
Particle streak velocimetry using Ensemble Convolutional Neural Networks

Alexander V. Grayver · Jerome Noir

Received: - / Accepted: -

Abstract Insert your abstract here.

Keywords Streak analysis · Neural Networks · Turbulent Flows

1 Introduction

Particle Image Velocimetry (PIV) is arguably the most widely used technique to quantitatively study experimental flows [13]. A common work-flow would consist of seeding flow with luminescent particles and taking pairs of pictures with a known short time separation to capture an instantaneous flow state. By splitting a pair of images into (possibly overlapping) windows and cross-correlating between them allows one to infer the direction and magnitude of the flow. With modern computers and digital cameras, PIV has experienced wide adoption in the scientific community. For cross-correlation to work properly, one has to ensure that the exposure time is short enough such that particles do not move more than a few pixels and the two images have clearly identifiable correlations. Violating this condition, for instance because of insufficient laser intensity or too fast flow, results in so called streaks - traces of particles.

Streaks in the PIV images are commonly considered an experimental failure since they render cross-correlation techniques less efficient or even inapplicable. When flow velocity imposes constraints for which the camera and light source

A. V. Grayver
Institute of Geophysics, ETH Zurich
Soneggstrasse 5
8092 Zurich, Switzerland
Tel.: +41-44-6333154
E-mail: agrayver@erdw.ethz.ch

J. Noir
Institute of Geophysics, ETH Zurich
Soneggstrasse 5
8092 Zurich, Switzerland
Tel.: +41-44-6337593
E-mail: jerome.noir@erdw.ethz.ch

at hand cannot capture instantaneous particle positions, one either has to use a higher speed camera or/and a stronger light source. Both of these solutions quickly hit financial and safety constraints, which often cannot be overcome in academia. Therefore, there exists a need for processing technique. Although not suitable for conventional cross-correlation methods, streak images do contain information about the flow. Whereas an image with streaks does not contain information about direction of the underlying flow, information on the mean velocity magnitude and azimuth can be inferred from streaks themselves. To the best of authors' knowledge, there are no reliable algorithm to recover quantitative information from streaks images , which motivated us to design a method for extracting information about the flow from streaks.

To this end, we applied an ensemble of Convolutional Neural Networks (CNN) trained on streak images to draw a statistical prediction of the displacement magnitude and corresponding azimuth. The potential of CNNs has long been recognized for applications related to the face and handwriting recognition [8, 14], although their power has been fully discovered only recently when deeper (that is, with more layers) networks became feasible to train within reasonable amount of time, mostly due to the emergence of Graphic Processor Units (GPUs) [7, 6]. An interested reader is referred to a recent overview of the deep learning with important development milestones listed [9].

Before we proceed to describing methodology and results, it is worth to mention that a crucial ingredient to a success of CNNs is a suitably large training set, which a CNN is supposed to learn from without facing an over-fitting. From this perspective, streak analysis appears almost an ideal problem. Similar to PIV, we make a reasonable assumption that within a sufficiently small interrogation window velocity exhibits simple functional form (for instance, constant or linear function). This allows us to quickly generate an unlimited number of training images. Adding variability in number of streaks per window, their thickness and intensity will mimic a realistic experimental scenario sufficiently well. Therefore, we can exhaustively sample the parameter space and feed the network with as many sample as needed to attain an acceptable accuracy. The specific numbers and parameters will be discussed in the corresponding sections below.

2 Methods

2.1 Problem setup

We aim at inferring the displacement and azimuth in an streak image of $N \times N$ pixels, typically as illustrated in figure 1.

Similarly to how PIV is applied to a pair of images, we split the image into ($n \times n$, with $n=48$ in this study) sub-windows with an overlap (50% in our case). Assuming the velocity be uniform enough over the $(n \times n)$, we aim at determining the mean length, the displacement over the exposure time, and azimuth over each sub-window.

A closer look at the image shows that the number of streaks, their intensity and their width, vary within each sub-window. In addition, in many cases, the streaks leave the $(n \times n)$ interrogation area.

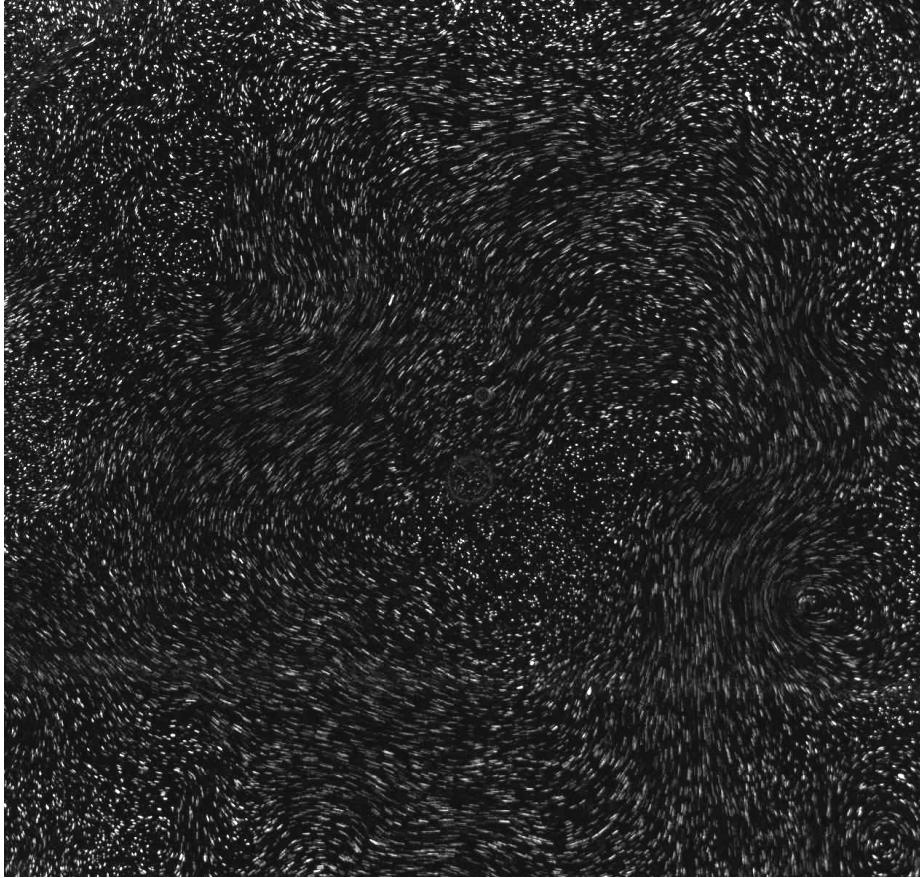


Fig. 1 Streaks image (1024×1024 pixels) from a longitudinal libration experiment in a cylindrical cavity. See section 3.3 for details

In order to make our CNN resilient to this variability, we generate train and validation images of (48×48 pixels) with a randomly chosen number of streaks varying between 2 and 10, a random color intensity between 90 and 255 and a random thickness between 2 and 4 pixels. The center of each streak is chosen randomly over the entire sub-window, such that streaks are allowed to go outside the image.

To generate train and validation images, we sampled Δ and ϕ from the uniform distributions

$$\Delta \sim U(1, \Delta_{\max}) \quad (1)$$

and

$$\phi \sim U(-\pi/2 + \delta, \pi/2), \quad (2)$$

respectively. Here, $\phi = 0$ corresponds to the positive x -axis and increases clockwise. Note that limits for ϕ cover only half of the circle minus $\delta = 5^\circ$ on one end to avoid ambiguity with respect to the direction. Once streaks are generated, the within streak intensity is perturbed (**explain briefly how you do it**) and the image is

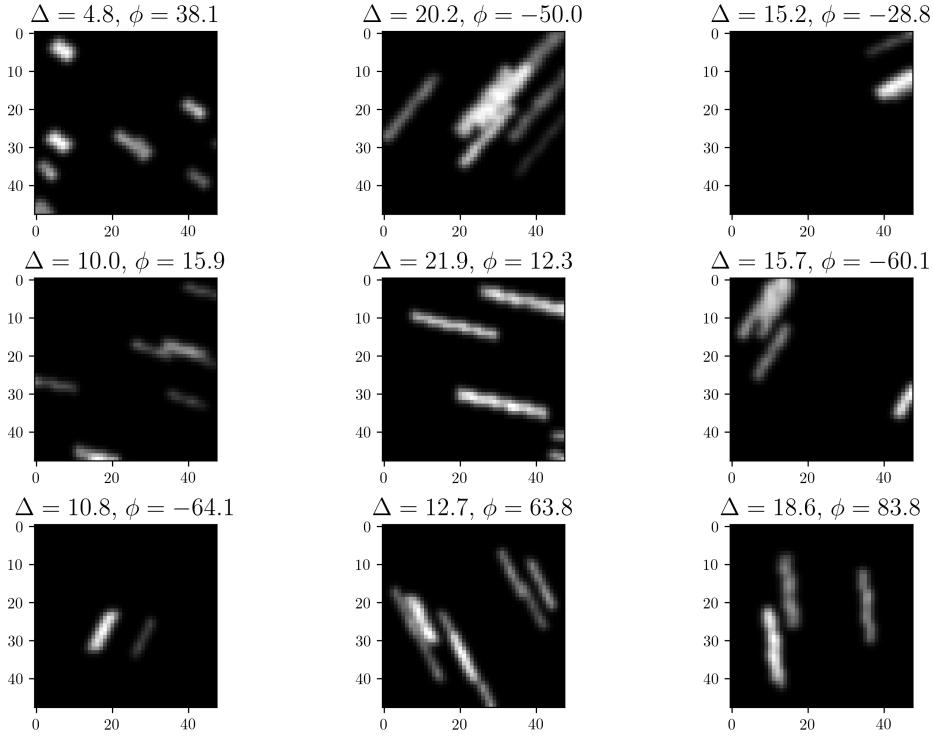


Fig. 2 Subset of generated streak images used for training and validation with the corresponding displacements (in pixels) and azimuth angles (in degrees) given in the titles.

blurred with a Gaussian kernel of 3×3 pixels. The maximum displacement, Δ_{\max} , of a particle within an image is assumed to be half of the chosen window size, that is $r_{\max} = 24$ pixels in our case. Each streak is thus characterized by two numbers Δ, ϕ , denoting displacement and azimuth, respectively. Figure 2 shows a selection of streak images generated using the stated procedure. The outlined procedure enables generation of millions of images in just a few minutes on a regular PC.

2.2 Network architecture and implementation

We aim at building a regression CNN that, given a single window image, outputs two real numbers, corresponding to the displacement and angle. To this end, we have created a network with the architecture shown in Figure 3. As can be seen, four convolutional units form the core of the network, each consisting of a convolutional layer, activation layer (in this case, rectified linear unit), average pooling layer used to downsample the input followed by a batch normalization layer. An increasing depth of the convolutional layers is a common practice which is aimed at giving a network freedom to learn more complex features from the input.

To prevent overfitting and improve training speed, we used batch normalization at the end of each convolution unit [5] and dropout layer with the 30% drop fraction [15]. Additionally, following common practices, the outputs are standar-

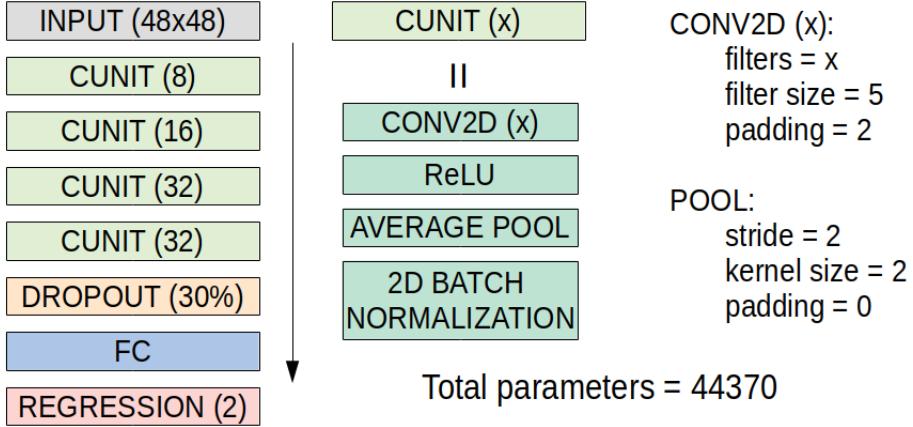


Fig. 3 Architecture of the network. Left: sequence of layers and corresponding parameters in the brackets. Each CUNIT layer consists of four actual layers listed in the middle column. Relevant layer parameters are given on the right. Total number of learnable parameters in the network is 44370.

tized and centered. The input grayscale images have a single eight bits channel. We normalized them by 255 to stay in the [0, 1] range.

Whereas there may exist more efficient architectures, we found that for our problem increasing depth of the network has not improved overall performance significantly, nor did increasing and/or decreasing filter sizes. Additionally, we found that, in contrast to majority of reported CNN architectures, using the average pooling instead of the max pooling results in higher accuracy.

To increase accuracy and robustness of the prediction, we built an ensemble consisting of ten networks [3]. All networks have the same architecture and parameters, but were trained using different randomly chosen initial weights [1, 2] and random shuffling to form a unique sequence of batches supplied to an optimizer. The outputs of all ensemble members are averaged to get a the final prediction, also enabling us to calculate a variance, thereby providing a proxy for the uncertainty of a prediction.

Finally, the network was implemented using the PyTorch library [11]. The complete implementation can be found using the link below.

3 Results

3.1 Network training

We generated one million images (e.g., Figure 2) of which 75% and 25% were used for training and validation, respectively. Each network in the ensemble has been trained for 100 epochs. The ADAM optimizer with an initial learning rate value of 1e-3 was used. The learning rate was halved if no sufficient decrease in the validation loss was observed for the past 10 epochs. To make the training more resistant to potential outliers (e.g. imagine a situation when all streaks leave the window), we chose to minimize the Huber loss [4] rather than conventional least-squares functional. Huber loss represent a hybrid $L_2 - L_1$ distribution with a high

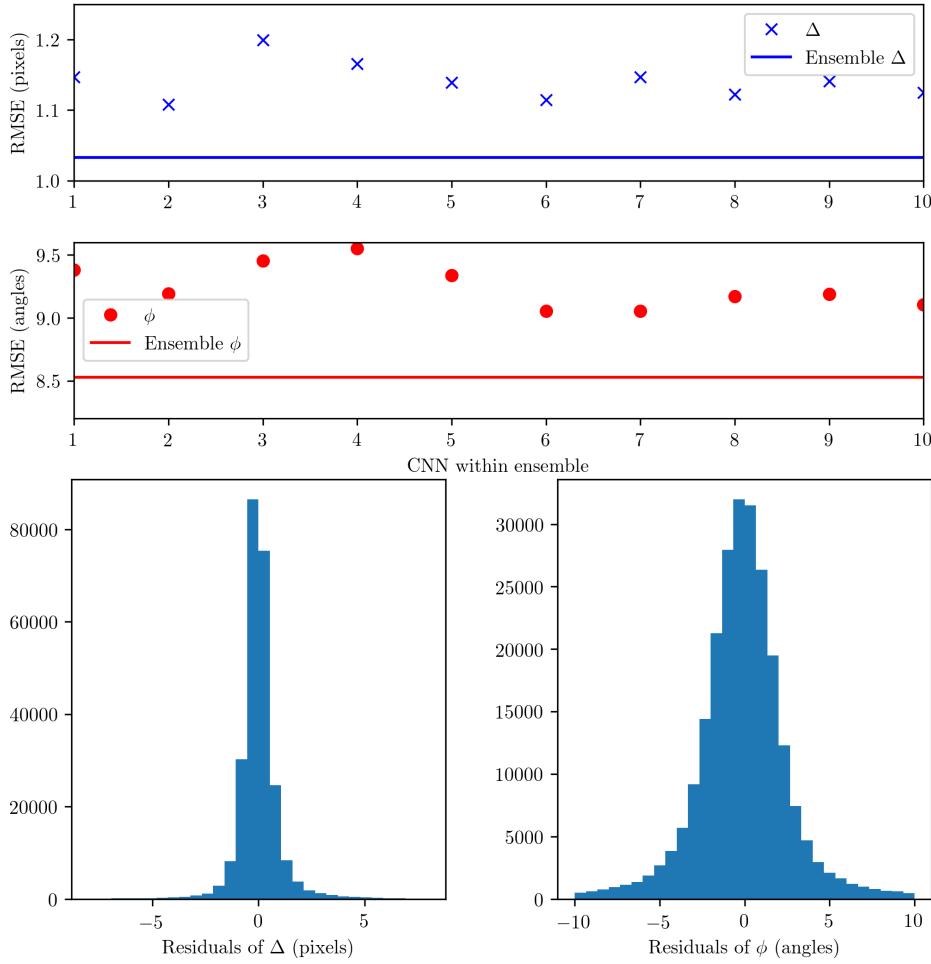


Fig. 4 Top: RMSE values of individual networks (markers) and ensemble (line) for displacement Δ and angle ϕ calculated on the validation set. Bottom: histograms of the residuals for the displacement and angle.

tolerance to the presence of long tails in the original distribution. No pathologies between training and validation loss were observed during the training, specifically the validation loss remains a bit higher than the training loss, indicating that the network learns some generic features of the dataset. Training of each network took 90 minutes on a single CPU-GPU system. This modest time suggests that training networks for window sizes larger than the one adopted here is feasible.

Figure 4 shows RMSE values for all networks separately as well as the RMSE of the ensemble. It is evident that the ensemble performs up to 20% better than a single network. CNN ensemble achieves the accuracy of 1.05 pixels for the displacement and 8.5° for the angle predictions, respectively. The latter may seem like a large error, however one should realize that having a 48×48 pixels window inevitably leads to limitations in recognizing small rotations, especially when streaks are short.

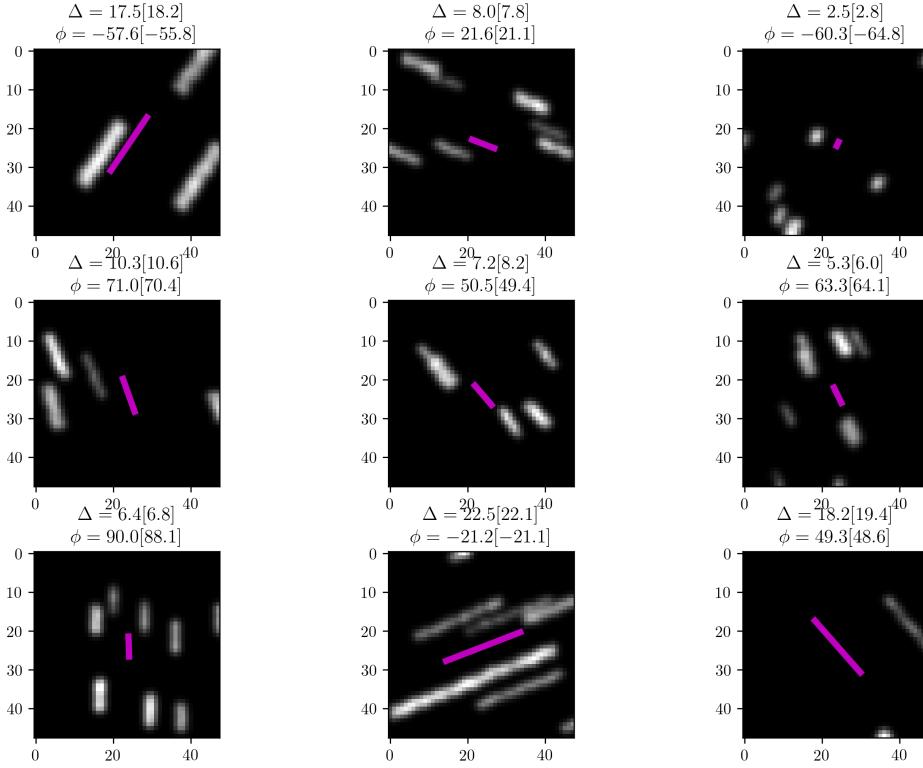


Fig. 5 Images with the predicted displacement and angle depicted as magenta lines. The true values are given in titles and predicted values in brackets.

Figure 5 shows a random selection of validation images with the corresponding CNN ensemble predictions and true values. We see that the ensemble CNN has learned to make accurate prediction even for complex situations when some streaks overlap, join or leave the window. Additionally, Figure 6 shows nine images which produced the worst predictions of the displacement. Clearly, most of these cases are associated with situations where originally longer streaks have all been displaced outside the window or multiple streaks coincidentally joined and formed what appears to be a longer streak. Note that such situations do not represent a failure of the CNN, but rather limitation of the window-based approach adopted in this work. Similar situations are likely to occur in experimental images and unfortunately they leave little chance for any sliding window approach to make an accurate prediction. We will discuss possible ways to mitigate this in the Conclusions section. It is worth to note that despite inaccurate displacement prediction, the angle is predicted accurately in most of these cases.

3.2 Application to a numerical simulation

After confirming the excellent performance of the network on the validation set, we turn to a discrete numerical simulation (3D-DNS) of a complex thermal convection

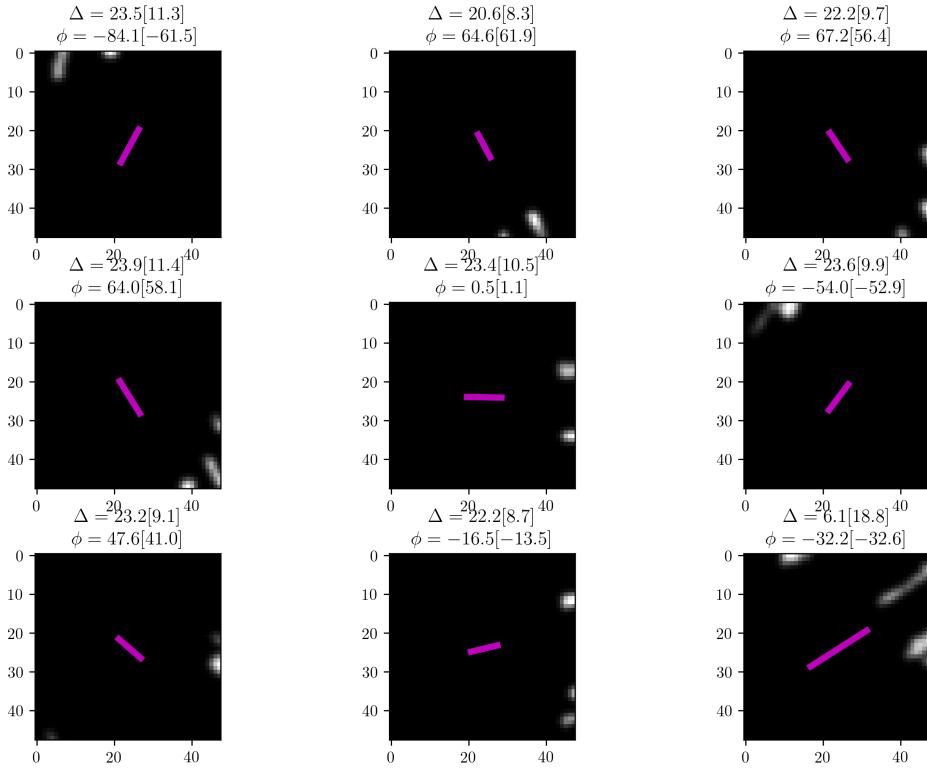


Fig. 6 Same as previous, but showing nine worst prediction in terms of displacement.

turbulent flow presented recently by [12]. In their paper, the authors simulate rapidly rotating thermal convection in a square box, heated from below and cooled from the top with a vertical gravitational field pointing downward and a vertical axis of rotation. To validate our CNN algorithm, we used a simulation obtained for $E = 10^{-7}$, where E is the Ekman number representing the ratio of viscous to rotational effect, and a Raleigh number 90 times super critical. Each velocity field is 192^3 . For the purpose of this study, we extract the horizontal component of the velocity in a horizontal plane a mid-depth. Due to the limited size of the calculation each velocity component has been interpolate on a 1024×1024 pixels grid. Figure 7-left shows the displacement map obtained from the full DNS for an arbitrary exposure time. Since the CNN will act on (48×48) interrogation windows with 50%, we apply a sliding window averaging with (48×48) and 50% overlap (7-right) The window averaged displacement is the best reconstruction we can possibly achieve with a 48×48 pixels window and a 50% overlap used in this study.

To reduce subjectivity, we generated 30 randomly seeded streak images using the original DNS velocity (Figure 7, left) and varying streak density, intensity and thickness as formerly stated. The trained CNN ensemble was applied to all images and the resulting averaged reconstruction of both displacement and angle are shown in Figure 8. We see that the reconstructed images recover the general structure of the flow very well. The only negative effect is that the ensemble

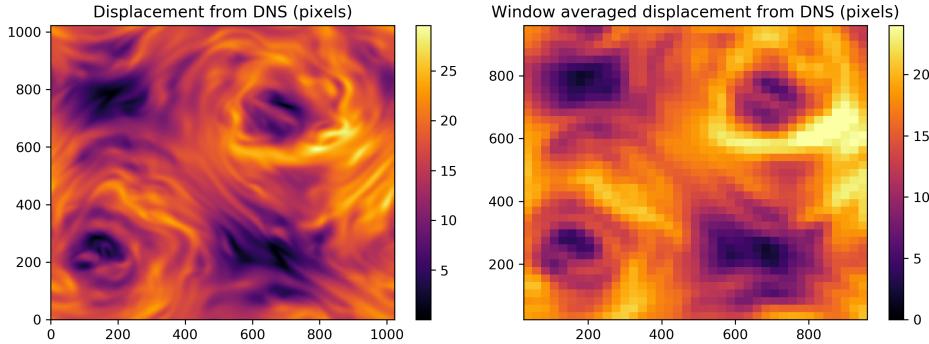


Fig. 7 Original (left) and window averaged (right) displacement maps of the DNS [12] used for validation.

CNN struggles to correctly predict large displacements. When the displacement reaches half of the window size, CNN systematically predicts smaller displacement. The reason for this is not a deficiency of the CNN, but rather the fact that for displacements as large as half the window size it becomes very likely that most of the streaks go outside the window, thus a smaller displacement is a rational prediction. To solve this problem, one can apply a larger window (for instance, 64×64 pixels), although this will reduce the overall resolution and may render constant velocity assumption less accurate. In real experimental settings, it may be advantageous to apply several window sizes and either choose one or combine obtained results, depending on a particular experimental setup.

Finally, Figure 9 shows one of the streak images generated from DNS with the true and predicted streaks for each window position. Note that while the whole velocity field is used to generate the background streak images, the overlying red streaks are reconstructed at the location of the CNN grid point using the averaged displacement and angle maps (figure 7). Generally, ensemble CNN performs well. There is, however, accuracy reduction in regions of large displacements for reasons described above and across areas of vertical shearing, where angle predictions become locally less accurate. Interestingly, these vertical shear regions also lead to abnormal streaks orientation (close to horizontal) when using the window averaged field values (red streak in figure 7).

3.3 Application to the flow driven by longitudinal libration in laboratory experiment

Finally, we apply our algorithm to the laboratory experiment depicted in figure 10. The apparatus consists of a straight cylinder filled with water set in rotation on a turntable at 1Hz. The cavity is 286mm high with a radius of 140mm. The bottom lead is cover with topography made of blocks ($64\text{mm} \times 64\text{mm} \times 8\text{mm}$). A second motor is used to oscillate the container on the turntable, resulting in the so-called longitudinal libration. For the purpose of this experimental validation we have set the libration frequency to 1Hz and the amplitude to 34° .

To obtain the streak images, we use a 1W continuous diode laser to illuminate a horizontal plane of the fluid seeded with fluorescent particles. We record movies of

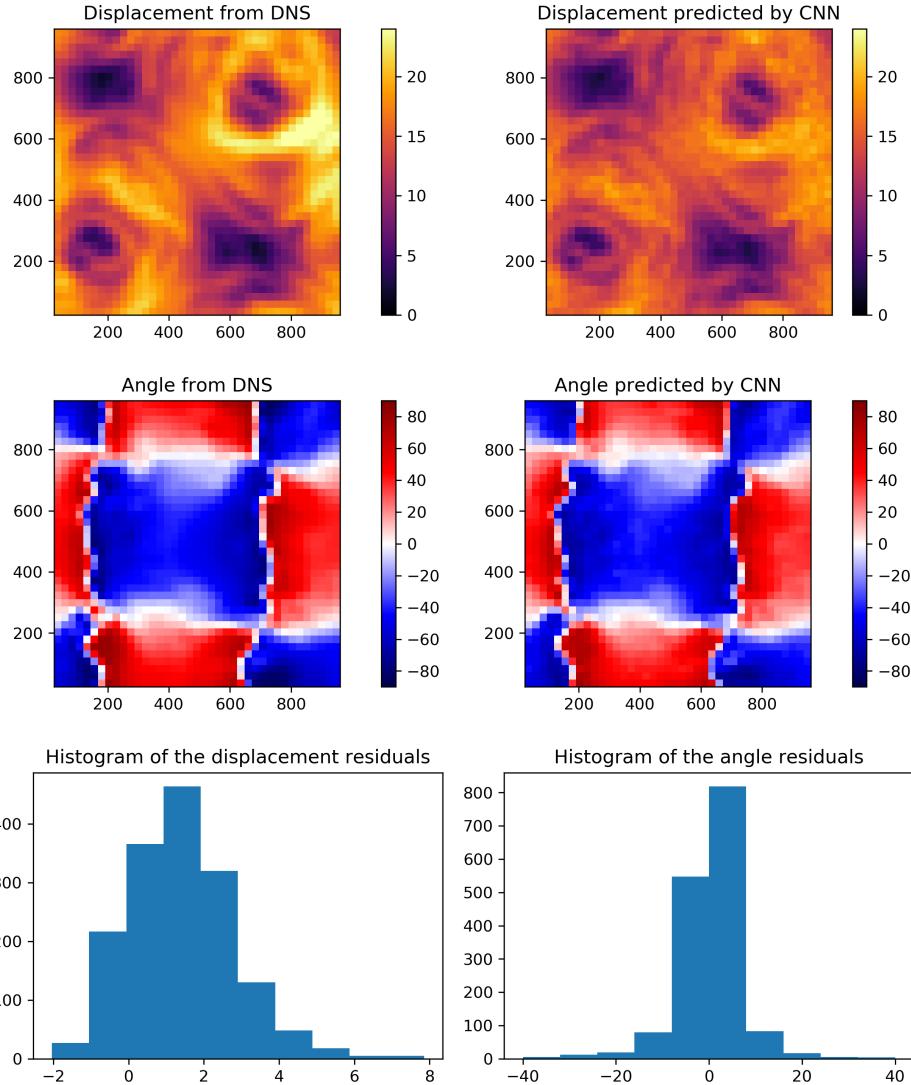


Fig. 8 Original (left) and window averaged (right) displacement maps of the DNS [12] used for validation.

the particles motion at a fixed exposure time of 21ms with a sampling frequency of 30 images per second. We extracted a frame corresponding to a maximum apparent amplitude of the swirling flow (1).

A pre-processing step was necessary for real images to filter out the CCD matrix noise. We achieved this by a simple thresholding of an image at intensity of 50. After that, image was split into overlapping 48×48 pixels windows, resulting in a total of 2400 window images. The ensemble CNN was then applied to predict displacement and angles as well as their standard deviations with the latter serving to be a proxy for the output uncertainties.

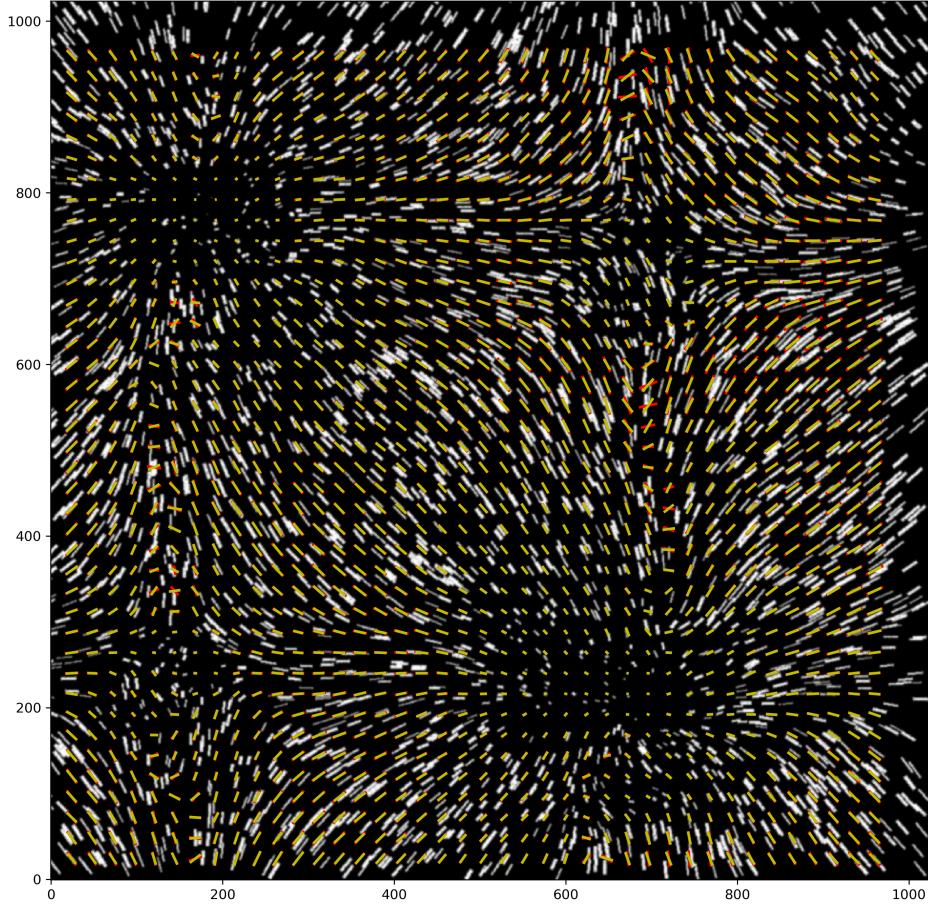


Fig. 9 Streak image generated from the DNS with true and predicted displacement and angle shown as red and yellow lines, respectively.

Figures 11-12 show the predictions in the form of maps and histograms, respectively. Figure 13 shows the actual experimental image we used and window-centered predictions drawn as lines. First of all, we see that experimental flow exhibits rather involved behaviour, with regions where streaks show non-linear behaviour of the flow within a window, which violates our initial assumptions. Nonetheless, the ensemble CNN produces coherent and spatially correlated predictions (Figure 11). It is also interesting to analyse the standard deviation maps (Figure 12). Large absolute errors in displacement appear to be correlated with regions with fast flow. On the other hand, large uncertainties in the angle are around regions with vertical velocity shearing, as observed with DNS synthetic images, specifically where velocity seems to flip the direction.

In terms of performance, ensemble CNN attains a speed of seven full images (each totaling to 2400 windows) per second on a NVIDIA GoForce GTX 1070Ti powered PC. This level of performance allows us to comfortably process long sequences of images taken during experiment. Using a more modern GPU will

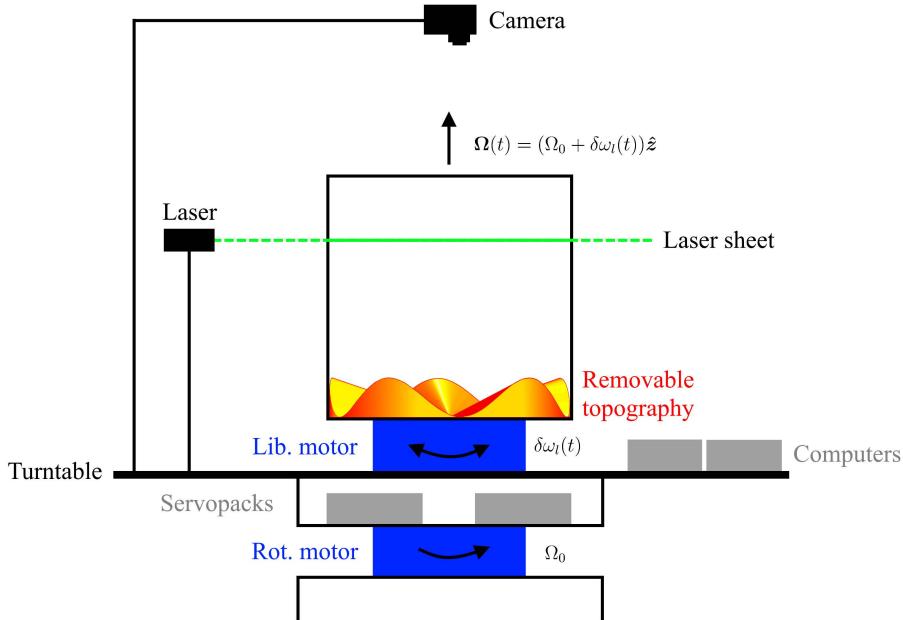


Fig. 10 Experimental setup of the longitudinal libration experiment [10].

likely also enable real-time predictions at rates of a few dozens of images per second.

4 Conclusions

We presented an open-source ensemble CNN aimed at quantitative analysis of complex experimental flows using streak images. The presented approach can be applied in situations when classical PIV is inaccurate or not possible due to experimental setup or hardware limitations. The advantage of the presented approach is the ease of training/validation set generation as well as recovery of the prediction uncertainty through application of the ensemble method. We foresee that the approach may need an experiment-dependent tailoring to achieve the best performance. According to our experience, this only takes a single day on an average PC equipped with a modern GPU and should not represent a practical limitation.

The potential extensions of the presented approach may include the use of multiple window sizes applied to the same image, more complex functional parameterizations of the displacement and angle fields within a window (e.g., polynomial). Finally, a completely different approach based on an image segmentation and object localization and/or tracking can be applied to extract information about individual streaks and their evolution in time. We expect to investigate these directions in details in our future studies.

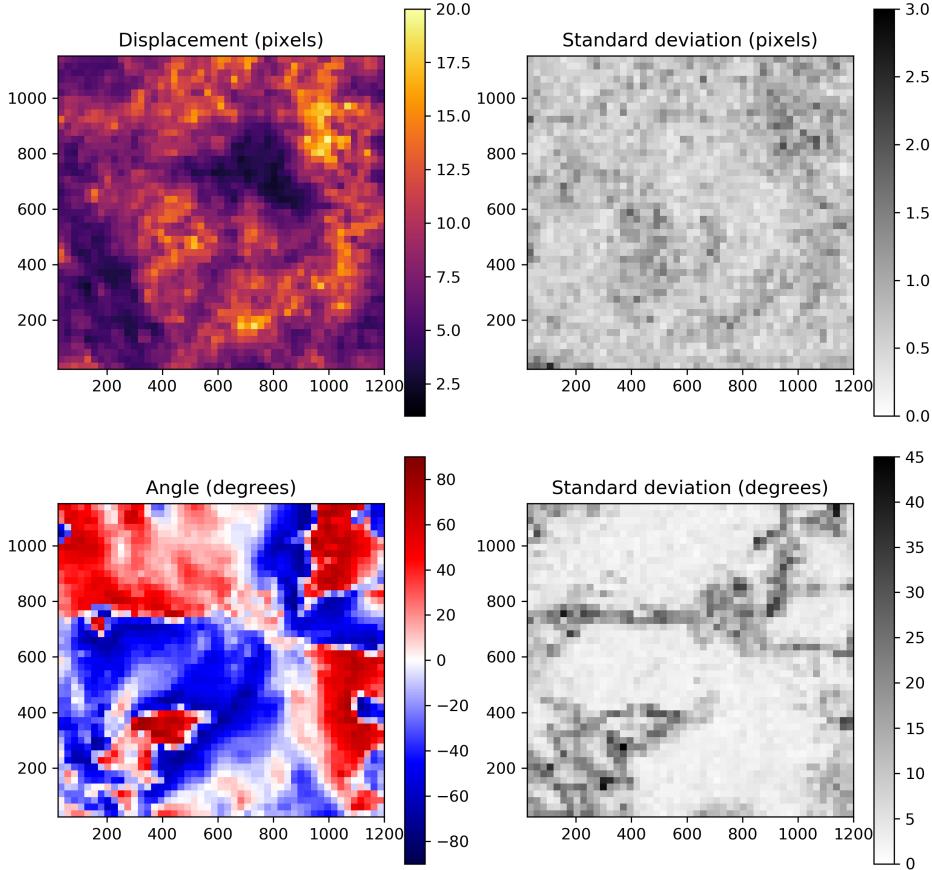


Fig. 11 Predicted displacement and angle (left) along with their standard deviations (right) inferred from the ensemble CNN for an experimental image shown in Figure 13.

Acknowledgements We thank authors of PyTorch, matplotlib and h5py libraries for making them available, Meredith Plumley for sharing her DNS data, Adrian Tasistro-Hart for a number of insightful discussions on CNNs. The Jupyter Python notebooks of all programs used in this study can be found here <https://github.com/agrayver/streakcnn>.

References

1. Glorot, X., Bengio, Y.: Understanding the difficulty of training deep feedforward neural networks. In: Proceedings of the thirteenth international conference on artificial intelligence and statistics, pp. 249–256 (2010)
2. He, K., Zhang, X., Ren, S., Sun, J.: Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: Proceedings of the IEEE international conference on computer vision, pp. 1026–1034 (2015)
3. Hinton, G., Vinyals, O., Dean, J.: Distilling the knowledge in a neural network. arXiv preprint arXiv:1503.02531 (2015)

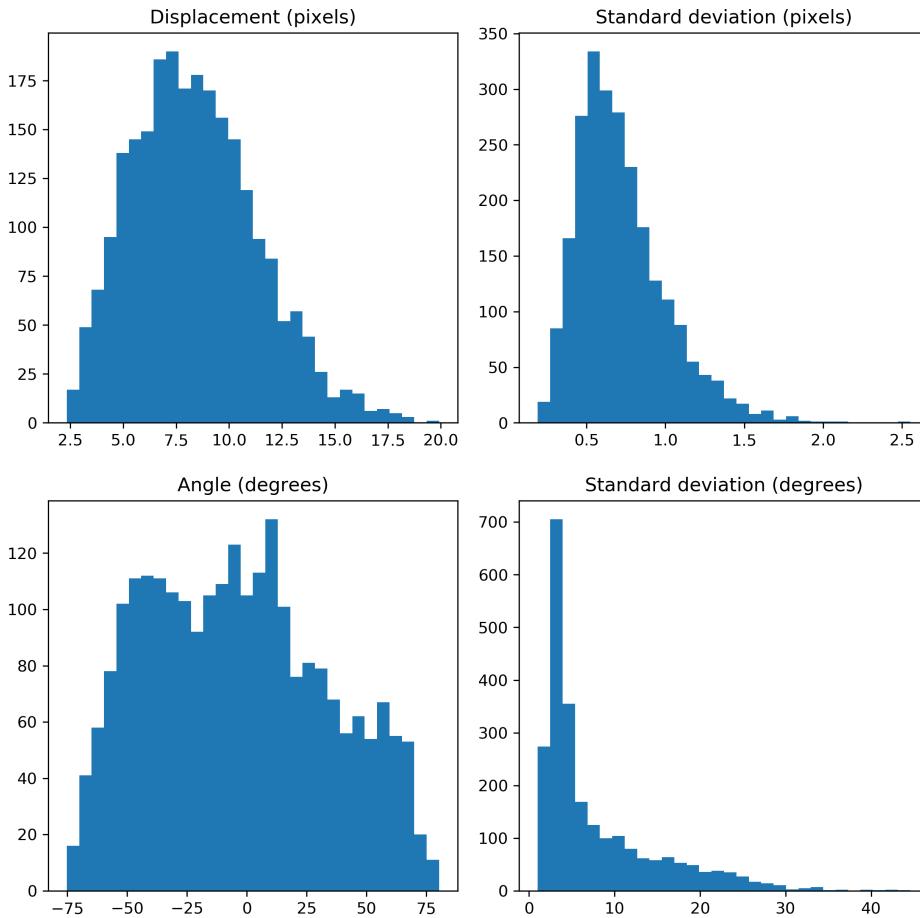


Fig. 12 Same as above, but in a form of histograms.

4. Huber, P.J., et al.: Robust regression: asymptotics, conjectures and monte carlo. *The Annals of Statistics* **1**(5), 799–821 (1973)
5. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
6. Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L.: Large-scale video classification with convolutional neural networks. In: Proceedings of the IEEE conference on Computer Vision and Pattern Recognition, pp. 1725–1732 (2014)
7. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp. 1097–1105 (2012)
8. Lawrence, S., Giles, C.L., Tsoi, A.C., Back, A.D.: Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks* **8**(1), 98–113 (1997)
9. LeCun, Y., Bengio, Y., Hinton, G.: Deep learning. *nature* **521**(7553), 436 (2015)

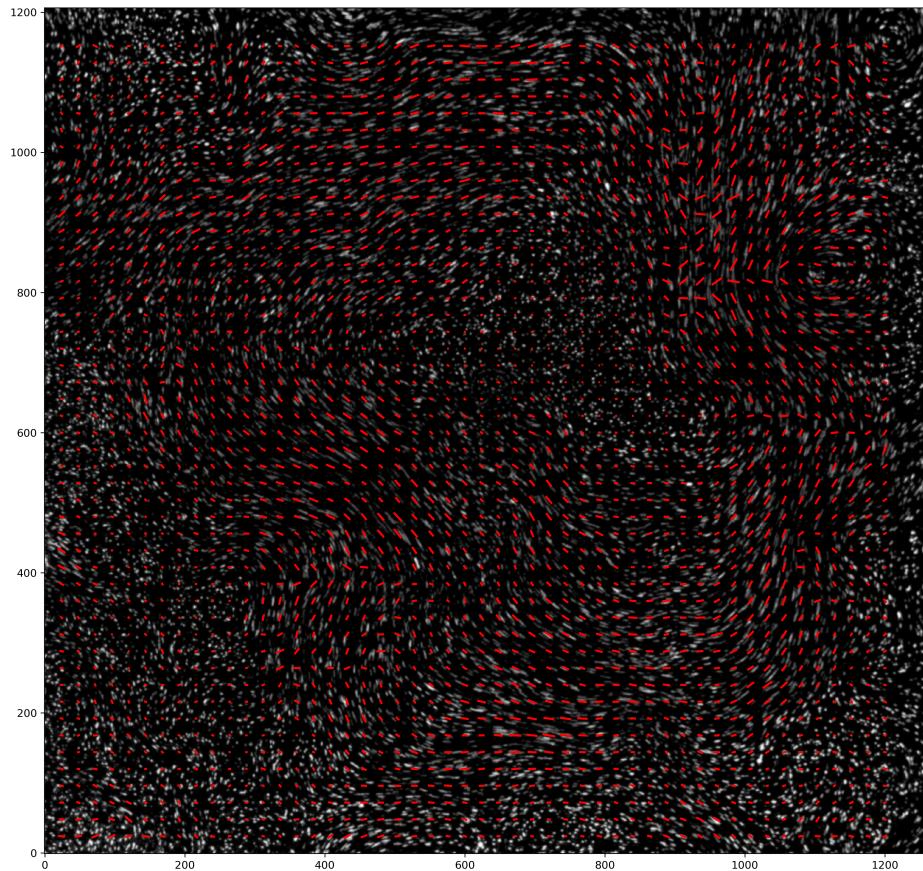


Fig. 13 Streak image taken during an experiment overlain by the predicted displacement and angle shown as red lines.

10. Meister, A.: Effects of boundary topography on the flow forced by libration in longitude in a rotating cylinder, master thesis, ethz, (2018)
11. Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., Lerer, A.: Automatic differentiation in PyTorch. In: NIPS-W (2017)
12. Plumley, M., Julien, K., Marti, P., Stellmach, S.: The effects of Ekman pumping on quasi-geostrophic Rayleigh–Bénard convection. *Journal of Fluid Mechanics* **803**, 51–71 (2016)
13. Raffel, M., Willert, C.E., Scarano, F., Kähler, C.J., Wereley, S.T., Kompenhans, J.: Particle image velocimetry: a practical guide. Springer (2018)
14. Simard, P.Y., Steinkraus, D., Platt, J.C., et al.: Best practices for convolutional neural networks applied to visual document analysis. In: Icdar, vol. 3 (2003)
15. Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.: Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research* **15**(1), 1929–1958 (2014)