This is the version of the article before peer review or editing, as submitted by an author to International Journal of Social Robotics.

Title: Machine learning for flow field measurements: a perspective

Authors: Stefano Discetti - Yingzheng Liu

Version of record available online at
https://doi.org/10.1088/1361-6501/ac9991
Machine learning for flow field measurements: a perspective

Stefano Discetti¹ and Yingzheng Liu²

¹ Aerospace Engineering Research Group, Universidad Carlos III de Madrid, Leganés, Spain
² Gas Turbine Research Institute/School of Mechanical Engineering, Shanghai Jiaotong University, Shanghai, China

E-mail: sdiscett@ing.uc3m.es, yzliu@sjtu.edu.cn

Abstract

Advancements in machine-learning techniques are driving a paradigm shift in image processing. Flow diagnostics with optical techniques is not an exception. Considering the existing and foreseeable disruptive developments in flow-field measurement techniques, we elaborate this perspective, particularly focused to the field of Particle Image Velocimetry. The driving forces for the advancements in machine-learning methods for flow diagnostics in recent years are reviewed, and possible routes for further developments are highlighted.

Keywords: machine learning, flow-field measurements, image processing, particle image velocimetry

1 Introduction

Machine learning (ML) is a subfield of artificial intelligence that aims at using data to perform tasks without human intervention. Examples of such tasks are pattern identification, predictive analytics, data mining, and filtering. In recent years, ML techniques have had a disruptive effect on a wide variety of scientific and engineering fields. The widespread development of ML methods has led to remarkable advancements in several disciplines, disclosing new research pathways whose merit is clearly apparent. ML encompasses a range of techniques that can be classified into three groups:

- Supervised learning, which generally is aimed at identifying the mapping between a set of input data with (discrete or continuous) output labels. Classification is a classic example of a task performed using supervised learning methods. An artificial neural network (ANN) can be trained using labeled data to identify the class to which an unlabeled input belongs. Supervised learning is also often used for regression tasks, in which the value of an output variable is determined based on statistical correlation with input parameters.
- Unsupervised learning, which identifies relations between data without any form of supervision. Such techniques include association and clustering algorithms that group unlabeled input data based on similarities. Data compression and dimensionality-reduction techniques also belong to this group. These techniques leverage the concept that the information contained in input data typically lies in a low-dimensional space. Additionally, unsupervised learning techniques include generative models that generate samples after having observed the statistical distribution of training data.
- Reinforcement learning, in which an agent learns, through its interactions with an environment, strategies to maximize rewards and minimize penalties according to a metric associated with a certain goal to be achieved.
Among others, the field of computer vision has significantly benefited from the advancements in ML in the past decade. ML techniques for tasks such as automatic pattern recognition and tracking, image segmentation, filtering and quality enhancement, and super-resolution are being continuously developed. These innovations have been flowing progressively through the wide scientific community working on flow-field measurement techniques, such as particle image velocimetry (PIV), flow visualization, schlieren imaging, particle image thermometry, and any other technique involving the optical measurement of field quantities. The merits of ML methods for flow diagnostics have rapidly become apparent, leading to the emergence of promising research pathways.

This article summarizes the recent advancements in ML methods for flow-field measurements, focused on PIV, and highlights future developments in this discipline. Section 2 describes the techniques for image preprocessing. Section 3 focuses on data processing and conditioning techniques, specifically, direct extraction of velocity fields using ML (Section 3.1), resolution enhancement in space (Section 3.2) and time (Section 3.3), and data augmentation and filtering (Section 3.4). Section 4 presents the concluding remarks.

## 2 ML methods for PIV image preprocessing

The accuracy of PIV depends on the quality of the images to be processed. In addition to eliminating undesired laser reflections and maximizing the contrast of the particles with respect to the background, the image quality can be significantly enhanced using preprocessing techniques. Image preprocessing is particularly relevant for near-surface measurements. The importance of image preprocessing in PIV has increased with the development of volumetric methods [1]. For example, in tomographic PIV [2] and Lagrangian particle tracking (LPT) [3] approaches, the background intensity must be minimized to reduce the probability of reconstruction artifacts. Traditional preprocessing techniques used in PIV include background removal—subtraction of a reference image obtained through e.g. low-pass filtering in space [4][5] or time [6], computation of the historical minimum [7], or ensemble averaging [8]—and level normalization to increase the contrast of the particles with respect to the background, such as through intensity capping [9] or min/max filtering [10].

In recent years, the interest in automatic unsupervised image preprocessing has been increasing owing to the need to reduce human intervention and minimize the dependence on the user. Mendez et al. [11] proposed a method based on proper orthogonal decomposition (POD) of the PIV snapshots to leverage the difference in the space/time coherence of the background and particles. This method, known as eigenbackground subtraction, is based on the hypothesis that the images originate from the superposition of a low-rank contribution from large objects (i.e., the eigenbackground) and a sparse contribution from small moving objects. This unsupervised dimensionality-reduction technique was first introduced in the computer vision community to identify moving objects in videos [12]. In PIV images, the background (in particular, laser reflections) typically exhibits a high degree of spatial/temporal coherence, whereas the particles randomly seed the images. A low-rank approximation of the images, rearranged in the form of snapshot matrix (as in the POD snapshot method), provides an approximation of the background for each snapshot. Although the traditional methods also involve the statistical extraction of the background, the eigenbackground method inherits the characteristic of “learning from experience” typical of ML methods. Consequently, the approximation accuracy increases with the increase in the dataset size.

Additionally, deep learning techniques have been introduced in PIV preprocessing for image masking. Vennemann and Rösgen [13] proposed different methods based on convolutional autoencoders to perform automatic dynamic image masking. An autoencoder is a type of ANN that consists of an encoder and a decoder and has a framework similar to that of nonlinear principal component analysis [14]. Figure 1 shows a schematic of the architecture. The encoder maps the input data onto a low-dimensional latent space, and these data are used by the decoder to obtain an output (ideally) identical to the input. This architecture seeks to identify low-dimensional representations of the images to maximize the preservation of the relevant image content. Large coherent objects are retained, and the scattered particles are deleted, thereby identifying the background to be removed from the images. Autoencoder techniques provide a valuable solution for single-frame masking and can address moving backgrounds and semi-transparent objects with a competitive computational cost.

Recent advancements in ML techniques have also facilitated the deblurring of and distortion removal from PIV images. Particle image blurring typically occurs in the presence of fast-moving particles with respect to the image exposure time and duration of the light pulse. Motion blurring leads to correlation peak broadening in the direction of motion, and in certain cases, asymmetry owing to changes in the index of refraction [15]. Oh et al. [16] recommended the deblurring of PIV images by using generative adversarial networks (GANs, [17]). GANs involve two competing networks: a generator and a discriminator that contest each other in a zero-sum game. In the method reported by Oh et al. [16] the generator is trained using synthetic images to generate deblurred images in a wide range of blurring conditions. This approach can reduce the uncertainty and number of outliers. Gao et al. [18] proposed an actuator-less distortion-correction method based on a multiple-input deep convolutional
neural network (CNN) and Hartmann–Shack sensing. The method can overcome the typical bandwidth limitations of actuators and reduce the uncertainty introduced by image distortion.

Although ML-based contributions in preprocessing tasks for PIV are limited at present, advancements in computer vision in this field are expected to promote novel developments in this direction in the near future. Progress in this direction would extend the boundaries of PIV applicability, for instance, to industrial-flow measurements in which high image quality is typically difficult to achieve.

Figure 1 Autoencoder architecture adopted by Vennemann and Rosgen [13]. Reprinted by permission from Springer Nature Customer Service Centre GmbH: Springer, Experiments in Fluids [13].

3 Data processing and conditioning

This section describes the use of ML methods for PIV processing and potential for use in postprocessing. First, we review the main techniques being used to realize and enhance cross-correlation (CC) and particle tracking analysis using ML. Recent developments in the field of three-dimensional (3D) PIV are summarized. Subsequently, we spotlight the methods to enhance the spatial and temporal resolutions and increase the measurement accuracy a posteriori.

3.1 PIV processing through ML methods

3.1.1 Extraction of velocity fields using neural networks

The recent interest in ML methods to directly extract velocity fields from particle image sequences has been fueled by advancements in artificial intelligence in computer vision. However, the idea of using ML methods such as neural networks (NNs) for PIV and particle tracking velocimetry (PTV) image analysis can be traced to literature dating back nearly 30 years. Early studies used NNs for two-frame particle tracking [19][20][21][22][23][24] and feature identification/classification [25].

Readers interested in the initial research involving the application of NNs to planar PTV can refer to the review paper by Grant and Pan [26]. Later, researchers focused on developing NNs and genetic algorithms for multi-pulse [27] and 3D PTV [28][29][30][31][32]. Interestingly, NNs were also used as classifiers for 3D particle tracking with colored particles [33] to reduce the number of ambiguities in particle reconstruction.

In the last five years, many frameworks have emerged based on the adaptation of ML algorithms to PIV, fostered by hardware improvement, availability of open codes, and introduction of new concepts in the computer vision community. We focus on the following three types of these frameworks:

In the first category, NNs are trained to reproduce the outcome of CC. Lee et al. [34] proposed the use of cascaded deep CNNs (DNNs) [35] for PIV analysis. This architecture involves four levels, with the lowest level trained to estimate the bulk of the displacement and the higher levels used to extract the fine details. This approach also incorporates the discrete window offset. Rabault et al. [36] investigated the performance of CNNs and fully connected NNs in performing an “end-to-end” PIV
analysis. Both types of ANN were trained on synthetic images. These methods [34][36] achieved promising performances in terms of accuracy and spatial resolution, and the algorithms were computationally efficient when implemented over graphical processing units. However, CNNs cannot readily incorporate image deformation, which reduces their accuracy in the presence of velocity gradients. Recently, Morrell et al. [37] attempted to train Bayesian CNNs (BCNNs) to simultaneously extract the velocity fields and uncertainty. The authors tested different architectures that were simultaneously and separately fed image interrogation regions and CC maps. The BCNNs fed only CC maps outperformed the other configurations. This study represented the first implementation of probabilistic NNs in PIV and provided a promising pathway for ML-based PIV analysis with directly embedded uncertainty estimation. Notably, the abovementioned approaches require a large number of training images, potentially covering all conditions that may be encountered in the experimental application. Training is typically performed using synthetic images, which raises concerns regarding the generalizability of these methods. This problem can potentially be alleviated by creating vast standardized datasets.

The second category involves optical flow methods based on CNNs. Optical flow [38] techniques leverage the apparent motion of brightness patterns in image sequences and additional regularization constraints (e.g., smoothness) to obtain velocity
fields. Readers interested in these techniques can refer to the review paper by Heitz et al. [39] on fluid flow measurement with optical flow. Cai et al. [40] proposed a modified version of FlowNetS, a CNN-based optical flow model introduced by Dosovitskiy et al. [41], to estimate dense velocity fields estimation from PIV images. The algorithm, known as PIV-NetS, is trained offline with synthetic images. Figure 2 shows the architecture of FlowNetS and examples of velocity fields reconstructed using the proposed optical flow method trained on synthetic data. Cai et al. [42] introduced a modified version of LiteFlowNet [43], a more efficient version of FlowNet, for the same purpose. To avoid the requirement of large labeled datasets for training, Zhang et al. [44], inspired by the studies of Yu et al. [45] and Meister et al. [46], proposed an unsupervised version of LiteFlowNet, named UnLiteFlowNet, in which the training is based on photometric loss (i.e., the difference between the two frames and corresponding forward–backward warped images), smoothness, and consistency (the forward flow must be the inverse of the backward flow obtained by implementing the sequence backward in time). The main challenges associated with CNN-based methods pertain to their generalization and robustness in the presence of noise. To address these issues, Gao et al. [47] combined the capabilities of CNNs in achieving high resolution with the robustness of CC against noise. The method, referred to as CC-FCN, embodies two inputs, i.e., the particle images and low-resolution field obtained by CC. Lagemann et al. [48] proposed an NN-based approach known as recurrent all-pairs field transforms (RAFT) [49]. RAFT is composed of three parts: a feature encoder that extracts features on a pixel scale; a correlation layer that produces a four-dimensional correlation volume for all pairs of pixels; and (3) an update operator based on convolutional gated recurrent units that iteratively updates the flow field. Overall, the abovementioned optical flow approaches based on deep learning have achieved encouraging results with enhanced performance metrics in recent years.

The third category pertains to particle tracking, which has long attracted the interest of researchers as a potential application of ML methods. Grayver and Noir [50] explored the use of an ensemble of CNNs to extract the displacement and azimuth of streaks in particle streak velocimetry. This approach is particularly relevant, given the increasing use of LEDs for illumination, which often leads to streaky particle images owing to the larger exposure time. Furthermore, advancements in recurrent NNs (RNNs) in tasks such as time-series forecasting, speech, and handwriting recognition have fostered novel particle-tracking concepts. Architectures based on long short-term memory (LSTM) RNNs have been proposed [51][52] to predict particle positions given the past trajectory. Liang et al. [53] introduced an architecture based on a modified version of the end-to-end scene flow estimator FlowNet3D [54]. FlowNet3D contains three modules: hierarchical point feature learning based on PointNet++ [55], flow embedding based on two consecutive views of the point cloud, and upsampling and refinement. The authors modified the original algorithm to incorporate a convection architecture to iteratively enhance particle matching and refine the flow field.

In general, advances in computer vision are showing potential to improve PIV-based image analysis, especially the emergence of deep NN end-to-end motion estimators that use image sequences. Although NNs have been mainly used for feature matching in the last two decades, architectures based on CNNs have fueled innovation in the last few years, and continuous progress in computer vision is expected to promote the development of novel algorithms. Nevertheless, the significant challenges of blending ML methods with particle-based measurements must be addressed. The most relevant barrier to be overcome is the generalization of these approaches: How can the robustness of ML-based algorithms in experiments be enhanced if the training is based solely on synthetic images? A potential solution is to incorporate physical laws in the process to enhance the flow prediction. Moreover, probabilistic NNs that can directly estimate the uncertainty in addition to the velocity fields represent an interesting pathway.

### 3.1.2 Three-dimensional PIV with NNs

Recently, researchers have attempted to embed robust image-processing tools from the computer vision domain to extend PIV to 3D. The reconstruction of 3D particle distributions and/or flow features from a finite set of two-dimensional projections of particle images is a challenging task that is typically realized using robust object identification and/or unambiguous mapping to 3D. The exploitation of ML methods has opened promising avenues in this domain. The most recent approaches can be divided into the following three categories.

The first category aims to enhance the reconstruction of particles in the classic multi-camera arrangement used in tomographic PIV and LPT. Gim et al. [56] proposed the use of a shallow neural network to reconstruct 3D particle positions for particle tracking instead of performing traditional line-of-sight-based particle triangulation. The main advantage is the reduced computational cost. Using algebraic reconstruction techniques in tomography, Gao et al. [57] introduced a 3D CNN to regularize the output intensity distribution of a multiplicative line-of-sight framework [58]. This network is trained to minimize the backprojection error, and it can outperform two-frame reconstruction based on a spatial-filtered version [59] of the multiplicative algebraic reconstruction technique (MART, [60]). Liang et al. [61] indicated that the MART tomographic
reconstruction can be filtered using an encoder-decoder CNN to suppress ghost particles and regularize the reconstructed particle shape.

The second category pertains to the feature-based 3D reconstruction of particle images, for instance, in astigmatism PIV [62][63] and holographic PIV [64][65]. König et al. [66] used a cascaded DNN for particle detection and position estimation in astigmatism PIV. Although the DNN exhibits a higher error for the depth position than classical evaluation methods in synthetic test cases, the uncertainty is comparable to that in the experimental results. The authors argued that classical techniques tend to perform worse when the experimental conditions are inconsistent with the assumptions incorporated in such methods, whereas DNNs directly use the particle images for training. Several other approaches have been proposed to address problems related to particle overlap and low intensity [67]. Barnkob et al. [68] systematically compared the particle image model functions (based on shape or edge detection), normalized CC between measured particle images, and reference templates from optical calibration and DNNs to decode the depth position of particles from the defocusing patterns. The results showed that DNN-based methods are not adequately mature, but the fast-paced progress of ML may change this landscape in the near future. In the case of holographic PIV, most of the efforts have been focused on exploiting the existing CNN architectures to reconstruct particle images from holograms. In recent studies [69][70], U-net, a CNN architecture for medical image segmentation [71][72], was used to accomplish this task. The abovementioned methods can outperform state-of-the-art conventional reconstruction methods for holograms. However, the flow-field estimation capabilities of these approaches remain to be assessed.

In the third category, physical constraints are introduced in the reconstruction process. Most recent approaches use physics-informed NNs (PINNs, [73]) that are trained to minimize a loss function based on the residual of the governing equations of the problem. The potential of PINNs in regularizing the tomographic reconstructions of scalar quantities from a limited set of projections [74][75] has been demonstrated, and these architectures are expected to undergo significant development in the near future. Given the widespread use of PINNs to increase the data reach instead of enhancing the tomographic reconstruction quality, Section 3.4.2 further discusses their implications.

In summary, the influence of ML techniques on 3D PIV is gradually becoming significant. The robust feature extraction capabilities of existing architectures and simplicity in adding constraints (based on either the consistency of projections or physics-based regularization) are expected to drive future research in this direction.

3.2 Spatial resolution enhancement through ML

Standard PIV techniques exhibit limited resolution owing to the spatial averaging of the particle data. Thus, super-resolution is a research hotspot in the PIV community. Super-resolution capabilities can increase the dynamic range and enable the complete resolution of the turbulent scales. This requirement becomes stricter as the Reynolds number increases [76]. Recent years have witnessed significant progress in super-resolution has made giant leaps in the last years thanks to deep learning, thus such progresses are expected to flow (and are currently flowing) to the PIV field. The critical concern in this case, which differs fundamentally from those of conventional spatial interpolations, is the involvement of fine-scale information in the learning database. In early studies, POD was widely used to transfer the information from a precursor database to new complemented flow vectors. With the development of artificial-intelligence techniques, ML has emerged as a power train, not only because of the algorithmic flexibility but also its ability to enhance the flow dynamic features.

PIV super-resolution reconstruction typically relies on building a database containing rich small-scale information. POD or ML techniques can create a mapping between this database and the flow measurements to be enhanced. Gunes and Rist [77] developed a direct numerical simulation (DNS)-based technique for the super-resolution reconstruction of stereo-PIV measurements in a transitional boundary layer flow. The main hypothesis was that the POD modes computed using the DNS data exhibit a high spatial resolution and contain abundant fine-scale flow structures. The time-varying mode coefficients were determined by cross-projecting the instantaneous PIV measurements onto the POD modes. Notably, this approach requires precise DNS data with a consistent flow configuration. Furthermore, the availability of comprehensive DNS data may raise questions regarding the need for conducting PIV experiments. Considering these aspects, He and Liu [78] used high-resolution (HR) pulsed PIV data instead of DNS data to establish the super-resolution database. The coefficients were computed by projecting the low-resolution (LR) time-resolved (TR) PIV fields onto the HR POD modes. This approach can be applied in most flow conditions by increasing the resolution to the standard double-pulse (shortened to “pulsed” henceforth) PIV level without the loss of high sampling frequency, after performing the HR pulsed PIV and LR TR-PIV on the same flow.
POD approaches must transfer the small-scale information between HR and LR fields during cross-projection. In this case, part of the small scales may be lost even if least-squares fine-to-coarse mapping is performed [78]. By establishing a dataset using the method described above, ML techniques can be used to achieve super-resolution using different frameworks, for example, GANs [17]. As mentioned above, GANs consist of (generally DNNs): generator G and discriminator D. Agent G is fed by LR fields and trained to generate HR fields that are as similar as possible to the ground truth (typically, a simultaneously measured HR dataset), whereas D is trained to distinguish the generated HR fields from the original fields. After sufficient training, D cannot distinguish the output results from the ground truth, which means that G can generate high-fidelity HR fields. Deng et al. [79] proposed a general strategy using the super-resolution GAN (SRGAN, [80]) and enhanced SRGAN, with deep NNs to enhance the system performance by allowing complex mappings. The generator is trained using the LR fields and their corresponding HR counterparts to learn the relationship between these entities. Unlike this static CNN (SCNN) approach, in which only the individual instantaneous LR snapshots is used as the input, a multiple temporal paths CNN (MTPC) [81] can simultaneously consider spatial and temporal information by using a time series of LR velocity fields as the input. The
hypothesis behind this approach is that the turbulent processes are nonlocal in both time and space. Figure 3 schematically illustrates the MTPC and a comparison of the LR, bicubic interpolations, SCNN, MTPC, and DNS fields of a turbulent channel flow at a resolution upscaling factor equal to 4. The MTPC consists of three temporal paths and a weight path. Each temporal path consists of a feature extraction module and reconstruction module, and the weight path concatenates all the features extracted in the temporal paths as the input of the reconstruction module. Liu et al. [81] reported that the MTPC cannot exactly reproduce the spatial structures with a kinetic energy several orders of magnitude smaller than the total energy or and showed the limits in recovering precisely recover all the physical properties associated with the anisotropic turbulence.

Recently, Fukami et al. [82] incorporated sparse sensor data in a CNN by using a Voronoi tessellation to obtain a structured representation. This process can in principle be applied to vector data extracted by PTV. In these approaches, the CNN estimator requires complete HR fields for training.

Although ML-based super-resolution methods have achieved promising results, labeled HR data are required for training, which is the widest source of criticism and the largest limitation for their application and generalization ability. In practice, the training database already contains all the data that we want to produce using either POD reconstruction [77] or ML techniques [79][81][82]. Therefore, subsequent research is expected to be focused on eliminating the requirement of labeled LR–HR pairs, and several encouraging results have already been obtained. Cortina-Fernandez et al. [83] leveraged gappy POD for super-resolution reconstruction, using a principle similar to that of ensemble PTV. The process flow of this algorithm is shown in Figure 4. The original LR fields are used to compute a POD temporal basis. The POD coefficients are cross-projected onto the PTV fields obtained on the same images to obtain the HR POD spatial modes. Super-resolution is thus realized by the reconstruction based on the POD time coefficients from the LR dataset and HR POD spatial modes. The novelty of this approach is that an HR dictionary is not required in advance, and super-resolution is achieved through direct construction based on the PIV snapshots and particle images. This framework endows flexibility in application to measurement data in settings in which benchmarks are not available. This strategy can also be applied with ML techniques. Güemes et al. [84] proposed a randomly seeded super-resolution GAN (RaSeedGAN) to produce HR fields, with the input and output being LR flow fields and sparse

Figure 4 Process flow of the data-enhanced PTV proposed by Cortina-Fernandez et al. [83]. Reprinted from [83] with permission from Elsevier.
PTV measurements, respectively. The principle is that the particles directly sample the HR field with minimal spatial filtering, and thus, the PTV data provide incomplete (a.k.a. gappy) views of the underlying HR fields.

Although ML has emerged as a promising tool in image and multimedia super-resolution processing in recent years [85][86][87][88], its application to fluid mechanics remains challenging. ML super-resolution techniques have been preliminary applied in the turbulence community, but large databases are required for training. Moreover, the weak generalization properties of ML models bring certain resistance to wide application. The use of existing and easily accessible data in the training process instead of a full set database can help introduce new paradigms in the fluid measurement community.

### 3.3 Temporal resolution enhancement (TRE)

TRE is a key aspect in PIV measurements owing to the limited sampling frequency of the hardware. TRE is analogous to the frame interpolation procedure, which is commonly used in computer vision and video processing [89][90][91][92]. However, achieving TRE in flow measurement scenarios is challenging owing to the high dimensionality of turbulence and its direct link to physical constraints, as TRE is not simply aimed at achieving reasonably realistic time dynamics but instead reflecting real physics laws.

TRE is typically achieved by introducing an advection model [93][94][95]. Nearly all data-driven TRE strategies are based on flow reconstruction using high-frequency probe signals. In recent approaches, ML methods have been used to estimate the time dynamics, with roots pertaining to linear stochastic estimation (LSE, [96]) and extended POD (EPOD, [97]). We briefly describe LSE as a solid start for ML, which has experienced rapid development in the last five years.

The nonlocal properties of turbulent flows represent the foundation of LSE for TRE. Discrete probes, which are inexpensive and can realize high-frequency-sampling measurements, capture not only the temporal information but also (partially) the spatial information in flow fields. In POD-based approaches, LSE was typically performed over single-/multipoint measurements and POD coefficients of the synchronized PIV measurements to estimate the time-varying LSE-POD coefficients [98][99][100]. The spatiotemporally varying low-order flow field was reconstructed using the first most energetic POD mode and estimated coefficients. This process was later improved by using the multi-time-delay LSE-POD approach and Kalman smoother [101][102][103]. This approach relies on the coupling mechanism between the time-accurate information of the pointwise quantity and flow fields and is suitable for flows with compact POD spectra, such as periodic or quasi-periodic flows. The resultant time-resolved flow fields are built targeting the energetic large-scale flow structures represented by the leading POD modes. Notably, this approach is prone to overfitting as it amplifies the singular eigenvalues with small amplitudes. Recently, Podvin et al. [104] proposed a new variant of the LSE technique in which overfitting is prevented through the direct mapping of the POD amplitudes from the measurement space to the state space. This method has been successfully applied in the flow reconstruction of a turbulent boundary layer.

Hosseini et al. [105] proposed a multi-time-delay embedding, assuming explicitly that the temporal information collected by the probes contains spatial information in the case of advection-dominated flows. This procedure is known in meteorology as multichannel singular spectrum analysis [106]. Discetti et al. [107] extended this work by introducing a robust truncation criterion to remove the uncorrelated part of the signal from the reconstructed flow fields. This LSE approach can manage turbulent flows without any dominant shedding frequency and reduce the time jittering in the rebuilt history of the time coefficients.

Notably, a natural limitation still exists in the LSE when the number of high-frequency probes is considerably smaller than the degree-of-freedom of the flow fields. This aspect limits the number of POD modes that can be used in the LSE reconstruction and may induce errors in amplitude computation. This limitation becomes significant when the flow is broadband and the leading POD modes contain a small amount of kinetic energy. Therefore, He and Liu [108] used a data assimilation technique to recover the high-order dynamic features of the separated flow from the reduced-order noisy LSE reconstruction. This model-based data-driven prediction strategy combined the measurement data and predictive model and could reproduce dynamic information that was initially absent in the measurement data. Recently, Chen et al. [109] have demonstrated that TRE based on a combination of fast probes and non-time-resolved PIV can enable the estimation of derivative quantities such as the pressure field.
With the rapid spreading of ML methods, numerous studies on TRE based on DNN architectures have been conducted in the last three years. Unlike LSE and EPOD, which are linear techniques, ML can introduce nonlinear mappings, enabling the exploitation of the relation between the time-resolved probe data and non-time-resolved field data. Deng et al. [110] reconstructed the time-resolved turbulent flow from high-frequency discrete point measurements and non-time-resolved PIV data by using an LSTM NN. An LSTM [111] is a type of RNN designed to prevent the NN output for a given input from either decaying or exploding as it cycles through the feedback loops. LSTMs outperform other RNN architectures by alleviating the vanishing gradient problem. Figure 5 shows the architecture of the LSTM-based POD model built to learn the relationship between the downsampled velocity signals and POD coefficients of the PIV data. The time-resolved POD coefficients were estimated by feeding the LSTM with high-frequency velocity signals. The results demonstrated that the LSTM-based POD model can realize time-series reconstruction by successfully recovering the temporal evolution of the POD coefficients for even high-order POD modes. Giannopoulos and Aider [112] used a fully connected focused time-delay neural network (FTDNN) for reduced-order system identification using local upstream probes. FTDNN is a standard feedforward architecture including a tapped constant time-delay in the input. This model predicts the global instantaneous flow fields without requiring time-consuming full-field decompositions and intrusive measurement devices. Jin et al. [113] noted that the time sequence of the velocity contained information regarding the spatial distribution of velocity. In other words, the time sequence of the velocities in the spatial domain could be replaced by a time series of the velocity to alleviate the ill-conditioned problem in TRE reconstruction. Consequently, bidirectional RNNs with gated recurrent units were designed to learn the time-resolved POD coefficients of the first few modes based on both PIV and probe measurement data. This approach achieved higher accuracy than the extended POD, especially in cases involving large Reynolds numbers and high-order coefficients.
Overall, both LSE and ML techniques for TRE rely on dimensionality reduction of the flow-field features to simplify training. At present, POD is the most commonly used approach. ML-based dimensionality-reduction architectures, such as convolutional autoencoders, are promising candidates to enhance the robustness of TRE based on probe data. Owing to the nonlinear properties of NNs, ML techniques generally yield more accurate results than LSE, but their training process is computationally intensive and requires large databases. Alternative POD-based TRE strategies can be established using least-squares regression or compressed sensing [114], but with similar high-order mode truncation. Model-based assimilation can be performed using the advection equation [115] or Navier–Stokes (N–S) equations [108]. These approaches are expected to recover rich turbulent structures without sharp truncation. Future work can be focused on developing novel ML-based TRE approaches, for example, by combining data assimilation with predictive models.

3.4 PIV data enhancement

Data enhancement includes data augmentation, i.e. the process of artificially increasing the amount of data by generating new data from existing data, and data improvement by reducing the error and noise of the original dataset. This section describes the ML methods for flow-field filtering, outlier correction, and increasing the data reach by enforcing physical constraints.

3.4.1 Filtering and outlier correction

The flow fields measured by PIV/PTV contain both systematic and random errors. Systematic errors attributable to the finite spatial resolution represent an intrinsic limitation of the processing algorithm and can be alleviated by using advanced processing techniques. Random errors can be mitigated through filtering, albeit at the expense of the spatial resolution. Another source of error is the presence of outliers, which are vectors that are significantly different from the surrounding vectors, resulting from the local failure of the image processing algorithm. Although median validation is an established process for isolated outlier removal, the removal of cluster of outlier vectors represents a challenge.

To address random errors, data-driven methods have been widely used for the feature-oriented filtering of PIV data. POD, as a valuable tool for noise reduction, exploits the truncation when discarding the high-order mode in the reconstruction. In this context, determining the threshold for the truncation is important. Several criteria for mode selection have been proposed based on POD truncation [116][117][118][119], such as singular value decay, signal-to-noise ratio, and intrinsic errors of PIV measurements. The objective is to minimize the PIV noise while retaining the kinetic energy of the real velocity fluctuations in the flow. POD-based filtering approaches outperform the commonly used constant energy rejection criterion in POD reconstruction (for example, using the leading POD modes containing 95% of the total kinetic energy) and standard frequency-based filtering. However, the spatial modes selected for reconstruction are sensitive to noise, as they are computed from the original noisy data. In this context, filtering processes directly imposed on the PIV algorithm or fusions of flow governing equations [120] are more promising alternatives to address random errors.

Outlier detection is widely performed using the statistics of the surrounding vectors, such as the local mean and local median [121][122]. However, these methods may lead to misdetections when spurious vectors are present in the neighborhood. Raben et al. [123] introduced a gappy POD procedure to replace the identified outliers. In this process, reconstruction was performed using a large number of modes, and the stopping criterion was based on the field smoothness. Wang et al. [124] proposed a POD-based outlier correction method (POD-OC) to dynamically approximate the original pure velocity field. This method reconstructed a reference velocity field using low-order POD modes to detect and correct outliers. POD-OC could not only efficiently identify the scattered spurious vectors but also detect and correct the clustered vectors. Although POD provides a principled approach to decompose a high-dimensional fluid flow into a hierarchy of orthogonal modes, the least-squares regression underlying these approaches is susceptible to outliers and gappy data.

Scherl et al. [125] proposed a robust principal component analysis (RPCA) for the model decomposition of corrupt fluid flows. Figure 6 schematically illustrates the RPCA filtering framework applied to corrupt flow-field data to remove noise and outliers in the flow past a cylinder. The RPCA uses sparse optimization to separate a data matrix into a low-rank matrix containing correlated structures and a sparse matrix containing spurious entries and is thus a promising method for outlier detection and correction in PIV measurements.
The limited use of DNN architectures for postprocessing of PIV data can be explained by the wider margin of improvements that can be achieved by directly embedding ML in the processing stage. The large number of solutions explored in this direction (Section 3.1) are a direct consequence of the availability of off-the-shelf solutions for image processing tools in the computer vision domain. Although advancements in data regularization are foreseeable, we envision stronger efforts in developing data augmentation techniques based on physical principles, as described in the following section.

### 3.4.2 Increasing the data reach

In complex geometries, harsh flow conditions, and difficult optical access for laser illumination and image acquisition, PIV/PTV measurements may be severely flawed. Data corruption can manifest as data gappiness (for instance, because of non-uniform seeding or illumination source shadows) or the measurement of flow features that are highly discrepant with physical laws.

Several data-driven methods for repairing gappiness are based on mode decompositions such as gappy POD [123] and data fusion [126]. These approaches establish a global modal base that can be used for the gappy flow reconstruction using mode coefficients computed from the local flows. Bui-Thanh et al. [127] applied gappy POD to the flow around an airfoil and showed that the flow field could be successfully reconstructed from snapshots with 30% data missing. However, the gappy POD assumes that the missing portions are randomly distributed in different snapshots, and thus, it cannot be applied in situations in which the missing region is fixed (e.g., in the presence of a shadow owing to illumination obstruction). Data fusion [126] can address these fixed missing regions by acquiring multiple datasets with different camera layouts. Among the recent trends in reconstruction based on ML, its application to field repairing is particularly interesting. Morimoto et al. [128] proposed a CNN model to estimate the velocity fields from particle images with missing regions. Figure 7 presents an overview of the ML-based experimental flow data estimation. The authors used artificial particle images as the input data (mimicking the PIV images), and the output data were the velocity fields. Notably, labeled data for training were provided through DNS, and thus, the method could be applied only low Reynolds number flows. The CNN model estimated the velocity fields with considerably denser data.
than those processed using the conventional CC method and could obtain finer and hidden structures of the flow field that could not be resolved using the conventional CC method.

Figure 7 Overview of the machine-learning-based experimental flow data estimation. (a) Preparation of training data set. The artificial particle image (API) at time instant \( q_r \) is generated by randomly distributing particles. The API with time increment \( q' \) can be obtained using the velocity data \( u_{DNS} \). (b) Schematic of the API used for the input to the machine-learning model, \( q = q_r + q' \). (c) Training of machine-learning model \( F \) for full data. Reprinted from [128], with the permission of AIP Publishing.

Figure 8 Architecture of a PINN. The inputs are space coordinates \( X = (x,y,z) \) and time \( t \), and the outputs are velocity \( U = (u,v,w) \), and pressure \( p \). The physical laws are represented by incompressible N–S equations and expressed using automatic differentiation operators [132]. Reprinted from [132], with the permission of AIP Publishing.

ML techniques have also brought new impetus for PIV data augmentation by enforcing physical constraints. PINNs [73] are architectures trained to estimate the flow quantities by adding the known differential equations directly into the loss function during training. PINNs can predict solutions located between experimental data points and do not need labeled data. Therefore,
the performance of these architectures is expected to be superior to those of simple DNN flow estimators. PINNs have been used to reconstruct the velocity and pressure distributions from flow visualizations [129] or to solve the incompressible N–S equations based on the velocity–pressure and velocity–vorticity forms [130][131]. Cai et al. [75] used PINNs to infer the full 3D velocity and pressure fields from snapshots of 3D temperature fields obtained by tomographic background-oriented schlieren imaging. The PINNs could integrate the governing equations and temperature data, similar to adjoint-based data assimilation methods [120]. Wang et al. [132] used a PINN to reconstruct dense velocity fields with considerably lower noise levels than those of sparse tomography PIV data and to predict the pressure fields. Figure 8 shows a classical architecture of PINNs for velocity and pressure reconstruction with space coordinates and time. Although PINNs cannot replace traditional methods at present owing to the large training cost and limited accuracy, they represent a novel data assimilation approach and have received increasing attention in the fluid community.

4 Discussion and conclusions

This perspective summarizes the advancements in data-driven and ML methods to enhance flow-field measurements. Recent years have witnessed widespread emergence of approaches based on ML techniques. This progress is expected to continue in subsequent years given the potential impact of ML in several branches of science and technology, and especially the growing interest of the fluid measurement community. The characteristic time between the innovation chain of computer vision and fluid flow measurements through optical techniques has been significantly shortened.

In our perspective, we envision the following challenges in the future development of ML methods of fluid flow measurements:

- **Alleviating the need for labeled training data.** Several approaches for fundamental tasks such as spatiotemporal resolution enhancement, image processing, and data corruption mitigation require labeled training data, which cannot be acquired in most practical cases. When these data are available, the need for advanced processing is reduced or eliminated. Efforts to develop unsupervised or semi-supervised approaches have been made, and we expect further research in this direction. Fluid-targeted design of processing techniques based on ML is expected to have more disruptive potential than the direct application of ML tools to fluid measurements.

- **Improving generalizability and robustness.** ML methods are often criticized for their poor capability of generalizing beyond the training dataset or training conditions. This issue is not novel for developers of PIV/PTV algorithms. The generation of synthetic images that are representative of the real experimental conditions has been a longstanding challenge. For ML–based methods the problem has transformed to the development of robust test cases and establishment of novel standards to be accepted and shared by the community. Joint efforts such as the PIV Challenge [133] might play an important role in this direction in the future.

- **Understanding and quantifying the resolution and accuracy limitations.** The backbone of PIV and PTV is a consolidated elegant theory that allows the prediction of the resolution limits with analytical tools. In the recent decade, the uncertainty estimation of PIV has reached a mature state, with an arsenal of techniques now available for direct use and uncertainty propagation. ML methods introduce a novel degree of complexity, which necessitates the development of tools to estimate the limits of the processing algorithms. Physical consistency is expected to be a key player in this process.

- **Bridge the gap between measurement and data analysis.** Although measurement science in fluid mechanics has been traditionally targeted to obtain high-fidelity as-complete-as-possible data for flow characterization, the portfolio of ML tools opens the scenario to different approaches, directly targeting to data reduction embedded in the process.

5 Acknowledgments

Stefano Discetti acknowledges funding from the European Research Council (ERC) under the European Union’s Horizon 2020 research and innovation programme (grant agreement No 949085). Yingzheng Liu acknowledges financial support from the National Natural Science Foundation of China (11725209).

6 References


