Complete Deep Computer-Vision Methodology for Investigating Hydrodynamic Instabilities

Re'em Harel[1,2], Matan Rusanovsky[2,3], Yehonatan Fridman[2,3], Assaf Shimony[4], and Gal Oren[3,4*]

[1] Department of Physics, Bar-Ilan University, Ramat-Gan, Israel
[2] Israel Atomic Energy Commission, P.O.B. 7061, Tel Aviv, Israel
[3] Department of Computer Science, Ben-Gurion University of the Negev, P.O.B. 653, Be'er-Sheva, Israel
[4] Department of Physics, Nuclear Research Center-Negev, P.O.B. 9001, Be'er-Sheva, Israel

reemharel22@gmail.com, matanru@post.bgu.ac.il, fridyeh@post.bgu.ac.il, shimonya@gmail.com, orenw@post.bgu.ac.il

Abstract. In fluid dynamics, one of the most important research fields is hydrodynamic instabilities and their evolution in different flow regimes. The investigation of said instabilities is concerned with highly non-linear dynamics. Currently, three main methods are used for understanding of such phenomena – namely analytical and statistical models, experiments, and simulations – and all of them are primarily investigated and correlated using human expertise. This work demonstrates how a major portion of this research effort could and should be analysed using recent breakthrough advancements in the field of Computer Vision with Deep Learning (CVDL, or Deep Computer-Vision). Specifically, this work targets and evaluates specific state-of-the-art techniques – such as Image Retrieval, Template Matching, Parameters Regression and Spatiotemporal Prediction – for the quantitative and qualitative benefits they provide. In order to do so, this research focuses mainly on one of the most representative instabilities, the Rayleigh-Taylor instability (RTI). We include an annotated database of images returned from simulations of RTI (*RayleAI*). Finally, adjusted experimental results and novel *physical loss* methodologies were used to validate the correspondence of the predicted results to actual physical reality to evaluate the model efficiency. The techniques which were developed and proved in this work can serve as essential tools for physicists in the field of hydrodynamics for investigating a variety of physical systems. Some of them can be easily applied on already existing simulation results, while others could be used via Transfer Learning to other instabilities research. All models as well as the dataset that was created for this work, are publicly available at: https://github.com/scientific-computing-nrcn/SimulAI.

Keywords: Fluid Dynamics, Hydrodynamic Instabilities, Rayleigh-Taylor Instability, Computer Vision, Deep Learning, Image Retrieval, Template Matching, Regressive Convolutional Neural Networks, Spatiotemporal Prediction

*Corresponding author

1 Introduction

The Rayleigh-Taylor instability occurs in an interface between two fluids with different densities in which the lighter fluid pushes the heavier fluid [1]. RTI is found in many hydrodynamic experiments and natural phenomena such as water suspended in oil in earth's gravity, Inertial Confinement Fusion (ICF), astrophysical systems, and many more [2]. Due to its importance, numerous experiments studying the growth of the instability and its effects on other phenomena are performed constantly all over the world [3–6]. Generally speaking, there are two types of experimental platforms for investigating the evolution of RTI: Liquid or gas systems (for example [3–5]) and High-Energy-Density Physics (HEDP) systems, in which the fluids are in plasma state after being heated by powerful lasers (for example [7,8]). In the former systems, it is difficult to control the initial perturbation, but due to the fact that between consecutive frames (2D images, the main diagnostics in the experiments) the time difference is short (compared to the duration of the experiment), namely it typically varies from milliseconds to seconds. Therefore, it is feasible to obtain with a fast camera tens or more frames per experimental shot. In the latter systems, the initial perturbation can be machined precisely prior to the laser drive – while the materials are in a solid state – but the time scales are much shorter (about tens of nanoseconds) and only one or at most few frames can be obtained from a single experimental shot. Therefore, the experimental data from both types of experimental systems contain a partial reflection of the instability – either the exact initial perturbation or the detailed temporal evolution of the instability is missing, while both are crucial for understanding the phenomenon. The growth of the perturbation in RTI depends on numerous variables such as viscosity, ablation, surface tension, small density gradients, and more. These variables are of different importance in different physical and experimental systems. In this work the case of two incompressible and immiscible fluids and a single-mode sinusoidal initial perturbation is considered. In this case, the early growth is exponential in time (T) and is proportional to (via linear stability theory):

$$e^{\sqrt{\mathcal{A}kg}T} \tag{1}$$

where k is the wave number $k = 2\pi/\lambda$, λ is the wavelength, g is earth's gravity (and in the general case the acceleration of the system), and $\mathcal{A}$ is the well known Atwood number, given by:

$$\mathcal{A} = \frac{\rho_1 - \rho_2}{\rho_1 + \rho_2} \tag{2}$$

where ρ_1 is the density of the heavier fluid and ρ_2 of the lighter one. In the late non-linear growth of such a single-mode perturbation, bubbles of the lighter fluid penetrate into the heavy fluid and spikes of heavy fluid penetrate into the light fluid at constant velocities, given by [9]:

$$u_{b/s} = \sqrt{\frac{2\mathcal{A}g\lambda}{c_d\left(1 \pm \mathcal{A}\right)}} \tag{3}$$

where u_b and u_s are the velocities of the bubble and the spike, respectively. The quantity c_d is the drag coefficient, which equals to 6π and 3π for 2D and 3D, respectively. As the perturbation grows, the shear velocities on the sides of the bubbles and the spikes create vortices due to the Kelvin-Helmholtz instability (KHI), in which the two materials mix in small length scales. Needless to say that in reality (experiments), knowing the exact conditions of the density and the viscosity of the two fluids is unrealistic. A possible way to bridge this gap is via simulations [10] (which are much cheaper than performing additional experiments): Given the correct initial parameters one can simulate the experiment and extract the missing time frames. However, initiating the simulation with the exact initial parameters is impossible due to experimental uncertainties. Nevertheless, it is still a viable solution as one can run parameter-sweep and select the most similar simulation in comparison to the experiment. However, this solution might be difficult as there are many different parameters (both physical parameters with uncertainties and parameters in the analysis of the experimental results), which makes this process hard for a human. Nowadays, the usage of different deep learning techniques in this scientific area is growing [11–19], and as it progresses, the above problem might also be solved using Computer Vision with Deep Learning (CVDL, or Deep Computer-Vision), since the introduction of deep learning techniques to the Computational Fluid Dynamics (CFDs) field proved to yield excellent results [20–26]. Therefore, we first define and devise several key Computer Vision (CV) problems, which collaboratively will enhance the understanding of RTI and other physical phenomena as follows:

1. Given a diagnostic *image* from a simulation/experiment, *sort* a database in accordance with an image similarity score to the input.
2. Given a diagnostic *image* from simulation/experiment, *extract the parameters* of the simulation that yields a best match to an image in a database.
3. Given a *partial template* of a phenomena from a simulation/experiment, *find* matches in a database and *sort* them in accordance with an image similarity score to the input.
4. Given a *set of images* that correspond to a set of time steps, and a time parameter T, *predict* future non-existing time steps images.

#	**Task**	**Quantitative & Qualitative Advantages**
I.	Database sorting by an image similarity score.	• Meaningful order. • Extraction of non-regressive parameters.
II.	Regressive parameter extraction.	• Physical parameters that fit the experimental data. • Uncertainty margins of the model training.
III.	Find and sort partial templates by an image similarity score.	• Meaningful order. • An extensive and reliable matching survey which decrease the uncertainty margins for the template assumption compared to full images analysis only.
IV.	Temporal inter/extrapolation of experiments.	• Data augmentation for low-data experiments. • Assurance and extension of a model.

Table 1: Quantitative and qualitative advantages for defined problems using CVDL.

The completion of the tasks above using CVDL techniques have both quantitative and qualitative advantages over classical optimization techniques, as summarized in Table 1:

I. **Image Retrieval**: A database (in this case, images) is sorted by an image similarity score, that corresponds to a meaningful physical similarity ordering score. In addition, the physical simulation parameters that yield the maximal image similarity (for example, with the experimental results) can be used as a non-regressive optimization. Similarly, uncertainty margins for the physical parameters can be calculated by defining a minimal required similarity factor.
II. **Parameter Regression**: Given experimental results and corresponding simulations with a set of free parameters, the technique provides the values of the free parameters for a best fit to the experimental results *iff* the model training and validation loss converges to approximately the same (small) value (thus, it successfully generalized the problem). The advantage over the parameter extraction in the first technique results from the regression process, which is significant when the simulation database is incomplete.
III. **Partial Template Matching**: Generally, it is valuable for physical problems in which there is a measurable pattern that is sensitive to any of the physical parameters. For example, in [27], the evolution of vortices, created by supersonic KHI, was measured and compared with hydrodynamic simulations. The analysis in [27] was based on the large-scale structures (i.e. the widths of the vortices). The medium-scale structures (i.e. the roll-ups within the vortices) were measured in the experiment and were compared to the simulations qualitatively. A more detailed analysis of the templates of the roll-ups could provide additional physical insights such as the effect of viscosity in the experimental conditions. Another example, that is relevant to the evolution of RTI, is originated from morphology differences between experiments in HEDP platforms and simulations of them. A detailed analysis of the morphology of the bubbles and the spikes can provide insights on magnetic effects due to the plasma conditions in these experiments [28] or perhaps other physical effects. A third example is the measurement of ablative RTI [29,30] that is relevant to astrophysical systems. Similar to the KHI example, The ablation effects were analyzed by the width of the mixing zone. The morphology of the spikes was affected by the ablation as clearly seen from the experimental images and simulations. A detailed analysis of the structure of the spikes can provide further insights on the ablation effects. In each of the examples above, the Partial Template Matching technique might provide a meaningful image similarity order between the sub-patterns within the simulation images to the template input of the experimental image. In addition, this technique provides a more extensive and more reliable matching survey, which can decrease the uncertainty margins for the physical parameters, compared to analyses of the full images only. Therefore, it can serve as a convenient method for analyzing the physical effects and their significance.

IV. **Spatiotemporal Prediction**: Provides a temporal interpolation/extrapolation of experimental results. First, it can be useful especially when the time-step ranges of the simulation and the experiment differs and when a prediction of experimental results at additional times are needed. Moreover, It can be used as a data augmentation methodology, by predicting images from some already known time-steps. This technique can also be used for self assurance: In cases of series of experimental images, one can provide images from a reduced range of time-steps. One possible way to evaluate the model is by predicting the images from time-steps outside the given range, and comparing them to the ones at hand.

We note that the techniques above can be applied on existing simulations databases, since physicists usually perform parameter surveys in order to analyze experimental data using simulations. In other words, it can be applied on previous works without running any additional simulation.

The rest of the paper is organized as follows: First, in Section 2 the state-of-the-art RTI database – *RayleAI* – that contains images with different parameters of a RTI simulation is introduced. Then, the different CVDL methods that are used in order to solve the key problems presented in Section 1 – specifically Information maximizing Generative Adversarial Network (InfoGAN), Parameters Regression (pReg), Quality-Aware Template Matching (QATM) and Predictive Recurrent Neural Network (PredRNN) – are covered with correspondence to the tasks above. Afterwards, a new evaluation methodology for the tasks above, named *physical loss* is introduced in Section 4. Next, the results of the different CVDL methods and evaluations are presented. Finally, future directions are proposed.

2 *RayleAI* – Database Characteristics

In order to implement the CVDL techniques described above, a state-of-the-art annotated simulation images database named *RayleAI* [31] was formed. The simulations were performed using the DAFNA hydrodynamic code [32], a multi-material Eulerian code with interface tracking and Adaptive Mesh Refinement (AMR) capabilities, which was validated for hydrodynamic instabilities [27,33]. The code solves the Euler equations, which are relevant to the turbulent flow in the simulated experimental system (Reynolds number of ∼6000 [6]). The initial perturbation is in the shape of a sinusoidal single-mode given by $y = h\cos(2\pi x/\lambda)$ where h is the initial amplitude height. The simulated domain is x =2.7cm (half of the wavelength, i.e. $\lambda = 5.4$cm) and y =5.4cm (h varies between simulations as detailed below), with reflected walls boundary conditions on all edges. The spatial resolution is 64x128 cells, which was found converged and yields the smallest experimentally measured patterns. Each fluid follows the equation of state of an ideal gas (with adiabatic index $\gamma = 5/3$) and a hydrostatic equilibrium was set adiabatically with a pressure of 1 bar on the interface. The simulation input consists of three free parameters: Atwood number, gravitational

acceleration and the amplitude of the perturbation (as well as additional time parameter). The database contains 101,250 images produced by 1350 different simulations, 75 time steps for each simulation, where the stride is 0.01s, with unique set of the free parameters per each simulation. The format of the repository is built upon directories, each represents a simulation performance with the directory name indicating the parameters of the specific simulation, and each directory holds several images corresponding to each relevant time step. For example, the directory *gravity_ 750_ amplitude_ 0.5_ atwood_ 0.16* is a simulation with $g = 750$ cm/s^2, initial amplitude of 0.5cm, and $\mathcal{A} = 0.16$. The ranges of the Atwood number, gravity and initial perturbation are presented in Table 3. Table 2 shows the simulation images from DAFNA compared to the experimental images, with $g = 750$ cm/s^2 and $\mathcal{A} = 0.16$.

The choice of these exact parameters were derived from the well known experimental results [6]. The physical parameters in the experiment were $\mathcal{A} = 0.155$, $g = 740$ cm/s^2 and $h = 0.54$cm. The initial amplitude is not given but can be estimated from the first image by about $h = 0.1$cm. Thus, one can deduce that the simulation in the database with $\mathcal{A} = 0.16$, $g = 750$ cm/s^2 and $h = 0.1$cm should produce the closet result to the experimental one (as shown in Table 2). An example of an experimental image is shown in Figure 1. The experiment images were originally taken in grayscale. For optimal results, each image was processed with two methods (Erode-Dilate vs. Histogram Equalization), and the most fitting result that resembles the interface between the two fluids (by expert opinion) was selected, cropped and resized, and then binarized by a threshold. Those images are also included in the database under the *experiment* folder.

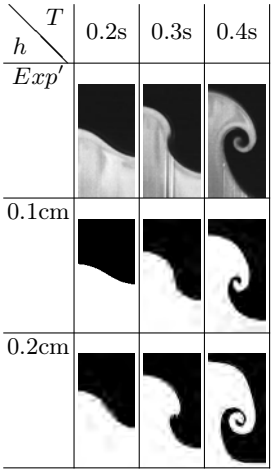


Table 2: Diagnostic of RTI in different T and h values from *RayleAI*.

Parameter	From	To	Stride
Atwood number ($\mathcal{A}$)	0.02	0.5	0.02
Gravity (g) [cm/s^2]	600	800	25
Amplitude (h) [cm]	0.1	0.5	0.1
X [cm]	2.7	2.7	0.0
Y [cm]	5.4	5.4	0.0

Table 3: Simulation parameters.

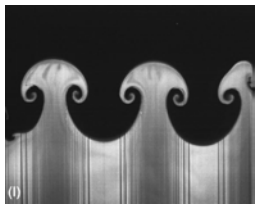


Fig. 1: The full image from the experiment (T=0.4s).

3 Deep Computer-Vision Methods

3.1 Task I: Image Retrieval using InfoGAN

Generative Advreserial Network (GAN) [34] is a framework capable of learning a *generator* network G that transforms noise variable z from some noise distribution into a generated sample $G(z)$, while the training of the generator is optimized against a *discriminator* network D, which targets to distinguish between real samples with generated ones. The fruitful competition of both G and D, allows G to generate samples such that D will have difficulty with distinguishing real samples from them. In the context of game theory, this competition might be referred as a variation of a MinMax game in which each player tries to to maximize the minimum gain. In this game for example, G maximizes its minimum gain by generating images with higher resemblance to real images, while D improves its distinguishing capabilities, thus hardening the generation process of G. The ability to generate indistinguishable new data in an unsupervised manner is one example of a machine learning approach that is able to understand an underlying deep, abstract and generative representation of the data. InfoGAN [35] utilizes latent code variables c_i, which are added to the noise variable. These noise variables are randomly generated from a user-specified domain. The latent variables impose an information theory regularization term to the optimization problem, which forces G to preserve the information stored in c_i through the generation process. This allows learning interpretative and meaningful representations of the data, with a negligible computation cost, on top of a GAN. The high-abstract-level representation can be extracted from the discriminator (e.g. the last layer before the classification) into a feature vector. These features can be used in order to measure the similarity between some input image to any other image, by applying some distance function (e.g. l_2 norm) on the features of the input to the features of the other image. This methodology provides the ability to order images similarity to a given input image [36].

In order to evaluate InfoGAN performances over *RayleAI*, the classic CV technique of LIRE [37] is used for comparison. LIRE is a library that provides image retrieval based on image characteristics among other classic features. LIRE is built on top of the open-source text search engine *Lucene* [38]. LIRE takes numeric images descriptors, which are mainly vectors or sets of vectors, and stores them inside a *Lucene* index as text along with the image path within a *Lucene* document. For the evaluation of the similarity of two images, one can calculate their distance in the space they were indexed to. Many state-of-the-art methods for extracting features can be used, such as Gabor Texture Features [39], Tamura Features [40], or FCTH [41]. For the purposes of this work, the Tamura Features method is the most suitable method that LIRE provides, as it indexes *RayleAI* images in a more dispersed fashion. The Tamura feature vector of an image is an 18 double values descriptor that represents texture features in the image that correspond to human visual perception.

3.2 Task II: Parameters Regression using Convolutional Neural Networks – *pReg*

Many Deep Learning techniques obtain state-of-the-art results for regression tasks, in a wide range of CV applications [42] such as Pose Estimation, Facial Landmark Detection, Age Estimation, Image Registration and Image Orientation [43,44]. Most of the deep learning architectures used for regression tasks on images are Convolutional Neural Networks (ConvNets), which are usually composed of blocks of Convolutional layers followed by a Pooling layer, and finally Fully-Connected layers. The dimension of the output layer depends on the task and its activation function is usually linear or sigmoid.

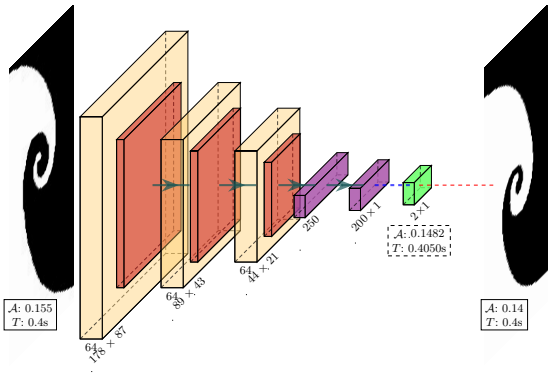


Fig. 2: ConvNet model, named *pReg*, for $\mathcal{A}$ and T parameters regression. Yellow layers are convolutional layers (5×5, with ReLU activation), orange are pooling layers, purple are dense layers with ReLU activation and the green layer is the output layer, activated by sigmoid. Green connections with arrows indicate connections with batch normalization, and the blue dashed connection is a regular connection. The left image is an example of an experimental input, with the real parameters of $\mathcal{A}$: 0.155, T: $0.4s$, which the model predicts for the values $\mathcal{A}$: 0.1482, T: $0.4050s$ as can be seen under the green layer. The red dotted trend indicates the similarity search operation that quantifies the distance of images from *RayleAI*, based on the l_2 distance over $\mathcal{A}$ and T.

ConvNets can be used for retrieving the parameters of an experiment image, via regression. The presented model (henceforth *pReg*) (Figure 2) consists of 3 Convolutional layers with 64 filters, with a kernel size 5×5, and with l_2 regularization, each followed by a Max-Pooling layer, a Dropout of 0.1 rate, and finally Batch Normalization. Then, there are two Fully-Connected layers of 250 and 200 features, which are separated again by a Batch Normalization layer. Finally, the Output layer of the network has 2 features (as will described next), and is activated by sigmoid which maps the values to $[0, 1]$, to prevent the exploding gradients problem. This is important, as in regression tasks gradients might grow rapidly and eventually explode, if there are no regularizations on the output features. Since the most significant parameters for describing each image frame are Atwood number and time – which *pReg* is trained to predict – only a subset of *RayleAI* was used for the training set, namely images with the following pa-

rameters: $\mathcal{A} \in [0.08, 0.5]$ (with a stride of 0.02), $g \in \{625, 700, 750, 800\}$ cm/s^2, $h \in [0.1, 0.5]$cm (with a stride of 0.1cm) and $T \in [0.1, 0.6]$s (with a stride of 0.01s). A small amount of values for gravity and for amplitude was fixed, so that the network will not try to learn the variance that these parameters impose while expanding the database with as minimal noise as possible. The value ranges of Atwood number and time were chosen in order to expose the model to images with both small and big perturbations such that the amount of the latter case ones will not be negligible. The reduced training set consists of ∼16K images and the validation set consists of ∼4K images. Nonetheless, for increasing generalization and for decreasing model overfitting, data augmentation was employed. Since there is high significance for the perspective from which each image is taken, the methods of data augmentation should be carefully chosen: Rotation, shifting and flipping methods may generate images such that the labels of the original parameters do not fit for them. Therefore, the training set was augmented with only zooming in/out (zoom range=0.1) via TensorFlow [45] preprocessing.

3.3 Task III: Quality-Aware Template Matching using QATM

One variation of the Template Matching problem is defined as follows: Given an exemplar image E, find the most similar region of interest in a target image S [46]. Classical template matching methods often use Sum-of-Squared Differences (SSD) and Normalized Cross-Correlation (NCC) [47] to asses the similarity score between a template and an underlying image. These approaches work well when the transformation between the template and the target search image is simple. However, with non-rigid transformations, which are common in real-life, they tend to fail. The Quality-Aware Template Matching (QATM) [48] algorithm is inspired by assessing the matching quality of source and target patches. It takes the uniqueness of pairs into consideration rather than simply evaluating matching score. It defines the $QATM(e, s)$ - measure as the product of likelihoods that a patch $s \in S$ is matched in E and a patch $e \in E$ is matched in S. Once $QATM(e, s)$ is computed, one can compute the template matching map for the template image E and the target searched image S, and eventually, can find the best-matched region R^* which maximizes the overall matching quality. Therefore, the technique is of great need when templates are complicated and targets are noisy. Thus, it is most suitable for RTI images from simulations and experiments.

3.4 Task IV: Time Series Prediction using PredRNN

Learning the evolution of the RTI in order to predict future time or gap frames requires both understanding of the spatial aspects of each time frame (e.g. the interface between the fluids), and understanding of temporal development: As the time progresses, the simulation tends to be more and more chaotic. Convolutional Long Short Term Memory networks (CLSTMs) [49] is a class of algorithms which are able to predict future image states by past and present image states based on

training sequences of images. The architecture of this network is based on a two-dimensional grid of units that pass spatial information vertically (upwards), and temporal information horizontally (rightwards). However, the standard CLSTMs architectures lack the capability of preserving the temporal information for long terms, since the spatial information that is learned via the top unit in a specific time step is not passed to the bottom unit in the next time step, leading to the loss of important information. PredRNN [50] is a state-of-the-art Long Short Term Memory (LSTM) Recurrent Neural Network (RNN) for predictive learning. PredRNN memorizes both spatial appearances and temporal variations in a unified memory pool. Unlike standard LSTMs, and in addition to the standard memory transition within them, memory in PredRNN can travel through the whole network in a zigzag direction, i.e, from the top unit of some time step to the bottom unit of the other. Thus, PredRNN is able to preserve the temporal as well as the spatial memory for long-term motions. In this work, PredRNN was used for predicting future time steps of simulations as well as experiments, based on a given sequence of time steps.

4 Evaluation Methodology

In order to test how the discussed above techniques perform on physical simulations as well as experiments, new task-specific test methods that quantify how well each technique operates on a concrete database are proposed. Novel evaluation methodologies are presented for the techniques discussed in Sections 3.1, 3.2 and 3.3, based on a suitable corresponding loss measure for the first two tasks, and a sophisticated clustering-visualization method for the third (QATM). The evaluation of last technique presented in Section 3.4 will be discussed separately.

The first evaluation method, namely *physical loss*, quantifies *how meaningful the results of the technique are*, i.e., whether the results of the technique are reflected in the (physical) annotations of the data. For example, in the case of Image Retrieval, it is inconclusive to decide whether the results are sufficient solely based on visual examination, since it is a difficult task for humans to determine the correct ordering of lots of results. Thus, it is hard to establish whether the technique is satisfactory. Therefore, this work suggests to measure how each input image is *physically close* to each of the returned image outputs, based on some or all of their parameter labels. Thus, for Image Retrieval (and for Parameters Regression, explained later in Section 5.2) each output image gets two scores – one from the technique at hand, e.g. similarity score, and one from the difference between its parameters to the parameters of the input image. In the case of high correlations of these scores, one can infer that indeed the results of the technique are of a meaningful (physical) value. The Relative Error was chosen in this work as the parameter difference function, although it may be calculated via any desired error function. Note that since the ranges of the parameters are scaled differently, it is suggested to normalize them beforehand.

Furthermore, one important aspect that results from *physical loss* is the ability to identify the parameters which are likely to produce a small impact on

the simulation results (depending on time). For example, in the case of a small ratio between the amplitude and the wavelength of the perturbation (up to a few percent), RTI grows linearly according to Eq. 1 and approximately preserves its initial shape. Therefore, two simulations that differ only by their initial small amplitudes will practically result in the same late evolution up to a constant time shift. As a result, it is expected from physical considerations that if one produces an amplitude-based *physical loss* methodology for later time steps, the CVDL techniques will generate semi-random values of error as the amplitude hardly affects the simulation in later time steps, or in other words, once the two fluids are mixed and form a chaotic mixture of the two. A similar result is also expected for the gravity parameter since for incompressible fluids (a good approximation in this case), two simulations that differ only by the gravity parameter will practically result in the same evolution as a function of the normalized time. A useful definition of the normalized time is $\tilde{T} = \sqrt{\mathcal{A}g/\lambda}T$ as also reflected from Eq. 1. Concluding the above physical influence of the initial amplitude and the gravity parameters, only the Atwood number and time parameters should have a significant impact on the results and are expected to be identified using the *physical loss* methodology.

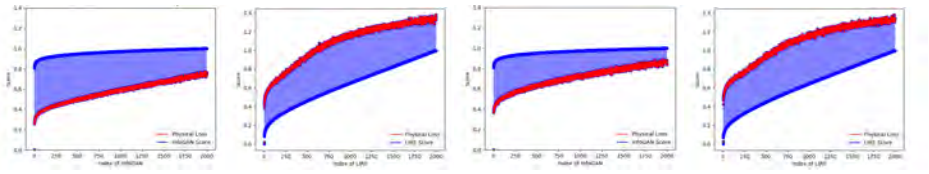
Another new evaluation method was developed for cases where there are no meaningful (physical) annotations. Specifically in the case of Template Matching, where some partial template is searched through a database. Unlike the *physical loss* case, the physical parameters of the returned partial region of interest have no unique physical labels, since it might be expected to find this template in images from a wide range of different parameters. To this end, a relaxed-evaluation method, that quantifies how well the technique at hand separates similar images from dissimilar images is presented. Similarly to the *physical loss* methodology, two values for each output are used: The score from the technique, and a cluster number – returned from some unsupervised clustering algorithm. Scenarios in which continuous sequences from the results of the technique are from the same cluster might indicate the ability of the technique to perform a proper distinction between classes of similarity to a given input template. Alternatively, cases of sequences of results from mixed clusters, especially in the first and most similar regions, might prove that the technique did not succeed in separating the most similar images from the rest. K-Means was used as the clustering algorithm, after extracting the main features from each image using Principal Component Analysis (PCA) in order to achieve more precise results [51]. Section 5 presents the evaluation results of the techniques from Section 3.

5 Results and Discussion

5.1 Task I: Image Retrieval

In order to test the performance of InfoGAN against LIRE, two separate test cases were studied: *Random* and *Complex*. In the first general test case, 13K random input test images were chosen from *RayleAI*. In the latter, approximately 10K input images with $T > 0.25$s were chosen in order to pick the most complex

images, as the RTI is more chaotic and dominant in this regime. Then, InfoGAN and LIRE were performed on the entire *RayleAI* dataset for both test cases. Then, for each tested image and for each tool, the results were sorted according to the similarity scores that were given by the model. To evaluate the results and quantify how well the tools performed, the *physical loss* methodology, introduced in Section 4, was employed over the Atwood number parameter. Then, for each tool and test case, the average *physical loss* was calculated.



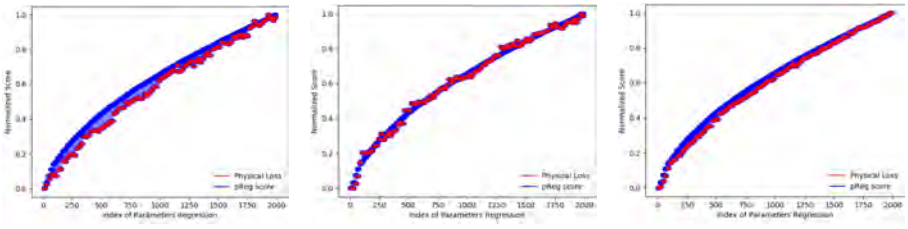
(a) *Complex* InfoGAN **(b)** *Complex* LIRE **(c)** *Random* InfoGAN **(d)** *Random* LIRE

Fig. 3: LIRE and InfoGAN averaged *physical loss* methodology over Atwood number.

Figure 3 presents the *physical loss* using a comparison between the technique score [in blue] and the *physical loss* [in red] per each index. A thin blue line is drawn to correlate them. We present only the Atwood number parameter, as it is the most significant parameter. Figures 3a and 3b, show that InfoGAN outperformed LIRE on the complex images test, as the averaged *physical loss* of the first – and most important – indices of InfoGAN is $\sim$0.25, in contrast to $\sim$0.4 of LIRE. Furthermore, InfoGAN outperforms LIRE along the entire 2K first examined indices, showing many powerful capabilities in the complex data case. Figures 3c and 3d, presents that InfoGAN and LIRE perform quite the same in the first indices, with averaged *physical loss* of $\sim$0.4. Yet, when focusing on the entire 2K first indices, it can be seen that InfoGAN starts to outperform LIRE with smaller *physical loss* values. Additionally, it seems that there is a higher correlation between the scores of InfoGAN to their corresponding *physical loss* values (blue and red lines act accordingly) in each of the test cases, which indicates again the InfoGAN has a better ability to learn the underlying physical pattern of the data. Another important aspect in which InfoGAN outperforms LIRE in both test-cases is the width of the *physical loss* trend: As the red trend is thinner, there is less noise and therefore the results have more physical sense. As one can notice, the red lines of InfoGAN are thinner than the red lines of LIRE. Note, that in all figures, the blue lines are normalized by the min-max normalization method, contrary to the red lines which are presented as raw values. The overwhelming superiority of InfoGAN is somehow expected and can be explained as the ability of a deep learning model to learn complex patterns from the new tailor-made and trained database. Although LIRE provides decent results without requiring to be trained on a specific organized database (which obtaining is not always an easy task), it is still a classical image processing tool, which lacks the learning capabilities that will allow it to understand deep patterns from the data. Therefore, for image retrieval applications with suitable databases, InfoGAN should be used.

5.2 Task II: Parameters Regression

In order to test the performances of the *pReg* network, evaluation tests that are similar to the tests presented in Section 4 were employed. *pReg* was used to predict the – activated by the sigmoid – parameters: $\mathcal{A}$ and T for 2K random images. Then, for each image a search was performed through *RayleAI* for the 2K images with the lowest scores, based on their l_2 distance between their $\mathcal{A}$ and T activated by sigmoid parameters, to that of the input image.



(a) Averaged T **(b)** Averaged $\mathcal{A}$ **(c)** Averaged $\mathcal{A}$ & T

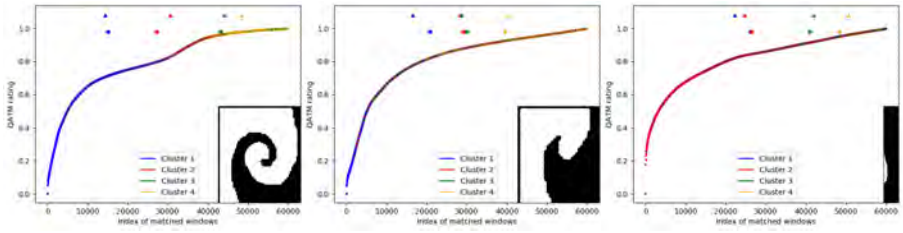
Fig. 4: *pReg* averaged *physical loss* over the Atwood number and the time.

As can be seen in Figure 4a, the model explains the dependence on time well, especially in the lowest ($\leq$ 500) and highest ($\geq$ 1500) indices, since the red dots of the normalized *physical loss* over the time, and the blue dots of the normalized l_2 distance from the predicted parameters, act similarly. The slightly higher difference between the dots in the middle of the scale ($500 <$ indices < 1500) is somehow expected, as it is harder for models to predict the parameters accurately in cases where there is a 'mild' physical difference. Yet, in cases where there is a high resemblance or significant difference with respect to the *physical loss*, it is more likely that the model will predict similar parameters or dissimilar parameters, respectively. Figure 4b shows that the model explains dependence on the Atwood number parameter even better, as the graphs are almost the same with some small noise. This can be explained by the significance and importance of the Atwood number parameter. Figure 4c shows that the combination of Atwood number and time greatly outperforms the former two cases, since the red trend almost converges to the blue trend. Note that as the predicted parameters in *pReg* are only $\mathcal{A}$ and T, and the difference is calculated only over them, for each input image there are lots of images in *RayleAI* with the same calculated distance – same $\mathcal{A}$ and T, but different g or h. Therefore the trends in Figs. 4b, 4a, 4c might have dense blocks of dots – with almost the same scores. Finally, since images with the same $\mathcal{A}$ and T but different g or h have the same loss (over $\mathcal{A}$ and T), they are ordered arbitrarily. Therefore, the *physical loss* over all parameters does not explain the similarity order of *pReg*, because of the arbitrariness that g and h impose.

5.3 Task III: Template Matching

For the evaluation of the QATM algorithm 16 meaningful templates were cropped from the experimental images. For each template, the following pro-

cedure was employed: The QATM algorithm was performed on each image in *RayleAI* and a matched sub-figure was found. Then, the results were sorted in accordance to the QATM matching scores. For results evaluation, the loss methodology of PCA and K-Means was employed as described in Section 4.



(a) Unique Template **(b)** Semi-Unique Template **(c)** Non-Unique Template

Fig. 5: PCA and K-Means clustering methodology made on QATM results.

Figure 5 presents the results of three representative templates, while in each the normalized score results of the algorithm are sorted in an increasing order. The color of every point represents the cluster of the corresponding template, achieved by the K-Means algorithm. The triangles and the circles lying above the curves represent the median and the average of the indices of each cluster, respectively. To keep the trends readable, only one of each 30 dots is presented. As can be seen in Figure 5a, the clustering algorithm divides the indices into four separate and distinct areas: There is pure congestion of blue dots in the first thousands of indices, without any rogue non-blue dots. This indicates that the algorithm understood the template successfully and found lots of significantly similar sub-figures. This powerful result might be explained because the searched template is of a 'unique' shape, which helps QATM extract lots of uncommon features and correlate them to *RayleAI*. Figure 5b, shows a pure congestion of blue dots in the first hundreds of indices, and a mixture of blue and red dots, with the unignorable presence of green dots in the right following indices. This mixture of clusters that appears in relatively small indices indicates that the algorithm's results start to be less meaningful after a couple of hundreds of indices. This can be explained by the search template that is less unique than the previous template. Finally, in Figure 5c the blue and red clusters seem to be inseparable all along the index axis. This indicates that the algorithm did not understand well the template and has difficulties to bring quality matched templates, as expected from the lack of uniqueness of the input template.

5.4 Task IV: Spatiotemporal Prediction

The PredRNN model was trained on *RayleAI* sequences of 0.01s time steps. However, RTI experiment's parameters are the following: $g = 740$ cm/s^2, estimated amplitude, and $\mathcal{A} = 0.155$ and it contains 12 black and white images with an interval of $\sim$0.033s between each couple of consecutive frames, while the time steps of *RayleAI*'s simulation are of 0.01s. In order to fill the missing time

steps, PredRNN was used to predict the missing time intervals of the experiment. The missing time steps were filled in an iterative manner, by predicting a single future time step at a time. Furthermore, the quality of the prediction of a simulation was tested with the following parameters: $g = 725$ cm/s^2, $\mathcal{A} = 0.14$ and $h = 0.3$cm. As an input, the first 10 images of the simulation were given, while predicting a total of 49 time steps.

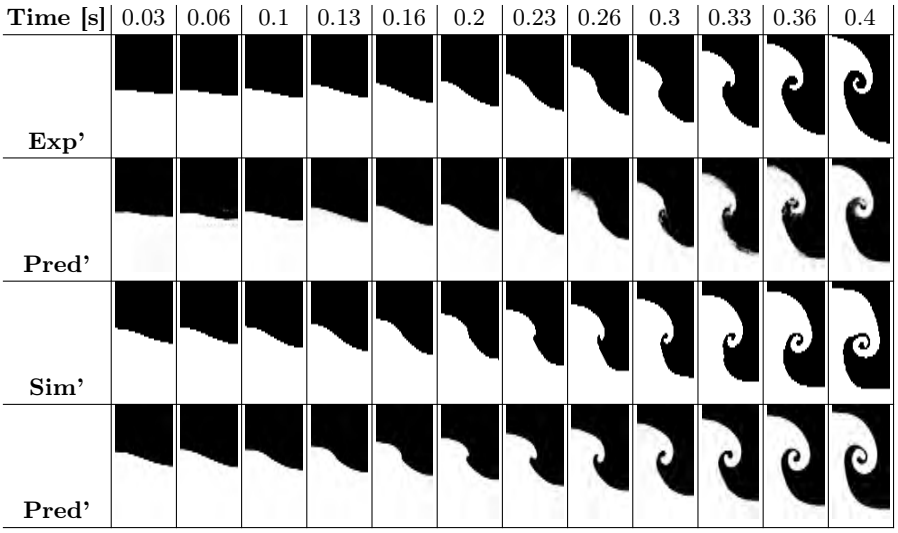
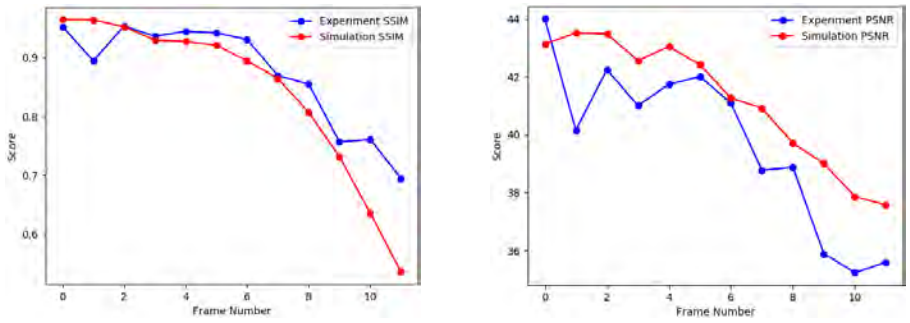


Table 4: PredRNN prediction of the experiment and simulation.



(a) PredRNN SSIM **(b)** PredRNN PSNR

Fig. 6: Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR) scores of the predicted experimental and simulated results.

The results of PredRNN prediction are shown in Table 4. The columns represent time steps ranging from the initial time of $T_{init} = 0.03$s to the final time of $T_{fin} = 0.4$s. The first and third rows of images represent the images of the corresponding time steps of the experiment and the simulation respectively, and as such can be considered to be Ground Truths (GTs). The second and the fourth rows of images represent the prediction of PredRNN on the corresponding time

steps of the *filled* experiment and the simulation, respectively. As one can see, PredRNN produces almost identical predictions to the GT images. To evaluate the quality of the image produced by PredRNN, Peak Signal-to-Noise Ratio score was used. PSNR measures the quality of reconstruction of loss compression codecs. When comparing the different images, PSNR is an approximation to human perception of reconstruction quality. Complementary, Structural Similarity Index Measure (SSIM) that predicts the perceived quality of the images [52] was used. In Figure 6 quantifies the quality of the predicted images using PSNR and SSIM evaluation tests between the produced PredRNN images to their corresponding GT images, similarly to [50]. Both scores measure the similarity between the predicted frame and its corresponding GT frame. These scores decrease as time progresses, due to the expected difficulty of the model to predict the distant future. However, its worth noting that a simple image sharpening on the predicted results can dramatically increase both SSIM and PSNR scores.

6 Conclusions and Future Work

This work presents a state-of-the-art complete CVDL methodology for investigating hydrodynamic instabilities. First, the problems were defined and their significance was emphasised. Second, a new comprehensive tagged database, which contains simulated diagnostics for training, and experimental ones for testing, was created for the needed learning process. Third, it was shown how the novel methodology targets the main acute problems in which CVDL can aid in the current analysis process, namely using deep image retrieval; regressive deep convolutional neural networks; quality aware deep template matching; and deep spatiotemporal prediction. Fourth, a new *physical loss* and evaluation methodology was formed. This methodology enables to compare the performances of the model against the physical reality, and by such to validate its predictions. At last, the usage of the methods on the trained models was exemplified and their performances were exemplified using the new *physical loss* methodology. In all of the four tasks, excellent results, which prove the methodology suitability to the problem domain were achieved. Thus, it is stressed that the proposed methodology can and should be an essential part of the hydrodynamic instabilities investigation toolkit, along with analytical and statistical models, experiments and simulations.

In regard to future work, an extension of the methodology might be useful for solving the discrepancy between simulations and experiments when it is clear that the initial parameters of the simulation does not cover the *entire* physical scope. Since in many cases a classical parameter sweep does not yield the desired results, an extension of the model – in the form of unmodelled parameters regression – should be used. For example, it might be useful for the physical problem presented in [29, 30], in which the simulation results do not cope with the experimental ones. Thus, an extended deep regressive parameter extraction model should be applied in a new form such that unknown parameters – i.e. parameters which are not part of the simulation initiation – could

be discovered and formulated. This is crucial, as numerous current efforts suggest that often there is a missing part in the understanding of the simulated results. Thus, preventing any traditional method to match the simulated results to the experimental ones. Once discovered, in order to understand and formulate said unknown parameters, an extensive Explainable AI (XAI) methodology [53] should be performed.

Another strength of the presented methodology is that it can be applied on an already existing data in case that parameter sweep was previously performed on other physical data. Therefore, it might yield physical insights without running any additional simulation. In addition, the toolkit can be easily suited to the physical problem. For example, if the width of vortices was investigated in a previous research [27,33,54], template matching would be useful for investigating the inner structure of the vortices; With series of experimental images from different times at hand, the spatiotemporal prediction can be used for prediction of results at unmeasured times. Finally, the models presented in this work might be invaluable for learning physical problems with less training data and a more complex form, using Transfer Learning [55].

Acknowledgments

This work was supported by the Lynn and William Frankel Center for Computer Science. Computational support was provided by the NegevHPC project [56].

References

1. David H Sharp. Overview of rayleigh-taylor instability. Technical report, Los Alamos National Lab., NM (USA), 1983.
2. Philip G Drazin. *Introduction to hydrodynamic stability*, volume 32. Cambridge university press, 2002.
3. KI Read. Experimental investigation of turbulent mixing by rayleigh-taylor instability. *Physica D Nonlinear Phenomena*, 12:45–58, 1984.
4. Stuart B Dalziel. Rayleigh-taylor instability: experiments with image analysis. *Dynamics of Atmospheres and Oceans*, 20(1-2):127–153, 1993.
5. Guy Dimonte and Marilyn Schneider. Turbulent rayleigh-taylor instability experiments with variable acceleration. *Physical review E*, 54(4):3740, 1996.
6. 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.
7. JP Knauer, R Betti, DK Bradley, TR Boehly, TJB Collins, VN Goncharov, PW McKenty, DD Meyerhofer, VA Smalyuk, CP Verdon, et al. Single-mode, rayleigh-taylor growth-rate measurements on the omega laser system. *Physics of Plasmas*, 7(1):338–345, 2000.
8. Bruce A Remington, Hye-Sook Park, Daniel T Casey, Robert M Cavallo, Daniel S Clark, Channing M Huntington, Carolyn C Kuranz, Aaron R Miles, Sabrina R Nagel, Kumar S Raman, et al. Rayleigh–taylor instabilities in high-energy density settings on the national ignition facility. *Proceedings of the National Academy of Sciences*, 116(37):18233–18238, 2019.

9. VN Goncharov. Analytical model of nonlinear, single-mode, classical rayleigh-taylor instability at arbitrary atwood numbers. *Physical review letters*, 88(13):134502, 2002.
10. David L Youngs. Numerical simulation of turbulent mixing by rayleigh-taylor instability. *Physica D: Nonlinear Phenomena*, 12(1-3):32–44, 1984.
11. Brian K Spears, James Brase, Peer-Timo Bremer, Barry Chen, John Field, Jim Gaffney, Michael Kruse, Steve Langer, Katie Lewis, Ryan Nora, et al. Deep learning: A guide for practitioners in the physical sciences. *Physics of Plasmas*, 25(8):080901, 2018.
12. Kelli D Humbird, Jayson Luc Peterson, BK Spears, and RG McClarren. Transfer learning to model inertial confinement fusion experiments. *IEEE Transactions on Plasma Science*, 2019.
13. A Gonoskov, Erik Wallin, A Polovinkin, and I Meyerov. Employing machine learning for theory validation and identification of experimental conditions in laser-plasma physics. *Scientific reports*, 9(1):7043, 2019.
14. Gonzalo Avaria, Jorge Ardila-Rey, Sergio Davis, Luis Orellana, Benjamín Cevallos, Cristian Pavez, and Leopoldo Soto. Hard x-ray emission detection using deep learning analysis of the radiated uhf electromagnetic signal from a plasma focus discharge. *IEEE Access*, 7:74899–74908, 2019.
15. Kelli Denise Humbird. *Machine Learning Guided Discovery and Design for Inertial Confinement Fusion*. PhD thesis, 2019.
16. Jim A Gaffney, Scott T Brandon, Kelli D Humbird, Michael KG Kruse, Ryan C Nora, J Luc Peterson, and Brian K Spears. Making inertial confinement fusion models more predictive. *Physics of Plasmas*, 26(8):082704, 2019.
17. Bogdan Kustowski, Jim A Gaffney, Brian K Spears, Gemma J Anderson, Jayaraman J Thiagarajan, and Rushil Anirudh. Transfer learning as a tool for reducing simulation bias: Application to inertial confinement fusion. *IEEE Transactions on Plasma Science*, 2019.
18. Yeun Jung Kim, Minsoo Lee, and Hae June Lee. Machine learning analysis for the soliton formation in resonant nonlinear three-wave interactions. *Journal of the Korean Physical Society*, 75(11):909–916, 2019.
19. A Gonoskov. Employing machine learning in theoretical and experimental studies of high-intensity laser-plasma interactions.
20. Maziar Raissi, Paris Perdikaris, and George E Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019.
21. Maziar Raissi, Zhicheng Wang, Michael S Triantafyllou, and George Em Karniadakis. Deep learning of vortex-induced vibrations. *Journal of Fluid Mechanics*, 861:119–137, 2019.
22. Arvind T Mohan and Datta V Gaitonde. A deep learning based approach to reduced order modeling for turbulent flow control using lstm neural networks. *arXiv preprint arXiv:1804.09269*, 2018.
23. Zheng Wang, Dunhui Xiao, Fangxin Fang, Rajesh Govindan, Christopher C Pain, and Yike Guo. Model identification of reduced order fluid dynamics systems using deep learning. *International Journal for Numerical Methods in Fluids*, 86(4):255–268, 2018.
24. Kjetil O Lye, Siddhartha Mishra, and Deep Ray. Deep learning observables in computational fluid dynamics. *Journal of Computational Physics*, page 109339, 2020.

25. J Nathan Kutz. Deep learning in fluid dynamics. *Journal of Fluid Mechanics*, 814:1–4, 2017.
26. Hengfeng Huang, Bowen Xiao, Huixin Xiong, Zeming Wu, Yadong Mu, and Huichao Song. Applications of deep learning to relativistic hydrodynamics. *Nuclear Physics A*, 982:927–930, 2019.
27. WC Wan, Guy Malamud, A Shimony, CA Di Stefano, MR Trantham, SR Klein, D Shvarts, CC Kuranz, and RP Drake. Observation of single-mode, kelvin-helmholtz instability in a supersonic flow. *Physical review letters*, 115(14):145001, 2015.
28. B Fryxell, CC Kuranz, RP Drake, MJ Grosskopf, A Budde, T Plewa, N Hearn, JF Hansen, AR Miles, and J Knauer. The possible effects of magnetic fields on laser experiments of rayleigh–taylor instabilities. *High Energy Density Physics*, 6(2):162–165, 2010.
29. Carolyn C Kuranz, H-S Park, Channing M Huntington, Aaron R Miles, Bruce A Remington, T Plewa, MR Trantham, HF Robey, Dov Shvarts, A Shimony, et al. How high energy fluxes may affect rayleigh–taylor instability growth in young supernova remnants. *Nature communications*, 9(1):1–6, 2018.
30. CM Huntington, A Shimony, M Trantham, CC Kuranz, D Shvarts, CA Di Stefano, FW Doss, RP Drake, KA Flippo, DH Kalantar, et al. Ablative stabilization of rayleigh-taylor instabilities resulting from a laser-driven radiative shock. *Physics of Plasmas*, 25(5):052118, 2018.
31. RayleAI Database. `https://github.com/scientific-computing-nrcn/RayleAI`. [Online].
32. Ygal Klein. Construction of a multidimensional parallel adaptive mesh refinement special relativistic hydrodynamics code for astrophysical applications. *Master's Thesis*, 2010.
33. WC Wan, Guy Malamud, A Shimony, CA Di Stefano, MR Trantham, SR Klein, D Shvarts, RP Drake, and CC Kuranz. Observation of dual-mode, kelvin-helmholtz instability vortex merger in a compressible flow. *Physics of Plasmas*, 24(5):055705, 2017.
34. Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.
35. Xi Chen, Yan Duan, Rein Houthooft, John Schulman, Ilya Sutskever, and Pieter Abbeel. Infogan: Interpretable representation learning by information maximizing generative adversarial nets. In *Advances in neural information processing systems*, pages 2172–2180, 2016.
36. Gan-Image-Similarity code repository. `https://github.com/marcbelmont/gan-image-similarity`. [Online].
37. Savvas A. Chatzichristofis Mathias Lux. Lire: Lucene image retrieval - an extensible java cbir library.
38. The Apache Lucene project. `https://lucene.apache.org`. [Online].
39. Dengsheng Zhang, Aylwin Wong, Maria Indrawan, and Guojun Lu. Content-based image retrieval using gabor texture features. *IEEE Transactions Pami*, 13, 2000.
40. Dr. Antony Selvadoss Thanamani K. Haridas. Well-organized content based image retrieval system in rgb color histogram, tamura texture and gabor feature.
41. Savvas A Chatzichristofis and Yiannis S Boutalis. Fcth: Fuzzy color and texture histogram-a low level feature for accurate image retrieval. In *2008 Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, pages 191–196. IEEE, 2008.

42. Stéphane Lathuilière, Pablo Mesejo, Xavier Alameda-Pineda, and Radu Horaud. A comprehensive analysis of deep regression. *IEEE transactions on pattern analysis and machine intelligence*, 2019.
43. Philipp Fischer, Alexey Dosovitskiy, and Thomas Brox. Image orientation estimation with convolutional networks. In *German Conference on Pattern Recognition*, pages 368–378. Springer, 2015.
44. Siddharth Mahendran, Haider Ali, and René Vidal. 3d pose regression using convolutional neural networks. In *Proceedings of the IEEE International Conference on Computer Vision Workshops*, pages 2174–2182, 2017.
45. Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: A system for large-scale machine learning. In *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, pages 265–283, 2016.
46. Roberto Brunelli. *Template matching techniques in computer vision: theory and practice*. John Wiley & Sons, 2009.
47. Raof R.A.A A.B A. Nazren N.M.Wafi M.B. Hisham, Shahrul Nizam Yaakob. Template matching using sum of squared difference and normalized cross correlation. In *2015 IEEE Student Conference on Research and Development (SCOReD)*. IEEE, 2015.
48. Wael Abd-Almageed Premkumar Natarajan Jiaxin Cheng, Yue Wu. Qatm: Quality-aware template matching for deep learning. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 2019.
49. SHI Xingjian, Zhourong Chen, Hao Wang, Dit-Yan Yeung, Wai-Kin Wong, and Wang-chun Woo. Convolutional lstm network: A machine learning approach for precipitation nowcasting. In *Advances in neural information processing systems*, pages 802–810, 2015.
50. Yunbo Wang, Mingsheng Long, Jianmin Wang, Zhifeng Gao, and S Yu Philip. Predrnn: Recurrent neural networks for predictive learning using spatiotemporal lstms. In *Advances in Neural Information Processing Systems*, pages 879–888, 2017.
51. Chris Ding and Xiaofeng He. K-means clustering via principal component analysis. In *Proceedings of the twenty-first international conference on Machine learning*, page 29, 2004.
52. Alain Hore and Djemel Ziou. Image quality metrics: Psnr vs. ssim. In *2010 20th International Conference on Pattern Recognition*, pages 2366–2369. IEEE, 2010.
53. Marco Tulio Ribeiro, Sameer Singh, and Carlos Guestrin. "why should i trust you?" explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2016.
54. A Shimony, WC Wan, SR Klein, CC Kuranz, RP Drake, D Shvarts, and G Malamud. Construction and validation of a statistical model for the nonlinear kelvin-helmholtz instability under compressible, multimode conditions. *Physics of Plasmas*, 25(12):122112, 2018.
55. Chuanqi Tan, Fuchun Sun, Tao Kong, Wenchang Zhang, Chao Yang, and Chunfang Liu. A survey on deep transfer learning. In *International conference on artificial neural networks*, pages 270–279. Springer, 2018.
56. NegevHPC Project. `http://www.negevhpc.com`. [Online].