

UNSUPERVISED RECURRENT ALL-PAIRS FIELD TRANSFORMS FOR PARTICLE IMAGE-VELOCIMETRY

C. Lagemann^{1*}, M. Klaas¹, W. Schröder¹

¹ RWTH Aachen University, Chair of Fluid Mechanics and Institute of Aerodynamics, Aachen, Germany

* c.lagemann@aia.rwth-aachen.de

Abstract

Convolutional neural networks have been successfully used in a variety of tasks and recently have been adapted to improve processing steps in Particle-Image Velocimetry (PIV). Recurrent All-Pairs Fields Transforms (RAFT) as an optical flow estimation backbone achieve a new state-of-the-art accuracy on public synthetic PIV datasets, generalize well to unknown real-world experimental data, and allow a significantly higher spatial resolution compared to state-of-the-art PIV algorithms based on cross-correlation methods. However, the huge diversity in dynamic flows and varying particle image conditions require PIV processing schemes to have high generalization capabilities to unseen flow and lighting conditions. If these conditions vary strongly compared to the synthetic training data, the performance of fully supervised learning based PIV tools might degrade. To tackle these issues, our training procedure is augmented by an unsupervised learning paradigm which remedy the need of a general synthetic dataset and theoretically boosts the inference capability of a deep learning model in a way being more relevant to challenging real-world experimental data. Therefore, we propose URAFT-PIV, an unsupervised deep neural network architecture for optical flow estimation in PIV applications and show that our combination of state-of-the-art deep learning pipelines and unsupervised learning achieves a new state-of-the-art accuracy for unsupervised PIV networks while performing similar to supervisedly trained LiteFlowNet based competitors. Furthermore, we show that URAFT-PIV also performs well under more challenging flow field and image conditions such as low particle density and changing light conditions and demonstrate its generalization capability based on an out-of-the-box application to real-world experimental data. Our tests also suggest that current state-of-the-art loss functions might be a limiting factor for the performance of unsupervised optical flow estimation.

1 Introduction

Particle-image velocimetry (PIV) is one of the key techniques in modern experimental fluid mechanics used to determine the velocity components of flow fields in a wide range of complex engineering problems. Current processing tools usually compute the most probable particle displacement of two consecutive particle images based on the cross-correlation between corresponding interrogation windows. This, in fact, always yields a spatially averaged optical flow output since a single displacement vector is estimated for an entire interrogation window. State-of-the-art algorithms additionally use a wide range of other elements including subpixel interpolation, multigrid correlation schemes, automatic outlier detection, and window deformation according to local velocity gradients. Usually, these approaches fully compensate for the loss-of-correlation due to in-plane motion if the flow within the final interrogation window is homogeneous or linearly varying. However, if the displacement is more complex due to unresolved fluctuations, non-constant velocity gradients, or out-of-plane displacement, the correlation peak is broadened and its intensity is reduced. The estimated mean field matches the ground truth fairly well, but velocity fluctuations are usually underestimated. A similar bias error can be observed in cases of inhomogeneously distributed tracer particles which is typical for near-wall flows.

Motivated by the limitations of current, cross-correlation based approaches, PIV analysis based on new ideas of deep learning in end-to-end optical flow applications was proposed to effectively learn dense displacement fields going far beyond the spatial resolution of the current gold-standard. In contrast to existing methods, these approaches are general, near-automated, and yield per-pixel flow estimates. These methods side-step the problem of manually designing an analytical pipeline by defining an end-to-end network whose

output is the dense per-pixel optical flow field. Thus, fine flow structures can be resolved which alternatively are smoothed due to the spatial averaging inherent to traditional cross-correlation based methods. The first end-to-end application using CNNs for PIV processing was introduced in Rabault et al. (2017). They trained different shallow convolutional and fully connected neural networks to predict the particle displacement of various synthetic test cases. However, the proposed networks were only applied to relatively simple test flows and ultimately did not achieve a competitive accuracy compared to available state-of-the-art PIV algorithms. A different evaluation scheme called PIV-DCNN was proposed in Lee et al. (2017). It consisted of a four-level regression convolutional neural network where each level was trained to predict a velocity vector from two input image patches. The network was verified to achieve similar results compared to standard PIV methods based on a single cross-correlation pass including window deformation. Due to its stacked architecture, PIV-DCNN suffers from large computational cost and low efficiency. In Cai et al. (2019b) a dense particle motion estimator was developed, PIV-FlowNetS, which was mainly based on FlowNet, a deep optical flow architecture introduced in Dosovitskiy et al. (2015). This motion estimator extracts feature maps of particle images and predicts a dense displacement field for synthetic and experimental particle images. It achieves a good accuracy with a higher spatial resolution compared to standard correlation-based PIV algorithms. Follow-up work adopted an advanced LiteFlowNet architecture proposed in Hui et al. (2018) which allowed to significantly improve the accuracy. Recently, a new PIV processing network called *RAFT-PIV* was presented in Lagemann et al. (2021) outperforming existing neural based PIV methods significantly. The underpinning optical flow backbone of this approach are *Recurrent All-Pairs Field Transforms (RAFT)* proposed in Teed and Deng (2020). RAFT differs from other optical flow networks in that it operates at a single resolution using a large number of lightweight, recurrent update operators. First empirical results demonstrate clear improvements of RAFT-PIV on challenging synthetic benchmark and experimental examples, relative to both classical approaches and existing optical flow learners.

All these approaches share the main idea of using supervised training on labeled synthetic PIV images. However, supervised learning of robust optical flow estimation requires a sufficiently large dataset of training images alongside ground truth optical flow information. The computation of reliable ground truth data for real image sequences within a reasonable time is almost impossible yielding inevitably the generation of synthetic datasets. However, the huge diversity in fluid flows and varying particle image conditions in experimental environments probably outmatch the data distribution which can be covered by artificially rendered PIV images resulting in an inherent distribution mismatch between training and test time domain. One potential solution to tackle this mismatch is the application of an unsupervised loss objective. In general, unsupervised learning paradigms have a major advantage since the loss objective is purely based on geometric penalty terms and hence, no ground-truth is required. As a consequence, unsupervised networks can be trained directly on real experimental data and therefore, might remedy the need of a general synthetic dataset while boosting the inference capability of a deep learning model such that the neural method can deal with arbitrary challenging real-world experimental data. First, Zhang and Piggott (2020) exploited the unsupervised loss formulation of Meister et al. (2017) and extended their unsupervised learning strategy to PIV application using a LiteFlowNet network design. Their loss consists of a photometric loss between two consecutive image frames, a consistency loss in bidirectional flow estimates, and a spatial smoothness loss. It is demonstrated that this method achieves competitive results compared to classical PIV algorithms and supervised architectures. However, a comprehensive application to various real-world experimental data is still missing. Therefore, we study the effectiveness of unsupervised learning paradigms in the context of RAFT-PIV and introduce an unsupervised RAFT model denoted *URAF-T-PIV*.

2 Method

The baseline model of our approach is RAFT-PIV introduced in Lagemann et al. (2021). Similar to the original architecture, RAFT-PIV extracts per-pixel features, computes a 4 level multi-scale correlation volume for all pairs of pixels, and iteratively updates a flow field using a convolutional gated recurrent unit. Details can be found in Teed and Deng (2020). Compared to other optical flow networks, it is unique in the sense that it operates at a single resolution using a large number of lightweight, recurrent update operators.

Given a pair of grayscale particle images, $\mathbf{I}_1, \mathbf{I}_2$, optical flow estimation requires to predict a dense displacement field $(\mathbf{f}^1, \mathbf{f}^2)$ mapping each pixel of \mathbf{I}_1 to its corresponding coordinates in \mathbf{I}_2 . RAFT mainly

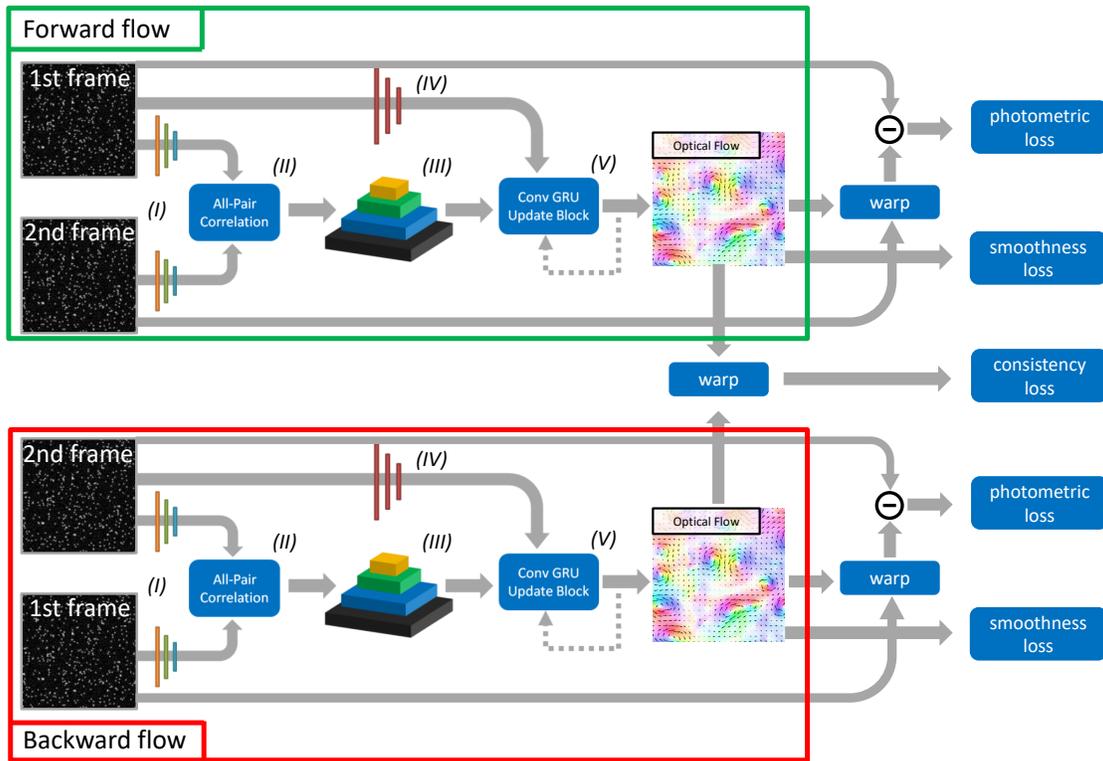


Figure 1: Schematic visualization of URAFT-PIV and its main components: A shared feature encoder (I) extracts per-pixel features from both input images. Based on the All-Pairs correlation (II) a 4D correlation volume is computed and subsequently stacked to form a correlation pyramid (III) by pooling the last two dimensions from level to level. The context encoder (IV) sharing the topology of the feature encoder computes a context map of the first image frame. The convolutional GRU (V) takes context map and correlation volume as input and recurrently updates the optical flow estimation. The loss objective of URAFT-PIV comprises a photometric, forward/backward consistent, and smoothness oriented penalty term. The photometric loss penalizes the photometric difference between the initial and the subsequent image which is warped according to the local flow prediction. An individual forward and backward flow is estimated by reversing the order of the input images. The consistency loss penalizes differences between forward and backward flow. Finally, a smoothness penalty term is applied to forward and backward flow separately.

consists of three stages, a feature extracting block, the computation of a full correlation volume between all pairs, and iterative updates based on a convolutional Gated Recurrent Unit (Conv GRU). A schematic visualization of URAFT-PIV can be found in Fig. 1. Furthermore, we introduced a cost volume normalization and waive the spatial downsampling within the feature extraction block. This allows a state-of-the-art performance while showing a strong generalization ability in direct real world applications.

To achieve an unsupervised learning scheme, we apply a loss objective which comprises photometric, forward/backward consistent, and smoothness oriented penalty terms. The photometric loss encourages the optical flow network to align image patterns by penalizing the photometric difference between the initial and the subsequent image which is warped according to the local flow prediction. In fact, a bi-directional photometric loss based on the generalized Charbonnier loss is used which simply can be achieved reversing the order of the input data. Thus, an individual forward and backward flow is estimated. To account for occluded pixels, which by definition do not have a valid counterpart in the other image, we compute an occlusion mask based on a forward/backward flow check. The consistency loss penalizes differences between forward and backward flow, again applying a generalized formulation of the Charbonnier loss to the difference between forward and backward flow. To address the aperture problem, e.g., motion estimation of regions with insufficient image structure as present in sections in-between particles, we apply an edge-aware

first-order accurate smoothness function to forward and backward flow. Extensive hyperparameter studies are performed to identify proper weights for the specific objective penalty terms. Furthermore, the loss of each optical flow iteration is weighted exponentially forming the final sequence loss. The sequence loss reads

$$L = \sum_{i=1}^N \gamma^{N-i} (L_{photo}^i + L_{con}^i + L_{smooth}^i) \quad (1)$$

where $L_{photo}^i, L_{con}^i, L_{smooth}^i$ denote the photometric, forward/backward consistent, and smoothness oriented penalty term of the iteration i and γ is the exponential weight. Similar to Teed and Deng (2020), we choose $\gamma = 0.8$. The evaluation metric is the Averaged Endpoint Error (AEE) representing the Euclidean distance between the final estimated (es) of the N -th iteration and ground truth (gt) optical flow of the test case being averaged over all pixels and reads

$$AEE = \|\mathbf{f}_{es,N} - \mathbf{f}_{gt}\|_1. \quad (2)$$

The computational graph of URAFT-PIV is implemented in the open source framework PyTorch (Paszke et al. (2017)). During training, we apply an Adam-Optimizer (Kingma and Ba (2014)) starting at an initial learning rate $\epsilon_0 = 0.0001$. Furthermore, the learning rate is reduced by a factor of five once the evaluation metrics stopped improving for 15 consecutive epochs. The minimum learning rate is set to $\epsilon_{min} = 10^{-8}$. All computations are run on multiple GPU nodes simultaneously each equipped with four nvidia A100.

We compare the test results of URAFT-PIV to our state-of-the-art inhouse code PascalPIV. To allow particle shifts greater than half the interrogation window size, the image evaluation uses a multi-grid approach with integer window shift to get an initial displacement field. Then, the displacement field is refined using an iterative predictor-corrector scheme with subpixel accurate image deformation according to the procedure described in Astarita and Cardone (2005). The initial displacement is interpolated for each pixel of the image using a third-order B-Spline interpolation. Both images are deformed by half the displacement to get a second-order accurate estimate of the displacement field. The image interpolation uses Lanczos resampling, i.e., Lanczos windowed cardinal sine interpolation, incorporating the neighboring $8 \times 8 \text{px}^2$. An integral velocity predictor is used to ensure convergence of the iterative scheme (Schrijer and Scarano (2008)). Hence, the predictor is the weighted average of the per-pixel displacement over the interrogation window. The corrector is determined by evaluating the cross-correlation function between both exposures with a 3-point Gaussian peak estimator (Raffel et al. (2018)). The initial window size for the multi-grid evaluation is $128 \times 128 \text{px}^2$ and the window size used for the iterative PIV evaluation is $32 \times 32 \text{px}^2$ with 75% overlap. The windows of the iterative PIV evaluation are weighted by a Gaussian window with $\sigma = 0.4$. Between the iterations, outliers in the vector field are detected using a normalized median test and are replaced by interpolated values. A total of three multi-grid steps and five steps of the iterative evaluation are performed.

3 Results

In this section, we highlight the performance of URAFT-PIV based on two learning tasks representing different image and flow conditions. During the first task, we trained the network on a synthetic particle image dataset consisting of five categories which was introduced in Cai et al. (2019b). While this data is interesting and serves as a benchmark, we note that experimental images barely achieve this idealised quality in realistic PIV experiments. As a result, inference runs of networks trained on this dataset hardly predict correct displacements for real-world measurements, likely due to the strong mismatch between training and test time distribution. To study the performance using images like those obtained in many real-world applications, we trained our network of the second learning task on a more realistic dataset introduced in Lagemann et al. (2021) and compare inference results of URAFT-PIV on synthetic and experimental PIV images to neural PIV processing competitors and our cross-correlation based algorithm PascalPIV.

3.1 Learning task I: Idealised PIV database

The networks were trained on synthetic particle image datasets consisting of five categories representing well-known and realistic flow cases: (1) Direct numerical simulations (DNS) of isotropic turbulence; (2)

flows along a backward facing step; (3) two-dimensional flows past a cylinder; (4) DNS of a turbulent channel flow; and (5) simulations of a sea surface flow driven by a Surface-Quasi-Geostrophic (SQG) model. This data is from a public database and is used to benchmark neural PIV processing methods. In total, the resource contains 15,050 particle image pairs with corresponding ground truth flow fields and is divided into 12,000 training and 3,050 test images across all flow categories. Further characteristics include a very high particle density, a maximum particle displacement of ± 10 px, and particle peak intensities ranging from 200 to 255 counts within an 8-bit grayscale resembling images in perfect experimental conditions. Details can be found in Cai et al. (2019b).

Table 1 illustrates inference results on the test dataset of learning task I for various neural network and cross-correlation based PIV processing methods. In the context of unsupervised learning, URAFT-PIV can outperform its LiteFlowNet based competitor in all test cases and achieves a slightly higher error compared to the supervised PIV-LiteFlowNet-en proving the effectiveness of this learning paradigm. However, it cannot match the performance of its supervised counterpart RAFT32-PIV, most likely due the simple but yet effective supervised loss objective which is based on the l_1 -norm between ground truth and optical flow estimation. In contrast, the unsupervised loss of URAFT-PIV is entirely based on geometric penalty terms including a photometric, forward/backward consistency and smoothness oriented loss. We note that networks solely trained on photometric differences tend to predict highly inconsistent displacement fields which are mainly characterized by local extrema in the neighborhood of image patterns while regions of less visual image texture do not contribute to the optical flow estimation. Considering the fact that optical flow estimators aim at matching extracted image features between subsequent images rather than learning the most probable, physical displacement which in fact is not possible, a single photometric loss will not necessarily converge to the physically correct minimum. In this light, PIV images pose an even greater challenge on unsupervised optical flow networks since they contain many, but tiny and almost identical image features - the particles - and hence, provide similar image patterns within the local neighborhood impeding the prediction of the physical correct displacement.

From a high-level perspective, smoothness oriented losses target this ambiguity since they penalize the prediction of strong gradients and encourage colinearity of neighboring flows to achieve a more effective regularization. However, this also means that displacement fields characterized by physically correct, strong gradients are usually underestimated resulting in higher error values. To illustrate this drawback, Fig. 2 compares the ground-truth and predictions of supervised and unsupervised networks for two flow fields side-by-side. Especially in gradient dominated flows, URAFT-PIV under-/overestimates the local displacement compared to its supervised counterpart since the unsupervised loss objective encourages the network to regularize the optical flow. As a result, the endpoint error of URAFT-PIV is one order of magnitude higher

Methods	Back-step	Cylinder	JHTDB Channel	DNS turbulence	SQG
WIDIM [1]	3.4	8.3	8.4	30.4	45.7
HS Optical Flow [1]	4.5	7.0	6.9	52.5	15.6
PIVNetS-noRef [2]	{13.9}	{19.4}	{24.7}	{52.5}	{52.5}
PIV-NetS [2]	{5.9}	{7.2}	{15.5}	{28.2}	{29.4}
PIV-LiteFlowNet [1]	{5.6}	{8.3}	{10.4}	{19.6}	{20.0}
PIV-LiteFlowNet-en [1]	{3.3}	{4.9}	{7.5}	{12.2}	{12.6}
RAFT32-PIV [4]	{0.4}	{1.8}	{1.1}	{2.8}	{2.1}
UnLiteFlowNet-PIV [3]	10.1	7.8	9.6	13.5	19.7
URAFT-PIV (present)	6.5	6.6	8.1	12.5	13.2

Table 1: Averaged Endpoint Error (AEE) for all test cases of the synthetic PIV database introduced in Cai et al. (2019b) of learning task I. URAFT-PIV outperforms its unsupervised competitor in all test cases and achieves a similar performance compared to a supervised LiteFlowNet based network (PIV-LiteFlowNet-en). Its supervised counterpart RAFT32-PIV still achieves the lowest endpoint error by quite a margin. The error unit is set to pixel per 100 pixels for easier comparison. Values in brackets correspond to supervised networks. References: [1] Cai et al. (2019a), [2] Cai et al. (2019b); [3] Zhang and Piggott (2020); [4] Lagemann et al. (2021)

compared to RAFT32-PIV, but still achieves a higher accuracy than UnLiteFlow-PIV which additionally shows strong prediction noise. In contrast, if the underlying displacement field is more smooth, the endpoint error decreases significantly. These findings are in line with literature (Jonschkowski et al. (2020)) and suggest that the loss functions currently used might be a limiting factor for the performance of unsupervised optical flow estimation. Current state-of-the-art unsupervised loss objectives are useful, but by far not as effective as supervising the network based on a ground-truth, i.e., more sophisticated loss objective can significantly boost the accuracy of unsupervised optical flow estimation.

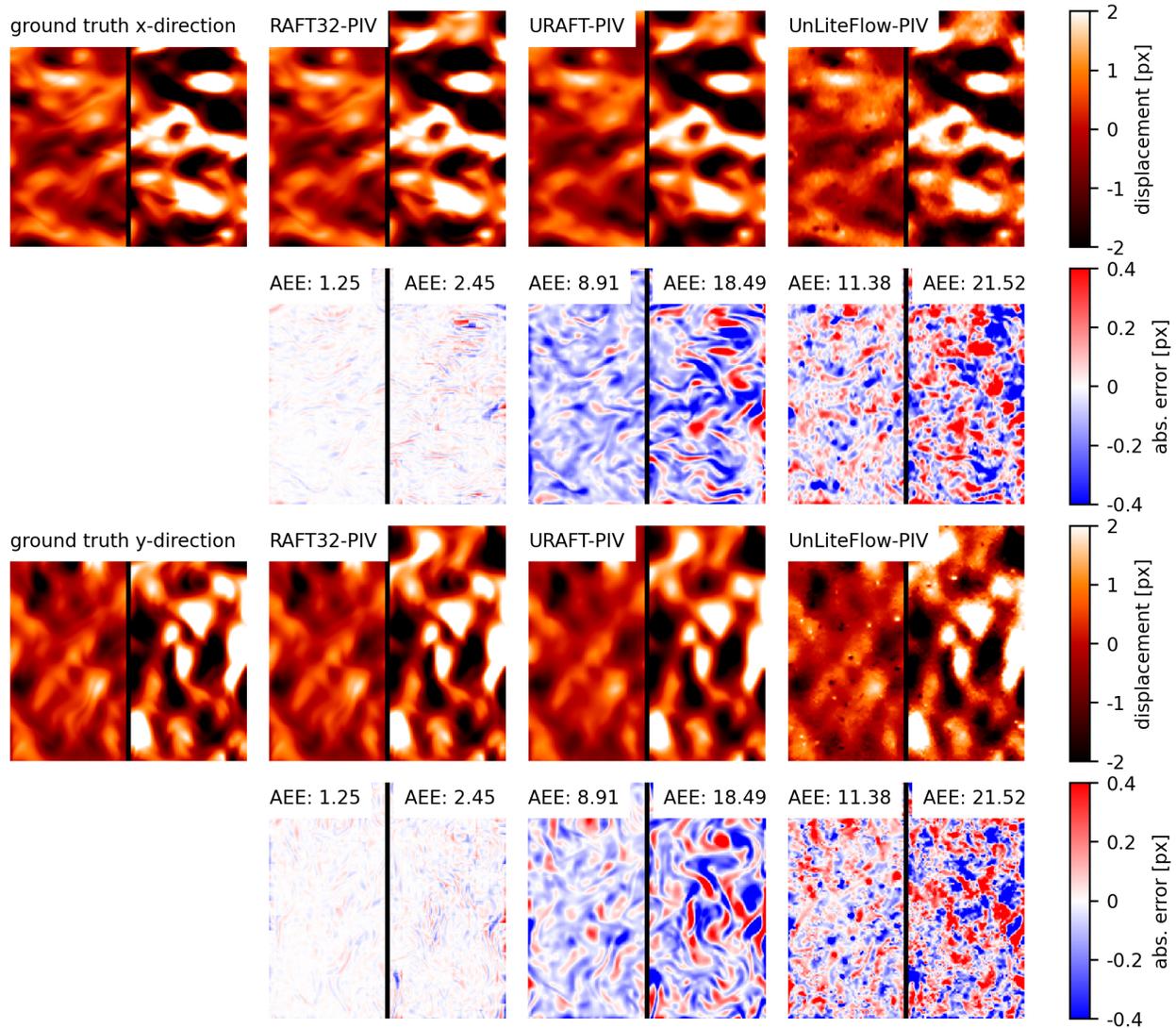


Figure 2: Optical flow prediction of different network architectures and absolute error between ground truth flow and network predictions. Each image depicts two individual flow fields characterized by medium and strong gradients. The first two rows illustrate the displacement and error distribution of the horizontal direction while the last two rows show estimates for the vertical axis. Especially in flow fields dominated by strong gradients (right half), URAFT-PIV under-/overestimates the local displacement compared to its supervised counterpart since the unsupervised loss objective encourages the network to regularize the optical flow. However, URAFT-PIV still achieves a higher accuracies compared to UnLiteFlowNet-PIV and a lower prediction noise.

3.2 Learning task II: Realistic synthetic and experimental PIV images

While the data above are interesting and useful, we note that it is almost impossible to obtain images of this quality in practical applications since PIV setups are very sensitive to external and internal sources of noise, e.g., reflections on side walls or surfaces, light refraction at glass surfaces, slight misalignments in the setup, or density gradients as present in supersonic flows. To study the performance using images like those obtained in many real-world applications, an additional database with an increased particle displacement up to ± 24 px, a reduced particle density and signal-to-noise-ratio (SNR), an increased variance of the particle diameter, and camera noise was used to train the networks of learning task II. Details of this dataset can be found in Lagemann et al. (2021).

First, we study the performance of URAFT-PIV in evaluating synthetic images based on a DNS of a laminar and a fully turbulent boundary layer. Results are depicted in Fig. 3 illustrating the displacement prediction of URAFT-PIV and its supervised counterpart alongside comparisons of displacement profiles at different positions. In case of a laminar boundary layer, barely any differences become visible between RAFT32-PIV and URAFT-PIV. Compared to the ground-truth, however, local flow feature appear to be smoothed and less sharp. We assume that this is directly related to the fact that the particle images of learning task II comprise significantly less particles compared to learning task I and consequently, less information of the underlying flow can be evaluated while more regions with low texture occur. The displacement profile confirms these findings highlighting that the RAFT inspired approaches and our cross-correlation algorithm closely match the ground-truth. In contrast, UnLiteFlowNet-PIV shows some slight deviations and a more noisy distribution similar to previous findings.

Test runs on a turbulent boundary layer confirm these results. Overall, one notices that RAFT32-PIV follows the ground-truth most accurately only deviating slightly in regions of local extrema. The unsupervised URAFT-PIV similarly matches the overall trend of the ground-truth, but cannot reach the accuracy of its supervised counterpart. However, it still matches the performance of our cross-correlation based method and proves its effectiveness in realistic particle and flow conditions.

In our final test case, we apply our URAFT-PIV model trained on the dataset of Lagemann et al. (2021) to real-world experimental PIV data. This test case consists of experimental PIV measurements dealing with a turbulent wavy channel flow as shown in Rubbert et al. (2019). Together with flow field predictions of unsupervised PIV networks and our RAFT models, we analyze the images using our in-house code. Generally, Fig. 4 evidences that both RAFT-PIV approaches - supervised and unsupervised - perform likewise state-of-the-art cross-correlation based PIV methods and hence, can serve as direct substitute. However, we noted that the prediction results based on URAFT-PIV show some spurious estimations in the area of high displacements (≈ 12 px). This is potentially based on the unsupervised loss formulation which might not be suitable for high displacements since similar patterns also occur for UnLiteFlowNet-PIV, but further analysis is required. Moreover, it is noteworthy that URAFT-PIV clearly reduces the prediction noise due to its recurrent nature compared to its LiteFlowNet based competitor. Please note that our RAFT based approaches do not involve any post-processing steps but nevertheless achieve at least an equal noise level compared to gold-standard PIV algorithms. For instance, PascalPIV performs a spatial multigrid cross-correlation scheme in a first step before computing the local displacement field of the final interrogation window in 25 iteration steps. Prior to every iteration step, several validation criteria are applied to detect outliers and spurious values are replaced using a Lanczos interpolation scheme. Thus, a smooth displacement field is finally achieved. In contrast, the RAFT model resembles a single-shot approach which solely operates on a fixed input window without taking further neighboring information into account. Considering this key difference, the low noise level of RAFT-PIV is quite astonishing and further proves for the first time that unsupervised learning is a viable alternative when processing arbitrary real-world PIV images. Especially the possibility of training respectively fine-tuning existing networks on real-world experimental data states a key advantage for unsupervised learning tasks. However, detailed studies on new loss objectives and their corresponding effect on the accuracy are necessary.

4 Conclusions

We studied URAFT-PIV, an unsupervised deep neural network architecture for optical flow estimation in PIV applications. URAFT-PIV achieves a new state-of-the-art accuracy on a public PIV database for unsupervised learning and performs likewise supervised LiteFlowNet based PIV networks. URAFT-PIV also performs well under more challenging flow field and image conditions such as low particle density and

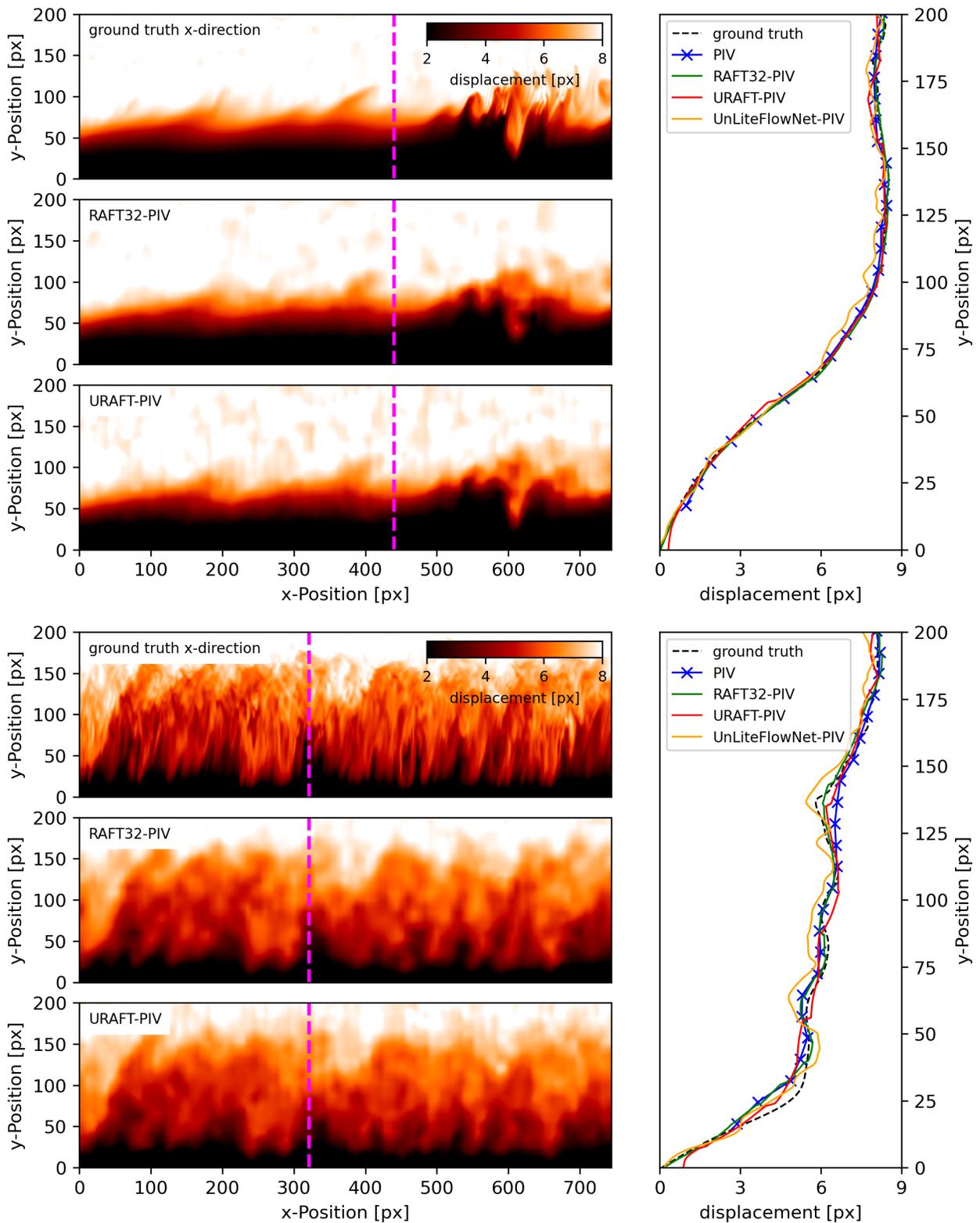


Figure 3: Comparison of displacement fields and profiles for a laminar and turbulent boundary layer. In case of a laminar boundary layer, RAFT based networks and PascalPIV match the ground-truth well. UnLiteFlowNet-PIV follows the overall trend, but predicts noisy results. The turbulent case reveals some smoothing behaviour of our RAFT approaches, but still match the ground-truth very accurately as does the traditional PIV algorithm. Similar to the laminar flow field, UnLiteFlowNet-PIV can roughly predict the ground-truth distribution but shows significant noise.

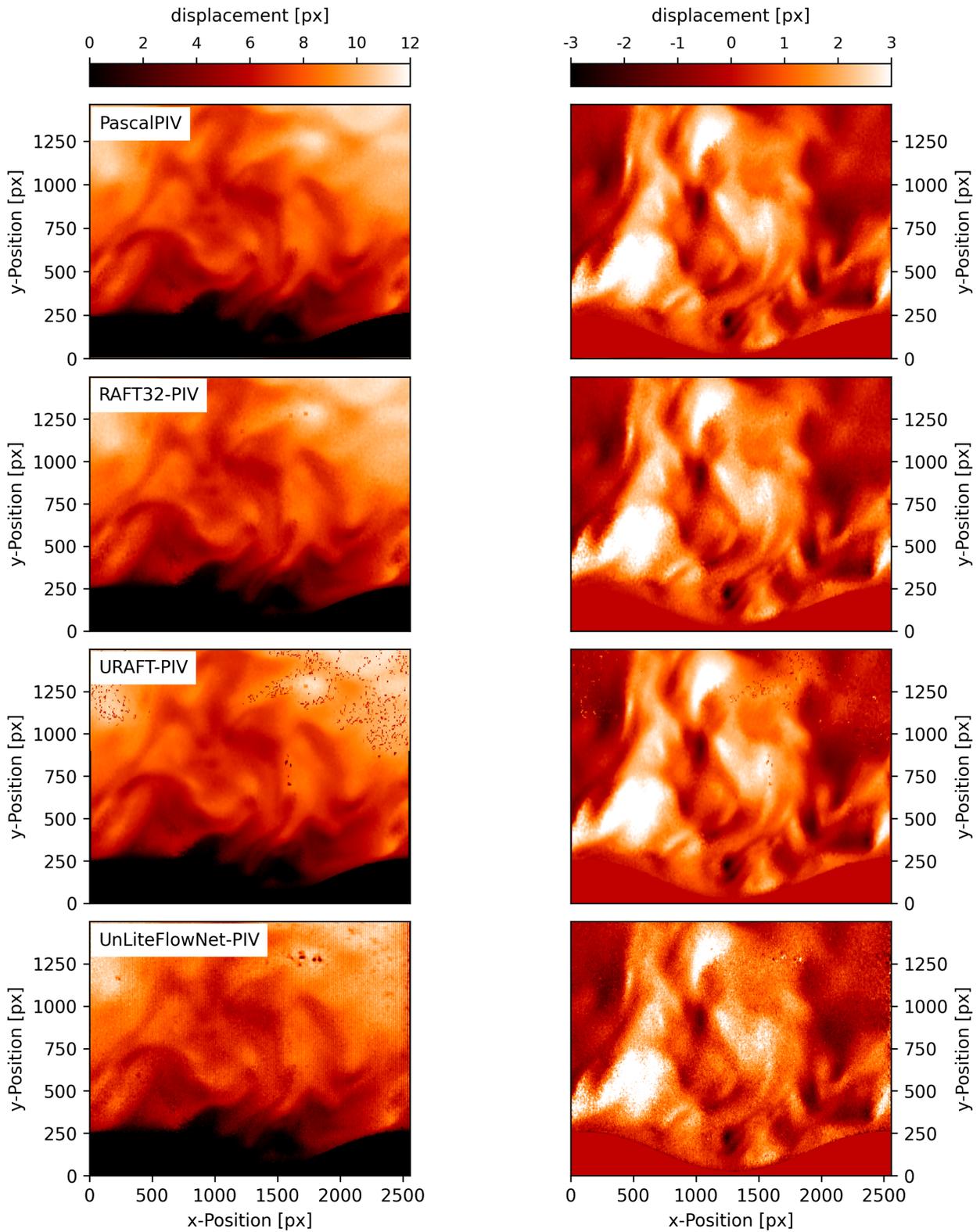


Figure 4: Visual comparison of RAFT-PIV models with state-of-the-art PIV algorithms as well as existing PIV networks. The left column represents the results of all available methods w.r.t. the horizontal optical flow component and the right column depicts the predictions of the displacement in the vertical direction. RAFT32-PIV and URAFT-PIV perform likewise with our high-performance code PascalPIV while significantly increasing the spatial resolution of the displacement field. Further note that these neural methods match the noise level of PascalPIV without interpolating spurious displacement vectors using neighboring data points and hence, resemble single-shot methods.

changing light conditions that are important for many real-world applications. Our tests show that URAFT-PIV accurately predicts displacements while significantly reducing the noise level, most likely due to its recurrent nature. We also noticed that our unsupervised model under-/overestimates the local displacement in regions dominated by strong gradients since the unsupervised loss objective encourages the network to regularize the optical flow. These findings suggest that the loss functions currently used might be a limiting factor for the performance of unsupervised optical flow estimation. Current state-of-the-art unsupervised loss objectives are useful, but by far not as effective as supervising the network based on a ground-truth meaning that more sophisticated loss objective can significantly boost the accuracy of unsupervised optical flow estimation. Applying URAFT-PIV in an out-of-the-box fashion to experimental PIV data demonstrates its generalization capabilities and its ability to significantly improve the spatial resolution while otherwise matching state-of-the-art PIV algorithms. Future work will incorporate further studies targeting the development of more suitable penalty terms.

Acknowledgements

The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time on the GCS Supercomputers HAWK at Höchstleistungsrechenzentrum Stuttgart (www.hlr.de) and Juwels at the Forschungszentrum Jülich (www.fz-juelich.de).

References

- Astarita T and Cardone G (2005) Analysis of interpolation schemes for image deformation methods in PIV. *Experiments in fluids* 38:233–243
- Cai S, Liang J, Gao Q, Xu C, and Wei R (2019a) Particle image velocimetry based on a deep learning motion estimator. *IEEE Transactions on Instrumentation and Measurement* 69:3538–3554
- Cai S, Zhou S, Xu C, and Gao Q (2019b) Dense motion estimation of particle images via a convolutional neural network. *Experiments in Fluids* 60:73
- Dosovitskiy A, Fischer P, Ilg E, Hausser P, Hazirbas C, Golkov V, Van Der Smagt P, Cremers D, and Brox T (2015) FlowNet: Learning optical flow with convolutional networks. in *Proceedings of the IEEE international conference on computer vision*. pages 2758–2766
- Hui TW, Tang X, and Change Loy C (2018) LiteflowNet: A lightweight convolutional neural network for optical flow estimation. in *Proceedings of the IEEE conference on computer vision and pattern recognition*. pages 8981–8989
- Jonschkowski R, Stone A, Barron JT, Gordon A, Konolige K, and Angelova A (2020) What matters in unsupervised optical flow. *arXiv preprint arXiv:2006.04902* 1:3
- Kingma DP and Ba J (2014) Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*
- Lagemann C, Lagemann K, Mukherjee S, and Schröder W (2021) Deep recurrent optical flow learning for particle image velocimetry data. *Nature Machine Intelligence*
- Lee Y, Yang H, and Yin Z (2017) PIV-DCNN: cascaded deep convolutional neural networks for particle image velocimetry. *Experiments in Fluids* 58:171
- Meister S, Hur J, and Roth S (2017) Unflow: Unsupervised learning of optical flow with a bidirectional census loss. *arXiv preprint arXiv:1711.07837*
- Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, Lin Z, Desmaison A, Antiga L, and Lerer A (2017) Automatic differentiation in pytorch
- Rabault J, Kolaas J, and Jensen A (2017) Performing particle image velocimetry using artificial neural networks: a proof-of-concept. *Measurement Science and Technology* 28:125301
- Raffel M, Willert CE, Scarano F, Kähler CJ, Wereley ST, and Kompenhans J (2018) *Particle image velocimetry: a practical guide*. Springer
- Rubbert A, Albers M, and Schröder W (2019) Streamline segment statistics propagation in inhomogeneous turbulence. *Physical Review Fluids* 4:034605
- Schrijer F and Scarano F (2008) Effect of predictor–corrector filtering on the stability and spatial resolution of iterative PIV interrogation. *Experiments in fluids* 45:927–941
- Teed Z and Deng J (2020) RAFT: Recurrent All-Pairs Field Transforms for Optical Flow. in *European Conference on Computer Vision*. pages 402–419. Springer
- Zhang M and Piggott MD (2020) Unsupervised Learning of Particle Image Velocimetry. *arXiv preprint arXiv:2007.14487*