



City Research Online

City St George's, University of London

Citation: Ani, A., Nakka, R., Subhash, G., Molinari, J-F. & Ponnusami, S. A. (2026). Machine learning for computational fracture and damage mechanics—Status and perspectives. *Engineering Fracture Mechanics*, 332, 111778. doi: 10.1016/j.engfracmech.2025.111778

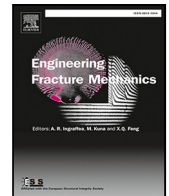
This is the published version of the paper.

This version of the publication may differ from the final published version. To cite this item please consult the publisher's version.

Permanent repository link: <https://openaccess.city.ac.uk/id/eprint/36408/>

Link to published version: <https://doi.org/10.1016/j.engfracmech.2025.111778>

Copyright and Reuse: Copyright and Moral Rights remain with the author(s) and/or copyright holders. Copies of full items can be used for personal research or study, educational, or not-for-profit purposes without prior permission or charge, unless otherwise indicated, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way. For full details of reuse please refer to [City Research Online policy](#).



Review

Machine learning for computational fracture and damage mechanics— Status and perspectives

Allamaprabhu Ani ^{a, }, Rajesh Nakka ^a, Ghatu Subhash ^c, Jean-François Molinari ^d,
Sathiskumar Anusuya Ponnusami ^{a, b, *}

^a Department of Engineering, City St. George's, University of London, United Kingdom

^b School of Engineering and Materials Science, Queen Mary University of London, United Kingdom

^c Department of Mechanical and Aerospace Engineering, University of Florida, USA

^d Institute of Civil Engineering, Institute of Materials, École Polytechnique Fédérale de Lausanne (EPFL), Switzerland

ARTICLE INFO

Keywords:

Computational fracture mechanics

Machine learning

Damage modelling

Finite element simulation

ABSTRACT

Fracture and damage mechanics have evolved remarkably from simple, yet useful Linear Elastic Fracture Mechanics (LEFM) to relatively modern techniques such as Cohesive Zone Model (CZM) and phase-field approaches. The advent of computational power allowed researchers and engineers to conduct high-fidelity numerical simulations to model complex fracture mechanisms in advanced materials and structures. Nonetheless, large-scale fracture simulations remain computationally intensive, particularly under loading conditions such as impact and extreme environments. In this context, Machine Learning (ML) techniques have seen a surge in their use for mechanics and computational simulations. In this perspective article, we review the existing research landscape in the recent literature on the application of ML to fracture and damage modelling across different material and structural classes. Specific focus is placed on classifying the ML approaches adopted to model or predict fracture behaviour, followed by an extensive discussion on the challenges and limitations of such approaches. Future directions are proposed with an emphasis on the generality, interpretability and reliability of the ML models. We believe the article serves as a guidance document for engineers and scientists involved in the developmental process of Artificial Intelligence (AI)-driven fracture modelling tools.

Contents

1.	Introduction	2
2.	Overview of fracture modelling techniques	3
2.1.	Continuum damage mechanics modelling	3
2.2.	Phase field damage modelling	4
2.3.	Cohesive zone modelling	4
2.4.	Discrete element method	4
3.	Overview of machine learning approaches	4
3.1.	Machine learning methods	4
3.2.	Neural networks	5
3.2.1.	Data	5
3.2.2.	Neural network architecture	6
3.2.3.	Loss functions and metrics	6

* Corresponding author at: Department of Engineering, City St. George's, University of London, United Kingdom.

E-mail address: s.a.ponnusami@qmul.ac.uk (S.A. Ponnusami).

3.2.4.	Optimisation algorithms.....	7
4.	Status review of AI applied to fracture problems.....	7
4.1.	Data-driven emulators and surrogates.....	8
4.2.	Graph neural networks.....	10
4.3.	CNN and U-Net based approaches.....	12
4.3.1.	Sequential frameworks.....	12
4.3.2.	Iterative spatiotemporal prediction.....	14
4.3.3.	Tackling data scarcity and interpretability.....	15
4.4.	Generative models.....	17
4.5.	Data-driven reduced-order modelling.....	18
4.6.	Physics-informed neural networks.....	19
4.7.	Surrogate modelling with neural operators.....	24
4.8.	Material-specific models.....	26
5.	Challenges and limitations in ML-driven fracture modelling.....	28
6.	Way forward: Perspectives and opportunities.....	29
7.	Conclusions.....	30
	CRedit authorship contribution statement.....	30
	Declaration of competing interest.....	30
	Acknowledgement.....	30
	Appendix.....	30
	Data availability.....	31
	References.....	31

1. Introduction

Understanding fracture and damage behaviour is an integral requirement of the design of advanced materials and structures. Experimental and computational methods have been extensively used over the past decades to establish the underlying mechanics of fracture, in turn aiding the design process. While experimental methods have widely been used for characterising fracture properties and mechanisms in materials, computational techniques primarily based on the finite element method are used for predicting fracture behaviour to inform structural design. Despite the success of computational modelling in simulating fracture in advanced materials and intricate structural configurations, the associated computational cost and modelling challenges, especially for large-scale problems, remain high and often inhibit practical implementation in industrial scenarios.

Further, loading conditions such as impact and other extreme environments make the design process non-trivial as it involves complex time and rate-dependent damage evolution mechanisms. These mechanisms demand a thorough understanding of material behaviour under diverse experimental setups and high-fidelity, large-scale impact simulations. Real-life design against fracture is still a challenge, primarily due to the coupled influence of multi-physics and multi-scale (both spatially and temporally) processes and insufficient understanding of the failure evolution in extreme environments. Again, high-fidelity deterministic computational techniques such as the Finite Element Analysis (FEA) serve as excellent design tools; however, their computational costs and time in predicting the failure prohibit their full-scale deployment.

From the above perspective, the use of ML architectures for computational mechanics has been exponentially increasing in the last decade. The key enablers have been primarily the advent of computational power and a whole range of capabilities offered by such intelligent data-driven approaches.

Several efforts have been taken, as seen in the recent literature, to develop ML-enabled techniques to achieve faster simulations or physical predictions. The recent detailed review article of FEA summarises the 80 years of developments with a special section on mechanistic data science (MDS)-based and ML-driven FEA approaches [1]. Researchers have applied data-driven machine-learning techniques for accelerated multi-scale modelling [2], micromechanical homogenisation [3], constitutive modelling of composite materials [4] and damage detection. Numerical solutions of high-dimensional partial differential equations (PDEs) using physics-informed or enhanced neural networks [5–7] have also gained traction due to their suitability in cases where there is minimal or even no prior training data. Multiple works have focused on ML-driven fracture mechanics, but are heavily focused on either damage monitoring [8] or property characterisation [9]. A recent review by Jin, Zhang and Espinosa [10] provides an overview of machine learning algorithms and terminologies used in experimental solid mechanics. Their review highlights the increasing application of ML techniques in fracture mechanics, particularly for predicting fracture toughness and analysing complex crack propagation paths across various materials. There is a growing body of research focused on applying ML approaches to fracture simulations and predictive modelling across various material systems and structural configurations. This article aims to examine the range of ML architectures explored in the literature for modelling and simulating fracture evolution.

The article is organised as follows: In Section 2, we present a short overview of the computational fracture mechanics techniques and general ML landscape covering the various ML architectures including classical regression algorithms to the latest generative ML architectures in Section 3. These two short introductory sections set the stage for Section 4 dealing with a detailed status review of fracture mechanics problems addressed so far in the literature using ML methods. Challenges and limitations of the proposed models are extensively discussed in Section 5, followed by the authors' perspective on the way forward for research in this domain of ML-driven computational fracture and damage mechanics in Section 6.

Nomenclature

Symbol	Description
<i>Fracture Mechanics</i>	
σ	Cauchy Stress tensor
ϵ	Strain tensor
C	Elastic stiffness tensor
$D, d(\mathbf{x})$	Scalar damage/phase-field variables, $D, d \in [0, 1]$
ℓ	Length-scale parameter for crack regularisation
K_I, K_{II}	Mode I and Mode II Stress Intensity Factors (SIFs)
u_1, u_2	Enriched displacement components
$f_{i,I}, f_{i,II}$	Crack tip enrichment functions
<i>Machine Learning & Neural Networks</i>	
$\mathcal{M}(\theta), \mathcal{N}(\theta)$	ML model/neural network with parameters θ
$\theta = \{\mathbf{W}^l, \mathbf{b}^l\}$	Trainable weights and biases for all layers
$\mathbf{z}^l, \mathbf{a}^l$	Pre- and post-activation vectors at layer l
$\sigma(\cdot)$	Element-wise nonlinear activation function
$h(\mathbf{x}, t)$	Hidden feature representation evolving in pseudo-time
$\mathcal{L}$	Total composite loss function
$\mathcal{L}_{\text{data}}, \mathcal{L}_{\text{reg}}, \mathcal{L}_{\text{phys}}$	Data, regularisation, and physics-informed loss components
$\lambda, \gamma, \lambda_1, \lambda_2$	Hyperparameters or weights for loss terms
F, F^{-1}	Fourier and inverse Fourier transforms
R_v	Learnable frequency-domain representation of a convolution kernel
κ	Learnable convolutional kernel in the frequency domain
$\mathcal{I}$	Nonlocal convolution operator
$\tilde{\chi}(\mathbf{x}), \chi(\mathbf{x}, t)$	Mask/characteristic functions for non-physical regions or time-evolving domains
W, c	Learnable linear operator and bias
Δt	time step size

2. Overview of fracture modelling techniques

Design of materials and structures against failure is historically conducted using strength or energy-based techniques. Using simple effective stress-based measures like the maximum principal stress against the material strength, engineers have designed components that are supposed to lie within the linear elastic range. However, the presence of flaws in structures led to the advent of the fracture mechanics principles whereby the critical load required for the crack extension or the critical flaw size under a given load is estimated using LEFM approaches invoking Stress Intensity Factor (SIF) or energy release rate as the crack driving force parameters. While LEFM approaches have proven to be highly successful and practically useful, complex fracture mechanisms are often difficult to treat within the framework of LEFM. LEFM struggles with crack branching, coalescence and fracture in heterogeneous media. To address such limitations, a variety of advanced computational fracture modelling approaches have been developed, including Continuum Damage Mechanics (CDM), CZM, Phase-Field Method (PFM) and Discrete Element Method (DEM). These methods can be broadly classified as either continuum or discrete approaches: in CDM, damage evolves smoothly as a degradation in material stiffness, while discrete models explicitly represent sharp crack surfaces.

This review examines these four modelling approaches crucial for applying ML to fracture mechanics, which primarily serve to generate data sources or provide foundational physics priors for hybrid ML models. This review does not attempt an exhaustive survey of all fracture modelling techniques. Classical approaches such as level set methods, variational formulations [11–14] and peridynamics [15–17] are excluded here for brevity and due to their limited use in the ML-integrated workflows.

2.1. Continuum damage mechanics modelling

CDM offers a thermodynamically consistent framework to model the progressive deterioration of material stiffness and strength by introducing internal variables, commonly denoted by a scalar $D \in [0, 1]$, that represent the extent of damage in the material [18]. The constitutive behaviour is typically modified through damage-dependent stiffness degradation, such as in the effective stress concept: $\sigma = (1 - D)\mathbb{C} : \epsilon$, where $\mathbb{C}$ is the elastic stiffness tensor and ϵ the strain tensor. Damage evolution is governed by energy-based or strain-driven criteria, often derived from potential functions ensuring thermodynamic admissibility. Originally pioneered by Kachanov [19] and further developed by Rabotnov [20] and Lemaitre [21], CDM models have found wide application in metals, composites and cementitious materials for simulating fatigue, creep and ductile damage accumulation [22,23]. The smeared damage representation enables efficient modelling of microcracking in quasi-brittle materials such as concrete [24] and composites [25]. CDM inherently lacks the ability to localise damage sharply along a discrete crack path. To address this limitation and to overcome mesh sensitivity during strain softening, gradient-enhanced or nonlocal formulations have been introduced [26–28].

2.2. Phase field damage modelling

Phase-field models can accommodate crack initiation, branching, coalescence and irregular topologies without requiring predefined crack paths or remeshing [29]. The PFM for fracture models crack evolution using a continuous scalar damage field $d(\mathbf{x}) \in [0, 1]$, where $d = 0$ denotes intact material and $d = 1$ corresponds to fully broken regions [11]. This approach enables a diffuse representation of cracks within a variational framework. A regularisation length scale ℓ governs the width of the diffused crack zone, effectively capturing the material's fracture process zone and eliminating the need for explicit crack tracking. The governing equations arise from the minimisation of an energy functional comprising elastic and fracture surface energies [12,30]. The variational formulation builds upon Griffith's energy criterion and has been extended to model dynamic fracture, ductile damage and fatigue [31–33]. The crack surface energy is approximated using a regularised functional that spreads the sharp crack over a finite width governed by a length scale parameter ℓ . The total energy functional includes elastic strain energy and a surface energy penalty and its minimisation yields coupled Euler–Lagrange equations for displacement and damage fields. Despite their flexibility, phase-field methods are computationally intensive due to the need for fine mesh resolution near cracks and small time steps for stability, especially in dynamic settings [34].

2.3. Cohesive zone modelling

CZMs simulate fracture as a gradual decohesion process across a predefined or adaptively inserted fracture process zone ahead of the crack tip, where cohesive tractions act as a function of the relative displacement between crack faces [35–37]. By replacing the singular crack tip with a cohesive zone, CZMs remove stress singularities typical of classical fracture models and enables an explicit energy-based treatment of crack initiation and propagation. The core of CZMs lies in the definition of the traction-separation law, which governs the mechanical response within the cohesive zone. A cohesive law typically relates the traction vector $\mathbf{t}$ across the interface to the relative displacement δ , incorporating damage initiation, evolution and final decohesion. The foundational work of Barenblatt [36] and Dugdale [37] established the theoretical basis for cohesive fracture. Subsequent computational implementations in finite elements were pioneered by Xu and Needleman [35,38,39] and Camacho and Ortiz [40], whose widely cited studies enabled efficient simulations of dynamic fracture including in three dimensions and under large-scale plastic deformations [41]. The incorporation of cohesive surfaces into finite elements and later into Extended Finite Element Method (XFEM) frameworks was critical to handling arbitrary crack initiation and propagation paths.

CZMs are widely used in simulating delamination in layered composites [42], interfacial debonding and crack growth in ductile materials. The method has been extended to arbitrary crack paths using adaptive insertion of cohesive elements or through XFEM [43]. Rate-dependent and mixed-mode CZMs have also been developed to handle dynamic fracture and complex loading scenarios [44,45]. Their ability to represent fracture mechanisms in a computationally tractable manner has led to widespread adoption. Their effectiveness depends on accurate calibration of cohesive parameters and resolution of the cohesive zone, which can be computationally expensive without adaptive schemes or enrichment techniques [46].

2.4. Discrete element method

DEM represents materials as an assembly of distinct particles or rigid elements interacting via contact forces and cohesive bonds. In fracture modelling, DEM is employed for granular and quasi-brittle media, where it enables direct simulation of crack initiation, propagation and fragmentation based on bond breakage or contact force thresholds. Unlike continuum models, DEM captures discontinuities and discrete events naturally, offering detailed insight into micromechanical failure processes.

Originally proposed by Cundall and Strack [47], DEM has seen applications in scenarios ranging from rock burst simulations [48] to brittle fracture under indentation loading in glasses [49], as well as the particle-scale modelling of sand grain crushing [50]. The method facilitates the study of phenomena such as comminution [51], crack network evolution and particle-scale interactions. DEM requires careful calibration of micro-level parameters to reflect macroscopic behaviour and often demands considerable computational resources for large-scale or high-resolution simulations. Hybrid FE–DEM methods have been developed to leverage the accuracy of continuum mechanics in bulk regions with the discrete capabilities of DEM in fracture zones [52].

3. Overview of machine learning approaches

3.1. Machine learning methods

Machine learning, a branch of artificial intelligence, has grown substantially over the last two decades due to its potential in addressing many real world problems. These developments are made possible with state-of-the-art computational capabilities, innovative algorithms and ever increasing data availability. In this section, various machine learning methods are reviewed, with more focus on deep learning models, in the context of mechanics problems.

Let us represent a ML model as, $\mathcal{M}(\theta)$ where θ is a set of learnable (or adjustable) parameters. In the first step of model development, these parameters are typically initialised to some random state or the one determined by specialised algorithms [53–55]. During training, these model parameters are iteratively updated aiming to reach the optimal state $\mathcal{M}^*$, where the model minimises error on the training data samples. It is desirable to have the models that perform well on both the unseen data and out-of-distribution data with reasonably good accuracy, in addition to the training data.

In general, a ML model can be trained in three different ways depending on the availability of data and nature of the task, as follows:

1. **Supervised learning** of a model $\mathcal{M} : \mathcal{X} \rightarrow \mathcal{Y}$ where each training sample contains an input $x \in \mathcal{X}$ and its corresponding output label $y \in \mathcal{Y}$. The learning process can be regression or classification, depending on the nature of y being continuous or discrete, respectively. Predicting fatigue life as number of cycles until failure (a continuous label) of a structure from its material properties and loading conditions (input) can be treated as regression [56–58].
2. **Unsupervised training** is performed to find an optimal ML model $\mathcal{M}^*$, that finds patterns, structures or relationships within the data that does not contain the label information. A range of tasks belong to this category, like clustering, dimensionality reduction and probability density estimation of the given data, among many applications. One can detect structural damage [59] by finding the changes in the patterns of sensor data (input).
3. **Reinforcement Learning (RL)** does not require input and/or output labels. Instead, the agent (i.e., model) learns through trial and error, interacting with its environment. The goal is to train the agent to take actions that maximise cumulative rewards over time. As the training process progresses, the agent focuses on those actions which historically have resulted in positive rewards. Hence, training the RL model is a continuous process of balancing ‘exploration of new actions’ and ‘exploitation of rewarding actions’. Consider a task of developing an optimal control system for fluid flow through a network of pipes and valves. The control system (agent) can be trained by exploring multiple control parameters (actions), like pump speeds and valve openings. Over time, the agent learns to prioritise actions that optimise flow rates and pressure distributions (rewards), gradually improving system performance [60–62].

In addition to these, there also exists the semi-supervised learning that combines the best of both supervised and unsupervised training by using small amounts of labelled data and large amounts of unlabelled data. This method is very attractive for cases where sample labelling is very expensive or time-consuming.

3.2. Neural networks

In general, machine learning methods like linear and logistic regression, decision trees, Support Vector Machine (SVM) and k-Nearest Neighbours (k-NN) are commonly used in a wide range of domains, including computational mechanics [9,63,64]. While these methods offer advantages, the complexity of the datasets and underlying systems requires more sophisticated approaches, such as neural networks. Unlike traditional machine learning methods, neural networks can automatically learn intricate relations within raw data, eliminating the need for explicit feature engineering. Consequently, neural networks have become the default choice for the majority of problems, including mechanics, due to their ability to handle large and complex datasets backed by universal function approximation theorem [65] and leveraging parallel computational capabilities. Hence, we proceed to review deep learning methods encompassing neural networks and various approaches aimed at solving complicated problems in the scientific computational domain.

A neural network, $\mathcal{N}(\theta) : \mathbb{R}^m \rightarrow \mathbb{R}^n$, is a differentiable function composed of several functions organised in a series of layers that maps $\mathbf{x} \in \mathbb{R}^m$ of input layer to $\mathbf{y} \in \mathbb{R}^n$ of output layer, through one or more hidden layers. Here, $\theta = \{\mathbf{W}^l, \mathbf{b}^l\}_{l=2}^L$ is a collection of learnable parameters, containing weights ($\mathbf{W}^l \in \mathbb{R}^{d_l \times d_{l-1}}$) and biases ($\mathbf{b}^l \in \mathbb{R}^{d_l}$) from the first hidden layer to output layer. At each layer l (excluding first layer), these parameters are used to transform inputs $\mathbf{a}^{l-1}$ coming from the previous $(l-1)$ th layer into outputs $\mathbf{z}^l$, followed by an element-wise application of non-linear activation functions σ as shown in Equation (1). Note that deep neural networks are a class of neural networks with more than one hidden layer.

$$\begin{aligned} \mathbf{z}^l &= \mathbf{W}^l \mathbf{a}^{l-1} + \mathbf{b}^l \\ \mathbf{a}^l &= \sigma(\mathbf{z}^l) \end{aligned} \quad (1)$$

Neural networks (or Artificial Neural Networks) are known for their capability in capturing highly complex and non-linear relations. The layered representation of the network enables hierarchical learning, where each successive layer extracts features from the previous layer with increasing complexity towards the last layer. For example, to predict the buckling behaviour of a column, initial layers of a neural network may learn to identify basic features like length, loads, and cross-sectional area. Then, as the data passes through successive layers, the model learns high-level details like load distribution, boundary conditions, and material non-linearities, finally estimating the full buckling response of the column. Explainable AI (XAI) is being developed to decode this kind of hierarchical learning, providing insights into how neural networks, often considered black boxes, arrive at their predictions.

The development and training of a neural network depends on four basic elements: (i) the nature and representation of the data, (ii) the network architecture, (iii) the loss function, and (iv) the optimisation algorithm. The implementation of these components is determined by characteristics of available data and the task-specific requirements, such as classification, regression, sequence prediction, physics constraints, and data generation.

3.2.1. Data

Data containing the characteristic information of the system can be obtained through experimental measurements or generated synthetically using high-fidelity physics-based simulations. The latter approach has become increasingly important due to the expensive and time-consuming nature of experimental data collection in engineering applications, particularly for complex mechanical systems. For mechanics problems, data typically represents physical quantities such as strain/stress fields, displacement time series, damage patterns, or discretised geometries. These quantities are computationally represented in various forms, such as tensors for regular grid data, graphs for mesh kind of data, and point clouds for irregular geometries, with the choice depending on the specific physical problem and computational requirements.

A dataset comprises multiple samples or training examples, each containing input features and/or output labels. It is worth noting that while increasing the number of samples in a dataset may enhance learning, the quality and physical representativeness of the data play a crucial role. The dataset is typically split into training, validation, and test sets, a process that requires careful attention to maintaining the representative distribution of the whole dataset across various subsets to ensure reliable model evaluation.

Raw data often requires several processing steps before being used to train the model. These steps include normalising the data to a common scale, dimensionality reduction, reducing the class imbalance, handling missing or noisy data, feature engineering, and data augmentation. While operations like normalisation are essential for stable training, others like feature engineering and data augmentation are optional and can be custom-designed based on the problem and nature of the data at hand. Data augmentation, for instance, can generate new training samples by applying physically plausible transformations to existing data, such as elastically distorting the geometry of a component or leveraging the symmetry inherent in a system's boundary conditions. Before proceeding to training, it is important to ensure the correctness, quality, and representativeness of the data by performing thorough data cleaning and evaluation through various sanity checks for physical validity, completeness, and consistency.

3.2.2. Neural network architecture

A Neural network architecture defines the hierarchical arrangement of layers and neurons to form a differentiable function approximator. The selection of architecture is primarily determined by the structure of the input data structure and the specific task requirements, such as regression, classification, sequence prediction, dimensionality reduction, and data generation.

For low-dimensional problems, a simple Multilayer Perceptron (MLP) used with multiple fully connected layers is sufficient. However, MLPs scale poorly to high-dimensional spatial data. For such data, particularly image-based representations like full-field strain maps, Convolutional Neural Networks (CNNs) are the dominant architecture. CNNs learn hierarchical spatial features through local connectivity and weight sharing, making them best suited for problems like image based microstructure–property prediction. U-Net, a successful variant of CNN, with its encoder–decoder structure and skip connections, is widely used for pixel-wise predictions, such as damage identification or stress fields estimation. For data on irregular grids such as finite element meshes, GNNs provide a more natural framework. GNNs operate on graph structures, propagating information between nodes and edges to learn representations that inherently respect the domains' topology, making them a powerful choice for mesh based simulations.

Sequence prediction problems, such as modelling damage evolution or analysing experimental sensor data, require architectures that capture temporal dependencies. Recurrent Neural Networks (RNNs) represent the first generation of architectures designed for sequence prediction tasks. However, their ability to capture long-range dependencies is limited, and suffer from vanishing gradient problems, which hinders effective learning for longer sequences. These issues are addressed with advanced architectures like Long Short-Term Memory (LSTM), which control the information flow through cell states gating mechanisms. Sequential computations in these architectures limit parallelisation, making them less suitable for large-scale tasks. Transformer architecture overcomes this by using a self-attention mechanism to capture global dependencies and parallelise computation. For spatiotemporal data, such as prediction of evolving physical fields, Convolutional LSTM (ConvLSTM) combines the advantages of LSTM and CNN to capture both spatial and temporal dependencies.

For data generation tasks such as creating synthetic microstructure, generative models like Generative Adversarial Networks (GANs), Variational Autoencoders (VAEs) and diffusion models are often employed. Generative models fundamentally learn to map the underlying data distribution to a known latent space distribution such as a Gaussian. Standard GANs utilise an adversarial objective that is susceptible to 'mode collapse'. This phenomenon occurs when the generator converges to a limited subset of outputs, failing to capture the full statistical diversity of the target distribution. This limitation is especially critical in fracture mechanics, where material failure and structural reliability depend on these stochastic defects. A collapsed model that overlooks these stochastic tails yields a statistically biased dataset that discards the very features that govern material failure. To address this limitation, recent research has pivoted towards architectures that enforce physical consistency through conditional inputs and preserve local texture through patch-based discrimination.

For surrogate or Reduced Order Model (ROM) development, architectures typically follow data-driven or physics-informed paradigms. Data-driven methods, such as those employing Autoencoder (AE) for latent space compression, rely on patterns within large datasets to construct efficient models. In contrast, physics-informed approaches like Physics-Informed Neural Network (PINN) directly integrate governing equations into the training loss to enforce physical consistency, proving valuable when data is sparse. Neural operator learning, implemented by architectures such as Deep Operator Network (DeepONet) and Fourier Neural Operator (FNO), represents a powerful advancement in recent times that can leverage both approaches; it learns mappings between function spaces, enabling zero-shot generalisation across new boundary conditions and geometries without retraining.

A fundamental element across all architectures is the activation function, which introduces non-linearity and differentiability, and makes the learning of non-linear relationships possible. The choice of function influences how the model learns patterns and should be guided by the underlying physics. The piecewise linearity of ReLU makes it suitable for capturing sharp discontinuities like crack fronts, while its simplicity aids stable training. The smooth, differentiable nature of softplus [66] is preferred in PINNs for problems requiring higher-order derivatives. Saturating functions like tanh and sigmoid provide bounded outputs [53], making them ideal for constraining predictions to physically plausible ranges, such as normalised damage indices or stresses.

3.2.3. Loss functions and metrics

The training and evaluation of a model are guided by two distinct but related concepts: the loss function and evaluation metrics. The loss function, which can be considered as the cost or objective function, quantifies the discrepancy between the model's predictions and the true values, for a given batch of training data, into a single scalar value. This measure of error is directly

minimised by an optimisation algorithm to update the model's parameters θ . Some of the common loss functions for regression tasks include Mean Square Error (MSE) and Mean Absolute Error (MAE), while cross-entropy is standard for classification. To enhance generalisation and prevent overfitting, the base data loss $\mathcal{L}_{\text{data}}$ is often augmented with additional terms. A regularisation loss $\mathcal{L}_{\text{reg}}$ discourages overly complex models, while a physics-informed loss $\mathcal{L}_{\text{phys}}$ ensures adherence to known physical principles by penalising violations of governing equations, computed via automatic differentiation. The total composite loss is a weighted sum:

$$\mathcal{L} = \mathcal{L}_{\text{data}} + \lambda \mathcal{L}_{\text{reg}} + \gamma \mathcal{L}_{\text{phys}} \quad (2)$$

where λ and γ are hyperparameters controlling the influence of each term.

In contrast, evaluation metrics are typically used after training to assess the model's performance on unseen test data. These metrics are designed to be interpretable and often align closely with the final engineering objective. It is worth noting that, while a metric may sometimes be the same as the loss function (e.g., MSE), this is not always the case. Most common evaluation metrics reported in the literature include MAE, Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and R^2 for regression tasks; precision, recall, and the F1 score for classification tasks (see Appendix Table 4 and Table 3 for definitions).

3.2.4. Optimisation algorithms

Optimisers define the algorithm by which a neural network's weights (i.e., learnable parameters) are iteratively updated to minimise a given loss function. This process is fundamentally enabled by backpropagation, which efficiently computes the gradients of the loss with respect to all network parameters via the chain rule. The optimiser's global or parameter-wise learning rate controls the step size of the update. The choice of optimiser is critical, as it directly influences convergence behaviour and numerical stability. This is especially important in problems such as fracture modelling involving highly non-linear physical interactions, sharp spatial gradients or tightly coupled field equations that are sensitive to small perturbations.

A review of prevalent practices reveals a clear pattern. Adam [67], with its adaptive learning rates and momentum-based updates, is the predominant choice for stochastic and mini-batch training. Conversely, for Physics-Informed Neural Networks (PINNs), the Limited-memory Broyden–Fletcher–Goldfarb–Shanno (L-BFGS) optimiser [68] is often favoured. Its full-batch, second-order approach is well-suited to the characteristically smooth but complex loss landscapes generated by physics-informed constraints [69]. While less common in modern deep learning, RProp [70] remains a relevant benchmark for certain scenarios involving shallow networks and potentially unstable gradients. Ultimately, the optimiser selection is a deliberate trade-off between convergence speed, computational efficiency, numerical robustness, complexity of the loss landscape and adherence to the underlying physical constraints of the problem.

4. Status review of AI applied to fracture problems

Building on the growing integration of machine learning into computational mechanics, this section presents a focused review of ML methods applied to fracture and damage modelling. We examine how different ML architectures have been employed in fracture mechanics, categorised by their learning methods, modelling principles and suitability for specific tasks. The rapid adoption of ML in fracture problems has led to a range of strategies — from purely data-driven surrogates to hybrid physics-informed frameworks. Supervised deep learning models, such as CNN and GNN, have shown strong performance in learning mappings from input conditions (e.g., material properties, geometry and loading) to outputs like crack paths or damage fields. Once trained, these models offer substantial speed-ups, achieving $6\times$ to $25\times$ faster inference than XFEM and up to $\sim 2800\times$ over phase-field methods [71,72], enabling near real-time predictions with minimal loss in accuracy.

In this section, machine learning architectures are explored alongside specific fracture modelling objectives such as crack path prediction, energy minimisation and damage field reconstruction. This integrated approach is a deliberate methodological choice necessitated by the interdependency between model architecture and physical modelling goals observed in the literature. In many cases, the selection or design of an architecture is directly influenced by the nature of the fracture problem being addressed; consequently, a rigid separation would misrepresent this relationship. Readers are therefore encouraged to view the following subsections as a dual narrative: a structured classification of learning architectures and a contextual analysis of how each aligns with different classes of fracture mechanics problems.

We begin with classical data-driven surrogates that emulate physics-based fracture solvers using relatively simple architectures such as feed-forward networks, k-nearest neighbours, regression forests and LSTMs. Applied across phase-field, discrete and boundary element models, these approaches offer strong predictive performance and computational speed-ups, but typically function as black-box regressors without encoding geometric structure or physical constraints. The following subsections explore specialised architectures. GNNs leverage mesh topology to model irregular and evolving crack patterns, while CNN-based U-Nets treat fracture prediction as an image-to-image translation task. Generative models such as GANs extend this capability to stochastic fracture scenarios by synthesising realistic crack patterns. ROMs, developed using auto-encoders or tensor decompositions, offer efficient approximations of high-dimensional fracture fields. While effective, these purely data-driven models require extensive experimental/simulation data, lack interpretability and physical consistency, which limits their generalisability. To address this, a class of methods emerged that utilise physics to guide the Neural networks (NNs). Among these methods, PINNs have attracted significant attention from the community due to their ability to use governing differential equations in learning and reducing data requirements thereby enabling better extrapolation across physical settings. We then review operator learning frameworks such as DeepONet and the FNO, which learn mappings between infinite-dimensional function spaces [73]. These models can generalise

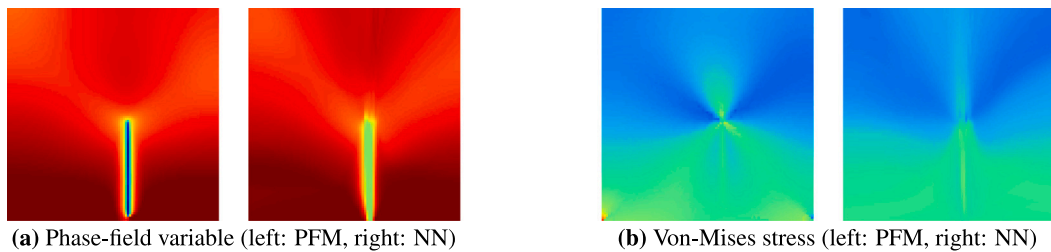


Fig. 1. Comparison between the phase-field model (PFM; left) and neural network (NN; right) predictions of (a) the phase-field variable ϕ and (b) the von Mises stress field at the same time step [133]. The NN accurately reproduces the crack growth direction and stress pattern away from the crack tip. However, it underestimates the peak value of ϕ and fails to resolve sharp stress gradients near the crack tip, likely due to the coarse, uniform sampling grid. Improved spatial resolution may help recover these features more effectively.

across families of PDEs and boundary conditions, offering remarkable flexibility and efficiency for fracture simulations. Finally, we explore material-specific applications, where ML models are adapted to simulate fracture behaviour in complex systems ranging from nanostructure graphene to hyper-elastic soft tissues and metallic alloys.

The following subsections trace the evolution of ML in fracture mechanics from surrogate emulators to operator-based and physics-informed models. This trajectory points towards hybrid, physics-aware ML capable of delivering interpretable, generalisable and accelerated simulations. To support this overview, Table 1 compares existing ML techniques, while Table 2 surveys fracture modelling approaches alongside the ML methods employed.

4.1. Data-driven emulators and surrogates

Data-driven emulators are designed to learn the input–output relationships of complex physics-based simulations, offering significant computational speed-ups. For phase-field models, Moraes et al. [132] developed a supervised framework coupling a thermodynamically consistent damage-fatigue model with classification-based learners (k-NN and Artificial Neural Network (ANN)). Trained on time-series data from virtual sensing nodes, their model achieved over 90% for cases having Gaussian noise up to 20%. Similarly, Koeppe et al. [133] trained a four-layer LSTM network on 430 phase-field simulations of 2D notched plates, accurately predicting time-evolving damage and stress fields with a MSE in the order of (10^{-4}) and achieving a speed-up of 10× to 20× over traditional solvers. Fig. 1 presents a comparison between PFM and NN predictions of the damage and stress fields for a test geometry at the same time step. The NN demonstrates strong agreement with the reference simulation, accurately capturing the crack propagation path and stress distribution. In dynamic fracture and discrete modelling, the Hybrid Optimization Software Suite (HOSS) software suite developed at Los Alamos National Laboratory, [139] has enabled large-scale surrogate training. Feedforward networks trained on 200 Finite-Discrete Element Method (FDEM) simulations predicted dominant crack paths with 85% accuracy and 16% deviation in failure timing, achieving $10^3 \times$ – $10^4 \times$ speedups [115].

In parallel with the full-field, spatially aware surrogates described in this section, a separate line of research employs classical tabular machine learning to predict scalar fracture quantities directly from geometric, loading and material parameters. Data-driven methods have been implemented to predict mixed mode SIFs in heterogeneous materials [140]. For Boundary Element Method (BEM) based fracture simulations, Huang et al. [121] pioneered a regression forest model, eliminating the need for solving dense linear systems for crack opening displacements (CODs). Their model was trained with input vectors of SIFs (K_I , K_{II} , K_{III}), material toughness K_c and propagation distance r , predicting projected CODs with linear-time complexity. The model trained on simple cube geometries demonstrates generalisability by predicting fracture surfaces on the Stanford Armadillo model [141] showcased in their paper. Bahrami et al. [142] applied Gaussian process regression (GPR), decision tree ensemble-boosting, and shallow neural networks to predict fracture load from 631 experimental tests covering 20 brittle material types under mixed-mode I/II loading across 8 geometries and 4 blunt notch types (V-, U-, key-hole, and circular). They report R^2 values of 0.96 (GPR), 0.93 (ensemble-boosting), and 0.91 (ANN) on unseen data, with GPR providing the best interpolation and uncertainty quantification. The same group extended this line of work to notched specimens of poly(methyl methacrylate) (PMMA; a transparent brittle polymer) and rock-like materials in [143], reporting that 92% (GPR), 89% (XGBoost), and 88% (ANN) of the fracture load predictions on unseen notch geometries fell within $\pm 15\%$ of the measured experimental values. The recent review by Niu et al. [144] showcase benchmarks of Random Forest (RF), gradient-boosting methods, SVM and shallow MLP in prediction fracture toughness of quasi-brittle materials from mix-design and aggregate parameters. Ensemble models typically reduce error by 15%–20% compared to single learners and yield straightforward feature-importance rankings.

Hybrid strategies integrating classical physics-based methods with data-driven components have also emerged to leverage the strengths of both paradigms. Greve et al. [124] replaced the Marciniak–Kuczyński imperfection model with a feedforward NN trained on virtual data, enabling real-time fracture prediction in metallic sheets with experimental validation. Nguyen et al. [74] developed an adaptive strategy to optimise feedforward ANN architectures by iteratively adjusting the number of neurons and hidden layers during training. Model performance is evaluated using the absolute percentage error, which quantifies the relative difference between predicted and actual outputs for each observation. A control condition based on a minimal allowable gap guides

Table 1

Comparison of prominent AI techniques applied in fracture mechanics. Each technique offers different trade-offs in terms of accuracy, interpretability and computational cost.

AI technique	Applications in fracture mechanics	Strengths	Limitations	Notable variants/case studies	Accuracy & computational efficiency
Convolutional Neural Networks (CNNs) e.g., U-Net, FE-CNN, StressNet, ConvLSTM	- Full-field surrogates (stress, damage) for phase-field/cohesive fracture [72,74–76] - Predicting damage evolution from microstructure data [75] - Crack path prediction from images [77,78] - Local stress prediction using FE input (FE-CNN) [79] - End-to-end SIF/displacement prediction (StressNet) [80,81] - Learning fracture histories (ConvLSTM) [72]	- Captures spatial features effectively [72,75,79] - Fast inference for full-field predictions [76,80] - Can model spatiotemporal evolution (ConvLSTM) [72] - Good performance within training distribution [74,75,79]	- Requires large simulation/experimental datasets - May lack strict physics enforcement; potential for unphysical results [82] - Generalisation outside training data (e.g., unseen geometries) can be poor [74,79]	- U-Net: Predicted 2D crack maps in concrete (phase-field) [72,83]; Learned damage/stress evolution [82]. - FE-CNN: Local stress prediction [79]. - StressNet: SIF prediction [80]. - Concrete/Porous Media: Path prediction [72]. - Alloys/Steel: 3D fatigue crack prediction [84]; - Crystalline Solids: Graphene, MoS₂, Polycrystals [85–89]; - Composites: Damage/Stress prediction [74,75,83].	- Accuracy: High match with FEM (e.g., F1 ~0.9 reported in [80]). - Speed: Orders of magnitude faster inference than FEM/PDEs solvers.
Graph Neural Networks (GNNs) e.g., ACCURATE, EvolveGCN	- Crack growth, interaction and coalescence in discrete/networked cracks [71,90] - Predicting SIFs or crack tip activity [71,91] - Fracture in irregular domains/lattices (rocks, masonry) [90] - modelling node/element-wise fracture on meshes [92,93] - Learning entire fracture histories (spatiotemporal GNNs) [88,90,92]	- Handles arbitrary/complex crack configurations and mesh structures [71,92] - Learns interactions via message passing [71,90] - Potential for generalisation across different configurations [91] - Can capture spatiotemporal dynamics (DyFraNet, EvolveGCN) [88,90,94]	- Often requires large datasets (e.g., XFEM/FEM) [71] - Graph construction can be complex and critical [92] - May focus on discrete events (e.g., tip propagation) rather than continuous fields [71,90]	- Microcracks: GNN replicated XFEM paths/SIFs [71]. - Masonry/Rock: Cracking simulation [90]. - Temporal GNNs: EvolveGCN [90,94] for fracture evolution prediction.	- Accuracy: Low SIF errors (~1%–3%) vs. FEM/XFEM reported in [71]. - Speed: Significant speedup reported (e.g., 6–25x) compared to XFEM [71].
Physics-Informed Neural Networks (PINNs) incl. VPINN, VE-PINN, XPINN, I-FENN, Transfer Learning PINNs	- Solving forward/inverse fracture mechanics PDEs using physics loss [69,95,96] - Phase-field modelling via PDEs residual or energy minimisation (VPINN, Deep Ritz) [63,95,97–99] - Handling boundary conditions [96,100] - Surrogate modelling within FE loops (I-FENN) [101] - Accelerated simulation via transfer learning across steps [95]	- Requires minimal or no simulation data [69] - Enforces physical laws (PDEs, conservation laws) [63,69] - Can generalise better than purely data-driven models [69,102] - Variational forms align well with phase-field theory [63,100] - PBML principles often incorporated	- Training can be computationally expensive and challenging to converge [69,103] - Handling sharp discontinuities (cracks) can be difficult [96,104,105] - Often requires retraining for new boundary conditions or time steps (mitigated by transfer learning [95])	- VPINN/Deep Ritz: Phase-field simulation [63,98], improved BC handling [100]. - XPINN: Domain decomposition [106,107]. - Transfer Learning: Accelerated step-wise simulation [95]. - I-FENN: Coupled FEM-PINN solver [101,108]. - Brittle Fracture: Crack nucleation/branching studies [96,97,105].	- Accuracy: Generally within a few percent of FEM solutions [69,95,96,101]. - Efficiency: Fewer DOFs than FEM [101], but training is often slow [69]. Transfer learning improves step-wise efficiency [95].
Operator Learning e.g., DeepONet, FNO, V-DeepONet	- Learning solution operators (function-to-function maps) for fracture problems [109,110] - Fast prediction of displacement/damage fields from initial conditions/parameters [109,111] - Generalising across geometries and loading conditions [110] - Surrogate modelling for multi-query tasks (e.g., uncertainty quantification) [112]	- Can generalise across input functions (e.g., different initial defects) [109,110] - Extremely fast inference once trained [109,111] - Physics-informed variants (V-DeepONet) possible [109] - Suitable for tasks requiring many forward solves with varying inputs [112]	- Can require vast training data (function pairs) [109] - Training complexity depends on operator complexity - Ensuring physical constraints across the entire function space can be challenging [109]	- V-DeepONet: Learned displacement/damage fields from initial defects [109]. - FNO: Applied to various PDEs systems, potential for fracture modelling. - Quasibrittle Fracture: DeepONet predicted crack paths without explicit time stepping in some studies [109]. - Nucleation/Propagation: Prediction using DeepONets [110].	- Accuracy: High fidelity demonstrated in specific applications [109–111]. - Efficiency: Very fast inference, suitable for real-time or multi-query scenarios [109,112].
Autoencoders (AEs) incl. Variational AEs	- Surrogate modelling via latent space representations [112–114] - Dimensionality reduction of fracture data (e.g., crack networks, fields) [113,114] - Unsupervised feature extraction for damage states - Potential for inverse design of microstructures for fracture resistance [84]	- Efficient data compression/dimensionality reduction [113,114] - Can learn salient features without labels - Latent space can be used for generation or interpolation [84,113] - Fast prediction using the decoder part [112]	- Reconstruction may lose fine details (e.g., sharp crack tips) - Latent space often lacks direct physical interpretability [114] - Tuning required to balance compression and fidelity	- Microstructure evolution/representation: Latent space modelling [113]. - Generative modelling/Inverse design: Using latent spaces [84]. - Hydraulic Fracture/Rock Networks: (Less direct mapping in provided bib; general dimensionality reduction applicable [114])	- Accuracy: Can reproduce major features; quantitative error depends on compression level [113]. - Speed: Latent representation enables very fast surrogate predictions (milliseconds) [112].

the ANN configuration to refine the network predictions. This automated tuning process led to 180x acceleration compared to traditional phase-field simulations. LSTM-based recurrent networks have demonstrated high accuracy in predicting crack bifurcation in heterogeneous tensile specimens, with deviation in crack paths kept below 2% [126]. Probabilistic surrogates combining Extended Support Vector Regression (X-SVR) with Dirichlet feature mapping have been used to model damage field distributions under uncertainty [136].

Building on this overview, recent work traces the rise of data-driven approaches, moving from shallow surrogates and classical ML models to deep and hybrid frameworks. Increasingly expressive architectures such as LSTM, Graph Convolutional Network (gCN)

Table 2
Summary of ML/DL Techniques applied to various fracture modelling approaches.

Fracture methodology	ML/DL technique	Reference
BEM Brittle Fracture Rigid body dynamics	CNN + MLP	[78]
Brittle fracture	Reproducing Kernel Particle Method-NN	[105]
Cohesive Phase Field	CNN	[72]
CZM (FDEM HOSS)	ANN, RF, DT	[115]
CZM (FDEM HOSS)	CNN	[80]
CZM (FDEM HOSS)	GCN + RNN + LSTM	[116]
Configurational Forces	CNN	[117]
Configurational Forces	RNN + gCN + LSTM	[90]
CPD Constitutively informed Particle Dynamics	DeepONet	[118]
CZM (IGFEM)	CNN	[75]
DEM (unspecified)	ROM	[119]
FDEM(HOSS), Continuum-Scale Model (FLAG)	LSTM + RNN	[120]
Galerkin BEM	RF	[121]
Gurson–Tvergaard–Needleman model (Additive Mfg., Warp3D)	ANN	[122]
LEFM	PINN	[96]
*Likely PFM, CDM (High Fidelity Model)	ConvAutoEncoder	[123]
MD (Graphene, MoS ₂)	ConvLSTM	[85,86,88,89]
MD simulations	U-Net	[87]
Marciniak and Kuczyński Model, Enhanced Hill/Stören-Rice (eHSR) model	NN	[124]
CZM	LSTM + Hidden Markov Model	[125]
Modified Mohr–Coulomb ductile model	RNN + LSTM	[126]
Non-local gradient enhanced damage	PINN	[103]
Peridynamic model	Deep Implicit FNO	[127]
Peridynamics (brittle fracture)	Domain Agnostic FNO	[128]
PFM	ANN	[74]
PFM	cGAN	[129]
PFM	CNN	[76,83,130]
PFM	CNN + FNO	[131]
PFM	deepRitz VPINN	[98]
PFM	GNN	[92]
PFM	k-NN, ANN	[132]
PFM	LSTM, NN	[133]
PFM	NN	[134]
PFM	PINN	[95,99,101]
PFM	PINN, VPINN, VE-PINNs	[97]
PFM	ResNet, KANs, DeepONet	[110]
PFM	TCN	[108]
PFM	TT + GPR	[135]
PFM	VE-PINN - DeepONet	[109]
PFM	VExPINNs	[107]
PFM	X-SVR	[136]
PFM (AMR)	GNN	[93]
PFM (FFT)	CNN	[137]
XFEM (based on minimisation of total energy [138])	GNN	[71,91]

and attention-based models offer greater fidelity and adaptability, underscoring the need for specialised designs that can capture irregular geometries and evolving crack topologies. The next subsections examine these approaches, beginning with graph-based learning.

4.2. Graph neural networks

GNNs are well-suited for irregular mesh data and unstructured domains similar to finite element simulations. By encoding spatial relationships and topological connectivity, GNNs have been used to predict nodal damage, stress or displacement fields with high fidelity. This makes them ideal for scenarios involving mesh adaptivity, crack branching or evolving geometries. A significant body of work by Perera and Agrawal [71,91–93], among others, demonstrates a systematic progression in developing these frameworks to address key challenges like physical fidelity, model generalisability and the integration of dynamic meshing techniques.

Early work in applying graph-based deep learning to fracture dynamics was presented by Schwarzer et al. [90] where they reduced computational overhead by converting finite element data into a reduced-order “fracture graph” where nodes and edges

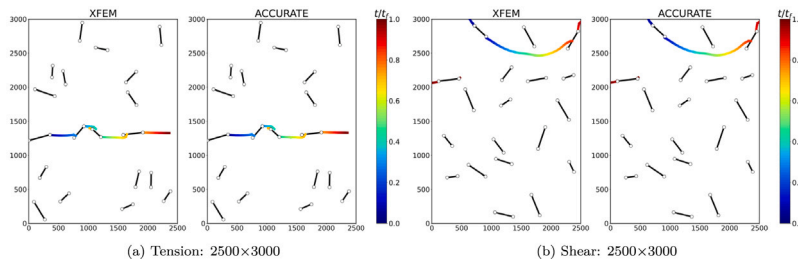


Fig. 2. Comparison of crack coalescence and growth by XFEM versus GNN ACCURATE models [71].

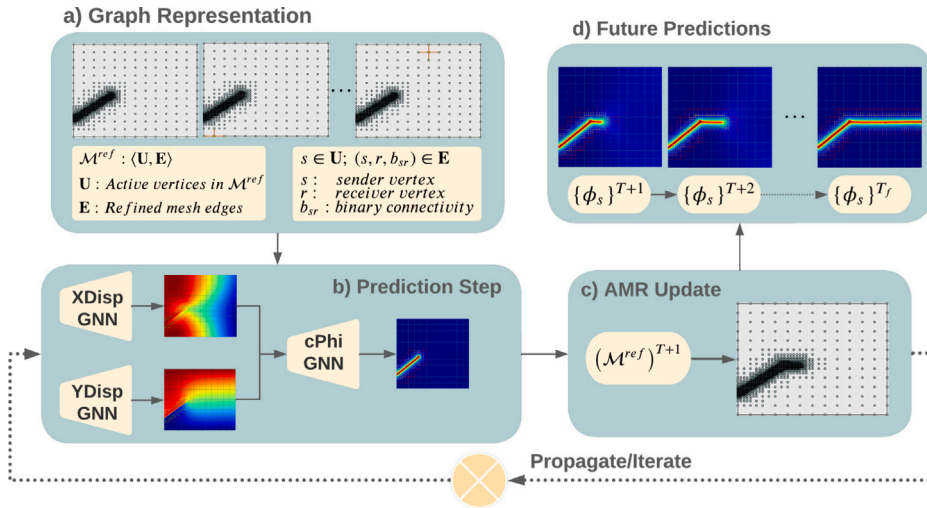


Fig. 3. The cyclical workflow of the ADAPT-GNN framework. It integrates AMR by (a) representing the refined mesh as a graph, (b) performing GNN-based prediction, (c) updating the mesh with AMR and (d) iterating for future time steps [92].

represented individual fractures and physical relationships respectively. A gCN combined with RNN was used to learn the system's spatio-temporal dynamics, achieving substantial acceleration while maintaining high accuracy.

Building on this, Perera and Agrawal established a GNN-based surrogate model - 'Microcrack-GNN', to emulate XFEM simulations of multiple interacting microcracks [71]. A key innovation was a sequential, physics-informed pipeline that treated crack-tip interactions as a graph prediction problem first estimating Mode I/II SIF, then identifying which cracks propagate and finally predicting the updated crack geometry. Using this sequential GNN approach, the crack propagation and coalescence of 5–19 microcracks was simulated, with results matching closely with XFEM solver. The GNN could handle variable crack configurations even on newer crack patterns previously unseen by the model, making it an attractive approach for simulating microcrack propagation with speed-ups of 6× to 25× compared to XFEM simulator.

To extend the model's capabilities, the authors developed *ACCURATE* — ACcelerated Universal fRACture Emulator [71]. The model leverages Transfer Learning (TL) to efficiently adapt the pre-trained Microcrack-GNN to new problem configurations using minimal additional data. As shown in Fig. 2, the predictions of crack coalescence and growth remain in close agreement with XFEM, while achieving speed-ups of up to 200×. The framework achieved reliable generalisation with as few as 20 simulations per adaptation step, demonstrating the effectiveness of transfer learning.

The research was extended from XFEM to PFM with Adaptive Mesh Refinement (AMR) which the authors called ADAPT-GNN. This framework was designed specifically to emulate fracture simulations by uniquely re-defining the GNN's graph structure as the refined mesh at each time step [92]. As shown in its workflow (Fig. 3), the framework operates in a cycle: the GNN predicts displacement and damage fields on the current mesh and the mesh is then updated via AMR for the next prediction step. This dynamic coupling of the GNN with AMR yielded speed-ups of up to 36× compared to the high-fidelity PFM [92].

Mesh-based GNNs, in their classical design, typically rely on fixed discretisation and require retraining when applied to problems with different geometries, boundary conditions or mesh topologies, which limits their generalisability and increases computational cost. To address these challenges, Perera and Agrawal [93] propose a multiscale GNN architecture integrated with bs-AMR to efficiently solve time-dependent problems with near-singular operators, such as phase-field fracture. As shown in Fig. 4, the framework processes the mesh derived from AMR, at multiple resolution levels. At each coarsening step, the highest refinement

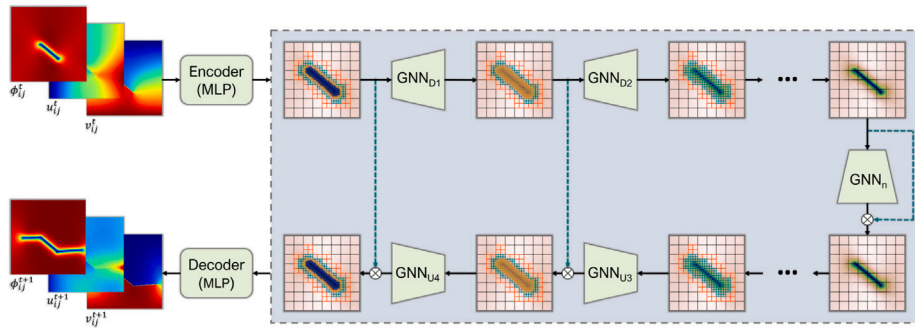


Fig. 4. Architecture of the multiscale GNN framework with bs-AMR [93]. The encoder uses an MLP to embed the input mesh graph, which is then coarsened progressively through multiple GNN blocks ($GNN_{D1}, GNN_{D2}, \dots$) by eliminating the highest refinement level at each stage. A corresponding upsampling pathway reconstructs the full-resolution graph through additional GNN stages ($GNN_{U3}, GNN_{U4}, \dots$), with skip connections linking matching downsampling and upsampling layers. The final embedding is decoded to predict crack fields and displacement evolution.

level is eliminated, resulting in progressively coarser graphs that preserve essential global information while reducing complexity. The encoded graph features are passed through a series of downsampling operations, followed by upsampling to reconstruct the fine-scale representation. Graph-transformer-based message passing is performed at each resolution level and skip connections bridge corresponding downsampling and upsampling stages to retain fine-scale spatial information and mitigate over-smoothing. Three coarsening strategies namely Four-Stage Refinement (FSR), Two-Stage Refinement (TSR) and Single-Stage Refinement (SSR) were evaluated with respect to prediction accuracy and computational efficiency. These differ in the number of mesh refinement levels removed per downscale step: FSR removes one level at a time, TSR removes two and SSR eliminates all four levels in a single operation. SSR was selected for subsequent Transfer Learning due to its lowest simulation time (3.91 h for 30 simulations) and high predictive accuracy (crack field error < 0.3%, x-displacement < 0.25%, y-displacement < 0.3%). Results suggest that for phase-field fracture, the coarsest mesh retains sufficient information for accurate prediction, rendering intermediate resolutions unnecessary. Initially, the framework is trained on Mode-I fracture cases with a single edge crack and later generalised to other configurations such as centre cracks, shear and mixed-mode loading through transfer learning. This multiscale GNN with bs-AMR yields high accuracy and up to 200 $\times$ speed-up compared to full-resolution phase-field simulations.

4.3. CNN and U-Net based approaches

While GNNs excel at modelling the irregular, mesh-based topologies inherent to fracture mechanics, a parallel and popular line of research frames the problem as an image-to-image translation task. This approach utilises CNNs and their variants to learn mappings from images to output fields such as stress, damage, or final crack patterns. Central to this approach is the U-Net architecture, whose encoder-decoder structure with skip connections has emerged as the dominant framework for these full-field prediction tasks. Key innovations include hierarchical frameworks using intermediate physical fields, the integration of self-attention mechanisms to capture non-local effects, and physics-informed inputs to reduce data-dependency. The application of CNNs, and its variants to fracture simulation and data-driven surrogate modelling is explored in this section.

4.3.1. Sequential frameworks

Contrary to the conventional approach of predicting fracture phenomenon in a single stage, some studies have developed hierarchical learning. In this approach, multiple networks are trained in sequence, where the output of a network is fed as input to the next network. It is observed that the intermediate representations are typically explainable physical fields like stress distribution in the domain of interest. This “physics-grafting” approach simplifies the learning process while helping us understand and explain how the inputs transform through the model in making the fracture prediction. Now we proceed to explore a few studies that used sequential learning.

Mohammadzadeh and Lejeune [76] proposed a surrogate modelling framework that begins with an initial dimensionality reduction stage followed by a predictive stage. First, a convolutional autoencoder (cAE) is trained in an unsupervised manner to learn a compact latent representation of high-resolution ground-truth, using only the binary damage data. This reduced-order map is later decoded to reconstruct the corresponding damage field with minimal loss of information. In the second stage, rather than directly predicting full-resolution damage fields, a regression model is trained to map material property distributions to their corresponding latent representations. This strategy simplifies the learning task by reducing it from a high-dimensional field-to-field translation to a lower-dimensional vector prediction, improving both efficiency and accuracy. To achieve this, the authors introduce the MultiRes-WNet, a deep convolutional architecture constructed by stacking two modified U-Net networks in series. Each U-Net is enhanced with multi-resolution convolutional blocks (MultiRes-Conv-Blocks), which enable multiscale feature extraction, and residual paths (ResPaths), which prevent direct concatenation of low- and high-level features. Additionally, standard convolutions are replaced

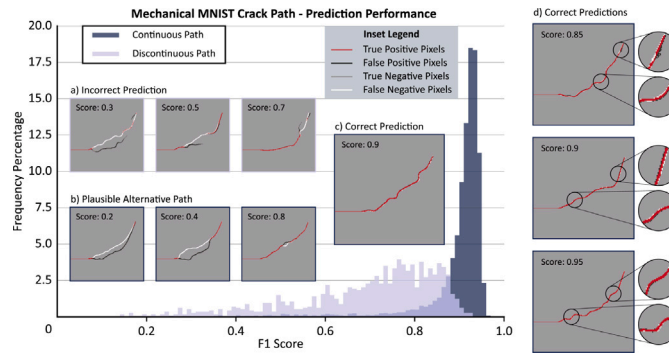


Fig. 5. Prediction performance of the MultiRes-WNet model on the Mechanical-MNIST Crack Path dataset [76]. The histogram shows the distribution of F1 scores for test samples, split into continuous (dark blue) and discontinuous (light blue) predicted crack paths. Insets: (a) Examples of incorrect predictions (discontinuous paths); (b) Plausible alternative paths (continuous but low F1 scores); (c) a high-confidence correct prediction; and (d) three correct predictions with F1 scores of 0.85, 0.90, and 0.95.

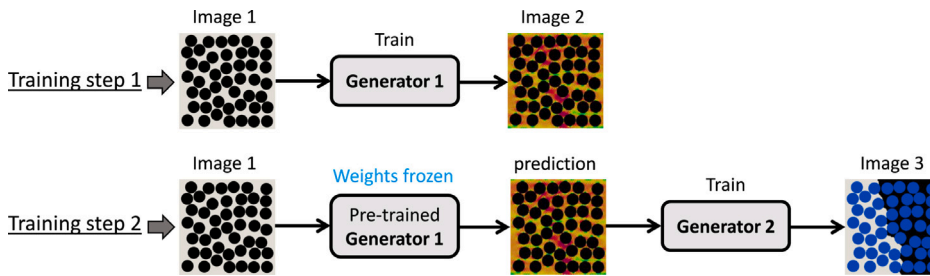


Fig. 6. Generator 1 is first trained to map the microstructure (Image 1) to the intermediate stress field (Image 2). Then, with its weights frozen, its predictions are used to train Generator 2, which maps the stress field to the final crack pattern (Image 3) [75].

with depthwise separable convolutions to improve parameter efficiency. During inference, the MultiRes-WNet predicts the latent damage representation from new input conditions, which is then decoded using the previously trained decoder to reconstruct the high-resolution damage field.

To evaluate prediction performance, the authors introduced a scoring framework based on the Mechanical-MNIST Crack Path dataset, classifying predictions as continuous or discontinuous and scoring them using F1 metric (calculated after binarising pixel values at a threshold of $\phi > 0.5$) shown in Fig. 5. Discontinuous predictions were deemed incorrect, while continuous ones with $F1 > 0.85$ were considered correct. This analysis yielded a mean F1 score of 0.87 overall, with most predictions classified as correct. This benchmark not only quantifies accuracy but also motivates the need for more structured learning strategies, such as hierarchical modelling, to further improve generalisation and interpretability.

On similar lines, Sepsadar et al. [75] developed a two-stage deep learning framework (Fig. 6) to predict both the nonlinear stress distribution and the final transverse failure pattern in fibre-reinforced composites. Their framework comprises two sequentially trained generator networks, Generator 1 and Generator 2, both based on a modified U-Net architecture. Generator 1 learns to map the microstructure to a full-field von Mises stress distribution at the early stage of damage initiation (ESoDI), while Generator 2 translates this stress field into the final crack pattern. This sequential approach outperformed direct microstructure-to-crack prediction and was trained using 4500 high-fidelity finite element simulations. The two generators share a fully convolutional encoder-decoder structure with reduced skip connections and dropout regularisation to mitigate overfitting. Generator 1 is responsible for learning intricate features from low-level geometric details, whereas Generator 2 performs a comparatively simpler task of predicting the crack pattern from the stress field generated by Generator 1. To improve learning in this more demanding stage, the authors employed an attention-based loss function instead of the standard MAE. This loss formulation helps concentrate training on physically important high-stress regions. Since the relevant high-stress zones are already embedded in the input, a standard MAE loss function was found sufficient for training this second network. The reported 90% accuracy is based on expert visual inspection rather than a standard quantitative metric. Pixel-wise measures such as MAE can be misleading for evaluating Generator 2, as visually incorrect crack patterns may still yield low MAE if the spatial differences are small. Predictions were manually classified as good, partly good, or bad based on visual agreement with the reference crack path, with partial credit assigned to intermediate cases. The visual inspection-based accuracy reflects physical correctness but remains subjective, with no clearly defined criteria for categorisation, thereby limiting the transparency and reproducibility of the evaluation.

Building on this two-stage approach, Chen et al. [137] used the self-attention mechanism as a layer acting on the latent space instead of as a loss function, for predicting both elastic stress fields and final fracture patterns in composite microstructures. This

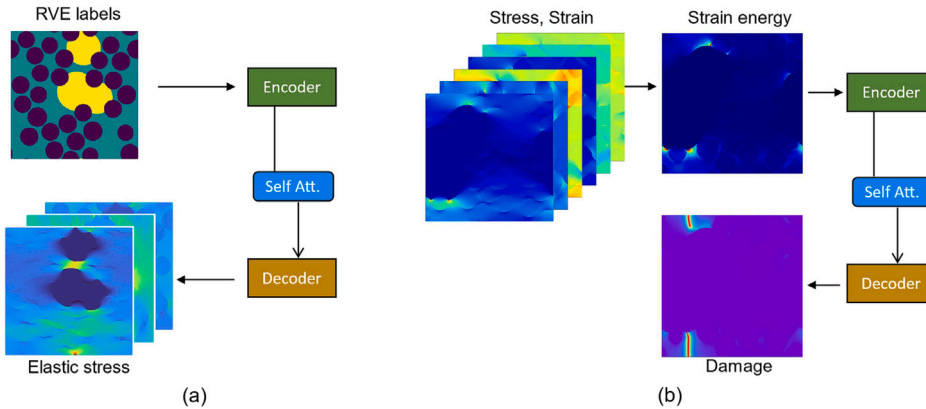


Fig. 7. Inputs and outputs for the two surrogate models proposed by Chen et al. [137]. (a) Elastic surrogate: input – phase label image (normalised), output – three stress components. (b) Crack surrogate: input – strain energy field (computed from stress and strain, then normalised), output – damage field highlighting cracks. Both networks share the same encoder–decoder CNN architecture with a self-attention mechanism.

architectural modification enhanced the ability to capture non-local dependencies like distant microstructure features influencing the stress field at the current point. Fig. 7 illustrates two surrogate frameworks. The first surrogate model predicts elastic stress fields from the input Representative Volume Element (RVE), represented as an image where each pixel corresponds to a discrete material phase label (e.g., matrix, inclusion, void). To standardise the input scale, the image is normalised by dividing all pixel values by the maximum phase label. In the second surrogate, the strain energy field is used as the input to predict crack patterns, as strain energy density (SED) quantifies the local thermodynamic driving force for fracture. By mapping the stored elastic energy distribution, the model is grounded to established fracture mechanics theory to identify energetic hotspots where the critical energy threshold is likely to be exceeded, creating conditions favourable for fracture. The authors initially considered an end-to-end surrogate pipeline, where the output of the first surrogate served as the input to the second. This led to inaccurate crack predictions due to the accumulation of errors from the elastic surrogate. Thus, a hybrid approach was adopted, in which the crack surrogate took its input directly from the high-fidelity Fast Fourier Transform (FFT) simulation. Despite not being used in the crack pipeline, the elastic surrogate remains useful for fast stress/strain evaluations in multiscale or multiphysics models. The training dataset, covering a range of fibre volume fractions and porosities, included about 7000 high-fidelity simulations which are further augmented by translation on the fly during the training. The crack prediction model achieved an F1 score of ~ 0.8 (evaluated after binarising the continuous damage field with a threshold of $\phi > 0.5$) and a speed-up of 3–4 orders of magnitude compared to the full field FFT solver.

4.3.2. Iterative spatiotemporal prediction

While predicting the final fracture state is valuable, a more ambitious goal is to predict the dynamic evolution of a crack over time. This transforms the problem from image-to-image translation to sequence prediction, capturing the full history of damage accumulation.

Xu et al. [83] developed Crack-Net, a spatiotemporal deep learning framework aimed at predicting the full fracture process by jointly modelling the crack propagation path and the corresponding stress–strain response. The architecture employs U-Net and operates in an iterative loop, where the state at each subsequent time step is predicted from the current state. A single latent vector encodes the system state and is used to simultaneously predict both the crack phase field and the stress response, thereby ensuring physical coupling. This is achieved by modelling the log-transformed increment of the crack phase field, which helps capture quasi-static variations more effectively.

The model was trained on a relatively small dataset comprising 554 phase-field simulations of particle-reinforced composites, yet demonstrated strong generalisation. Crack-Net accurately predicted fracture behaviour in interpenetrating phase composites (a structure not seen during training) suggesting that the model had learned underlying physical principles affecting the flow of fracture (Fig. 8). Furthermore, by leveraging transfer learning, the model could be adapted to new material systems using as few as 120 additional simulations. The authors claim a speed up of 95% compared to conventional finite element simulations. Due to its sequential/iterative nature, Crack-Net is prone to error accumulation during long-term predictions, which can lead to deviations from the ground truth over extended time horizons.

Addressing the problem of error accumulation in iterative methods, Hsu and Buehler [88] proposed a Deep Learning (DL) architecture capable of generating an entire fracture history in a single pass. Their 2D to 3D design processes a single 2D multi-channel input image with a convolutional encoder, a standard unidirectional (causal) LSTM for temporal embedding and a 3D deconvolutional decoder to reconstruct full spatio-temporal crack evolution. The LSTM is trained on complete forward sequences and therefore learns causal dynamics that are effectively invertible. This enables a single-pass reconstruction of the full history from any time point, i.e. both forecasting final crack pattern from initial states or backcasting the cause from final fracture patterns become possible without bidirectional processing or iterative drift. Gradient-weighted Class Activation Mapping (Grad-CAM) visualisation and interpretation demonstrated the model's focus on crack-tip regions during early stages, validating its physical reasoning. While

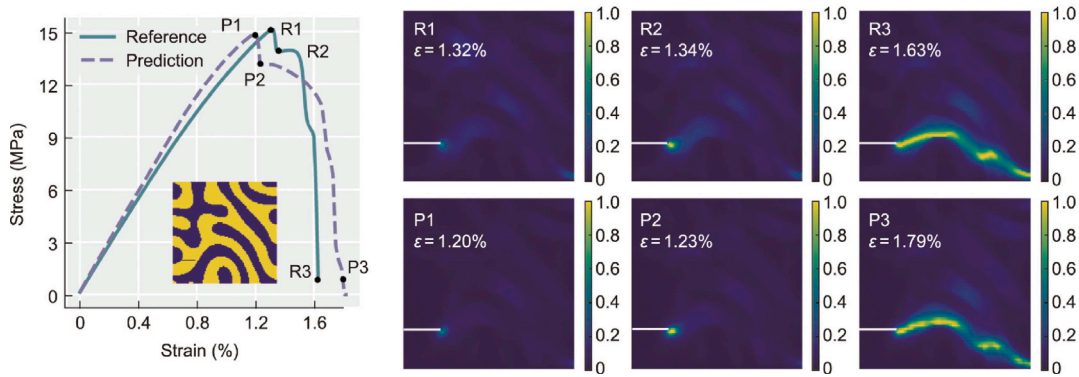


Fig. 8. Comparison of predicted and reference stress–strain responses for a heterogeneous material, with corresponding damage field visualisations at selected strain levels. The left plot shows the overall mechanical response, where solid and dashed curves denote reference and predicted behaviour, respectively. The inset illustrates the material’s microstructure. Right sub-figures display the damage evolution at representative strain points: R1–R3 for the reference and P1–P3 for the prediction. Colour maps indicate normalised damage values ranging from 0 (intact) to 1 (fully damaged) [83].

it represents a significant advancement in spatio-temporal modelling, its success depends on careful encoding of input conditions and may be less flexible to parameter adjustments.

On similar lines, another spatiotemporal U-Net framework was developed by Koopas et al. [72] for the full-field prediction of the entire fracture process in concrete meso-structures having explicitly modelled Interfacial Transition Zone (ITZ). To tackle the challenge of data dependency, they included the physical properties of the domain as an input to the NN. Instead of a simple image-to-image mapping, the U-Net is fed a four-channel tensor containing the spatial distribution of Young’s modulus, Ultimate Tensile Strength (UTS), fracture energy and the damage field from the previous time step. This approach enabled successful training on a minimal dataset of only 470 simulations. The authors claim a $2864\times$ speed-up over the full FEA, with an F1 score of 0.6 (Fig. 9) on unseen test cases. Here, the F1 metric was computed by binarising the continuous damage field using a threshold corresponding to the 99th percentile of the ground truth FEM damage distribution. By strictly isolating the top 1% of most damaged elements, this threshold focuses the evaluation exclusively on critical fracture localisation, filtering out negligible background deformations.

It is important to note that crack branching predictions in both Koopas et al. [72] and Xu et al. [83] were found to be sensitive to the mesh discretisation and the regularisation length parameter, suggesting limitations in generalising across different phase-field resolution settings.

To predict time-series quantities influenced by evolving spatial data, Wang et al. [80] developed StressNet, a dual-branch architecture that fuses two distinct data streams. One branch uses a Bidirectional LSTM (Bi-LSTM) to process historical stress data, while a second branch uses a custom Temporal Independent-CNN (TI-CNN) to extract spatial features from damage-field images each time step. While effective, this model operates as a black-box surrogate, learning correlative rather than causal relationships from the high-fidelity simulation data. The model’s demonstrated high accuracy (average MAPE of 2%) and its compelling computational gain predicting a full sequence in 20 s versus 4 h for the FDEM simulation is impressive but specific to the fixed geometry, material properties and uniaxial tensile loading conditions of the training dataset.

Feng et al. [130] approached this temporal damage evolution problem in a new light to estimate the Residual Useful Life (RUL) of a structure. Their framework first uses the phase-field method to generate a sequence of structural displacement fields that represent the entire crack propagation process. These fields are then saved as images and binned into discrete intervals, each corresponding to a specific stage of the structure’s useful life. A LeNet-5 CNN is trained to classify a given displacement image into the correct RUL bin. The training, validation and testing datasets are all derived from different time steps of a single simulation run with a fixed geometry and loading condition. The authors demonstrate their model’s ability to classify the state of a known system, but its capacity to generalise to new geometries, material properties or loading conditions is unevaluated.

4.3.3. Tackling data scarcity and interpretability

Data scarcity and the “black-box” nature of models remain two of the main barriers to adopting deep learning in engineering. Mohammadzadeh and Lejeune [76] addressed part of this challenge by introducing the Mechanical-MNIST Crack Path dataset, a public benchmark derived from high-fidelity simulations in which heterogeneous material properties were varied to generate diverse crack paths. Beyond this, the field is increasingly exploring strategies such as active learning, Explainable AI (XAI), and advanced architectures including LSTMs and transformer-based diffusion models to improve data efficiency, interpretability, and robustness against error propagation.

Huang et al. [117] developed a Hybrid Neural Network (HNet) that directly addresses both issues. To predict a crack path on a point-by-point basis, the hybrid input structure simultaneously processes a greyscale image representing the structure’s geometry and a numerical x -coordinate to predict the corresponding y -coordinate of the crack path. This method allows a single model to be applied

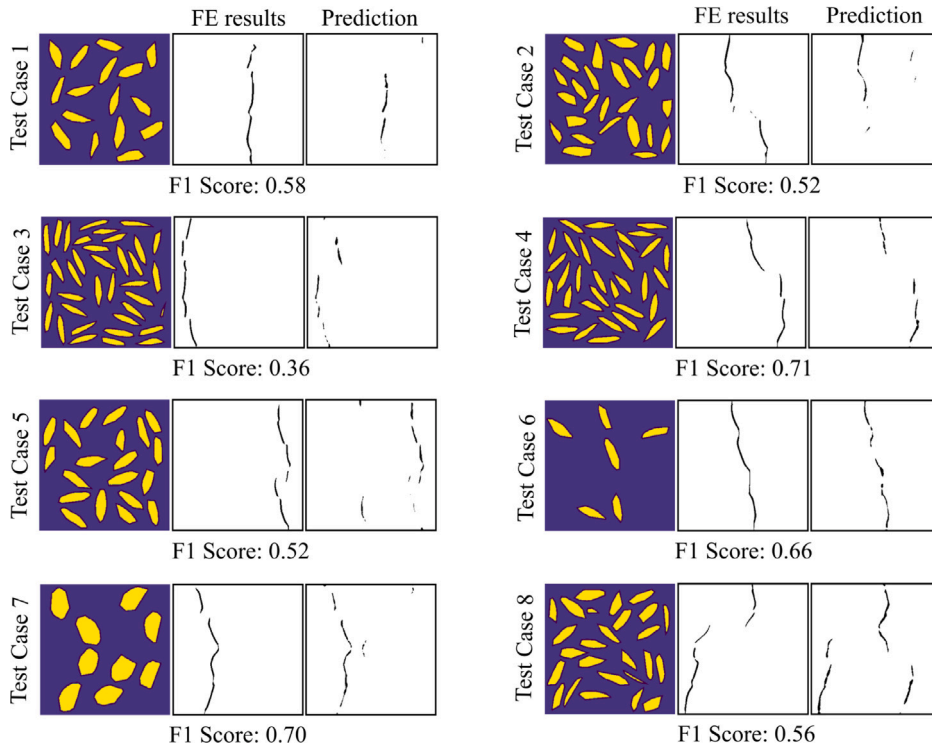


Fig. 9. Quantitative evaluation of the performance of the developed spatiotemporal surrogate model in the unseen test cases using the F1 score [72].

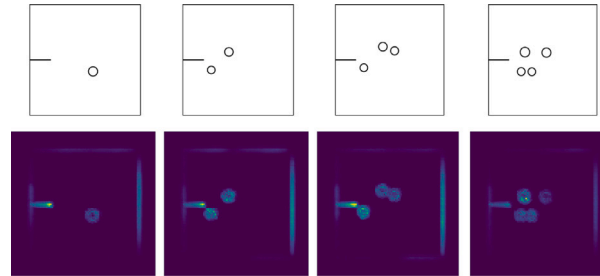


Fig. 10. Application of XAI by Huang et al. [117] The top row shows input images with different hole configurations. The bottom row shows the corresponding saliency maps, where brighter pixels indicate regions the HNet model identified as most important for its prediction, correctly highlighting the initial crack and the holes.

to various geometries and defect types without altering its input parameters. The architecture's dual-branch design comprising a CNN for spatial feature extraction and a fully connected pathway for coordinate conditioning, exemplifies a minimal yet meaningful hybridisation of image-based and scalar inputs. The authors claim to enhance transparency using XAI, where saliency maps confirm that the model correctly focused on physically important regions like the crack tip and nearby defects (Fig. 10). Training data scarcity is tackled using a query-by-committee active learning algorithm. Their approach is shown to achieve the same prediction accuracy as the baseline model while reducing the required training data from 60,000 to just 25,000 samples, a reduction greater than 50%. Although the hybrid design generalises well across geometries, its dependence on point-wise sequential inference may constrain scalability for full-field predictions or introduce compounding errors in long-range path forecasting. This limitation becomes crucial when the method is used to iteratively reconstruct an entire crack path, where local inaccuracies may accumulate over time.

While the reviewed studies demonstrate the abilities of CNN-based surrogate models in accelerating fracture prediction, several inherent challenges persist. A primary limitation is the substantial computational cost associated with generating the high-fidelity simulation datasets necessary for training these models. Methodologically, the reliance of these models on largely image-based data representations introduces the risk of information loss during dimensionality reduction processes such as autoencoding or when converting continuous simulation outputs into discretised image formats. Such transformations can obscure critical features related to crack initiation and propagation.

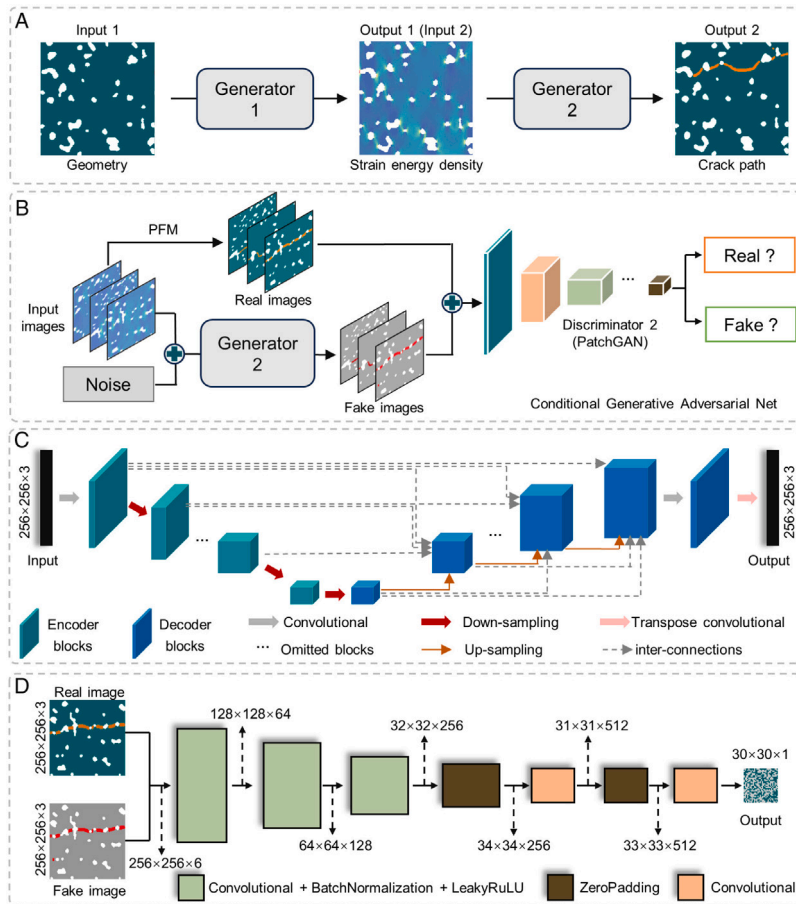


Fig. 11. (A) The two-step strategy, translating geometry to an intermediate SED field, then to a crack path. (B) The cGAN structure. (C) The U-Net 3+ architecture of the generators. (D) The PatchGAN architecture of the discriminators [129].

Additionally, the above-described two-stage or “grafting” strategy, where intermediate physical fields (e.g., elastic stress or strain fields) are first predicted and then used to infer fracture patterns while simplifying the learning objective, is vulnerable to error accumulation. Inaccuracies introduced during the initial prediction stage can propagate and amplify in the subsequent stages, potentially compromising the fidelity of the predicted fracture evolution.

4.4. Generative models

Generative models in fracture mechanics aim to capture the distribution of possible fracture patterns and produce physically plausible outcomes conditioned on material or structural parameters. He et al. [129] developed a cGAN-based framework to predict discontinuous crack paths in random porous media directly from microstructure images. In the first stage, the microstructure is transformed into an intermediate elastic strain energy density (SED) field, capturing the underlying physical behaviour of the material. This physically-informed field aids in predicting non-linear crack paths in the second stage. Using the intermediate SED field helped the model predict crack paths more accurately, improving overall accuracy from 81.25% to 92.75% compared to a model without the intermediate step. The final crack paths achieved 90.25% accuracy, and the predicted SED field had an average error of 0.015. Fig. 11 shows the two-step strategy, the cGAN structure, and the architectures of the generators and discriminators. Structurally, both stages utilise U-Net 3+ generators with full-scale skip connections that capture the multi-scale interactions and preserve low-level geometric details. To avoid the blurring effect of standard loss functions, the authors employed a Markovian ‘PatchGAN’ discriminator. Unlike standard global discriminators that assign a single scalar validity score to the entire image, this architecture evaluates local $N \times N$ patches to distinguish between ground truth samples and generated approximations. By applying this classification convolutionally across the domain and averaging the resulting validity map, the model enforces high-frequency textural fidelity rather than just global consistency. Notably, this strategy of conditioning the GAN on the intermediate physical SED field further mitigates the training instability and mode collapse issues often observed in GANs, ensuring that predicted crack paths remain physically consistent with the local stress state.

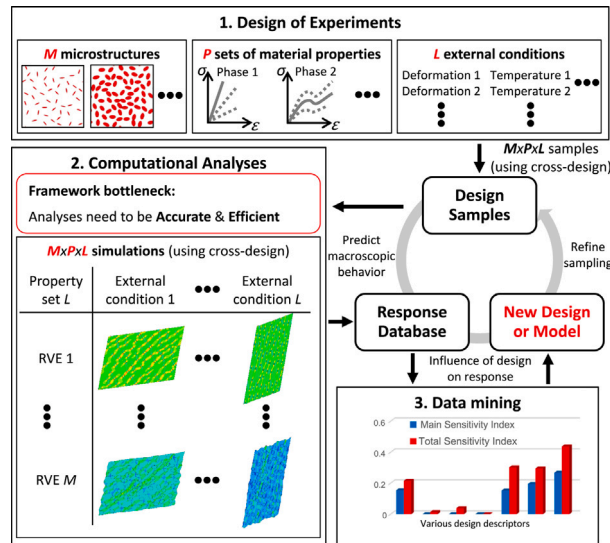


Fig. 12. System-level view of the data-driven reduced-order modelling framework by Bessa et al. [114]. The approach begins with a cross-design of M microstructures, P material properties and L external conditions to generate simulation samples. To overcome the computational bottleneck, SCA is used as a surrogate model, replacing full-field simulations with efficient cluster-based approximations. Simulation results are stored in a response database, which is mined to assess design sensitivities and guide new model generation. This closed-loop pipeline enables scalable exploration of high-dimensional design spaces in fracture and materials modelling.

Similarly, Huang and Kanai's DeepFracture framework [77] predicts 3D brittle fracture patterns from runtime collision dynamics, primarily for computer graphics, but it demonstrates the ability of generative models to learn complex fracture distributions, produce high-resolution geometries, and condition outputs on physical parameters such as collision impulses. Using a neural implicit representation and a GAN trained with a BEM-based loss, DeepFracture reduces pattern generation time from 18.7 min for BEM simulations to 6.19 s, while maintaining approximate physical plausibility. Together, these studies highlight how generative approaches can model stochastic fracture behaviour and efficiently generate diverse, physically informed outcomes applicable to engineering-focused fracture ML. Furthermore, by employing a neural implicit representation with discrete latent codes (VQ-VAE), this framework avoids the mode collapse limitations of continuous GANs, allowing it to capture distinct, multi-modal fracture patterns effectively.

4.5. Data-driven reduced-order modelling

ROMs aim to reduce the computational cost of high-fidelity simulations by constructing surrogates that capture essential system behaviour in a lower-dimensional space. Unlike end-to-end predictive models, ROMs are often embedded within simulation workflows or hybrid architectures and can leverage both classical numerical methods and machine learning for approximation. While traditional ROMs have relied on projection-based techniques such as Proper Orthogonal Decomposition (POD) and Galerkin projections, recent approaches incorporate data-driven strategies like clustering [145,146], autoencoding [112] and tensor decomposition [135,147]. These non-intrusive ROMs prove advantageous in fracture simulations where non-linearities, path dependencies and topological changes make classical linearisation insufficient. In this section, we survey four representative ROM frameworks that exemplify these advances: self-consistent clustering surrogates, graph-theoretic ROMs, tensor-decomposition models and deep auto-encoder based methods.

A primary challenge in data-driven modelling is the creation of a sufficiently large and diverse database, which can be prohibitively expensive for simulations involving history-dependent phenomena like plasticity and damage. Bessa et al. [114] address this by proposing a general-purpose data-driven framework centred on the Self-Consistent Clustering Analysis (SCA). SCA is a non-intrusive ROM that accelerates the analysis of heterogeneous materials by grouping regions of a high-fidelity model into a smaller number of "clusters" that share a similar mechanical response. The offline stage of SCA requires only a few linear elastic simulations to create this reduced representation, which can then be used to rapidly and accurately predict the full nonlinear, inelastic response under new loading conditions. By using SCA as an efficient database generator, the framework makes it feasible to apply machine learning algorithms to explore high-dimensional design spaces and derive data-driven constitutive models for complex fracture and failure processes. Fig. 12 provides a system-level overview of this integrated pipeline.

Hunter et al. [119] tackle the specific challenge of brittle fracture by leveraging graph representations of crack networks. They evaluated four ROM strategies for brittle fracture prediction using 185 simulations generated via a FDEM. The models reflected a range of modelling philosophies, including geometric simplification, heuristic process zone characterisation and physics-informed learning. These included the Orthogonal Projection (OP) model, which assumed horizontal crack growth; the Network-based Fracture

Process Zone (NFPZ) model, which identified likely failure paths using shortest-path algorithms guided by process zone theory; the Ellipse Process Zone (EPZ) model, which used learned heuristics to define elliptical fracture zones; and the Micro-Crack Pair Informed Coalescence (McPIC) model, which predicted failure through interaction dynamics between crack pairs.

Among the four, McPIC demonstrated the highest predictive fidelity. It employed a neural network trained on physically relevant features including crack length, orientation, spacing and SIFs to capture the dynamics of crack coalescence. This model correctly predicted full fracture paths in 45% of non-branching cases and provided time-to-failure estimates closely aligned with high-fidelity results. In contrast, the other models showed diminished accuracy, often collapsing to mean predictions. The study emphasises the need for integrating domain knowledge into the training process, showing that ROMs combining physics-informed features with trainable architectures outperform purely geometric or heuristic approaches.

For parametric studies where fracture behaviour is explored across a range of material properties or loading conditions, Duong et al. [135] propose a non-intrusive ROM approach that combines tensor-train decomposition and Gaussian Process Regression (GPR) to efficiently predict the evolution of fracture fields. The spatio-temporal evolution of crack solved using PFM along with initial state parameters of fracture energy and initial crack length formed the training data tensor. Tensor-train decomposition algebraically compresses this large data tensor into a product of much smaller, low-rank cores effectively decoupling the influence of each dimension. A GPR model is trained on this compressed data to map new parametric inputs to the corresponding low-dimensional tensor-train core representation. The full-field solution is then reconstructed instantly. This approach demonstrates up to 95% accuracy with a 100× speed-up and aids efficient surrogate modelling in parametric design spaces.

Shinde et al. [123] introduce a deep learning ROM based on cAE for dimensionality reduction and surrogate prediction in phase-field fracture mechanics. The model treats 2D displacement and damage fields as multi-channel images and uses a cAE to learn a compressed, low-dimensional latent representation of the system's state. The architecture separates the compression task (handled by the cAE) from the forward prediction task (handled by a deep neural network), allowing for modular training. A separate deep neural network is trained to act as a predictor, taking the current latent state and applied loading as input to predict the latent state at the next time step in a recurrent loop. This approach is validated under both proportional and non-proportional loading scenarios, demonstrating strong generalisation across varying crack paths and initiation points.

The reviewed data-driven approaches adopt fundamentally distinct strategies for reduced-order modelling in fracture mechanics. Clustering-based methods focus on grouping similar responses to construct efficient surrogates from representative samples, offering a compact physical model that enables rapid data generation well-suited for accelerating constitutive simulations and uncertainty quantification. Graph-based models encode the topological structure of discrete fracture networks, making them suitable for problems characterised by crack interaction, growth and coalescence. Hybridised GNN-ROMs are used for studying non-parametric geometries in structural dynamic problems [148]. Tensor-based ROMs specialise in high-dimensional parametric studies, leveraging low-rank decompositions to interpolate simulation outputs across varying material properties and boundary conditions. Auto-encoder based models address high-resolution, image-like field data, using deep learning to capture non-linear fracture dynamics and generalise to unseen geometries or loading scenarios. Collectively, these complementary methods define a versatile design space in which hybrid strategies combining topology-awareness, spatial compression and probabilistic inference can be tailored to the complexity and computational demands of modern fracture simulations. The future of ROMs in fracture mechanics likely lies in hybridising these complementary approaches – using a GNN to capture crack topology within an auto-encoder based framework for field compression – to create models that are both computationally efficient and physically robust.

4.6. Physics-informed neural networks

In the context of ML applications to fracture mechanics, data quality and quantity play a pivotal role in model development. While increasing the amount of training data is a conventional approach to enhance reliability, it is not without pitfalls. Beyond a certain threshold, additional data may lead to over-fitting or introduce spurious correlations that degrade model performance. These limitations have prompted researchers to explore hybrid strategies that incorporate governing physical laws directly into the learning process, either by integrating them in the loss function or directly in the neural network architecture [149]. In this direction, PINN have gained significant attention and are being applied, either directly or through their variants, to solve fracture mechanics problems. PINN-based methods have been applied to phase-field models to solve fracture problems by solving Cahn–Hilliard and Allen–Cahn equations where the damage field is the only unknown scalar [150]. It is challenging to design neural networks for predicting fracture, using phase-field models, due to the interlinked coupling between the vector-valued displacement field and the highly localised scalar phase field (driven by the regularisation length scale) that induces sharp spatial gradients near cracks. Recent efforts have addressed these challenges in several ways: PINNs have been integrated with discontinuous Galerkin methods to better represent crack interfaces [151]; Adaptive domain refinement has been employed to concentrate computational effort near crack tips, where solution gradients are highest [152]. Enhancements in boundary condition enforcement, such as the use of Lagrange multipliers [153] and constrained optimisation, have improved model robustness [154]. Additionally, peridynamic energy minimisation [155,156] and level-set representation [157] based approaches have extended the applicability of PINNs to more complex geometries and evolving crack topologies [158]. Despite these advances, PINNs continue to struggle with sharp discontinuities, remain computationally intensive in 3D applications, and face challenges in generalising across diverse material systems and loading scenarios. In this context, the remainder of this section synthesises key advances in PINN-based fracture modelling. The narrative will explore enriched architectures, hybrid solvers, energy-based formulations, and scalable learning frameworks. By grounding the discussion in both methodological innovations and application-specific insights, we aim to highlight the advances in physics-informed learning from theoretical constructs to practical tools for computational fracture mechanics.

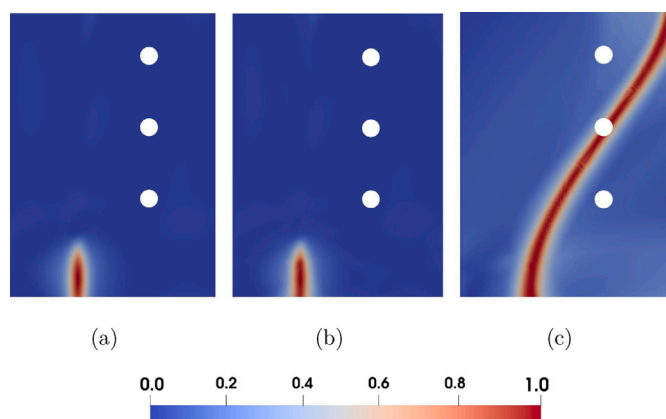


Fig. 13. Phase-field evolution for a 2D plate with three circular inclusions, predicted by the VE-PINN framework. Subfigures show crack initiation (a), intermediate step (b) and full propagation (c) at increasing displacement values. The crack naturally navigates around the inclusions, demonstrating VE-PINN's capability to resolve complex, geometry-induced fracture paths [95].

An early enhancement to the standard PINN framework for fracture mechanics was proposed by Goswami et al. [95], introducing several methodological advancements to develop a robust and efficient model. In this approach, instead of minimising the residual of the governing PDEs, the learning process minimises the system's total variational energy. This modification improves stability in the presence of discontinuities and reduces computational effort by relying on lower-order derivatives compared with conventional residual-based PINNs. To accurately and efficiently evaluate the total variational energy, the geometry is modelled using NURBS, and numerical integration is performed with Gauss–Legendre quadrature, offering better accuracy and efficiency than Monte Carlo sampling. Dirichlet boundary conditions are hard-enforced by embedding them directly into the network architecture, avoiding ill-conditioned optimisation problems and manual tuning of penalty terms in the loss function. Neumann boundary conditions, which involve derivative constraints, are more challenging to enforce and are therefore treated using a soft-penalty approach within the loss function. These methodological choices collectively enhance the stability, efficiency, and predictive accuracy of the framework for fracture mechanics applications. To make quasi-static fracture simulations computationally feasible, the authors propose a transfer learning strategy. Once the solution for a given load step is obtained, the network for the next step is initialised using the previously trained model. The weights of the early layers are frozen, and only the final layer is retrained, which accelerates convergence considerably and enables the optimiser to reach a lower energy state in fewer iterations. The authors illustrate the advantage of using transfer learning in Variational Energy-based PINN (VE-PINN) by comparing convergence of the loss function with and without parameter reuse from previous iterations. With transfer learning, the algorithm converges more rapidly and achieves a lower final loss (variational energy). The computational time per iteration is reduced by approximately half. Using six benchmark problems, the authors show that VE-PINN outperforms conventional residual-based PINNs in handling discontinuities across 1D, 2D, and 3D domains. Its ability to accurately capture a deflected crack path in a complex 2D geometry is visually demonstrated in Fig. 13.

Building on this, Motlagh et al. [97] conducted an extensive comparison of standard PINN, Variational PINN (V-PINN) and VE-PINN applied to phase-field fracture. VE-PINN emerged as the most accurate approach overall, reducing displacement errors (L^2 norm) by 30–34% and stress errors (semi- H^1 norm) by up to 68% compared with both standard PINN and V-PINN across 1D and 2D benchmarks (see Table 2 and Figures 9–11 in [97]). The study also explores different methods for imposing boundary conditions, including soft and hard enforcement and a self-adaptive PINN approach. A further contribution is the demonstration of using VE-PINN for a data-driven inverse problem to discover the material-specific phase-field length scale parameter l_0 with less than 2% error.

A critical observation here is that, VE-PINN exhibits significantly higher sensitivity to the choice of activation function than residual-based PINNs. The authors show that ReLU consistently yields physically realistic crack propagation speeds and paths that closely match high-fidelity finite-element reference solutions, whereas tanh (and to a lesser extent softplus) produces unphysically rapid crack growth and premature complete fracture as shown in Fig. 14. The variational energy formulation in VE-PINN directly minimises a global scalar potential that couples displacement and phase-field evolution. Smooth, saturating activations like tanh tend to over-regularise the highly non-convex energy landscape near sharp crack fronts, effectively smearing the fracture surface energy and allowing the optimiser to reduce the total energy by advancing the crack too aggressively. In contrast, ReLU's piecewise-linear nature introduces sharper transitions that better approximate the threshold-like tensile failure and irreversibility inherent in brittle fracture, resulting in propagation speeds and paths that respect the underlying physics. This extreme hyperparameter sensitivity is much more pronounced when variational principles are enforced through universal approximators. Superior theoretical convergence properties do not guarantee robustness unless the function class aligns well with the solution's regularity (or lack thereof). The observation therefore serves as both a practical warning and a valuable insight into the interplay between network architecture and thermodynamic consistency in energy-based physics-informed learning.

Chakraborty et al. [107] extended the VE-PINN framework to a domain-decomposed formulation Variational Energy-based XPINN (VExPINN). The framework partitioned the global domain into multiple non-overlapping subdomains (see Fig. 15), each

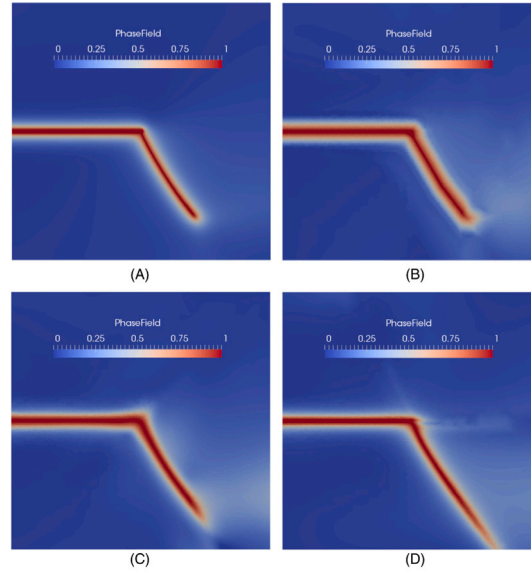


Fig. 14. Comparison of crack propagation paths using different activation functions in VE-PINNs. (A) Finite element ground truth. (B) Deep learning prediction with ReLU. (C) Prediction with softplus. (D) Prediction with tanh [97].

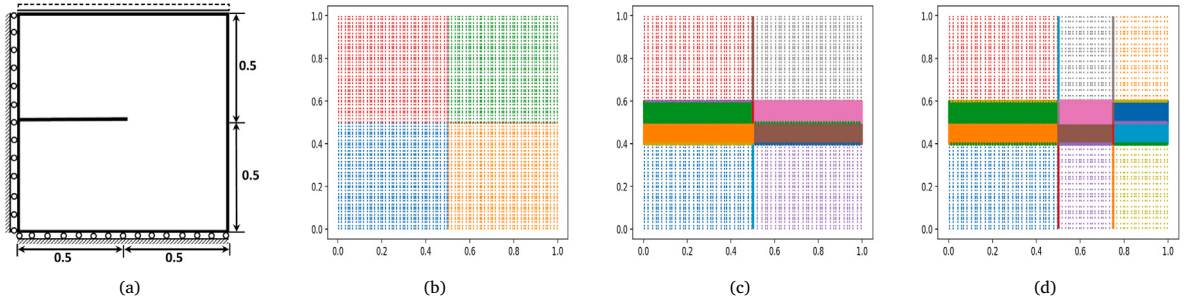


Fig. 15. Domain decomposition and adaptive refinement strategy of VExpINNs from Chakraborty et al. [107]. (a) shows geometry setup and boundary conditions for the single-edge notched tension test with units in mm. (b)–(d) Distribution of training data over 4, 8 and 12 subdomains.

trained independently to minimise the variational energy. This design enables parallelisation and localised refinement near crack tips, improving computational efficiency and preserving high-resolution predictions. Additionally, noting from their previous works [95], the authors implement adaptive activation function scaling and fourth-order phase-field models to further enhance solution regularity with fewer degrees of freedom, improving convergence and accuracy. The VExpINNs idea is demonstrated using numerical experiments on crack kinking, branching and coalescence without prior crack path knowledge.

Gu et al. [96] proposed enriched PINN (E-PINN) for 2D LEFM problems by embedding asymptotic crack-tip solutions directly into the network output. This approach analytically captures the singular components of the displacement field, allowing the neural network to learn only the regular part. By introducing the Mode I and II SIFs (K_I , K_{II}) as trainable parameters, the method enables direct computation of SIFs within a fully meshless framework.

The enriched displacement field, $\mathbf{u}$, is formulated as:

$$u_1(x, \theta_{NN}, K_I, K_{II}) = u_{1,NN}(x, \theta_{NN}) + K_I f_{1,I}(r, \theta) + K_{II} f_{1,II}(r, \theta) \quad (3)$$

$$u_2(x, \theta_{NN}, K_I, K_{II}) = u_{2,NN}(x, \theta_{NN}) + K_I f_{2,I}(r, \theta) + K_{II} f_{2,II}(r, \theta) \quad (4)$$

where u_{NN} is the output of a standard neural network with trainable parameters θ_{NN} (weights and biases) and $f(r, \theta)$ represents the analytical enrichment functions derived from LEFM.

The network learns the optimal SIF values simultaneously with the network weights and biases by minimising a loss function that enforces the governing PDEs and boundary conditions. This allows for the direct calculation of SIFs without post-processing. The reported relative error was as low as 1.42×10^{-3} , with strong accuracy in resolving crack-tip stress fields (Fig. 16).

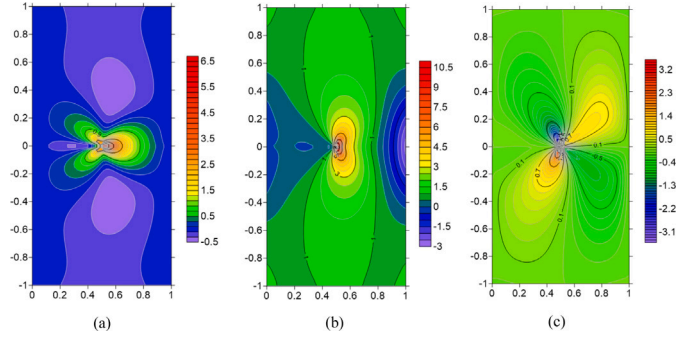


Fig. 16. Stress contours around the crack tip showing σ_{11} , σ_{22} and σ_{12} as captured by the enriched PINN from Gu et al. [96].

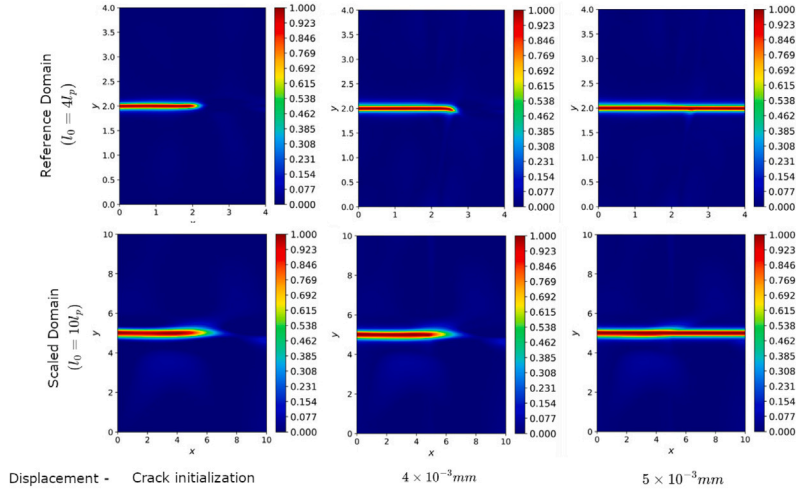


Fig. 17. Crack propagation in scaled domains using length-scale decoupling from Lian et al. [99]. The top row corresponds to a reference domain with $l_0 = 4l_p$; the bottom row shows a domain scaled 10× larger with $l_0 = 10l_p$. Columns indicate different loading stages: (Left) Crack initialization, (Centre) $d = 4 \times 10^{-3}$ mm, and (Right) $d = 5 \times 10^{-3}$ mm. The consistent damage evolution across scales validates the decoupling strategy.

Lian et al. [99] proposed a physics-informed approach within the VE-PINN framework to explicitly decouple the physical and numerical length scales involved in fracture modelling. Central to their method is a novel degradation function, controlled by a tuneable parameter q , which allows the physical length scale l_p associated with material fracture behaviour to be treated independently of the numerical length scale l_0 , which is typically constrained by mesh resolution. When $q \approx 1$, the model maintains accurate peak stress predictions even in cases where the mesh size (and hence l_0) is much larger than l_p . This decoupling enables the use of significantly coarser meshes while preserving physical realism. As a result, the method allows simulations over domains four to ten times larger without increasing the number of integration points, substantially reducing computational cost. This advancement addresses a major limitation in large-scale fracture simulations and extends the capabilities of previous VE-PINN approaches by enabling scalable and accurate modelling across length scales. The effectiveness of this decoupling strategy is illustrated in Fig. 17, which shows that fracture patterns remain consistent across domain sizes and resolutions.

Pantidis et al. [101] proposed the Integrated Finite Element Neural Network (I-FENN), combining Helmholtz-regularised finite element damage models with neural networks to account for history effects. For improved handling of time-dependent behaviour, temporal convolutional network (TCN) were integrated to produce I-FENN-TCN [108]. The enhanced model achieved errors under 3% in peak load and 5% in energy dissipation predictions.

To optimise the performance of PINNs within this framework, a systematic study on the effect of depth, width, and number of neurons in a neural network was explored [103] by varying the aspect ratio of the network (defined as width divided by depth) from 0.2 (very deep and narrow) to 5.0 (very shallow and wide), while keeping the total number of trainable parameters fixed. The experiments revealed that networks with near unity aspect ratios (i.e., width-to-depth between 0.8 to 1.2) consistently provided the best trade-off between numerical accuracy and computational cost. NN constructions that are either very deep (aspect ratio = 0.2) or very wide (aspect ratio = 5) exhibited slower convergence, higher prediction errors, and increased training loss. Fig. 18 presents results for networks with $n = 420$ trainable neurons. The full study [103] includes additional results for $n = 300$ and $n = 560$, which

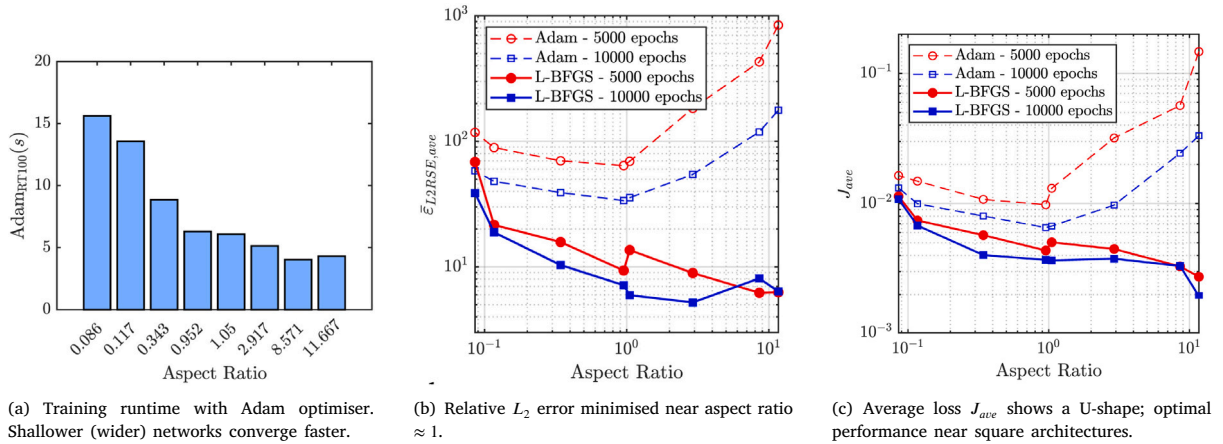


Fig. 18. Impact of network aspect ratio on training time, prediction error and loss for networks with $n = 420$ trainable parameters. Results highlight trade-offs in PINN architectures between depth and width [103].

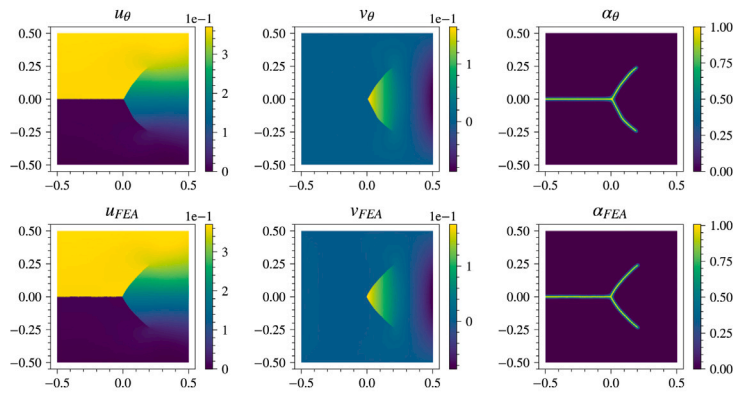


Fig. 19. Crack branching as predicted by DRM compared with FEMs reference solutions from Manav et al. [98].

showed consistent trends. These findings offer practical guidelines for designing efficient PINN architectures for non-local damage modelling. Further to this, in their recent work [159] they employed physics-informed convolutional neural network (PICNN) to predict the phase-field variable ϕ from strain energy density maps, while finite element methods handle equilibrium equations. The highlight of this work being that the training took only 5 min while offering substantial computational speed-ups scaling with mesh resolution. In a more recent work, Konale et al. [160] employed PINN for gradient-free fracture modelling, formulating the loss to enforce energy balance and irreversibility directly without explicit damage gradients. Applied to quasi-brittle materials under tensile loading, the model predicts crack paths and stress fields by minimising residuals of the governing PDEs. The authors demonstrate qualitative agreement with phase-field benchmarks on 2D notched specimens, highlighting reduced computational cost over traditional solvers while also noting limitations in handling complex 3D geometries. This work underscores the potential of pure PINNs for interpretable fracture surrogates, paving the way for integration with operator-learning frameworks discussed next.

Building on the Deep Ritz method (DRM) introduced by Yu et al. [161], which reformulates variational problems through neural network approximators, Manav et al. [98] employed DRM to model a wide range of fracture phenomena including nucleation, propagation, kinking, branching, and coalescence. In this work, variational fracture energy is minimised using a unified neural network that jointly learns displacement and damage fields. Key methodological innovations include parametrically scaled ReLU activations, finite element-inspired numerical gradient evaluations, and the adoption of the Rprop [70] optimiser to address training instabilities. Despite its strong agreement with finite element results and its ability to capture complex crack topologies without predefined paths or notches, the DRM remains computationally expensive due to the large number of quadrature points required, and crack initiation in notch-free domains still depends on artificial scaling of the regularisation length. The high visual similarity observed in crack branching (Fig. 19) justifies its potential for future developments in parametric fracture modelling and uncertainty quantification. This is particularly evident when complemented by physics-informed encoder-decoder networks that achieve over 90% spatial accuracy with orders-of-magnitude faster inference [116].

These developments collectively mark the evolution of PINN-based methods for fracture mechanics from early formulations minimising residuals to energy-based, enriched, and scalable frameworks. While challenges remain in 3D generalisation, discontinuity handling, and optimiser robustness, the reviewed methods demonstrate growing maturity in bridging physics and deep learning for fracture prediction and simulation. In the following section, we explore neural operators that are designed to learn the underlying physics for a range of boundary and initial conditions.

4.7. Surrogate modelling with neural operators

Neural operators are a class of surrogate modelling tools that have been widely explored for engineering applications. Unlike conventional neural networks that map between finite-dimensional spaces, they learn mappings between infinite-dimensional function spaces. While traditional networks approximate a function $y = f(x)$, neural operators learn an entire family of solutions, making them well-suited for parametric studies and solving PDEs. By directly approximating the solution operator of a PDEs system, (solution = $\mathcal{G}(\text{initial condition, parameters})$), predictions can be made rapidly for a wide range of initial conditions and parameters. This section reviews applications of neural operators to fracture problems.

Initial explorations into applying neural operators for fracture established two distinct philosophies: leveraging physical laws to guide learning in data-sparse environments and harnessing the architectural strengths of data-driven operators to capture the global nature of fracture. A key physics-informed approach was introduced by Goswami et al. [109] with the Variational Deep Operator Network (V-DeepONet). The authors propose a hybrid training methodology that combines the governing physical laws with data-driven learning by minimising a hybrid loss function. This function is a weighted sum of a data-driven loss and a physics-driven loss derived from the variational energy of the governing PDEs, which naturally handles the discontinuities inherent in fracture problems.

$$\mathcal{L}(\theta) = \lambda_1 \mathcal{L}_{\text{data}}(\theta) + \lambda_2 \mathcal{L}_{\text{var}}(\theta) \quad (5)$$

where $\mathcal{L}_{\text{var}}$ is the total variational energy of the system. As a result, the model can fit both interpolation and extrapolation tasks. For a Mode-I tensile failure problem, the model achieved an average prediction error 1.85% for an out-of-distribution crack length. In a more complex Mode-II shear failure scenario, the model predicted an out-of-distribution case with a low relative mean error of 3.17% using only 85 training samples, demonstrating its data efficiency.

The proposed framework employs a history-aware input structure by incorporating strain energy from the two preceding time steps to inform future predictions. This temporal context enables the model to capture crack progression dynamics under varying initial defects and loading conditions. The architecture and generalisation capability of V-DeepONet, including its performance on unseen cases, are illustrated in Fig. 20. Overall, the approach demonstrates a robust and generalisable surrogate modelling strategy for complex, history-dependent fracture phenomena.

Zhiheng He [131] explored a purely data-driven approach using the FNO for modelling brittle fracture. The study focuses on predicting fracture paths in plates with randomly distributed cavities, emphasising the architectural strengths of the FNO. The authors argue that FNOs are inherently more suitable for fracture modelling than conventional architectures such as CNN, as fracture propagation is governed by global stress distributions. By operating in the frequency domain, the FNO captures these global dependencies more effectively, whereas the locality of CNN kernels poses a fundamental limitation in this context.

A Fourier layer's core operation can be expressed as:

$$(\mathcal{K}(v)h)(x) = \mathcal{F}^{-1}(R_v \cdot (\mathcal{F}h))(x) \quad (6)$$

where $\mathcal{F}$ is the Fourier transform and R_v is the learnable frequency-domain representation of the convolution kernel.

The trained FNO delivers a massive acceleration, predicting a fracture path for a 512×512 image in 2 s, which is a 750× speed-up over the 25 min required by the FEMs solver. When trained on the same dataset of 2080 input–output pairs, the FNO outperformed a CNN, achieving an accuracy of 72.62%, which was one-third higher than the CNN's 54.76%.

Building on these foundations, recent research has tackled more complex challenges such as spatiotemporal prediction and problems involving irregular or evolving geometries - a critical step towards modelling realistic fracture dynamics. An advancement in spatiotemporal prediction is presented by Kiyani et al. [118], who explore the use of DeepONet variants to predict the time-dependent evolution of crack propagation. A key distinction of this work is its use of training data generated from Constitutively-informed Particle Dynamics (CPD), a discrete particle-based method that captures crack nucleation and evolution without the constraints of continuum assumptions. The study compares a standard basic DeepONet against the more advanced Fusion DeepONet. The Fusion DeepONet architecture consistently outperforms the basic DeepONet with final training loss (MSE) of approximately 10^{-5} , two orders of magnitude lower than the vanilla DeepONet (loss of 10^{-3}). This was achieved with a simpler architecture (3 hidden layers vs. 5) and faster training times.

A primary limitation of FNOs is their reliance on the FFT, which inherently assumes periodicity and rectangular domains. This makes them unsuitable for problems involving irregular geometries or evolving topologies. Liu et al. [128] address this critical gap by introducing the Domain Agnostic Fourier Neural Operator (DAFNO), a novel architecture designed to handle irregular geometries and evolving domains with topology change. The core idea is to explicitly encode geometric information into the FNO architecture

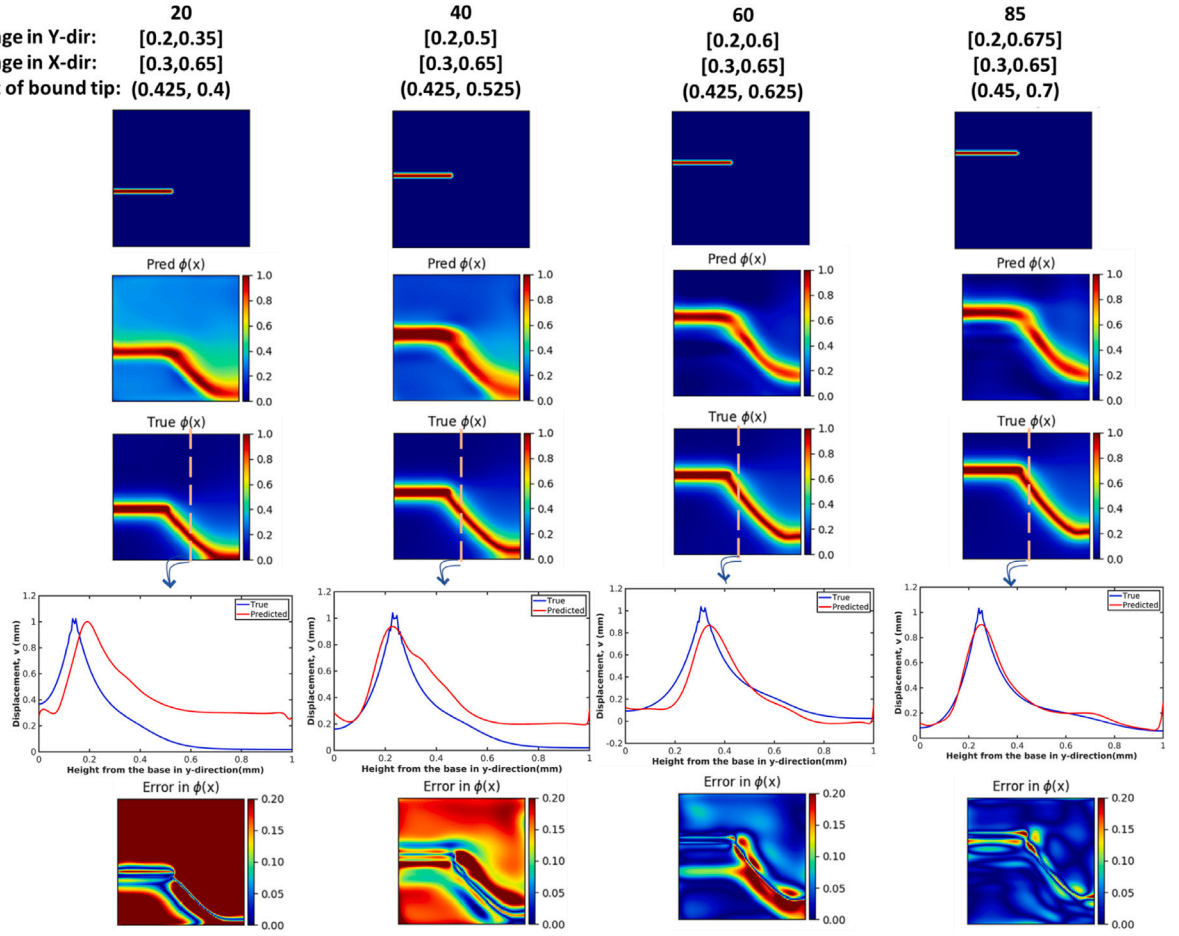


Fig. 20. Out-of-distribution prediction for shear failure using V-DeepONet, from Goswami et al. [109]. The figure shows how the prediction accuracy for an unseen crack location (out of bound tip) improves as the number of training samples (n) increases from 20 to 85. With just 85 samples, the model achieves high accuracy (3.17% error), demonstrating its generalisation capability and highlighting the model's data efficiency.

by incorporating a smoothed characteristic function ($\tilde{\chi}$) into the integral kernel layer. The DAFNO layer update is expressed as:

$$J[H] = \sigma(\tilde{\chi}(x)(I(\tilde{\chi}h) - hI(\tilde{\chi})) + Wh + c) \quad (7)$$

where I denotes a nonlocal convolution operator, $\tilde{\chi}(x)$ masks out non-physical regions and σ is a nonlinear activation function. The term $I(\tilde{\chi}h) - hI(\tilde{\chi})$ encodes domain-aware filtering, ensuring that information propagates only within the physical domain while maintaining consistency near complex boundaries.

For time-dependent problems like fracture, the characteristic function evolves with the domain, $\chi(x, t)$, keeping the operator informed of the changing topology. Trained on a dataset containing snapshots from just one crack simulation, the authors claim that the model was able to accurately predict qualitatively different and unseen fracture patterns, demonstrating that it learned the true underlying constitutive behaviour rather than merely fitting to a specific pattern.

Addressing the challenge of performance degradation of standard FNOs with the increase in network depth, You et al. [127] proposed the Implicit Fourier Neural Operators (IFNO) architecture. Rather than computing the layer output directly, IFNO models the incremental update between consecutive layers. This design interprets the network as a discretised time-dependent nonlocal equation that mimics a fixed-point solver, with the update rule given by:

$$h(x, (l+1)\Delta t) := h(x, l\Delta t) + \Delta t \sigma(Wh(x, l\Delta t) + F^{-1}[F[\kappa] \cdot F[H]](x) + c) \quad (8)$$

Here, F and F^{-1} denote the forward and inverse Fourier transforms, κ is a learnable convolutional kernel in the frequency domain and Δt is a pseudo-time step. This formulation enables stable deep architectures by interpreting the network as a discretised fixed-point solver, suitable for simulating long-horizon, nonlocal dynamics such as fracture evolution.

This structure fixes the memory cost regardless of depth and enables stable “shallow-to-deep” training. In a brittle fracture problem demonstrated in their article, the best IFNO (with $L=8$ layers) achieved a low test error of 12.8% for a 15-step-ahead

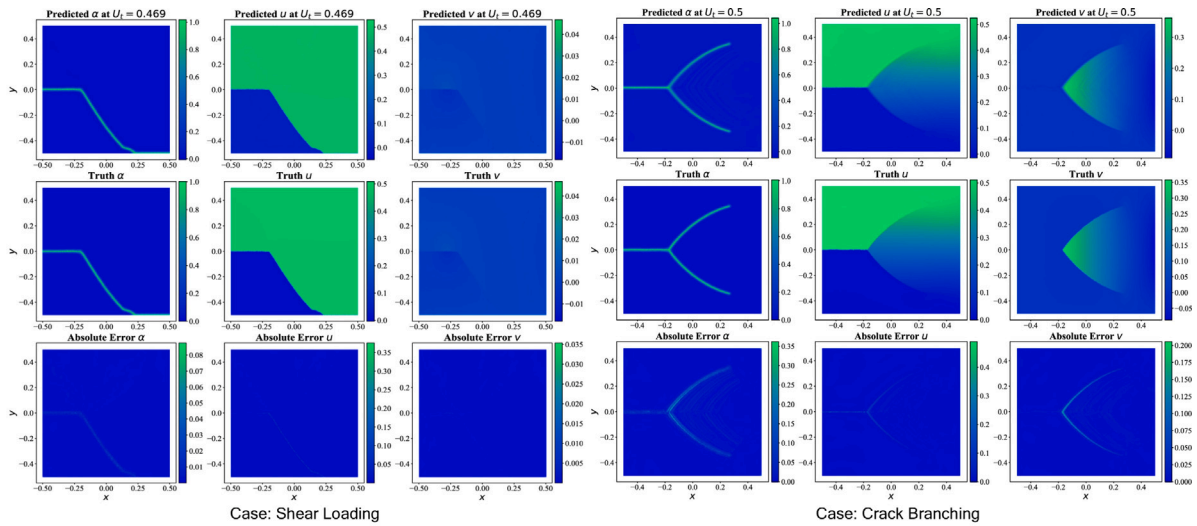


Fig. 21. Predicted vs. ground-truth results and absolute error maps for (left) Shear loading using the two-step DeepONet and (right) Crack branching using DeepOKAN [110].

prediction, whereas the best-performing FNO ($L=4$ layers) had a much higher error of 25.7%. The IFNO's error consistently decreased with depth, while the FNO's performance deteriorated considerably for $L > 4$.

Addressing the high computational cost of phase-field fracture modelling and the error accumulation inherent in sequential operator learning approaches, Kiyani et al. [110] investigate optimising the trunk network architecture within DeepONet with three variants: a data-driven two-step DeepONet, a physics-informed two-step DeepONet incorporating the energy functional into the loss function and DeepOKAN, which replaces the trunk's MLP with a Kolmogorov–Arnold Networks (KAN). A key contribution in their work is the formulation of a direct mapping from initial conditions to the final state at a given load step which avoids the cumulative error accumulation during sequential prediction. The physics-informed variant demonstrates remarkable data efficiency with a MAE of approximately 6.6×10^{-4} for the phase field under shear loading, using only 10 training samples, compared to 6.8×10^{-4} for the data-driven model trained on 4.5 times more data. These architectural variants and their comparative performance are illustrated in Fig. 21.

While the approaches surveyed in Sections 4.1–4.7 described the explored state of the art in architecture setups, the following subsection examines studies that applied these architectures to specific material systems by incorporating crystallographic and mechanism constraints to address multiscale modelling challenges. By tailoring network components, loss terms, or training data to the material at hand, the resulting models often achieve predictive accuracy and physical interpretability that generic approaches cannot match.

4.8. Material-specific models

As seen in the earlier subsections, much of the literature on deep learning for fracture simulation focuses on solver-agnostic architectures. A parallel thread of research applies machine learning to uncover material-specific fracture phenomena. These studies leverage tailored datasets and models to accelerate discovery, enhance physical insights and facilitate inverse design in materials [162,163] ranging from nanostructure graphene to metallic beryllium [85,89]. In this subsection, we review several representative works that integrate ML with material-specific mechanics, each contributing unique capabilities to the broader landscape of computational fracture modelling.

Hsu et al. [89] employed ConvLSTM models to predict fracture evolution in crystalline solids subjected to mixed-mode (Mode I and II) loading. Trained on Molecular Dynamics (MD) simulations across diverse lattice orientations, the model learned spatiotemporal crack propagation patterns and associated energy release rates. It generalised effectively to gradient-oriented bicrystals despite being trained on discrete single-crystal data. However, the model's scalability to polycrystalline or 3D domains remains an open challenge (Fig. 22).

Fernández-Godino et al. [120] addressed the multiscale modelling challenge in high-strain-rate beryllium by training a machine learning emulator to bridge mesoscale fracture simulations (using HOSS) and macroscale constitutive response (in Finite-volume Arbitrary Lagrangian–Eulerian Code (FLAG)). Using only the maximum crack length and peak stress as inputs, their recurrent model reproduced the key dynamics of crack growth with orders-of-magnitude computational speedup. Importantly, the emulator

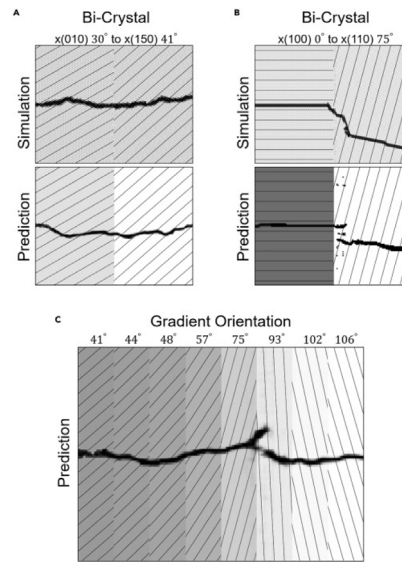


Fig. 22. ConvLSTM model predictions vs. MD simulations for crack paths and energy metrics in crystalline materials [89].

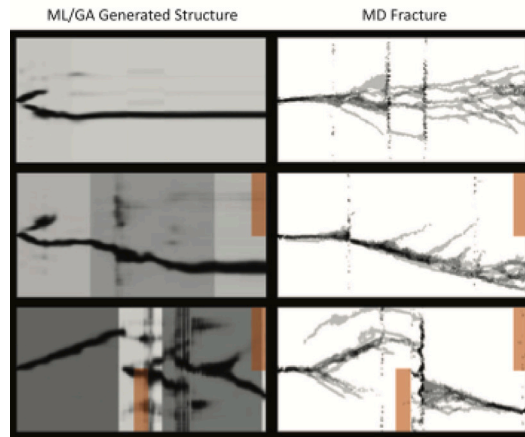


Fig. 23. ML-GA framework applied to polycrystalline MoS_2 . Left: ML-predicted fracture paths avoiding orange-marked weak zones. Right: Corresponding MD validation [85].

generalised to experimental Velocity Interferometer System of Any Reflector (VISAR) data, showcasing robustness and opening avenues for hybrid data-model calibration workflows.

Lew et al. [86] developed a convolutional RNN to simulate fracture evolution in monolayer graphene. Trained solely on MD data for armchair and zigzag orientations, the model accurately captured spatiotemporal crack propagation and stress-strain responses, demonstrating the potential of deep sequence models as efficient surrogates for atomistic simulations. Building on this foundation, Lew and Buehler [85] integrated the trained network with a Genetic Algorithm (GA) to enable inverse design of polycrystalline 2D materials with enhanced fracture resistance. Impressively, by sampling a dense range of grain boundary orientations, the hybrid ML-GA framework revealed narrow peaks in fracture toughness at 19° and 41° orientations associated with maximised crack branching and resistance despite being trained only on single-crystal data. This capability is further demonstrated in Fig. 23, where the framework successfully generates MoS_2 microstructures with tailored crack trajectories that avoid predefined vulnerable regions.

Worthington and Chew [122] proposed a data-driven framework for predicting and designing crack trajectories in ductile materials with complex void morphologies, commonly encountered in additive manufacturing. Their model enables inverse design by optimising void placements to steer crack paths towards desired shapes, while also revealing that visually similar trajectories can emerge from distinct microstructures with noticeably different fracture energies. Furthermore, as illustrated in Fig. 24, the

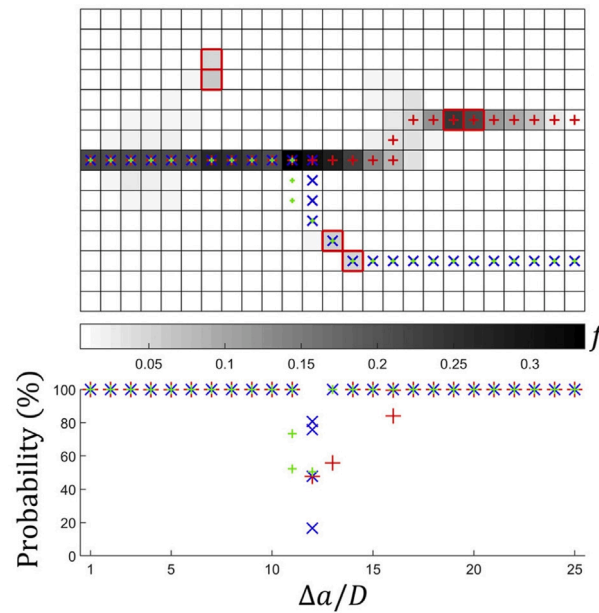


Fig. 24. Probabilistic crack path prediction from Worthington and Chew [122]. The model outputs a distribution of plausible crack trajectories (grey lines) and predicts associated energy dissipation, colour-coded from low (blue) to high (red). The probability distribution highlights the stochastic nature of fracture propagation.

framework captures the inherent stochasticity of fracture by outputting a distribution of likely crack paths and their associated energy dissipation.

The recent work of Dammaß et al. [164] proposes a novel hybrid model tailored for soft, hyper-elastic solids. Their approach integrates the phase-field method for predicting crack propagation with a Physics-Augmented Neural Network (PANN) that acts as a high-fidelity constitutive model for the bulk material. PANN enables a decoupled description of the material's volumetric and isochoric responses, allowing for a weakening of the incompressibility constraint in fracturing zones. The authors ensure the physical soundness of their model by designing the PANN to be polyconvex by construction, which guarantees material stability and enhances extrapolation capabilities. The paper demonstrates that this hybrid framework, one of the first of its kind for soft solids, can be successfully trained on experimental data and used to accurately simulate complex fracture phenomena, showcasing the potential for neural networks to replace traditional constitutive models within established fracture mechanics frameworks.

5. Challenges and limitations in ML-driven fracture modelling

The preceding review draws attention to a consistent trend of increasing adoption of AI architectures, such as CNNs, GNNs, PINNs, neural operators, ROMs, and generative models, for modelling fracture and damage phenomena. Despite their strong performance on curated benchmark datasets, broader adoption remains hindered by several interconnected challenges such as limited training data, poor generalisation to unseen microstructures and loading conditions, insufficient physical consistency, handling of complex geometries and boundary conditions, computational scalability, lack of interpretability (understanding internal representations and learned physics) and lack of explainability (providing human understandable reasons for specific predictions).

Foremost is data scarcity, as producing both experimental and high-fidelity simulation data are expensive and time-consuming [132,136]. Simulation methods such as XFEM, PFM, or CZM require considerable computational resources, while experiments for material-specific fracture characterisation often yield data that is limited in scope and resolution [32,120]. This scarcity hinders supervised training and contributes to poor model robustness in extrapolative scenarios. Moreover, the manual annotation of crack paths, damage fields, or stress concentrations introduces further bottlenecks [125]. The challenge is magnified in 3D, where the cost of generating and storing datasets becomes higher. Existing models do not scale well with mesh size or geometric complexity [71,120] and generative models like GANs or VAEs often produce unrealistic crack geometries [78].

Closely tied to data scarcity is the challenge of generalisation. Many models perform well on familiar datasets but deteriorate markedly when evaluated under new geometries, boundary conditions or materials. CNNs that work on image-based inputs like U-Net variants are fundamentally constrained by their requirement for regular grids, often requiring lossy projection of FEMs data [72] or capturing full 3D fracture morphology. GNNs handle unstructured domains more naturally, but are sensitive to graph construction and resolution [71]. In dynamic simulations involving long time sequences or fine meshes, architectures modelled with

message-passing schemes can accumulate numerical errors and cause over-smoothing.

PINNs, while attractive for embedding governing equations, have not consistently outperformed classical numerical methods. Benchmarking studies show that classical numerical methods are generally faster and more accurate for a range of PDEs at comparable resolutions [165]. The high computational cost of PINNs stems from the large number of optimiser steps needed to solve non-convex training problems. While they can offer fast evaluation in high-dimensional problems where FEMs becomes infeasible [166,167], this benefit only emerges after an expensive offline phase. In low- to mid-dimensional problems where traditional solvers are well-established, they are often less efficient [168]. Furthermore, they struggle to accurately capture singularities near crack tips or sharp displacement jumps due to the smoothness of their function approximators [69,106], a known issue with sharp gradients [100]. The sensitivity of these models to the weighting of physics and data terms further complicates stable convergence [107].

Another limitation is representing geometric complexity. Fracture involves evolving, irregular fields that defy standard grid-based representations like those used in CNNs. While GNNs show promise in handling irregularity, their scalability to large-scale 3D domains remains an open challenge. ROMs, which enhance computational efficiency by simplifying fracture dynamics, may oversimplify critical behaviour and typically fail to capture localised crack propagation unless enriched or coupled with full-order solvers. FNOs are architecturally biased towards fixed domains and periodic structures, making them unsuitable for arbitrary crack patterns without extensive preprocessing or architectural redesign. Neural operators are still an immature frontier with limited benchmarking in fracture applications and face difficulty handling discontinuities. Implementing these methods successfully requires further theoretical and algorithmic innovations to robustly model fracture processes that encompass sudden jumps and evolving discontinuities.

Despite an evolving literature, the industrial adoption of deep learning models for fracture assessment remains minimal, with no documented case studies demonstrating integration into production-scale pipelines or safety certification processes. Most models focus on benchmarks rather than operational needs and lack interpretability, as post-hoc analysis like SHAP or Grad-CAM rarely align with physically meaningful quantities such as crack driving forces, fracture energy release rates or toughness parameters [169,170]. Validation practices further raise concerns, as these models are typically benchmarked against outputs from classical numerical methods rather than physical experiments, risking the propagation and amplification of errors during training. Without experimental validation or interpretable predictions, claims of accuracy remain provisional, undermining trust in critical decision-making where failure mechanisms are poorly understood.

6. Way forward: Perspectives and opportunities

Integration of ML into computational fracture mechanics is rapidly evolving from accelerating simulations to enabling predictive, interpretable and generative capabilities. To fully realise its potential, ongoing research must address key challenges including data scarcity, generalisation, physical inconsistency and lack of interpretability. Based on current trends, several promising directions emerge:

Physics-Integrated and Hybrid Models: Embedding physical laws such as conservation principles, frame invariance and thermodynamic consistency directly into ML architectures can enhance generalisation and reduce reliance on large datasets. Variational, enriched and operator-informed models show strong potential for producing physically consistent and generalisable results. Pre-training large models on diverse simulation datasets and fine-tuning them for new tasks can reduce data requirements for novel applications.

Uncertainty Quantification and Reliability: For ML models to be adopted in safety-critical engineering applications, it is essential that they not only provide accurate predictions but also quantify their associated uncertainty. Integrating methods such as Bayesian Neural Networks, ensemble learning and conformal prediction allows models to output calibrated confidence intervals, identify out-of-distribution inputs and support decision-making under uncertainty. In fracture modelling, ensemble GNNs have been used to predict crack paths and highlight regions of high uncertainty near crack tips [71]. Similarly, Feng et al. [136] employed probabilistic surrogates to predict damage fields under uncertain loading, demonstrating the importance of uncertainty-aware design.

Inverse Design and Generative modelling: Generative models namely GANs, VAEs and diffusion models offer powerful capabilities for exploring and optimising complex fracture phenomena. When coupled with ML surrogates, these models can rapidly generate microstructures or geometries tailored to exhibit specific fracture responses, enabling inverse design for targeted material performance. For example, He et al. [129] employed a conditional GAN to generate crack paths conditioned on porous microstructures, while DeepFracture [77] used neural implicit representations to synthesise 3D fracture surfaces from collision inputs. Beyond design, generative models can also assist in industrial applications by predicting the next step of crack evolution directly from images of fractured surfaces. This image-to-evolution capability enables fast post-failure analysis and damage prognosis, supporting preventive maintenance and real-time monitoring in high-risk engineering environments.

Open Datasets and Benchmarking for Reproducible Research: The existing studies rely on datasets with custom mesh resolutions, geometries, material models and boundary conditions which inhibit reproducibility and comparison across different ML models. While early efforts such as the Mechanical-MNIST Crack Path dataset [76] are a step towards openness, most studies do

not share their data generation processes or training pipelines publicly which limits reproducibility and community benchmarking. To support meaningful progress, we advocate for community driven initiatives to create well-documented, diverse datasets that span material systems, loading modes and crack phenomena. These should include benchmarking protocols and specific evaluation metrics for the predicted crack path, energy release rate or damage field. These efforts naturally align with open science principles [171] and the FAIR framework (Findability, Accessibility, Interoperability and Reusability) [172]. Benchmark datasets have already proven transformative in other domains such as ImageNet in computer vision [173], QM9 in quantum chemistry [174], the AlphaFold Protein Structure Database in structural biology [175] and a similarly curated suite for fracture mechanics could catalyse reproducibility, model development and broader adoption of ML-based tools in this field. Such infrastructure would enable systematic model comparison, foster generalisable architectures, and lay the groundwork for foundation models in mechanics.

7. Conclusions

This review presented a perspective on the growing role of ML in computational fracture and damage mechanics. We charted the progression from traditional physics-based solvers to advanced ML architectures, including surrogate models, GNNs, CNNs, PINNs, neural operators and generative frameworks. There is an evident clear shift from black-box emulators towards hybrid, physics-aware models that balance accuracy, interpretability and efficiency. Techniques such as VE-PINNs, DeepONets, multiscale GNNs and enriched operator networks exemplify this trend, offering faster simulations while preserving physical consistency. Despite this progress, several challenges persist. Data generation remains expensive, models often struggle to generalise beyond training domains and interpretability is still limited. Furthermore, robust deployment in real-world applications requires advances in uncertainty quantification, causal reasoning and scalable learning across geometries and materials. The potential impact of ML in fracture mechanics is profound. These tools can drastically reduce computational costs, enable real-time predictions and open new avenues for inverse design and multiscale modelling. Moving forward, AI is poised not to replace but to augment traditional simulation workflows and enable faster engineering design cycles.

CRedit authorship contribution statement

Allamaprabhu Ani: Writing – original draft, Investigation, Formal analysis, Data curation. **Rajesh Nakka:** Writing – review & editing. **Ghatu Subhash:** Writing – review & editing. **Jean-François Molinari:** Writing – review & editing. **Sathiskumar Anusuya Ponnusami:** Writing – review & editing, Supervision, Project administration, Investigation, Funding acquisition, Conceptualisation.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: Sathiskumar Anusuya Ponnusami reports financial support was provided by Engineering and Physical Sciences Research Council. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to acknowledge the funding support from UKRI- Engineering and Physical Sciences Research Council (EPSRC), United Kingdom through the grant EP/Y004671/1.

Appendix

Notation:

- y_i : True value of the i th sample
- $\hat{y}_i$: Predicted value of the i th sample
- $\bar{y}$: Mean of true values ($\bar{y} = \frac{1}{N} \sum y_i$)
- N : Number of samples

Table 3
Confusion matrix for binary classification.

	Predicted	
Actual	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Table 4
Common evaluation metrics in deep learning.

Metric	Definition	Formula
Accuracy	Overall proportion of correct predictions.	$\frac{TP + TN}{TP + TN + FP + FN}$
Precision	Proportion of predicted positives that are true positives.	$\frac{TP}{TP + FP}$
Recall Sensitivity	Proportion of actual positives correctly predicted.	$\frac{TP}{TP + FN}$
F1 Score	Harmonic mean of precision and recall.	$\frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$
MAE	Mean Absolute Error: average absolute deviation.	$\frac{1}{N} \sum_{i=1}^N y_i - \hat{y}_i $
MSE	Mean Squared Error: average of squared deviations.	$\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$
RMSE	Root Mean Squared Error: interpretable in true units.	$\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$
RE	Relative Error (element-wise).	$RE_i = \frac{ y_i - \hat{y}_i }{ y_i }$
MRE	Mean Relative Error.	$\frac{1}{N} \sum_{i=1}^N \frac{ y_i - \hat{y}_i }{ y_i }$
L^2 Norm Error	Euclidean norm of prediction error.	$\ y - \hat{y}\ _2 = \sqrt{\sum (y_i - \hat{y}_i)^2}$
Relative L^2 Norm	L^2 error normalised by ground truth norm.	$\frac{\ y - \hat{y}\ _2}{\ y\ _2}$
R^2 (Coefficient of Determination)	Proportion of variance explained by the model.	$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$

Data availability

No data was used for the research described in the article.

References

- [1] Liu WK, Li S, Park HS. Eighty years of the finite element method: Birth, evolution, and future. *Arch Comput Methods Eng* 2022;29(6):4431–53.
- [2] Peng GC, Alber M, Buganza Tepole A, Cannon WR, De S, Dura-Bernal S, Garikipati K, Karniadakis G, Lytton WW, Perdikaris P, et al. Multiscale modeling meets machine learning: What can we learn? *Arch Comput Methods Eng* 2021;28:1017–37.
- [3] Bishara D, Xie Y, Liu WK, Li S. A state-of-the-art review on machine learning-based multiscale modeling, simulation, homogenisation and design of materials. *Arch Comput Methods Eng* 2023;30(1):191–222.
- [4] Liu Y, et al. Review of machine learning applications in composite materials. *Compos Part B: Eng* 2021;219:108892. <http://dx.doi.org/10.1016/j.compositesb.2021.108892>.
- [5] Karniadakis GE, et al. Physics-informed machine learning. *Nat Rev Phys* 2021;3(6):422–40. <http://dx.doi.org/10.1038/s42254-021-00049-9>.
- [6] Cuomo S, Di Cola VS, Giampaolo F, Rozza G, Raissi M, Piccialli F. Scientific machine learning through physics-informed neural networks: Where we are and what's next. *J Sci Comput* 2022;92(3):88.
- [7] Pestourie R, Mroueh Y, Rackauckas C, Das P, Johnson SG. Physics-enhanced deep surrogates for partial differential equations. *Nat Mach Intell* 2023;5(12):1458–65.
- [8] Avci O, Abdeljaber O, Kiranyaz S, Hussein M, Gabbouj M, Inman DJ. A review of vibration-based damage detection in civil structures: From traditional methods to machine learning and deep learning applications. *Mech Syst Signal Process* 2021;147:107077.
- [9] Liu X, Athanasiou CE, Padture NP, Sheldon BW, Gao H. A machine learning approach to fracture mechanics problems. *Acta Mater* 2020;190:105–12.
- [10] Jin H, Zhang E, Espinosa HD. Recent advances and applications of machine learning in experimental solid mechanics: A review. *Appl Mech Rev* 2023;75(6):061001.
- [11] Francfort GA, Marigo J-J. Revisiting brittle fracture as an energy minimisation problem. *J Mech Phys Solids* 1998;46(8):1319–42.
- [12] Bourdin B, Francfort GA, Marigo J-J. The variational approach to fracture. *J Elasticity* 2008;91:5–148.

- [13] Miehe C, Gürses E. A robust algorithm for configurational-force-driven brittle crack propagation with r-adaptive mesh alignment. *Internat J Numer Methods Engrg* 2007;72(2):127–55.
- [14] Hofacker M, Miehe C. Continuum phase field modeling of dynamic fracture: variational principles and staggered fe implementation. *Int J Fract* 2012;178:113–29.
- [15] Silling SA, Askari E. A meshfree method based on the peridynamic model of solid mechanics. *Comput Struct* 2005;83(17–18):1526–35.
- [16] Ha YD, Bobaru F. Studies of dynamic crack propagation and crack branching with peridynamics. *Int J Fract* 2010;162(1):229–44.
- [17] Agwai A, Guven I, Madenci E. Predicting crack propagation with peridynamics: a comparative study. *Int J Fract* 2011;171(1):65–78.
- [18] Lemaitre J. A course on damage mechanics. Springer science & business media; 2012.
- [19] Kachanov L. On time to rupture in creep conditions. *Izv an SSSR OTN* 1958;8:26–31, [in russian].
- [20] Rabotnov YN. Creep rupture. In: *Applied mechanics: proceedings of the twelfth international congress of applied mechanics, stanford university, august 26–31, 1968*. Springer; 1969, p. 342–9.
- [21] Lemaitre J. Coupled elasto-plasticity and damage constitutive equations. *Comput Methods Appl Mech Engrg* 1985;51(1–3):31–49.
- [22] Kachanov L. In: *Introduction to continuum damage mechanics*, vol. 10, Springer Science & Business Media; 2013.
- [23] Murakami S. In: *Continuum damage mechanics: a continuum mechanics approach to the analysis of damage and fracture*, vol. 185, Springer Science & Business Media; 2012.
- [24] Bazant ZP, Planas J. *Fracture and size effect in concrete and other quasibrittle materials*. Routledge; 2019.
- [25] Kim E-H, Rim M-S, Lee I, Hwang T-K. Composite damage model based on continuum damage mechanics and low velocity impact analysis of composite plates. *Compos Struct* 2013;95:123–34.
- [26] Peerlings RH, de Borst R, Brekelmans WM, de Vree J. Gradient enhanced damage for quasi-brittle materials. *Internat J Numer Methods Engrg* 1996;39(19):3391–403.
- [27] Forest S. Micromorphic approach for gradient elasticity, viscoplasticity, and damage. *J Eng Mech* 2009;135(3):117–31.
- [28] Gonçalves GF, Lopes IAR, Carneiro AMC, Pires FMA. A critical comparison of gradient and integral nonlocal damage models: Formulation, numerical predictions and computational aspects. *Finite Elements Anal Des* 2025;248:104358.
- [29] Miehe C, Welschinger F, Hofacker M. Thermodynamically consistent phase-field models of fracture: Variational principles and multi-field fe implementations. *Internat J Numer Methods Engrg* 2010;83(10):1273–311.
- [30] Bourdin B, Francfort GA, Marigo J-J. Numerical experiments in revisited brittle fracture. *J Mech Phys Solids* 2000;48(4):797–826.
- [31] Ambati M, Gerasimov T, De Lorenzis L. Phase-field modeling of ductile fracture. *Comput Mech* 2015;55:1017–40.
- [32] Borden MJ, Verhoosel CV, Scott MA, Hughes TJ, Landis CM. A phase-field description of dynamic brittle fracture. *Comput Methods Appl Mech Engrg* 2012;217:77–95.
- [33] Li P, Li W, Li B, Yang S, Shen Y, Wang Q, Zhou K. A review on phase field models for fracture and fatigue. *Eng Fract Mech* 2023;289:109419.
- [34] Sarmadi N, Mousavi Nezhad M, Fisher QJ. On the numerical and mesh-dependent parameters in a computationally enhanced phase-field fracture model coupled with a novel mesh refinement strategy. *Eng Comput* 2023;39(6):3973–92.
- [35] Xu X-P, Needleman A. Numerical simulations of fast crack growth in brittle solids. *J Mech Phys Solids* 1994;42(9):1397–434.
- [36] Barenblatt GI. The mathematical theory of equilibrium cracks in brittle fracture. *Adv Appl Mech* 1962;7:55–129.
- [37] Dugdale DS. Yielding of steel sheets containing slits. *J Mech Phys Solids* 1960;8(2):100–4.
- [38] Xu X-P, Needleman A. Void nucleation by inclusion debonding in a crystal matrix. *Modelling Simul Mater Sci Eng* 1993;1(2):111.
- [39] Xu X-P, Needleman A. Numerical simulations of dynamic crack growth along an interface. *Int J Fract* 1996;74(4):289–324.
- [40] Camacho GT, Ortiz M. Computational modelling of impact damage in brittle materials. *Int J Solids Struct* 1996;33(20–22):2899–938.
- [41] Ortiz M, Pandolfi A. Finite-deformation irreversible cohesive elements for three-dimensional crack-propagation analysis. *Internat J Numer Methods Engrg* 1999;44(9):1267–82.
- [42] Tvergaard V, Hutchinson JW. The relation between crack growth resistance and fracture process parameters in elastic–plastic solids. *J Mech Phys Solids* 1992;40(6):1377–97.
- [43] Moës N, Dolbow J, Belytschko T. A finite element method for crack growth without remeshing. *Internat J Numer Methods Engrg* 1999;46(1):131–50.
- [44] Giraldo-Londoño O, Spring DW, Paulino GH, Buttlar WG. An efficient mixed-mode rate-dependent cohesive fracture model using sigmoidal functions. *Eng Fract Mech* 2018;192:307–27.
- [45] Yang D, He X, Liu X, Deng Y, Huang X. A peridynamics-based cohesive zone model (pd-czm) for predicting cohesive crack propagation. *Int J Mech Sci* 2020;184:105830.
- [46] Linke M, Lammering R. On the calibration of the cohesive strength for cohesive zone models in finite element analyses. *Theor Appl Fract Mech* 2023;124:103733.
- [47] Cundall PA, Strack OD. A discrete numerical model for granular assemblies. *Geotechnique* 1979;29(1):47–65.
- [48] Procházka P. Application of discrete element methods to fracture mechanics of rock bursts. *Eng Fract Mech* 2004;71(4–6):601–18.
- [49] André D, Jebahi M, Iordanoff I, Charles J-I, Néauport J. Using the discrete element method to simulate brittle fracture in the indentation of a silica glass with a blunt indenter. *Comput Methods Appl Mech Engrg* 2013;265:136–47.
- [50] Cil MB, Alshibli KA. 3D assessment of fracture of sand particles using discrete element method. *Géo Lett* 2012;2(3):161–6.
- [51] Weerasekara NS, Powell MS, Cleary P, Tavares LM, Evertsson M, Morrison R, Quist J, Carvalho R. The contribution of dem to the science of comminution. *Powder Technol* 2013;248:3–24.
- [52] Oñate E, Rojek J. Combination of discrete element and finite element methods for dynamic analysis of geomechanics problems. *Comput Methods Appl Mech Engrg* 2004;193(27):3087–128, computational Failure Mechanics for Geomaterials.
- [53] Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: *Proceedings of the thirteenth international conference on artificial intelligence and statistics, JMLR workshop and conference proceedings*. 2010, p. 249–56.
- [54] He K, Zhang X, Ren S, Sun J. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In: *Proceedings of the IEEE international conference on computer vision*. 2015, p. 1026–34.
- [55] Kumar SK. On weight initialization in deep neural networks. 2017, <http://dx.doi.org/10.48550/ARXIV.1704.08863>.
- [56] Bao H, Wu S, Wu Z, Kang G, Peng X, Withers PJ. A machine-learning fatigue life prediction approach of additively manufactured metals. *Eng Fract Mech* 2021;242:107508. <http://dx.doi.org/10.1016/j.engfractmech.2020.107508>.
- [57] Yang J, Kang G, Liu Y, Kan Q. A novel method of multiaxial fatigue life prediction based on deep learning. *Int J Fatigue* 2021;151:106356. <http://dx.doi.org/10.1016/j.ijfatigue.2021.106356>.
- [58] Zhang X-C, Gong J-G, Xuan F-Z. A deep learning based life prediction method for components under creep, fatigue and creep-fatigue conditions. *Int J Fatigue* 2021;148:106236. <http://dx.doi.org/10.1016/j.ijfatigue.2021.106236>.
- [59] Wang Z, Cha Y-J. Unsupervised deep learning approach using a deep auto-encoder with a one-class support vector machine to detect damage. *Struct Health Monit* 2020;20(1):406–25. <http://dx.doi.org/10.1177/1475921720934051>.
- [60] Pivot C, Cordier L, Mathelin L. A continuous reinforcement learning strategy for closed-loop control in fluid dynamics. 2017, <http://dx.doi.org/10.2514/6.2017-3566>.
- [61] Garnier P, Viquerat J, Rabault J, Larcher A, Kuhnle A, Hachem E. A review on deep reinforcement learning for fluid mechanics. *Comput Fluids* 2021;225:104973. <http://dx.doi.org/10.1016/j.compfluid.2021.104973>.

- [62] Estiri E, Mirinejad H. Closed-loop control of fluid resuscitation using reinforcement learning. *IEEE Access* 2023;11:140569–81. <http://dx.doi.org/10.1109/ACCESS.2023.3341036>.
- [63] Samaniego E, Anitescu C, Goswami S, Nguyen-Thanh VM, Guo H, Hamdia K, Zhuang X, Rabczuk T. An energy approach to the solution of partial differential equations in computational mechanics via machine learning: Concepts, implementation and applications. *Mech Eng* 2020;362:112790.
- [64] Brunton SL, Noack BR, Koumoutsakos P. Machine learning for fluid mechanics. *Annu Rev Fluid Mech* 2020;52(1):477–508.
- [65] Hornik K, Stinchcombe M, White H. Multilayer feedforward networks are universal approximators. *Neural Netw* 1989;2(5):359–66.
- [66] Dugas C, Bengio Y, Bélisle F, Nadeau C, Garcia R. Incorporating second-order functional knowledge for better option pricing. *Adv Neural Inf Process Syst* 2000;13.
- [67] Kingma DP, Ba J. Adam: A method for stochastic optimization. 2014, arXiv preprint arXiv:1412.6980.
- [68] Liu DC, Nocedal J. On the limited memory bfgs method for large scale optimization. *Math Program* 1989;45(1):503–28.
- [69] Raissi M, Perdikaris P, Karniadakis GE. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *J Comput Phys* 2019;378:686–707.
- [70] Riedmiller M, Braun H. A direct adaptive method for faster backpropagation learning: The rprop algorithm. In: *IEEE international conference on neural networks*. IEEE; 1993, p. 586–91.
- [71] Perera R, Guzzetti D, Agrawal V. Graph neural networks for simulating crack coalescence and propagation in brittle materials. *Comput Methods Appl Mech Engrg* 2022;395:115021.
- [72] Koopas RN, Rezaei S, Rauter N, Ostwald R, Lammering R. A spatiotemporal deep learning framework for prediction of crack dynamics in heterogeneous solids: Efficient mapping of concrete microstructures to its fracture properties. *Eng Fract Mech* 2025;314:110675.
- [73] Lu L, Jin P, Karniadakis GE. Learning nonlinear operators via deepnet based on the universal approximation theorem of operators. *Nat Mach Intell* 2021;3:218–29. <http://dx.doi.org/10.1038/s42256-021-00302-5>.
- [74] Nguyen H-Q, Le B-A, Tran B-V, Vu T-S, Bui T-L. Deep artificial neural network-powered phase field model for predicting damage characteristic in brittle composite under varying configurations. *Mach Learn: Sci Technol* 2024;5(2):025062.
- [75] Sepasdar R, Karpatne A, Shakiba M. A data-driven approach to full-field nonlinear stress distribution and failure pattern prediction in composites using deep learning. *Comput Methods Appl Mech Engrg* 2022;397:115126.
- [76] Mohammadzadeh S, Lejeune E. Predicting mechanically driven full-field quantities of interest with deep learning-based metamodels. *Extrem Mech Lett* 2022;50:101566.
- [77] Huang Z, Chang D, Yang X, Zuo H. A deep learning-based approach for crack damage detection using strain field. *Eng Fract Mech* 2023;293:109703.
- [78] Huang Y, Kanai T. Deepfracture: A generative approach for predicting brittle fractures with neural discrete representation learning. In: *Computer graphics forum*. Wiley Online Library; 2025, e70002.
- [79] Rezafat M, Hogan JD. A finite element-convolutional neural network model (fe-cnn) for stress field analysis around arbitrary inclusions. *Mach Learn: Sci Technol* 2023;4(4):045052.
- [80] Wang Y, Oyen D, Guo W, Mehta A, Scott CB, Panda N, Fernández-Godino MG, Srinivasan G, Yue X. Stressnet-deep learning to predict stress with fracture propagation in brittle materials. *Npj Mater Degrad* 2021;5(1):6.
- [81] Moghtaderi SH, Jedi A, Ariffin AK, Thamburaja P. Application of machine learning in fracture analysis of edge crack semi-infinite elastic plate. *Fract Struct Integr* 2024;18(68):197–208.
- [82] Yonekura K, Maruoka K, Tyou K, Suzuki K. Super-resolving 2d stress tensor field conserving equilibrium constraints using physics-informed u-net. *Finite Elem Anal Des* 2023;213:103852.
- [83] Xu H, Fan W, Ruan L, Shi R, Taylor AC, Zhang D. Crack-net: A deep learning approach to predict crack propagation and stress-strain curves in particulate composites. *Engineering* 2025.
- [84] Park D, Lee J, Lee H, Gu GX, Ryu S. Deep generative spatiotemporal learning for integrating fracture mechanics in composite materials: inverse design. *Discov Optim Mater Horizons* 2024;11(13):3048–65.
- [85] Lew AJ, Buehler MJ. A deep learning augmented genetic algorithm approach to polycrystalline 2d material fracture discovery and design. *Appl Phys Rev* 2021;8(4).
- [86] Lew AJ, Yu C-H, Hsu Y-C, Buehler MJ. Deep learning model to predict fracture mechanisms of graphene. *Npj 2D Mater Appl* 2021;5(1):48.
- [87] Buehler MJ. Predicting mechanical fields near cracks using a progressive transformer diffusion model and exploration of generalization capacity. *J Mater Res* 2023;38(5):1317–31.
- [88] Hsu Y-C, Buehler MJ. Dyfrant: Forecasting and backcasting dynamic fracture mechanics in space and time using a 2d-to-3d deep neural network. *APL Mach Learn* 2023;1(2).
- [89] Hsu Y-C, Yu C-H, Buehler MJ. Using deep learning to predict fracture patterns in crystalline solids. *Matter* 2020;3(1):197–211.
- [90] Schwarzer M, Rogan B, Ruan Y, Song Z, Lee DY, Percus AG, Chau VT, Moore BA, Rougier E, Viswanathan HS, et al. Learning to fail: Predicting fracture evolution in brittle material models using recurrent graph convolutional neural networks. *Comput Mater Sci* 2019;162:322–32.
- [91] Perera R, Agrawal V. A generalized machine learning framework for brittle crack problems using transfer learning and graph neural networks. *Mech Mater* 2023;181:104639.
- [92] Perera R, Agrawal V. Dynamic and adaptive mesh-based graph neural network framework for simulating displacement and crack fields in phase field models. *Mech Mater* 2023;186:104789.
- [93] Perera R, Agrawal V. Multiscale graph neural networks with adaptive mesh refinement for accelerating mesh-based simulations. *Comput Methods Appl Mech Engrg* 2024;429:117152.
- [94] Pareja A, Domeniconi G, Chen J, Ma T, Suzumura T, Kanezashi H, Kaler T, Schardl T, Leiserson C. Evolvegn: Evolving graph convolutional networks for dynamic graphs. In: *Proceedings of the AAAI conference on artificial intelligence*, vol. 34, 2020, p. 5363–70.
- [95] Goswami S, Anitescu C, Chakraborty S, Rabczuk T. Transfer learning enhanced physics informed neural network for phase-field modeling of fracture. *Theor Appl Fract Mech* 2020;106:102447.
- [96] Gu Y, Zhang C, Zhang P, Golub MV, Yu B. Enriched physics-informed neural networks for 2d in-plane crack analysis: Theory and matlab code. *Int J Solids Struct* 2023;276:112321.
- [97] Ghaffari Motlagh Y, Jimack PK, de Borst R. Deep learning phase-field model for brittle fractures. *Internat J Numer Methods Engrg* 2023;124(3):620–38.
- [98] Manav M, Molinaro R, Mishra S, De Lorenzis L. Phase-field modeling of fracture with physics-informed deep learning. *Comput Methods Appl Mech Engrg* 2024;429:117104.
- [99] Lian H, Zhao P, Zhang M, Wang P, Li Y. Physics informed neural networks for phase field fracture modeling enhanced by length-scale decoupling degradation functions. *Front Phys* 2023;11:1152811.
- [100] Kharazmi E, Zhang Z, Karniadakis GE. hp-vpinns: Variational physics-informed neural networks with domain decomposition. *Comput Methods Appl Mech Engrg* 2021;374:113547.
- [101] Pantis P, Mobasher ME. Integrated finite element neural network (i-fenn) for non-local continuum damage mechanics. *Comput Methods Appl Mech Engrg* 2023;404:115766.
- [102] Mishra S, Molinaro R. Estimates on the generalization error of physics-informed neural networks for approximating pdes. *IMA J Numer Anal* 2023;43(1):1–43.

- [103] Pantidis P, Eldababy H, Tagle CM, Mobasher ME. Error convergence and engineering-guided hyperparameter search of pinns: towards optimized i-fenn performance. *Comput Methods Appl Mech Engrg* 2023;414:116160.
- [104] Zhao L, Shao Q. Denms: Discontinuity-embedded neural networks for fracture mechanics. *Comput Methods Appl Mech Engrg* 2025;446:118184.
- [105] Baek J, Chen J-S. A neural network-based enrichment of reproducing kernel approximation for modeling brittle fracture. *Comput Methods Appl Mech Engrg* 2024;419:116590.
- [106] Jagtap AD, Karniadakis GE. Extended physics-informed neural networks (xpinns): A generalized space–time domain decomposition based deep learning framework for nonlinear partial differential equations. *Commun Comput Phys* 2020;28(5).
- [107] Chakraborty A, Anitescu C, Goswami S, Zhuang X, Rabczuk T. Variational energy based xpinns for phase field analysis in brittle fracture. 2022, arXiv preprint arXiv:2207.02307.
- [108] Pantidis P, Eldababy H, Abueidda D, Mobasher ME. I-fenn with temporal convolutional networks: Expediting the load-history analysis of non-local gradient damage propagation. *Comput Methods Appl Mech Engrg* 2024;425:116940.
- [109] Goswami S, Yin M, Yu Y, Karniadakis GE. A physics-informed variational deepoNet for predicting crack path in quasi-brittle materials. *Comput Methods Appl Mech Engrg* 2022;391:114587.
- [110] Kiyani E, Manav M, Kadivar N, De Lorenzis L, Karniadakis GE. Predicting crack nucleation and propagation in brittle materials using deep operator networks with diverse trunk architectures. *Comput Methods Appl Mech Engrg* 2025;441:117984.
- [111] Rashid MM, Pittie T, Chakraborty S, Krishnan NA. Learning the stress–strain fields in digital composites using fourier neural operator. *Iscience* 2022;25(11).
- [112] Kontolati K, Goswami S, Em Karniadakis G, Shields MD. Learning nonlinear operators in latent spaces for real-time predictions of complex dynamics in physical systems. *Nat Commun* 2024;15(1):5101.
- [113] Oommen V, Shukla K, Goswami S, Dingreville R, Karniadakis GE. Learning two-phase microstructure evolution using neural operators and autoencoder architectures. *Npj Comput Mater* 2022;8(1):190.
- [114] Bessa MA, Bostanabad R, Liu Z, Hu A, Apley DW, Brinson C, Chen W, Liu WK. A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality. *Comput Methods Appl Mech Engrg* 2017;320:633–67.
- [115] Moore BA, Rougier E, O'Malley D, Srinivasan G, Hunter A, Viswanathan H. Predictive modeling of dynamic fracture growth in brittle materials with machine learning. *Comput Mater Sci* 2018;148:46–53.
- [116] Chen S, Feng S, Huang Y, Lei Z, Jia X, Lin Y, Rougier E. Hossnet: An efficient physics-guided neural network for simulating micro-crack propagation. *Comput Mater Sci* 2024;236:112846.
- [117] Huang Z, Liu Q, Liu R, Chang D, Yang X, Zuo H, Dong Y. A hybrid neural network–based approach to predict crack propagation paths. *Fatigue Fract Eng Mater Struct* 2025;48(3):1098–111.
- [118] Kiyani E, Ananchaperumal V, Peyvan A, Uchimali M, Li G, Karniadakis GE. Crack path prediction with operator learning using discrete particle system data generation. 2025, arXiv preprint arXiv:2506.01976.
- [119] Hunter A, Moore BA, Mudunuru M, Chau V, Tchoua R, Nyshadham C, Karra S, O'Malley D, Rougier E, Viswanathan H, et al. Reduced-order modeling through machine learning and graph-theoretic approaches for brittle fracture applications. *Comput Mater Sci* 2019;157:87–98.
- [120] Fernández-Godino MG, Panda N, O'Malley D, Larkin K, Hunter A, Hafika RT, Srinivasan G. Accelerating high-strain continuum-scale brittle fracture simulations with machine learning. *Comput Mater Sci* 2021;186:109959.
- [121] Huang Y, Yu Y, Kanai T. Predicting brittle fracture surface shape from a versatile database. *Comput Animat Virtual Worlds* 2019;30(6):e1865.
- [122] Worthington M, Chew HB. Crack path predictions in heterogeneous media by machine learning. *J Mech Phys Solids* 2023;172:105188.
- [123] Shinde K, Itier V, Mennesson J, Vasiukov D, Shakoov M. Dimensionality reduction through convolutional autoencoders for fracture patterns prediction. *Appl Math Model* 2023;114:94–113.
- [124] Greve L, Schneider B, Eller T, Andres M, Martinez J-D, van de Weg B. Necking-induced fracture prediction using an artificial neural network trained on virtual test data. *Eng Fract Mech* 2019;219:106642.
- [125] Nguyen-Le DH, Tao Q, Nguyen V-H, Abdel-Wahab M, Nguyen-Xuan H. A data-driven approach based on long short-term memory and hidden markov model for crack propagation prediction. *Eng Fract Mech* 2020;235:107085.
- [126] Weg BVD, Greve L, Andres M, Eller T, Rosic B. Neural network-based surrogate model for a bifurcating structural fracture response. *Eng Fract Mech* 2021. <http://dx.doi.org/10.1016/j.engfracmech.2020.107424>.
- [127] You H, Zhang Q, Ross CJ, Lee C-H, Yu Y. Learning deep implicit fourier neural operators (ifnos) with applications to heterogeneous material modeling. *Comput Methods Appl Mech Engrg* 2022;398:115296.
- [128] Liu N, Jafarzadeh S, Yu Y. Domain agnostic fourier neural operators. *Adv Neural Inf Process Syst* 2023;36:47438–50.
- [129] He Y, Tan Y, Yang M, Wang Y, Xu Y, Yuan J, Li X, Chen W, Kang G. Accurate prediction of discontinuous crack paths in random porous media via a generative deep learning model. *Proc Natl Acad Sci* 2024;121(40):e2413462121.
- [130] Feng S, Xu Y, Han X, Li Z, Inceci A. A phase field and deep-learning based approach for accurate prediction of structural residual useful life. *Comput Methods Appl Mech Engrg* 2021;383:113885.
- [131] He Z. Application of neural operators to the phase field modeling of brittle fractures. In: 2024 conference on AI, science, engineering, and technology. IEEE; 2024, p. 311–5.
- [132] de Moraes EAB, Salehi H, Zayernouri M. Data-driven failure prediction in brittle materials: A phase field-based machine learning framework. *J Mach Learn Model Comput* 2021;2(1).
- [133] Koeppel A, Bamer F, Hernandez Padilla CA, Markert B. Neural network representation of a phase-field model for brittle fracture. *PAMM* 2017;17(1):253–4.
- [134] Li X, Zhang X, Feng W, Wang Q. Machine learning-based prediction of fracture toughness and path in the presence of micro-defects. *Eng Fract Mech* 2022;276:108900.
- [135] Duong PLT, Hussain S, Jhon MH, Raghavan N. Data driven prognosis of fracture dynamics using tensor train and gaussian process regression. *IEEE Access* 2020;8:222256–66.
- [136] Feng Y, Wang Q, Wu D, Luo Z, Chen X, Zhang T, Gao W. Machine learning aided phase field method for fracture mechanics. *Internat J Engrg Sci* 2021;169:103587.
- [137] Chen Y, Dodwell T, Chuaqui T, Butler R. Full-field prediction of stress and fracture patterns in composites using deep learning and self-attention. *Eng Fract Mech* 2023;286:109314.
- [138] Sutula D, Kerfriden P, van Dam T, Bordas SP. Minimum energy multiple crack propagation (parts i, ii & iii). *Eng Fract Mech* 2018;191:205–76.
- [139] Rougier E. Hoss: Hybrid optimization software suite. Tech. rep., Los Alamos, NM (United States): Los Alamos National Laboratory (LANL); 2020.
- [140] Zhang X, Zhao T, Liu Y, Chen Q, Wang Z, Wang Z. A data-driven model for predicting the mixed-mode stress intensity factors of a crack in composites. *Eng Fract Mech* 2023;288:109385.
- [141] Krishnamurthy V, Levoy M. Fitting smooth surfaces to dense polygon meshes. In: Proceedings of the 23rd annual conference on computer graphics and interactive techniques. stanford Computer Graphics Laboratory; 1996, p. 313–24.
- [142] Bahrami B, Talebi H, Ayatollahi MR, Khosravani MR. Artificial neural network in prediction of mixed-mode i/ii fracture load. *Int J Mech Sci* 2023;248:108214.
- [143] Talebi H, Bahrami B, Daneshfar M, Bagherifard S, Ayatollahi MR. Data-driven based fracture prediction of notched components. *Phil Trans R Soc A* 2024;382(2264):20220397.

- [144] Niu L, Li N, Guan J, Xie C, Li L, Chen S. Machine learning-based prediction approaches for the fracture toughness of quasi-brittle materials. *Eng Fract Mech* 2025;111449.
- [145] Deng S, Apelian D, Bostanabad R. Adaptive spatiotemporal dimension reduction in concurrent multiscale damage analysis. *Comput Mech* 2023;72(1):3–35.
- [146] Deng S, Mora C, Apelian D, Bostanabad R. Data-driven calibration of multifidelity multiscale fracture models via latent map gaussian process. *J Mech Des* 2023;145(1):011705.
- [147] Huynh P-T, Cao Y, Hoang T-T-P. Fast and accuracy-preserving domain decomposition methods for reduced fracture models with nonconforming time grids. *J Sci Comput* 2023;96(1):23.
- [148] Matray V, Amlani F, Feyel F, Néron D. A hybrid numerical methodology coupling reduced order modeling and graph neural networks for non-parametric geometries: applications to structural dynamics problems. *Comput Methods Appl Mech Engrg* 2024;430:117243.
- [149] Toscano JD, Oommen V, Varghese AJ, Zou Z, Ahmadi Daryakenari N, Wu C, Karniadakis GE. From pinns to pikans: Recent advances in physics-informed machine learning. *Mach Learn Comput Sci Eng* 2025;1(1):1–43.
- [150] Zhao CL. Solving Allen-Cahn and Cahn-Hilliard equations using the adaptive physics informed neural networks. *Commun Comput Phys* 2020;29(3).
- [151] Franck E, Michel-Dansac V, Navoret L. Approximately well-balanced discontinuous galerkin methods using bases enriched with physics-informed neural networks. *J Comput Phys* 2024;512:113144.
- [152] Tu J, Liu C, Qi P. Physics-informed neural network integrating pointnet-based adaptive refinement for investigating crack propagation in industrial applications. *IEEE Trans Ind Inform* 2022;19(2):2210–8.
- [153] Csati Z, Moës N, Massart TJ. A stable extended/generalised finite element method with lagrange multipliers and explicit damage update for distributed cracking in cohesive materials. *Comput Methods Appl Mech Engrg* 2020;369:113173.
- [154] Berrone S, Pieraccini S, Scialo S. A pde-constrained optimization formulation for discrete fracture network flows. *SIAM J Sci Comput* 2013;35(2):B487–510.
- [155] Ning L, Cai Z, Dong H, Liu Y, Wang W. Physics-informed neural network frameworks for crack simulation based on minimized peridynamic potential energy. *Comput Methods Appl Mech Engrg* 2023;417:116430.
- [156] Eghbalpoor R, Sheidaei A. A peridynamic-informed deep learning model for brittle damage prediction. *Theor Appl Fract Mech* 2024;131:104457.
- [157] Ventura G, Xu J, Belytschko T. A vector level set method and new discontinuity approximations for crack growth by efg. *Internat J Numer Methods Engrg* 2002;54(6):923–44.
- [158] Hu H, Qi L, Chao X. Physics-informed neural networks (pinn) for computational solid mechanics: Numerical frameworks and applications. *Thin-Walled Struct* 2024;112495.
- [159] Pantidis P, Svolos L, Abueidda D, Mobasher ME. Integrated finite element neural network (ifenn) for phase-field fracture with minimal input and generalised geometry-load handling. *Comput Methods Appl Mech Engrg* 2026;448:118485.
- [160] Konale A, Srivastava V. A physics-informed neural network for modeling fracture without gradient damage: Formulation, application, and assessment. *J Mech Phys Solids* 2025;106395.
- [161] Yu B, et al. The deep ritz method: a deep learning-based numerical algorithm for solving variational problems. *Commun Math Stat* 2018;6(1):1–12.
- [162] Badini S, Regondi S, Pugliese R. Unleashing the power of artificial intelligence in materials design. *Materials* 2023;16(17):5927.
- [163] Chen C-T, Gu GX. Generative deep neural networks for inverse materials design using backpropagation and active learning. *Adv Sci* 2020;7(5):1902607.
- [164] Dammaß F, Kalina KA, Kästner M. Neural networks meet phase-field: A hybrid fracture model. *Comput Methods Appl Mech Engrg* 2025;440:117937.
- [165] Krishnapriyan A, Gholami A, Zhe S, Kirby R, Mahoney MW. Characterizing possible failure modes in physics-informed neural networks. *Adv Neural Inf Process Syst* 2021;34:26548–60.
- [166] Kunisch K, Walter D. Semiglobal optimal feedback stabilization of autonomous systems via deep neural network approximation. *ESAIM Control Optim Calc Var* 2021;27:16.
- [167] Hu Z, Shukla K, Karniadakis GE, Kawaguchi K. Tackling the curse of dimensionality with physics-informed neural networks. *Neural Netw* 2024;176:106369.
- [168] Grossmann TG, Komorowska UJ, Latz J, Schönlieb C-B. Can physics-informed neural networks beat the finite element method? *IMA J Appl Math* 2024;89(1):143–74.
- [169] Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-cam: visual explanations from deep networks via gradient-based localization. *Int J Comput Vis* 2020;128:336–59.
- [170] Zhou S, Yang B, Xiao S, Yang G, Zhu T. Interpretable machine learning method for modelling fatigue short crack growth behaviour. *Met Mater Int* 2024;30(7):1944–64.
- [171] Hrynaskiewicz I, Li P, Edmunds S. Open science and the role of publishers in reproducible research. In: *Implementing reproducible research*. Chapman and Hall/CRC; 2018, p. 383–418.
- [172] Wilkinson MD, Dumontier M, Aalbersberg IJ, Appleton G, Axton M, Baak A, Blomberg N, Boiten J-W, da Silva Santos LB, Bourne PE, et al. The fair guiding principles for scientific data management and stewardship. *Sci Data* 2016;3(1):1–9.
- [173] Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. Imagenet: A large-scale hierarchical image database. In: *2009 IEEE conference on computer vision and pattern recognition*. Ieee; 2009, p. 248–55.
- [174] Ramakrishnan R, Dral PO, Rupp M, Von Lilienfeld OA. Quantum chemistry structures and properties of 134 kilo molecules. *Sci Data* 2014;1(1):1–7.
- [175] Varadi M, Bertoni D, Magana P, Paramval U, Pidruchna I, Radhakrishnan M, Tsenkov M, Nair S, Mirdita M, Yeo J, et al. Alphafold protein structure database in 2024: providing structure coverage for over 214 million protein sequences. *Nucleic Acids Res* 2024;52(D1):D368–75.