Filtering enhanced tomographic PIV reconstruction based on deep neural networks

Jiaming Liang\(^1\), Shengze Cai\(^1\), Chao Xu\(^1\) \(\&\) Jian Chu\(^{1,2}\)

\(^1\)Institute of Cyber-Systems and Control, College of Control Science and Engineering, Zhejiang University, Hangzhou, People’s Republic of China

\(^2\)Ningbo Industrial Internet Institute (NIII), Ningbo, Zhejiang, People’s Republic of China

\(\&\) E-mail: cxu@zju.edu.cn

Abstract: Tomographic particle image velocimetry (Tomo-PIV) has been successfully applied in measuring three-dimensional (3D) flow field in recent years. Such technology highly relies on the reconstruction technique which provides the spatial particle distribution by using images from multiple cameras at different viewing angles. As the most popular reconstruction method, the multiplicative algebraic reconstruction technique (MART) has advantages in high computational speed and high accuracy for low particle seeding reconstruction. However, the accuracy is not satisfactory in the case of dense particle distributions to be reconstructed. To overcome this problem, a symmetric encode–decoder fully convolutional network is proposed in this paper to improve the reconstruction quality of MART. The input of the neural network is the particle field reconstructed by the MART approach, while the output is the regenerated image with the same resolution. Numerical evaluations indicate that these blurred or irregular particles can be significantly refined by the trained neural network. Most of the ghost particles can also be removed by this filtering method. The reconstruction accuracy can be improved by more than 10% without increasing the computational cost. Experimental evaluations indicate that the trained neural network can also provide similar satisfactory reconstruction and improved velocity fields.

1 Introduction

From fish to birds to insects, locomotion is a fundamental feature of life [1]. Nowadays, the locomotion though fluid media (e.g. air and water) has attracted more and more attention from researchers. In such locomotion (e.g. swimming and flapping), special structures and kinematic strategies have been evolved to generate propulsion and navigate complex flow fields, which is associated with evolutionary advantages including energy savings [2–5] and flow enhancement [6]. Visualisation of the complex fluid media [7] induced by the biological locomotion is a promising way to yield intriguing insight into the physical mechanisms of biological intelligence. The discovered physical mechanisms can help to guide the design of intelligent biologically inspired robots [8, 9].

Particle image velocimetry (PIV) [10, 11] is a non-intrusive optical technique, which provides global velocity measurements and plays an important role in fluid mechanics. By seeding particles to the fluid flow, PIV obtains the instantaneous velocity estimation of the particles. The seeded particles are assumed to faithfully follow the flow. Therefore, the velocity estimation of the flow is obtained. In recent years, tomographic PIV (Tomo-PIV) [12, 13] has made significant progress in three-dimensional (3D) flow field measurements. Tomo-PIV is widely used in various fields due to the accurate and efficient 3D three-component velocity measurements. The technique first reconstructs the distribution of the spatial particles by multiple cameras imaging at different viewing angles and then uses the cross-correlation method to calculate the velocity field.

Particle reconstruction plays a key role in Tomo-PIV. The development of reconstruction techniques almost promotes the development of Tomo-PIV. The iterative algebraic methods [14] were first applied to tomographic reconstruction since they are simple, effective, and can be able to deal with large-scale, undersampled data with noise. Subsequently, the multiplicative algebraic reconstruction technique (MART) [12] was introduced into Tomo-PIV, which greatly improved the effect of 3D flow field measurement and became the most popular reconstruction method. MART is exponential iterated, which is improved from algebraic reconstruction technique (ART) [15]. Such improvement contributes to the fast convergence speed and high accuracy for low particle seeding measurement. There followed the simultaneous multiplicative ART (SMART) [16], which significantly reduces computation time and memory cost without loss of accuracy. Although the above proposed reconstruction techniques have gradually improved the accuracy and efficiency of 3D particle reconstruction [17], the existence of ghost particles [18] and the uncertainty of image particle density [19] still greatly hindered the development of reconstruction techniques and Tomo-PIV.

The main research direction of particle reconstruction is proposing new techniques by optimising MART, which aims to improve the reconstruction accuracy. Recently, more attention is paid to the features of the reconstructed object itself. Observing the particle images, it is easy to find some priori information that leads to the improvement of MART. For example, the particle intensity distribution is sparse, most pixels in the image (>95%) have zero intensity. Considering the sparsity, multiplicative first guess (MFG) [20] and multiplicative line-of-sight (MLOS) [21] are proposed to initialise the particle distribution before iteration. The sparse initial particle distribution accelerates the convergence speed of iteration and reduces the computational cost.

It is realised that each individual particle is supposed to have a spherical intensity distribution [10]. Based on the special feature, MART and SMART put forward a reasonable physical assumption that the 3D object to be reconstructed has a spherical and uniform shape. In addition, according to the laser scattering characteristics from particles, Gaussian distribution is commonly a priori to approximate the particle intensity distribution of each particle [22]. For instance, Ye et al. [23] developed a new technique called dual-basis pursuit, which combines a Gaussian intensity distribution template with a correcting template as a priori to reconstruct the particle intensity field with sub-voxel accuracy.

The usage of the sparsity and spherical intensity distribution as a priori knowledge greatly improves MART and gives rise to higher accuracy for low particle seeding measurement. However, the introduction of the above priori knowledge greatly increases the
computational time and memory usage. Besides, the reconstruction accuracy is still highly related to the number of cameras, the viewing angle of cameras, and the particle density. Especially, even the improved MART cannot deal with the elongation of true particles and ghost particles very well. More a priori knowledge needs to be taken into account in Tomo-PIV reconstruction. To overcome these problems, we try to use the deep learning method to refine the particle field reconstructed by the MART approach. It is desired that the deep learning method should obtain a particle intensity distribution with higher accuracy.

The traditional flow visualisation plays an important role in the research of the locomotion through fluid media. However, it may not be able to extract enough fine structures of the complex flow, due to the coarse information process. In the past decade, artificial intelligence, especially the deep learning technology [24], has developed rapidly and been applied to many complex problems such as classification, recognition, and image generation. As one of the most powerful tools, convolutional neural networks (CNNs) [25] have succeeded in many computer vision tasks, which is also expected to show a promising improvement in PIV. Gao et al. [26] firstly proposed a CNN model to refine the particle reconstruction from a coarse initial guess of distribution. They obtained significant results with higher reconstruction quality than traditional methods. Inspired by their work, we try to refine and regenerate the particle field reconstructed by MART with newly-developed deep learning techniques. Image segmentation and image generation methods in deep learning have attracted our attention. A fully convolutional network (FCN) [27] is a milestone of image segmentation. What FCN has done is replacing the fully connected layer with a deconvolution layer for pixel-level classification. The U-Net [28], which is a derivative of FCN network, combines multi-scale features to achieve excellent performance in biomedical image segmentation and has been widely used. Fully convolution DenseNet [29] incorporates a special network structure DenseNet, which establishes dense connections between all previous network layers and the following network layer. The special structure makes the network easier to train and getting smaller network parameters. Pyramid scene parsing network [30] uses the pyramid pooling structure to fuse context information from different regions, and also obtain excellent segmentation result. SegNet [31] is a novel fully CNN for semantic pixel-wise segmentation. This architecture is more competitive in both memory usage and computational time than other networks, which highly inspired our work. Considering the outstanding performance in pixel-level extraction, a symmetric encoder–decoder FCN is proposed in this study to refine and regenerate the particle field reconstructed by MART.

The paper is organised as follows. The basic concept of the MART is provided in Section 2. In Section 3, we describe the Tomo-PIV reconstruction technique on the basis of a symmetric encoder–decoder FCN and the generation of synthetic training data. Both numerical and experimental evaluations are demonstrated in Section 4, followed by the conclusions in Section 5.

2 Reconstruction technique

2.1 Reconstruction problem

As mentioned before, Tomo-PIV first reconstructs the spatial particle field based on the simultaneous projections defined in the image space and then extracts the displacement of the corresponding particle pair by using the 3D cross-correlation method. Reconstruction technique shows great room for improvement and novelty in Tomo-PIV and attracts the most concentration of recent research. The aim of the reconstruction problem is to produce the particle intensity distribution based on the mapping relationship between the image space and the physical space, which is obtained from stereoscopic camera calibration.

Assuming that particles in physical space $E$ are simultaneously projected onto four camera planes at different viewing angles, the mapping function can also be obtained in advance with a calibration plate being imaged by the four fixed cameras. Schematic of the imaging model used for tomographic reconstruction presented in Fig. 1. Discretising the physical space of particle distribution into an array of voxels which have the same size as pixels, we can write the projection function as a linear equation

$$W_i f = I$$

where $I$ denotes the pixel intensity projected at the image plane and can also be viewed as the integration of spatial particle intensity along the line of sight. $W_{ij}$ is the weighting matrix function, which represents the contribution of the $j$th voxel to the $i$th pixel intensity. Also, $E$ refers to the volume intensity of the voxels. Therefore, the reconstruction technique can be regarded as an inverse problem to solve $E$ with known $I$ and $W_{ij}$.

However, as the weighting matrix is usually sparse, the linear equation (1) is underdetermined in general. To cope with this problem, optimisation techniques are applied by introducing some objective functions, which will be demonstrated in the next section.

2.2 Multiplicative ART (MART)

As one of the most widely accepted approaches for Tomo-PIV, MART is designed to maximise the entropy to solve the underdetermined problem. MART is proved to have a significant effect in the reconstruction method and has impressive advantages in terms of computational efficiency and accuracy for low particle seeding measurement. Besides, it is also the prototype of many MART-based methods in the PIV community. MART updates the reconstructed intensity (namely the particle distribution) by iteration. The iteration equation is given as follows:

$$E^{k+1}_v = E^k_v (I^{WE^k} E^k)^{WE^k}$$

(2)

where $k$ denotes the $k$th iteration, $0 \leq u \leq 1$ is a scalar to determine the convergence rate. The fraction $(I^{WE^k})^*E^k$ represents the difference of the currently estimated intensity projection $WE^k$ with the initially measured pixel intensity $I$. Thus, the volume intensity $E$ will converge to the optimal value successively according to the difference. Before the iteration process, an initial particle intensity distribution $E^0$ should be given by MFG [20] or MLOS [21] method, which is mentioned above.

As the reconstructed particles may only occupy little voxels of the total volume, most of the voxels can be ignored and will not take part in the reconstruction process. The initial distribution given by MLOS can pick out the 'non-zero' voxels, which leads to fewer computational operations and less memory usage. After getting the list of non-zero voxels, simultaneous MART (SMART) corrects every voxel using the simultaneous projection to every pixel on different camera planes. The correction of each non-zero voxel is performed as below:

Fig. 1 Schematic of the imaging model used for tomographic reconstruction
Moreover, a novel loss function is required, which guarantees the
2.3 Filtering improved MART via a neural network (NN)
by the MART algorithm should be further improved.

The main idea for NN is illustrated in Fig. 2. The proposed NN
and obtain a spatial particle distribution with better accuracy, we
of the input data, CNN can approximate the training dataset
the elongation of true particles and ghost particles), which greatly
refined reconstruction with higher accuracy. In order to achieve all
particle distribution with better quality, particle field reconstructed
be able to implement pixel-level refinement and regeneration.

2.3 Filtering improved MART via a neural network (NN)

CNN is one of the most powerful artificial NNs in image
processing. Owing to the capability to learn representative features
of the input data, CNN can approximate the training dataset
accurately and provide satisfactory results. A typical deep NN is
composed of multiple layers (including input, output, and hidden
layers). Each layer consists of a number of neurons and

\[ E^{k+1} = E^k \prod_{i}^{N} \left[ \frac{1}{WE^k} \right]^{iN} \]  

(3)

where \( N \) denotes the number of pixels, which correspond to the
observed voxel.

MART and SMART greatly improve reconstruction accuracy
and become the most representative techniques in Tomo-PIV. However, the reconstruction results are greatly related to the
number of cameras, the viewing angle of cameras, and the particle
density of the observed domain. In particular, the existence of
ghost particles and the elongation of true particles still limit the
reconstruction accuracy of MART and SMART. Mathematically,
the formation of ghost particles is caused by the non-uniqueness
solution of the underdetermined problem given in (1). As the
information of spatial particle position is limited, estimated
particles may appear at any intersection of all lines of sight, where
there is no actual particle. On the other hand, due to the multiple
viewing angles and the imaging out of focus, the projected
intensity profile deviates further from the Gaussian shape. This is
generally referred to as the particle elongation phenomenon. In
other words, the reconstructed particle, which is supposed to be
spherical, is now elongated along the axial line of the camera (i.e.
the depth direction). To solve these problems and obtain a spatial
particle distribution with better quality, particle field reconstructed
by the MART algorithm should be further improved.

The input of the NN is the particle intensity distribution
of the above requirements, a CNN structure is designed and more
details of the network are demonstrated in the next section.

3 NN design and training

In this section, we present the structure of the symmetric encoder–
decoder FCN at the beginning. The NN model is used for the
refinement of the particle field reconstructed by MART, which is
equivalent to a spatial filtering operation. Then, we also present
how to generate a dataset for network training. Finally, we show
more details about the network training.

3.1 Network architecture

The main idea for the NN is illustrated in Fig. 2. The full CNN
structure is adopted, and offline trained on the synthetic dataset.
The input of the NN is the particle intensity distribution \( E_{\text{MART}} \)
reconstructed by the MART approach, while the output is the
regenerated particle intensity distribution \( E_t \) with the same
resolution.

Considering the generic FCN architecture in conducting the
refinement of particle field, a symmetric encoder–decoder FCN is
proposed. The network architecture is demonstrated in Fig. 3 and a
brief introduction is given as below. Some basic layers are
successively stacked together with skip connections to form a
complex structure. Original images will be transformed into one-
channel images and then fed into the network as input. The encoder
(left half of the network) extracts features and compresses the
spatial information into feature maps with multiple channels and
reduced resolution. The multiple feature maps can maximise the
use of multiple image features, which make it possible to obtain
satisfying refinement even trained on the small-scale dataset.
Subsequently, to enable per pixel refinement and regeneration, the
unsampling (de-convolutional) layer is applied to the decoder
(right half of the network). This network architecture is inspired by
U-Net [28].

Different from simply using the de-convolutional layer to
reverse the forward convolution and resemble an upsampling
interpolation in FCN, the decoder applies several convolutional
layers closely after each de-convolutional layer to fuse the
concatenated multi-scale features. The concatenation comes from
many skip connections (black arrows in Fig. 3), which transfers
multi-scale feature maps directly from a level of the encoder to the
corresponding level of the decoder. By fusing the particle field
feature maps at the multi-scale spatial resolution, the network is
capable of realising multi-scale refinement and deep supervision
[33].

The deep blue block in Fig. 3 consists of a convolutional layer
which is stride with a factor of 2 (performs as a pooling layer as
well), a batch normalisation layer and a leaky rectified linear unit
(ReLU) layer, of which more details will be presented later. Each
pooling layer (red block) after two deep blue blocks mainly plays a
role of data downsampling, which make it possible to extract multi-
scale features of particle fields. The aggregation of information
in large particle field areas also makes the processing more effective.
Specifically, the Max pooling method is adopted, which takes
maximal values in a local region. The batch normalisation [34]
layer normalises the passing values in the hidden layers by

\[
E_{t} = E_{\text{MART}} \prod_{i}^{N} \left[ \frac{1}{WE^{t}} \right]^{iN}
\]
adjusting and scaling the activations, which accelerates the NNs training and reduces the sensitivity to network initialisation. The leaky ReLU layer provides non-linearity to the network. Instead of being zero with negative input, leaky ReLU still have a small negative slope, which avoids the appearance of silent neuron. Besides, the dropout layer [35] is incorporated into NN to implement regularisation and address over the fitting problem. Dropout randomly drops out some neural units to form different network architectures each time, and therefore prevents the co-adaptations between specific neurons.

As mentioned before, the particle intensity distribution (input of network) is simple and sparse, high-level features and low-level features are almost playing the same important role in refinement. Specifically, the high-resolution but low-level features are used to refine the particles in blurred contour or irregular shape and implement precise particle contour extraction. While low-resolution but high-level features are almost playing the same important role in refinement.

Table 1 Parameter selection for particle image generation

| Parameter          | Range    | Unit |
|--------------------|----------|------|
| seeding density $\rho$ | 0.05–0.30 | ppm  |
| peak intensity $F_0$   | 180–200  | grey value |
| particle diameter $d$  | 1        | pixel |
| horizontal location $x_0$ | 0–992   | pixel |
| vertical location $y_0$  | 0–192   | pixel |

The spatial locations of particles are randomly distributed, and the number of particles in an image is determined by particle seeding density $\rho$ (unit: particle per pixel, ppm). The intensity distribution of every single particle can be described by 3D Gaussian function

$$F(x, y, z) = F_0 \exp \left( \frac{(x-x_0)^2}{(1/8)d^2} + \frac{(y-y_0)^2}{(1/8)d^2} + \frac{(z-z_0)^2}{(1/8)d^2} \right)$$

where $x, y, z$ represent the 3D spatial location of a pixel; $d$ denotes the particle diameter; $F_0$ is the peak intensity in the Gaussian centre, and $(x_0, y_0, z_0)$ refers to the centre position of the particle. We consider the 2D particle intensity distribution, $z$ is ignored and $x, y$ are preserved. The detailed parameter selection for particle image generation is given in Table 1. Some of the parameters are randomly selected in a proper range. Given specific values of $(x_0, y_0), F_0, d$, particle intensity distribution images can be produced by repeating the data generation steps mentioned before. Then, thousands of images are generated for network training and testing in this study. A magnified view of some generated images is presented in Fig. 4. Note that in this study, we only use 800 particle images with $\rho = 0.15$ ppm for the CNN training. The other data items consist of 200 testing images. The resolution of all images is $192 \times 960$ pixels.

3.3 Network training

After building up network architecture and generating a synthetic dataset, the parameters of the network are supposed to be optimised by minimising a given loss function. A loss function is used for evaluating how well the trained network models the dataset or how well the network refinement meets the expectations more specifically. Typically, as the refinement loss, the FCN uses mean square error (MSE), which is the squares of the deviations between the refined generation and the ground truth, averaged over all pixels

$$\text{LOSS}_{\text{MSE}} = \frac{1}{mn} \sum_{i=1}^{m} \sum_{j=1}^{n} (y_{ij} - \hat{y}_{ij})^2$$

where $y_{ij}$ is every single pixel of the ground-truth image $E_i$ and $\hat{y}_{ij}$ is a single pixel of the CNN refined image. $n$ denotes the total pixels in our image, and $m$ refers to the image batch size in each iteration. Considering the quality evaluation of reconstructed intensity distributions in MART, we try to combine the normalised correlation coefficients $Q$ [3] with the MSE, which is a standard metric in the experimental fluid mechanics community. The definition of $Q$ is given below

$$Q = \frac{\sum_{i,j} (y_{ij} - \bar{y}) (\hat{y}_{ij} - \bar{\hat{y}})}{\sum_{i,j} (y_{ij} - \bar{y})^2 \sum_{i,j} (\hat{y}_{ij} - \bar{\hat{y}})^2}$$

where $\bar{y}$ is the average of the ground-truth image $E_i$, and $\bar{\hat{y}}$ is the average of the CNN refined image.
\[ Q = \frac{1}{m} \sum_{i=1}^{m} \frac{Y^T_{gi} Y_{pi}}{\sqrt{Y^T_{gi} Y_{gi}} (Y^T_{pi} Y_{pi})} \]  

where \( Y_{gi} \) and \( Y_{pi} \) denote column vectors of ground-truth image and refined image, respectively. The column vector is reshaped from an image matrix with a size of \( n \times 1 \). The superscript T means a transpose operation. Finally, the combined loss function is presented below

\[ \text{Loss}_{\text{combined}} = \text{Loss}_{\text{MSE}} + \lambda(1 - Q) \]

where \( \lambda \) denotes a hyper-parameter to weigh the importance of \( Q \) with regard to MSE. To illustrate the effect of the inclusion of \( Q \) values to loss function and the selection of hyper-parameter \( \lambda \), we use various \( \lambda \) from 0 to 1 with every 0.25 interval for training. The loss after training and \( Q \) on a validation set with different \( \lambda \) are listed in Table 2. The validation set is split from the original training set with a proportion of 12.5%. The factor \( Q \) of the refined result is obviously higher in the case of the inclusion of \( Q \) values to the loss function. Comparing the result with different \( \lambda \), we set \( \lambda = 0.75 \) in our experiment.

Considering the huge amount of training data, the employed network is trained using a stochastic optimisation method named Nesterov-accelerated adaptive moment estimation (Nadam) [36] to accelerate the training efficiency. Nadam is a special optimisation algorithm combining Nesterov accelerated gradient descent [36] and adaptive moment estimation (Adam) [37] method. Based on mini-batch gradient descent, a mini-batch of data is used to backpropagate loss and update network parameters for each iteration process. In order to improve the stability, the batch size starts with four images and then multiplied by two every 2000 iterations. The training schedule of batch size is shown in Fig. 5.

The proposed network converges after 6000 iterations. The convergence of loss function and quality factor \( Q \) over training data is presented with a log scale in Fig. 6. The network is trained on the Ubuntu system using the tensor-flow platform with the graphics processing unit (GPU) mode, which can highly accelerate the computation. Therefore, it only takes about 3 h until the training loss converges. More details of the hardware system can be seen in Section 4.1.
4 Evaluations on experiments

In order to illustrate the theoretical results, both numerical and real experiments of the proposed NN method are carried out and discussed in the next two sections. We first show the experiment setting and evaluation flow. Next, reconstruction results of the proposed reconstruction method and comparison of quality factor $Q$ for different reconstruction techniques with various particle seeding densities are shown in Section 4.1. Quantitatively, the quality factor $Q$ and execution time are considered as the most important criteria. Finally, in order to validate the effectiveness of the proposed approach in real experiments, the proposed method is evaluated by real experimental data in Section 4.2. Since the ultimate goal of particle reconstruction is to get accurate velocity fields, we also show the improvement in the estimated motion fields after the refinement of the NN.

4.1 Numerical experiments results and assessments

A typical 2D experiment is used to assess the performances of the proposed reconstruction method based on NN. Firstly, we synthesise 1D projections of the particles imaged on several virtual cameras. Next, 2D images are preliminary reconstructed using the MART method, by which particle distribution is iteratively estimated according to the known image and weight function $W$. The iteration form is shown in (2). Successively, using the preliminary reconstructed image as input, the trained full CNN refines and regenerates an optimised reconstruction with higher accuracy. Note that once the network is trained as introduced before the parameters of the network will be fixed. That is to say, the network will output same refinement given the same input without any manual tuning process. The total evaluation flow of the proposed reconstruction method is shown in Fig. 7.

Fig. 7 Evaluation flow of the proposed reconstruction method

Consider the tremendous influence of particle seeding density to reconstruction, we generate several particle intensity distributions for evaluation with various particle seeding densities $\rho$ from 0.05 to 0.3 ppp. The density values with every 0.05 ppp interval. The reconstruction accuracy in this experiment is mainly described by the normalised correlation coefficients $Q$ given in (6). Experiment results and assessments are described in the following section.

The trained network shows impressive results on not only the training data but also the testing data. A particle field image with seeding density $\rho = 0.15$ ppp is given as input for observation and comparison. Significant improvement is obtained by the proposed deep learning reconstructor and can be seen in Fig. 8. Randomly seeded particles are shown as black dots in the ground-truth particle field (Fig. 8a). Then, particle field reconstructed by MART is presented in Fig. 8b. Eventually, the particle field refined by the trained network is illustrated in Fig. 8c. From the comparison of Figs. 8b and c, it is obvious that the network refines the particles in blurred contour or irregular shape and removes nearly most of the ghost particles. In addition, the final refinement (Fig. 8c) is almost the same as the ground-truth particle field (Fig. 8a).

As shown in Fig. 8, some ‘ghost’ particles, which do not exist in the ground-truth image, seem to be intensified and more obvious. To illustrate the impact of this observation, we show the probability density functions (PDFs) of the peak intensities of ghost and actual particles with seeding density $\rho = 0.15$ ppp on the test set in Fig. 9. It is obvious that only little ghost particles are intensified after the refinement of the network. Most ghost particles are weaker and have high probability densities for small intensity fields using our trained NN. Therefore, the enhancement of little ‘ghost’ particles may not lead to further problems during the reconstruction. In addition, nearly 70% of the ghost particles are removed after the refinement.

In addition, particle field image with seeding density $\rho = 0.2$ ppp is also evaluated. The corresponding results are shown in Fig. 10. Particles with seeding density $\rho = 0.2$ ppp are shown in the ground-truth particle field (Fig. 10a). For more observation and comparison, the particle field is reconstructed by SMART and presented in Figs. 10b and d). Similarly, the network refinement of particle field reconstructed by SMART is also presented in Figs. 10c and e, respectively. However, there are still many particles in blurred contour or irregular shape and ghost particles with particle field reconstructed by SMART (Fig. 10d), which is similar to that of MART (Fig. 10b). Surprisingly, the proposed deep learning reconstructor can also refine the particle field reconstructed by SMART and even provide better result (Fig. 10c) than the refined reconstruction of MART (Fig. 10c). Comparing the reconstructed results in Figs. 8 and 10, the deep learning reconstructor can also deal with particle field images with higher seeding density. Besides, we can also find the powerful generalisation of our deep learning reconstructor. As mentioned above, only data with $\rho = 0.15$ ppp is used for CNN training. The deep learning reconstructor can perform very well on particle field images with various seeding densities, most of which do not even appear in the training dataset. As a spatial filter, the deep learning
reconstructor is only trained on particle field reconstructed by the MART method. However, it can also be used to refine particle field reconstructed by other methods (e.g. SMART).

To quantitatively illustrate the effectiveness of the proposed deep learning reconstructor, quality factor $Q$ of both MART (or SMART) reconstruction and NN refinement after MART (or SMART) with respect to various seeding densities are shown in Fig. 11. By comparison, the reconstruction accuracy can be improved by the proposed NN method by >10%. For the particle field with seeding density $\rho < 0.1$ ppp, the proposed NN method could reconstruct particle field with $Q > 0.99$. Although with the increase of seeding density, it is too dense to reconstruct particle field with both methods, the proposed NN method still shows a satisfactory improvement.

In addition, the method runs in a high efficiency without any extra computational burden. The experiment is evaluated on Ubuntu and Windows systems. The Ubuntu system is equipped with Dual E5-2678 processor and TITAN graphics card, which make it possible to run networks on the GPU mode and accelerate the computation. Averagely, it takes 0.15 s to refine every particle field image. Even on the Windows system, which is equipped with a Core i7-6700 CPU processor, the refinement only requires 0.43 s on average.

### 4.2 Evaluation of experimental data

To investigate the performance in real experiments, the proposed NN is applied to real experimental data. The experimental setup will be illustrated in the next paragraph. Then, the velocity fields of two reconstructed consecutive particle images are estimated to explore the possible impacts of our proposed method on velocity measurements. The evaluation flow is similar to that of the numerical experiments, as described in Section 4.1.

Many fish swim in the water by creating body undulations or tail oscillations, which will generate a series of vortex rings in the wake [38]. Such kinds of vortex rings can produce high thrust for propulsion and improve the sustained propulsive efficiency of other fish in schooling formations [4]. Therefore, the visualisation of fluid flow induced by such biological locomotion is worth exploring and can help to discovery the physical mechanisms of biological intelligence. In this experiment, macro fibre composites (MFCs) [39] are used as actuators to mimic the oscillation
behaviours of a fish tail because of the thin film, excellent flexibility, strong actuating force, and waterproof characteristic. The method of fabricating standard MFC actuators is commercialised by the Smart Material Corp [40]. Such actuators consist of piezoelectric ceramic fibres, placed between layers of epoxy adhesive and copper electrodes. With the high potential in underwater locomotion, miniature biomimetic robotic swimmers using MFC actuators are attracting increasing attention.

The experimental setup is shown in Fig. 12a, which is conducted in a water tank. The biomimetic fish tail is fixed with a plexiglas clamp and placed in the middle of the water tank to avoid the wall effect. A magnified view of the tail is presented in the lower-right corner of Fig. 12a. The laser sheet shown in Fig. 12a is produced by a continuous laser generator. The biomimetic tail oscillates horizontally. We also show the top view of the underwater tail in the upper-right corner of Fig. 12a. The substrate of the biomimetic tail is made of 1106 aluminium alloy. Two MFC actuators are attached to both sides of the substrate. A schematic diagram of fish tail actuated by MFC is shown in Fig. 12b. A sinusoidal voltage signal with different frequencies is applied to the MFC actuators to mimic the oscillation behaviours. The vortices generated by the corresponding oscillations are also shown in the possible positions according to observation in [38].

Continuous particle images are captured from the top of the biomimetic tail. One of the particle images is shown in Fig. 12c. The white line in Fig. 12c is the oscillating tail, and it oscillates from the right to left at the moment. In this experiment, a region of interest (ROI) is extracted with the red dotted box, which may contain the possible vortex. Then, the ROI is projected on several virtual cameras to get the 1D projections. Next, 2D particle distributions are reconstructed from the projections using the MART method and our proposed method. The corresponding reconstructions are presented in Figs. 12d and e, respectively. From the comparison of Figs. 12d and e, the network obviously refines...
the particles as with the result in numerical experiments, which indicates the effectiveness of the proposed approach in real experiments. The same operation is performed in the next frame. Finally, the velocity fields are estimated using the successive reconstructed images. Velocity fields of reconstructed particle images using the MART method and our proposed method are presented in Figs. 12f and g, respectively. The estimation is performed using a coarse-to-fine Horn-Schunck (HS) optical flow method [41] with the same parameters. From the comparison of Figs. 12f and g, impressive improvement can be found in the estimated motion fields after the refinement of the NN. As shown in Fig. 12g, a distinct clockwise vortex emerges as the tail oscillates from the right to left, which is consistent with the observation [38].

5 Conclusion
Reconstruction technique plays an important role in Tomo-PIV. Several reconstruction techniques including the most popular MART method were reviewed. However, the above approaches cannot deal with the elongation of true particles and ghost particles very well. Therefore, a deep learning method for particle field reconstruction and refinement is proposed to overcome these problems and obtain a spatial particle distribution with higher accuracy. The proposed symmetric encoder–decoder fully CNN is developed from segmentation networks. The input of the NN is the particle field reconstructed by the MART approach, while the output is the refined image with the same resolution. Moreover, the NN is trained using a synthetic particle field dataset described in Section 3.2. The parameters of the NN are optimised by minimising a special loss function, which combines the traditional MSE with the normalised correlation coefficients Q introduced in MART.

Both numerical and experimental evaluations demonstrate the significant performance of the proposed NN reconstructor and show it is superior to the MART. Impressive experimental comparisons also indicate that the network refines the particles in blurred contour or irregular shape and filters nearly most of the ghost particles. Besides, the influence of particle seeding density is also investigated to verify the effectiveness of the proposed NN reconstructor. Most importantly, the reconstruction accuracy can be significantly improved by >10% without increasing additional execution time. Although the network is trained from a synthetic dataset, the impressive improvement of the estimated velocity fields further validates the effectiveness of real experimental data. The deep learning reconstructor shows a promising future and can be applied to various applications.

There are several on-going works, which could be presented in the future. Firstly, the deep learning reconstructor should be evaluated on more experiments. We may test the effect of random noise on reconstruction. Secondly, a more powerful NN model will be taken into consideration. Keeping track of the progress of deep learning techniques is also a great way to improve our deep learning reconstructor. Thirdly, we will upgrade the NN model for 3D reconstruction. Finally, we may extend the encoder–decoder CNN to a sequential network such as recurrent NNs, which can make use of the temporal structure of multi-frame particle images. This kind of network can utilise the continuity of particle motion to remove some ghost particles and improve the reconstruction accuracy.

6 Acknowledgments
This work was supported in parts by the National Natural Science Foundation of China under grant no. 61973270, the Foundation for Innovative Research Groups of the National Natural Science Foundation of China under grant no. 61621002, and the Fundamental Research Funds for Central Universities.

References
[1] Hosoi, A.E., Lauga, E.: ‘Mechanical aspects of biological locomotion’, Exp. Mech., 2010, 50, pp. 1259–1261
[2] Reddy, G., Wong-Ng, J., Celani, A., et al.: ‘Glider soaring via reinforcement learning in the field’, Nature, 2018, 562, p. 236
[3] Reddy, G., Celani, A., Sejnowski, T.J.: ‘Learning to soar in turbulent environments’, Proc. Natl. Acad. Sci., 2016, 113, pp. 4877–4884
[4] Verma, S., Novati, G., Kounoutsakos, P.: ‘Efficient collective swimming by harnessing vortices through deep reinforcement learning’, Proc. Natl. Acad. Sci. U.S.A., 2018, 115, pp. 5849–5854
[5] Becker, A.D., Masoud, H., Newbolt, J.W., et al.: ‘Hydrodynamic schooling of flapping swimmers’, Nat. Commun., 2018, 8, p. 8514
[6] Gravish, N., Perlin, K., Ju, Y., et al.: ‘Collective flow enhancement by tandem flapping wings’, Phys. Rev. Lett., 2015, 115, p. 188101
[7] Gao, S., Xu, C., Yu, K.F.C.: ‘Visual inference of flow flux via free surface PDE model and image sequence assimilation’, IET Cyber-Syst. Robot., 2019, 1, pp. 20–27
[8] Chen, L., Zhang, Y., Zhou, C., et al.: ‘Aerodynamic mechanisms in bio-inspired micro air vehicles: a review in the light of novel compound layouts’, IET Cyber-Syst. Robot., 2019, 1, pp. 1–22
[9] Shank, B.Y., Cheng, B.: ‘Efficient thrust generation in robotic fish caudal fins using policy search’, IET Cyber-Syst. Robot., 2019, 1, pp. 38–44
[10] Raffel, M., Willert, C.E., Wersley, S.T., et al.: ‘Efficient flow field velocimetry: a practical guide’ (Springer-Verlag, Berlin Heidelberg, 2007)
[11] Adrian, R.J., Westerweel, J.: ‘Particle image velocimetry’ (Cambridge University Press, Cambridge, England, 2011)
[12] Elisgina, G.E., Scarano, F., Wieneke, B., et al.: ‘Tomographic particle image velocimetry’, Exp. Fluids, 2006, 41, pp. 933–947
[13] Scarano, F.: ‘Tomographic PIV: principles and practice’, Meas. Sci. Technol., 2013, 24, p. 022001
[14] Herman, G.T., Lent, A.: ‘Iterative reconstruction algorithms’, Comput. Biol. Med., 1976, 6, pp. 273–294
[15] Gordon, R., Bender, P., Herman, G.T.: ‘Algebraic reconstruction techniques (ART) for three-dimensional electron microscopy and X-ray photography’, Theory Biol., 1970, 29, pp. 471–481
[16] Atkinson, C., Soria, J.: ‘An efficient simultaneous reconstruction technique for tomographic particle image velocimetry’, Exp. Fluids, 2009, 47, pp. 553–568
[17] Gao, Q., Wang, H.P., Shen, G.X.: ‘Review on development of volumetric particle image velocimetry’, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2015, pp. 4541–4556
[18] Elisgina, G., Westerweel, J., Scarano, F., et al.: ‘On the velocity of ghost particles and the bias errors in tomographic-PIV’, Exp. Fluids, 2011, 50, pp. 825–838
[19] Michaelis, D., Novara, M., Scarano, F., et al.: ‘Comparison of volume reconstruction techniques at different particle densities’. 15th Int. Symp. on Applications of Laser Techniques to Fluid Mechanics, Lisbon, Portugal, 2010, pp. 3–17
[20] Worth, N.A., Nickels, T.B.: ‘Acceleratio of Tomo-PIV by estimating the initial volumen interior distribution’, Exp. Fluids, 2008, 45, pp. 847–856
[21] Lanarche, F., Leroy, C., Grapperon, P., et al.: ‘ImageNet classification with deep convolutional neural networks’, Proc. Adv. in Neural Information Processing Systems, Lake Tahoe, American, December 2012, pp. 1097–1105
[22] Gao, Q., Li, Q., Pan, S., et al.: ‘Particle reconstruction of volumetric particle image velocimetry with strategy of machine learning’, arXiv preprint arXiv:1909.07815, 2019
[23] Long, J., Shelhamer, E., Darrell, T.: ‘Fully convolutional networks for semantic segmentation’, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Columbus, Ohio, 2014, vol. 39, pp. 640–651
[24] Ronneberger, O., Fischer, P., Brox, T.: ‘U-Net: convolutional networks for biomedical image segmentation’, Int. Conf. on Medical Imaging Computer and Computer-assisted Intervention, Munich, Germany, 2015, pp. 234–241
[25] Jégou, S., Drozdzal, M., Vazquez, D., et al.: ‘The one hundred layers tiramisu: fully convolutional DenseNets for semantic segmentation’, IEEE Conf. on Computer Vision and Pattern Recognition Workshops (CVPRW), Honolulu, HI, USA, 2017
[26] Zhao, H., Shi, J., Qi, X., et al.: ‘Pyramid scene parsing network’, IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017
[27] Badrinarayanan, V., Kendall, A., Cipolla, R.: ‘SegNet: a deep convolutional encoder-decoder architecture for image segmentation’, IEEE Trans. Pattern Anal. Mach. Intell., 2017, 39 (12), pp. 2481–2495
[28] Ducett, S., Natale, A., Astaño, T.: ‘Spatial filtering improved tomographic PIV’, Exp. Fluids, 2013, 54, (4), p. 1505
[29] Huang, G., Liu, Z., Laurens, V.D.M., et al.: ‘Densely connected convolutional networks’, Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Las Vegas, Nevada, USA, 2016
[30] Ioffe, S., Szegedy, C.: ‘Batch normalization: accelerating deep network training by reducing internal covariate shift’. Int. Conf. on Machine Learning, Lille, France, 2015
[31] Hinton, G.E., Srivastava, N., Krizhevsky, A., et al.: ‘Improving neural networks by preventing co-adaptation of feature detectors’, arXiv preprint arXiv:1207.0580, 2012
[32] Ruder, S.: ‘An overview of gradient descent optimization algorithms’, arXiv preprint arXiv:1609.04747, 2016
[37] Kingma, D., Ba, J.: ‘Adam: a method for stochastic optimization’. Proc. Int. Conf. on Learning Representations (ICLR), Banff, Canada, 2014

[38] Zhu, J., White, C., Wainwright, D. K., et al.: ‘Tuna robotics: a high-frequency experimental platform exploring the performance space of swimming fishes’ Sci. Robot., 2019, 4, (34), p. eaax4615

[39] Lou, J., Yang, Y., Chen, T., et al.: ‘Oscillating performance and propulsion mechanism of biomimetic underwater oscillatory propulsion by resonant actuation of macro fiber composites’. Proc. Inst. Mech. Eng. C, 2019, p. 0954406219894352

[40] High, J. W.: ‘Method of fabricating NASA-standard macro-fiber composite piezoelectric actuators’ (National Aeronautics and Space Administration, Langley Research Center, 2003)

[41] Brox, T., Bruhn, A., Papenberg, N., et al.: ‘High accuracy optical flow estimation based on a theory for warping’. Proc. Eur. Conf. on Computer Vision, Berlin, Heidelberg, 2004, pp. 25–36