Volumetric reconstruction of flow particles through light field

particle image velocimetry and deep neural network

Xiaoyu Zhu (朱效宇),[1] Mengxi Fu (傅梦希),[1] Chuanlong Xu (许传龙),[1,a] Md. Moinul Hossain,[2] Boo Cheong Khoo[3]

[1] National Engineering Research Center of Power Generation Control and Safety, School of Energy and Environment, Southeast University, Nanjing 210096, China

[2] School of Engineering, University of Kent, Canterbury, Kent, CT2 7NT, UK

[3] Department of Mechanical Engineering, National University of Singapore, 119260, Singapore

[a]Authors to whom correspondence should be addressed: *chuanlongxu@seu.edu.cn*

Abstract: Tomographic reconstruction of three-dimensional tracer particle distributions through light field particle image velocimetry (LF-PIV) faces challenges in low reconstruction resolution owing to the elongation effect and extensive computational cost incurred by the iterative process. To resolve these challenges, this study proposes a deep neural network-based volumetric reconstruction approach to alleviate the reconstruction elongation and enhance the reconstruction efficiency. A tailored deep learning model (namely LF-DNN) incorporating residual neural network architecture and a novel hybrid loss function is established to reconstruct the particle distributions through LF images. The parallax information of the flow field decoded from the raw LF data is leveraged as the input features of the network model. Comparative studies between the proposed method and the traditional tomographic reconstruction algorithms (multiplicative algebraic reconstruction technique, MART and pre-recognition MART, PR-MART) are performed through synthetic datasets. Experiments on a cylinder wake flow are further conducted to validate the performance of the proposed LF-DNN. The results indicate that the LF-DNN outperforms MART and PR-MART in terms of the reconstruction quality, mitigation of elongation effect and noise resilience. The LF-DNN also improves the reconstruction efficiency which is 9.6 and 7.1 times higher than the MART and PR-MART, respectively. The relative error of the cylinder wake flow achieved by the LF-DNN is 2% lower than the MART. It suggests that the LF-DNN can facilitate accurate volumetric particle reconstruction and hence the three-dimensional flow measurement by single camera-based LF-PIV.

Keywords: Three-dimensional flow measurement, Light field PIV, Deep learning neural network, Tomographic reconstruction, Tracer particle distribution

1. Introduction

Volumetric particle image velocimetry (PIV) has now emerged as a powerful technique for characterizing three-dimensional (3D) unsteady flow structures [1,2]. The fully three-dimensional three-component (3D-3C) flow field can be acquired by employing 3~6 cameras to record the flow structures from different viewing angles [3]. The multi-camera-based volumetric velocimetry techniques provide high spatial resolution and are demonstrated in investigating the fine flow structures of turbulent flows [4-6]. Nonetheless, these techniques are constrained in industrial internal flows (e.g., tip-leakage flow [7]) due to limited optical access. Recently, the single-camera-based light field PIV (LF-PIV) offers a potential solution to the space-constraint flow measurements [8,9]. Incorporating a dense microlens array (MLA) into a conventional camera (light field camera, LFC) allows the capture of a complete 3D-3C flow field via a single perspective.

A key process of the LF-PIV is to retrieve the 3D tracer particle distributions from the recorded LF images, which involves solving an inverse problem [10]. Conventionally, this can be tackled by tomographic reconstruction algorithms [11] incorporating mathematical constraints (e.g., maximum entropy, minimum norm, etc.) to acquire an optimal solution. However, the tomographic reconstruction suffers lower spatial resolution due to the severe elongation effect [12]. This is mainly attributed to the finite viewing angle of the camera lens, leading to a lower imaging resolution along the optical axis direction [13]. The elongation effect is also demonstrated in the holographic PIV [14] where a single camera is employed to record the particle holograms. Various algorithms have been proposed to mitigate the elongation issue. Some approaches attempt to formulate an initial estimation of the particle distribution as a priori to remove the reconstruction elongation [15,16]. Others incorporate the spectral information of the light source to enhance the spatial resolution of LF imaging, thereby alleviating the elongation [17]. Post-processing based on line-of-the-sight has been developed to filter the elongation [18]. Although these approaches can partially suppress the elongation effect, their efficacy significantly degrades with seeding densities and calibration errors. Efforts have also been made to expand the viewing angle of the imaging system by adding an extra LFC [19,20]. This strategy has shown the potential to eliminate the elongation effect at high seeding densities, albeit at the cost of generating ghost particles. Moreover, the extra LFC makes the experimental apparatus complex and necessitates greater optical access.

In addition to the issue of reconstruction elongation, the computational cost is another challenge in tomographic reconstruction. The operation of a large sparse weight matrix always takes a longer time and demands a larger computer storage space. Although advanced initialization strategies [e.g., multiplicative first guess (MFG) [21], multiplicative line-of-the-sight (MLOS) [22]] have been incorporated to either reduce the iteration number or to accelerate the iteration process, disparity persists in attaining real-time volumetric velocimetry. Recently, the graphics processing unit (GPU) has been used to expedite the particle reconstruction process [23] by utilizing a fine-grained parallelism strategy where the reconstruction task is divided into numerous sub-tasks and executed on thousands of GPU threads. This approach demonstrated the capability of acceleration ratio over 20 times faster compared to the conventional CPU-based methodologies. Nonetheless, the finite memory capacity of GPUs requires more memory-efficient strategies to accommodate the weight matrix.

The expeditious development of deep learning (DL) approaches presents a promising path for addressing the challenge of 3D particle reconstruction from a finite set of 2D projections [24-27]. The DL-based methods have been applied to the realm of computational imaging and PIV for

achieving volumetric reconstruction [28,29]. Specifically, in the tomographic PIV, a 3D convolutional neural network (CNN) incorporating the geometrical information of the imaging system was developed to refine the coarse reconstruction produced by the traditional tomographic reconstruction algorithms [30]. The deep CNN shows the superiority in terms of reconstruction accuracy, efficiency and robustness to noise over the traditional tomographic algorithms. Moreover, CNN is employed as a filter to remove ghost particles produced by tomographic algorithms and regularize anomalous shapes of the particles [31]. The DL methods were also applied in holographic PIV where a modified U-net architecture was utilized to retrieve the sizes and positions of tracer particles from holograms [32]. The U-net facilitates a higher particle extraction ratio and enhanced particle positioning accuracy in contrast to the analytical methods [33], particularly in dense seeding scenarios. Additionally, an integration of the physical models (e.g., free-space point spread function) into the neural networks was considered to augment the generalizability of the DL methods and diminish the requisite volume of training data [34]. Recently, a 3D U-net neural network was proposed for the tracer particle reconstruction in LF-PIV [35]. By utilizing a digital refocusing algorithm, a stack of LF-refocused images of particles was generated as a coarse reconstruction, which was then refined using the 3D U-net model. Experimental results show that the reconstruction time with the 3D U-net is ~10 times faster than that of the tomographic reconstruction algorithms.

While the ability of DL strategies has been demonstrated to tackle the inverse problem of 3D particle reconstruction, some challenges remain when these methods are applied to the LF-PIV. For instance, (*i*) unlike conventional imaging where the scene is captured by a single lens, the incorporation of a MLA increases the complexity of the imaging model. Consequently, deploying neural networks to understand the underlying physics of LF imaging may fail. A potential solution to this issue involves extracting the perspective information encoded in the LF raw data and integrating it into the neural network models as prior knowledge, (*ii*) the LF-PIV features a relatively sparse seeding concentration, leading to most discrete voxels having a zero intensity. Therefore, volumetric particle reconstruction is a task of sparse target reconstruction, which necessitates a tailored network architecture and loss function to prevent overfitting and instability during the model training process and (*iii*) the current DL-based approaches usually require a coarse guess of particle distribution, typically obtained through traditional tomographic algorithms. End-to-end methods that are capable of directly predicting the particle distributions from LF images remain to be explored to improve the reconstruction efficiency.

To address these above challenges, this study proposes a DL-based framework for achieving an accurate volumetric reconstruction of particle distributions in LF-PIV. A tailored DL model (named LF-DNN) integrating the residual network structure and a novel hybrid loss function is employed to reconstruct the particle distributions seeded in the flow field. The parallax information of the flow field decoded from the LF raw data is leveraged as the geometrical input features of the network. Synthetic tests of particle reconstructions and experiments on a 3D cylinder wake flow are conducted to assess the performance of the LF-DNN. Comprehensive comparative analysis between the proposed method and traditional tomographic reconstruction algorithms along with mitigation of elongation effect, reconstruction quality and efficiency, and robustness to noise are performed. The principle and implementation details of the LF-DNN are given and findings from the synthetic and experimental tests are presented and discussed.

3

2. Principle of tomographic LF-PIV

In computational photography, the propagation of the light rays in the free space can be described by a 4D plenoptic function, encompassing 2D spatial and 2D angular coordinates of rays. A full capture of the plenoptic function allows the recovery of a 3D scene of interest, but it is challenging through a conventional camera because the angular information of the incoming rays is lost owing to the focusing of the lens unit. To capture a more complete picture of plenoptic function, a MLA composed of densely packed microlens with μm-level pitch is inserted in front of the image sensor to modulate the incoming rays. The incoming rays are initially focused on the MLA through the main lens and are further redirected by the microlens to different image pixels. Consequently, each pixel records the intensity of rays emanating from a particular direction. The LFC can therefore preserve both the spatial and angular information of rays, which contains the 3D spatial information of the scene.
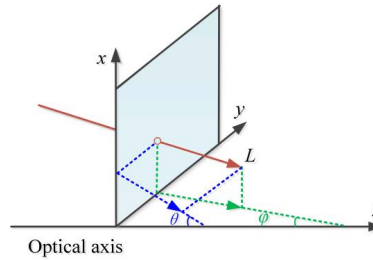
Fig. 1 Parameterization of the LF through a 4D plenoptic function $L(x, y, \theta, \varphi)$

In computational photography, the propagation of the light rays in free space can be described by a 4D plenoptic function $L(x, y, \theta, \varphi)$, which encompasses the 2D spatial and 2D angular coordinates of rays. As illustrated in Fig. 1, (x, y) and (θ, φ) define the intersection of the ray with a plane perpendicular to the optical axis and the propagation direction of the ray, respectively. Capturing the full plenoptic function enables the recovery of a 3D scene of interest, however, this is challenging with a conventional camera because the main lens focuses the incoming rays, leading to the loss of angular information, as shown in Fig. 2(a). To capture a complete plenoptic function, a microlens array (MLA) composed of densely packed microlenses with μm-level pitch is inserted in front of the image sensor to modulate the incoming rays. Depending on the distance between the MLA and the image sensor, the LFC can be classified into unfocused and focused configurations. In this study, the unfocused type is employed where the incoming rays are initially focused on the MLA through the main lens and are further redirected by the microlens to different image pixels, as illustrated in Fig. 2(b). Consequently, each pixel records the intensity of rays emanating from a particular direction. The LFC can therefore preserve both the spatial and angular information of rays, which contains the 3D spatial information of the scene. The propagation trajectory of the light ray can be described using a ray transfer matrix based on the principle of geometric optics, as

$$
\begin{pmatrix} x' \\ y' \\ \theta' \\ \phi' \end{pmatrix} = \overbrace{\begin{bmatrix} 1 & 0 & s_\mu & 0 \\ 0 & 1 & 0 & s_\mu \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}^{\text{MLA to sensor}} \left(\overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1/f_m & 0 & 1 & 0 \\ 0 & -1/f_m & 0 & 1 \end{bmatrix}}^{\text{Through MLA}} + \begin{pmatrix} 0 \\ 0 \\ s_x/f_m \\ s_y/f_m \end{pmatrix} \right) \overbrace{\begin{bmatrix} 1 & 0 & s_i & 0 \\ 0 & 1 & 0 & s_i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}^{\text{Main lens to ML}} \overbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ -1/F_m & 0 & 1 & 0 \\ 0 & -1/F_m & 0 & 1 \end{bmatrix}}^{\text{Through main lens}} \overbrace{\begin{bmatrix} 1 & 0 & s_o & 0 \\ 0 & 1 & 0 & s_o \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}^{\text{Object to main lens}} \begin{pmatrix} x \\ y \\ \theta \\ \phi \end{pmatrix} \quad (1)
$$

where $(x', y', \theta', \varphi')$ is the 4D coordinate of the light ray at the image sensor plane. The parameters s_o and s_i denote the object and image distances, respectively. F_m and f_m are the focal lengths of the main lens and microlens, respectively. s_x and s_y represent the distances from the microlens center to the optical axis in the x and y directions. s_μ is the distance from the MLA to the image sensor.
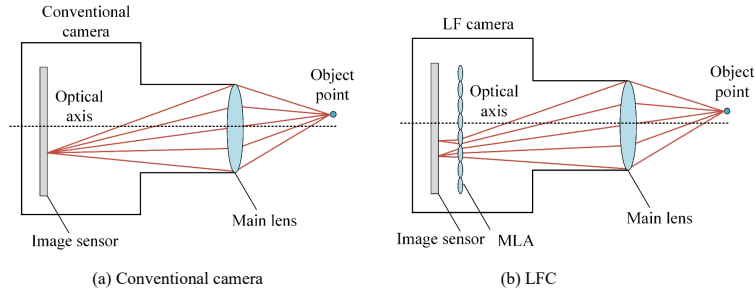

(a) Conventional camera   (b) LFC

Fig. 2 Principle of conventional and LF imaging. (a) Conventional camera. (b) LFC.

Leveraging the ability to capture the 3D flow field, the LFC employs to PIV technique to measure 3D flow fields. In LF-PIV, a crucial step is to reconstruct the tracer particle distribution from the LF projection images for flow velocity estimation. The particle reconstruction can be solved by a linear equation and defined as

$$I_i = \sum_{j=1}^{N_i} w_{i,j} E_j \tag{2}$$

where E_j and I_i denote the intensity of j_{th} discrete voxel and i_{th} image pixel, respectively, N_i is the number of voxels whose intensity contributes to the pixel, $w_{i,j}$ denotes the weight coefficient describing the intensity contribution of j_{th} voxel to i_{th} pixel and can be obtained through volumetric calibration techniques. Eq. (2) is typically underdetermined and the tomographic reconstruction algorithm incorporating mathematical constraints such as the minimum norm, the maximum entropy, the least absolute shrinkage and selection operator regularization can be utilized to acquire a unique solution. Various tomographic reconstruction algorithms are developed to solve Eq. (2). Among them, the multiplicative algebraic reconstruction technique (MART) employing the maximum entropy strategy has been demonstrated as a standard algorithm. The details of MART can be found elsewhere in [36].

A major challenge in the tomographic reconstruction of flow particles is the elongation effect. Specifically, spherical tracer particles are distorted into ellipsoidal shapes during the reconstruction, with elongations predominantly along the optical axis (i.e., depth direction). This effect is mainly attributed to the finite aperture of the camera lens, which restricts the viewing angle and consequently leads to a relatively lower spatial resolution in the depth than in lateral directions. The elongation effect degrades the estimation accuracy of the flow velocity, particularly in the depth velocity component [13]. To tackle this issue, efforts have been made to optimize the tomographic reconstruction algorithms. In particular, a pre-recognition (PR) method based on the line-of-the-sight technique is proposed to eliminate the potential zero-intensity voxels in the measurement volume [16]. This PR method is then integrated with the MART to alleviate the elongations and accelerate the reconstruction. Nonetheless, the efficacy of the PR method gradually decreases when

5

the particle concentration increases.

Figure 3(a) illustrates the random particle distribution generated via simulation at a seeding concentration (C) of 0.1 particles per microlens (ppm). The reconstructed particle volumes achieved by the MART and PR-MART are shown in Fig. 3(b) and (c), respectively. It can be observed that the particles with the actual size of 0.1 mm are elongated to 1 mm for the MART reconstruction. While PR-MART succeeds in partially alleviating this elongation, the length of the reconstructed particle remains four times larger than the true size. Therefore, further study is imperative to resolve the elongation issue in the reconstruction of flow particles. This study introduces a deep learning-based volumetric reconstruction approach to mitigate the particle elongation effect of single-camera LF-PIV.

(a) True particle distribution
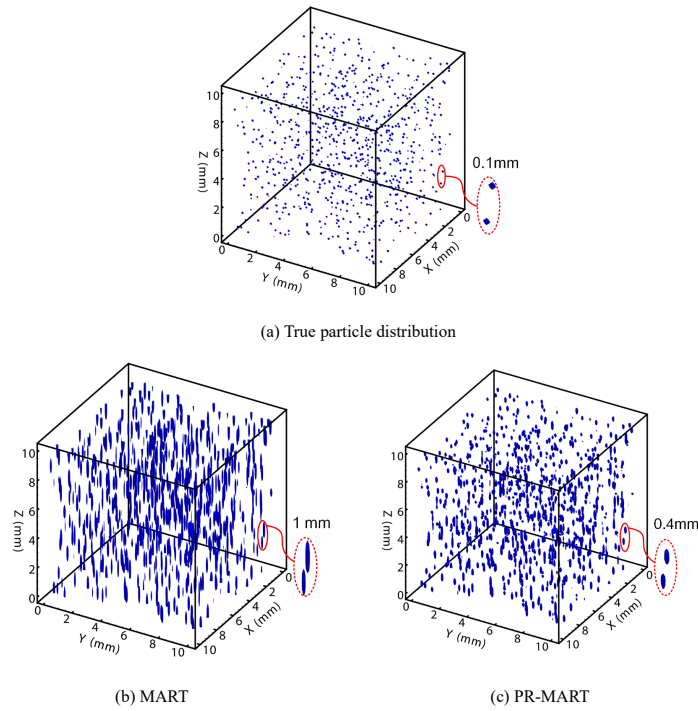
(b) MART   (c) PR-MART

Fig. 3 An illustration of the reconstructed particle distribution along with the elongation effect achieved through the MART and PR-MART. (a) True particle distribution. (b) MART. (c) PR-MART.

3. Deep neural network-based particle reconstruction

3.1 Network architecture

The deep learning (DL) method is an alternative to the tomographic reconstruction algorithms for the 3D particle reconstruction. Instead of relating the voxels to pixels through the weight matrix, the mapping between the 3D particle volume and 2D particle image is established by utilizing a supervised deep neural network (DNN). The DNN architecture consists of multiple hidden layers

between the input and output layers [37]. It can be a variety of architectures including fully connected layers, convolutional layers, recurrent layers, etc. Among them, the convolutional layers are specially designed for processing structured array data, such as images. A DNN with convolutional layers can automatically and adaptively learn the features of spatial hierarchies of input images [38]. This allows the input images to be transformed into new output images with some specific features preserved while others are altered or enhanced.

In general, deeper networks with more layers have the potential for higher accuracy because they can learn more complex features [39]. However, such networks often encounter the vanishing gradient problem [40], where the gradient of the loss function decreases exponentially when it propagates backwards from the output layer to the input layers. To mitigate this issue, residual blocks are incorporated into the DNN. The structure of a residual block is shown in Fig. 4. It is composed of primary and secondary convolutional layers. A shortcut connection that bypasses one or more layers is incorporated to enhance the flow of gradients throughout the network [41]. The underlying principle of a residual layer can be mathematically described as,

$$y = H(x, w_h) + x \tag{3}$$

where x is the input, $H(x, w_h)$ represents the non-linear operations and y is the output. The non-linear operations are executed by two convolutional layers, each utilizing 3×3 kernels.
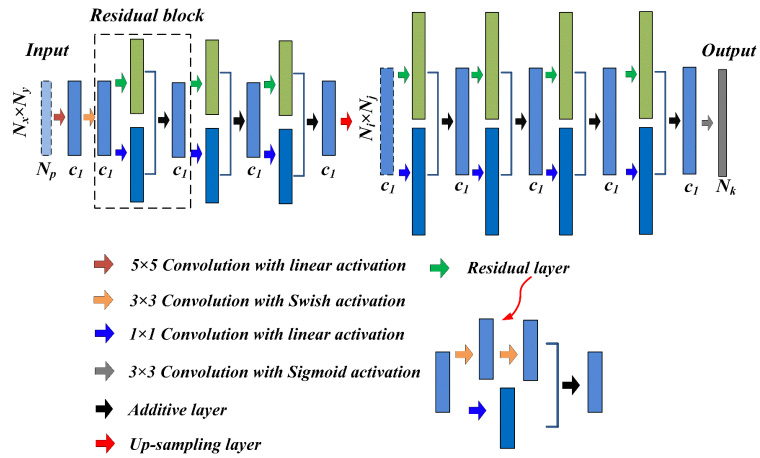

Fig. 4 Architecture of the proposed DNN model for the 3D reconstruction of flow particle distribution.

The architecture of the proposed DNN (named LF-DNN) is illustrated in Fig. 4. The perspective-shift images containing the geometry-informed features of the flow field are taken as the input of the network. The training target is a good approximation of the 3D particle distribution. The size of the input image is $N_x\times N_y$, where N_x and N_y represent the resolution of sub-aperture images. The channel of the input image is the number of viewing perspectives N_p (sub-apertures). The spatial information of input images is encoded into feature maps through the encoding part of the LF-DNN, which is mainly composed of multiple convolutional layers and residual blocks. The feature maps are subsequently scaled up by an up-sampling layer and further decoded by the

decoding part of the network, which consists of several residual blocks. The output particle volume is generated with the dimensions of $N_i \times N_j \times N_k$.

3.2 Novel hybrid loss function

For a DNN model, the loss function is defined to quantify the error of a model's predictions, serving as crucial feedback for the training process. Specifically, the gradient of the loss function indicates the direction in which the network parameters should be adjusted and is calculated using the gradient descent optimizer. The primary objective of the training process is to fine-tune the network parameters to minimize the loss. A frequently used loss function for regression tasks is the mean squared error (MSE), which quantifies the average squared discrepancy between predicted outcomes and actual values. However, initial tests suggest that the MSE might not be adequate for tasks involving sparse target reconstruction. The limitation of the MSE lies in its uniform penalization of all errors, irrespective of their scale, rendering it less effective in situations where certain errors carry more significance. This issue becomes apparent in scenarios with sparse seeding concentration, where a majority of discrete voxels in the measurement volume have zero intensity. Under this condition, a model that predicts zero intensity across all voxels can achieve a lower MSE. This leads to the model mixing the non-zero signals for noise and ignoring them during the learning process. Consequently, a significant divergence between the reconstructed and actual particle distributions can be presented.

To tackle this issue, a novel hybrid loss function is designed for the LF-DNN. The overall loss arises from two factors, i.e., the weighted mean absolute error ($MAE_{weighted}$) and reconstruction quality factor error (Q). The $MAE_{weighte}$ is defined as,

$$MAE_{weighted} = \frac{1}{N}\left(\sum_{i=1}^{N-N_n} \left| y_{true}^i - y_{pred}^i \right| + \sum_{j=1}^{N_n} w_n \left| y_{true}^j - y_{pred}^j \right| \right) \qquad (4)$$

where N_n is the number of non-zero intensity voxels, N is the total voxel number, y_{true} and y_{pred} denote the true and predicted voxel intensity, respectively, and w_n denotes the weighting given to errors arising from the voxels with non-zero intensity. As aforementioned, the non-zero intensity signals are of more significance for network learning, therefore an adaptive weight coefficient larger than 1 is assigned to the non-zero voxels. Specifically, a maximum weight of 100 is applied for the sparse seeding of $C = 0.1$ ppm, while a weight of 5 is employed for the seeding of $C = 1$ ppm.

The metric Q assesses the similarity between the reconstructed and true images [3]. It is defined as,

$$Q = \frac{\sum_{k=1}^{N} y_{true}^k \cdot y_{pred}^k}{\sqrt{\sum_{k=1}^{N} \left(y_{true}^k \right)^2 \cdot \sum_{k=1}^{N} \left(y_{pred}^k \right)^2}} \qquad (5)$$

A high similarity between the reconstructed and actual intensity distributions yields a Q value close to 1, while a Q value of 0 indicates different particle distributions. The total loss is formulated as the sum of $MAE_{weighted}$ and Q losses, expressed as,

$$Loss = MAE_{weighted} + \alpha \left(1 - Q \right) \qquad (6)$$

where α denotes the weighting of the Q loss and is optimized as a constant of 2. In this hybrid loss function, the $MAE_{weighted}$ and $1-Q$ serve as the standard and custom loss components, respectively.

The inclusion of 1-Q enhances the particle reconstruction accuracy because it aligns with the particle reconstruction task perfectly. However, relying solely on the Q loss could potentially lead to model overfitting, as this custom loss function is highly specific to certain datasets. The integration of $MAE_{weighted}$ helps to reduce this risk. The $MAE_{weighted}$ exhibits strong robustness against outliers and thus can ensure a balanced and effective guidance for the network's training process.

3.3 Data generation

(a) 0.1 ppm   (b) 0.5 ppm   (c) 1 ppm

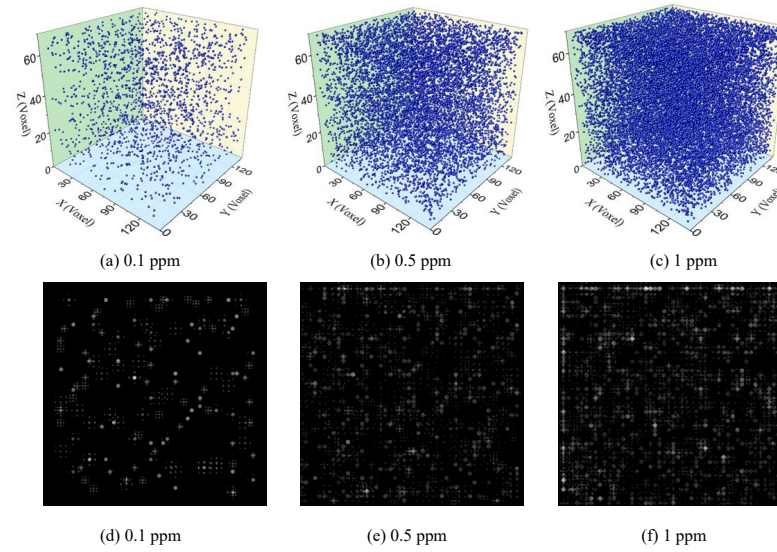(d) 0.1 ppm   (e) 0.5 ppm   (f) 1 ppm

Fig. 5 Random particle distributions [(a), (b), (c)] at different C and corresponding synthetic LF images [(d), (e), (f)] are generated for a training dataset.

    The proposed LF-DNN is trained using a dataset that comprises 3D particle distributions alongside their LF images. Since the true particle distributions are difficult to acquire in real-world experiments, thus synthetic data with labeled particle locations is employed to construct the training dataset. Random particle fields are generated using the particle generator described in [42] and corresponding LF images are generated through a ray tracing technique based on geometrical optics. The C varies from 0.1 to 1 ppm (as illustrated in Fig. 5), which reflects the typical seeding concentrations usually observed in the LF-PIV experiments. The intensity of synthetic LF images is normalized to a scale of 0 to 1. With the consideration of training performance and computational cost, a total of 20000 random particle distributions at different C accompanied by their LF images are generated. The resolution of generated grayscale LF raw images is 3001×3001 pixels. The dataset is split into three subsets for training, validation and testing with a ratio of 70%, 15% and 15%. Data augmentation is performed to enhance the noise resistance of the model and prevent it from overfitting. This includes the addition of Gaussian noise to the LF raw images. The mean value of Gaussian noise (μ) varies from 0.002 to 0.02 and a large μ indicates a high noise level.

    Unlike the DL frameworks that utilize the PIV raw images as direct inputs to neural networks [43], this study employs a preprocessing step to decode the LF raw images before model training.

This is because the MLA increases the complexity of the imaging model, making it challenging to solely rely on a neural network to fully capture the underlying physics. Consequently, to develop a more robust model that can effectively learn how light rays are modulated by the MLA, it is crucial to integrate additional physical insights into the training process. To achieve that, the LF decoding is carried out to extract the perspective shift images from the LF raw data.
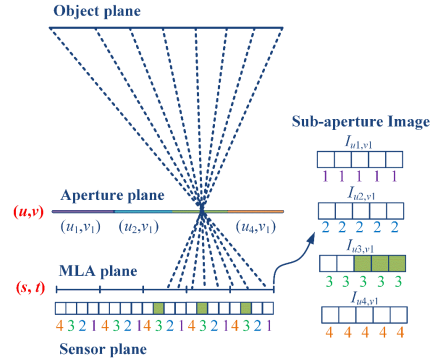
(a) Two-plane parametrization of light field and principle of LF decoding
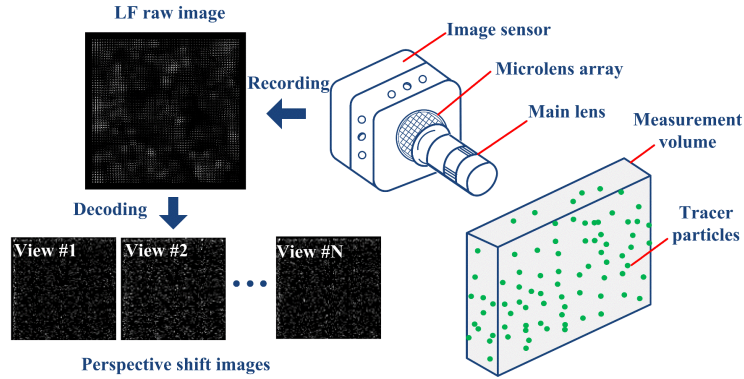
(b) Extraction of sub-aperture images of test flow field from LF raw images

Fig. 6 Decoding of raw LF data for acquiring parallax-informed network input features. (a) Two-plane parametrization of light field and principle of LF decoding. (b) Extraction of sub-aperture images of test flow field from LF raw images.

As depicted in Fig. 6(a), the main aperture is conceptualized as consisting of multiple sub-apertures, each offering a unique viewing angle. By extracting the images captured by these different sub-apertures, parallax information concerning the particle field is obtained. This parallax information is more readily interpretable and learnable by the neural network compared to the unprocessed raw data. The decoding process of a LF raw image into sub-aperture images is illustrated in Fig. 6(a). The 4D plenoptic function $L(x, y, \theta, \varphi)$, which describes the light rays in a scene, can also be expressed as $L(s, t, u, v)$ using the two-plane parameterization. Here, (s, t)

represents the spatial coordinates on the image plane, and (u, v) represents the angular coordinates on the lens plane. Mathematically, a sub-aperture image can be obtained by fixing the angular coordinates (u, v) and varying the spatial coordinates (s, t), as

$$I_{u,v}(s,t) = L(s,t,u,v) \tag{7}$$

where $I_{u,v}(s,t)$ is the sub-aperture image corresponding to the angular position (u, v). By selecting different pairs of (u, v) coordinates, a set of sub-aperture images can be generated, each corresponding to a different perspective of the scene [as shown in Fig. 6(b)]. These images represent the scene as viewed from different angles, simulating a perspective shift, as

$$I_{u_i,v_i}(s,t) \qquad i = 1, 2, ..., n_p^2 \tag{8}$$

where n_p is the number of pixels covered by each microlens. The resolution of a sub-aperture image is $n_x \times n_y$, where n_x and n_y denote the number of microlens in the x and y-directions. For the synthetic LF image generation, n_x and n_y are 251, and n_p is calculated to be 18. Therefore, the size of the LF sub-aperture image, which serves as the input of LF-DNN, is 251×251 pixels. In the dataset, the total number of sub-aperture images is 6,480,000.

3.4 Model establishment

To refine the model parameters and avoid local minima, the cosine annealing strategy [44] is employed for dynamically modulating the learning rate during the training process. The learning rate schedule of cosine annealing is depicted in Fig. 7(a). To optimize the hyperparameters of the DNN model, comparative analyses are conducted on the loss associated with different initial learning rates, batch sizes and optimizers. The variations of the validation loss of the DNN model are illustrated in Figs. 7(b)-(d). It can be observed that using the Adam optimizer [45] with an initial learning rate of 0.0001 and a batch size of 32 yields the lowest loss.

To justify the proposed network architecture, the validation losses of the network with and without residual blocks are calculated and shown in Fig. 8(a). After 50 training epochs, the validation loss is 0.001 and 0.01 with and without residual blocks, respectively, demonstrating that the residual structure improves the training performance significantly. Fig. 8(b) illustrates the loss at different numbers of residual blocks, where the legend 'E1' indicates one residual block in the encoding and 'D2' indicates two residual blocks in the decoding. As shown in Fig. 8(b), the optimal architecture features contain three residual blocks in the encoding part and four in the decoding part. The DNN with additional residual blocks degrades the performance. Because the network learns excessive noise features from the training data. Fig. 8(c) depicts the loss at different numbers of convolutional kernels. It can be found that increasing the kernel count from 64 to 256 enhances the training performance, whereas further increasing it to 512 does not affect the training performance. Thus, the optimal network structure is obtained based on 7 residual blocks with 256 convolutional kernels.

(a) Variation of learning rate with training epochs

(b) Losses at different initial learning rates

(c) Losses at different batch sizes

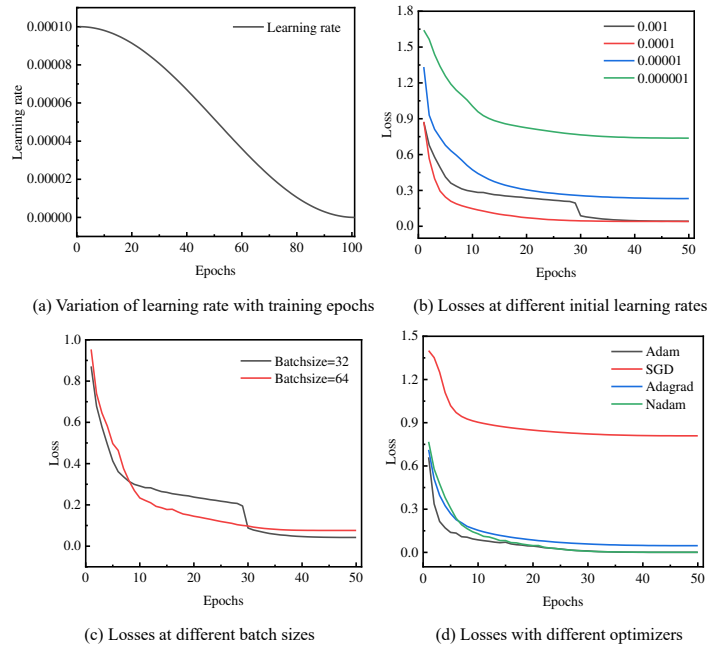(d) Losses with different optimizers

Fig. 7 Validation loss of the proposed DNN model at different model hyperparameters. (a) Variation of learning rate with training epochs. (b) Losses at different initial learning rates. (c) Losses at different batch sizes. (d) Losses with different optimizers.

The Sigmoid-weighted linear unit (Swish) is chosen as the activation function for all layers except the final layer. Swish is distinguished by its non-monotonic and smooth characteristics, which facilitate precise adjustments to the network weights during the optimization process. Additionally, Swish's continuous gradient across zero and negative inputs helps to reduce the occurrence of dying neurons during network training. This feature is especially valuable in applications that involve sparse target reconstruction. The Swish function is defined as,

$$Swish(x) = x \cdot Sigmoid(\beta x) = \frac{x}{1 + e^{-\beta x}} \tag{9}$$

where x represents the input, β denotes a constant that modulates the function's curvature and $Swish(x)$ denotes the output. The Swish activation function normalizes the output to a range of 0 to 1, aligning well with the expected output domain. The validation loss of the DNN model under the Swish and commonly used Rectified Linear Unit (ReLU) functions is shown in Fig. 8(d). The result shows that the loss with Swish is approximately 2.5 times smaller than the ReLU, demonstrating that Swish is more effective for sparse particle reconstruction. The architecture and optimized hyperparameters of the DNN model are summarized in Table 1.

12

(a) Losses with and without residual blocks

(b) Losses at different numbers of residual blocks

(c) Losses at different numbers of convolutional kernels   (d) Losses at different activation functions
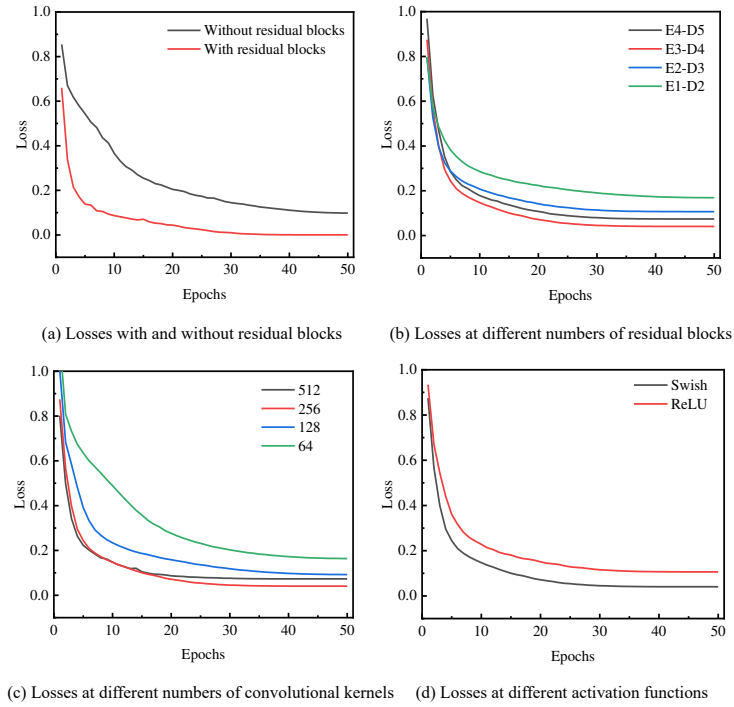
Fig. 8 Variation of validation loss of the DNN model with training epochs at different model architectures. (a) Losses with and without residual blocks. (b) Losses at different numbers of residual blocks. (c) Losses at different numbers of convolutional kernels. (d) Losses at different activation functions.

Table 1: Architecture and hyperparameters of the proposed deep neural network

Model parameter	Value
Number of residual layers	7
Size of the convolution kernel	256×256
Activation function	Swish
Learning strategy	Cosine annealing
Initial learning rate	0.0001
Optimizer	Adam
Batch size	32
Ratio of Training, validation and testing	70%, 15%, 15%

In contrast to the 3D convolution kernels, 2D kernels are chosen for the DNN model with consideration of memory efficiency and training data requirements. The 2D kernels require fewer parameters and therefore the network is less prone to be overfitting. Besides, the proposed DNN model is adaptable to varying resolutions of input images without compromising the model's performance. This flexibility is attributed to the nature of convolution operations and is further enhanced by the network's exclusion of pooling and fully connected layers. A schematic of the proposed LF-DNN framework for particle reconstruction is shown in Fig. 9.
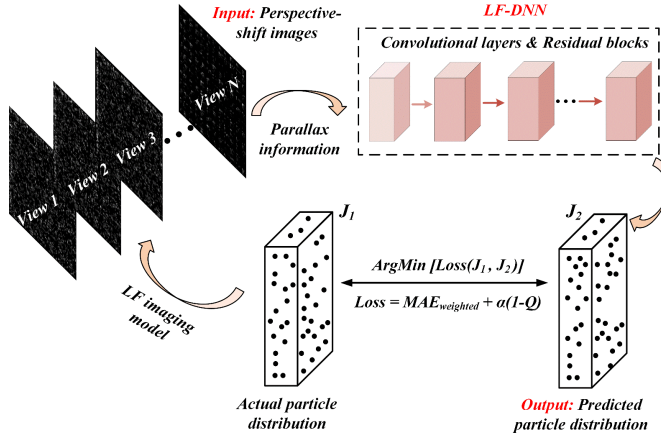
13

Fig. 9 Schematic diagram of proposed LF-DNN model for reconstructing the 3D particle distributions.

3.5 Performance assessment

To verify the performance of the proposed LF-DNN, numerical reconstructions of the 3D particle field are performed by generating synthetic LF images. The 3D measurement volume that contains the particle field has a physical size of 14.1 mm×14.1 mm×7.1 mm and is discretized into the voxels with a resolution of 0.1 mm×0.1 mm×0.1 mm. Particles with a diameter of 10 μm are dispersed randomly in the measurement volume with seeding concentration (C) from 0.1 to 1 ppm with an interval of 0.1 ppm. Corresponding LF images of these particle volumes are synthetically generated using the ray tracing technique. The optical parameters employed in the synthetic image generation are listed in Table 2. The generated LF raw images are subsequently processed through the method described in Section 3.3 to extract perspective shift images, which are then input to the LF-DNN for reconstructing the particle distributions.

Table 2: Optical parameters for the synthetic LF image generation

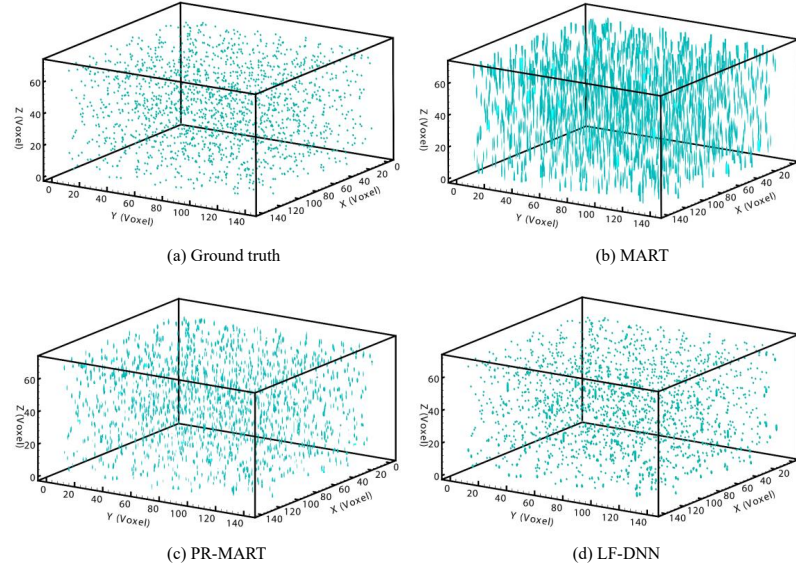| Symbol	Parameter	Value
M	The magnification ratio of the main lens	-1
F_m	The focal length of the main lens	100 mm
$F/\#$	f-number of the main lens	4
f_m	The focal length of the microlens	800 μm
$f/\#$	f-number of microlens	8
-	Microlens geometry	Square
p_m	Microlens pitch	100 μm
n_x, n_y	Number of microlens on MLA	251×251
p_x	Pixel size	5.5 μm
N_x, N_y	Camera resolution	3001×3001

14

Fig. 10 3D particle distributions reconstructed by the MART, PR-MART and LF-DNN at $C = 0.1$ ppm. (a) Ground truth. (b) MART. (c) PR-MART. (d) LF-DNN.

In addition to the LF-DNN, the conventional MART and pre-recognition MART (PR-MART) are also employed to reconstruct the particle distributions for comparative analysis. Figure 10 shows the particle reconstruction achieved by these three methods at $C = 0.1$ ppm. Compared with the actual particle distribution, severe particle elongations along the depth direction are observed for the MART. This effect is to some extent mitigated by the PR-MART, but distortions in particle shape persist. In contrast, the particle distribution reconstructed by the LF-DNN is similar to the actual distribution with a significant reduction of the elongation and a discernible enhancement of spatial resolution of reconstruction.

To assess the reconstruction accuracy quantitatively, the reconstruction quality Q along with the Recall and Ghost rates are calculated under various concentrations (C) and levels of image noise (μ). For the Recall and Ghost rates calculations, the reconstructed particle volumes are initially converted into 3D binary images through the application of an adaptive intensity threshold. Subsequently, distinct sub-regions within the binary images are identified based on a connected-component labeling algorithm [46]. The regional maxima within these sub-regions are identified to determine the location of the reconstructed particles. With known locations of the reconstructed particles, the Recall rate is calculated by,

$$\text{Recall rate} = \frac{\#VD}{T_p} \tag{10}$$

where a valid detection (VD) is characterized by the identification of a reconstructed particle situated within a one-voxel chessboard distance from an actual particle, # symbolizes the count of VD instances, T_p denotes the number of actual particles. The Recall rate is in the range of 0 to 1 and a

15

larger value implies that a greater proportion of particles are being reconstructed at the correct locations. The Ghost rate is calculated by,

$$\text{Ghost rate} = \frac{T_d}{T_p} \tag{11}$$

where T_d denotes the cumulative number of the particles that have been reconstructed. A lower Ghost rate indicates fewer spurious particles generated in the reconstruction result.

Figures 11(a), (c) and (e) illustrate the variations of Q, Recall and Ghost rates with Cs under a noise-free case for the MART, PR-MART and LF-DNN. It can be seen that the LF-DNN outperforms the MART and PR-MART. Specifically, for LF-DNN, the Q index exceeds 0.75 which is almost twice the MART. The Recall rate for LF-DNN remains persistent at 1 for various Cs, indicating that all particles have been reconstructed at their original locations, even in scenarios of high seeding density. Conversely, the Recall rates of the MART and PR-MART decrease with the increasing seeding concentration. Furthermore, the Ghost rate achieved by the LF-DNN is lower than the MART and PR-MART, indicating that the elongation effect can largely be alleviated by the LF-DNN.
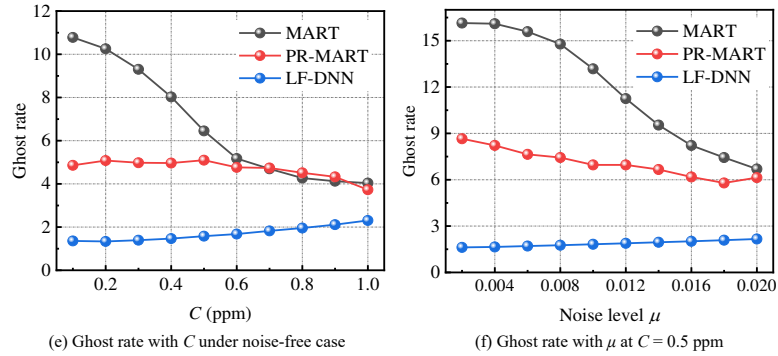

(a) Q with C under noise-free case

(b) Q with μ at $C = 0.5$ ppm

(c) Recall rate with C under noise-free case

(d) Recall rate with μ at $C = 0.5$ ppm

(e) Ghost rate with C under noise-free case   (f) Ghost rate with μ at $C = 0.5$ ppm
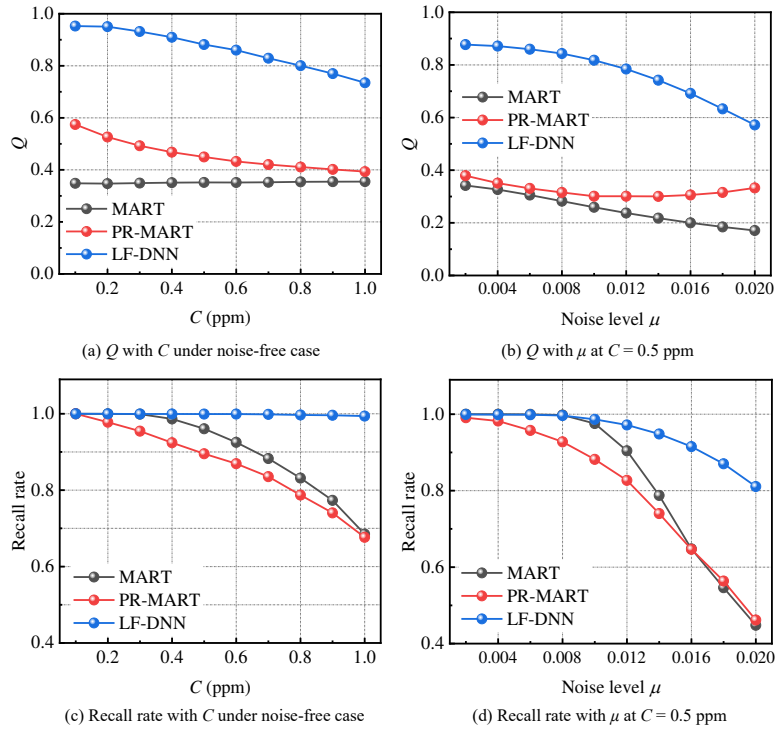
Fig. 11 Comparison of Q, Recall rate and Ghost rate achieved by the MART, PR-MART and LF-DNN under different particle concentrations C and noise levels μ of LF images. (a) Q with C under noise-free case. (b) Q with μ at $C = 0.5$ ppm. (c) Recall rate with C under noise-free case. (d) Recall rate with μ at $C = 0.5$ ppm. (e) Ghost rate with C under noise-free case. (f) Ghost rate with μ at $C = 0.5$ ppm.

Comparative analysis is conducted under different noise levels μ and the results are illustrated in Figs. 11(b), (d) and (f). It is observed that the Q and the Recall rate are reduced with an increase of μ from 0.002 to 0.02, indicating that an increase in noise of the raw LF image deteriorates the particle reconstruction accuracy. In comparison to the MART and PR-MART, the LF-DNN demonstrates superior resilience to noise interference. Notably, at a higher $\mu = 0.2$, the Recall and Ghost rates of 0.81 and 2.16 are observed respectively for the LF-DNN whereas 0.44 and 6.68 are observed for the MART. Note that the higher Recall rate achieved by the LF-DNN is due to the influence of noise that has been considered in the generation of the training dataset.

Computational cost also stands as a critical factor alongside the reconstruction quality and noise resilience. Therefore, the computational demands of these three methods for particle reconstruction are quantified. The computational cost of MART and PR-MART are evaluated on a 12-core CPU (Intel Core i7-13700) and the proposed LF-DNN is deployed on both CPU (Intel Core i7-13700) and GPU (Nvidia RTX 3090 GPU) using the PyTorch backend. The computational costs achieved for these three methods are listed in Table 3. The LF-DNN has a total of approximately 30 million parameters and the storage size of the model is 150 MB.

Table 3: Computational costs of 3D particle reconstruction for the MART, PR-MART and LF-DNN

Computational time	MART	PR-MART	LF-DNN-CPU	LF-DNN-GPU
Weight matrix calculation	2.26 hrs	2.26 hrs	-	
Pre-recognition step	-	0.16 hrs	-	
MART iterations	0.38 hrs	0.12 hrs	-	
Model training	-	-	50.75 hrs	15.92 hrs
Reconstruction/LF image	0.38 hrs	0.28 hrs	3.7 s	1.2 s
Total time cost for 100 image pairs	154.26 hrs	114.26 hrs	51.16 hrs	16.05 hrs

The results show that the MART and PR-MART require 2.26 hours (hrs) to compute the weight matrix, which is not required for the LF-DNN. However, the LF-DNN experiences approximately 50.75 hrs of training. This duration is reduced to 15.92 hours with GPU acceleration but the training cost is still high. The time required for the model training is a common issue for the DL-based

17

methods. In addition to the weight matrix calculation, the reconstruction times required for the MART and PR-MART are 0.38 hrs/LF image and 0.28 hrs/LF image, respectively. In contrast, the LF-DNN takes only 3.7 seconds/LF image on a CPU and 1.2 seconds/LF image on a GPU. For the unsteady flow structures, statistical measurements of the flow velocity field are usually required and several hundred repetitions of the particle reconstruction process are needed. For instance, in the reconstruction of 200 PIV image pairs (400 LF images), the accumulated times required for the MART and PR-MART are 154.26 and 114.26 hrs, respectively, whereas that for the LF-DNN is 51.16 and 16.05 hrs by CPU and GPU. The acceleration ratios achieved by the LF-DNN are 9.6 and 7.1 compared to the MART and PR-MART. This evidences that the LF-DNN requires less computational time compared to traditional iterative reconstruction approaches and thus improves the computational efficiency significantly.

4. Experimental validation

4.1 Experimental setup

The proposed method is further validated by carrying out experiments on a 3D flow around a circular cylinder. As shown in Fig. 12, a cylinder with a diameter of $d = 2$ mm is installed inside a water channel. The channel is 4.6 m long and the size of a cross-section is 20 mm $\times$ 20 mm. The cylinder is positioned at a section where the flow is fully developed. The axis of the cylinder is aligned with the central height of the channel. The flow climbs over the cylinder, forming the wake flow with shedding vortices. The flow rate is controlled by a Proportional-Integral-Derivative (PID) controller, maintaining the Reynolds number Re_d (based on d) of 140, 100 and 70 for different test cases.

(a) Experimental setup   (b) Schematic of the measurement volume
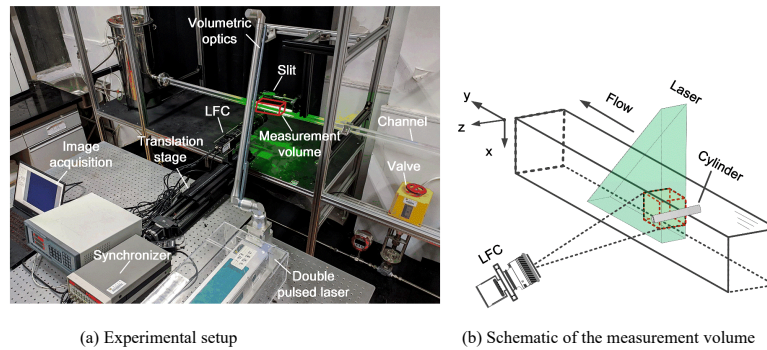
Fig. 12 Experimental setup of the 3D flow around a circular cylinder. (a) Experimental setup. (b) Schematic of the measurement volume.

The 3D velocity field of the wake flow is measured by the LF-PIV incorporating the proposed DNN model. Polyamide particles with a diameter of 10 μm and density of 1.02 g/cm^3 are immersed in the test flow to serve as the tracer particles. The seeding concentration is maintained at ~0.6 ppm. To visualize the movement of the tracers, a double-pulse laser with a peak energy of 200 mJ and wavelength of 532 nm is used to provide illumination. Images of the particles are captured using a bespoke LFC configured to operate in double-trigger mode. To guarantee that the LF-DNN model

trained on synthetic data is effective in reconstructing experimental flow particles, the optical settings of the main lens and MLA of the LFC are maintained identically to those used in the simulations. With the current optical settings, the measurement domain spans $5d \times 6.5d \times 6d$ in the x, y, and z-directions, respectively.

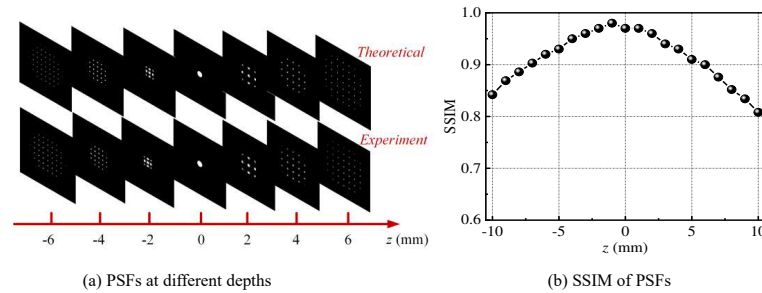(a) PSFs at different depths   (b) SSIM of PSFs

Fig. 13 Comparison of theoretical and experimental PSFs of the LFC. $z = 0$ mm represents the focal depth of the LFC. Positive and negative z values denote the far-focal and near-focal zone, respectively. (a) PSFs at different depths. (b) SSIM of PSFs.

A test is conducted to ascertain the alignment between numerical and experimental optical settings. This involves the acquisition of the point spread functions (PSFs) of the LFC by capturing a small pinhole (diameter of 5 μm). The pinhole is moved along the optical axis of the LFC by a high-resolution translation stage. The captured PSFs are then compared with those derived from theoretical calculations using the same optical parameters. The structural similarity index (SSIM) [47] is employed to evaluate the similarity. Fig. 13 presents the SSIM values between the theoretical and experimentally captured PSFs at different depths. It can be observed that the captured PSFs match well with the theoretical computing results. The SSIM value exceeds 0.8 at the depth of -10 mm $< z <$ 10 mm, indicating there is a good agreement between the experimental and numerical optical settings. Therefore, the LF-DNN model that is trained using synthetic LF data proves to be applicable for processing experimental data for the reconstruction of particle distributions.

The reconstructed particle volumes achieved by the LF-DNN are subsequently processed by a cross-correlation algorithm to estimate the 3D flow velocity. Utilizing the $16 \times 16 \times 16$ interrogation window and an overlap of 50%, $25 \times 32 \times 30$ vectors are generated in the velocity estimation. To acquire the statistical measures of the cylinder wake flow, 240 LF-PIV image pairs are reconstructed by the DNN for subsequent flow velocity estimation. For a comparative study, the MART algorithm is applied to reconstruct the particle distribution. For each LF image, 5 iterations are performed to reconstruct the particle volume. To evaluate the measurement accuracy of the proposed LF-DNN model, planar PIV measurements are also carried out in a similar configuration. A conventional camera is employed to capture the velocity field in the central depth plane of the channel. The lens magnification on the regular camera is adjusted to -1, forming a field of view of $5d \times 6.5d$. For cross-correlation calculation, the size and overlap of the interrogation window are set to 32×32 pixels and 50%, respectively, resulting in 109×146 vectors in the final velocity field. A total of 500 planar PIV image pairs are processed to derive the time-averaged velocity and compared with the proposed method.

4.2 Results and discussions

The mean velocity field of the cylinder wake flow at Re_d of 140, 100 and 70 are measured by the LF-PIV and shown in Fig. 14. The spanwise, streamwise and wall-normal directions are denoted by x, y and z-axes, respectively. The axis of the cylinder is positioned at $y = 0$. The velocity distributions on the x-y slices at $z = -2.7d$, $-1.5d$, $0d$, $1.5d$ and $2.7d$ are presented. The contour illustrates the magnitude of the streamwise velocity component (v).

(a) Re_d =140, MART   (b) Re_d =140, LF-DNN

(c) Re_d =100, MART   (d) Re_d =100, LF-DNN

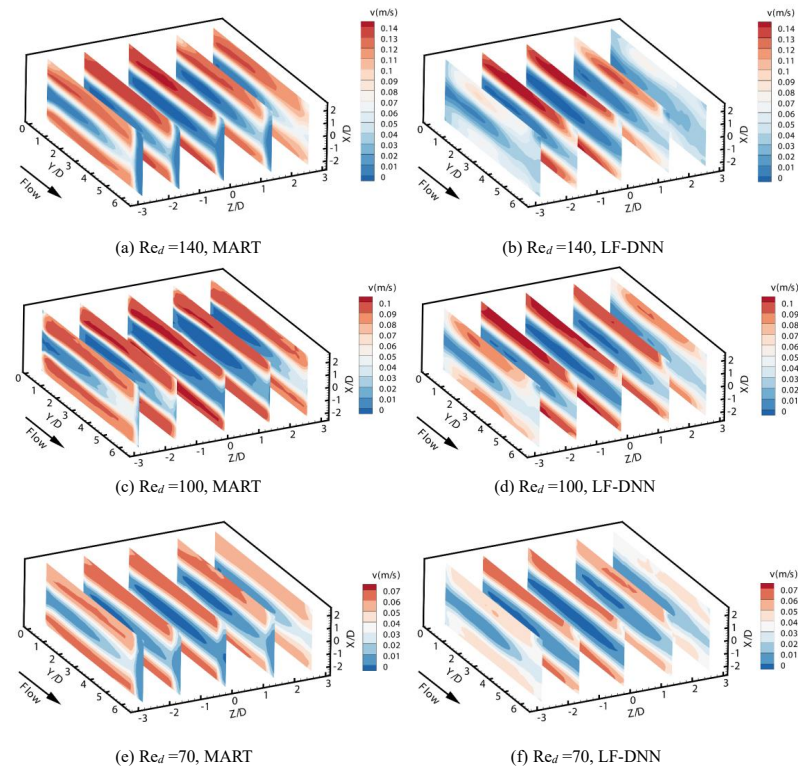(e) Re_d =70, MART   (f) Re_d =70, LF-DNN

Fig. 14 Time-averaged 3D velocity of the cylinder wake flow at Reynolds number Re_d of 140, 100 and 70 reconstructed by traditional MART and proposed LF-DNN methods. (a) Re_d =140, MART. (b) Re_d =140, LF-DNN. (c) Re_d =100, MART. (d) Re_d =100, LF-DNN. (e) Re_d =70, MART. (f) Re_d =70, LF-DNN.

From Fig. 14, it can be found that a low-speed zone is formed downstream of the cylinder. The velocity magnitude is significantly lower than the surrounding mainstream, approximately an order of magnitude less. The length of this zone gradually shrinks as the slice transits from the center to the wall. A comparative analysis is undertaken between the MART and LF-DNN. In the case of the central three slices, the velocity distribution obtained by the LF-DNN closely aligns with those from the MART. Nonetheless, discrepancies emerge on the outermost slices (z = -2.7d and 2.7d), where the velocity magnitude measured by the LF-DNN is ~20% smaller than the MART at the same Re_d. This is because the MART produces particles with severe elongations along the wall-normal

20

direction. Consequently, some voxels that contain no tracers are assigned with false intensity. This brings additional noises in the cross-correlation analysis and thus unable to detect the correlation peak. The accuracy of velocity estimation is therefore degraded. For the LF-DNN, fewer elongations can be shown in the reconstruction, as a result, the estimated velocity obtained by the cross-correlation algorithm provides a higher accuracy.

Furthermore, the instantaneous velocity field at the central depth plane of the channel is extracted from the 3D data, as shown in Fig. 15. Both the MART and LF-DNN provide the downstream shedding vortices of the cylinder. The outer vortices are in good agreement with these two approaches in terms of shapes and locations.

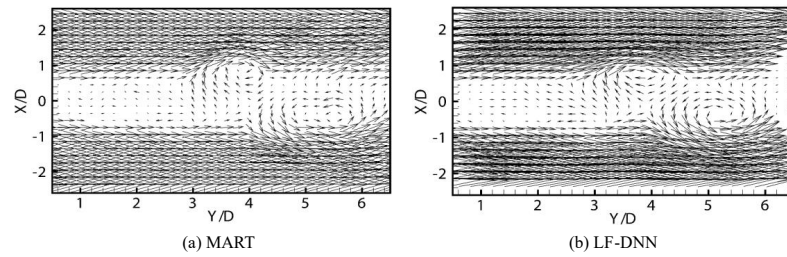

(a) MART   (b) LF-DNN

Fig. 15 Instantaneous velocity field at central depth plane of the channel measured by MART and LF-DNN. (a) MART. (b) LF-DNN.

Quantitative comparison between the LF-PIV and planar PIV is performed by extracting the mean streamwise velocity distribution along the spanwise direction on the central depth plane of the channel ($z = 0$, $y = 3d$). The result is illustrated in Fig. 16. It can be observed that streamwise velocity shows a U-shape distribution across the spanwise direction. A fast decrement of v is presented at the spanwise locations of $-1.5d < x < -0.5d$ and $0.5d < x < 1.5d$. Concurrently, the maximum streamwise velocity of the mainstream reduces from 0.16 m/s to 0.08 m/s with the decrease of Re_d from 140 to 70. The measured velocities achieved by the MART and LF-DNN are very close and generally match the trends of the planar PIV results. The observed discrepancies between the LF-PIV and planar PIV can be attributed to the difference in the spatial resolution of the velocity measurement.

Since the planar PIV has been developed as a standard technique for 2D flow diagnostics, therefore the measurement of planar PIV is considered a baseline to evaluate the measurement accuracy of the MART and the LF-DNN. The relative measurement error ε is calculated as,

$$\varepsilon = \frac{\sum_{i=1}^{i_t} \left| v_{l(i)} - v_{p(i)} \right|}{v_{mean}} \times 100\% \tag{12}$$

where $v_{l(i)}$ and $v_{p(i)}$ denote the streamwise component of the i_{th} velocity vector measured by the LF-PIV and planar PIV, respectively, v_{mean} is the averaged flow velocity estimated from the flow rate and size of the channel. The ε achieved at $Re_d = 140$, 100 and 70 are 11.9%, 9.6% and 6.9%, respectively for the LF-DNN. In contrast, the ε achieved for the MART is 13.7%, 10.3% and 9.1% with the same Re_d. For the LF-DNN, the computational cost for obtaining the mean 3D velocity is 10 mins for reconstructing the 240 LF image pairs, while the MART took about 7 days for the same measurement. Therefore, it indicates that the LF-DNN outperforms the MART in terms of reconstruction accuracy and efficiency. Despite the advantages of the LF-DNN over the conventional tomographic reconstruction algorithms, the limitations of the proposed model still

21

remain. Firstly, the improvement of the velocity accuracy is lower than the improvement of the reconstruction efficiency. This discrepancy may be attributed to the calibration errors of the optical system and experimental noise. Secondly, the model training time is still longer even with the GPUs. Therefore, it is imperative to explore approaches to reduce the training time.
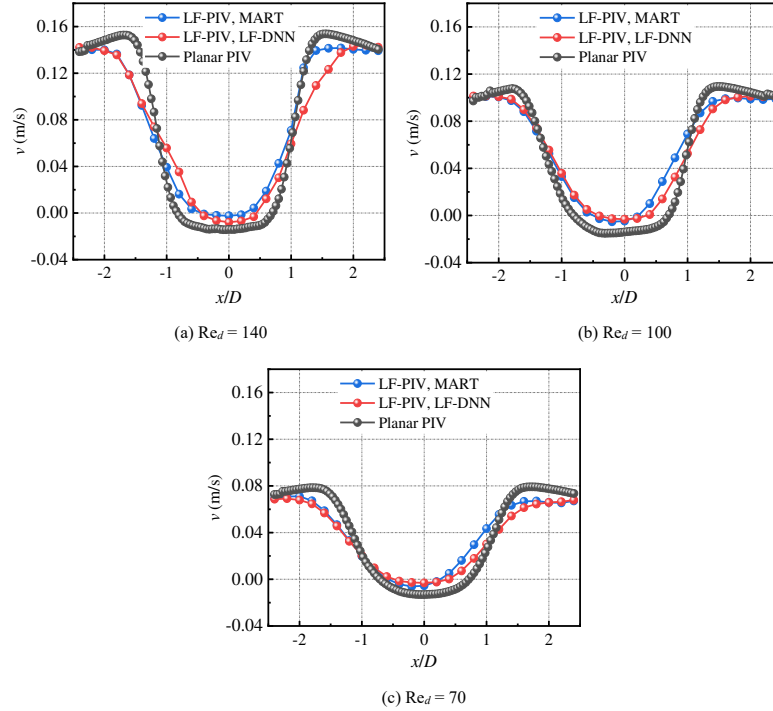

(a) $Re_d = 140$

(b) $Re_d = 100$

(c) $Re_d = 70$

Fig. 16 Streamwise velocity along the spanwise direction on the central depth of channel ($z = 0$, $y = 3d$) at different Reynolds numbers Re_d. (a) $Re_d = 140$. (b) $Re_d = 100$. (c) $Re_d = 70$.

5. Conclusion

This study proposes a deep learning-based volumetric reconstruction method for enhancing particle reconstruction accuracy and efficiency in LF-PIV. A tailored deep neural network model (namely LF-DNN) is employed to reconstruct the particle distributions from the LF data. Comparative studies between the proposed method and the state-of-the-art tomographic reconstruction algorithms (MART and PR-MART) are performed using the synthetic and experimental data. The key findings are summarized as follows:

- The LF-DNN outperforms the MART and the PR-MART in terms of particle reconstruction quality, Recall and Ghost rates. Specifically, the reconstruction quality index exceeds 0.75 for LF-DNN which is approximately twice that of the MART. For LF-DNN, the Recall rate maintains at 1 across different seeding concentrations, whereas the Recall rate decreases with

the increasing seeding concentration for the MART and PR-MART. Furthermore, the Ghost rate for the LF-DNN is significantly lower than the other two methods.

- The LF-DNN demonstrates better noise resilience compared to the MART and the PR-MART. With the integration of Gaussian noise into the LF images (mean noise value of $\mu = 0.2$), the Recall and Ghost rates of 0.81 and 2.16 are achieved, respectively for the LF-DNN, whereas 0.44 and 6.68 are observed for the MART.

- The LF-DNN provides a lower computational cost compared to the MART and the PR-MART. For reconstructing a single LF image, the LF-DNN necessitates merely 3.7 and 1.2 seconds on CPU and GPU platforms, respectively, whereas the MART and PR-MART require 0.38 and 0.28 hours.

- For LF-DNN, the relative measurement errors of 3D cylinder wake flow are 11.9%, 9.6% and 6.9% at Reynolds numbers of 140, 100 and 70. In comparison, the relative errors for the MART are 13.7%, 10.3% and 9.1% with the same Reynolds numbers. This demonstrates that the LF-DNN can facilitate accurate volumetric particle reconstruction and hence the 3D flow measurement by single-camera LF-PIV.

In future work, efforts will be made to investigate approaches to improve the calibration accuracy of the optical system and incorporate the experimental data into the training dataset. Additionally, flow governing equations will be integrated with the flow velocity data to improve the spatial resolution of the velocity measurement.

Author Declaration

Conflict of interest

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

Acknowledgements

References

[1] Westerweel J, Elsinga G E, Adrian R J. Particle image velocimetry for complex and turbulent flows[J]. Annual Review of Fluid Mechanics, 2013, 45: 409-436.

[2] Schröder A, Schanz D. 3D Lagrangian particle tracking in fluid mechanics[J]. Annual Review of Fluid Mechanics, 2023, 55: 511-540.

[3] Scarano F. Tomographic PIV: principles and practice[J]. Measurement Science and Technology, 2013, 24(10): 012001.

[4] Crane R J, Popinhak A R, Martinuzzi R J, et al. Tomographic PIV investigation of vortex shedding topology for a cantievered circular cylinder[J]. Journal of Fluid Mechanics, 2022, 931: R1.

[5] She W, Gao Q, Zuo Z, et al. Experimental study on a zigzagging bubble using tomographic particle image velocimetry with shadow image reconstruction[J]. Physics of Fluids, 2021, 33(8): 083313.

[6]   Qu Q, Cao Z, Xu L, et al. Reconstruction of two-dimensional velocity distribution in scramjet by laser absorption spectroscopy tomography[J]. Applied Optics, 2019, 58: 205-212.

[7]   Cheng H, Ji B, Long X, et al. A review of cavitation in tip-leakage flow and its control[J]. Journal of Hydrodynamics, 2021, 33(2): 226-242.

[8]   Zhao Z, Ding J, Shi S, et al. Volumetric flow characterization of a rectangular orifice impinging synthetic jet with single-camera light-field PIV[J]. Experimental Thermal and Fluid Science, 2021, 123(5), 110327.

[9]   Zhu X, Xu C, Hossain M M, et al. Approach to select optimal cross-correlation parameters for light field particle image velocimetry[J]. Physics of Fluids, 2022, 34(7): 073601.

[10]   Cao L, Zhang B, Hossain M M, et al. Tomographic reconstruction of light field PIV based on a backward ray tracing technique [J]. Measurement Science and Technology, 2021, 32: 044007.

[11]   Thomas L, Trembiais B, David L. Optimization of volume reconstruction for classical Tomo-PIV algorithms (MART, BIMART and SMART): synthetic and experimental studies[J]. Measurement Science and Technology, 2014, 25(2): 035303.

[12]   Hong L, Chamorro L P. A fast, non-iterative ray-intersection approach for three-dimensional microscale particle tracking[J]. Lab on a chip, 2022, 22(5): 964-971.

[13]   Deem E A, Zhang Y, Cattafesta L N, et al. On the resolution of plenoptic PIV[J]. Measurement Science and Technology, 2016, 27: 084003.

[14]   Toloui M, Hong J. High fidelity digital inline holographic method for 3D flow measurements[J]. Optics Express, 2015, 23(21): 27159-27173.

[15]   Shi S, Ding J, New TH, et al. Light-field camera-based 3D volumetric particle image velocimetry with dense ray tracing reconstruction technique[J]. Experiments in Fluids, 2017, 58: 78.

[16]   Zhu X, Wu Z, Li J, et al. A pre-recognition SART algorithm for the volumetric reconstruction of the light field PIV[J]. Optics and Lasers in Engineering, 2021, 143: 106625.

[17]   Feng X, Wang D, Tan H, et al. High-resolution light field particle imaging velocimetry with color-and-depth encoded illumination[J]. Optics and Lasers in Engineering, 2024, 173(2): 107921.

[18]   Fahringer T W, Thurow B S. Filtered refocusing: a volumetric reconstruction algorithm for plenoptic-PIV[J]. Measurement Science and Technology, 2016, 27(8): 094005.

[19]   Mei D, Ding J, Shi S, et al. High resolution volumetric dual-camera light-field PIV[J]. Experiments in Fluids, 2019, 60: 132.

[20]   Zhu X, Xu C, Hossain M M, et al. Fast and accurate flow measurement through dual-camera light field particle image velocimetry and ordered-subset algorithm[J]. Physics of Fluids, 2023, 35(6): 63603.

[21]   Worth N A, Nickels T B. Acceleration of Tomo-PIV by estimating the initial volume intensity distribution[J]. Experiments in Fluids, 2008, 45(5): 847-856.

[22]   Atkinson C, Soria J. An efficient simultaneous reconstruction technique for tomographic particle image velocimetry[J]. Experiments in Fluids, 2009, 47: 553-568.

[23]   Zeng X, He C, Liu Y. GPU-accelerated MART and concurrent cross-correlation for tomographic PIV[J]. Experiments in Fluids, 2022, 63: 91.

[24]   Wang G, Ye J C, Man B D. Deep learning for tomographic image reconstruction[J]. Nature Machine Intelligence, 2020, 2: 737-748.

[25]   Yu C, Bi X, Fan Y. Deep learning for fluid velocity field estimation: A review[J]. Ocean Engineering, 2023, 271(3): 113693.

[26]   Barnkob R, Cierpka C, Chen M, et al. Defocus particle tracking: a comparison of methods based on model functions, cross-correlation, and neural networks[J]. Measurement Science and Technology, 2021, 32(9): 094011.

[27] Yu C, Bi X, Fan Y, et al. LightPIVNet: An effective convolutional neural network for particle image velocimetry[J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 1-15.

[28] Gim Y, Jang D K, Sohn D K, et al. Three-dimensional particle tracing velocimetry using shallow neural network for real-time analysis[J]. Experiments in Fluids, 2020, 61: 26.

[29] Himpel M, Melzer A. Fast 3D particle reconstruction using a convolutional neural network: application to dusty plasmas[J]. Machine Learning: Science and Technology, 2021, 2(4): 045019.

[30] Gao Q, Pan S, Wang H, et al. Particle reconstruction of volumetric particle image velocimetry with the strategy of machine learning[J]. Advances in Aerodynamics, 2021, 3(9): 28.

[31] Liang J, Cai S, Xu C, et al. Filtering enhanced tomographic PIV reconstruction based on deep neural networks[J]. IET Cyber-Systems and Robotics, 2020, 2(1): 43-52.

[32] Shao S, Mallery K, Kumar S S, et al. Machine learning holography for 3D particle field imaging[J]. Optics Express, 2020, 28(3): 2987-2999.

[33] Shao S, Mallery K, Hong J, et al. Machine learning holography for measuring 3D particle distribution[J]. Chemical Engineering Science, 2020, 225(11): 115830.

[34] Chen N, Wang C, Heidrich W. Holographic 3D particle imaging with model-based deep network[J]. IEEE Transactions on Computational Imaging, 2021, 7(3): 288-296.

[35] Cao L, Hossain M M, Li J, et al. Three-dimensional particle image velocimetry measurement through three-dimensional U-Net neural network[J]. Physics of Fluids, 2024, 36(4): 047136.

[36] Herman G T, Lent A. Iterative reconstruction algorithms[J]. Computers in Biology and Medicine, 1976, 6(4): 273-294.

[37] Baldi P. Gradient descent learning algorithm overview: a general dynamical systems perspective[J]. IEEE Transactions on Neural Network, 1995, 6(1): 182-195.

[38] Sze V, Chen H, Yang J, et al. Efficient processing of deep neural networks: A tutorial and survey[J]. Proceedings of the IEEE, 2017, 105(12): 2295-2329.

[39] Li Z, Liu F, Yang W, et al. A survey of convolutional neural networks: analysis, applications, and prospects[J]. IEEE Transactions on Neural Networks and Learning, 2022, 33(12): 6999-7019.

[40] Liu W, Wang Z, Liu X, et al. A survey of deep neural network architectures and their applications[J]. Neurocomputing, 2017, 234: 11-26.

[41] He, K, Zhang, X, Ren, S, et al. Deep Residual Learning for Image Recognition[C]. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, USA, 2016, 770–778.

[42] Cai S, Zhou S, Xu C, et al. Dense motion estimation of particle images via a convolutional neural network[J]. Experiments in Fluids, 2019, 60: 73.

[43] Cai S, Liang J, Gao Q, et al. Particle image velocimetry based on a deep learning motion estimator[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(6): 3538-3554.

[44] Chen H, Wang M, Zhao X. A multi-strategy enhanced sine cosine algorithm for global optimization and constrained practical engineering problems[J]. Applied Mathematics and Computation, 2020, 369: 124872.

[45] Kingma D, Ba J. Adam: a method for stochastic optimization [C]. In Proceedings of the 3rd international conference on learning representations (ICLR), 2014.

[46] He L, Chao Y, Suzuki K, et al. Fast connected-component labeling[J]. Pattern Recognition, 2009, 42(9): 1977-1987.

[47] Wang Z, Bovcik A C, Sheikh H R, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600-612.