ClariPhy: Physics-Informed Image Deblurring with Transformers for Hydrodynamic Instability Analysis

Shai Stamler-Grossman*
SNRC
Israel
shayst@soreq.gov.il

Nadav Schneider*
Ben-Gurion University, IAEC
Israel
nadavsch@post.bgu.ac.il

Gershon Hanoch
Ben-Gurion University, NRCN
Israel
hanochge@post.bgu.ac.il

Gal Oren
Stanford University, Technion
United States
galoren@stanford.edu

Abstract

Recent advances in deep learning have greatly improved image deblurring for natural scenes. However, applying these methods to physical experiments, especially those involving rapid, complex dynamics like hydrodynamic instabilities, remains challenging. Unlike conventional deblurring tasks, these scenarios involve motion blur tied to evolving physical processes, complicating image restoration. We propose ClariPhy, a transformer-based approach utilizing the Restormer model, fine-tuned on a novel deblurring dataset derived from Rayleigh-Taylor instability simulations. This dataset features pairs of sharp and artificially spatial and temporal blurred images, reflecting the real-world conditions of physical experiments. Leveraging the self-attention mechanism of transformers, ClariPhy effectively captures spatiotemporal dependencies crucial for deblurring images of dynamic phenomena. Our results show that ClariPhy outperforms the original SOTA Restormer model, providing enhanced clarity and accuracy in time-sensitive physical experiments.[1]

1 Introduction

Image deblurring, a critical task in the field of computer vision, involves the restoration of sharp images from blurred ones, which can result from various factors such as camera shake, object motion, or defocus during image capture [28]. Over the years, significant advancements have been made in this domain, particularly with the advent of deep learning techniques and models achieving remarkable success in natural scene deblurring [8, 30].

Nonetheless, these models often fall short when applied to images captured in physical experiments, particularly in scenarios involving rapid, complex dynamics [12]. Unlike natural scenes where objects move predictably, physical experiments frequently involve phenomena that evolve swiftly and unpredictably, such as hydrodynamic instabilities. These phenomena introduce unique challenges in imaging, where motion blur is not merely the result of a moving object but is intrinsically tied to the evolving physical process [18]. Moreover, the limited number of images that can be captured in these experiments, due to the rapid dynamics, necessitates highly precise deblurring techniques.

*Equal contribution.
[1]Source code and dataset: https://github.com/Scientific-Computing-Lab/ClariPhy

Machine Learning and the Physical Sciences Workshop, NeurIPS 2024.

One specific domain where the challenges of image deblurring become particularly evident is X-ray radiography, a powerful imaging technique widely used in medical diagnostics, non-destructive testing, and physical experiments, such as studies of hydrodynamic instabilities in High-Energy-Density Physics (HEDP) experiments [3, 17, 1, 19, 20, 4]. In these applications, capturing sharp radiographic images is crucial, as they are the basis for further processing steps, including 3D reconstruction, extraction of initial experimental conditions, and assessment of material composition [27]. The rapid and complex dynamics in these experiments demand precise deblurring methods to ensure that the quantitative physical properties extracted from the images are accurate and reliable [25].

While X-ray radiography is extensively used in medical imaging [15], and several deblurring methods have been developed for these scenarios [16], their application in physical experiments involving instabilities presents distinct challenges. In studies of phenomena like the Rayleigh-Taylor instability (RTI) [9] the systems under observation are highly dynamic, allowing only a single or few radiographs to be taken. This rapid and unpredictable movement causes significant motion blur, as the boundaries of the measured objects shift during imaging. Additionally, both the measurement apparatus and the experimental setup introduce further spatial blurring and noise into the radiographic images.

Consequently, there is a critical need for advanced computational tools that can effectively remove both motion and spatial blur from these images. While classical deconvolution methods can address simple blurring and noise, they fall short in reconstructing sharp images when dealing with the complex combination of motion and spatial blur inherent in these physical experiments [2]. Effective deblurring techniques are therefore essential to enable accurate quantitative analysis and extraction of meaningful physical properties from radiographic data in such challenging environments.

1.1 Related Work

Deep-learning Deblurring: Traditional deblurring techniques primarily relied on deconvolution methods [6]. However, these methods often struggled with real-world images, particularly when the blur was spatially variant or when the noise levels were high. Deep learning has significantly shifted the paradigm for image deblurring by moving away from explicit blur kernel estimation. Instead, deep neural networks (DNNs) learn to map a blurred image to its sharp counterpart directly [23]. Later, generative adversarial network (GAN) methods were introduced [10], improving the perceptual quality of deblurred images and setting a new standard for visual plausibility in deblurring tasks. Vision Transformers (ViTs) have recently emerged as a powerful alternative to CNNs for various image-processing tasks, including deblurring. Unlike CNNs, which rely on local receptive fields, Vision Transformers leverage self-attention mechanisms to capture long-range dependencies and global context, which are crucial for effectively handling spatially variant blur [13, 26, 29].

Deblurring for HEDP and Hydrodynamic Instabilities: Until recently, denoising and deblurring radiographic images from High Energy Density Physics (HEDP) experiments primarily relied on classical methods [31, 14]. Recent studies have introduced CNNs for these tasks. For example, [11] uses a deep convolutional neural network with a modified DenseNet architecture, assuming a noise and blur model based on Poisson noise with additional Gaussian blur factors. However, due to the scarcity of physical simulation images, the model is mainly trained on natural images with synthetic noise. Another study [21] presents a direct 3D reconstruction of Inertial Confinement Fusion (ICF) experiments for Richtmyer–Meshkov instability using an attention-based transformer network, without prior deblurring. In contrast, our study leverages a large-scale dataset of full simulation images of the RTI, to which realistic temporal and spatial blur effects have been added. This enables training a model for physical deblurring based on the physical phenomenon accurate development as described in the simulations, resulting in a robust physically-aware approach.

Contributions. Contrary to prevalent deblur tasks when a static object is captured through time, physical experiments often include a dynamic, place, and structural changing object through time. While plenty of models address the first task, many fail when tackling the second. In this study, we have created a dataset comprising pairs of physical images, specifically RTI, when one image is sharp and the other has motion and spatial blur. We have developed an effective vision-transformer model based on fine-tuning Restormer [29] for deblurring physical experiments in general, and demonstrating good results on CLARIPHY dataset (section 2).

The main contributions of this paper are:

- CLARIPHY dataset creation — Spatial and temporal blur has been applied to the RayleAI dataset to create pairs of blur and original examples (section 2).
- Rayleigh-Taylor instability Restoration model, named CLARIPHY — We fine-tuned the pre-trained Restormer model, a ViT model that is trained for image restoration tasks (section 3).

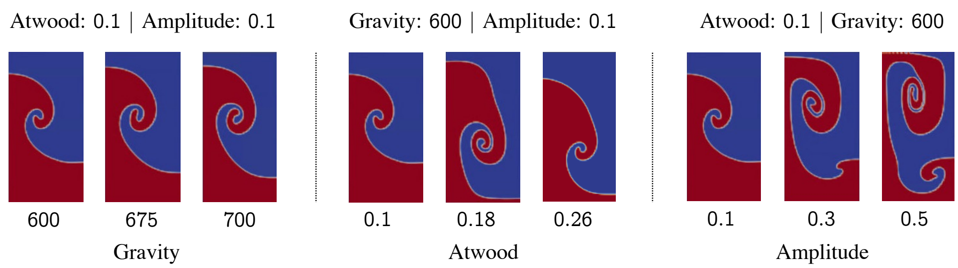


Figure 1: RayleAI dataset examples by alternating gravity (g) [cm/s^2] (negative values), Atwood ($\mathcal{A}$), and Amplitude (h) [cm]. While the time step presented is fixed at 0.55, it is clear that the different parameters influence the instability shape significantly.

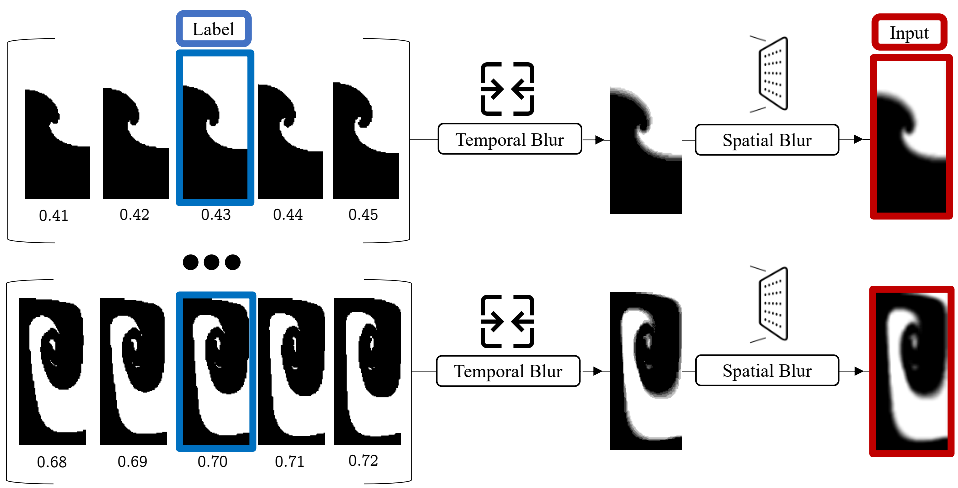


Figure 2: CLARIPHY dataset formation process based on RayleAI dataset. The new database is created through physical parameter alterations such as amplitude, gravity, and Atwood. This, in turn, results in different Rayleigh-Taylor instabilities' processes, which are then pre-processed. Pairs of temporal-spatial blur and the original process are produced. Temporal blur is achieved using an average of 5 following images in a time-set, while spatial is achieved by a convolution with a sphere.

2 Deblurred Labeled Dataset of Rayleigh-Taylor Hydrodynamic Instability

RTI occurs at the interface between two fluids of differing densities when a lighter fluid exerts pressure on a heavier fluid, leading to an unstable configuration [22]. This instability is observed in a wide range of hydrodynamic experiments and natural phenomena, such as ICF, the behavior of water suspended on oil under Earth's gravity, and various astrophysical systems [5].

In a prior study [7], a comprehensive dataset named RayleAI was developed, based on hydrodynamic simulations of RTI inspired by the well-known experiment conducted by Waddell et al. [24] (see examples in Figure 1). The RayleAI dataset was specifically designed to support the analysis of RTI evolution using advanced computer vision techniques. It includes over 100,000 images generated from 1,350 different simulations, with each simulation captured at 75 distinct time intervals. These simulations encompass variations in critical physical parameters that influence RTI, such as density ratio, acceleration, and initial perturbation amplitude.

As previously explained, in real-world experiments, and in contrast to the clear simulation images in RayleAI, temporal and spatial blur occur. Thus, we leveraged the RayleAI dataset to create a new dataset consisting of pairs of input clear images as ground truth and artificially created blurred images, simulating potential outcomes of physical experiments. Spatial and temporal blur has been chosen to imitate real physical experiment measurements. The spatial blur is created by sphere convolution at 4 different levels (sphere radius of 3,5,7 or 9 pixels), and the temporal blur by averaging images pixel-wise, each in a different timestamp at 5 different levels (averaging 3,5,7,9, or 13 consecutive time-points) (Figure 2).

3 ClariPhy Model and Experimental Results

The ClariPhy model is a fine-tuned version of the Restormer architecture, specifically trained to deblur input blurred images, restoring them to their corresponding GT images, which are unblurred. ClariPhy was compared against the original Restormer model inference and ClariPhy-tiny—a variant of this Restormer fine-tune trained on just 10% of the data. The model training was carried out with a batch size of 8, 30 epochs, an L1 loss, and an Adam optimizer. The model was trained and executed on a 40GB NVIDIA A100 GPU using PyTorch.

The performance of the models is illustrated in Figure 3 (over time complexity in the dataset) and Figure 4 (over epochs during training evaluation), using two key metrics: Structural Similarity Index (SSIM) and Peak Signal-to-Noise Ratio (PSNR). The ClariPhy model consistently outperforms the original Restormer model throughout the experiment, maintaining a higher structural similarity to the ground truth across all timestamps, with only minor fluctuations as the complexity of the examples increases over time. The higher SSIM values compared to PSNR highlight SSIM's ability to detect subtle quality improvements through structural and perceptual analysis, which PSNR, focusing on pixel-wise errors, may overlook. Interestingly, ClariPhy-tiny, despite being trained on only a fraction of the data, closely tracks the performance of the full ClariPhy model, suggesting that our models are highly efficient learners.

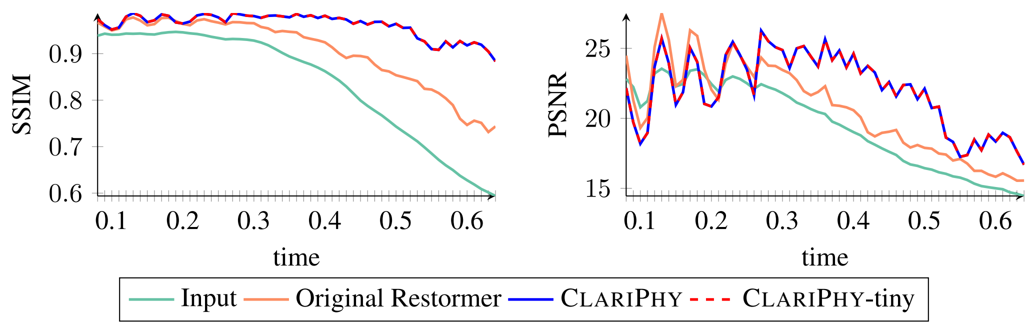


Figure 3: Deblurring models performances in SSIM and PSNR metrics as a function of time, while every timestamp is another, different, harder example to deblur. Our models' trends — ClariPhy, and ClariPhy-tiny, trained only on 10% of the database, show convergence. This implies that a tiny portion of our database is sufficient for qualitative learning. As time passes examples are more complex, hence, metrics are decreasing. Yet, our model remains superior to SOTA Restormer.

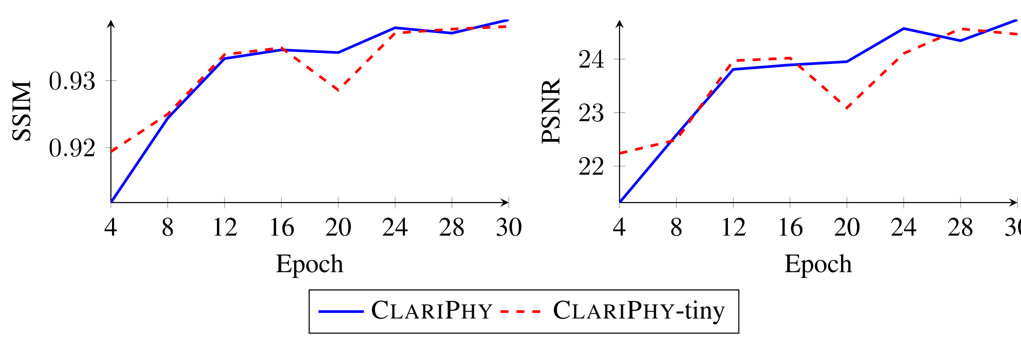


Figure 4: Metrics trends on validation set for both ClariPhy and ClariPhy-tiny are the same, thus, demonstrating the 10% of the dataset is sufficient for learning.

To provide a qualitative assessment, Table 1 showcases the CLARIPHY model's output across varying levels of difficulty: easy, medium, and hard. In the easy examples, the differences between the input and the reconstructed output are minimal, with the original Restormer and CLARIPHY accurately approximating the ground truth. As the difficulty increases, the challenge becomes apparent, particularly for the original Restormer, which struggles to capture the finer details of the RTIs. However, CLARIPHY demonstrates remarkable resilience, maintaining its performance even in the most complex cases.

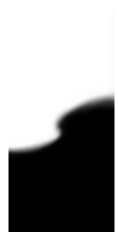
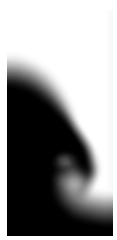
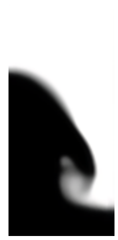
Difficulty	Input	Original Restormer	CLARIPHY	Ground Truth
Easy				
Medium				
Medium				
Hard				
Hard				

Table 1: Comparison of different methods (Input, Original Restormer, CLARIPHY, Ground Truth) across varying difficulty levels (Easy, Medium, Hard).

Acknowledgments

This research was supported by the Pazy Foundation and the Lynn and William Frankel Center for Computer Science. Computational support was provided by the NegevHPC project.

References

[1] Adrianna M Angulo, Sabrina R Nagel, Channing M Huntington, Christopher Weber, Harry F Robey, Gareth N Hall, Louisa Pickworth, and Carolyn C Kuranz. Design of a high-resolution rayleigh-taylor experiment with the crystal backlighter imager on the national ignition facility. *Journal of Instrumentation*, 17(02):P02025, 2022.

[2] Ashwini M Deshpande and Sampa Roy. An efficient image deblurring method with a deep convolutional neural network for satellite imagery. *Journal of the Indian Society of Remote Sensing*, 49(11):2903–2917, 2021.

[3] A Do, AM Angulo, GN Hall, SR Nagel, N Izumi, BJ Kozioziemski, T McCarville, JM Ayers, and DK Bradley. X-ray imaging of rayleigh–taylor instabilities using fresnel zone plate at the national ignition facility. *Review of Scientific Instruments*, 92(5), 2021.

[4] R Paul Drake. High-energy-density physics. *Physics Today*, 63(6):28–33, 2010.

[5] Philip G Drazin. *Introduction to hydrodynamic stability*, volume 32. Cambridge university press, 2002.

[6] Rob Fergus, Barun Singh, Aaron Hertzmann, Sam T Roweis, and William T Freeman. Removing camera shake from a single photograph. In *Acm Siggraph 2006 Papers*, pages 787–794. 2006.

[7] Re'em Harel, Matan Rusanovsky, Yehonatan Fridman, Assaf Shimony, and Gal Oren. Complete deep computer-vision methodology for investigating hydrodynamic instabilities. In *High Performance Computing: ISC High Performance 2020 International Workshops, Frankfurt, Germany, June 21–25, 2020, Revised Selected Papers 35*, pages 61–80. Springer, 2020.

[8] Jaihyun Koh, Jangho Lee, and Sungroh Yoon. Single-image deblurring with neural networks: A comparative survey. *Computer Vision and Image Understanding*, 203:103134, 2021.

[9] Hans-Jörg Kull. Theory of the rayleigh-taylor instability. *Physics reports*, 206(5):197–325, 1991.

[10] Orest Kupyn, Volodymyr Budzan, Mykola Mykhailych, Dmytro Mishkin, and Jiří Matas. Deblurgan: Blind motion deblurring using conditional adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 8183–8192, 2018.

[11] Joseph M Levesque, Elizabeth C Merritt, Kirk A Flippo, Alexander M Rasmus, and Forrest W Doss. Neural network denoising of x-ray images from high-energy-density experiments. *Review of Scientific Instruments*, 95(6), 2024.

[12] Haoying Li, Ziran Zhang, Tingting Jiang, Peng Luo, Huajun Feng, and Zhihai Xu. Real-world deep local motion deblurring. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 37, pages 1314–1322, 2023.

[13] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, pages 1833–1844, October 2021.

[14] Tuwanda McKenzie and Naima Naheed. Deblurring icf images using hybrid iterative regularization techniques. In *2023 International Conference on Computational Science and Computational Intelligence (CSCI)*, pages 1271–1274. IEEE, 2023.

[15] Maria Pia Morigi and Fauzia Albertin. X-ray digital radiography and computed tomography, 2022.

[16] Van-Giang Nguyen. Digital radiography with a consumer camera: Image denoising and deblurring. *IEIE Transactions on Smart Processing & Computing*, 10(5):398–406, 2021.

[17] G Rigon, B Albertazzi, T Pikuz, P Mabey, V Bouffetier, N Ozaki, T Vinci, F Barbato, E Falize, Y Inubushi, et al. Micron-scale phenomena observed in a turbulent laser-produced plasma. *Nature communications*, 12(1):2679, 2021.

[18] Denys Rozumnyi, Martin R Oswald, Vittorio Ferrari, Jiri Matas, and Marc Pollefeys. Defmo: Deblurring and shape recovery of fast moving objects. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3456–3465, 2021.

[19] Nadav Schneider, Matan Rusanovsky, Raz Gvishi, and Gal Oren. Determining hedp foams' quality with multi-view deep learning classification. In *2022 IEEE/ACM International Workshop on Artificial Intelligence and Machine Learning for Scientific Applications (AI4S)*, pages 19–25. IEEE, 2022.

[20] Nadav Schneider, Muriel Tzdaka, Galit Strum, Guy Lazovski, Galit Bar, Gilad Oren, Raz Gvishi, and Gal Oren. Explainable deep multi-view methodology for hedp foam quality assessment. In *AI for Accelerated Materials Design-Vienna 2024*.

[21] Daniel A Serino, Marc L Klasky, Balasubramanya T Nadiga, Xiaojian Xu, and Trevor Wilcox. Reconstructing richtmyer-meshkov instabilities from noisy radiographs using low dimensional features and attention-based neural networks. *arXiv preprint arXiv:2408.00985*, 2024.

[22] DH Sharp. Overview of rayleigh-taylor instability (no. la-ur-83–2130; conf-8305110-2), 1983.

[23] Jian Sun, Wenfei Cao, Zongben Xu, and Jean Ponce. Learning a convolutional neural network for non-uniform motion blur removal. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 769–777, 2015.

[24] JT Waddell, CE Niederhaus, and Jeffrey W Jacobs. Experimental study of rayleigh–taylor instability: low atwood number liquid systems with single-mode initial perturbations. *Physics of Fluids*, 13(5):1263–1273, 2001.

[25] Zhehui Wang, Andrew FT Leong, Angelo Dragone, Arianna E Gleason, Rafael Ballabriga, Christopher Campbell, Michael Campbell, Samuel J Clark, Cinzia Da Vià, Dana M Dattelbaum, et al. Ultrafast radiographic imaging and tracking: An overview of instruments, methods, data, and applications. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, page 168690, 2023.

[26] Zhendong Wang, Xiaodong Cun, Jianmin Bao, Wengang Zhou, Jianzhuang Liu, and Houqiang Li. Uformer: A general u-shaped transformer for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 17683–17693, June 2022.

[27] Bradley T Wolfe, Zhizhong Han, Jonathan S Ben-Benjamin, John L Kline, David S Montgomery, Elizabeth C Merritt, Paul A Keiter, Eric Loomis, Brian M Patterson, Lindsey Kuettner, et al. Neural network for 3d inertial confinement fusion shell reconstruction from single radiographs. *Review of Scientific Instruments*, 92(3), 2021.

[28] Sudha Yadav, Charu Jain, and Aarti Chugh. Evaluation of image deblurring techniques. *International Journal of Computer Applications*, 139(12):32–36, 2016.

[29] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 5728–5739, 2022.

[30] Kaihao Zhang, Wenqi Ren, Wenhan Luo, Wei-Sheng Lai, Björn Stenger, Ming-Hsuan Yang, and Hongdong Li. Deep image deblurring: A survey. *International Journal of Computer Vision*, 130(9):2103–2130, 2022.

[31] Zong-Qing Zhao, Wei-Hua He, Jian Wang, Yi-Dan Hao, Lei-Feng Cao, Yu-Qiu Gu, and Bao-Han Zhang. An improved deconvolution method for x-ray coded imaging in inertial confinement fusion. *Chinese Physics B*, 22(10):104202, 2013.