FoMR
Frontiers of Multidisciplinary Research

Neural Approaches to Computational Solid Mechanics: A Critical Review

Ankit Tyagi[a], Mariya Mamajiwala[b], *Debasish Roy[a]

ABSTRACT

The rapid evolution of physics-informed and data-driven neural networks has transformed scientific computing, offering new possibilities for solving and inferring complex physical systems. Yet, within the broad realm of computational solid mechanics (CSM), these advances have produced relatively modest outcomes. Despite their theoretical appeal and other advantages such as smooth field representation and the potential to overcome numerical pathologies such as locking, physics-informed neural networks (PINNs) implemented in the strong form have not demonstrated clear advantages over established finite element methods (FEM). This limited success arises from the distinctive challenges of solid mechanics: the multi-field and differential-algebraic nature of its governing equations, the non-smooth and history-dependent character of inelastic responses, and the severe conditioning issues that accompany residual-based training.

In contrast, data-driven approaches, where neural networks augment or replace constitutive laws, have seen rapid progress and widespread adoption. Graph neural networks (GNNs) and operator-learning architectures have also emerged as promising frameworks that encode mesh topology and local physical interactions directly within message-passing rules, thereby bridging the gap between classical discretizations and modern learning paradigms. Nevertheless, these graph-based methods face their own bottlenecks–most notably in efficient backpropagation, physical interpretability, and thermodynamic consistency.

This article presents a critical and integrative review of neural approaches to computational solid mechanics. We analyze why physics-informed strategies have underperformed relative to their success in other fields. We also summarize key developments in data-driven and graph-based mechanics, and discuss how emerging probabilistic and variational formulations based on optimal transport of probabilities, weak forms, and variance-controlled loss functionals, may help overcome the current limitations. The review concludes with a perspective on future directions for building neural solvers that are physically consistent, computationally scalable, and thermodynamically grounded. This should help laying the conceptual foundations for a new generation of mechanics-informed deep learning frameworks.

Received date: 27.10.2025
Accepted date: 08.12.2025

Keywords: Physics informed neural nets; Computational solid mechanics; Weak forms; Variance reduction; Locking; Optimal transport

https://doi.org/10.65601/FoMR.2026.1.1.4

***Corresponding Author:**

[a]*Computational Mechanics Lab, Department of Civil Engineering, Indian Institute of Science, Bangalore, India*
email: royd@iisc.ac.in

[b]*Department of Mathematical Sciences, University of Nottingham, UK*

INTRODUCTION

In recent years, physics-informed neural networks (PINNs) and other physics-embedded machine learning techniques have emerged as an attractive paradigm for solving partial differential equations (PDEs) (Cai et al., 2021a,b; Cuomo et al., 2022; Karniadakis et al., 2021; Kim and Lee, 2024; Lawal et al., 2022; Luo et al., 2025; Toscano et al., 2025). By enforcing the governing equations of physics as soft or hard constraints within a neural network loss functional, these approaches promise to bypass meshing, interpolate across sparse data, and provide continuously differentiable surrogates of physical fields. Since the early demonstrations of PINNs on canonical transport and fluid problems, a vast literature has developed on neural PDE solvers and data-driven scientific computing (Brandstetter et al., 2022, 2023; Brunton and Kutz, 2023; Cao et al., 2024; Kovachki et al., 2023; Li et al., 2021, 2020; Lippe et al., 2023; Lu et al., 2021a; Pfaff et al., 2021; Raissi et al., 2019; Sanchez-Gonzalez et al., 2020). Yet, in spite of this rapid expansion, the field of computational solid mechanics (CSM) has witnessed relatively limited and uneven progress with physics-informed neural methods (Haghighat et al., 2021; Hu et al., 2024a; Mouratidou et al., 2024; Nguyen-Thanh et al., 2020).

This asymmetry is striking. Classical finite element methods (FEM), which form the backbone of computational mechanics, have matured into highly optimized frameworks capable of handling large-scale nonlinear, multi-field, and path-dependent solid mechanics problems. Neural network solvers, in contrast, are still grappling with basic questions of robustness, accuracy, and computational efficiency when applied to the same equations of elasticity, plasticity, or damage (Basir and Senocak, 2022; Krishnapriyan et al., 2021; Maddu et al., 2022; Wang et al., 2024b, 2021a,b, 2022). Early studies of PINNs in elasticity have shown that, for standard benchmark problems, the accuracy of neural approximations rarely surpasses that of a moderately refined FEM discretization, while the associated computational cost can be orders of magnitude higher (Grossmann et al., 2024; Markidis, 2021). This outcome appears paradoxical given the expressive power of deep neural networks, their potential to learn complex multi-scale patterns, and their mesh-free adaptability that should, in principle, mitigate phenomena such as volumetric and shear locking or strain localization.

The roots of this limited success may be traced to several distinctive features of solid mechanics. Firstly, the governing PDEs are almost always multi-field in nature: equilibrium equations couple with constitutive and kinematic relations, often involving auxiliary fields such as stress, strain, internal variables, or Lagrange multipliers for incompressibility. The resulting differential-algebraic structure poses severe conditioning and scaling challenges for residual-based training. Secondly, solid mechanical responses are characteristically multi-scale, sometimes history-dependent, and often non-smooth – features that are not easily captured by smooth neural approximators optimized via stochastic gradient descent. These aspects have collectively hindered the straightforward translation of PINN successes in fluid and heat transfer problems to solid mechanics (Cai et al., 2021a,b).

At the same time, the data-driven branch of computational mechanics has advanced notably. Neural networks have been widely adopted to learn constitutive laws directly from experimental or simulation data, to perform model reduction, and to accelerate finite element simulations (As'ad et al., 2022; Bleyer, 2025; Fuhg et al., 2025; Hartman and Mestha, 2017; Lee and Carlberg, 2020; Liang et al., 2018; Linka et al., 2021; Wu et al., 2020). These efforts, though largely empirical, have shown tangible advantages when data are abundant and physical constraints can be imposed locally or through hybrid FEM–NN frameworks. The contrast between the relatively stagnant progress in physics-informed approaches and the momentum in data-driven mechanics underscores the need for a balanced and critical assessment of where the fundamental bottlenecks lie.

A related development has been the emergence of graph neural networks (GNNs) and other message-passing architectures that operate naturally on mesh or particle graphs (Corso et al., 2024; Wu et al., 2021; Xu et al., 2019; Zhou et al., 2020). GNNs offer a geometric and topological representation well aligned with discretized continua, enabling the enforcement of locality and interaction laws (Battaglia et al., 2018; Bronstein et al., 2017; Kipf and Welling, 2017). Nevertheless, their deployment in physics-informed settings introduces new computational challenges, particularly in efficient backpropagation through large unstructured graphs and in maintaining physical consistency across message updates

(Gao et al., 2022). Difficulty in efficient backpropagation also poses a hurdle in enforcing Neumann boundary conditions. Understanding these trade-offs is crucial if GNNs are to serve as scalable surrogates for solid mechanics (Zhao et al., 2024).

In view of these diverse trends, this article aims to present a critical and integrative review of neural approaches to CSM. We begin with a concise overview of PINN formulations, emphasizing their adaptation to elasticity problems and highlighting key limitations observed in the literature. We then examine why, despite their theoretical advantages, these methods have shown little evidence of outperforming established finite element schemes. Next, we review the rapid advances in data-driven mechanics and graph-based learning, identifying lessons that may cross-fertilize the physics-informed paradigm. The latter part of the paper turns to emerging alternative routes—probabilistic formulations based on optimal transport, weak-form or variational neural schemes, and variance-controlled loss constructions— that hold promise for overcoming some of the current obstacles. The discussion closes with a perspective on open problems and a research roadmap toward more reliable and thermodynamically consistent neural solvers for solid mechanics.

Overall, this critical review does not aim merely to catalogue existing works, but rather to trace the conceptual roots of both the limitations and opportunities that characterize the intersection of neural computation and solid mechanics. By critically analyzing the state-of-the-art landscape, we seek to clarify the paths along which future progress is most likely to occur.

A BRIEF PRIMER ON PINNs AND RELATED NEURAL PDE SOLVERS

Strong-form PINNs

Strong-form physics-informed neural networks (PINNs) implement the governing PDEs directly as pointwise residuals inside the neural-network loss (Raissi et al., 2019). Let $\Omega \subset R^d$ be the solid's domain, $\boldsymbol{u} : \Omega \to R^d$ the displacement field (say, in the small deformation setting), and denote by $N_\theta : \Omega \to R^d$ a neural network with parameters θ used to approximate $\boldsymbol{u}$. For a prototypical linear-elastic problem with body force $\boldsymbol{b}(\boldsymbol{x})$ and traction $\boldsymbol{t}$ on the Neumann boundary Γ_{N}, the strong-form residual

$$\mathcal{R}_{\mathrm{eq}}(x;\theta) \;=\; \nabla \cdot \sigma(\mathcal{N}_\theta(x)) + b(x) \qquad (1)$$

is enforced (in a mean-square sense) together with boundary residuals on Dirichlet and Neumann portions Γ_{D}, Γ_{N}. Here $\boldsymbol{\sigma}$ denotes the so called stress tensor (in the small deformation setting, we do not distinguish between Cauchy and Piola stresses; nor is $\boldsymbol{\sigma}$ a proper tensor for the small deformation and isotropic linear elastic setup). Let $\boldsymbol{u_D}$ and $\boldsymbol{n}$ respectively denote the specified displacement over Γ_{D} and unit outward normal on Γ_{N}. The canonical PINN loss then takes the form

$$\mathcal{L}(\theta) \;=\; \underbrace{\frac{1}{N_\Omega}\sum_{i=1}^{N_\Omega} \|\mathcal{R}_{\mathrm{eq}}(x_i;\theta)\|^2}_{\mathcal{L}_{\mathrm{int}}} + \alpha_D \underbrace{\frac{1}{N_D}\sum_{j=1}^{N_D} \|\mathcal{N}_\theta(x_j) - u_D(x_j)\|^2}_{\mathcal{L}_D} + \alpha_N \underbrace{\frac{1}{N_N}\sum_{k=1}^{N_N} \|\sigma(\mathcal{N}_\theta)\cdot n - t\|^2_{x_k}}_{\mathcal{L}_N}, \qquad (2)$$

where $\{\boldsymbol{x_i}\}$ are collocation points in the domain and $\{\boldsymbol{x_j}\}$, $\{\boldsymbol{x_k}\}$ are points on the boundaries. α_D and α_N are appropriately chosen penalty parameters. Automatic differentiation provides the divergence $\nabla\cdot\boldsymbol{\sigma}$ and other derivatives required for R_{eq} (Baydin et al., 2018).

Strong-form PINNs have compelling conceptual simplicity and have been successfully demonstrated on many scalar and low-order vector PDEs (Antonion et al., 2024; Farea et al., 2024). However, when applied to the multi-field, multi-scale, or path-dependent PDEs common in CSM, they encounter several systematic obstacles (Arora et al., 2022; Fuhg and Bouklas, 2022; Haghighat et al., 2022; Sharma et al., 2023; Wang et al., 2024c). We enumerate and analyze some of these issues below.

Multi-field coupling and loss imbalance: Many CSM problems are not single-field as the displacement itself has three components in 3D solids; additionally, they couple the displacement vector $\boldsymbol{u}$ with auxiliary variables such as pressure p (in incompressible or nearly-incompressible elasticity), stress $\boldsymbol{\sigma}$ when treated explicitly in mixed methods, plastic or damage-related internal variables $\boldsymbol{z}$, or temperature T. A typical mixed (saddle-point) formulation uses unknowns $(\boldsymbol{u}, p)$ satisfying

$$\begin{aligned} \nabla \cdot \sigma(u,p) + b &= 0, \\ \mathcal{C}(u,p) &= 0, \end{aligned} \qquad (3)$$

where $\mathcal{C}$ enforces incompressibility or other constraints. In a strong-form PINN, this produces multiple residual terms with very different magnitudes and sensitivities (scales). For example, in linear isotropic elasticity of St. Venant-Kirchhoff type (Jog, 2015), the Lamé constants λ, μ set dominant scales: as the material approaches incompressibility $(\lambda \gg \mu)$, the equilibrium residual becomes highly stiff in the volumetric part while the constraint residual (enforcing $\nabla \cdot u = 0$ or pressure compatibility) must be satisfied to machine precision to avoid locking. When residuals of different fields are naively summed (even with scalar weights α's, see equation 2), gradient directions from the different terms can interfere, producing slow or stalled training and solutions that satisfy some residuals well whilst severely violating others. This loss-scale mismatch is therefore a structural issue for multi-field PINNs (Khadijeh et al., 2025; Li and Feng, 2022; Pu and Feng, 2022; Xiang et al., 2022).

Conditioning and the effective stiffness of the optimization problem: Enforcing differential operators as pointwise residuals creates an optimization landscape whose curvature reflects the PDE operator spectrum. High-frequency modes or stiff volumetric responses correspond to large eigenvalues of the linearized operator; those directions induce steep gradients in parameter space and require very small learning rates, many iterations, or second-order preconditioning to resolve (Wang et al., 2021b, 2022). In contrast, FEM involves no pointwise residuals and handles such stiffness by targeted spatial discretization and specialized linear solvers (e.g. preconditioned Krylov methods) whose numerical linear-algebra structure is tailored to the operator. PINN training, driven by generic first-order optimizers, lacks an equally effective, problem-informed preconditioning mechanism by default.

Need for high-order derivatives and associated numerical noise: Many continuum constitutive relations involve derivatives (strain, strain rate) and sometimes higher-order gradients (e.g. strain gradient plasticity). Strong-form PINNs compute these via automatic differentiation through the neural mapping $\boldsymbol{x} \rightarrow N_\theta(\boldsymbol{x})$. While exact in symbolic differentiation, practical evaluation at many collocation points produces numerical instabilities: deep networks with complicated activations can yield poorly behaved derivative estimates (large variance across points), especially for higher derivatives. This contributes to noisy gradient estimates for evaluating the loss and can slow or destabilize training (Basir, 2023; Gladstone et al., 2025; Song et al., 2024).

Sampling issues and collocation density requirements: Accurate enforcement of PDE residuals requires adequate sampling of $\boldsymbol{\Omega}$, including near boundaries, corners, and material interfaces where fields exhibit steep gradients or singular behaviour. For multi-field CSM problems (with heterogeneous material properties, sharp damage fronts), the number of collocation points needed to achieve a given error tolerance may be very large, erasing the purported mesh-free advantage. Moreover, uniform random sampling is inefficient: unless carefully designed curricula or adaptivity (importance sampling) are used. In other words, regions of high error may remain under-resolved (Daw et al., 2023; Gao et al., 2023; Nabian et al., 2021; Wu et al., 2021).

Path-dependence and non-smooth evolution laws: Plasticity and damage models are history-dependent and often involve non-smooth constitutive rules (yield surfaces, rate-independent flow rules, variational inequalities). Strong-form PINNs, which aim to learn spatial fields for a fixed time or enforce time-dependent residuals, struggle to represent the required discontinuous updates or the correct dependence on internal state history without explicitly embedding update algorithms (Bleyer, 2025). This makes direct application of naïve strong-form PINNs to inelastic solids problematic (Eghbalian et al., 2023; Rezaei et al., 2023).

Computational and memory cost of backpropagating through differential operators: Each residual evaluation typically requires computing first and second order gradients of network outputs w.r.t. inputs. For large networks and many collocation points, this turns out to be expensive in both time and memory: backpropagation must retain intermediate activations for automatic differentiation, leading to high GPU memory usage. When multi-field unknowns are used, the dimensionality of outputs grows and the demand on memory increases further (Cho et al., 2023; Hu et al., 2024b).

Main points at a glance: Strong-form PINNs are simple and broadly applicable in principle, but their naïve and casual application to multi-field CSM exposes several interrelated weaknesses: (i) loss-scale mismatch and gradient interference

across fields and response regimes; (ii) gradient-based optimization conditioning issues tied to PDE stiffness; (iii) high collocation sampling costs and derivative noise; (iv) difficulty in encoding path-dependent, non-smooth constitutive behaviour; and (v) substantial computational/memory overhead. These challenges explain why many early PINN studies in solid mechanics reporting success on toy problems do not yet translate into reliable, scalable alternatives to well-established FEM workflows.

Remarks on mitigations: Several partial remedies have been proposed in the literature. These include adaptive loss weighting (Gao et al., 2025), operator-aware normalization (Maldonado et al., 2023), mixed and/or variational formulations (Uriarte et al., 2025), domain decomposition (Klawonn et al., 2024), and hybrid FEM-NN schemes (Mitusch et al., 2021). Each mitigates some weak spots, but none, to the best of our knowledge, provides a universal fix. In particular, recasting PDE constraints into weak or variational forms (which more naturally respect conservation and provide built-in preconditioning via test-function weighting) addresses many of the conditioning and loss-balancing issues discussed above. This motivates the subsequent subsection on variational/mixed PINNs.

Variational / Weak-Form and Mixed PINNs

An alternative to the strong-form formulation of PINNs is to adopt the variational or weak form of the governing equations (Kharazmi et al., 2021; Zang et al., 2020). In the classical finite element setting, the weak form arises from multiplying the equilibrium equations by admissible test functions and integrating over the domain, thereby transferring spatial derivatives from the trial field to the test field through integration by parts. This process defines a bilinear or nonlinear functional whose stationary point corresponds to the physical solution. When translated to a neural framework, the field variables could be represented by neural trial functions, while the residual of the weak form (or equivalently, the total potential energy functional) provides the loss to be minimized. This variational viewpoint naturally incorporates boundary conditions and enables energy-based training objectives that align closely with the principles of virtual work and thermodynamic consistency.

For multi-field problems, such as incompressible elasticity, mixed elasticity, or coupled electro-mechanical systems, one may extend the same basic idea to a mixed variational formulation. Here separate neural networks approximate the different primary variables (e.g. displacement, pressure, or electric potential), while the coupling arises through a common weak functional that enforces balance and compatibility in a variational sense. This class of mixed PINNs or weak-form neural solvers, if developed, would generalize the Galerkin method to a mesh-free, differentiable framework.

The potential advantages of such formulations are significant. First, weak-form and mixed PINNs inherit much of the favorable conditioning of Galerkin discretizations, since residuals are evaluated in an averaged rather than pointwise sense. This often leads to smoother gradients and more stable optimization compared with strong-form counterparts. Second, constraints such as incompressibility, symmetry of stress, or boundary traction conditions can be handled naturally through the choice of test spaces and variational structure, obviating the need for artificial penalty terms. Third, energy-based losses endow the scheme with physical interpretability: the optimization seeks the stationary point of a potential or complementary energy, rather than the arbitrary minimization of squared residuals. Collectively, these features suggest that variational and mixed PINNs might offer a principled route toward reliable and physically consistent neural solvers for solid mechanics, particularly in settings where strong-form PINNs exhibit poor convergence or spurious constraint violations.

Probabilistic PINNs and Bayesian / Optimal Transport Flavors

A conceptually distinct line of work, currently under development by our group, reinterprets physics-informed learning through a probabilistic or measure-theoretic lens. Instead of treating the neural network output as a single deterministic field satisfying the governing equations, one may view it as parameterizing a probability distribution over the admissible physical states, which in turn define the sample space. The learning task then is one of evolving this distribution so that it concentrates near configurations consistent with both the governing equations and available data. This shift from deterministic regression to probabilistic inference introduces an intrinsic mechanism for quantifying uncertainty and regularizing ill-posed or data-sparse problems that are ubiquitous in CSM.

From the popular Bayesian perspective, the governing equations may be looked upon as priors or constraints that shape the posterior over model parameters and field realizations. Training in this framework corresponds to performing approximate Bayesian inference, where data and physics (the latter is considered as generalized data) jointly update the belief about the system's state. Such formulations naturally connect PINNs with stochastic filtering and data assimilation methods that have long been used to reconcile model predictions with observations.

An alternative yet mathematically related viewpoint arises from the optimal transport (OT) theory. Here, the evolution of the probability measure associated with the network output can be seen as a transport process driven by a gradient flow in a suitable metric space, typically the Wasserstein space of probability measures. The physics residuals define a potential whose minimization guides the transport toward equilibrium. Entropic or Sinkhorn-type regularizations (Cuturi, 2013) could be introduced to ensure numerical tractability and smooth convergence. This OT interpretation transforms physics-informed training from a pointwise loss minimization into a global measure-matching problem, offering a natural route to stability, regularization, and physical interpretability.

Together, these Bayesian and OT-inspired formulations point toward a unifying probabilistic foundation for physics-informed learning in CSM. By embedding uncertainty, dissipation, and measure evolution directly within the learning framework, they promise enhanced robustness, interpretability, and adaptability – features that are essential for modeling complex, multi-scale solid mechanical systems. The detailed mathematical structures underlying these formulations remain an active topic of ongoing research and will be taken up elsewhere.

STATE-OF-THE-ART ON PINNs APPLIED TO CSM

As noted, when used to solve CSM problems, PINN often struggles with limited accuracy and computational efficiency. That is perhaps why many PINN-based applications are only on toy problems. They generally involve PDEs with scalar-valued unknown fields, e.g. Burgers equation, Poisson's equation, the reaction-diffusion equation etc. (Cai et al., 2020; Hao et al., 2023; Raissi et al., 2019; Sirignano and Spiliopoulos, 2018; Weinan and Yu, 2018). This may be contrasted with several instances of deep learning data-driven methods applied, with a measure of success, to non-trivial CSM problems, specifically using GNNs (Pfaff et al., 2021; Zhao et al., 2024). Haghighat et al. (2021) apply PINN to linear elasticity and elasto-plastic problems, only to find no advantages in PINN as a forward solver either in accuracy or in performance. So, they instead use it for inversion and surrogate modeling. They observe that using a separate multi-layer perceptron (MLP) for each field variable is more accurate than a single network with all the variables as its output. Also, both stresses and displacements are used as unknowns, similar to mixed methods in FEM. The study does not compare the performance of the neural schemes against more conventional alternatives.

More recently, it is shown that inverse problems in physics and engineering could be solved faster and more accurately via a learning approach that does not use neural networks (Karnakov et al., 2024). Here, the loss function is constructed using conventional grid-based discrete approximations of PDEs. The problem remains sparse, making Hessian based optimization schemes with quadratic convergence rate applicable. The method inherits the accuracy, convergence, and conservation properties from grid-based discretizations of PDEs. Although the method uses automatic differentiation, it outperforms PINN by several orders of magnitude in computational speed, and has better accuracy and convergence rates. In the context of surrogate modeling with neural networks parameterized with material properties, it seems to be realizable only for small parameter spaces, especially owing to the amount of data required and is unlikely to replace established methods (Baker et al., 2019).

Roy et al. (2023) also use a mixed approach and separate MLPs to solve 2D linear elasticity benchmark problems, including the Airy solution to elasticity and the Kirchhoff–Love plate problem. However, the loss function includes data-driven fitting terms across randomly selected collocation points, making the approach hybrid. No potential benefits of this approach are shown against the established methods. Accuracy and computational costs are reported, but not compared with the alternative

approaches. The loss function also includes many hyperparameters, whose tuning is not discussed. Like Haghighat et al. (2021) and Roy et al. (2023), Mouratidou et al. (2024) also use an ensemble of neural networks to solve 2D elasticity problems and compare the solutions with those obtained using an FE software. As reported in other studies, the ensemble works better than a single network, and yet its accuracy and computational efficiency lag far behind FE solutions.

The studies discussed so far use an aggregate of mean squared error (MSE) of PDE residuals as the loss function, as proposed in the original work by Raissi et al. (2019). Nguyen-Thanh et al. (2020) propose deep energy method (DEM) to solve large deformation hyperelasticity problems, which minimize the total potential energy serving the loss function. It has the benefit of a lower order of derivative in the loss function and an automatic satisfaction of traction-free boundary conditions. Although the method has been shown to solve multiple problems satisfactorily, its computational efficiency and accuracy have not been quantified and compared. Samaniego et al. (2020) use both forms of PINN —based on MSE of PDE residuals and 'deep energy' — to solve several solid mechanics problems, e.g. linear elasticity, hyperelasticity, beam and plate bending, elastodynamics, and phase-field modeling of fracture. The study does show the potential of PINN in solving various types of CSM problems, but there are no comparisons with the existing methods. While the main focus is on the deep energy method, it is not quite general as a variational principle is not always available. Bai et al. (2023) use PINN with a modified least square weighted residual as the loss function (LSWR) and an integration scheme based on Delauny triangulation to solve 2D and 3D problems in solid mechanics. They show that its performance is better than the PINN using MSE of PDE residuals or the free energy as the loss function. However, comparisons with the FEM and computational and hyperparameter tuning costs are not reported. In the context of dissipative problems, Eghbalian et al. (2023) present a novel architecture called elasto-plastic neural network (EPNN), which is embedded with the additive decomposition of strains into elastic and plastic parts. It is simply a surrogate for classical elasto-plastic constitutive relations and performs better than the vanilla network surrogates.

Imposing Dirichlet boundary conditions exactly has always been a challenge in PINN. Generally they are imposed in weak sense by minimizing their mean squared error. Wang et al. (2023) propose EPINN (Exact Dirichlet boundary PINN) to solve solid mechanics problems based on the principle of least work. Dirichlet boundary conditions are imposed using distance functions. Only nodal displacements are used as the trainable parameters, and FE shape functions and tensor decompositions are used to reduce the number of parameters. EPINN shows better overall performance than PINN, even though it is less general. M-PINN (Wang et al., 2024a) (mesh-based PINN) goes one step further and use the finite element function-space itself in the solution approximation. It reduces the difficulty of optimization of PINN by constraining the neural functional approximation with the finite element data distribution. M-PINN is shown to outperform traditional PINN on plate-with-a-hole problems with 2D linear elastic material. Gao et al. (2022) make degrees of freedom over the FE mesh the nodal features of a graph and use a GNN to solve elasticity problems among others. The spatial derivatives of the field variables are still obtained using FEM shape functions, as they are not directly available in a GNN, unlike the standard PINN. They implicitly solve for the degrees of freedom by optimizing the network parameters. The method imposes Dirichlet boundary conditions exactly, and it automatically satisfies Neumann boundary conditions through the weak form. Using weak form leads to lower order of derivatives in the loss function; moreover these derivatives are obtained from the FEM shape functions, making the computation faster. This method certainly has a few advantages over the standard PINN, but it is not clear what advantages it might have over the FEM.

In an interesting review, Li et al. (2025) highlight the problems faced by PINN when applied to CSM and identify two major issues: PINN providing infinite domain solutions and using Euclidean geometry. They propose to remedy these by encoding finite geometry into the neural network inputs through Laplace-Beltrami operator (LBO) eigen-functions. Both strong and weak formulations of the method have been implemented and shown to outperform PINN on several problems, especially in computational efficiency. Le-Duc et al. (2024) propose hierarchically normalized PINN (hnPINN) to overcome common ill-conditioning issues of

the standard PINN. The crux of the method is a proper normalization and non-dimensionalization of the PDEs and gradients. The method is applied to several CSM problems to showcase its better performance when compared with the standard PINN. Wang et al. (2025) propose a self-adaptive weighing strategy using either the gradient normalization algorithm or augmented Lagrangian to ameliorate the loss-imbalance issue of PINN when applied to CSM. The method does show overall better performance than the vanilla PINN, but lacks in details about computational efficiency and a comparison with the established methods. Hu et al. (2024a) compare different kinds of PINN, DEM, and neural operator methods with FEM on several CSM benchmark problems and find them lacking in both accuracy and computational efficacy. Among neural methods, neural operator methods (Kovachki et al., 2023) perhaps have the most expensive training and the lowest accuracy. Manav et al. (2024) solve phase-field models of brittle fracture using deep Ritz method (Weinan and Yu, 2018) to simulate quasistatic crack nucleation, propagation, kinking, branching, and coalescence. Although they report a measure of success with the numerical work, there are many challenges vis-à-vis the network architecture, activation function, boundary conditions, optimization method, automatic differentiation, and weight regularization. There has also been some theoretical work computing error bounds on PINN solutions. Mishra and Molinaro (2022) compute generalization error bounds for a data assimilation inverse problem solved using PINN. For numerical demonstration, they use Poisson, heat, wave and Stokes equations. Similarly, De Ryck and Mishra (2022) compute error bounds for Kolmogorov type equations solved using PINN. Among other things, these error bounds presuppose a well trained PINN and use of sufficient number of collocation points.

WHAT AILs PINNs Vis-á-Vis FEM IN CSM: A CRITICAL ANALYSIS

Despite the rapid proliferation of physics-informed neural networks (PINNs) across diverse fields of computational physics, their impact on CSM has remained limited and, in many cases, disappointing. As already noted, while PINNs offer an attractive mesh-free formulation and a flexible representation of field variables, these apparent advantages have not translated into superior accuracy, robustness, or efficiency compared to the finite element method (FEM). This section examines the structural, numerical, and physical reasons underlying this gap, focusing on five interrelated aspects: (i) conditioning and scaling of residual losses, (ii) multi-field coupling and constraint enforcement, (iii) energy and equilibrium inconsistency, (iv) lack of adaptive resolution and variational grounding, and (v) computational inefficiency and reproducibility issues.

ill-conditioning and residual imbalance

At the heart of any PINN lies the minimization of a loss functional composed of PDE residuals and boundary-condition penalties. For a mechanical system governed by

$$\nabla \cdot \sigma + b = 0, \quad \sigma = \mathbb{C} : \varepsilon(u), \tag{4}$$

the residual norm,

$$\mathcal{L}_{\mathrm{PDE}} = \int_{\Omega} \|\nabla \cdot (\mathbb{C} : \varepsilon(u_\theta)) + b\|^2 \, \mathrm{d}\Omega, \tag{5}$$

contains second derivatives of the neural approximation $\boldsymbol{u}_\theta$. These derivatives amplify even minor fluctuations in the learned displacement field and create loss landscapes that are highly ill-conditioned. Gradient-based training then suffers from exploding or vanishing sensitivities, leading to slow convergence and spatially oscillatory solutions. In contrast, the FEM operates on weak forms that involve only first derivatives and produce symmetric positive-definite stiffness matrices—an implicit form of preconditioning absent in the strong-form PINN.

Moreover, mechanical PDEs are typically anisotropic: residual magnitudes vary by orders of magnitude between stress equilibrium, kinematic compatibility, and constitutive relations. Naively weighted residual norms give disproportionate importance to certain components (e.g. normal stresses) at the expense of others (e.g. shear or volumetric modes). Attempts to balance these through heuristic weighting or adaptive loss rescaling (Wang et al., 2021a) only partially mitigate the problem. The resulting optimization remains fragile and problem-specific.

Multi-field coupling and constraint enforcement

Most formulations in solid mechanics are multi-field systems: displacement-pressure pairs in mixed elasticity, stress–strain–internal variable triplets in plasticity, or displacement– damage fields in fracture mechanics. These couplings are governed by differential–algebraic relations that impose compatibility, incompressibility, or yield-related constraints. Enforcing such relations through soft penalties in a PINN loss function inevitably breaks the underlying saddle-point structure of the problem. The loss minimization then drives the network toward configurations that may locally satisfy one constraint while violating another, viz. achieving equilibrium but losing compatibility.

In FEM, these couplings are handled through mixed or hybrid variational formulations that guarantee stability via inf-sup conditions and consistent interpolation spaces. No such theoretical safeguards exist for PINNs. The neural approximation spaces for multiple fields are typically independent and not constructed to satisfy compatibility or inf-sup stability, leading to spurious pressure oscillations in nearly incompressible materials and loss of equilibrium accuracy in mixed formulations. This structural mismatch between the algebraic topology of CSM equations and the parameter topology of neural networks is perhaps the single most significant reason for their underperformance.

Energy and equilibrium inconsistency

Strong-form PINNs directly minimize the squared residual of the PDE, not the potential energy or any thermodynamically meaningful functional. As a result, they do not guarantee pathwise energy consistency or mechanical equilibrium in the sense of virtual work. Even when the PDE residual becomes small in the L^2-norm, the resulting field may violate the discrete balance of internal and external work or exhibit spurious stress fluctuations that have no energetic interpretation. For inelastic or dissipative systems, this inconsistency becomes more severe: the network may satisfy the instantaneous flow rule but violate the global dissipation inequality.

In contrast, variational formulations, on which FEM is fundamentally based, minimize an energy functional that encodes the structure of the governing laws. The Galerkin approximation ensures that the weak residual is orthogonal to the admissible space, preserving energy consistency at the discrete level. The absence of such variational grounding in PINNs deprives them of the mathematical structure that underpins both stability and convergence aspects of the FEM.

Lack of adaptive resolution and hierarchical refinement

Mechanical problems often involve localized deformation modes: stress concentrations, strain localization, or sharp gradients in plastic zones and cracks. FEM accommodates these naturally through mesh refinement and adaptive remeshing, guided by error estimators derived from the variational structure. By contrast, PINNs rely on global function approximation with uniform sampling. Without explicit adaptivity or multi-scale representations, the network wastes representational capacity on smooth regions and fails to resolve singular or localized features. Adaptive collocation strategies (Lu et al., 2021b) offer partial remedies, but these remain ad hoc and computationally expensive. The absence of a built-in notion of spatial hierarchy or error localization prevents PINNs from achieving the spatial efficiency and accuracy that mesh-based methods routinely deliver.

Computational inefficiency and reproducibility

Training a PINN for a three-dimensional elasticity problem typically requires millions of collocation points and thousands of gradient-descent iterations. Automatic differentiation, while convenient, scales poorly with network depth and with the need to compute second-order derivatives. The resulting computational cost often exceeds that of assembling and solving the FEM stiffness matrix by several orders of magnitude. Moreover, the stochastic nature of optimization (random initialization, batch sampling) leads to solutions that are not reproducible, unlike deterministic FEM solutions. Hyperparameter tuning (network architecture, learning rate, residual weights) further adds to the uncertainty and hinders systematic verification and validation.

A summary of the critical factors

To summarize, the failure of PINNs to outperform FEM in CSM arises from a confluence of structural and numerical mismatches, as enumerated below.

- The use of strong-form residuals introduces high-order derivatives and ill-conditioning absent in variational formulations.

- Multi-field coupling is treated through heuristic penalties rather than stable mixed formulations.
- Energy and equilibrium consistency, fundamental to mechanical fidelity, are not enforced by design.
- Absence of adaptive spatial refinement limits accuracy in localized or multiscale phenomena.
- Training remains computationally prohibitive and lacks reproducibility guarantees.

These shortcomings do not imply that neural approaches are intrinsically unsuitable for solid mechanics. Rather, they highlight that a direct transposition of strong-form PINN ideas from scalar or single-field PDEs to coupled and constrained mechanical systems overlooks essential structural properties of continuum mechanics. The way forward, therefore, lies not in the incremental tuning of strong-form networks, but in rethinking their foundational formulation altogether, embedding variational, probabilistic, and physically constrained structures that align with the governing laws of solids. The subsequent sections of this review discuss emerging strategies that move in this direction.

HOW DATA-DRIVEN APPROACHES HAVE ADVANCED MECHANICS

While PINNs have struggled to achieve decisive advantages in computational solid mechanics, the broader family of data-driven approaches has made much more tangible progress. These methods replace, augment, or accelerate parts of traditional simulation workflows by exploiting data to infer constitutive behaviour, reduce computational cost, or construct efficient surrogates. In contrast to PINNs, which attempt to satisfy governing equations directly through residual minimization, datadriven mechanics often preserves the established numerical infrastructure of finite elements, finite volumes, or particles, and confines the role of learning to sub-models or mappings where data are abundant, but physical closure is uncertain. This section reviews the major directions of progress, highlighting why data-driven methods have found wider adoption and what lessons they hold for physics-informed learning.

Data-driven constitutive modelling

One of the earliest and most successful avenues of machine learning in solid mechanics has been the identification of constitutive relations from data. Classical approaches require selecting a functional form (e.g. hyperelastic, viscoplastic) and calibrating a small set of parameters. Neural-network-based constitutive models relax this assumption by learning the stress-strain or stress versus strain-rate relationship directly from experimental, multiscale, or simulation data.

A typical workflow employs neural networks as nonparametric regressors

$$\sigma = \mathcal{N}_\theta(\varepsilon, \dot{\varepsilon}, T, \ldots), \tag{6}$$

trained on stress-strain data pairs (T denotes the ambient temperature). Variants have been developed (Garanger, 2024) to enforce frame indifference, material symmetries, and thermodynamic consistency by constructing invariant inputs or embedding potential-based architectures where $\boldsymbol{\sigma} = \partial W_\theta / \partial \boldsymbol{\varepsilon}$. Such constitutive neural networks have been incorporated into finite element codes as drop-in material subroutines, often achieving excellent interpolation accuracy and significant speed-up in multiscale simulations where high-fidelity microscale computations are replaced by a trained surrogate. For example, deep material networks that learn the material constitution (Dey et al., 2024) and tensor basis neural models (Bhatia et al., 2025) have been used to represent the homogenized response of composites or polycrystals, offering both accuracy and interpretability.

A key reason for the success of data-driven constitutive learning is its modularity: the equilibrium equations and numerical solvers remain untouched, so the stiff multi-field coupling that plagues strong-form PINNs does not appear. Moreover, the training data are typically local (at the integration point level) and abundant, ensuring efficient learning. The cost of network evaluation at runtime is marginal compared with the global FEM assembly and solve, making the approach computationally viable.

Model-free and hybrid data-driven solvers

Beyond learning constitutive laws, data-driven methods have been developed to bypass constitutive modelling entirely, treating the stress and strain data themselves as elements of an admissible set. The seminal formulation by Kirchdoerfer and Ortiz (2016) recasts equilibrium as a nearest-neighbour problem: among all strain– stress pairs consistent

with compatibility and equilibrium, select those closest (in an energetic norm) to the experimental data set. This model-free approach has since been extended to inelastic, path-dependent, and stochastic settings (Eggersmann et al., 2019; Prume et al., 2023). Although not neural in the traditional sense, its conceptual influence is profound: it demonstrates that data can directly substitute for phenomenological models whilst preserving mechanical structure.

Neural versions of this idea use autoencoders or latent-variable networks to parameterize admissible stress-strain manifolds, enabling interpolation and denoising of sparse experimental databases. Hybrid variants combine these data-driven concepts with physics-based constraints: the global equilibrium and compatibility equations are enforced as in FEM, while the local material response is provided by a learned mapping or database query. These hybrid solvers have shown robust accuracy and scalability across diverse problems, from hyperelastic membranes to complex composites, and have been implemented in both CPU- and GPU-based finite element frameworks (Dettmer et al., 2024; Korzeniowski and Weinberg, 2021; Vlassis et al., 2020).

Reduced-order and surrogate modelling

Another highly active direction is the use of neural networks for reduced-order modelling and fast surrogates of high-fidelity simulations. Techniques such as proper orthogonal decomposition (POD) and reduced basis methods have long been used to accelerate parametric analyses; neural networks, especially autoencoders and recurrent architectures, now serve as nonlinear generalizations of these classical approaches. By learning low-dimensional manifolds of system response, they can predict global displacements or stress fields at new loading conditions orders of magnitude faster than conventional solvers (Chen et al., 2021; Fresca et al., 2021; Fresca and Manzoni, 2022).

A related line of work employs convolutional or graph-based networks as surrogates that map material geometry or load parameters to global responses (e.g. maximum stress, failure probability) (Khorrami et al., 2023; Maurizi et al., 2022; Pfaff et al., 2021; Zhu et al., 2019). When combined with transfer learning, such surrogates provide rapid exploration tools for design and uncertainty quantification. Because they are trained on precomputed simulation databases, their deployment incurs no new numerical stiffness or constraint-enforcement difficulties.

Why data-driven methods have progressed faster

The relative maturity of data-driven approaches compared with physics-informed solvers stems from several structural advantages:

- **Modular integration:** Data-driven components can be embedded within established numerical infrastructures, preserving the stability and conditioning of the global solver.
- **Local learning tasks:** Constitutive models and surrogates operate on low-dimensional, local data (stress–strain pairs or reduced modes) rather than high-dimensional PDE fields, resulting in smoother loss landscapes and more stable optimization.
- **Abundant and high-quality data:** Experimental or simulated databases of material behaviour are increasingly available, allowing data-driven models to train efficiently without complex residual weighting schemes.
- **Ease of benchmarking and validation:** Because data-driven models are plugged into FEM workflows, their performance can be quantitatively compared with conventional material models on standard benchmarks.

By contrast, PINNs require solving the full inverse problem of minimizing field residuals, which introduces multi-field coupling, conditioning, and computational barriers.

Lessons for physics-informed learning

Despite their empirical orientation, data-driven approaches provide valuable guidance for advancing physics-informed networks:

- Local learning within a global constraint structure, akin to hybrid FEM-NN schemes, may combine the best of both paradigms.
- Embedding physical symmetries, invariances, and thermodynamic consistency at the architectural level substantially improves generalization and interpretability.
- Modular training strategies that decouple local constitutive learning from global equilibrium

enforcement could inspire more stable multi-field PINN formulations.

- Finally, the success of data-driven mechanics underscores the importance of open, standardized datasets and benchmark problems, resources that are still largely missing for PINNs in CSM.

Main points at a glance: The achievements of data-driven mechanics demonstrate that machine learning can add real value to solid mechanics when used judiciously within established numerical frameworks. The contrast with the slower progress of strong-form PINNs is instructive: where the latter struggle with stability and multi-field coupling, the former succeed by exploiting modularity, locality, and data availability. Future advances in physics-informed learning for CSM will likely emerge from hybrid strategies that inherit these advantages while reinstating the governing physics as soft or variational constraints.

GRAPH NEURAL NETWORKS AND MESH/ GRAPH REPRESENTATIONS

The challenges faced by strong-form PINNs in enforcing equilibrium, compatibility, and inter-field coupling have led many researchers to explore graph-based neural representations as a more structured alternative. Graph neural networks (GNNs) offer a natural way to encode the mesh topology of continuum discretizations, enabling direct message-passing of mechanical quantities (such as forces, displacements, and stresses) along the edges of finite element or finite volume meshes. In this framework, each node or element acts as a vertex in a computational graph, and the governing equations are implicitly represented through the structure of message propagation rather than through residual minimization.

Graphs as a generalization of meshes

Traditional discretizations in CSM—finite elements, finite volumes, boundary elements—can all be viewed as special cases of graphs endowed with geometric and physical attributes. Nodes represent degrees of freedom, edges carry connectivity and adjacency information, and element-wise or face-wise data encode constitutive and geometric features. This observation, emphasized in works such as Battaglia et al. (2018) and Pfaff et al. (2021), provides a direct route to constructing graph-based mechanical models without discarding the numerical infrastructure of classical solvers.

In a GNN, the evolution of node features h_i typically follows a message-passing rule

$$h_i^{(k+1)} = \phi\left(h_i^{(k)}, \sum_{j\in\mathcal{N}(i)} \psi(h_i^{(k)}, h_j^{(k)}, e_{ij})\right), \quad (7)$$

where $\boldsymbol{e_{ij}}$ represents the edge attributes (geometry, stiffness, or physical coupling), $\mathcal{N}(i)$ denotes the neighbours of node i, and ϕ and ψ denote differentiable functions such as MLPs. Through such updates, local equilibrium relations are propagated across the graph, mimicking the assembly of element-level contributions in the FEM. This message-passing paradigm provides an inductive bias that preserves physical locality, improves generalization across meshes, and mitigates the ill-conditioning that often hinders PINN training.

Learning mechanical behaviour through message passing

Graph-based learning has achieved notable successes in both forward and inverse problems in mechanics. For forward simulations, GNNs have been trained to approximate the solution operator of PDEs, mapping nodal positions, boundary conditions, and material parameters to displacements or stress fields, without explicitly solving the governing equations at runtime. Notable examples include the MeshGraphNets framework by Pfaff et al. (2021), which demonstrated accurate predictions of nonlinear elastic and fluid dynamics trajectories on complex meshes, and the physics-aware GNN models proposed by Horie and Mitsume (2022) and Sanchez-Gonzalez et al. (2020), where message functions are tailored to encode momentum conservation and frame invariance.

For inverse problems, GNNs have been used to infer material parameters or boundary conditions from partial field measurements (Zhao et al., 2022). Because graph architectures respect the locality and sparsity of mechanical systems, they can efficiently integrate observational data while preserving structural connectivity, leading to reconstructions that are both physically consistent

and computationally scalable. Unlike PINNs, which require solving a dense residual system during training, GNNs learn through local updates, resulting in linear or near-linear computational scaling with the number of nodes.

Physics-informed and hybrid GNNs

An emerging class of physics-informed GNNs (PI-GNNs) blends message-passing architectures with physically derived losses. Here, the network outputs field quantities that are constrained by global energy or residual principles:

$$\mathcal{L}_{\mathrm{PI}} = \sum_{e} \|\nabla \cdot \sigma_e + b_e\|^2 + \lambda \|\varepsilon_e - \nabla^s u_e\|^2, \qquad (8)$$

where ∇^s denotes the symmetric part of the gradient and each term above is evaluated locally on the graph edges or elements. This design permits partial enforcement of PDE constraints while retaining the flexible, data-driven feature learning of GNNs. Recent works (Thangamuthu et al., 2023; Zhao et al., 2024) show that such hybrids outperform both pure PINNs and purely data-driven GNNs when moderate amounts of data and partial physical knowledge are available. Moreover, the graph framework provides a convenient scaffold for coupling multiple physical fields (thermal–mechanical, electro–mechanical) through multiplexed or multilayer graphs, each representing a distinct field variable and connected by inter-layer message-passing channels.

Relation to finite element operators

A central insight unifying GNNs and FEM is that the stiffness matrix assembly can be viewed as a single message-passing layer, where messages are linear functions of nodal displacements weighted by local stiffness coefficients. Thus, classical FEM can be interpreted as a deterministic, linear GNN with known message functions. Learning-based extensions generalize these functions to nonlinear or data-dependent forms. This perspective (Brandstetter et al., 2022, 2023) reveals why GNNs naturally inherit the stability, symmetry, and conservation properties of FEM discretizations, offering a theoretically grounded path toward machine-learning-based solvers that remain physically interpretable.

Advantages and limitations

Compared with strong-form PINNs, GNN-based solvers exhibit several distinct advantages as noted below.

- **Structural consistency:** Graph connectivity mirrors mesh topology, ensuring that local equilibrium and compatibility relations are implicitly encoded.
- **Computational scalability:** Message passing operates locally, with complexity scaling linearly in the number of edges.
- **Data efficiency:** The inductive bias toward physical locality enables training on small datasets and generalization to unseen geometries or boundary conditions.
- **Multi-field coupling:** Graph multiplexing provides a principled mechanism to handle coupled PDE systems, such as thermoelasticity or poromechanics.

However, GNNs are not without limitations. The learned message functions are often opaque and may thus violate strict physical principles (e.g. energy conservation, momenta balances) unless constrained explicitly. Moreover, training stability can deteriorate for highly anisotropic or nearly incompressible materials where the stiffness contrast leads to numerical imbalance, challenges reminiscent of those in classical FEM. Finally, while GNNs offer strong generalization within a family of mesh topologies, their extrapolation to radically different geometries remains an open problem.

Outlook

Graph-based learning represents one of the most promising directions for integrating machine learning into computational mechanics. By treating traditional meshes as graphs, these methods provide a common language between numerical analysis and data-driven inference. Future research is likely to focus on:

- Developing variational GNNs that exactly reproduce FEM energy functionals through learnable but physically constrained message functions.
- Extending graph-based representations to higher-order continuum theories (e.g. strain gradient, micropolar or Cosserat mechanics) via enriched node and edge features.
- Incorporating probabilistic inference on graphs to handle uncertainty quantification and stochastic material behaviour.

Together, these developments may bridge the gap between data-driven learning and physics-based simulation, offering a principled route to scalable, interpretable, and thermodynamically consistent neural solvers.

POSSIBLE WAYS TO MAKE PHYSICS-INFORMED NETWORKS WORK FOR CSM

The preceding sections have emphasized the structural deficiencies of strong-form PINNs when applied to computational solid mechanics. Yet these shortcomings need not be terminal. A growing body of theoretical and algorithmic developments is beginning to reformulate the very foundation of physics-informed learning, seeking to reconcile the expressive flexibility of neural representations with the rigorous variational and energetic structure of continuum mechanics. This section outlines several possible and conceptually distinct pathways that hold promise for overcoming the current barriers. Each approach targets one or more of the critical limitations identified in Section 4, while preserving the data-driven adaptability that motivates neural formulations in the first place. The discussion remains deliberately schematic, as many of these ideas are still in active development.

Weak-form and variational reformulations

One of the most decisive shifts is in moving from the strong-form residual minimization of standard PINNs to weak-form or variational formulations. By projecting the governing equations onto appropriate test spaces, these approaches would inherit the stability, symmetry, and conditioning properties that make the finite element method reliable. When combined with neural trial functions that reproduce low-order polynomials (either globally or piecewise), they should be able to recover the exact finite element solution in the limit of sufficient basis enrichment. This hybridization provides an interpretable bridge between the deterministic variational framework of FEM and the statistical learning viewpoint of PINNs, while allowing for an essentially mesh-free representation and straightforward handling of complex geometries. In essence, weak-form learning reintroduces the energy-consistent backbone that strong-form networks lack.

Probabilistic and optimal-transport formulations

As we have already noted, a complementary direction reinterprets physics-informed learning through a probabilistic lens. Here, the evolving network output is viewed not as a deterministic field but as a probability measure that optimally evolves (say, through a gradient flow structure) to minimize a residual potential. The dynamics of this measure, formulated as a Wasserstein or Sinkhorn-regularized gradient flow, encode the PDE constraints as thermodynamic or optimal-transport dissipation mechanisms. Such formulations bring several benefits: they enforce global consistency through measure transport rather than pointwise residual minimization; they regularize training by controlling the entropy of the learned distribution; and they open the door to quantifying model uncertainty directly from the evolution of the measure. These probabilistic extensions have begun to demonstrate superior convergence and robustness, particularly when combined with neural normalizing flows that parameterize the transport maps.

Variance-reduced and energy-consistent loss formulations

A persistent difficulty in physics-informed training is the high variance of gradient estimates arising from disparate residual magnitudes across the domain and between loss components. Variance-reducing strategies motivated by stochastic optimization theory and generalized least squares, aim to stabilize training by adaptively controlling the contribution of each residual term. These techniques could be interpreted as introducing a local, data-driven preconditioning of the loss landscape, enabling more consistent convergence toward equilibrium states. When the variance control is coupled with energy-based loss components, the resulting formulation implicitly respects the virtual-work principle and thermodynamic admissibility, bridging the gap between optimization dynamics and physical structure.

Game-theoretic and multi-objective weighting

Another promising route treats the composite PINN loss as a multi-objective optimization problem in which the different physical components (e.g. force equilibrium, boundary conditions, constitutive laws) correspond to distinct agents. Nash or Stackelberg game formulations then determine the equilibrium weighting between these objectives in a self-consistent manner, obviating the need for ad hoc tuning of penalty coefficients. This game-theoretic weighting has the potential to restore balance among

competing constraints, particularly in coupled-field problems such as nearly incompressible elasticity or inelastic deformation. Beyond improving numerical stability, it may also provide an interpretable mechanism for quantifying the relative importance of each physical law during training, which may inform adaptive model refinement.

Towards hybrid and structure-preserving architectures

The final, and perhaps most far-reaching, avenue involves embedding the structural features of CSM directly into the architecture rather than the loss function. Examples include neural representations that respect objectivity and frame indifference, architectures that enforce compatibility by construction (Dhas et al., 2022; Kumar et al., 2024), or mixed formulations that pair neural displacements with analytically derived stress fields. At a broader level, combining the above ingredients (variational training, probabilistic regularization, variance control, and Nash-type balancing) could yield a new class of structure-preserving physics-informed networks. Such networks would not merely approximate PDE solutions but would embody the geometric and energetic symmetries that define solid mechanics itself.

Summary

These discussions mark a conceptual turning point: from heuristic penalty-based learning toward formulations that are variational, geometric, probabilistic, and game-theoretic in nature. Each of these directions addresses a distinct dimension of the gap between neural and finite-element thinking – conditioning, structure, uncertainty, and adaptability. Although empirical demonstrations are still emerging, the theoretical coherence of these ideas suggests that a synthesis of optimal-transport, weak-form, and structure-preserving strategies could eventually deliver the long-sought mesh-free and data-assimilative mechanics solver that PINNs initially promised.

A PROPOSED RESEARCH AGENDA AND ROAD-MAP

The preceding discussions highlight that the limited success of current PINN formulations in computational solid mechanics (CSM) arises not from a lack of expressive power in neural networks, but from a fundamental misalignment between the structure of continuum mechanics and the loss-based optimization frameworks that have so far dominated physics-informed learning. Closing this gap calls for a coordinated research agenda that integrates insights from numerical analysis, stochastic optimization, and nonequilibrium thermodynamics into the design of next-generation learning architectures. The roadmap outlined below aims to provide such a coherent direction.

Stage I: Foundational analysis and benchmarking

The first stage must focus on building a rigorous analytical and numerical foundation for physics-informed learning in CSM. Three immediate goals are essential.

- **Conditioning and convergence analysis:** Develop systematic criteria for assessing the well-posedness, conditioning, and convergence of PINN and weakform networks in comparison with classical Galerkin discretizations. This includes establishing spectral equivalence between neural residual minimization and variational projections, and quantifying the effect of network depth, activation nonlinearity, and sampling density on error propagation
- **Canonical benchmarks:** Establish a hierarchy of benchmark problems – ranging from linear elasticity and nearly incompressible materials to representative inelastic and multi-scale systems – serving as a common platform for evaluating neural and hybrid solvers under reproducible conditions.
- **Standardized training diagnostics:** Introduce physics-aware diagnostics that monitor balance of forces, energy consistency, and residual variance during training, enabling quantitative comparison across formulations and architectures.

Stage II: Variational, probabilistic, and hybrid formulations

Having established the analytical foundation, the next step is to operationalize the emerging paradigms introduced in Section 7.

- **Weak-form PINNs and energy-based training:** Formalize and test weak or mixed-form neural formulations that combine the variational rigor of the finite element method with the flexibility of deep approximators. These studies should

clarify the precise role of test and trial space design, and the possibility of local polynomial reproduction for improved convergence.

- **Probabilistic optimal-transport solvers:** Explore formulations where the neural model evolves a probability measure rather than a deterministic field, and where PDE satisfaction is recast as a Wasserstein or Sinkhorn-regularized gradient flow for the optimal transport of the evolving probability measures. This direction tethers physics-informed learning to measure-theoretic optimal transport, stochastic filtering, and thermodynamic consistency.
- **Variance-controlled and multi-objective optimization:** Implement and evaluate adaptive, variance-reduced, or Nash–multi-task-learning-based weighting strategies that can automatically balance residual terms across multiple physical constraints and scales.

Stage III: Integration with data and uncertainty

Once the physics-consistent backbone is established, the third stage should emphasize integration with experimental or simulation data and the quantification of uncertainty.

- **Data assimilation and inverse mechanics:** Couple probabilistic PINN formulations with Bayesian or filtering-based inference schemes (e.g. ensemble Kalman or particle-filter variants (Arulampalam et al., 2002; Evensen, 2003; Fearnhead and Künsch, 2018; Fujii, 2013; Hunt et al., 2007) to enable dynamic updating of model parameters and boundary conditions from sparse data.
- **Hybrid physics–data solvers:** Develop architectures that blend mechanistic weak-form constraints with neural constitutive surrogates, yielding hybrid solvers that remain predictive even outside the training domain.
- **Uncertainty quantification and reliability:** Introduce rigorous metrics for epistemic and aleatoric uncertainty in learned mechanical responses, enabling physics-informed neural models to be used in design, safety, and control contexts.

Stage IV: Toward structure-preserving and scalable implementations

The final stage envisages physics-informed learning as a practical and scalable alternative to conventional solvers for complex, multi-scale mechanics.

- **Structure-preserving architectures:** Embed geometric and energetic symmetries (such as frame invariance, objectivity, and thermodynamic consistency) directly into the network architecture, ensuring physically valid predictions by construction.
- **Parallel and differentiable implementations:** Exploit modern differentiable programming frameworks and GPU/TPU acceleration to scale variational PINN and probabilistic formulations to three-dimensional, large deformation, and path-dependent problems.
- **Open-source ecosystems:** Establish open benchmarks, repositories, and reproducible workflows for physics-informed learning in CSM, analogous to what exists in computational fluid dynamics and materials modeling.

Anticipated impact

If successfully executed, this research roadmap could transform the landscape of computational solid mechanics. By integrating variational structure, probabilistic reasoning, and adaptive optimization into a unified learning framework, it would enable solvers that are simultaneously mesh-free, data-assimilative, and physically interpretable. Beyond improved accuracy or efficiency, the greater promise lies in reshaping the conceptual interface between mechanics and machine learning: from algorithmic approximation towards a genuinely symmetry-abiding and thermodynamically consistent theory of learning in physical systems. Such a synthesis could bridge numerical analysis, stochastic thermodynamics, and information geometry, positioning CSM as a natural proving ground for the next generation of physics-informed artificial intelligence.

CONCLUSIONS

Over the past decade, neural network-based methods have profoundly influenced scientific computing and the modeling of physical systems. Within this broader landscape, however, the field of computational solid mechanics (CSM) stands

out as both a beneficiary and a skeptic. While data-driven approaches have achieved tangible progress – particularly in constitutive modeling, reduced-order simulations, and parameter identification – the application of physics-informed neural networks (PINNs) to solid mechanics has yielded outcomes far less decisive or useful. The reasons for this asymmetry are deeply structural: the multi-field, mixed, and often differential-algebraic nature of solid mechanical equations does not lend itself readily to strong-form residual minimization or direct neural regression of physical fields.

This review has examined these limitations from multiple viewpoints. The strong-form PINN, though conceptually elegant, suffers from severe conditioning, imbalance between field residuals, and difficulties in enforcing compatibility and boundary conditions across coupled variables. These issues have hindered its scalability and accuracy, especially in nonlinear or nearly incompressible elasticity, plasticity, and damage mechanics. Graph neural networks (GNNs) and operator-learning architectures, by contrast, have capitalized on the natural topological structure of discretized continua, demonstrating improved stability and generalization. Similarly, data-driven frameworks that hybridize classical solvers with neural representations of constitutive behavior have emerged as practical and physically interpretable alternatives. Yet, even these advances remain limited by their dependence on abundant training data and their partial adherence to thermodynamic principles.

Looking forward, the path to a robust neural paradigm for CSM likely lies in reconciling the structural rigor of continuum mechanics with the representational flexibility of deep networks. Three directions appear especially promising. First, variational and weak-form neural formulations can restore balance between kinematic and kinetic fields while improving numerical conditioning. Second, probabilistic and optimal-transport-based formulations reinterpret the learning problem as one of measure evolution, offering natural regularization, uncertainty quantification, and a direct link to thermodynamic consistency. Third, variance-controlled and energyconsistent loss functionals can stabilize training and encode the dissipation structure inherent to inelastic processes. Collectively, these approaches suggest that the next generation of physics-informed learning in solid mechanics will not merely mimic numerical solvers, but will instead realize a genuinely variational, probabilistic, and thermodynamically grounded synthesis of mechanics and learning theory.

In conclusion, the integration of neural computation into solid mechanics remains a frontier in both opportunity and challenge. Progress will depend not only on algorithmic innovation, but also on a deeper alignment between the mathematical structure of continuum physics and the learning architectures that seek to approximate it. By tracing the conceptual roots of current limitations and pointing toward emerging remedies, this review aims to help chart a coherent path toward physically faithful, data-efficient, and interpretable neural models of deformable solids.

Supplementary Materials

A few PINN-based solutions of problems in CSM and some other relevant material are available at https://github.com/Debasish-Roy-IISc/neural-CSM-review.

Conflict of interest

The authors have no conflicts of interest to declare. All co-authors have seen and agree with the contents of the manuscript and there is no financial interest to report.

REFERENCES

Antonion, K., Wang, X., Raissi, M., and Joshie, L. (2024). Machine Learning Through Physics–Informed Neural Networks: Progress and Challenges. *Academic Journal of Science and Technology,* 9(1):46–49.

Arora, R., Kakkar, P., Dey, B., and Chakraborty, A. (2022). Physicsinformed neural networks for modeling rate- and temperature-dependent plasticity. arXiv:2201.08363 [cond-mat].

Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on signal processing,* 50(2):174–188.

As'ad, F., Avery, P., and Farhat, C. (2022). A mechanics-informed artificial neural network approach in data-driven constitutive modeling. *International Journal for Numerical Methods in Engineering,* 123(12):2738–2759. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.6957.

Bai, J., Rabczuk, T., Gupta, A., Alzubaidi, L., and Gu, Y. (2023). A physicsinformed neural network technique based on a modified loss function for computational 2D and 3D solid mechanics. *Computational Mechanics,* 71(3):543–562.

Baker, N., Alexander, F., Bremer, T., Hagberg, A., Kevrekidis, Y., Najm, H., Parashar, M., Patra, A., Sethian, J., Wild, S., Willcox, K., and Lee, S. (2019). Workshop Report on Basic Research Needs for Scientific Machine Learning: Core Technologies for Artificial Intelligence. Technical Report None, 1478744.

Basir, S. (2023). Investigating and Mitigating Failure Modes in Physicsinformed Neural Networks (PINNs). *Communications in Computational Physics,* 33(5):1240–1269. arXiv:2209.09988 [cs].

Basir, S. and Senocak, I. (2022). Critical Investigation of Failure Modes in Physicsinformed Neural Networks. arXiv:2206.09961 [cs].

Battaglia, P. W., Hamrick, J. B., Bapst, V., Sanchez-Gonzalez, A., Zambaldi, V., Malinowski, M., Tacchetti, A., Raposo, D., Santoro, A., Faulkner, R., Gulcehre, C., Song, F., Ballard, A., Gilmer, J., Dahl, G., Vaswani, A., Allen, K., Nash, C., Langston, V., Dyer, C., Heess, N., Wierstra, D., Kohli, P., Botvinick, M., Vinyals, O., Li, Y., and Pascanu, R. (2018). Relational inductive biases, deep learning, and graph networks. arXiv:1806.01261 [cs].

Baydin, A. G., Pearlmutter, B. A., Radul, A. A., and Siskind, J. M. (2018). Automatic Differentiation in Machine Learning: a Survey. *Journal of Machine Learning Research,* 18(153):1–43.

Bhatia, A., Saggi, M., and Kais, S. (2025). Application of quantum-inspired tensor networks to optimize federated learning systems. *Quantum Machine Intelligence.*

Bleyer, J. (2025). Learning elastoplasticity with implicit layers. *Computer Methods in Applied Mechanics and Engineering,* 444:118145.

Brandstetter, J., Welling, M., and Worrall, D. E. (2022). Lie Point Symmetry Data Augmentation for Neural PDE Solvers. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., and Sabato, S., editors, *Proceedings of the 39th International Conference on Machine Learning,* volume 162 of *Proceedings of Machine Learning Research,* pages 2241–2256. PMLR.

Brandstetter, J., Worrall, D., and Welling, M. (2023). Message Passing Neural PDE Solvers. arXiv:2202.03376 [cs].

Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., and Vandergheynst, P. (2017). Geometric Deep Learning: Going beyond Euclidean data. *IEEE Signal Processing Magazine,* 34(4):18–42.

Brunton, S. L. and Kutz, J. N. (2023). Machine Learning for Partial Differential Equations. arXiv:2303.17078 [cs].

Cai, S., Mao, Z., Wang, Z., Yin, M., and Karniadakis, G. E. (2021a). Physicsinformed neural networks (PINNs) for fluid mechanics: a review. *Acta Mechanica Sinica,* 37(12):1727–1738.

Cai, S., Wang, Z., Wang, S., Perdikaris, P., and Karniadakis, G. E. (2021b). Physics- Informed Neural Networks for Heat Transfer Problems. *Journal of Heat Transfer,* 143(060801).

Cai, Z., Chen, J., Liu, M., and Liu, X. (2020). Deep least-squares methods: An unsupervised learning-based numerical method for solving elliptic PDEs. *Journal of Computational Physics,* 420:109707.

Cao, Q., Goswami, S., and Karniadakis, G. E. (2024). Laplace neural operator for solving differential equations. *Nature Machine Intelligence,* 6(6):631–640.

Chen, W., Wang, Q., Hesthaven, J. S., and Zhang, C. (2021). Physics-informed machine learning for reduced-order modeling of nonlinear problems. *Journal of Computational Physics,* 446:110666.

Cho, J., Nam, S., Yang, H., Yun, S.-B., Hong, Y., and Park, E. (2023). Separable Physics-Informed Neural Networks. In Oh, A., Naumann, T., Globerson, A., Saenko, K., Hardt, M., and Levine, S., editors, *Advances in Neural Information Processing Systems,* volume 36, pages 23761–23788. Curran Associates, Inc.

Corso, G., Stark, H., Jegelka, S., Jaakkola, T., and Barzilay, R. (2024). Graph neural networks. *Nature Reviews Methods Primers,* 4(1):17. Publisher: Nature Publishing Group.

Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., and Piccialli, F. (2022). Scientific Machine Learning Through Physics–Informed Neural Networks: Where we are and What's Next. *Journal of Scientific Computing,* 92(3):88.

Cuturi, M. (2013). Sinkhorn distances: Lightspeed computation of optimal transport. In Burges, C., Bottou, L., Welling, M., Ghahramani, Z., and Weinberger, K., editors, *Advances in Neural Information Processing Systems,* volume 26. Curran Associates, Inc.

Daw, A., Bu, J., Wang, S., Perdikaris, P., and Karpatne, A. (2023). Mitigating Propagation Failures in Physics-informed Neural Networks using Retain-Resample- Release (R3) Sampling. arXiv:2207.02338 [cs].

De Ryck, T. and Mishra, S. (2022). Error analysis for physics-informed neural networks (PINNs) approximating Kolmogorov PDEs. *Advances in Computational Mathematics,* 48(6):79.

Dettmer, W. G., Muttio, E. J., Alhayki, R., and Perić, D. (2024). A framework for neural network based constitutive modelling of inelastic materials. *Computer Methods in Applied Mechanics and Engineering,* 420:116672.

Dey, A., Welschinger, F., Schneider, M., Kobler, J., and Bohlke, T. (2024). On the effectiveness of deep material networks for the multi-scale virtual characterization of short fiber-reinforced thermoplastics under highly nonlinear load cases. *Archive of Applied Mechanics.*

Dhas, B., N., J. K., Roy, D., and Reddy, J. N. (2022). A mixed variational principle in nonlinear elasticity using Cartan's moving frames and implementation with finite element exterior calculus. *Computer Methods in Applied Mechanics and Engineering,* 393:114756.

Eggersmann, R., Kirchdoerfer, T., Reese, S., Stainier, L., and Ortiz, M. (2019). Model-Free Data-Driven inelasticity. *Computer Methods in Applied Mechanics and Engineering,* 350:81–99.

Eghbalian, M., Pouragha, M., and Wan, R. (2023). A physics-informed deep neural network for surrogate modeling in classical elasto-plasticity. *Computers and Geotechnics,* 159:105472.

Evensen, G. (2003). The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics,* 53(4):343–367.

Farea, A., Yli-Harja, O., and Emmert-Streib, F. (2024). Understanding Physics- Informed Neural Networks: Techniques, Applications, Trends, and Challenges. *AI,* 5(3):1534–1557. Publisher: Multidisciplinary Digital Publishing Institute.

Fearnhead, P. and Künsch, H. R. (2018). Particle filters and data assimilation. *Annual Review of Statistics and Its Application,* 5(1):421–449.

Fresca, S., Dede', L., and Manzoni, A. (2021). A Comprehensive Deep Learning- Based Approach to Reduced Order Modeling of Nonlinear Time-Dependent Parametrized PDEs. *Journal of Scientific Computing,* 87(2):61.

Fresca, S. and Manzoni, A. (2022). POD-DL-ROM: Enhancing deep learningbased reduced order models for nonlinear parametrized PDEs by proper orthogonal decomposition. *Computer Methods in Applied Mechanics and Engineering,* 388:114181.

Fuhg, J. N., Anantha Padmanabha, G., Bouklas, N., Bahmani, B., Sun, W., Vlassis, N. N., Flaschel, M., Carrara, P., and De Lorenzis, L. (2025). A Review on Data-Driven Constitutive Laws for Solids. *Archives of Computational Methods in Engineering,* 32(3):1841–1883.

Fuhg, J. N. and Bouklas, N. (2022). The mixed Deep Energy Method for resolving concentration features in finite strain hyperelasticity. *Journal of Computational Physics,* 451:110839.

Fujii, K. (2013). Extended kalman filter. *Refernce Manual,* 14:41.

Gao, B., Yao, R., and Li, Y. (2025). Physics-informed neural networks with adaptive loss weighting algorithm for solving partial differential equations. *Computers & Mathematics with Applications.*

Gao, H., Zahr, M. J., and Wang, J.-X. (2022). Physics-informed graph neural Galerkin networks: A unified framework for solving PDE-governed forward and inverse problems. *Computer Methods in Applied Mechanics and Engineering,* 390:114502.

Gao, Z., Yan, L., and Zhou, T. (2023). Failure-Informed Adaptive Sampling for PINNs. *SIAM Journal on Scientific Computing,* 45(4):A1971–A1994. Publisher: Society for Industrial and Applied Mathematics.

Garanger, K., Kraus, J., & Rimoli, J. J. (2024). Symmetry-enforcing neural networks with applications to constitutive modeling. *Extreme Mechanics Letters.*

Gladstone, R. J., Nabian, M. A., Sukumar, N., Srivastava, A., and Meidani, H. (2025). FO-PINN: A First-Order formulation for Physics-Informed Neural Networks. *Engineering Analysis with Boundary Elements,* 174:106161.

Grossmann, T. G., Komorowska, U. J., Latz, J., and Schönlieb, C.-B. (2024). Can physics-informed neural networks beat the finite element method? *IMA Journal of Applied Mathematics,* 89(1):143–174.

Haghighat, E., Amini, D., and Juanes, R. (2022). Physics-informed neural network simulation of multiphase poroelasticity using stress-split sequential training. *Computer Methods in Applied Mechanics and Engineering,* 397:115141.

Haghighat, E., Raissi, M., Moure, A., Gomez, H., and Juanes, R. (2021). A physicsinformed deep learning framework for inversion and surrogate modeling in solid mechanics. *Computer Methods in Applied Mechanics and Engineering,* 379:113741.

Hao, Z., Liu, S., Zhang, Y., Ying, C., Feng, Y., Su, H., and Zhu, J. (2023). Physics- Informed Machine Learning: A Survey on Problems, Methods and Applications. arXiv:2211.08064 [cs].

Hartman, D. and Mestha, L. K. (2017). A deep learning framework for model reduction of dynamical systems. In *2017 IEEE Conference on Control Technology and Applications (CCTA),* pages 1917–1922.

Horie, M. and MITSUME, N. (2022). Physics-Embedded Neural Networks: Graph Neural PDE Solvers with Mixed Boundary Conditions. In Koyejo, S., Mohamed, S., Agarwal, A., Belgrave, D., Cho, K., and Oh, A., editors, *Advances in Neural Information Processing Systems,* volume 35, pages 23218–23229. Curran Associates, Inc.

Hu, H., Qi, L., and Chao, X. (2024a). Physics-informed Neural Networks (PINN) for computational solid mechanics: Numerical frameworks and applications. *Thin-Walled Structures,* 205:112495.

Hu, Z., Shi, Z., Karniadakis, G. E., and Kawaguchi, K. (2024b). Hutchinson Trace Estimation for high-dimensional and high-order Physics-Informed Neural Networks. *Computer Methods in Applied Mechanics and Engineering,* 424:116883.

Hunt, B. R., Kostelich, E. J., and Szunyogh, I. (2007). Efficient data assimilation for spatiotemporal chaos: A local ensemble transform kalman filter. *Physica D: Nonlinear Phenomena,* 230(1-2):112–126.

Jog, C. S. (2015). *Continuum mechanics: foundations and applications of mechanics.* Cambridge-IISc Press.

Karnakov, P., Litvinov, S., and Koumoutsakos, P. (2024). Solving inverse problems in physics by optimizing a discrete loss: Fast and accurate learning without neural networks. *PNAS Nexus,* 3(1):pgae005.

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

Khadijeh, M., Cerqueglini, V., Kasbergen, C., Erkens, S., and Varveri, A. (2025). Multistage physics informed neural network for solving coupled multiphysics problems in material degradation and fluid dynamics. *Engineering with Computers.*

Kharazmi, E., Zhang, Z., and Karniadakis, G. E. (2021). hp-VPINNs: Variational physics-informed neural networks with domain decomposition. *Computer Methods in Applied Mechanics and Engineering,* 374:113547.

Khorrami, M. S., Mianroodi, J. R., Siboni, N. H., Goyal, P., Svendsen, B., Benner, P., and Raabe, D. (2023). An artificial neural network for surrogate modeling of stress fields in viscoplastic polycrystalline materials. *npj Computational Materials,* 9(1):37. Publisher: Nature Publishing Group.

Kim, D. and Lee, J. (2024). A Review of Physics Informed Neural Networks for Multiscale Analysis and Inverse Problems. *Multiscale Science and Engineering,* 6(1):1–11.

Kipf, T. N. and Welling, M. (2017). Semi-Supervised Classification with Graph Convolutional Networks. arXiv:1609.02907 [cs].

Kirchdoerfer, T. and Ortiz, M. (2016). Data-driven computational mechanics. *Computer Methods in Applied Mechanics and Engineering.*

Klawonn, A., Lanser, M., and Weber, J. (2024). Machine learning and domain decomposition methods - a survey. *Computational Science and Engineering.*

Korzeniowski, T. F. and Weinberg, K. (2021). A multi-level method for datadriven finite element computations. *Computer Methods in Applied Mechanics and Engineering,* 379:113740.

Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., and Anandkumar, A. (2023). Neural Operator: Learning Maps Between Function Spaces With Applications to PDEs. *Journal of Machine Learning Research,* 24(89):1–97.

Krishnapriyan, A., Gholami, A., Zhe, S., Kirby, R., and Mahoney, M. W. (2021). Characterizing possible failure modes in physics-informed neural networks. In Ranzato, M., Beygelzimer, A., Dauphin, Y., Liang, P. S., and Vaughan, J. W., editors, *Advances in Neural Information Processing Systems,* volume 34, pages 26548–26560. Curran Associates, Inc.

Kumar, N. J., Reddy, J. N., Srinivasa, A. R., & Roy, D. (2024). A new mixed variational approach for kirchhoff shells and C0 discretization with finite element exterior calculus. *Computer Methods in Applied Mechanics and Engineering,* 432:117351.

Lawal, Z. K., Yassin, H., Lai, D. T. C., and Che Idris, A. (2022). Physics-Informed Neural Network (PINN) Evolution and Beyond: A Systematic Literature Review and Bibliometric Analysis. *Big Data and Cognitive Computing,* 6(4):140. Publisher: Multidisciplinary Digital Publishing Institute.

Le-Duc, T., Lee, S., Nguyen-Xuan, H., and Lee, J. (2024). A hierarchically normalized physics-informed neural network for solving differential equations: Application for solid mechanics problems. *Engineering Applications of Artificial Intelligence,* 133:108400.

Lee, K. and Carlberg, K. T. (2020). Model reduction of dynamical systems on nonlinear manifolds using deep convolutional autoencoders. *Journal of Computational Physics,* 404:108973.

Li, H., Miao, Y., Khodaei, Z. S., and Aliabadi, M. H. (2025). Finite-PINN: A physics-informed neural network with finite geometric encoding for solid mechanics. *Journal of the Mechanics and Physics of Solids,* 203:106222.

Li, S. and Feng, X. (2022). Dynamic Weight Strategy of Physics-Informed Neural Networks for the 2D Navier–Stokes Equations. *Entropy,* 24(9):1254.

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., and Anandkumar, A. (2021). Fourier Neural Operator for Parametric Partial Differential Equations. arXiv:2010.08895 [cs].

Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Stuart, A., Bhattacharya, K., and Anandkumar, A. (2020). Multipole Graph Neural Operator for Parametric Partial Differential Equations. In Larochelle, H., Ranzato, M., Hadsell, R., Balcan, M. F., and Lin, H., editors, *Advances in Neural Information Processing Systems,* volume 33, pages 6755–6766. Curran Associates, Inc.

Liang, L., Liu, M., Martin, C., and Sun, W. (2018). A deep learning approach to estimate stress distribution: a fast and accurate surrogate of finite-element analysis. *Journal of The Royal Society Interface,* 15(138):20170844. Publisher: Royal Society.

Linka, K., Hillgärtner, M., Abdolazizi, K. P., Aydin, R. C., Itskov, M., and Cyron, C. J. (2021). Constitutive artificial neural networks: A fast and general approach to predictive data-driven constitutive modeling by deep learning. *Journal of Computational Physics,* 429:110010.

Lippe, P., Veeling, B., Perdikaris, P., Turner, R., and Brandstetter, J. (2023). PDERefiner: Achieving Accurate Long Rollouts with Neural PDE Solvers. In Oh, A., Naumann, T., Globerson, A., Saenko, K., Hardt, M., and Levine, S., editors, *Advances in Neural Information Processing Systems,* volume 36, pages 67398– 67433. Curran Associates, Inc.

Lu, L., Jin, P., Pang, G., Zhang, Z., and Karniadakis, G. E. (2021a). Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators. *Nature Machine Intelligence,* 3(3):218–229.

Lu, L., Meng, X., Mao, Z., and Karniadakis, G. E. (2021b). DeepXDE: A Deep Learning Library for Solving Differential Equations. *SIAM Review,* 63(1):208– 228. Publisher: Society for Industrial and Applied Mathematics.

Luo, K., Zhao, J., Wang, Y., Li, J., Wen, J., Liang, J., Soekmadji, H., and Liao, S. (2025). Physics-informed neural networks for PDE problems: a comprehensive review. *Artificial Intelligence Review,* 58(10):323.

Maddu, S., Sturm, D., Müller, C. L., and Sbalzarini, I. F. (2022). Inverse Dirichlet weighting enables reliable training of physics informed neural networks. *Machine Learning: Science and Technology,* 3(1):015026.

Maldonado, S., Vairetti, C., Jara, K., Carrasco, M., and Lopez, J. (2023). Owadapt: An adaptive loss function for deep learning using owa operators. *Knowledge-Based Systems.*

Manav, M., Molinaro, R., Mishra, S., and De Lorenzis, L. (2024). Phase-field modeling of fracture with physics-informed deep learning. *Computer Methods in Applied Mechanics and Engineering,* 429:117104.

Markidis, S. (2021). The Old and the New: Can Physics-Informed Deep-Learning Replace Traditional Linear Solvers? *Frontiers in Big Data,* 4:669097.

Maurizi, M., Gao, C., and Berto, F. (2022). Predicting stress, strain and deformation fields in materials and structures with graph neural networks. *Scientific Reports,* 12(1):21834. Publisher: Nature Publishing Group.

Mishra, S. and Molinaro, R. (2022). Estimates on the generalization error of physicsinformed neural networks for approximating a class of inverse problems for PDEs. *IMA Journal of Numerical Analysis,* 42(2):981–1022.

Mitusch, S., Funke, S., and Kuchta, M. (2021). Hybrid fem-nn models: Combining artificial neural networks with the finite element method. *Journal of Computational Physics.*

Mouratidou, A. D., Drosopoulos, G. A., and Stavroulakis, G. E. (2024). Ensemble of physics-informed neural networks for solving plane elasticity problems with examples. *Acta Mechanica,* 235(11):6703–6722.

Nabian, M. A., Gladstone, R. J., and Meidani, H. (2021). Efficient training of physics-informed neural networks via importance sampling. *Computer-Aided Civil and Infrastructure Engineering,* 36(8):962–977.

Nguyen-Thanh, V. M., Zhuang, X., and Rabczuk, T. (2020). A deep energy method for finite deformation hyperelasticity. *European Journal of Mechanics - A/Solids,* 80:103874.

Pfaff, T., Fortunato, M., Sanchez-Gonzalez, A., and Battaglia, P. (2021). Learning Mesh-Based Simulation with Graph Networks. *In International Conference on Learning Representations.*

Prume, E., Reese, S., and Ortiz, M. (2023). Model-free Data-Driven inference in computational mechanics. *Computer Methods in Applied Mechanics and Engineering,* 403:115704.

Pu, R. and Feng, X. (2022). Physics-Informed Neural Networks for Solving Coupled Stokes–Darcy Equation. *Entropy,* 24(8):1106. Publisher: Multidisciplinary Digital Publishing Institute.

Raissi, M., Perdikaris, P., and Karniadakis, G. (2019). 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.

Rezaei, S., Moeineddin, A., and Harandi, A. (2023). Learning solution of nonlinear constitutive material models using physics-informed neural networks: COMMPINN. arXiv:2304.06044 [cs].

Roy, A. M., Bose, R., Sundararaghavan, V., and Arróyave, R. (2023). Deep learningaccelerated computational framework based on Physics Informed Neural Network for the solution of linear elasticity. *Neural Networks,* 162:472–489.

Samaniego, E., Anitescu, C., Goswami, S., Nguyen-Thanh, V. M., Guo, H., Hamdia, K., Zhuang, X., and Rabczuk, T. (2020). An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. *Computer Methods in Applied Mechanics and Engineering,* 362:112790.

Sanchez-Gonzalez, A., Godwin, J., Pfaff, T., Ying, R., Leskovec, J., and Battaglia, P. (2020). Learning to Simulate Complex Physics with Graph Networks. In *Proceedings of the 37th International Conference on Machine Learning,* pages 8459–8468. PMLR. ISSN: 2640-3498.

Sharma, P., Evans, L., Tindall, M., and Nithiarasu, P. (2023). Stiff-PDEs and Physics-Informed Neural Networks. *Archives of Computational Methods in Engineering,* 30(5):2929–2958.

Sirignano, J. and Spiliopoulos, K. (2018). DGM: A deep learning algorithm for solving partial differential equations. *Journal of Computational Physics,* 375:1339– 1364.

Song, C., Park, Y., and Kang, M. (2024). How does PDE order affect the convergence of PINNs? In Globerson, A., Mackey, L., Belgrave, D., Fan, A., Paquet, U., Tomczak, J., and Zhang, C., editors, *Advances in Neural Information Processing Systems,* volume 37, pages 73–131. Curran Associates, Inc.

Thangamuthu, A., Kumar, G., Bishnoi, S., Bhattoo, R., Krishnan, N. M. A., and Ranu, S. (2023). Unravelling the Performance of Physics-informed Graph Neural Networks for Dynamical Systems. arXiv:2211.05520 [cs].

Toscano, J. D., Oommen, V., Varghese, A. J., Zou, Z., Ahmadi Daryakenari, N., Wu, C., and Karniadakis, G. E. (2025). From PINNs to PIKANs: recent advances in physics-informed machine learning. *Machine Learning for Computational Science and Engineering,* 1(1):15.

Uriarte, C., Bastidas, M., Pardo, D., Taylor, J., and Rajas, S. (2025). Optimizing variational physics-informed neural networks using least squares. *Computers and Mathematics with Applications.*

Vlassis, N. N., Ma, R., and Sun, W. (2020). Geometric deep learning for computational mechanics Part I: anisotropic hyperelasticity. *Computer Methods in Applied Mechanics and Engineering,* 371:113299.

Wang, F., Zhai, W., Zhao, S., and Man, J. (2025). A novel unsupervised PINN framework with dynamically self-adaptive strategy for solid mechanics. *Journal of Computational Physics,* 542:114373.

Wang, J., Mo, Y. L., Izzuddin, B., and Kim, C.-W. (2023). Exact Dirichlet boundary Physics-informed Neural Network EPINN for solid mechanics. *Computer Methods in Applied Mechanics and Engineering,* 414:116184.

Wang, L., Liu, G., Wang, G., and Zhang, K. (2024a). M-PINN: A mesh-based physics-informed neural network for linear elastic problems in solid mechanics. *International Journal for Numerical Methods in Engineering,* 125(9):e7444. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/nme.7444.

Wang, S., Sankaran, S., and Perdikaris, P. (2024b). Respecting causality for training physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering,* 421:116813.

Wang, S., Teng, Y., and Perdikaris, P. (2021a). Understanding and Mitigating Gradient Flow Pathologies in Physics-Informed Neural Networks. *SIAM Journal on Scientific Computing,* 43(5):A3055–A3081. _eprint: https://doi.org/10.1137/20M1318043.

Wang, S., Wang, H., and Perdikaris, P. (2021b). On the eigenvector bias of Fourier feature networks: From regression to solving multi-scale PDEs with physics-informed neural networks. *Computer Methods in Applied Mechanics and Engineering,* 384:113938.

Wang, S., Yu, X., and Perdikaris, P. (2022). When and why PINNs fail to train: A neural tangent kernel perspective. *Journal of Computational Physics,* 449:110768.

Wang, Y., Yao, Y., Guo, J., and Gao, Z. (2024c). A practical PINN framework for multi-scale problems with multi-magnitude loss terms. *Journal of Computational Physics,* 510:113112.

Weinan, E. and Yu, B. (2018). The Deep Ritz Method: A Deep Learning- Based Numerical Algorithm for Solving Variational Problems. *Communications in Mathematics and Statistics,* 6(1):1–12.

Wu, P., Sun, J., Chang, X., Zhang, W., Arcucci, R., Guo, Y., and Pain, C. C. (2020). Data-driven reduced order model with temporal convolutional neural network. *Computer Methods in Applied Mechanics and Engineering,* 360:112766.

Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., and Yu, P. S. (2021). A Comprehensive Survey on Graph Neural Networks. *IEEE Transactions on Neural Networks and Learning Systems,* 32(1):4–24.

Xiang, Z., Peng, W., Liu, X., and Yao, W. (2022). Self-adaptive loss balanced Physics-informed neural networks. *Neurocomput.,* 496(C):11–34.

Xu, K., Hu, W., Leskovec, J., and Jegelka, S. (2019). How Powerful are Graph Neural Networks? arXiv:1810.00826 [cs].

Zang, Y., Bao, G., Ye, X., and Zhou, H. (2020). Weak adversarial networks for high-dimensional partial differential equations. *Journal of Computational Physics,* 411:109409.

Zhao, Q., Lindell, D. B., and Wetzstein, G. (2022). Learning to solve PDEconstrained inverse problems with graph networks. In Chaudhuri, K., Jegelka, S., Song, L., Szepesvari, C., Niu, G., and Sabato, S., editors, *Proceedings of the 39th International Conference on Machine Learning,* volume 162 of *Proceedings of Machine Learning Research,* pages 26895–26910. PMLR.

Zhao, Y., Li, H., Zhou, H., Attar, H. R., Pfaff, T., and Li, N. (2024). A review of graph neural network applications in mechanics-related domains. *Artificial Intelligence Review,* 57(11):315.

Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., and Sun, M. (2020). Graph neural networks: A review of methods and applications. *AI Open,* 1:57–81.

Zhu, Y., Zabaras, N., Koutsourelakis, P.-S., and Perdikaris, P. (2019). Physics- Constrained Deep Learning for High-dimensional Surrogate Modeling and Uncertainty Quantification without Labeled Data. *Journal of Computational Physics,* 394:56–81. arXiv:1901.06314 [physics].