reduce the need to fully learn the required physics.

In addition, our architecture employs residual connections and the Swish activation function to reduce the likelihood of the training becoming trapped in local minima or producing a large number of dead neurons. We use two types of loss functions – Huber loss and TV-regularized MSE loss – for the output channels of particle depth and particle centroids, respectively, in order to improve the prediction accuracy, produce a smooth background, and reduce ghost particles.

Lastly, a transfer learning approach is adopted to reduce the training requirements for new hologram datasets. Through an assessment of synthetic holograms, our approach is suitable for processing much denser particle concentrations than prior approaches, with a 94% extraction rate at a concentration (6.1×10−2 ppp) 305 times higher than previously demonstrated with a machine learning approach [44] and 4 times higher than the 90% extraction limit for a state-of-the-art analytical method [31]. This improvement to the maximum concentration comes while also achieving improved positioning accuracy (error of <3.2 voxels). Validating the proposed method on experimental holograms with a concentration of 0.020 ppp results in an extraction rate over 90% with a positioning error of 5.24 voxels (13.1 μm) for the depth measurement.

These assessments demonstrate the unique power of machine learning for particle hologram reconstruction with a broad range of particle concentrations. Finally, our learning-based hologram reconstruction is more than 30 times faster than the analytical RIHVR method, even though minimal effort has been undertaken to optimize the speed of the current approach.

Future Goals

We suggest that the proposed method can be generalized for other sparse field imaging tasks such as imaging individual brain neuron activities, particle localization with synthetic aperture or defocusing imaging, and imaging through diffusive media.While our machine learning approach has equal or superior performance compared to state-of_x0002_the-art conventional hologram reconstruction methods, there remains room for improvement.

Ongoing work aims to synthesize holograms with sufficient fidelity to train a model suitable for processing experimental images. Such an approach can substantially reduce the cost of collecting ground truth measurements and has been proven effective for image classification tasks and 2D shadow image particle segmentation . Additionally, although our learning-based approach has achieved significant speed improvement in comparison to conventional reconstruction methods, more than 90% of our total processing time is consumed in pre- and post-processing steps.

Code for 3D computer generated holography using deep learning

```
#importing libraries
import tensorflow as tf
import os
import numpy as np
import matplotlib.pyplot as plt
from time import time
!git clone https://github.com/Pushker-stark/DeepCGH.git
os.chdir('DeepCGH')
from deepcgh import DeepCGH_Datasets, DeepCGH
from utils import GS3D, display_results, get_propagate

retrain = True
coordinates = False

data = {'path' : 'DeepCGH_Datasets/Disks',# path to store the data
        'shape' : (512, 512, 3),# shape of the holograms. The last dimension determines the number of depth planes
        'object_type' : 'Disk',# shape of the object in simulated images, can be disk, square, or line
        'object_size' : 10,# has no effect if object type is 'Line'
        'object_count' : [27, 48],# number of random objects to be created
        'intensity' : [0.2, 1],# the (range of) intensity of each object. If a range is specified, for each object the intensity is randomly determined
        'normalize' : True,# if the data is 3D, it'll normalize the intensities from plane to plane (see manuscript fot more info)
        'centralized' : False,# avoids putting objects near the edges of the hologram (useful for practical optogenetics applications)
        'N' : 100, # number of sample holograms to generate
        'train_ratio' : 90/100,# the ratio of N that will be used for training
        'compression' : 'GZIP',# tfrecords compression format
        'name' : 'target',# name of the dictionary that contains the targets (leave as "target" if you're not changing the structure of network input)
        }

dset = DeepCGH_Datasets(data)
dset.getDataset()
```

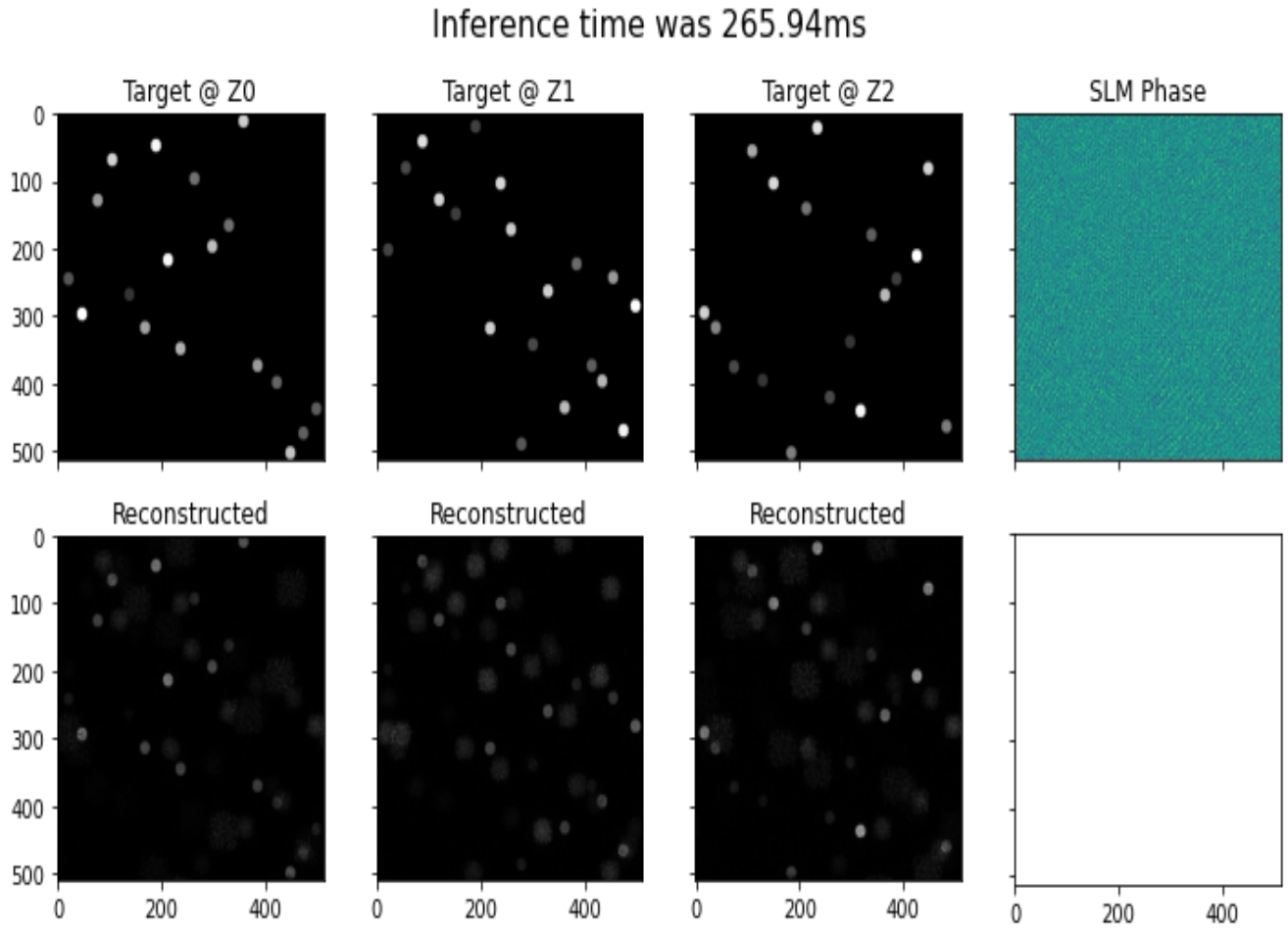
```
model = {'path' : 'DeepCGH_Models/Disks',# the path to the saved model
        'plane_distance':0.005,# the physical distance between depth planes when we're doing 3D holography
         # physical setup parameters
        'wavelength':1e-6,# the wavelength of the laser (both simulations and experiments)
        'pixel_size':0.000015,# size of the SLM pixel sizes
        'int_factor':16,# the interleaving factor
         # CNN model and training parameters
        'n_kernels':[ 64, 128, 256],# the number of kernels in the U-Net model (see the manuscript)
        'input_name':'target',# name of the input layer in the U-Net model
        'output_name':'phi_slm',# name of the output layer
        'lr' : 1e-7,# learning rate of the optimizer
        'batch_size' : 8,
        'epochs' : 1,
        'shuffle' : 8,# determine how many samples are going to be shuffled
        'token' : '64',# string to be attached to the name of the model to differentiate it from similar models
        'max_steps' : 5000 # maxmimum number of batches/steps to be processed
        }

dcgh = DeepCGH(data, model)
dcgh.train(dset)

image = dset.get_randSample()[np.newaxis,...]
# making inference is as simple as calling the get_hologram method
phase = dcgh.get_hologram(np.zeros_like(image)) # the very first inference takes a long time (a known tensorflow characteristic)

t0 = time()
phase = dcgh.get_hologram(image)
t = time() - t0
```

```
#Now we simulate the hologram that this phase will generate:
propagate = get_propagate(data, model)
reconstruction = propagate(phase)
display_results(image, phase, reconstruction, t)
```



The code can be accesed through this link:(copy paste this url):
(https://colab.research.google.com/drive/1m3bbqukiqa7ilKAAgYL9aysVVqIzvJUi?usp=sharing#scrollTo=AbnUOTwn358U)

Acknowledgements

I here would like to thank Prof.Rakesh Kr. Singh Sir for providing me such a task to explore the further areas related to optics .

Also to learn such a advanced technologies like Deep learning,Computer Vision and its applications on a field related to optics.And to guide me to research on them by supervising us on our Exploratory Project.

References

- Research Paper on Machine learning holography for 3D particle field imaging by Siyao Shao,Kevin Mallery,Santosh Kumar,And Jiarong Hong.
- Machine learning and Deep learning by Andrew Ng on coursera.
- Recognizance machine learning event workshops done by IIT (BHU) Varanasi Electrical Dept.
- Udyam'21 workshops and events of Mossaic & Cassandra which is based on computer vision(CNN) and Data Science,conducted by IIT (BHU) Varanasi Electronics Dept.

Thankyou