Computational sensing, understanding, and reasoning: an artificial intelligence approach to physics-informed world modeling

Beatriz Moya[1*], Alberto Badias[2], David Gonzalez[3], Francisco Chinesta[1,4], Elias Cueto[3]

[1*]CNRS@CREATE LTD., CNRS, 1 CREATE Way, Singapore, 138602, Singapore.
[2]ETSIAE Polytechnic University of Madrid, C. de José Gutiérrez Abascal, 2, Madrid, 28006, Spain.
[3]ESI Group Chair, Aragon Institute of Engineering Research, Universidad de Zaragoza, Maria de Luna, s.n., Zaragoza, 50018, Spain.
[4]ESI Group Chair, PIMM Lab, ENSAM Institute of Technology, 151 Bvd. de l'Hôpital, Paris, 75013, France.

*Corresponding author(s). E-mail(s): beatriz.moya@cnrsatcreate.sg;
Contributing authors: alberto.badias@upm.es; gonzal@unizar.es; francisco.chinesta@ensam.eu; ecueto@unizar.es;

Abstract

This work offers a discussion on how computational mechanics and physics-informed machine learning can be integrated in the process of sensing, understanding, and reasoning of physical phenomena. A foundation in physics can leverage interpretability, data efficiency, and generalization of the sought models for the dynamics of complex physical systems. Consequently, this synergy results in promising approaches to develop world models that are capable for performing accurate and reliable simulations (reasoning) in low-data regimes. Among the possible alternative formulations, we highlight how thermodynamics offers a general framework to construct inductive biases, demonstrating its potential in applications where physics-consistent predictions are essential.

Keywords: Physics perception, Inductive biases, Thermodynamics machine learning, Augmented intelligence, Hybrid twins

1 Introduction

The recent appearance of chat bots based on the employ of Large Language Models (LLMs) is having a profound impact in the way citizens see artificial intelligence and its capabilities [1]. However, these systems exhibit severe limitations in aspects related to the construction of world models, despite some claims on the contrary [2]. Particularly true with regard to the construction of physical world models, there is now a consensus that, for an artificial intelligence to be able to understand the surrounding world, it is necessary to

- explain and understand what we see,
- imagine things we could see but haven not yet,
- problem solving and planning actions,
- build new models as we learn more about the world,

(cf. [3] and references therein). Cognition does not only consist in merely observing the world, but also in actively creating our perception of the surrounding reality. Our sense of reality is affected by our beliefs and intentions and, simultaneously, is limited by our knowledge and understanding of the world [4] [5]. The ability to sense and understand the physics of the environment is known as physical scene understanding. Through computer simulation, this discipline aims thus to provide machines with the basic capabilities of physical understanding [6]: (1) interpretation of the environment, (2) physics prediction, and (3) learning of newly observed phenomena.

The first task ("understand" what the machine sees, through computer vision, of course) is often referred to as machine perception, which is one of the main pillars of human cognition [3] [7]. To this end, the artificial intelligence must be able to construct a model during runtime—or, at least, to be able to update an existing one, as stated in the fourth bullet above—based on the data coming from its exploration of the surroundings.

The second bullet is related to prognosis: forecast what the future will look like, "reasoning". The interest on the development of learned simulators to this end is growing [8] [9] [10]. As a result, computer simulation may be the way to develop next-generation intelligent and autonomous systems capable of operating in any unknown and complex real-world scenario [11] [12].

In this context, it would be impossible for a person to cope with all the information they consume without using some type of knowledge bypass that favors making decisions, the growth of personal knowledge, and an adequate selection of what will be remembered [13].

This issue also applies to artificial intelligence. An excessive amount of information leads to the curse of big data, where the treatment of information becomes challenging and overwhelming. Without bias shortcuts, learning is deeply conditioned by data quality issues, and interpretation can be impractical [14].

In this work, we will disclose human perception and its approximation with artificial intelligence, which aims to replicate human cognition. We will analyze how physics-informed machine learning leads to a more efficient understanding of the physics of reality. In addition, we will expose the benefits of the development of augmented perception systems where additional hidden physical quantities of interest can be streamed [15] [16].

The possible applications of this type of artificial intelligence exceeds the possibility of developing embodied intelligences able to cope with the physical world. From an industrial point of view, for instance, the possibility of developing cognitive digital twins is of utmost interest [17] [18]. Beyond artificial intelligence, the interest of augmented intelligence is out of question [19].

2 Background

Within human cognition processes, physical scene understanding consists in performing predictions of real world dynamics. Based on our previously acquired knowledge, both forces acting upon objects and their consequences can be estimated. Theories on cognitive learning attribute this skill to experience [20]. Human learning is conceived then as a continuous on-the-loop process that gathers data obtained from experimentation. Based on this statement, several studies in computational cognitive science show proofs of the correlations between computational and machine learning processes and how people approximate physics principles [21] [22] [23].

The physics perception cycle is presented in Fig.1. There is a connection between cognition and reality through our senses, where data is perceived and interpreted. With the inferred knowledge about the state of the scenario, predictions are done with our beliefs of physics. In this step, one can add a learning bias to constraint the beliefs of the learnt physics, e.g., the existence of gravity. As predictions are made, the loop corrects and enriches by comparing the results with the ground truth to improve future predictions.

Physics perception can be split into several pieces that together create the desired physics engine. Knowledge inference puts emphasis in the data acquisition phase and interpretation of data [24]. Object recognition and tracking [25] [26], inference of physical states from images [27], or property estimation [28] [29] are some of the development paths of these group of techniques.

With this information, the system enters into the prediction stage [30][31]. This has been one of the most prolific fields since the rise of deep learning. In a range of methods that go from purely data-driven to completely knowledge-driven approximations, several solutions have aroused trying to replicate human physics abstraction [32][9] [33] [34]. Currently, one of the main challenges is to perform enrichment and correction of physics when there is a deviation from reality due to changes in the physics context [35] [18]. These systems must be implemented in active independent machine systems for action recognition and planning, human interaction [36] [37], and reinforcement learning [38] [39].

This review offers a vision mainly focused on one of the main targets of this field: the abstraction and rich understating of physics for simulation and correction of the forecast of the dynamical process. This work is structured as follows. In a general-to-specific composition, Section 3 provides an overview on the most recent developments and approaches in data interpretation and motion prediction. Section 4 dives into the

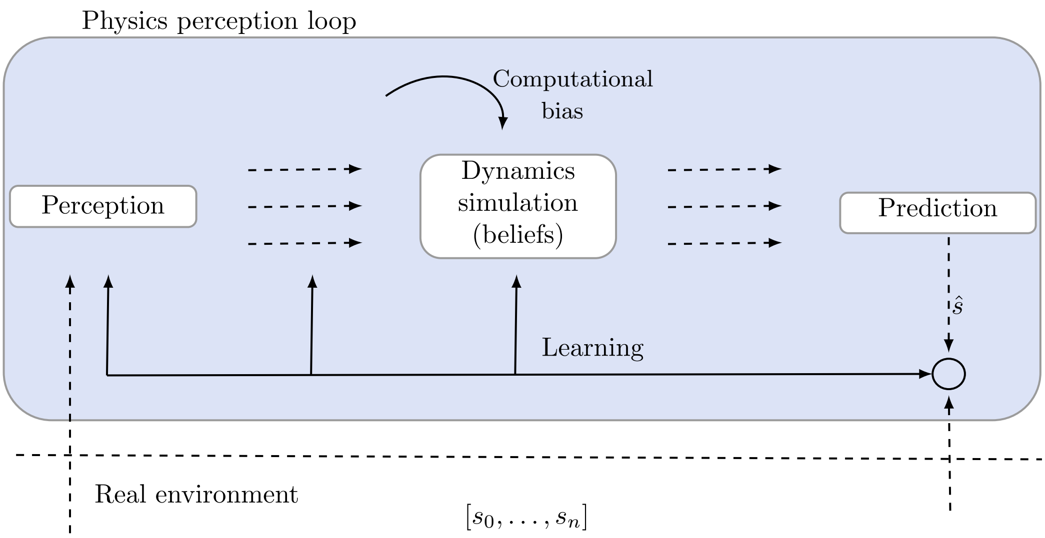

Fig. 1 Physics understanding loop for machine systems. Real measurements $s_0, \dots, s_n$ are the information available to perform both predictions and enrichment of the physics engine. The data is perceived and interpreted to enter a simulation loop constrained by some imposed inductive biases. Then, the prediction, and (possibly) some additional hidden variables are displayed. The result is compared with measurable quantities for correction.

use of thermodynamics-based inductive biases through the stages of physics perception. Section 5 analyses next challenges in the field and, finally, Section 6 offers a summary of the main conclusions that arise from this review.

3 Simulation as a tool for physics perception

We consider one or more target objects, which are tracked in a series of RGB-D images. While rigid objects can be reconstructed very accurately by using a moving RGB camera, deformable objects render the problem ill-posed, since these cameras record only one snapshot of the deformed object per frame. Therefore, the interest of using RGB-D cameras, able to provide with two different perspectives of the object at each time instant—very much like the human eyes work—is of great interest [40]. The goal is, on one hand, to obtain a rich understanding of the dynamics of the object(s), which moves across the scene and interacts with the environment, and, on the other, to predict its future states.

Perception systems are to be trained on a sequence of these RGB-D images where the dynamical features of the object (physical and dynamical properties) must be estimated. To this end, the physics perception engine must synthesize the available information, and transform it to augmented information of the object that can be used for more elevated purposes.

Machine learning will be the tool for learning a map from the visual and sensored inputs to the outputs. In this work, we will focus on data-driven approaches, and especially those based on deep neural networks.

In this setting, physics-based simulations are the means for developing and testing physics perception systems. Simulation is understood as the approximation of dynamic systems to replicate the behavior of both machine and environment in time, which results in a hybrid twin of reality. Black-box (physics-free) simulation is hence an unreliable approach for decision making: only trustworthy and interpretable results must be used, and this is achieved only if the simulation relies on physics. Nonetheless, physics-based simulation can be employed in several manners depending on the objective of the work.

3.1 Interpretation of the scene

The connection between real and virtual environments is made through the understanding, interpretation, and reconstruction of the surroundings from data acquired from diverse types of equipment, i.e. sensors and cameras. With this information, machines can perform both credible and physically valid scene reconstructions with which to predict future interactions between the target object and the environment. Hence, inferring latent representations of physics from data is the first step for the development of scene understanding algorithms. One of the applications of computer vision is the distillation of such physical knowledge from images and sensors. This capacity enables the actual awareness and understanding of the present objects and their interaction [41]. Figure 2 shows the pinhole approximation to project 3D information to 2D images, which is used to recover depth information. Projection correlation is described by

$$p = A[R|t]P_w,$$

where P_w is the 3D point that is to be projected to the 2D image plane, A the matrix of intrinsic parameters of the camera, and R and t the rotation and translation matrices respectively to reference the position of the camera to an initial coordinate system in the 3D space. The intrinsic parameters A stablish a correlation between the camera coordinate system and the image coordinate system

$$p = AP_c = \begin{bmatrix} f_x & 0 & C_x \\ 0 & f_y & C_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}.$$

The camera is described by the focal lenght f_x and f_y, and the center of the coordinate system of the image plane C_x, C_y.

Images and depth reconstruction is the information usually used in perception systems. Works oriented towards property prediction, such as material analysis, learn the main features of objects from this information to propose problem-specific models [42] [43]. In [44], authors propose a Bayesian Optimization approach to learn the properties of liquids from videos. They perform common manipulation tasks, such as stirring and pouring, comparing the reality with simulation based on the predicted properties. Other works focus on the perception of position and velocity fields from

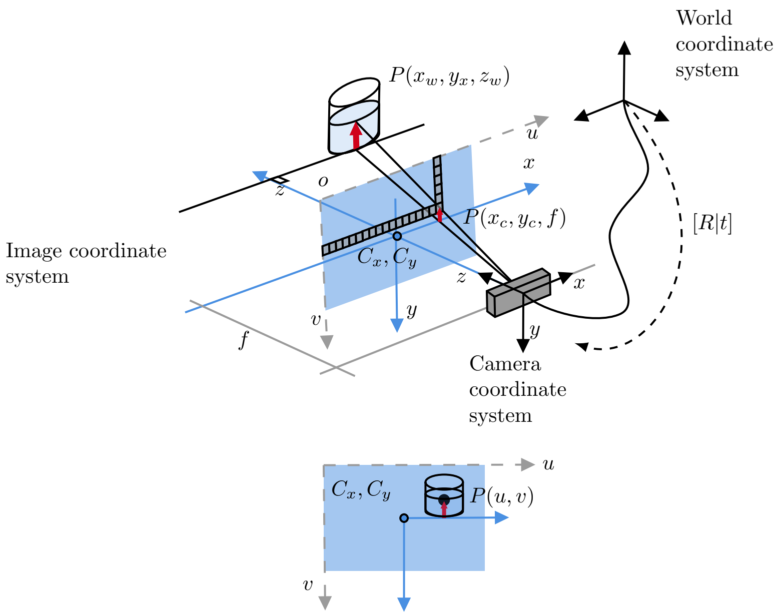

Fig. 2 Pinhole camera scheme. Projection from 3D to 2D.

images and sensors. [45] makes use of recurrent neural networks to link observed measurements with deformation states in soft robotics. Although fluids are the most difficult setting to be perceived, several techniques arise to reconstruct fluid volumes [46][47][26].

A different approach to learning state representations is to work with latent, or low-dimensional, manifolds, where physical attributes are embedded. From this manifold one can extract correlations to perform dynamics simulation. Several works combine analytical models with latent representations to perform an interpretation of sensed data and perform dynamical forecast [48] [49] [50] [51]. Kandukuri and coworkers [28] infer physical properties through self-supervised learning, and use transformers [52] for dynamics prediction.

One of the main difficulties that needs to be addressed in this type of approaches is the reconstruction of dynamical states from partial information. On the one hand we find reconstruction issues of information fields from sparse measurements. Purely data-driven deep learning is an appealing option for such task [53][54] [55]. However, most recent approaches rely on geometric or inductive biases to fulfill the compliance of the solution. Graph Neural Networks are a novel framework for these developments. In [56], the authors make use of graph neural networks to exploit the geometrical knowledge of the domain for information recovery. Authors in [57] employ a bayesian approach to account for the sparsity and noise of the measurements guided by the fluid governing equations, i.e. Navier Stokes. Gappy POD is also a popular approach for field reconstruction [58] [59][60] [61], especially in robot planning [62] [63] and fluid dynamics studies [64] [65]. Finally, super-resolution [66] is a growing field, specially in fluid dynamics [67][68] [69] and image reconstruction and enhancement [70] [71] [72], with potentially interesting applications for physics perception assemblies.

On the other hand, other works focus on the recovery of information of internal variables required to perform physically consistent approximations [47] [73] [74] [75] [53]. [76] and [77], for instance, include observed information (displacements and forces for the former, and fluid measurements for the latter) into a physically sound model to extract hidden, unmeasured, information. Also, works based on neural networks, such as Erichson et al. [78], show increasingly good performance for information recovery. In [79], the authors proposed a scheme based on Gated Recurrent Units (GRUs) to connect partial information of the free surface of fluids with the latent representation of the full dynamics. In a somehow similar spirit, Sun et al. reconstruct dynamical systems with partial information through the use of recurrent neural networks to recover missed information from the history analysis of data [80]. In spite of not being based on RNNs, Antonova et al. obtain the state information of a sequence of images by representing the deformed object's shape as a distribution embedding [81].

3.2 Computational reasoning about physical events: forecasted dynamics

Physics-free techniques look for motion correlations [82] [83] and feature extraction (position, distance and velocity fields) [84] [85] [86] to construct simulators from video sequences. Usually, the best performance approaches rely on Recurrent Neural Networks, such as LSTM [87] [88] [89] [90]. A step forward is the inclusion of human biases. For instance, [91] succeeds at learning general physics principles with a recurrent state space model (RSSM) [92] by imposing *development as stochastic optimization* and *development as complexity* constraints, which are hypotheses of human cognition learning. In spite of its good results, it does not implicitly fulfill any physics principle. In contrast, the inclusion of physics priors has demonstrated to foster the general comprehension and interpretation of any event [93] [94] [95].

There are several phenomena that are particularly complex in the context of physics perception. This is, for instance, the case of contact [96] [97] [98]. Contact is widely present in physics perception since objects are in continuous interaction with the surrounding scene. In Strecke et al., the authors propose an approach to face contact based on signed distance functions for differentiable physics simulation [99]. More recently, Hernandez et al. employ a Port-metriplectic formalism, a thermodynamics consistent framework for open systems that takes into account the possibility of analysing different objects in interaction [100].

Differentiable formulations are a popular approach in motion learning and planning [28]. This technique consists in incorporating physical equations of the phenomenon under study, with a chosen discretization, in the Deep Learning algorithm [101] [102] [103]. In other words, the networks are trained together with a numerical simulator instead of incorporating a residual physics loss. Li et al. [34] use particle-based interaction networks to learn the physics engine at a particle level of description, and use graph theory to propagate the interactions among the particles. From the same authors, [51] performs Visual Reasoning with Differentiable Physics (VRDP), which consists in learning together visual interpretation and reasoning, language concepts and properties of the objects involved in the simulations to perform the predictions. In this case, the simulator is phrased for rigid bodies to model, among others, collision.

In contrast, physics-informed machine learning directly includes general knowledge about the physics leading the evolution of a system to guide the learning process [104] [105]. This is done by the imposition of physics priors, or biases, to penalize those results that deviate from the fulfillment of the physics, in the case that the governing equations are known—something that is not always the case—[14]. As a result, the network learns more general, interpretable, consistent, and accurate representations of the dynamics for intuitive physics [106] [48] [107] [108]. These constraints can be fulfilled in hard or soft manners [109]. Hard constraints are restrictions that the method *must* fulfill [110] [111] [112] [113]. In [114], for instance, authors impose momentum conservation for learning fluid dynamics. This is done by a modification of the kernels of convolutional neural networks through a correlation between the internal forces and the residual positions.

However, soft constraints are the most commonly chosen approach. Contrary to hard constraints, soft constraints include the fulfillment of a restriction as a weighted penalty in the loss function [115]. One popular case is the use of Physics Informed neural networks (PINNs) [116], which include an additional term in the loss from the evaluation of the fulfilment of the PDE that describes the phenomenon under study and the interaction across environments [117] [118]. The latest trends, such as Huang et al. [119], combine Graph Neural Networks and ODEs to work over unstructured domains and partial observations of the physical domain for physics perception, such as spring motion and walking forecast.

In spite of its popularity in the computational community, PINNs are not an extended technique in physics perception given that, usually, the governing PDE is considered unknown. Hence, a different framework is required to perform the approximation to dynamical systems with neural networks by the imposition of specific structures that ensure the fulfillment of the physics laws of reality [120]. If the problem formulation is onservative, and therefore can be approximated as a Hamiltonian system, this formalism can be applied as a restriction to guide the training process [121][122] [123] [124].

Nevertheless, for dissipative, non-reversible systems, Hamiltonian formulations must be modified accordingly. Thermodynamic frameworks appear in this circumstance to provide an interpretable and general solution [125] [126] [127] [128]. Masi et al. [129] impose the compliance of the first and second laws of thermodynamics presented as PDEs through automatic differentiation. However, this method is based on the knowledge of internal variables required for the description of the thermodynamic state of the material, as is also the case in [130]. By using the Generalized Osanger Principle, Yu et al.[131] developed a learning scheme that describes the evolution of a system in terms of its energy and entropy potentials. The Generalized Osanger Principle is optimized for data-driven inference using a Runge-Kutta scheme. Huang et al present a similar approach [132] which, instead, performs optimization by minimizing the Rayleighian [133], an alternative formulation of this principle.

The two last mentioned approaches are basically based on the learning of two potentials—instead of one, the Hamiltonian—that account for the reversible and irreversible part of the dynamics. This takes the form of a metriplectic formulation since both metric and symplectic parts are present [134]. With the same spirit, the GENERIC formalism (General Equation for Non-Equilibrium Reversible-Irreversible Coupling) posses a metriplectic structure accompanied by the so-called degeneracy conditions to fulfill the first and second principles of thermodynamics [135]. Several successful works have been developed under the umbrella of GENERIC given its flexibility to be applied to either dissipative or conservative phenomena [136] [137] [10] [138], also for physics perception [79] and augmented reality [139]. Hernandez et al. combined the use of GENERIC as an inductive bias, and graph networks as geometrical biases, demonstrating the versatility of this approach [140].

In Section 4 we dive deep into the GENERIC formalism, given its importance in the development of an inductive bias that ensures thermodynamic correctness in the reasoning process.

3.3 Continuous learning

In the introduction above, we have highlighted that a crucial aspect for a system to be able to understand the physics of the surrounding environment is to be able to refine the models as it learns from observation. Hence, simulators and perception systems require a fine-tuning step to adapt to evolving environments and previously unseen situations [141] [142] [108] [143] [35]. TuneNet [144] performs parameter update with one-shot observations in perception systems. This is done by learning a network on pairs of simulations with known parameter differences. Lutter et al. [145] employ Deep Lagrangian Networks (DeLaN) to learn on-the-go the underlying dynamics with symmetry preserving networks, showing better performance than black-box based domain adaptation.

Techniques globally coined as Imitation Learning learn from observed demonstrations [146]. Generative Adversarial Imitation Learning, GAIL, employs generative adversarial networks (GANs) to train a discriminator that can discriminate between previously learnt and new trajectories of the system in phase space, to enrich the simulation loop [147]. In the same spirit, Chen et al. include differentiable physics simulators to update the policies to substitute the GAN discriminator [148].

As an obvious application of this rationale, hybrid twins (also known as cognitive twins) combine previous knowledge with new data [17] [149]. The original model may not account for complex effects, subtle dynamics, o environment changes, and the approximation is to be enriched with new data [106] [150]. In [18], the authors presented a hybrid twin that performs the correction of the strain prediction of a viscous-hyperelastic beam from sparse observations.

Transfer learning is an ad-hoc technique to perform adaptation [151]. If one approximation has properly learned generic features of the phenomenon, the new simulator just has to perform a fine-tuning of a few parameters, enabling correction with low-data regimes and partial measurements. This technique has been used in manifold adaptation for model order reduction [152], and dynamics correction [153] [154] [155] [156].

4 Thermodynamics-informed perception of physical phenomena

It goes without saying that the laws of thermodynamics apply to any known physical phenomenon, since they provide fundamental principles and laws governing the behavior of macroscopic systems at the most general level of description, that of invariant quantities. As a result, they constitute a very convenient framework to guide the scene understanding process from a computational perspective.

Considering the three steps of physical scene understanding, we propose a framework based on thermodynamics to reconstruct visual information, perform learned simulations [8], and achieve active adaptation and correction. We chose as a proof of concept to physics perception the sloshing dynamics of a general fluid, given its nonlinearity and complexity, along with its practical interest for embodied systems manipulating fluids. The proposed system will carry out perception in a continuous loop, integrating perception, reasoning and correction. To this end, we consider a glass of glycerine as a starting point of the system proposed (viscosity $\mu = 0.950\ Ns/m^2$ and density $\rho = 1261\ kg/m^3$).

Our perception system consist of the following steps:

1. The process starts by learning a mapping from partial visual measurements to a representation of the full dynamical state of the fluid. Note that the RGB-D camera is able to see the free surface of fluid only, it does not register any information concerning velocities (other than the one of the glass), stresses, etc. These measurements will therefore be correlated to a low dimensional manifold where the information is embedded. In addition, this reduced order representation of the dynamics will enable realtime simulation.
2. Then, a learned simulator is trained: instead of simply assimilate the data to a known governing equation, we propose a neural network that learns the operators of our thermodynamics framework at each time step.
3. Finally, a correction algorithm is proposed to build the hybrid twin that corrects, if necessary, the learnt model. This completes the perception loop.

4.1 Reconstruction of the perceived dynamics

The proposed machine learning strategy is based on the fulfillment of the principles of the thermodynamics. For this reason, we must be able to characterize the thermodynamic state of the system under study by means of an appropriate selection of state variables. However, the main drawback of this description is the impossibility of measuring some of these variables, i.e. internal (phenomenological) variables, by using commodity computer vision means.

In the absence of necessary experimental information, the system is first trained with synthetic data. To this end, the sloshing fluids are approximated with Smooth Particle Hydrodynamics [157], which is a convenient framework for fluid simulation in free surface settings. In addition, this level of description is shared with human thinking [158]. Bates et al. performed a study to analize human liquid reasoning to build a computational analogy. This work resulted in demonstrating that participants' predictions matched those performed with computational models built upon particle interactions in a coarse grained level. Thus, the use of this type of description is justifyed. . To fully describe the energy of the system we require information about the position $\boldsymbol{p}$, velocity $\boldsymbol{v}$, internal energy e, and stresses $\boldsymbol{\tau}$ for each particle at each time step to characterize the dynamical evolution of the system [159]. A database is prepared with computational simulations where we have access to the whole set of state variables required for a thermodynamics-informed description of the motion of the fluid. Then, this state will be reconstructed from partial measurements of the free surface.

We follow an approach based on latent physical representations, in which we learn a reduced representation of the sloshing dynamics. We make use of model order reduction techniques based on deep-learning for efficient compression. As a result, we not only obtain a latent representation of the phenomenon, but also foster the inductive learning process in this space and achieve realtime performance of the calculations [160].

We choose a stacked sparse autoencoder (AE) that learns a compressed representation of each state variable individually. The sparsity term will work similarly to the energy criteria in Proper Orthogonal Decompositions: it will penalize the activation of layers in the bottleneck to only have those that include most of the information, providing additional compression. The learning process is thus guided by the loss:

$$\mathcal{L}_{AE} = \frac{1}{N}\sum_{i=1}^{N}(\boldsymbol{s}_i - \hat{\boldsymbol{s}}_i)^2 + \lambda_{\text{reg}}\sum_{i=1}^{N}|\boldsymbol{x}_i|\,,$$

which consists in the evaluation of the Mean Squared Error of the reconstruction of the full system and a L_1-norm to account for the sparsity (this is equivalent to the well-known Occam's razor or parsimonious approximation [161]). The second term in the loss is weighted to give more importance to the reconstruction error.

Once the dynamical features are embedded on a low dimensional manifold, we propose extracting the dynamic features related to the evolution of the free surface by studying the history of its motion. To achieve this, we suggest utilizing a recurrent neural network (RNN) that correlates a reduced-order manifold where the features of the slosh are embedded with partial measurements of the free surface.

RNNs learn from sequences of information rather than individual snapshots. Traditional RNNs often face challenges such as vanishing and exploding gradients when applied to complex problems [162]. In this case, we propose the use of GRUs [163], which are more complex architectures that incorporate additional information flows to preserve long-term correlations and prevent the network from learning short-term information. The fundamental feature of GRUs is the hidden state ($\boldsymbol{h}_n$), which represents a summary of the features identified in previous sequences of the batch of snapshots. Provided that ordinary, fully-connected neural networks functions read as:

$$\boldsymbol{x}_n = \sigma(\boldsymbol{W}_n \times \boldsymbol{x}_{n-1} + \boldsymbol{b}_n),$$

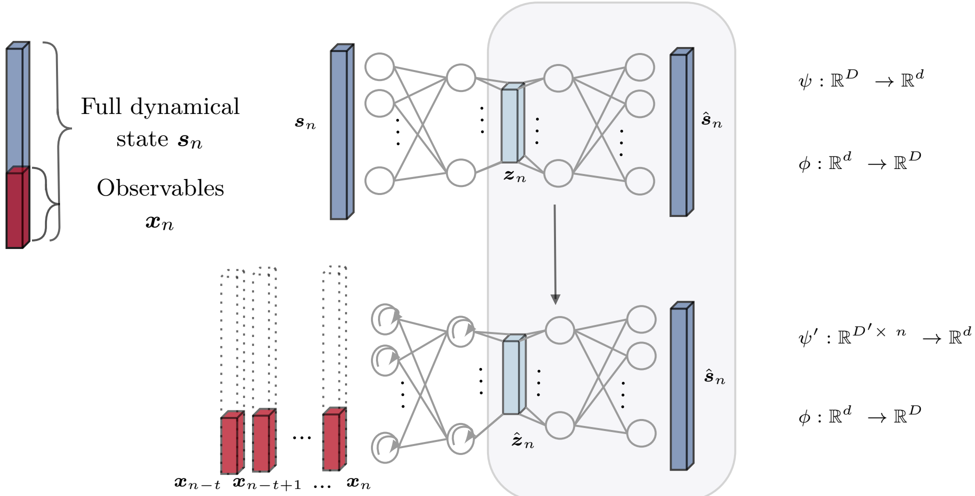

Fig. 3 Sketch of the proposed perception system. The encoder first learns a mapping ψ from the high-resolution synthetic data to a low dimensional manifold, and reconstructs the full state by applying an approximate inverse ϕ. Then, during runtime, since the camera is able to obtain only a portion of the data, highlighted in red, $\boldsymbol{x}_n$, the embedding ψ is substituted by a new one ψ', that operates on partial data only. This limitation is overcome by storing sequences of partial data $\boldsymbol{x}_{n-t}, \ldots, \boldsymbol{x}_n$. These feed a recurrent neural network that extracts information from the recent history of the fluid.

where x represents the input and output of the layer, W the weight, b the bias, and σ the activation function, the scheme of recurrent neural networks is expressed as follows:

$$\begin{aligned}
\boldsymbol{x}_n &= \sigma(\boldsymbol{W}_{\boldsymbol{x}_n} \times \boldsymbol{x}_{n-1} + \boldsymbol{U}_{\boldsymbol{x}_n} \times \boldsymbol{h}_{n-1} + \boldsymbol{b}_{x_n}),\\
\boldsymbol{r}_n &= \sigma(\boldsymbol{W}_{\boldsymbol{r}_n} \times \boldsymbol{x}_{n-1} + \boldsymbol{U}_{\boldsymbol{r}_n} \times \boldsymbol{h}_{n-1} + \boldsymbol{b}_{\boldsymbol{r}_n}),\\
\tilde{\boldsymbol{h}}_n &= \sigma(\boldsymbol{W}_{\boldsymbol{h}_n} \times \boldsymbol{x}_{n-1} + \boldsymbol{r}_n \times \boldsymbol{U}_{\boldsymbol{h}_n} \times \boldsymbol{h}_{n-1} + \boldsymbol{b}_{\boldsymbol{h}_n}),\\
\boldsymbol{h}_n &= \boldsymbol{x}_n \cdot \boldsymbol{h}_{n-1} + (1 - \boldsymbol{x}_n)\tilde{\boldsymbol{h}}_n.
\end{aligned}$$

The layer functions include the hidden state $\boldsymbol{h}_{n-1}$, and its corresponding weights $\boldsymbol{U}$, to calculate the output of the layer. The scheme includes a forget gate $\boldsymbol{r}_n$, which consists in the short-term, not important features, of the history. With the output of the layer and the reset information, the new hidden state is calculated to pass to the next layer.

We perform a from-many-to-one projection from a sequence of snapshots to a vector in the low dimensional manifold that compresses the dynamical information of the dynamical state we seek. In our case, the GRUs input comprises a sequence of vertical and horizontal positions of selected points on the free surface of the liquid. We track these points in real images from binary, black and white, frames. We train the mapping ψ' with a computational database of sloshing liquids. By tracking particles belonging to the free surface we interpolate a uniform point distribution to compare the free surface of different snapshots,

$$\mathcal{L}_{\text{GRU}} = \frac{1}{N_{\text{snap}}} \sum_{n=1}^{N_{\text{snap}}} (\boldsymbol{x}_n - \hat{\boldsymbol{x}}_n)^2.$$

The proposed architecture is therefore a sequence of GRU layers joined to a fully connected layer to achieve the desired projection. The network is trained with a mean squared error loss (MSE) that evaluates the accuracy of the mapping by comparing the predicted low-dimensional state and the ground truth.

4.2 Physics forecast with the learned simulators

The simulation framework proposed is also constrained by thermodynamic principles, not only to estimate dynamical features of the system, but to forecast future states in the interaction with the environment. The use of this physics constraint will trigger interpretability and generalization of the sloshing dynamics, preparing the simulation engine for future adaptation. This framework, however, required some internal state variables for its development. Hence, we propose to work in the low-dimensional manifold learnt from a full-dynamics dataset, which has been correlated to evaluations of the free-surface, to perform physically sound simulations. With this information it is possible to build the proposed thermodynamically-consistent framework.

Dissipative effects substantially dominate real-world systems. As indicated by the fluctuation dissipation theorem [164], here it is where entropy caused by the lack of knowledge arises. Then, a metriplectic framework can learn also the dissipation that underlies the data in the learning algorithm [165] to ensure the conservation of key variables (mass and momentum) and the thermodynamic feasibility of the system's evolution. The thermodynamic framework presented by GENERIC [166] was employed for describing the temporal evolution of a system with respect to the evolution of its energy and entropy. The GENERIC formulation is defined with the following structure:

$$\frac{\mathrm{d}\boldsymbol{z}}{\mathrm{d}t} = \boldsymbol{L}\nabla E + \boldsymbol{M}\nabla S.$$

Here, $\boldsymbol{z}$ represents the set of state variables used to describe the evolution of the dynamics. In the case of applying model order reduction, this vector will correspond to the one representing the reduced order state and we will denote the set of macroscopic variables by $\boldsymbol{s}$, as in Fig. 3. $E(\boldsymbol{z})$ and $S(\boldsymbol{z})$ are the energy and entropy of the system. $\boldsymbol{L}(\boldsymbol{z})$ and $\boldsymbol{M}(\boldsymbol{z})$ stand for the Poisson and friction matrices respectively. The structure-preserving scheme must ensure that $\boldsymbol{L}(\boldsymbol{z})$ is skew-symmetric and $\boldsymbol{M}(\boldsymbol{z})$ is symmetric positive semidefinite. In addition the *degeneracy conditions* must be fulfilled to ensure that the entropy S does not contribute to the conservation of energy, and E is not involved in the production of entropy:

$$\boldsymbol{L}\nabla S = \boldsymbol{M}\nabla E = \mathbf{0}.$$

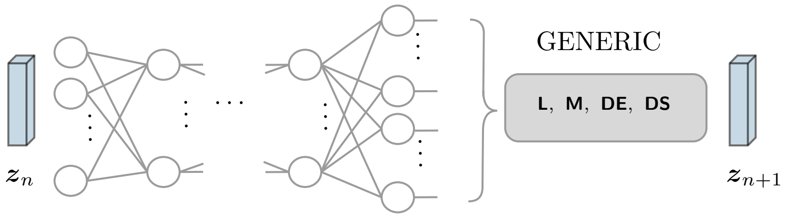

Fig. 4 Structure-preserving Neural Network. A set of fully connected layers learns the GENERIC operators as the output of the network to perform learned simulation through integration in time of the state of the system.

Structure-preserving neural networks (SPNN) embed GENERIC in a deep learning architecture to reveal the value of the main elements of equation [10] [138]. In this framework, the discretized gradients of energy and entropy are targets of the network optimization. For this purpose, a forward-Euler discretization of GENERIC is applied:

$$\boldsymbol{z}_{n+1} = \boldsymbol{z}_n + \Delta t \left(\mathbf{L}_n \mathbf{DE}_n + \mathbf{M}_n \mathbf{DS}_n \right),$$

constrained by the degeneracy conditions:

$$\mathbf{L}_n \mathbf{DS}_n = \mathbf{0}, \quad \mathbf{M}_n \mathbf{DE}_n = \mathbf{0}, \tag{1}$$

Although $\mathbf{L}$ and $\mathbf{M}$ are generally known in the literature, they are unknown in the low-dimensional latent manifold where we perform the simulation of the dynamics. Therefore, the network learns $\mathbf{L}, \mathbf{M}, \mathbf{DE}_n, \mathbf{DS}_n$ from data. This is represented in Fig. 4. The neural network is trained upon a weighed loss that includes an evaluation of the reconstruction of the dynamics $\mathcal{L}^{\text{mse}}_{\text{SPNN}}$ and the penalization imposed by the degeneracy conditions $\mathcal{L}^{\text{deg}}_{\text{SPNN}}$:

$$\mathcal{L}^{\text{mse}}_{\text{SPNN}} = \frac{1}{N_{\text{snap}}} \sum_{n=1}^{N_{\text{snap}}} (\boldsymbol{z}_{n+1} - \hat{\boldsymbol{z}}_{n+1})^2,$$

$$\mathcal{L}^{\text{deg}}_{\text{SPNN}} = \frac{1}{N_{\text{snap}}} \sum_{i=1}^{N_{\text{snap}}} (\mathbf{L}_n \mathbf{DS}_n)^2 + (\mathbf{M}_n \mathbf{DE}_n)^2,$$

$$\mathcal{L}_{\text{SPNN}} = \lambda^{\text{mse}}_{\text{SPNN}} \mathcal{L}^{\text{mse}}_{\text{SPNN}} + \mathcal{L}^{\text{deg}}_{\text{SPNN}}.$$

4.3 Transfer learning for active learning

The initial model of the fluid slosh is to be trained on one single liquid. Specifically, it has been designed to correctly simulate the physics of glycerine based on a set of high resolution synthetic data. From this knowledge, the physics perception system must be able to perceive a new liquid and still infer this new behavior so as to perform accurate predictions. The dynamics of the new liquid must be recovered and learned from the measurements of the free surface $\boldsymbol{x}_n \subset \boldsymbol{s}_n$ performed by the camera. Here, $\boldsymbol{s}_n \in \mathcal{S}$ represents the set of state variables that describe the full dynamics, and $\mathcal{S}$ its state space.

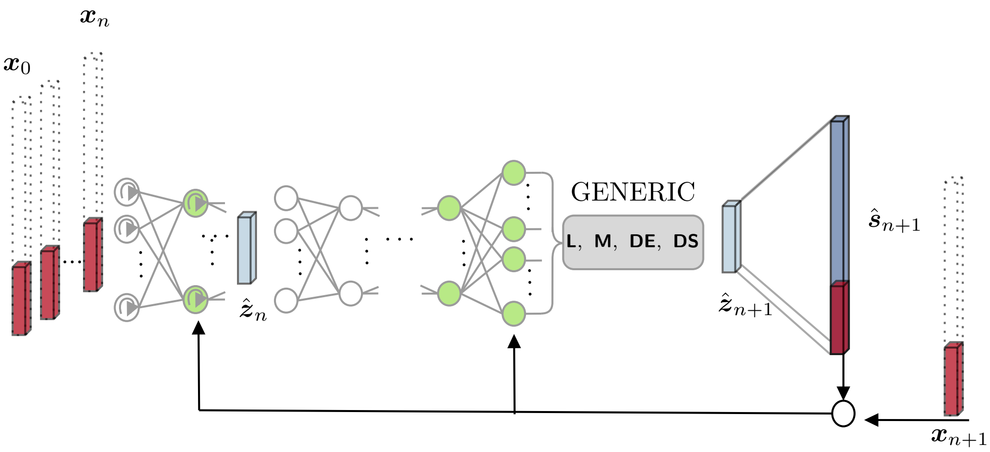

Fig. 5 Hybrid twin algorithm for the adaptation to new physics. The network as a whole adapts to the new data by fine-tuning only part of the layers, highlighted in green. The result of the integration algorithm is compared with the ground truth liquid by evaluating the reconstruction of the free surface in time. As a result the twin adapts when it detects any deviation between the prediction and the ground truth.

The proposed thermodynamics-informed correction algorithm achieves an adaptation of the hybrid twin of a liquid (which corresponds in this case to glycerine) to fit the new perceived physics. For this purpose, a technique based on transfer learning is applied. The premise of this approach is that the prediction algorithm, thanks to the thermodynamic biases, has learned some basic patterns of the slosh that will be shared by any liquid we study. Given the scarce data provided by the camera for learning the new liquid, this knowledge is to be preserved. This is achieved using transfer learning: the network already prepared will be retrained in a few selected layers to use the previously acquired knowledge of the slosh and only perform fine-tuning to the new liquid due to the different perceived properties. The correction is performed with the Mean Squared Error, that evaluates the accuracy of the reconstruction of the free surface. This loss is complemented, again, by the fulfillment of the degeneracy conditions established in the GENERIC formalism, to ensure that the changes belong to a thermodynamics admisible regime:

$$\mathcal{L}_{\text{correction}} = \lambda \frac{1}{N_{\text{snap}}} \sum_{n=1}^{N_{\text{snap}}} (\boldsymbol{x}_{n+1} - \hat{\boldsymbol{x}}_{n+1})^2 + \frac{1}{N} \sum_{N} \left(\|\mathsf{L}_n \mathsf{DE}_n\|^2 + \|\mathsf{M}_n \mathsf{DS}_n\|^2 \right).$$

4.4 Results

The source model has been trained according to the parameters of Table 1 for learning computational glycerine and assembling the main parts of the algorithm. This is the source model of the hybrid twin, the starting point to learn and adapt to unknown liquids.

Algorithm 1 Hybrid twin

Require: Free surface $\mathbf{x}_n \in X$, full state $\mathbf{s}_n \in S$, source model π_θ
Ensure: Next state $\mathbf{s}_{n+1}$
  for Iterations **do**
    for $n = 1$ to N sequences **do**
      Encode $\overline{\mathbf{z}}_n \leftarrow \phi_{\text{GRU}}(\mathbf{x}_n)$;
      Compute SPNN $a_n \leftarrow \pi_\theta(\overline{\mathbf{z}}_n)$, with $a_n = [\mathsf{L}_n, \mathsf{M}_n, \mathsf{DE}_n, \mathsf{DS}_n]$;
      Determine next integration step $\widehat{\mathbf{z}}_{n+1} \leftarrow \Delta t(\mathsf{L}_n\mathsf{DE}_n + \mathsf{M}_n\mathsf{DS}_n) + \overline{\mathbf{z}}_n$;
      Decode $\widehat{\mathbf{s}}_{n+1} \leftarrow \psi(\widehat{\mathbf{z}}_{n+1})$;
    end for
    Extract free surface of $\widehat{\mathbf{s}}_{n+1}$;
    Compute loss $\mathcal{L}_{\text{correction}}$;
    Update π_θ;
  end for
    return Optimized hybrid twin π_θ

	Learning rate	Weight decay	# Hidden layers	Size hidden layers	Input size	Output Size	Weight	Epochs
AE $\boldsymbol{q}$	1e-4	1e-6	2	120	6402	20	2000	10000
AE $\boldsymbol{v}$	1e-4	1e-5	4	200	6402	20	2000	10000
AE e	1e-4	1e-5	3	40	2134	10	2000	10000
AE $\boldsymbol{\sigma}$	1e-4	1e-5	3	200	2134	20	2000	10000
AE $\boldsymbol{\tau}$	1e-3	1e-6	3	200	6402	20	2000	10000
SPNN	1e-3	1e-5	13	195	13	195	1000	5000
GRU Recurrent layers	1e-3	1e-5	3	26	16 × 42	13	-	10000
GRU Fully connected layers			1	13	26			

Table 1 Training parameters of the source model

The correction is performed by comparing the predictions with new video recordings. Experiments have been made with four different liquids: water, beer, gazpacho (traditional Spanish tomato soup) and honey. Water and beer are less viscous than glycerine. Hence, with them the algorithm proves its efficacy to adapt to liquids that are out of the motion range of the original twin. In contrast, gazpacho and honey and more viscous (some mild non-Newtonian behavior could be expected for gazpacho, which is a suspension). As a result, the algorithm must learn that higher forces are required to start the motion of these liquids. These recordings are made with a stereo camera to have the pixel depth map of the points of the image. Fig. 6 illustrates how data about the position of the free surface have been acquired. The RGBD frames are converted to binary, black and white, images, where it is possible to track the free surface of the liquid due to the color gradient that appears in this transition.

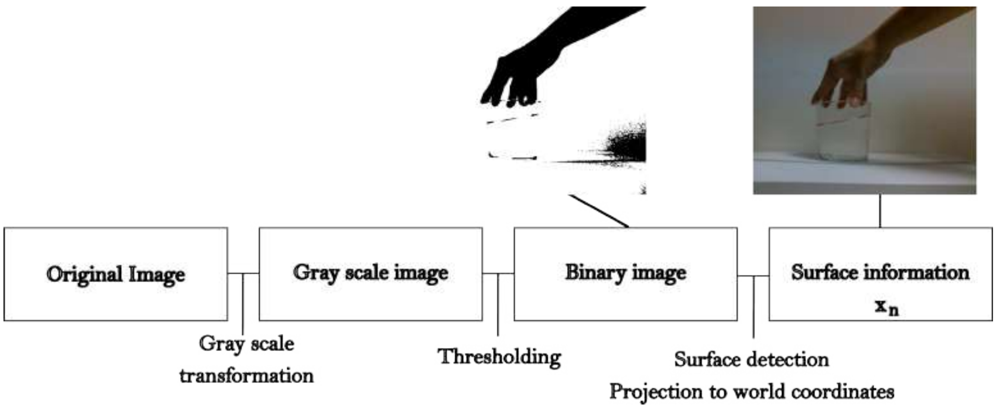

Fig. 6 Data processing transformations. The starting point is a collection of RGB-D images, which are first converted to gray scale, and then binarized. Frames in the black and white are used to detect and track the free surface from the gradient that appears in the phase transition. Given that the depth map of the image is provided, the projection to cartesian coordinates can be done.

		Activated layers for adaptation
GRU	Recurrent layers	1 last layer
	Fully connected layers	1 last layer
SPNN		5 last layers

Table 2 Layers activated for correction in the hybrid twin

When a systematic, biased, deviation of predictions from the observed reality is detected, only a few layers of the networks that are part of the algorithm activate for the fine-tuning to the new dynamics perceived. Table 2 details the layers activated for correction for each of the networks.

Two recordings of each liquid were made: the first for training and correction, and the second for testing the efficacy of the adaptation. Each recording is about 8 seconds long. This fact enables to evaluate in the accuracy of hybrid twin in the long term. As shown in Fig. 7, the relative error of the reconstruction of the motion of each liquid decreases after correction. We have compared the predicted free surface, that describes the fluid motion, with the ground truth from videos. The probability density functions show that the relative error greatly decreases after correction for all the liquids. In addition, there is less dispersion in the error values after adaptation. This is also represented in the boxplots. The mean error is lower for the predictions after adaptation. These conclusions are also applicable for the test recording. It is worth noting that this is new data that is uncorrelated with the training recordings.

A comparison of the reconstruction of the free surface before and after correction is presented in Fig. 8. Maximum slosh height is one of the targets in intelligent robotics control to perform tasks efficiently (i.e., to avoid liquid spillage). Water initially shows greater deviation than honey since it is less viscous than glycerine and the slosh height

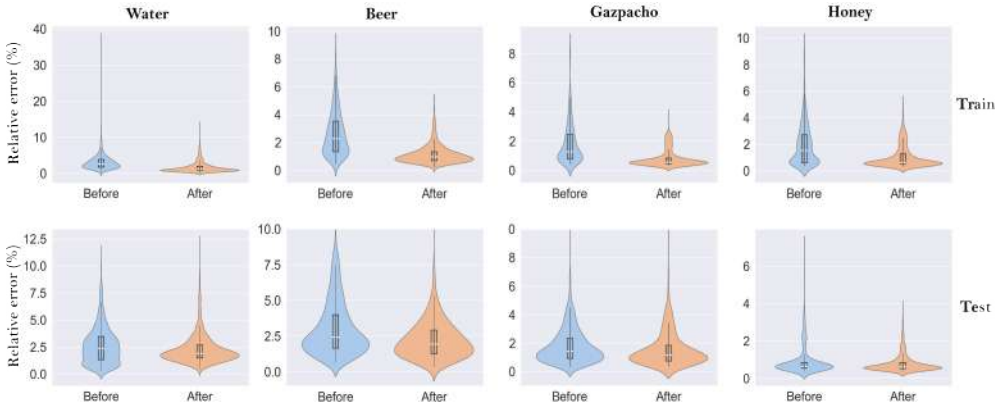

Fig. 7 Probability density function of relative errors combined with boxplots of the error distributions.

is greater than that of the source liquid. After correction, absolute errors decrease from up to 8 mm to 1 mm. Also, amplitude and frequency of the slosh improves.

The right hand side of Fig. 8 shows the final output of the hybrid twin. The physics scene understating system does not only provide the reconstruction of the whole fluid volume and its motion, but also some (possibly phenomenological, internal) variables such as stresses, that were reconstructed thanks to the thermodynamics framework. This set of variables cannot be evaluated with real liquids since they cannot be measured with a standard RGBD camera. However, the algorithm ensured that the adaptation and the reconstruction of these variables is physically consistent.

5 Next challenges in physics perception

5.1 Knowledge acquisition

Physics perception imitates human physics learning and understanding. However, the correlations and patterns acquired are more accessible in machine systems. Performing interpretation and appropriate analysis could lead to a deeper understating, or knowledge discovery, of complex physics. That is the case, for instance, of non-Newtonian fluids [167]. Symbolic descriptions obtained through symbolic learning could not only be used as a validation tool for the simulation algorithm learnt. It could unveil new characteristics in real systems for smart data processing [168].

5.2 Physics-informed computer vision

In spite of the good performance of computer vision algorithms due to the use of large databases, the results usually lack interpretability and physical meaning [169]. Conversely, the use of physics-informed machine learning enables to build robust and functional computer vision systems based in fundamental laws [170] [171]. These physics priors can be applied in the acquisition and pre-processing of data, model design and training, and inference. A detailed review of these applications can be found

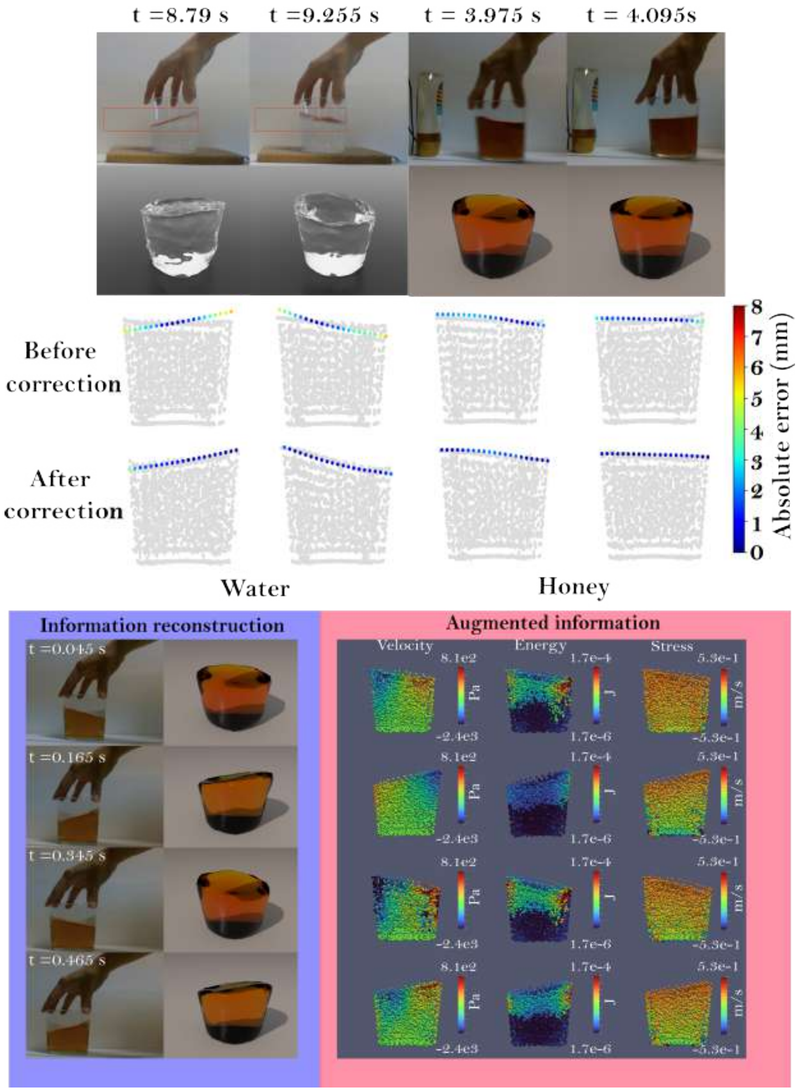

Fig. 8 Results of fluid motion prediction. Top, a comparison of liquid motion prediction with ground truth before and after correction is shown. Bottom, the hybrid twin outputs additional information related to the learned fluid slosh dynamics on top of the reconstruction of the fluid volume and its motion.

in [172] . Their application in perception and physical knowledge discovery (velocity fields, material properties etc.) might be of great impact to complement scene understanding.

5.3 General AI

In spite of the advances in physical scene understating, the developed applications are categorized into narrow, or specialized, AI. General AI (AGI) is considered the next step to be achieved in artificial intelligence developments. Here, machine systems are supposed to imitate human cognition independently and show a wide spectrum of high-level intellectual skills. This paradigm of learning entails a great abstraction, which a capacity of learning that current machine systems do not posses. [173]. Reinforcement Learning is seen as a possible path to create intelligent systems that can independently enrich themselves to reach evolving goals [174]. Regardless of whether it has real potential to deal with the great scope of the AGI proposal, it is certain that its combination with the physically sound stated techniques could lead to a more efficient learning of the desired policies.

6 Conclusion

Physical scene understanding imitates human cognition skills that make possible to acquire general concepts about physics, such as properties, interactions, and behaviors of objects to predict dynamics and interact with reality. As a key aspect of machine intelligence, physical scene understating predominantly involves systems that learn from data. Physics is thus an essential asset to acquired high quality knowledge in machine systems. Physics-infused machine learning is present in data interpretation, learning, and adaptation. Data interpretation is strongly influenced by physics to project data to new representations that represent features of the dynamics and reconstruct fields of data and, more specifically, internal variables required for developing interpretable simulation. Also, inductive biases are widely present in dynamics learning for approximating dynamics prediction, in the development or adaptation phases, from data. Since mathematical descriptions such as PDEs are sometimes not considered, inaccurate, or unknown, more general approaches can lead not only to the approximation of a wider variety of phenomena, but also to the discovery of new knowledge about the physics of reality. These approaches include inductive biases based on symmetries and basic laws of physics that must be fulfilled. These approaches that belong to thermodynamic frameworks can be applied to conservative and dissipative systems indistinctly. However, considering the noise and uncertainties of the measurements, these frameworks are greatly appealing for learning real dynamics. This statement has been demonstrated within a thermodynamics-aware application based on the GENERIC formalism. A physics scene understating engine capable of learning the physics of previously unseen liquids is provided, imitating the adaptability of human cognition. In addition, physics-informed constraints trigger its applicability to augmented intelligence. Not only the motion and volume of previously unseen fluid is provided. The system outputs a physically sound prediction of immeasurable state variable as additional information for machine systems or potential users that can be used in complex decision-making scenarios.

Acknowledgments

This research is part of the DesCartes programme and is supported by the National Research Foundation, Prime Minister Office, Singapore under its Campus for Research Excellence and Technological Enterprise (CREATE) programme.

This material is also based upon work supported in part by the Army Research Laboratory and the Army Research Office under contract/grant number W911NF2210271.

This work has been partially funded by the Spanish Ministry of Science and Innovation, AEI /10.13039/501100011033, through Grant number TED2021-130105B-I00.

The authors also acknowledge the support of ESI Group through the chairs at the University of Zaragoza and at ENSAM Institute of Technology.

Declarations

All authors declare they have contributed equally to this work and there is no interest conflict.

References

[1] Tamkin, A., Brundage, M., Clark, J., Ganguli, D.: Understanding the capabilities, limitations, and societal impact of large language models. arXiv preprint arXiv:2102.02503 (2021)

[2] Li, K., Hopkins, A.K., Bau, D., Viégas, F., Pfister, H., Wattenberg, M.: Emergent world representations: Exploring a sequence model trained on a synthetic task. arXiv preprint arXiv:2210.13382 (2022)

[3] Lake, B.M., Ullman, T.D., Tenenbaum, J.B., Gershman, S.J.: Building machines that learn and think like people. Behavioral and brain sciences **40**, 253 (2017)

[4] Kubricht, J.R., Holyoak, K.J., Lu, H.: Intuitive physics: Current research and controversies. Trends in cognitive sciences **21**(10), 749–759 (2017)

[5] McCloskey, M.: Intuitive physics. Scientific american **248**(4), 122–131 (1983)

[6] Reynolds, C.R., Fletcher-Janzen, E.: Encyclopedia of Special Education: A Reference for the Education of Children, Adolescents, and Adults with Disabilities and Other Exceptional Individuals, Volume 3 vol. 3. John Wiley & Sons, ??? (2007)

[7] Piloto, L.S., Weinstein, A., Battaglia, P., Botvinick, M.: Intuitive physics learning in a deep-learning model inspired by developmental psychology. Nature human behaviour **6**(9), 1257–1267 (2022)

[8] Allen, K.R., Lopez-Guevara, T., Stachenfeld, K., Sanchez-Gonzalez, A., Battaglia, P., Hamrick, J., Pfaff, T.: Physical design using differentiable learned simulators. arXiv preprint arXiv:2202.00728 (2022)

[9] Sanchez-Gonzalez, A., Godwin, J., Pfaff, T., Ying, R., Leskovec, J., Battaglia, P.: Learning to simulate complex physics with graph networks. In: International Conference on Machine Learning, pp. 8459–8468 (2020). PMLR

[10] Hernandez, Q., Badias, A., Gonzalez, D., Chinesta, F., Cueto, E.: Deep learning of thermodynamics-aware reduced-order models from data. Computer Methods in Applied Mechanics and Engineering **379**, 113763 (2021)

[11] Finn, C., Levine, S.: Deep visual foresight for planning robot motion. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 2786–2793 (2017). IEEE

[12] Liu, C.K., Negrut, D.: The role of physics-based simulators in robotics. Annual Review of Control, Robotics, and Autonomous Systems **4**, 35–58 (2021)

[13] Allen, K.R., Smith, K.A., Tenenbaum, J.B.: Rapid trial-and-error learning with simulation supports flexible tool use and physical reasoning. Proceedings of the National Academy of Sciences **117**(47), 29302–29310 (2020)

[14] Battaglia, P.W., Hamrick, J.B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., et al.: Relational inductive biases, deep learning, and graph networks. arXiv preprint arXiv:1806.01261 (2018)

[15] Zheng, N.-n., Liu, Z.-y., Ren, P.-j., Ma, Y.-q., Chen, S.-t., Yu, S.-y., Xue, J.-r., Chen, B.-d., Wang, F.-y.: Hybrid-augmented intelligence: collaboration and cognition. Frontiers of Information Technology & Electronic Engineering **18**(2), 153–179 (2017)

[16] Rui, Y.: From artificial intelligence to augmented intelligence. IEEE MultiMedia **24**(1), 4–5 (2017)

[17] Chinesta, F., Cueto, E., Abisset-Chavanne, E., Duval, J.L., Khaldi, F.E.: Virtual, digital and hybrid twins: a new paradigm in data-based engineering and engineered data. Archives of computational methods in engineering **27**, 105–134 (2020)

[18] Moya, B., Badías, A., Alfaro, I., Chinesta, F., Cueto, E.: Digital twins that learn and correct themselves. International Journal for Numerical Methods in Engineering **123**(13), 3034–3044 (2022)

[19] Cerf, V.G.: Augmented intelligence. IEEE Internet Computing **17**(5), 96–96 (2013)

[20] Bar, M.: The proactive brain: memory for predictions. Philosophical Transactions of the Royal Society B: Biological Sciences **364**(1521), 1235–1243 (2009)

[21] Hamrick, J., Battaglia, P., Tenenbaum, J.B.: Internal physics models guide probabilistic judgments about object dynamics. In: Proceedings of the 33rd Annual Conference of the Cognitive Science Society, vol. 2 (2011). Cognitive Science Society Austin, TX

[22] Battaglia, P.W., Hamrick, J.B., Tenenbaum, J.B.: Simulation as an engine of physical scene understanding. Proceedings of the National Academy of Sciences **110**(45), 18327–18332 (2013)

[23] Traylor, A., Feiman, R., Pavlick, E.: Can neural networks learn implicit logic from physical reasoning? In: The Eleventh International Conference on Learning Representations (2022)

[24] Murphy, K.P., Torralba, A., Freeman, W.: Using the forest to see the trees: A graphical model relating features, objects, and scenes. Advances in neural information processing systems **16** (2003)

[25] Gupta, A., Efros, A.A., Hebert, M.: Blocks world revisited: Image understanding using qualitative geometry and mechanics. In: Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV 11, pp. 482–496 (2010). Springer

[26] Schenck, C., Fox, D.: Detection and tracking of liquids with fully convolutional networks. arXiv preprint arXiv:1606.06266 (2016)

[27] Shen, B., Yan, X., Qi, C.R., Najibi, M., Deng, B., Guibas, L., Zhou, Y., Anguelov, D.: Gina-3d: Learning to generate implicit neural assets in the wild. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 4913–4926 (2023)

[28] Kandukuri, R., Achterhold, J., Moeller, M., Stueckler, J.: Learning to identify physical parameters from video using differentiable physics. In: DAGM German Conference on Pattern Recognition, pp. 44–57 (2020). Springer

[29] Wu, J., Yildirim, I., Lim, J.J., Freeman, B., Tenenbaum, J.: Galileo: Perceiving physical object properties by integrating a physics engine with deep learning. Advances in neural information processing systems **28** (2015)

[30] Mrowca, D., Zhuang, C., Wang, E., Haber, N., Fei-Fei, L.F., Tenenbaum, J., Yamins, D.L.: Flexible neural representation for physics prediction. Advances in neural information processing systems **31** (2018)

[31] Bender, J., Erleben, K., Trinkle, J.: Interactive simulation of rigid body dynamics in computer graphics. In: Computer Graphics Forum, vol. 33, pp. 246–270 (2014). Wiley Online Library

[32] Rath, L., Geist, A.R., Trimpe, S.: Using physics knowledge for learning rigid-body forward dynamics with gaussian process force priors. In: Conference on Robot Learning, pp. 101–111 (2022). PMLR

[33] Huang, S., Cheng, Z.-Q., Li, X., Wu, X., Zhang, Z., Hauptmann, A.: Perceiving physical equation by observing visual scenarios. arXiv preprint arXiv:1811.12238 (2018)

[34] Li, Y., Wu, J., Tedrake, R., Tenenbaum, J.B., Torralba, A.: Learning particle dynamics for manipulating rigid bodies, deformable objects, and fluids. arXiv preprint arXiv:1810.01566 (2018)

[35] Schenck, C., Fox, D.: Spnets: Differentiable fluid dynamics for deep neural networks. In: Conference on Robot Learning, pp. 317–335 (2018). PMLR

[36] Duchaine, V., Gosselin, C.: Safe, stable and intuitive control for physical human-robot interaction. In: 2009 IEEE International Conference on Robotics and Automation, pp. 3383–3388 (2009). IEEE

[37] Koppula, H.S., Saxena, A.: Anticipating human activities using object affordances for reactive robotic response. IEEE transactions on pattern analysis and machine intelligence **38**(1), 14–29 (2015)

[38] Liu, X.-Y., Wang, J.-X.: Physics-informed dyna-style model-based deep reinforcement learning for dynamic control. Proceedings of the Royal Society A **477**(2255), 20210618 (2021)

[39] Driess, D., Schubert, I., Florence, P., Li, Y., Toussaint, M.: Reinforcement learning with neural radiance fields. Advances in Neural Information Processing Systems **35**, 16931–16945 (2022)

[40] Badias, A., Alfaro, I., Gonzalez, D., Chinesta, F., Cueto, E.: Morph-dslam: Model order reduction for physics-based deformable slam. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(11), 7764–7777 (2021)

[41] Wei, B., Zhao, Y., Hao, K., Gao, L.: Visual sensation and perception computational models for deep learning: State of the art, challenges and prospects. arXiv preprint arXiv:2109.03391 (2021)

[42] Assen, J.J.R., Nishida, S., Fleming, R.W.: Visual perception of liquids: Insights from deep neural networks. PLoS computational biology **16**(8), 1008018 (2020)

[43] Zhang, Y., Dong, Z., Obaidat, M.S., Ban, X.: Non-newtonian fluid simulation

and reconstruction from monocular videos. Simulation Modelling Practice and Theory **123**, 102688 (2023)

[44] Lopez-Guevara, T., Pucci, R., Taylor, N.K., Gutmann, M.U., Ramamoorthy, S., Suhr, K.: Stir to pour: Efficient calibration of liquid properties for pouring actions. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 5351–5357 (2020). IEEE

[45] Della Santina, C., Truby, R.L., Rus, D.: Data–driven disturbance observers for estimating external forces on soft robots. IEEE Robotics and automation letters **5**(4), 5717–5724 (2020)

[46] Richter, F., Orosco, R.K., Yip, M.C.: Image based reconstruction of liquids from 2d surface detections. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 13811–13820 (2022)

[47] Schenck, C., Fox, D.: Perceiving and reasoning about liquids using fully convolutional networks. The International Journal of Robotics Research **37**(4-5), 452–471 (2018)

[48] Kloss, A., Schaal, S., Bohg, J.: Combining learned and analytical models for predicting action effects from sensory data. The International Journal of Robotics Research **41**(8), 778–797 (2022)

[49] Degrave, J., Hermans, M., Dambre, J., et al.: A differentiable physics engine for deep learning in robotics. Frontiers in neurorobotics, 6 (2019)

[50] Avila Belbute-Peres, F., Smith, K., Allen, K., Tenenbaum, J., Kolter, J.Z.: End-to-end differentiable physics for learning and control. Advances in neural information processing systems **31** (2018)

[51] Ding, M., Chen, Z., Du, T., Luo, P., Tenenbaum, J., Gan, C.: Dynamic visual reasoning by learning differentiable physics models from video and language. Advances in Neural Information Processing Systems **34**, 887–899 (2021)

[52] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)

[53] Rao, C., Sun, H., Liu, Y.: Physics-informed deep learning for computational elastodynamics without labeled data. Journal of Engineering Mechanics **147**(8), 04021043 (2021)

[54] Callaham, J.L., Maeda, K., Brunton, S.L.: Robust flow reconstruction from limited measurements via sparse representation. Physical Review Fluids **4**(10), 103907 (2019)

[55] Yang, T.-Y., Rosca, J., Narasimhan, K., Ramadge, P.J.: Learning physics constrained dynamics using autoencoders. Advances in Neural Information Processing Systems **35**, 17157–17172 (2022)

[56] Marisca, I., Cini, A., Alippi, C.: Learning to reconstruct missing data from spatiotemporal graphs with sparse observations. Advances in Neural Information Processing Systems **35**, 32069–32082 (2022)

[57] Sun, L., Wang, J.-X.: Physics-constrained bayesian neural network for fluid flow reconstruction with sparse and noisy data. Theoretical and Applied Mechanics Letters **10**(3), 161–169 (2020)

[58] Tong, Z., Li, Y.: Real-time reconstruction of contaminant dispersion from sparse sensor observations with gappy pod method. Energies **13**(8), 1956 (2020)

[59] Li, T., Buzzicotti, M., Biferale, L., Bonaccorso, F., Chen, S., Wan, M.: Data reconstruction of turbulent flows with gappy pod, extended pod and generative adversarial networks. arXiv preprint arXiv:2210.11921 (2022)

[60] Mainini, L., Willcox, K.: Surrogate modeling approach to support real-time structural assessment and decision making. AIAA Journal **53**(6), 1612–1626 (2015)

[61] Demo, N., Tezzele, M., Rozza, G.: A deeponet multi-fidelity approach for residual learning in reduced order modeling. arXiv preprint arXiv:2302.12682 (2023)

[62] Salam, T., Hsieh, M.A.: Adaptive sampling and reduced-order modeling of dynamic processes by robot teams. IEEE Robotics and Automation Letters **4**(2), 477–484 (2019)

[63] Rovina, H., Salam, T., Kantaros, Y., Hsieh, M.A.: Asynchronous adaptive sampling and reduced-order modeling of dynamic processes by robot teams via intermittently connected networks. In: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4798–4805 (2020). IEEE

[64] Ebert, C., Ruwisch, C., Weiss, J., Uijt De Haag, M., Silvestre, F.: Trajectory planning in windy urban environment–a gappy pod approach for wind field estimates with sparse sensors. In: AIAA AVIATION 2022 Forum, p. 3757 (2022)

[65] Raben, S.G., Charonko, J.J., Vlachos, P.P.: Adaptive gappy proper orthogonal decomposition for particle image velocimetry data reconstruction. Measurement Science and Technology **23**(2), 025303 (2012)

[66] Kelshaw, D., Rigas, G., Magri, L.: Physics-informed cnns for super-resolution of sparse observations on dynamical systems. arXiv preprint arXiv:2210.17319 (2022)

[67] Yu, L., Yousif, M.Z., Zhang, M., Hoyas, S., Vinuesa, R., Lim, H.-C.: Three-dimensional esrgan for super-resolution reconstruction of turbulent flows with tricubic interpolation-based transfer learning. Physics of Fluids **34**(12) (2022)

[68] Vinuesa, R., Brunton, S.L., McKeon, B.J.: The transformative potential of machine learning for experiments in fluid mechanics. arXiv preprint arXiv:2303.15832 (2023)

[69] Gao, H., Sun, L., Wang, J.-X.: Super-resolution and denoising of fluid flow using physics-informed convolutional neural networks without high-resolution labels. Physics of Fluids **33**(7) (2021)

[70] Wang, X., Xie, L., Dong, C., Shan, Y.: Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 1905–1914 (2021)

[71] Chen, H., He, X., Qing, L., Wu, Y., Ren, C., Sheriff, R.E., Zhu, C.: Real-world single image super-resolution: A brief review. Information Fusion **79**, 124–145 (2022)

[72] Saharia, C., Ho, J., Chan, W., Salimans, T., Fleet, D.J., Norouzi, M.: Image super-resolution via iterative refinement. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(4), 4713–4726 (2022)

[73] Nair, A., Chen, D., Agrawal, P., Isola, P., Abbeel, P., Malik, J., Levine, S.: Combining self-supervised learning and imitation for vision-based rope manipulation. In: 2017 IEEE International Conference on Robotics and Automation (ICRA), pp. 2146–2153 (2017). IEEE

[74] Nava, M., Paolillo, A., Guzzi, J., Gambardella, L.M., Giusti, A.: Uncertainty-aware self-supervised learning of spatial perception tasks. IEEE Robotics and Automation Letters **6**(4), 6693–6700 (2021)

[75] Yan, M., Zhu, Y., Jin, N., Bohg, J.: Self-supervised learning of state estimation for manipulating deformable linear objects. IEEE robotics and automation letters **5**(2), 2372–2379 (2020)

[76] Flaschel, M., Kumar, S., De Lorenzis, L.: Unsupervised discovery of interpretable hyperelastic constitutive laws. Computer Methods in Applied Mechanics and Engineering **381**, 113852 (2021)

[77] Lye, K.O., Mishra, S., Ray, D.: Deep learning observables in computational fluid dynamics. Journal of Computational Physics **410**, 109339 (2020)

[78] Shin, Y.-S., Kim, J.: Sensor data reconstruction for dynamic responses of structures using external feedback of recurrent neural network. Sensors **23**(5), 2737 (2023)

[79] Moya, B., Badias, A., Gonzalez, D., Chinesta, F., Cueto, E.: Physics perception in sloshing scenes with guaranteed thermodynamic consistency. IEEE Transactions on Pattern Analysis and Machine Intelligence **45**(2), 2136–2150 (2022)

[80] Sun, C., Karlsson, P., Wu, J., Tenenbaum, J.B., Murphy, K.: Stochastic prediction of multi-agent interactions from partial observations. arXiv preprint arXiv:1902.09641 (2019)

[81] Antonova, R., Yang, J., Sundaresan, P., Fox, D., Ramos, F., Bohg, J.: A bayesian treatment of real-to-sim for deformable object manipulation. IEEE Robotics and Automation Letters **7**(3), 5819–5826 (2022)

[82] Lim, V., Huang, H., Chen, L.Y., Wang, J., Ichnowski, J., Seita, D., Laskey, M., Goldberg, K.: Real2sim2real: Self-supervised learning of physical single-step dynamic actions for planar robot casting. In: 2022 International Conference on Robotics and Automation (ICRA), pp. 8282–8289 (2022). IEEE

[83] Li, Y., Torralba, A., Anandkumar, A., Fox, D., Garg, A.: Causal discovery in physical systems from videos. Advances in Neural Information Processing Systems **33**, 9180–9192 (2020)

[84] Bai, Z., Brunton, S.L., Brunton, B.W., Kutz, J.N., Kaiser, E., Spohn, A., Noack, B.R.: Data-driven Methods in Fluid Dynamics: Sparse Classification from Experimental Data. Springer, ??? (2017)

[85] Rodríguez-Ocampo, P., Ring, M., Hernández-Fontes, J., Alcérreca-Huerta, J., Mendoza, E., Gallegos-Diez-Barroso, G., Silva, R.: A 2d image-based approach for cfd validation of liquid mixing in a free-surface condition. Journal of Applied Fluid Mechanics **13**(5), 1487–1500 (2020)

[86] Bieker, K., Peitz, S., Brunton, S.L., Kutz, J.N., Dellnitz, M.: Deep model predictive flow control with limited sensor data and online learning. Theoretical and computational fluid dynamics **34**, 577–591 (2020)

[87] Thuruthel, T.G., Shih, B., Laschi, C., Tolley, M.T.: Soft robot perception using embedded soft sensors and recurrent neural networks. Science Robotics **4**(26), 1488 (2019)

[88] Tariverdi, A., Venkiteswaran, V.K., Richter, M., Elle, O.J., Tørresen, J., Mathiassen, K., Misra, S., Martinsen, Ø.G.: A recurrent neural-network-based real-time dynamic model for soft continuum manipulators. Frontiers in Robotics and AI **8**, 631303 (2021)

[89] Bonassi, F., Farina, M., Xie, J., Scattolini, R.: On recurrent neural networks for learning-based control: recent results and ideas for future developments. Journal of Process Control **114**, 92–104 (2022)

[90] Ehrhardt, S., Monszpart, A., Mitra, N.J., Vedaldi, A.: Taking visual motion prediction to new heightfields. Computer Vision and Image Understanding **181**, 14–25 (2019)

[91] Buschoff, L.M.S., Schulz, E., Binz, M.: The acquisition of physical knowledge in generative neural networks (2023)

[92] Doerr, A., Daniel, C., Schiegg, M., Duy, N.-T., Schaal, S., Toussaint, M., Sebastian, T.: Probabilistic recurrent state-space models. In: International Conference on Machine Learning, pp. 1280–1289 (2018). PMLR

[93] Rai, R., Sahu, C.K.: Driven by data or derived through physics? a review of hybrid physics guided machine learning techniques with cyber-physical system (cps) focus. IEEE Access **8**, 71050–71073 (2020)

[94] Achterhold, J., Tobuschat, P., Ma, H., Büchler, D., Muehlebach, M., Stueckler, J.: Black-box vs. gray-box: A case study on learning table tennis ball trajectory prediction with spin and impacts. In: Learning for Dynamics and Control Conference (L4DC) (2023). accepted

[95] Driess, D., Huang, Z., Li, Y., Tedrake, R., Toussaint, M.: Learning multi-object dynamics with compositional neural radiance fields. In: Conference on Robot Learning, pp. 1755–1768 (2023). PMLR

[96] Badías, A., González, D., Alfaro, I., Chinesta, F., Cueto, E.: Real-time interaction of virtual and physical objects in mixed reality applications. International Journal for Numerical Methods in Engineering **121**(17), 3849–3868 (2020)

[97] Zhong, Y.D., Han, J., Dey, B., Brikis, G.O.: Improving gradient computation for differentiable physics simulation with contacts. In: Learning for Dynamics and Control Conference, pp. 128–141 (2023). PMLR

[98] Jiang, Y., Sun, J., Liu, C.K.: Data-augmented contact model for rigid body simulation. In: Learning for Dynamics and Control Conference, pp. 378–390 (2022). PMLR

[99] Strecke, M., Stueckler, J.: Diffsdfsim: Differentiable rigid-body dynamics with implicit shapes. In: 2021 International Conference on 3D Vision (3DV), pp. 96–105 (2021). IEEE

[100] Hernández, Q., Badías, A., Chinesta, F., Cueto, E.: Port-metriplectic neural networks: thermodynamics-informed machine learning of complex physical systems. Computational Mechanics, 1–9 (2023)

[101] Um, K., Brand, R., Fei, Y.R., Holl, P., Thuerey, N.: Solver-in-the-loop: Learning from differentiable physics to interact with iterative pde-solvers. Advances in Neural Information Processing Systems **33**, 6111–6122 (2020)

[102] Wiewel, S., Kim, B., Azevedo, V.C., Solenthaler, B., Thuerey, N.: Latent space subdivision: stable and controllable time predictions for fluid flow. In: Computer Graphics Forum, vol. 39, pp. 15–25 (2020). Wiley Online Library

[103] Takahashi, T., Liang, J., Qiao, Y.-L., Lin, M.C.: Differentiable fluids with solid coupling for learning and control. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 35, pp. 6138–6146 (2021)

[104] Karniadakis, G.E., Kevrekidis, I.G., Lu, L., Perdikaris, P., Wang, S., Yang, L.: Physics-informed machine learning. Nature Reviews Physics **3**(6), 422–440 (2021)

[105] Ayensa-Jimenez, J., Doweidar, M.H., Sanz-Herrera, J.A., Doblare, M.: Prediction and identification of physical systems by means of physically-guided neural networks with meaningful internal layers. Computer Methods in Applied Mechanics and Engineering **381**, 113816 (2021)

[106] Ajay, A., Wu, J., Fazeli, N., Bauza, M., Kaelbling, L.P., Tenenbaum, J.B., Rodriguez, A.: Augmenting physical simulators with stochastic neural networks: Case study of planar pushing and bouncing. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 3066–3073 (2018). IEEE

[107] Zeng, A., Song, S., Lee, J., Rodriguez, A., Funkhouser, T.: Tossingbot: Learning to throw arbitrary objects with residual physics. IEEE Transactions on Robotics **36**(4), 1307–1319 (2020)

[108] Allevato, A., Pryor, M., Thomaz, A.L.: Multiparameter real-world system identification using iterative residual tuning. Journal of Mechanisms and Robotics **13**(3), 031021 (2021)

[109] Rudin, C., Chen, C., Chen, Z., Huang, H., Semenova, L., Zhong, C.: Interpretable machine learning: Fundamental principles and 10 grand challenges. Statistic Surveys **16**, 1–85 (2022)

[110] Lu, L., Pestourie, R., Yao, W., Wang, Z., Verdugo, F., Johnson, S.G.: Physics-informed neural networks with hard constraints for inverse design. SIAM Journal on Scientific Computing **43**(6), 1105–1132 (2021)

[111] Liu, S., Zhongkai, H., Ying, C., Su, H., Zhu, J., Cheng, Z.: A unified hard-constraint framework for solving geometrically complex pdes. Advances in Neural Information Processing Systems **35**, 20287–20299 (2022)

[112] Alkhadhr, S., Almekkawy, M.: Wave equation modeling via physics-informed neural networks: Models of soft and hard constraints for initial and boundary conditions. Sensors **23**(5), 2792 (2023)

[113] Wu, W., Daneker, M., Jolley, M.A., Turner, K.T., Lu, L.: Effective data sampling strategies and boundary condition constraints of physics-informed neural networks for identifying material properties in solid mechanics. Applied mathematics and mechanics **44**(7), 1039–1068 (2023)

[114] Prantl, L., Ummenhofer, B., Koltun, V., Thuerey, N.: Guaranteed conservation of momentum for learning particle-based fluid dynamics. Advances in Neural Information Processing Systems **35**, 6901–6913 (2022)

[115] Tang, J., Kim, B., Azevedo, V.C., Solenthaler, B.: Physics-informed neural corrector for deformation-based fluid control. In: Computer Graphics Forum, vol. 42, pp. 161–173 (2023). Wiley Online Library

[116] Raissi, M., Perdikaris, P., Karniadakis, G.E.: 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)

[117] Huang, Z., Sun, Y., Wang, W.: Generalizing graph ode for learning complex system dynamics across environments. arXiv preprint arXiv:2307.04287 (2023)

[118] Huang, Z., Sun, Y., Wang, W.: Coupled graph ode for learning interacting system dynamics. In: Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining, pp. 705–715 (2021)

[119] Huang, Z., Sun, Y., Wang, W.: Learning continuous system dynamics from irregularly-sampled partial observations. Advances in Neural Information Processing Systems **33**, 16177–16187 (2020)

[120] Weinan, E.: A proposal on machine learning via dynamical systems. Communications in Mathematics and Statistics **1**(5), 1–11 (2017)

[121] Wen, G., Li, D., Qin, F.: Learning symplectic dynamics via generating recurrent neural network. In: Proceedings of the 2022 5th International Conference on Machine Learning and Machine Intelligence, pp. 65–70 (2022)

[122] Jin, P., Zhang, Z., Zhu, A., Tang, Y., Karniadakis, G.E.: Sympnets: Intrinsic structure-preserving symplectic networks for identifying hamiltonian systems. Neural Networks **132**, 166–179 (2020)

[123] Chen, Z., Feng, M., Yan, J., Zha, H.: Learning neural hamiltonian dynamics: a methodological overview. arXiv preprint arXiv:2203.00128 (2022)

[124] Cranmer, M., Greydanus, S., Hoyer, S., Battaglia, P., Spergel, D., Ho, S.: Lagrangian neural networks. arXiv preprint arXiv:2003.04630 (2020)

[125] Trask, N., Huang, A., Hu, X.: Enforcing exact physics in scientific machine

learning: a data-driven exterior calculus on graphs. Journal of Computational Physics **456**, 110969 (2022)

[126] Patel, R.G., Manickam, I., Trask, N.A., Wood, M.A., Lee, M., Tomas, I., Cyr, E.C.: Thermodynamically consistent physics-informed neural networks for hyperbolic systems. Journal of Computational Physics **449**, 110754 (2022)

[127] Vlassis, N.N., Sun, W.: Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening. Computer Methods in Applied Mechanics and Engineering **377**, 113695 (2021)

[128] Kaltenbach, S., Koutsourelakis, P.-S.: Incorporating physical constraints in a deep probabilistic machine learning framework for coarse-graining dynamical systems. Journal of Computational Physics **419**, 109673 (2020)

[129] Masi, F., Stefanou, I., Vannucci, P., Maffi-Berthier, V.: Thermodynamics-based artificial neural networks for constitutive modeling. Journal of the Mechanics and Physics of Solids **147**, 104277 (2021)

[130] He, X., Chen, J.-S.: Thermodynamically consistent machine-learned internal state variable approach for data-driven modeling of path-dependent materials. Computer Methods in Applied Mechanics and Engineering **402**, 115348 (2022)

[131] Yu, H., Tian, X., Weinan, E., Li, Q.: Onsagernet: Learning stable and interpretable dynamics using a generalized onsager principle. Physical Review Fluids **6**(11), 114402 (2021)

[132] Huang, S., He, Z., Reina, C.: Variational onsager neural networks (vonns): A thermodynamics-based variational learning strategy for non-equilibrium pdes. Journal of the Mechanics and Physics of Solids **163**, 104856 (2022)

[133] Biot, M.: Variational principles in irreversible thermodynamics with application to viscoelasticity. Physical Review **97**(6), 1463 (1955)

[134] Morrison, P.J.: A paradigm for joined hamiltonian and dissipative systems. Physica D: Nonlinear Phenomena **18**(1-3), 410–419 (1986)

[135] Öttinger, H.C., Grmela, M.: Dynamics and thermodynamics of complex fluids. ii. illustrations of a general formalism. Physical Review E **56**(6), 6633 (1997)

[136] Lee, K., Trask, N., Stinis, P.: Machine learning structure preserving brackets for forecasting irreversible processes. Advances in Neural Information Processing Systems **34**, 5696–5707 (2021)

[137] Zhang, Z., Shin, Y., Em Karniadakis, G.: Gfinns: Generic formalism informed neural networks for deterministic and stochastic dynamical systems. Philosophical Transactions of the Royal Society A **380**(2229), 20210207 (2022)

[138] Hernández, Q., Badías, A., González, D., Chinesta, F., Cueto, E.: Structure-preserving neural networks. Journal of Computational Physics **426**, 109950 (2021)

[139] Hernández, Q., Badías, A., Chinesta, F., Cueto, E.: Thermodynamics-informed neural networks for physically realistic mixed reality. Computer Methods in Applied Mechanics and Engineering **407**, 115912 (2023)

[140] Hernández, Q., Badías, A., Chinesta, F., Cueto, E.: Thermodynamics-informed graph neural networks. arXiv preprint arXiv:2203.01874 (2022)

[141] Liu, J., Shen, H., Wang, D., Kang, Y., Tian, Q.: Unsupervised domain adaptation with dynamics-aware rewards in reinforcement learning. Advances in Neural Information Processing Systems **34**, 28784–28797 (2021)

[142] Liu, J., Zhang, H., Zhuang, Z., Kang, Y., Wang, D., Wang, B.: Design from policies: Conservative test-time adaptation for offline policy optimization. arXiv preprint arXiv:2306.14479 (2023)

[143] Short, E.S., Allevato, A., Thomaz, A.L.: Sail: simulation-informed active in-the-wild learning. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 468–477 (2019). IEEE

[144] Allevato, A., Short, E.S., Pryor, M., Thomaz, A.: Tunenet: One-shot residual tuning for system identification and sim-to-real robot task transfer. In: Conference on Robot Learning, pp. 445–455 (2020). PMLR

[145] Lutter, M., Peters, J.: Combining physics and deep learning to learn continuous-time dynamics models. The International Journal of Robotics Research **42**(3), 83–107 (2023)

[146] Hussein, A., Gaber, M.M., Elyan, E., Jayne, C.: Imitation learning: A survey of learning methods. ACM Computing Surveys (CSUR) **50**(2), 1–35 (2017)

[147] Ho, J., Ermon, S.: Generative adversarial imitation learning. Advances in neural information processing systems **29** (2016)

[148] Chen, S., Ma, X., Xu, Z.: Imitation learning as state matching via differentiable physics. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 7846–7855 (2023)

[149] Adams, M., Li, X., Boucinha, L., Kher, S.S., Banerjee, P., Gonzalez, J.-L.: Hybrid digital twins: A primer on combining physics-based and data analytics approaches. IEEE Software **39**(2), 47–52 (2021)

[150] Sancarlos, A., Cameron, M., Abel, A., Cueto, E., Duval, J.-L., Chinesta, F.: From rom of electrochemistry to ai-based battery digital and hybrid twin. Archives of

Computational Methods in Engineering **28**, 979–1015 (2021)

[151] Desai, S., Durugkar, I., Karnan, H., Warnell, G., Hanna, J., Stone, P.: An imitation from observation approach to transfer learning with dynamics mismatch. Advances in Neural Information Processing Systems **33**, 3917–3929 (2020)

[152] Mohebujjaman, M., Rebholz, L.G., Iliescu, T.: Physically constrained data-driven correction for reduced-order modeling of fluid flows. International Journal for Numerical Methods in Fluids **89**(3), 103–122 (2019)

[153] Goswami, S., Anitescu, C., Chakraborty, S., Rabczuk, T.: Transfer learning enhanced physics informed neural network for phase-field modeling of fracture. Theoretical and Applied Fracture Mechanics **106**, 102447 (2020)

[154] Laroche, R., Barlier, M.: Transfer reinforcement learning with shared dynamics. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 31 (2017)

[155] Guastoni, L., Güemes, A., Ianiro, A., Discetti, S., Schlatter, P., Azizpour, H., Vinuesa, R.: Convolutional-network models to predict wall-bounded turbulence from wall quantities. Journal of Fluid Mechanics **928**, 27 (2021)

[156] Moya, B., Badías, A., González, D., Chinesta, F., Cueto, E.: A thermodynamics-informed active learning approach to perception and reasoning about fluids. Computational Mechanics, 1–15 (2023)

[157] Monaghan, J.J.: Smoothed particle hydrodynamics. Annual review of astronomy and astrophysics **30**(1), 543–574 (1992)

[158] Bates, C., Battaglia, P.W., Yildirim, I., Tenenbaum, J.B.: Humans predict liquid dynamics using probabilistic simulation. In: CogSci (2015)

[159] Espanol, P., Serrano, M., Öttinger, H.C.: Thermodynamically admissible form for discrete hydrodynamics. Physical review letters **83**(22), 4542 (1999)

[160] Moya, B., Alfaro, I., Gonzalez, D., Chinesta, F., Cueto, E.: Physically sound, self-learning digital twins for sloshing fluids. PLoS One **15**(6), 0234569 (2020)

[161] Brunton, S.L., Proctor, J.L., Kutz, J.N.: Discovering governing equations from data by sparse identification of nonlinear dynamical systems. Proceedings of the national academy of sciences **113**(15), 3932–3937 (2016)

[162] Pascanu, R., Mikolov, T., Bengio, Y.: On the difficulty of training recurrent neural networks. In: International Conference on Machine Learning, pp. 1310–1318 (2013). Pmlr

[163] Chung, J., Gulcehre, C., Cho, K., Bengio, Y.: Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv preprint arXiv:1412.3555 (2014)

[164] Kubo, R.: The fluctuation-dissipation theorem. Reports on progress in physics **29**(1), 255 (1966)

[165] Kraus, M.: Metriplectic integrators for dissipative fluids. In: Geometric Science of Information: 5th International Conference, GSI 2021, Paris, France, July 21–23, 2021, Proceedings 5, pp. 292–301 (2021). Springer

[166] Grmela, M., Öttinger, H.C.: Dynamics and thermodynamics of complex fluids. i. development of a general formalism. Physical Review E **56**(6), 6620 (1997)

[167] Reyes, B., Howard, A.A., Perdikaris, P., Tartakovsky, A.M.: Learning unknown physics of non-newtonian fluids. Physical Review Fluids **6**(7), 073301 (2021)

[168] Cranmer, M., Sanchez Gonzalez, A., Battaglia, P., Xu, R., Cranmer, K., Spergel, D., Ho, S.: Discovering symbolic models from deep learning with inductive biases. Advances in Neural Information Processing Systems **33**, 17429–17442 (2020)

[169] Arrieta, A.B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., García, S., Gil-López, S., Molina, D., Benjamins, R., *et al.*: Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai. Information fusion **58**, 82–115 (2020)

[170] Hao, Z., Liu, S., Zhang, Y., Ying, C., Feng, Y., Su, H., Zhu, J.: Physics-informed machine learning: A survey on problems, methods and applications. arXiv preprint arXiv:2211.08064 (2022)

[171] Kadambi, A., Melo, C., Hsieh, C.-J., Srivastava, M., Soatto, S.: Incorporating physics into data-driven computer vision. Nature Machine Intelligence, 1–9 (2023)

[172] Banerjee, C., Nguyen, K., Fookes, C., Karniadakis, G.: Physics-informed computer vision: A review and perspectives. arXiv preprint arXiv:2305.18035 (2023)

[173] LeCun, Y.: A path towards autonomous machine intelligence version 0.9. 2, 2022-06-27. Open Review **62** (2022)

[174] Silver, D., Singh, S., Precup, D., Sutton, R.S.: Reward is enough. Artificial Intelligence **299**, 103535 (2021)