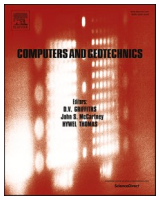
Computers and Geotechnics 185 (2025) 107310

Contents lists available at ScienceDirect

Computers and Geotechnics

journal homepage: www.elsevier.com/locate/compgeo



Review

Artificial intelligence for computational granular media

Tongming Qu [a,*], Jidong Zhao [a,*], Y.T. Feng [b]

[a] Department of Civil and Environmental Engineering, Hong Kong University of Science and Technology, Clearwater Bay, Kowloon, Hong Kong Special Administrative Region

[b] Zienkiewicz Centre for Computational Engineering, Faculty of Science and Engineering, Swansea University, Swansea, Wales SA1 8EP, UK

ARTICLE INFO

Keywords:
Granular materials
Machine learning
Discrete element method
Computational mechanics
Knowledge discovery
Inverse problems

ABSTRACT

Artificial intelligence (AI) has played a transformative role in accelerating scientific discovery and driving engineering innovations. Here we examine the primary applications of AI in computational granular materials over the past decades, focusing on three key objectives: (i) what machine learning (ML) can do in computational granular mechanics, (ii) how ML is integrated into routine computational simulations of granular media, and (iii) the opportunities and challenges that ML presents in this domain. The review highlights the key objectives of computational granular mechanics and the role of ML in bridging these critical research gaps. It systematically covers three aspects: (i) ML-accelerated computational modelling, (ii) ML-enabled pattern recognition and knowledge discovery, and (iii) ML-assisted inverse analysis in granular mechanics. Pertinent challenges are thoroughly discussed from the perspective of data and models. To promote the development of data-driven computational granular mechanics, we launched "Clear Data Bay", a metadata website tailored for domain data sharing and management. Despite ongoing challenges, data-driven approaches offer great potential in enabling computational granular mechanics models to tackle previously unattainable challenges.

1. Introduction

Granular materials, such as sand and food grains, are ubiquitous in engineering, industry, and nature. Although individual grains may seem simple, an assembly of granular particles can exhibit rich and intricate mechanical phenomena. Various computational granular mechanics (CGM) models have been developed as powerful tools to describe, explain, and predict the behaviour of granular materials across diverse environments and applications. The discrete element method (DEM) is a prominent technique for modelling granular materials which represents particles as discrete elements that interact through contact forces (Cundall and Strack, 1979). In contrast, the finite element method (FEM)

Abbreviations: AD, Automatic Differentiation; AE, Autoencoder; AI, Artificial Intelligence; AoR, Angle of Repose; BGNN, Boundary Graph Neural Network; BVPs, Boundary Value Problems; CFD, Computational Fluid Dynamics; cGANs, Conditional Generative Adversarial Networks; CGM, Computational Granular Mechanics; CNN, Convolutional Neural Network; DCGAN, Deep Convolutional GANs; DDCM, Data-Driven Constitutive Modelling; DDMM, Data-Driven Multiscale Modelling; DDP, Differential Dynamic Programming; DDPM, Denoising Diffusion Probabilistic Models; DE, Differential Evolution; DEM, Discrete Element Method; DGECNN, Dynamic Graph Edge Convolution Neural Network; FDEM, Combined Finite-Discrete Element Method; FEM, Finite Element Method; FI-GAN, Fractal Informed-GAN; GA, Genetic Algorithm; GAN, Generative Adversarial Model; GBDT, Gradient Boosting Decision Trees; GCNNs, Graph Convolutional Neural Networks; GNN, Graph Neural Network; GNS, Graph Network-Based Simulators; GRU, Gated Recurrent Unit; GSD, Grain Size Distribution; HMM, Hierarchical Multiscale Modelling; LDM, Latent Diffusion Model; LightGBM, Light Gradient Boosting Machine; LLM, Large Language Model; LSTM, Long Short-Term Memory; LVM, Large Vision Model; MFRNN, Multi-Fidelity Residual Neural Network; ML, Machine Learning; MLP, Multilayer Perceptron; MPM, Material Point Method; NNs, Neural Networks; OCR, Over-Consolidation Ratio; ODEs, Ordinary Differential Equations; OOD, Out-of-Distribution; PBEs, Population Balance Equations; PCA, Principal Component Analysis; PD, Peridynamics; PDEs, Partial Differential Equations; PFEM, Particle Finite Element Method; PINN, Physics-Informed Neural Network; POD, Proper Orthogonal Decomposition; R-CNN, Region Convolutional Neural Network; RF, Random Forests; RL, Reinforcement Learning; RNN, Recurrent Neural Network; RNNSR, Recurrent Neural Network with Stochastically Calculated Random Motion; ROMs, Reduced Order Models; RVEs, Representative Volume Elements; SA, Simulated Annealing; SAM, Segment Anything Model; SGN, Signed Distance Function-Based Graph Neural Network; SH, Spherical Harmonics; SHAP, SHapley Additive exPlanations; SPH, Smoothed Particle Hydrodynamics; SVM, Support Vector Machines; TCN, Temporal Convolutional Network; TGNNS, Temporal Graph Neural Network-Based Simulator; UCS, Uniaxial Compressive Strength; UDEC, Universal Distinct Element Code; VAE, Variational Autoencoder-Decoder; VFMs, Vision Foundation Models; ViT, Vision Transformer; XGBoost, Extreme Gradient Boosting; μCT, X-Ray Micro-Computed Tomography.

* Corresponding authors.
E-mail addresses: tongmingqu@ust.hk (T. Qu), jzhao@ust.hk (J. Zhao).

https://doi.org/10.1016/j.compgeo.2025.107310
Received 28 February 2025; Received in revised form 10 April 2025; Accepted 24 April 2025
Available online 7 May 2025

treats large-scale granular materials as continuous media and discretizes them into mesh elements for analysis (Zienkiewicz et al., 1977). The material point method (MPM) (Sulsky et al., 1994) and the particle finite element method (PFEM) (Oñate et al., 2004) are particularly effective for modelling large deformations in granular materials. Smoothed particle hydrodynamics (SPH) (Liu and Liu, 2010) excels at simulating the flow behaviour of granular materials, while Peridynamics (PD) (Silling, 2000) and the combined finite-discrete element method (FDEM) (Munjiza et al., 1995) are suitable for modelling cemented granular materials involving particle crushing or discontinuities (Chen et al., 2020).

Over the past decades, efforts in CGM can be broadly categorized into three main aspects: (1) **Improving Accuracy**: This involves developing reliable models and algorithms to predict the behaviour of granular materials under various mechanical, hydraulic, and thermal conditions (Vargas and McCarthy, 2007). Key advancements include modelling realistic grain shapes (Feng, 2023), particle breakage (de Bono and McDowell, 2016), large deformation (Chen and Qiu, 2012), fluid-particle interaction (Feng et al., 2007, Zhao and Shan, 2013), and phase transitions (Vidyapati and Subramaniam, 2012). (2) **Boosting Efficiency**: The goal is to design computational algorithms that accelerate numerical simulations. Techniques include coarse-graining (Nasato et al., 2015), parallel computations using CPU or GPU (Guo and Zhao, 2016b, Zhao et al., 2021), and reduced-order modelling (Zhong and Sun, 2018, Li et al., 2023b). (3) **Application of CGM**: Numerical simulations are utilized to examine phenomena related to granular materials, industrial processes and engineering challenges, intending to gain new insights and optimized solutions for practical problems (Zeng et al., 2025).

Despite these significant advancements in CGM, several crucial challenges remain unresolved in the field. Specifically, direct simulations of particle-scale interactions for large-scale problems are computationally expensive, while homogenization-based or phenomenological approaches often fail to accurately reproduce the particle-scale physics and underlying mechanisms of granular materials. Furthermore, measuring and calibrating particle-scale properties can be challenging (Qu et al., 2022b). Addressing these long-standing issues necessitates the exploration of new research paradigms.

Data-driven approaches represent a new research paradigm in the era of Artificial Intelligence (AI) and big data. The past decade has witnessed tremendous progress in applying AI models across various scientific and engineering domains, notably in areas such as image recognition, natural language processing, and material discovery. Recently, the potential of AI in CGM has garnered increasing attention, with innovative machine learning (ML)-based strategies introduced to enhance traditional CGM methods (Wang et al., 2024b). However, a comprehensive comprehension of how ML can transform CGM is still lacking, and the specific scenarios where ML can provide benefits remain unclear. This review aims to address these questions by systematically summarizing the applications of ML in CGM, identifying limitations, and exploring future trends in ML-enhanced CGM.

2. Overview of ML applications in CGM

2.1. Brief revisit to machine learning

Machine learning (ML) is a subfield of AI focusing on developing algorithms and statistical models that enable computers to learn from data without explicit programming. As shown in Fig. 1, ML can be broadly classified into three categories: (1) Supervised Learning: This approach is trained on a labelled dataset where each input is paired with corresponding outputs. Classification and regression are two types of supervised learning algorithms, where the goal of classification is to predict the categorical or discrete class labels for new, unseen instances based on input features, while regression seeks to learn underlying patterns or relationships to predict continuous outputs. These data pairs include one-to-one state mapping data, time series data, spatial structure data, and even non-structured data. Most ML models, such as multilayer perceptron (MLP), random forests (RF), and support vector machines (SVM), can adapt both classification and regression tasks. (2) Unsupervised Learning: It aims to uncover hidden patterns, structures, or relationships within unlabeled data. Examples include clustering analysis, dimension reduction, and generative models. Particularly, generative models, such as the Generative adversarial model (GAN) and Variational autoencoder-decoder (VAE), learn to create new data by discovering underlying distributions in unlabeled training data without requiring explicit target outputs. (3) Reinforcement Learning (RL): This approach involves an agent learning to make sequential decisions through interaction with an environment. Its distinctive strength lies in providing efficient solutions for combinatorial optimization problems with vast

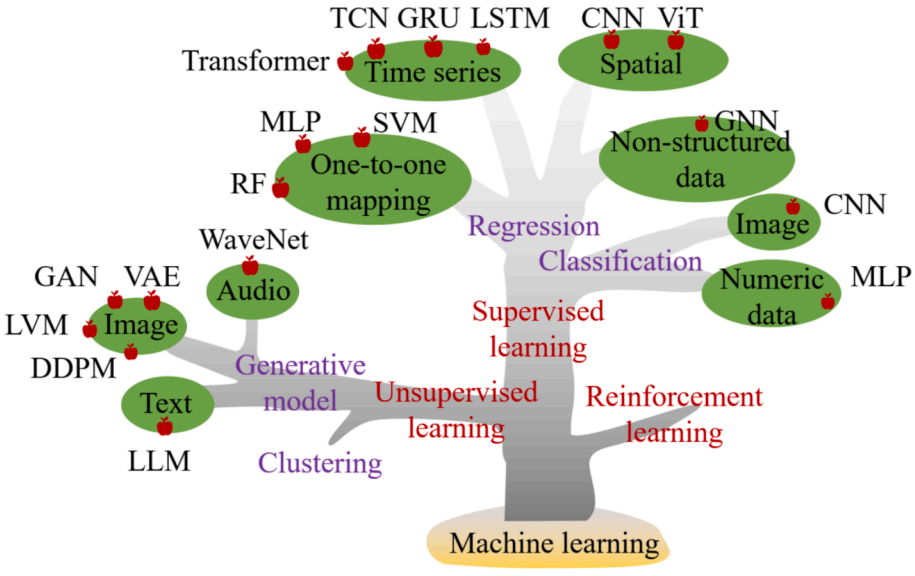


Fig. 1. A basic classification of machine learning (Note: i. Generative models are mainly unsupervised but can be supervised when using labelled data. ii. Abbreviations that are not mentioned in the aforementioned text are listed in alphabetical order. **CNN**: Convolutional neural network; **DDPM**: Denoising Diffusion Probabilistic Models; **GNN**: Graph neural network; **GRU**: Gated recurrent unit; **LLM**: large language model; **LSTM**: Long short-term memory; **LVM**: Large Vision Model; **TCN**: Temporal Convolutional Network; **ViT**: Vision Transformer).

potential combinations.

2.2. Overview of ML for CGM

As shown in Fig. 2, ML provides useful tools for data mining, surrogate modelling, , uncertainty quantification, and inverse analysis. These capabilities offer distinct advantages over physics-based computational models, empowering ML to address certain limitations in numerical simualtions. For instance, ML-based surrogate modelling can enhance computational efficiency, identify underlying patterns, and accelerate inverse analysis. Fig. 3 highlights three key applications of ML in CGM:

(1) **Accelerate CGM Computation**: ML serves as a surrogate model to replace part or all of certain processes during numerical computations. There are approximately seven types of ML-assisted acceleration approaches: (a) data-driven multiscale or material modelling; (b) MLP-based contact resolution in DEM; (c) CNN-enabled contact-based computation; (d) GNN-based simulators; (e) recurrent neural network (RNN)-based simulators, (f) ML-enhanced computational fluid dynamics (CFD)-DEM coupling simulations, and (g) case-specific surrogate modelling applications.

(2) **Uncover Hidden Physical Laws and Patterns**: ML helps reveal underlying physical laws and patterns in CGM simulation data. Typical applications include: (a) automatic recognition of grain shape and size distribution of granular materials; (b) AI-assisted 3D reconstruction of granular specimens; (c) AI-enabled generation of granular materials; (d) force chain network predictions; (e) description of granular material behaviour; (f) pattern recognition for granular dynamics; and (g) AI-assisted discovery in granular physics.

(3) **Solve Inverse Problems in CGM**: ML is utilized to address inverse problems in CGM, with three key applications: (a) efficient calibration of particle-scale parameters in DEM, (b) optimization problems in granular mechanics, and (c) physics-informed neural network (PINN) for inverse problems in granular mechanics.

3. Machine learning accelerates granular simulations

3.1. Data-driven multiscale modelling of granular materials

The macroscopic mechanical behaviour of granular media is fundamentally rooted in its microstructural evolution and particle-scale properties. Hierarchical multiscale modelling (HMM) that couples FEM and DEM is a well-recognized framework for simulating macroscale boundary value problems (BVPs) while preserving particle-scale information (Guo and Zhao, 2014). In this framework, FEM is used for BVP simulation, while DEM replaces traditional constitutive models through online simulations. During each loading step, FEM transmits deformation or strain gradients to DEM-based representative volume elements (RVEs) as boundary conditions, which then computes stress responses and tangent operators to send back to FEM (Guo and Zhao, 2016a, Qu et al., 2021c). This hybrid computational scheme allows for large-scale BVP simulation while capturing material responses from lower-scale interparticle interactions, thus bypassing the phenomenological assumptions inherent in conventional constitutive modelling (Wang and Sun, 2018). In addition to coupled FEM/DEM HMM, the MPM and PFEM offer alternative coupling approaches with DEM for multiscale modelling of large deformation problems (Liang and Zhao, 2019, Guo et al., 2021, Chen et al., 2023). However, as each Gauss point involved in the simulated continuum domain requires a corresponding DEM solution, a significant number of DEM RVEs must be computed simultaneously in a BVP, leading to substantial computational costs.

ML offers a strategy to accelerate the multiscale modelling of granular materials. This approach, referred to as data-driven multiscale modelling (DDMM), involves creating surrogate material models that link strain sequences to corresponding stress responses and tangent operators (Qu et al., 2023b), thereby eliminating the need for computationally intensive DEM computations.

Although the efficiency gains from data-driven approaches in multiscale modelling depend on factors such as particle-scale parameters in DEM and neural network architecture, previous studies have generally demonstrated one or two orders of magnitude accelerations in contrast to traditional FEM-DEM simulations.

The primary challenge in implementing DDMM is developing reliable surrogate material models using particle-scale simulation data. A related challenge is data-driven constitutive modelling (DDCM), which aims to use ML to build surrogate material models without relying on phenomenological assumptions. However, laboratory experiments often do not provide enough strain–stress data, leading both DDMM and DDCM approaches to depend on analytical or numerical data rather than experimental measurements. Given these common challenges and solutions, this section reviews advances in both DDMM and DDCM of granular media.

Granular media exhibits path-dependent and state-dependent behaviour. Path dependence indicates that the material's response depends not only on its current stress or strain state but also on its loading history. State dependency means that the material's response is

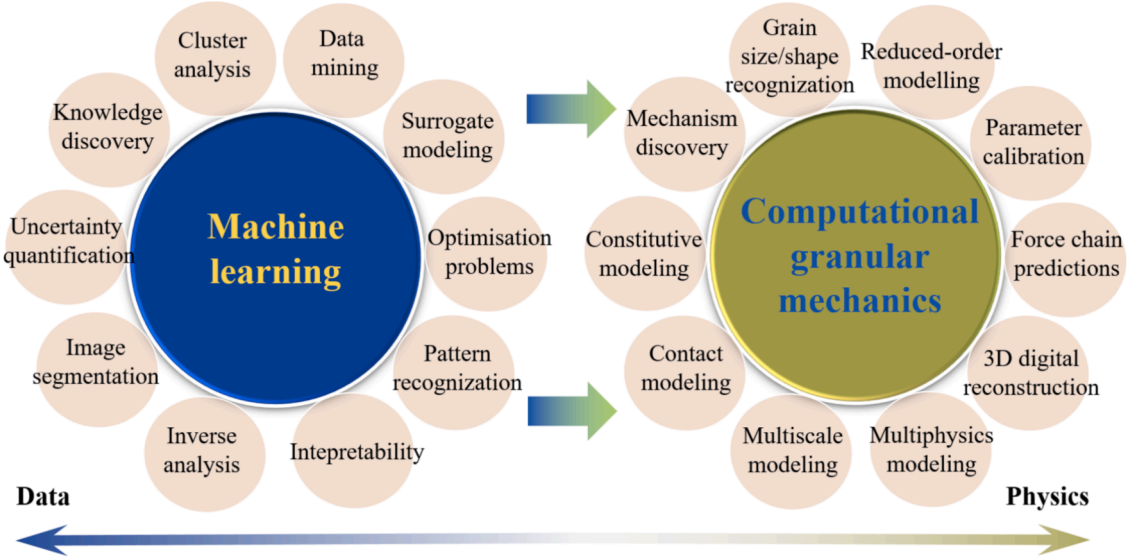


Fig. 2. Machine learning and computational granular mechanics.

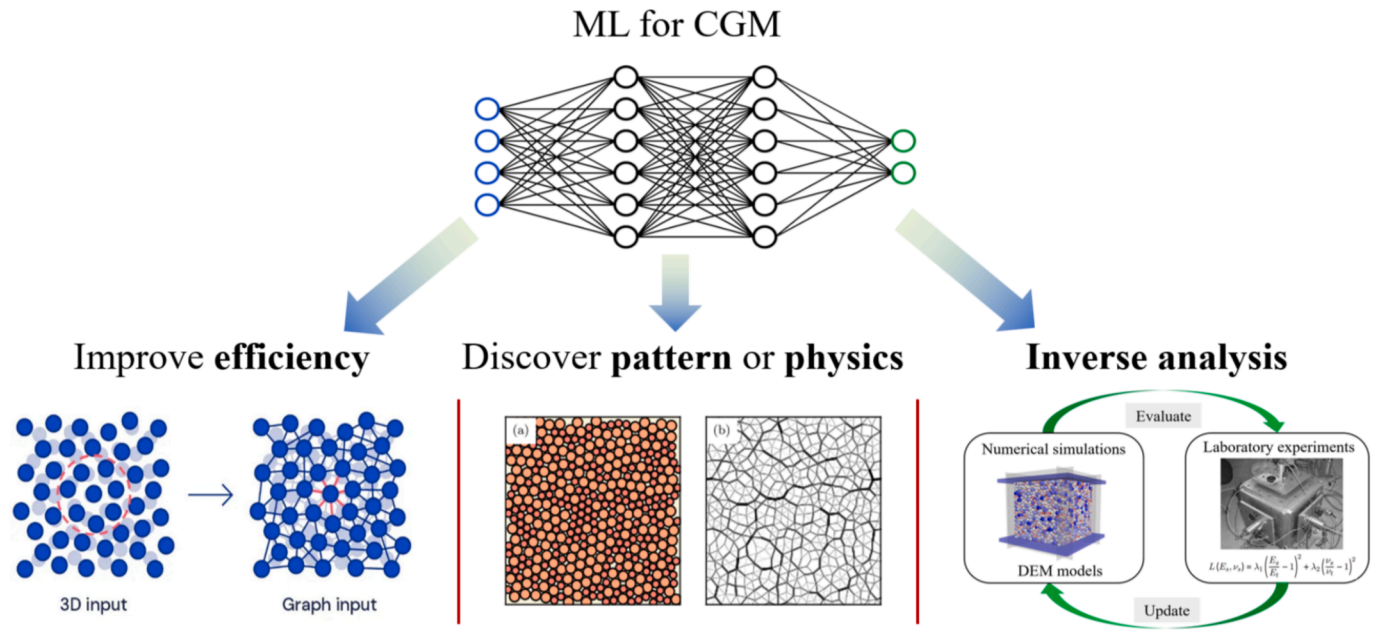


Fig. 3. A simple summary of applications of ML in computational granular mechanics (Credit: The figures on the left and right are adapted from references (Mayr et al., 2023) and (Mandal et al., 2022), respectively).

influenced by its current state, such as void ratio, confining pressure, or fabric structure. To tackle the path dependency of granular material, three strategies are commonly employed: (i) time-sequence ML models (Qu et al., 2021b), (ii) non-temporal ML models with extra inputs of previous strain states (Guan et al., 2024); and (iii) non-temporal ML models incorporating internal variables to encode loading history.

Time-sequence ML models, including LSTM networks and their variants (Zhang et al., 2020, Li et al., 2023a, Zhang et al., 2023b), GRU (Cho, 2014), and TCN (Wang et al., 2022), are effective at capturing patterns and relationships within sequential data over time. While these time-series models can effectively capture history-dependent behaviour, they typically require a large number of parameters for training, which can result in inefficient inference processes. In contrast, non-temporal ML models tend to be lighter and more efficient.

When using non-temporal ML models for stress–strain modelling of granular materials, one early-stage solution involves adding previous strain states as additional inputs (Ghaboussi et al., 1991). However, it remains unclear how many prior strain states are necessary to adequately capture stress responses. The other scheme is to introduce several variables to encode loading histories (Guan et al., 2023a). A challenge with non-temporal ML models is their difficulty in learning time-series relations, as the target relation represents a "one-to-many" mapping, lacking unique solutions for a given input. By introducing extra variables, this "one-to-many" mapping can be transformed into a unique "one-to-one" mapping, as illustrated in Fig. 4. Representative external variables may include the accumulation of absolute strain increments (Huang et al., 2020, Guan et al., 2023b) and the Frobenius norm of the strain tensor (Wang et al., 2024a). In addition, state descriptions can be tackled either by directly incorporating state parameters, such as initial void ratio (Ghaboussi and Sidarta, 1998), overconsolidation ratio (OCR) (Ellis et al., 1995), and anisotropic microstructure (Guan et al., 2024), or by using modified network architecture to store initial states of materials (Ma et al., 2022a).

The strain-based internal variables previously mentioned can be considered measurable and serve as direct inputs in model training to characterize external loading conditions. In contrast, some internal variables employed in constitutive modelling are not easily measurable using standard experimental setups. Directed graphs offer a solution by implicitly incorporating these hard-to-measure internal variables into stress–strain prediction. For example, by formulating constitutive models as information flow in directed graphs that connect critical internal variables related to stress–strain relations (see Fig. 5a), a deep reinforcement learning strategy can identify the optimal information flow that maximizes prediction accuracy (Wang and Sun, 2019). The

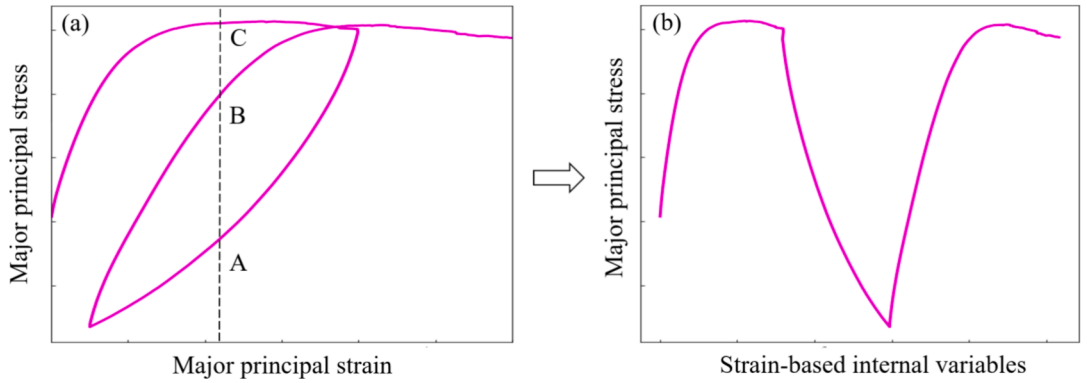


Fig. 4. Loading-history dependent stress–strain behaviour for granular materials: (a) "one-to-many" mapping where a strain state may correspond to multiple stress responses; (b) "one-to-one" mapping where an internal variable is used to encode loading history.

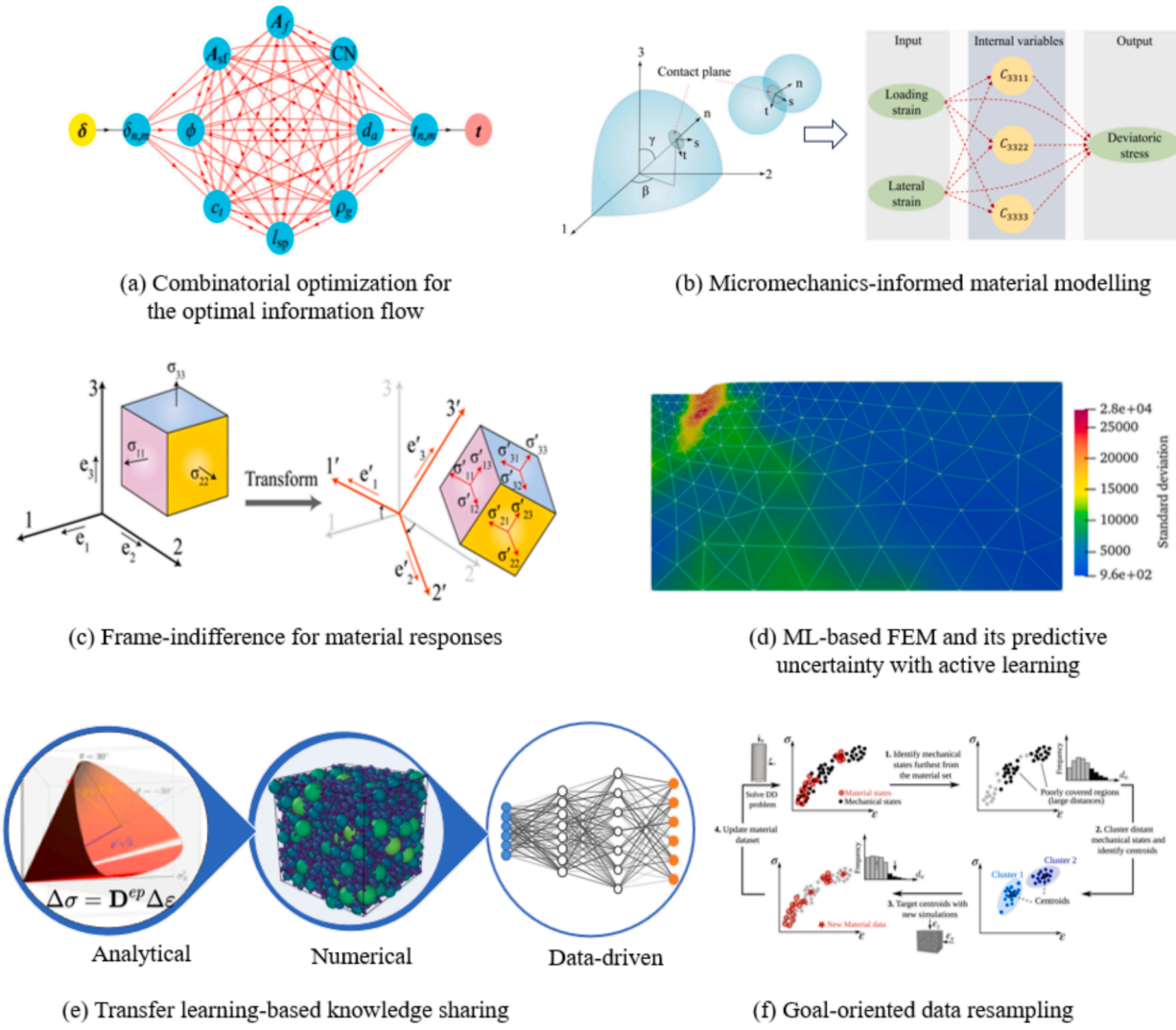


Fig. 5. Representative learning strategies in data-driven multiscale modelling. Note: Subfigures (a) and (f) adapted from (Wang and Sun 2019) and (Gorgogianni et al., 2023), respectively.

contact-scale microstructure is linked to macroscopic stress and force responses in granular media (Mital and Andrade, 2022). Directed graphs can leverage the evolution of microstructures in sheared granular materials to predict stress responses without requiring explicit inputs about the microstructures (see Fig. 5b) (Qu et al., 2021a).

Leveraging physics or domain knowledge to enhance DDMM is crucial for improving the model's predictive capability. There are four main strategies for incorporating physical knowledge: (i) Physics-constrained soft constraints: Examples include thermodynamic constraints (such as the non-convexity of the yield function and non-negative plastic work) used as penalty terms in the objective function (Vlassis and Sun, 2021, Su et al., 2024). (ii) Physics or knowledge-inspired hard constraint: For example, the positive-definite assumption of the stiffness matrix of material and the formulation of thermodynamics can be used to tailor network architectures as hard constraints (Xu et al., 2021b, Masi and Stefanou, 2023). (iii) Using physics to manage training data: For example, the principle of frame indifference (Fig. 5c) simplifies the general stress and strain tensor to their invariants (Ling et al., 2016, Heider et al., 2020), while the symmetry assumption of materials can reduce the dimensionality of the training data (Vlassis and Sun 2021). (iv) Integrating ML into established phenomenological models (Wang et al., 2025a): For example, within the framework of plasticity, ML can replace certain components, such as the yield surface (Vlassis and Sun, 2022, Nascimento et al., 2023), hardening laws (Li et al., 2019, Wen et al., 2021), internal state variables (Eghbalian et al., 2023), and stress-integration schemes (Fazily and Yoon, 2023). In thermodynamics-based constitutive models, energy potential and dissipation pseudo-potential can be direct outputs of neural networks, with final stress responses obtained by differentiating the neural network outputs with respect to their inputs using automatic differentiation (Masi et al., 2021, Zhang et al., 2025b). In addition, under the assumption of hyperelasticity, stress is derived from the strain energy density with respect to strain. These gradient relationships can be incorporated into model training using Sobolev training (Vlassis and Sun, 2021, Zhang et al., 2022d).

Data is essential for training reliable data-driven material models. One key effort is to ensure that the data closely reflects genuine material behaviour, such as by considering realistic particle morphology in discrete element modelling (Wu and Wang 2022a, Xiong et al., 2023). Additionally, there is a growing focus on developing algorithms that maximize data utilization. For instance, deep active learning can identify the most informative data for training ML models. By training a committee of ML models simultaneously, the variance in these predictions can provide a measure of uncertainty, aiding in interactive model training and identifying unreliable predictions (Qu et al., 2023a, Zhang et al., 2023d), as seen in Fig. 5d. Furthermore, transfer learning

techniques (Qu et al., 2023b, Xiong and Wang, 2024) and multi-fidelity networks (Zhang et al., 2022b, He et al., 2023, Su et al., 2023) can leverage phenomenological models or existing material data to assist in training new material models (see Fig. 5e).

Data-driven multiscale modelling also includes model-free approaches (Kirchdoerfer and Ortiz, 2016). This approach assigns stress responses to each Gauss point without using any ML models by identifying the closest strain sequence in a strain–stress pair dataset that satisfies conservation laws and constraints. This framework has been applied to the multiscale modelling of granular media, with material data sourced from discrete element modelling (Karapiperis et al., 2021). To improve prediction accuracy, a goal-oriented data sampling strategy is introduced to identify the most valuable data for model-free modelling (see Fig. 5f), where BVPs are formulated as distance minimization problems (Gorgogianni et al., 2023). A similar model-free approach has been adapted to account for the evolution of particle size distribution (PSD) due to particle breakage in granular materials (Ulloa et al., 2023). In addition, DDMM is applied to the thermal analysis of granular media using MLP to relate microstructure properties (porosity and fabric) to the dimensionless thermal conductivity tensor (Rangel et al., 2024b). Data-driven thermo-mechanical simulations of BVPs are also conducted using the Finite Volume Method (FVM), with trained data simulated from offline DEM computations (Rangel et al., 2024a).

3.2. MLP-based contact resolution for discrete element modelling

Fig. 6 shows a typical computation cycle for conventional DEM simulations, involving four key steps: (1) detecting and resolving contacts, (2) determining contact forces, (3) calculating particle motion, and (4) updating particle positions. In the first step, contact detection identifies if two particles or a particle and a boundary are in contact. Contact resolution involves calculating the geometric properties necessary to evaluate contact forces, including contact overlap $\delta_{\mathbf{n}}$, contact normal, and contact point (Fig. 6a). These contact geometric features are used to calculate contact forces following a chosen contact model (Fig. 6b). Kinematical information of particles is solved with Newton–Euler equations (Fig. 6c). The central difference time integration scheme is further used to update particle positions and orientations (Fig. 6d).

Non-spherical particle-based DEM is computationally intensive, primarily due to the contact detection and resolution steps. Since the solutions for steps (2)–(4) are well-established and already efficient for conventional DEM, researchers proposed using neural networks (NNs) to detect particle contacts and resolve contact geometric features (Hwang et al., 2022, Lai et al., 2022b), aiming to reduce computational costs for non-spherical particle-based DEM simulations. In the ML-based contact resolution framework, all standard computational procedures in DEM are retained, except that two NNs are designed for contact detection and resolution. As illustrated in Fig. 7, the first NN performs a classification task to determine whether two elements (particle–particle or particle–wall) are in contact. The second NN is a regression model used for computing contact geometric features. Both networks receive particle geometric descriptors as inputs. The classification network outputs a binary contact status (true or false), while the resolution

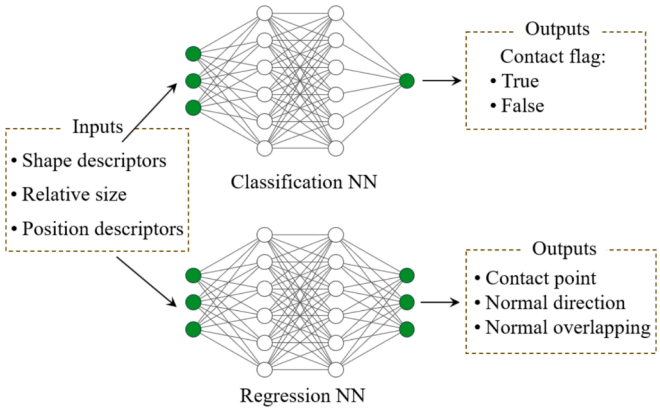


Fig. 7. MLP-based contact detection and resolution in DEM.

network calculates contact geometric features such as contact overlap, contact normal, and contact point.

The geometric descriptors of a particle or boundary can be categorized as three components: size, shape, and position. The size descriptor is defined by the diameter of a circle that has an equivalent area to the particle. The position descriptor includes the particle's centroid coordinates and its rotation angle from the initial alignment. The shape descriptors consist of a series of parameters derived from the associated shape function that defines the particle surface. In 2D DEM, boundary elements can be represented as open polylines or curves in a Cartesian coordinate system. A line can be described by the equation $x\sin\alpha - y\cos\alpha - d_o = 0$, where α and d_0 serve as position descriptors.

The data for training the NNs come from an artificial generation by considering all possible "particle–particle" and "particle–wall" contact scenarios. In the particle size aspect, a particle is set to be a unit size (object particle), whereas the other contacting particle (cue particle) is of random size in the specified range of relative particle size (e.g. 1–10). The range of shape descriptors is contingent upon the specified particle shape templates. For example, elliptical particles have a shape parameter represented by the aspect ratio (major axis divided by the minor axis), which is randomly selected within a specified range. For the position descriptor, different position scenarios are created by rotating different angles of particles and walls, and the distance from the particle to wall element.

Several demonstration cases show that MLP-based 2D DEM produces results comparable to conventional DEM while reducing computation time by 50 %. This method has recently been extended to 3D DEM scenarios (Huang et al., 2024). In contrast to traditional 3D DEM simulation methods, the 3D ML-based DEM can achieve around five times greater computational efficiency than the conventional DEM model in terms of the Cundall number.

Instead of predicting contact geometry, an MLP with residual structure is employed to directly output resultant interparticle forces based on particle positions and velocities (Zhou et al., 2025b). The loss function includes a physics term that enforces the zero-sum constraint of particle interaction forces. The approach achieves a 7–10 times speedup over traditional DEM models in 2D wall-bounded swirling flow scenarios

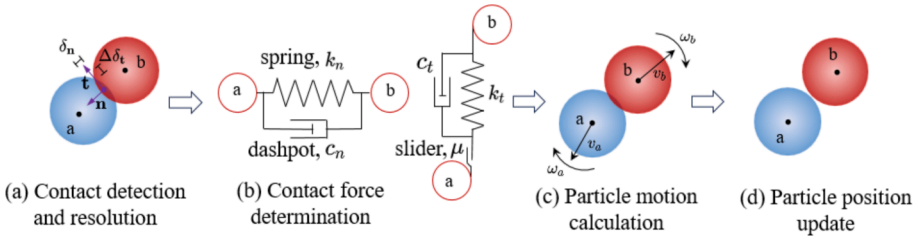


Fig. 6. A typical computation cycle for DEM.

with hundreds of particles. However, its applicability to more complex particle systems remains unexplored. MLP-based contact resolution can accelerate computations in DEM, but its effectiveness heavily depends on the quality and diversity of training data. Since contact interactions are inherently linked to particle shapes, any change in shape requires retraining the model with new data to account for new contact scenarios. Natural particle shapes vary significantly, resulting in a vast range of contact scenarios. To prevent unacceptable extrapolation errors, this method may need to be restricted to a limited number of particle shapes in simulation studies.

3.3. CNN-assisted discrete element modelling

CNNs are a type of ML method which employs the convolution operator to capture local geometrical patterns. They are a natural choice for learning local collision-related mechanical problems. A CNN-based numerical framework initially developed for fluid simulations (Ummenhofer et al., 2019) was subsequently extended to simulate DEM scenarios (Lu et al., 2021). Since DEM primarily involves collision-based interactions, the contact forces on a particle can be determined once the positions of its neighbouring particles are known. In this CNN-based DEM framework, simulated particles are termed "dynamic particles", while boundary conditions are treated as "static particles". Representing boundaries as static particles allows the NNs to extrapolate to new scenarios and learn collision interactions within a unified framework. Unlike traditional DEM simulations, this approach does not require contact detection, resolution, or the computation of contact forces.

Note that only collision-related interactions are considered while the non-contact forces, e.g. gravity force, are dealt with alternatively. There are two different processing measures. One strategy introduces the concepts of intermediate positions $\mathbf{x}_i^*$ and velocities $\dot{\mathbf{X}}_i^*$ of particles (Lu et al., 2021). These two intermediate variables are updated as shown in Eqs (1–2) by only considering gravity forces (without incorporating contact interactions):

$$\dot{\mathbf{X}}_i^* = \dot{\mathbf{X}}_i^n + \mathbf{g}\Delta t \tag{1}$$

$$\mathbf{x}_i^* = \mathbf{x}_i^n + \Delta t \frac{\mathbf{v}_i^* + \mathbf{v}_i^n}{2} \tag{2}$$

where n denotes time steps; i represents the i^{th} particle in the system; $\mathbf{x}_i^n$ denotes the position of the i^{th} particle at the n^{th} time step; $\mathbf{g}$ is gravitational acceleration, and Δt is the time step.

A neural network further is trained to establish the relation between the intermediate velocities $\dot{\mathbf{X}}_i^*$ and the position correction terms $\Delta\mathbf{x}_i$, which are added to the intermediate positions $\mathbf{x}_i^*$ to obtain final particle positions at the next time step.

$$\mathbf{x}_i^{n+1} = \mathbf{X}_i^* + \Delta\mathbf{x}_i \tag{3}$$

Then the final velocities are approximated via the backward finite difference method based on particle positions:

$$\dot{\mathbf{X}}_i^{n+1} = \frac{\mathbf{x}_i^{n+1} - \mathbf{x}_i^n}{\Delta t} \tag{4}$$

This method is criticized for being inconsistent with Newton's second law of motion, which states that overall particle acceleration should be determined first, followed by calculating the corresponding particle velocities and positions. The alternative strategy addresses this downside by predicting acceleration directly based on particles and their neighbouring positions (Xu and Shen, 2022).

$$\dot{\mathbf{X}}_i^{n+1} = \dot{\mathbf{X}}_i^n + \mathbf{g}\Delta t + \Delta t\ddot{\mathbf{X}}_{ext}^{n+1} \tag{5}$$

$$\mathbf{x}_i^{n+1} = \mathbf{x}_i^n + \Delta t \frac{\dot{\mathbf{X}}_i^{n+1} + \dot{\mathbf{X}}_i^n}{2} \tag{6}$$

where the extra acceleration $\ddot{\boldsymbol{X}}_{ext}^{n+1}$ is the sum of the acceleration other than gravity. Note that, the particle–particle and particle–wall collisions can be considered separately in this approach via:

$$\ddot{\boldsymbol{X}}_{ext}^{n+1} = \ddot{\boldsymbol{X}}_{p-p}^{n+1} + \ddot{\boldsymbol{X}}_{p-w}^{n+1} \tag{7}$$

where $\ddot{\boldsymbol{X}}_{p-p}^{n+1}$ and $\ddot{\boldsymbol{X}}_{p-w}^{n+1}$ represent the acceleration induced by particle–particle and particle–wall collisions, respectively.

As shown in Fig. 8(a) and (c), the inputs for these two CNN-based strategies are different. The first category involves continuous convolutions that process the features of static boundary particles and moving particles. The input parameters consist of the normal direction for static particles and the velocities of the moving particles. The second category comprises a fully connected layer that takes the features of each particle (e.g., density, diameter) as inputs. After several layers of convolution computations, the outputs are either the position correction terms $\Delta\mathbf{x}$ (for the first strategy) or the extra acceleration (for the second strategy).

Three key techniques are employed in the above CNN-assisted computational framework. First, 3-D continuous convolution (Ummenhofer et al., 2019) is employed to build NNs. Second, a function $\Lambda(\mathbf{x}_i\text{-}\mathbf{x})$ maps a unit sphere to a unit cube, to accommodate unstructured particle assemblies within a spherical domain and enable the values of the filter function to be stored on regular lattices in the original CNN framework. Finally, a multiscale and multistep loss function is utilized to improve the generalization of CNNs in capturing physical granular flow. "Multiscale loss" refers to both the particle-scale loss (the discrepancy between the predicted and ground-truth positions) and macro-scale loss (the difference between the geometry centres of the predicted and actual particle assemblies), as shown in Eq. (8).

$$L^{n+1} = \alpha\frac{1}{N}\sum_{i=1}^{N}\left\|\mathbf{X}_i^{n+1} - \widehat{\mathbf{X}}_i^{n+1}\right\|_2 + (1-\alpha)\left\|\frac{1}{N}\sum_{i=1}^{N}\left(\mathbf{X}_i^{n+1} - \widehat{\mathbf{X}}_i^{n+1}\right)\right\|_2 \tag{8}$$

where α is a weighting factor between 0 and 1 to adjust the contributions of both terms. The multi-step loss function involves using a weighted sum of the loss functions L^{sum} from multiple frames during training, as shown in Eq. (9):

$$L^{sum} = \sum_{i=1}^{F} w^i L^{n+i} \tag{9}$$

where F is the total number of frames utilized for the loss calculation; w^i denotes the weighting factor of the loss computed from the frame $n+i$.

The training data for the CNN-based framework comes from DEM simulation frames. The first strategy reportedly achieved a speedup of 78 times compared to traditional MFiXDEM with little loss of accuracy when modelling a hopper system. By contrast, the second strategy achieves a speedup of about 7 times faster than DEM. Although the acceleration from the second strategy is less pronounced, its predicted accuracy is generally higher than that of the first strategy, especially in capturing the granular flow patterns by considering both particle–particle and particle–wall collisions.

However, CNN-assisted surrogate DEM modelling suffers from some drawbacks, including sensitivity to hyperparameters and the need for retraining when particle geometry and gradation change. To develop a universally applicable model for various scenarios, a large number of input cases with diverse geometries and particle characteristics are necessary.

3.4. GNN-based simulators for particle-based computations

The MLP-based contact resolution introduced in Section 3.2 and the CNN-assisted DEM in Section 3.3 aim to replace several steps in standard DEM procedures. The graph network-based simulators (GNS) can substitute the entire DEM simulation process. The fundamentals of GNS

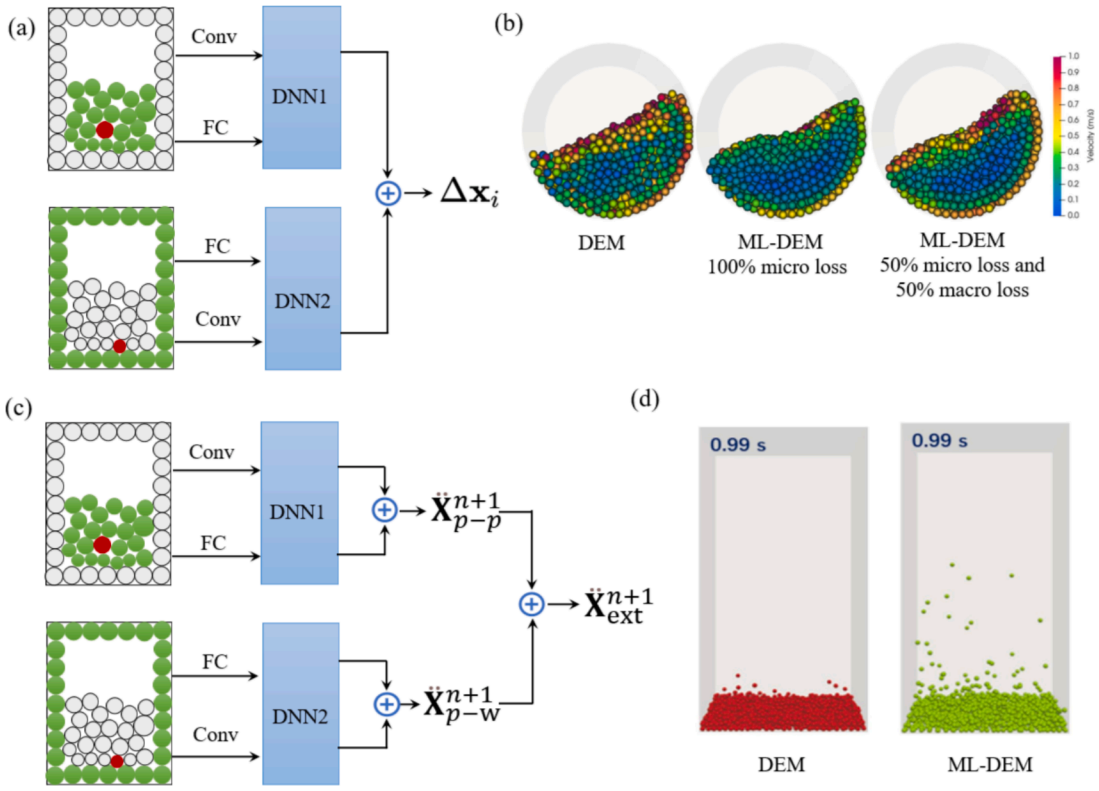


Fig. 8. CNN-assisted DEM simulations: (a) the first type of CNN-based DEM; (b) Comparison of traditional DEM and the first type of CNN-based DEM in rotating drum simulations; the image credits to (Lu et al., 2021); (c) the second type of CNN-based DEM; (d) Comparison of traditional DEM and the second type of CNN-based DEM in particle settlement simulations; the image credits to (Xu and Shen 2022). Note: FC and Conv represent fully connected and convolutional layers, respectively.

start from the concept of graph. In computer science, a graph **G** is a data structure consisting of vertices and edges, i.e. **G = (V, E)**, where **V** and **E** are sets of nodes and edges that connect nodes, respectively. GNN is a class of ANNs, originally developed for handling data that can be represented as graphs (Scarselli et al., 2008). As shown in Fig. 9, the basic structure of Graph Neural Networks (GNNs) follows a "graph-in, graph-out" architecture. It can make predictions at the node, edge, and entire graph levels. The embedding vectors for graph attributes (nodes, edges, global) can be updated by using trained MLPs.

GNS is an application of GNN originally designed to represent the state of a granular physical system (Sanchez-Gonzalez et al., 2020). In the framework of GNS, nodes represent the physical states of particles and learned message-passing through edges serves to depict interparticle interactions (Choi and Kumar, 2024a). As shown in Fig. 10, a typical GNS consists of a dynamics approximator d_θ for computing spatial interaction (including an encoder, a processor, and a decoder)

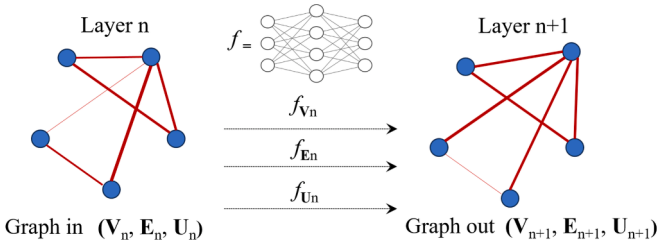


Fig. 9. Illustration of a single layer of a GNN. **V** represents vertex (or node) attributes, e.g., node identity, number of neighbours; **E** denotes edge (or link) attributes and direction, e.g., edge identity, edge weight; **U** is global (or master node) attributes, e.g., number of nodes, longest path. The n^{th} layer graph is used as input with each component (**V, E, U**) being updated with MLPs to produce the $(n + 1)^{th}$ layer graph. (Image is adapted from https://distill.pub/2021/gnn-intro/).

and an updater for time integration.

(a) The encoder serves to preprocess particle-based state representation as a graph structure G^0 by assigning a node to each particle and adding an edge between any two particles close enough below a threshold called *connectivity radius*. This connectivity radius represents the local interactions of particles and remains constant for each simulation. The raw input data are each particle's state vector $\boldsymbol{S}_i^t$, which includes: the particle position at the t^{th} step $\mathbf{X}_i^t$, a sequence of previous velocities (c is the number of time steps involved) $\left\{\dot{\mathbf{X}}_i^{t-c+1}, \cdots, \dot{\mathbf{X}}_i^t\right\}$ and features, $\boldsymbol{f}_i$, representing material properties (e.g. Young's modulus), and the global properties of the system (e.g. gravity).

$$\mathbf{S}_i^t = \left[\mathbf{X}_i^t, \dot{\mathbf{X}}_i^{t-c+1}, \cdots, \dot{\mathbf{X}}_i^t, \boldsymbol{f}_i\right] \tag{10}$$

The velocities of particles involved in Eq. (10) are computed via the finite difference of the position sequences as below:

$$\dot{\mathbf{X}}_i^t = \left(\mathbf{X}_i^t - \mathbf{X}_i^{t-1}\right) \Big/ \Delta t \tag{11}$$

(b) The processor is a GNN which simulates the node interactions via multiple steps of learned message-passing, to establish interactions for each edge in the graph, aggregate messages from all interacting edges for each node, and update both node and edge features. Normally, MLPs are used to learn these interactions. Note that the number of message-passing steps is a hyperparameter controlling the spatial information propagation within a single time step, and it differs from time discretization steps used in numerical integration.

(c) The decoder uses an MLP to predict accelerations $\ddot{\mathbf{X}}_i^t$ of particles using the updated node/edge embeddings from the processor. This decoder can be regarded as a postprocessor for the GNN.

When the spatial dynamics of the granular system are computed,

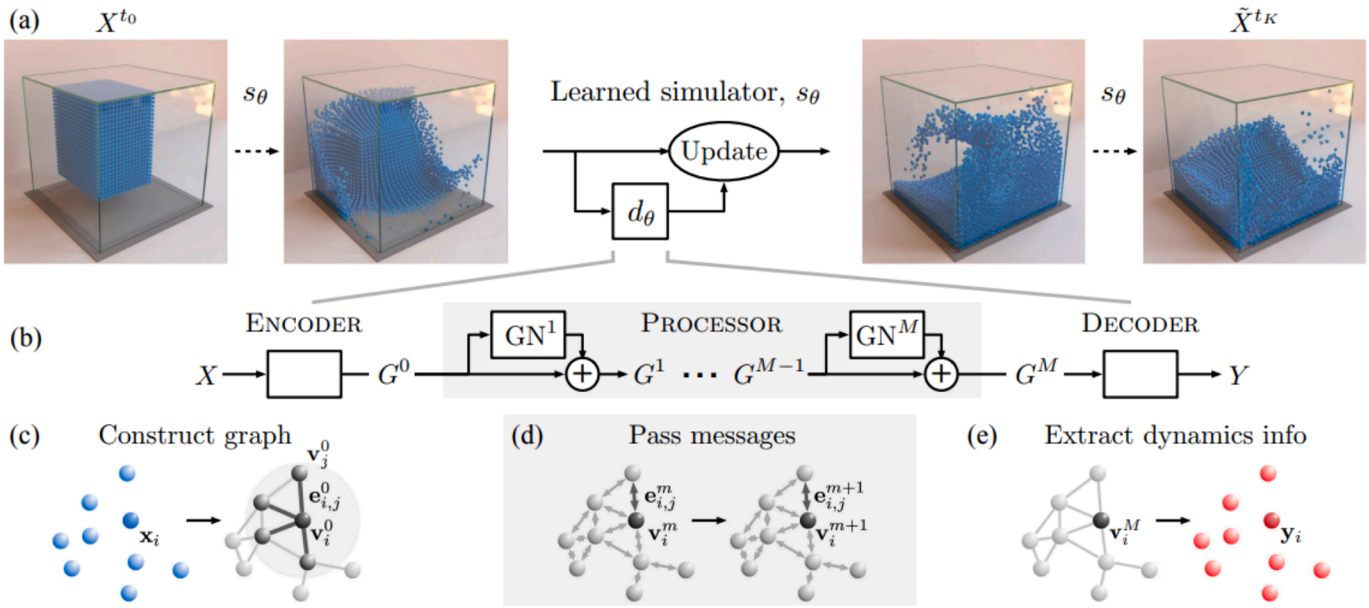


Fig. 10. The basic computational workflow of GNS: (a) GNS serves as a learned dynamics model, d_θ, to predict future granular state; (b) the d_θ uses an "encoder-process-decoder" architecture to compute dynamic information Y using input state X; (c) the encoder constructs a latent graph from the input state; (d) the processor performs message-passing over latent graphs; and (e) the decoder gives dynamic predictions based on the final latent graph [Image sourced from (Sanchez-Gonzalez et al., 2020)].

time integration is conducted to update the physical state at the next time step. For example, the predicted accelerations are then used to update velocities and displacements of the particles via Euler integration as shown in Eqs (12)-(13):

$$\dot{\mathbf{X}}_i^{t+1} = \dot{\mathbf{X}}_i^t + \ddot{\mathbf{X}}_i^t \Delta t \tag{12}$$

$$\mathbf{X}_i^{t+1} = \mathbf{X}_i^t + \dot{\mathbf{X}}_i^{t+1} \Delta t \tag{13}$$

The key advantage of GNS is its ability to significantly reduce the computational costs of conventional particle-based physical simulators (up to 10-100x faster) while maintaining comparable accuracy. Additionally, the fundamental concept behind GNS bears similarities to that of physics-based particle simulations, making it an attractive and widely accepted approach (Zhang et al., 2023a). However, as a supervised surrogate model, GNS requires a substantial amount of data from various granular flow scenarios to ensure reliable generalizability. Furthermore, the original GNS framework faces challenges in accurately simulating boundary conditions and demanding high training costs.

To enhance GNS's ability to simulate arbitrarily shaped boundaries, the boundary graph neural network (BGNN) was proposed to improve the learning of particle–wall interactions, as shown in Fig. 11. This is achieved by dynamically inserting virtual particle nodes when the distance between particles and walls falls below a specified threshold (Mayr et al., 2023). Additional features, such as wall normal vector information, are included in the GNN inputs. Each particle can interact

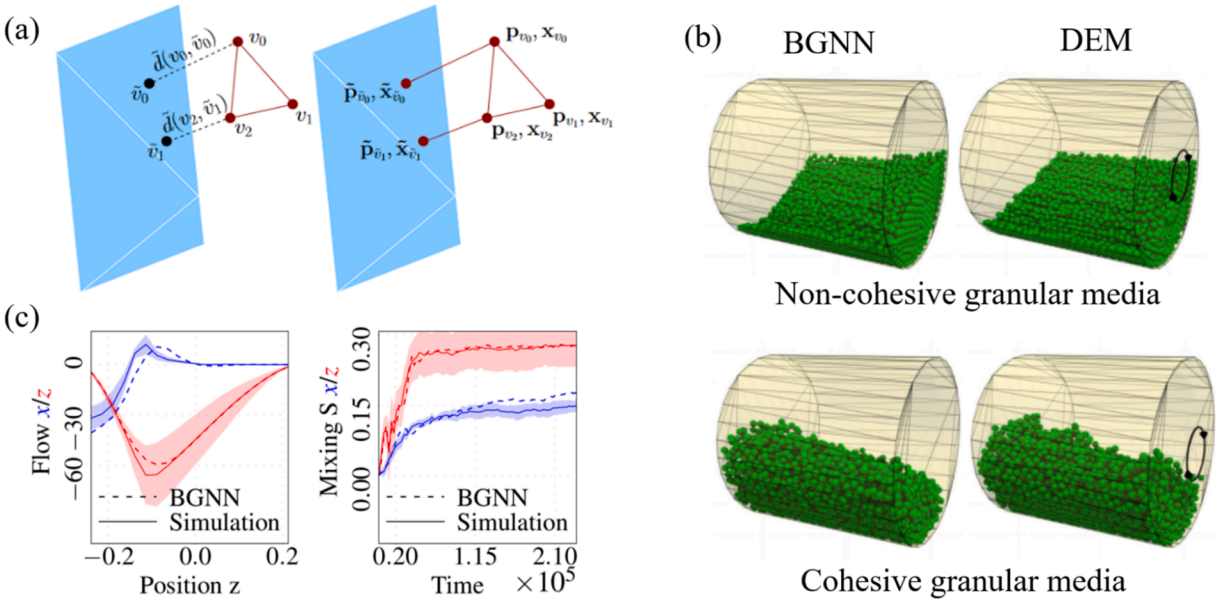


Fig. 11. BGNN-based simulations: (a) the virtual nodes method for simulating particle–wall interactions; (b) qualitative and (c) quantitative forecast performance comparisons of rotating drum simulations between BGNN and DEM (Mayr et al., 2023).

with at most one virtual particle, which is the closest to the real particle, representing the boundary surface. The other work to improve the simulation of boundaries is by using a signed distance function-based graph neural network (SGN). In this framework, particle dynamics are represented through graph structures, with arbitrarily shaped wall boundaries modelled using signed distance function fields. The SGN has demonstrated acceleration exceeding 100x in several simulations (Li and Sakai, 2024).

GNS typically requires substantial computational resources during training, particularly for cases involving a massive number of particles. To reduce its computational costs, principal component analysis (PCA) is utilized to perform dimensionality reduction on the trained graph data. GNS is trained and used in reduced data space. Subsequently, the predicted data is mapped back to the original full space. This type of PCA-assisted subspace graph network is found to improve 700x acceleration than general MPM-based machine-terrain interaction simulations (Haeri et al., 2024).

Another noteworthy advancement in GNS is the Temporal Graph Neural Network-based Simulator (TGNNS) (Zhao et al., 2025). In contrast to traditional GNS, TGNNS incorporates three improvements: (1) a node historical state to store the embedding information of particle dynamics; (2) dynamic graphs to capture the evolving nature of granular flow; and (3) the omission of edge information in inter-graph message passing, aligning with the computational procedures of MPM methods. The TGNNS is approximately 100x faster than GPU-based hierarchical MPM-DEM simulations.

3.5. RNN-accelerated DEM simulator

A recurrent neural network with stochastically calculated random motion (RNNSR) is developed to accelerate DEM simulations (Kishida et al., 2023). The main idea of this method includes: (1) DEM simulation data of individual particle trajectory is decomposed into local mean and local variability components; (2) RNN is used to learn the evolution of local mean particle dynamics; (3) the simulation system is divided into fixed Cartesian grids, and a Gaussian random (SR) model is used to predict the local variability component of particle trajectory within each grid; and (4) the time evolution of an individual particle motion is calculated by using the addition of local mean and local variability components. In contrast to the DEM simulations of powder mixing processes, RNNSR can achieve 355 times faster computational speed with comparable accuracy. This acceleration can be attributed to two factors: (1) RNNSR reads the current frame of particle positions to predict the positions at the next frame, bypassing complex computational procedures of DEM as outlined in Fig. 6; and (2) the time step in RNNSR can be much larger than the one used in DEM. Further research shows that the mean period of the original particle trajectory data can be used to determine an appropriate time step for the RNNSR (Kishida et al., 2024).

3.6. ML-accelerated CFD-DEM simulations

CFD-DEM is a prevailing method for simulating fluid-particle systems, with CFD modelling continuous fluid flow and DEM handling the motion and collision of particles. As shown in Fig. 10, CFD-DEM simulations can be roughly categorized into unresolved (Vidyapati and Subramaniam, 2012) and resolved approaches (Hager et al., 2014), based on the relative sizes of fluid grids and particles. The resolved CFD-DEM directly simulates fluid-particle interactions on fluid meshes that are much smaller than the particles (Hu et al., 2001). While theoretically rigorous and accurate, the resolved CFD-DEM is computationally demanding. In contrast, the unresolved CFD-DEM framework typically uses an empirical drag force model to represent the force exerted by the fluid on a particle, while the force exerted by particles on the fluid is calculated based on reaction forces or approximated using Darcy's law. Although the unresolved approach offers significant acceleration compared to the resolved CFD-DEM, it remains inefficient for engineering applications. Consequently, ML is explored to enhance the efficiency of both resolved and unresolved CFD-DEM simulations.

In the unresolved approach, the fluid cell should be about 3 to 5 times the particle size to ensure sufficient accuracy (Wang et al., 2019b). However, achieving realistic results often requires a smaller mesh, which increases computational demands. To address the challenges of using a large mesh in modelling and to take advantage of faster computations, ML models are employed to develop a mesh-independent drag force model for particles. Research has shown that ML-assisted coarse-scale coupling significantly improves efficiency and closely resembles fine-scale modelling, outperforming classical coarse-scale simulations (Davydzenka and Tahmasebi, 2022).

In the resolved approach, calculating drag force is computationally intensive. To improve computational efficiency, an MLP-based drag force model can be trained using data from a sufficient number of particle-resolved simulations (He and Tafti, 2019). A physics-guided model architecture (PhyNet) was designed based on the understanding that the pressure field and the velocity field around a particle directly affect the pressure component and shear component of the drag force. The pressure field and the velocity field, as two physical intermediate variables, can be incorporated into the PhyNet to enhance generalizability (Muralidhar et al., 2020). Additionally, the interaction between non-spherical particles and fluid has also been explored (Ashwin et al., 2022).

A significant challenge in ML-assisted particle–fluid flow scenarios is effectively describing the intricate shape and orientation of non-spherical particles using a limited number of digital descriptors. Spherical harmonics (SH) is a common method to reconstruct particle shapes (Xiong and Wang, 2021). The primary energy spectra of SH frequencies describe particle shape, while Euler angles specify particle orientation (Hu et al., 2024). These indicators, along with flow conditions, are directly input into an MLP to predict the drag coefficient. Alternatively, given the challenge of characterizing particle shapes using hand-crafted parameters, the ML-based data dimension reduction method can be a viable strategy for describing complex grain shapes. For example, in Fig. 12, a VAE model was used to extract latent geometric features (Hwang et al., 2021).

Although existing research has shown that ML models can predict particle drag force with accuracy comparable to the resolved approach while reducing computation time significantly, the incorporation of ML-based drag force models in resolved CFD-DEM simulations remains an emerging area. The inputs and outputs for predicting drag forces vary based on distinct training strategies. Table 1 gives an overview of current ML-based drag force prediction models in CFD-DEM simulations. In addition to ML models, other mathematical methods, such as proper orthogonal decomposition (POD), can be utilized for dimensionality reduction and data analysis to accelerate CFD-DEM simulations (Hajisharifi et al., 2023). However, the discussion of these statistical techniques is beyond the current scope and will not be included in this review.

3.7. State-based mapping of ML for accelerating predictions of granular flow

The previous sections reviewed the integration of ML models with process-based granular simulations. While incorporating complete physical processes improves an ML model's ability to generalize across various scenarios, it also significantly increases computational costs. Unlike process-based simulations of DEM, some studies employ ML for end-to-end state predictions (Wu et al., 2024) or for establishing direct mapping relationships among variables (Mahdi and Holdich, 2017). These approaches are limited to specific training scenarios but are often more computationally efficient. For example, one study used an MLP to predict the mass discharge rate from conical hoppers based on DEM simulation data, utilizing particle properties and hopper angle as inputs

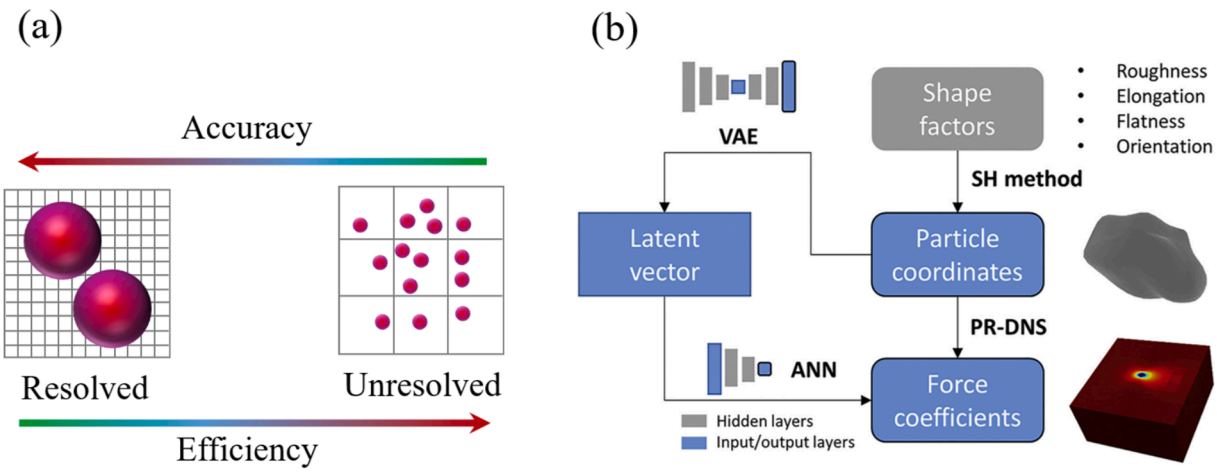


Fig. 12. ML for CFD-DEM simulations: (a) Resolved and unresolved CFD-DEM coupling methods; (b) an ML-assisted drag force prediction framework in CFD-DEM simulations with non-spherical particles, image credit: (Hwang et al., 2021).

(Kumar et al., 2018), without considering individual particle trajectories and the flow process. Another study applied ML to predict the passing flow rate of a vibrating screen, based on particle and screen information derived from DEM simulations (Arifuzzaman et al., 2022). As shown in Fig. 13, a hybrid CNN and LSTM network was utilized to capture spatial–temporal discharge patterns of granular flow in a wedge-shaped hopper, with the CNN part focusing on extracting image features, while the LSTM component for time-sequence predictions (Liao et al., 2021). The same CNN-LSTM model was also employed to predict mixing and segregation behaviours in a bidisperse solid − liquid fluidized bed (Xie et al., 2022). Furthermore, ML was used to predict the radial segregation of granular materials in a heap, demonstrating that incorporating the initial configuration of the mixture and DEM parameters in inputs can enhance prediction accuracy (Hadi et al., 2024).

3.8. A summary of ML-accelerated granular simulations

The application of ML to accelerate granular simulations is a relatively new topic, but the concept of reducing computational complexity to improve efficiency in physical simulations has been explored for decades. Similar to traditional reduced order models (ROMs), which are classified as intrusive or non-intrusive, the application of AI can also be categorized this way. Intrusive ROMs integrate the AI model directly into traditional computational workflow, as seen in DDMM (Section 3.1), MLP-based contact resolution (Section 3.2), CNN-based contact force determination in DEM (Section 3.3), and surrogate models of drag force in CFD-DEM coupling (Section 3.6). Non-intrusive ROMs use purely data-driven approaches for surrogate modelling, such as GNS (Section 3.4) and RNN-based granular simulations (Section 3.5). To facilitate a comparison of different ML acceleration methods, a summary is listed in Table 2. Note there is always a trade-off between computational efficiency and accuracy, but the combination of AI with physics-based simulations has the potential to improve efficiency while maintaining accuracy.

4. Machine learning enabled pattern recognition and granular physics discovery

Pattern recognition extracts meaningful structures from data to identify, categorize, and understand underlying patterns. The goal of using AI in pattern recognition is to uncover regularities within the data, providing insights that traditional methods alone may not achieve. In granular media applications, pattern recognition includes particle shape and size analysis, 3D reconstruction of granular materials, force-chain network prediction, and characterization of macroscopic behaviour. This section also reviews recent advances in AI applications for discovering granular physics.

4.1. Shape and gradation recognition for granular materials

Particle shape and size significantly influence the mechanical properties of granular materials (Zhang and Yin, 2021). The grain size distribution (GSD) of granular media is typically characterized by cumulative percentages passing through specific sieve sizes (e.g. d_1, d_{10}, …, d_{100}) (Zhang et al., 2021). Particle morphology is typically characterized using a range of global and local descriptors, as shown in Fig. 14. Common terms used to describe the shape of particles are defined below:

(1) **Elongation**: The ratio of the longest length (R_{max}) to the shortest length (R_{min}) from the particle centroid to its edge.
(2) **Slenderness**: The ratio of the major axis length (L_F) to the minor axis length (S_F) of a fitted ellipse of the particle.
(3) **Sphericity** (or **eccentricity**): A measure of the overall shape of a particle, indicating the similarity between the particle's length, height, and width.
(4) **Convexity**: The degree of concaveness of a particle, defined as the ratio of the particle's area to the combined area of the particle and its convex hull (see Fig. 14b).
(5) **Roundness** (or **angularity**): The sharpness of a particle's edges and corners.
(6) **Roughness** (or **smoothness**): A measure of the smallest surface irregularities on a particle.

Instance segmentation in computer vision offers an effective strategy for determining the shape and size of particles (Gao et al., 2024). ML models including U-net (Chow et al., 2022), Mask R-CNN (Lai et al., 2022a), and Inception-ResNet (Kim et al., 2022), are common choices for grain segmentation tasks. In addition to predicting shape and size distribution (Fan et al., 2022), ML models are also trained to automatically analyze grain angularity (Zheng et al., 2022) to enhance our understanding of granular behaviour. To identify particle shapes in 2D RGB images of packed specimens with unavoidable particle occlusion, a self-supervised scene de-occlusion model named PCNet-M (Zhan et al., 2020) was utilized to infer the complete contours of the occluded particles (Zhang et al., 2024a). The particle overlap issue in stacked gravelly sands was also examined using a deformable convolutional network (DCNv2) modified SOLOv2 model (Gong et al., 2025). Instead of using a single image, multi-view camera images are utilized as inputs in a 3D CNN model to predict exact shape descriptors for a particle (Liu et al., 2025a).

Table 1
ML-based drag force model in CFD-DEM simulations.

References	Inputs	Outputs	ML models	Data source
(He and Tafti 2019)	Reynolds number, solid fraction, the relative distance between the target particle and its nearest neighbouring particles	Drag forces	MLP	Resolved CFD-DEM simulations
(Muralidhar et al., 2020)	Reynolds number, solid fraction, neighbouring particle positions	Drag forces	CNN	Resolved CFD-DEM simulations
(Hwang et al., 2021)	129 features including the latent vector from the VAE and a Reynolds number	Drag forces and x, y, z-torque coefficients	VAE and MLP	Resolved CFD-DEM simulations
(Davydzenka and Tahmasebi 2022)	Particle velocities, fluid velocities, void fraction	Drag forces	MLP	Unresolved CFD-DEM simulations
(Ashwin et al., 2022)	Mean flow Reynolds number, solid fraction, the aspect ratio of particles and their orientation to flow direction	Drag forces	MLP	Resolved CFD-DEM simulations
(Ashwin et al., 2022)	Mean flow Reynolds number, solid fraction, the aspect ratio of particles and their orientation to flow direction, along with a 3D image of the distance function for the particle of interest and its immediate neighbours	Drag forces	CNN	Resolved CFD-DEM simulations
(Qu et al., 2022a)	A sequence of matrixes/images showing gas–solid flow state	A parameter to model gas–solid flow	RNN and CNN	Resolved CFD-DEM simulations
(Hwang et al., 2024)	VAE-based latent vectors for describing particle shapes, Reynolds number	Flow fields	VAE and MLP	Resolved CFD-DEM simulations
(Xiang et al., 2024)	Particle shape factor: sphericity, aspect ratio, and diameter ratio; fluid data: Reynolds number	Drag coefficient	MLP and Random Forest	DEM–LBM simulations
(Hu et al., 2024)	Primary energy spectra of SH frequencies, Euler angles, and flow conditions	Drag coefficient	MLP	Resolved CFD-DEM

Note that ML alone cannot provide exact size measurements from images. Only when the physical length per pixel is known, PSD can be predicted with accuracy comparable to manual sieving methods (Bai et al., 2021). A practical approach is to include an object of known physical size, allowing for the calculation of a scaling factor to convert pixels into real-world dimensions (Zhou et al., 2023). A summary of ML-based particle shape and size recognition research in granular media is provided in Table 3.

Data is crucial for accurately recognizing the shape and size of particles in images. Various strategies have been developed to increase the availability of high-quality labelled image data. One key approach is data augmentation, which includes: (1) rotating images by specific degrees, (2) flipping axes, (3) revising brightness, and (4) adding random noise (e.g. Gaussian noise). In addition to using real particle images, synthesized particles from DEM simulations are used to test and validate the segment anything model (SAM) (Kirillov et al., 2023) in recognizing the PSD of granular packings (Xiao et al., 2024). Furthermore, super-resolution techniques using GANs are employed to generate high-resolution images for assessing particle size and shape (Gong et al., 2024a). To simplify labelling efforts, a multi-patch strategy decomposes large images into smaller segments using sliding windows. These cropped small images are then processed by a trained mask R-CNN model before being reassembled into the original large image (Zhang et al., 2024b).

The description of particle shape is inherently a problem with infinite-dimensional function space. Traditional shape descriptors as shown in Fig. 11 are simply approximations and cannot uniquely define a shape. ML offers a new tool to generate shape descriptors. An example is to use VAE and latent diffusion model (LDM) to construct a latent space to represent particle shapes (Macedo et al., 2023, Zhong et al., 2024). Despite different methodologies, both VAE and LDM leverage latent representations for generative tasks. Once trained, these models can sample from the latent space to produce new images that are not present in the training datasets. This concept has also been applied to simulate the interactions of fluids and particles with complex shapes (Hwang et al., 2021) and can be used for the description of microstructures of granular materials. In addition, 3D CNN is used to characterize complex particle shapes directly by assigning different values in a 3D matrix, where points inside the particle are marked as 1 and those outside as −1. The flow and packing behaviour of these particles can then be inferred using 3D CNN (Hesse et al., 2021).

4.2. 3D reconstruction of granular specimens

Reconstructing DEM RVEs with realistic shape and size distributions typically involves using 3D X-ray computed tomography (X-RCT) images of granular media and extracting individual grains via explicit grain segmentation (Kawamoto et al., 2016). This process can be challenging due to diverse shapes of grains, complex microstructures, and imperfections in the images (Faessel and Jeulin, 2010, Liu et al., 2023). ML has emerged as a powerful tool for grain segmentation, with methods such as trainable Weka segmentation (Lai and Chen, 2019), random forest and particle swarm optimization (Zhang et al., 2022a, Li et al., 2024b) delivering promising results. In addition, CNN is used to identify incorrectly segmented grains, thereby expediting the manual inspection process, after image processing based on raw 3D X-RCT images (Cevallos et al., 2023). Recently, vision foundation models (VFMs), such as SAM, have been leveraged to segment 2D X-RCT images for 3D reconstruction purposes (Li et al., 2024a). Here a computational procedure for the VFM-assisted 3D grain reconstruction method is shown in Fig. 15. The reconstructed granular packing offers a digital foundation for conducting DEM simulations with realistic particle shapes (Kawamoto et al., 2018) or microstructures (He et al., 2024).

However, 3D X-RCT scanning is limited by sample size and equipment requirements. Researchers have explored alternative approaches using 2D projections or images to reconstruct the 3D geometrical features of granular media (Santos and Neto, 2023). Examples include 3D laser scanning (Su and Yan, 2018), structural light technique (Sun et al., 2019), dynamic imaging analysis (Wang et al., 2019a) and video tracking (Wang et al., 2023a). Additionally, a dynamic graph edge convolution neural network (DGECNN) was developed to utilize 2.5D point cloud data to characterize the roundness and sphericity of 3D particles (Xi et al., 2025). In recent years, there has been a growing interest in reconstructing 3D representations of objects from single-view or multi-view RGB images using ML. Inspired by the Pix2Vox method

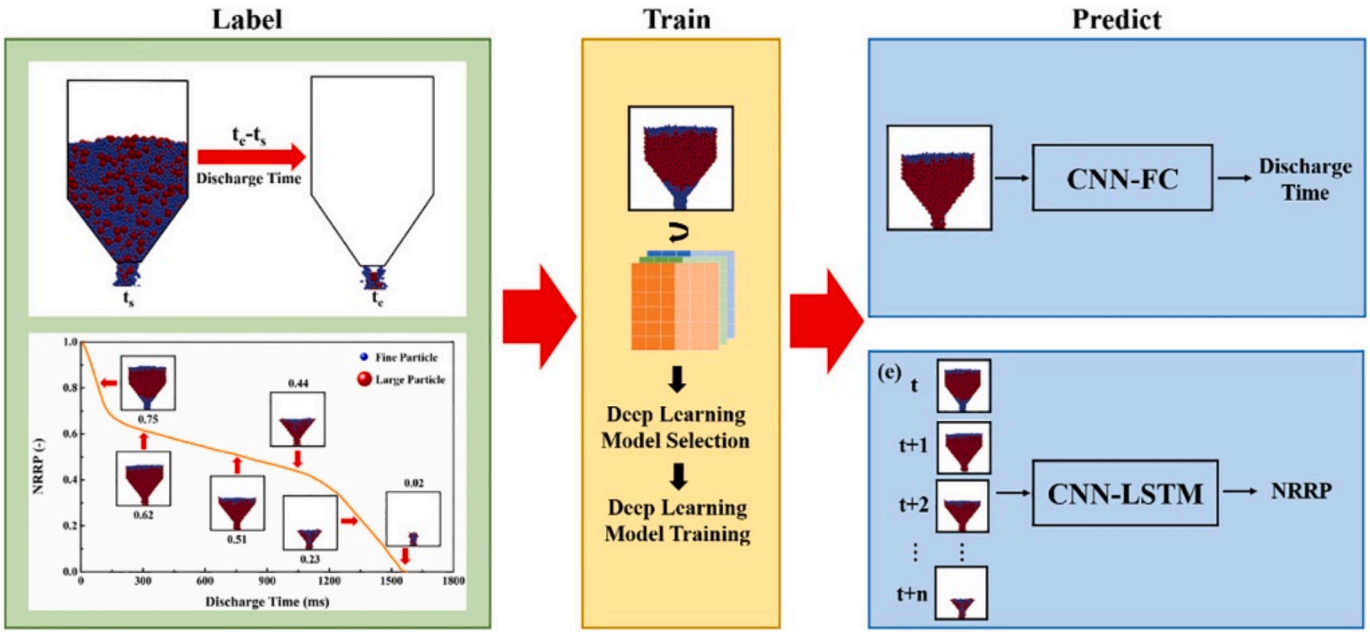


Fig. 13. Image-based prediction of granular flow (Liao et al., 2021).

Table 2
A summary of ML-accelerated granular simulations.

Methods	Role of MLs	Efficiency	Applicability	Notes
DDMM	Replaces online DEM computations in FEM × DEM HMM	50-210x acceleration	Confined granular simulations	Constitutive descriptions of granular material
MLP-based contact resolution in DEM	Contact detection and determination of contact force in DEM	2x (2D) and 5x (3D)	Limited particle shape types	–
CNN-assisted DEM	Contact detection and determination of contact force in DEM	74x (strategy 1) /7x (strategy 2)	Collision-based simulation	Two different strategies reported
GNN for DEM	Replaces the entire computational cycle	10-100x acceleration	Small-to-medium scale granular simulations	A 700x acceleration was reported when using PCA
RNN for DEM	Replaces the entire computational cycle	355x acceleration	Case-specific applications	–
ML-accelerated CFD-DEM	Replaces drag force model	43x acceleration	Fluid-particle interaction	Tested based on CPU time per particle drag force calculation
State-based mapping	Bridges the mappings between two states of granular flows	Several orders of acceleration	Case-specific applications	Efficient, yet limited in scenario extrapolation

(Xie et al., 2019), a 3D CNN with an encoder-decoder framework has been used to infer 3D voxel models from 2D particle images (Giannis et al., 2024). These generated voxel models are then converted into mesh models via the Marching Cubes algorithm. Similar pixels-to-voxels (PVP) models have been applied to reconstruct the 3D morphology of irregular sands (Zhao et al., 2024) and to determine the 3D positions of tracer particles in volumetric particle image velocimetry (Lin and Gao, 2024). However, the 2D-to-3D transformation remains a relatively new area that deserves further exploration.

4.3. *Granular material generation via generative AI*

Grain shape is important, but obtaining versatile, high-quality 3D grain geometry data can be costly. Research has been conducted to leverage generative models to artificially construct new granular specimens with different grain shapes. Generative models are AI models capable of generating new data instances that are similar to but not identical to the original data, by learning the underlying patterns and distribution characteristics from training data. Three types of generative models, including GAN, VAE, and DDPM, are widely used, as depicted in Fig. 16.

GANs typically consist of two neural networks, a generator that creates new data samples, and a discriminator that evaluates them. The goal of the generator is to create imitations of data to trick the discriminator, while the discriminator aims to identify real data from the fakes created by the generator (Fig. 16a). Both are trained simultaneously in a competitive learning process to improve the quality of the generated samples over time. An application of GANs in CGM is that Deep Convolutional GANs (DCGAN) are employed to synthesize X-ray CT images of partially saturated sands (Argilaga, 2023). To falsify CT images with similar factual dimensions as real sands, the pore fractal dimension is computed based on the real CT images with the box count method. The probability distribution metrics of fractal dimension are embedded into the original GAN loss function, called Fractal Informed-GAN (FI-GAN). The developed model can generate representative sand samples, as depicted in Fig. 17, potentially offering a cost-effective method for augmenting CT image data.

Conditional Generative Adversarial Networks (cGANs) extend traditional GANs by providing both the generator and discriminator with additional input information during training. This conditioning allows cGANs to produce more targeted and controlled outputs. Zhao et al., (Zhao et al., 2023c) leverage a cGAN to incorporate shape descriptors as explicit conditions for generation tasks. A feature of their work is the use of the Metaball-Imaging algorithm for dimensionality reduction on the X-RCT data of real particles, which helps reduce training time for the models.

A VAE generates new samples by combining neural networks and variational inference to learn a probabilistic representation of data. In

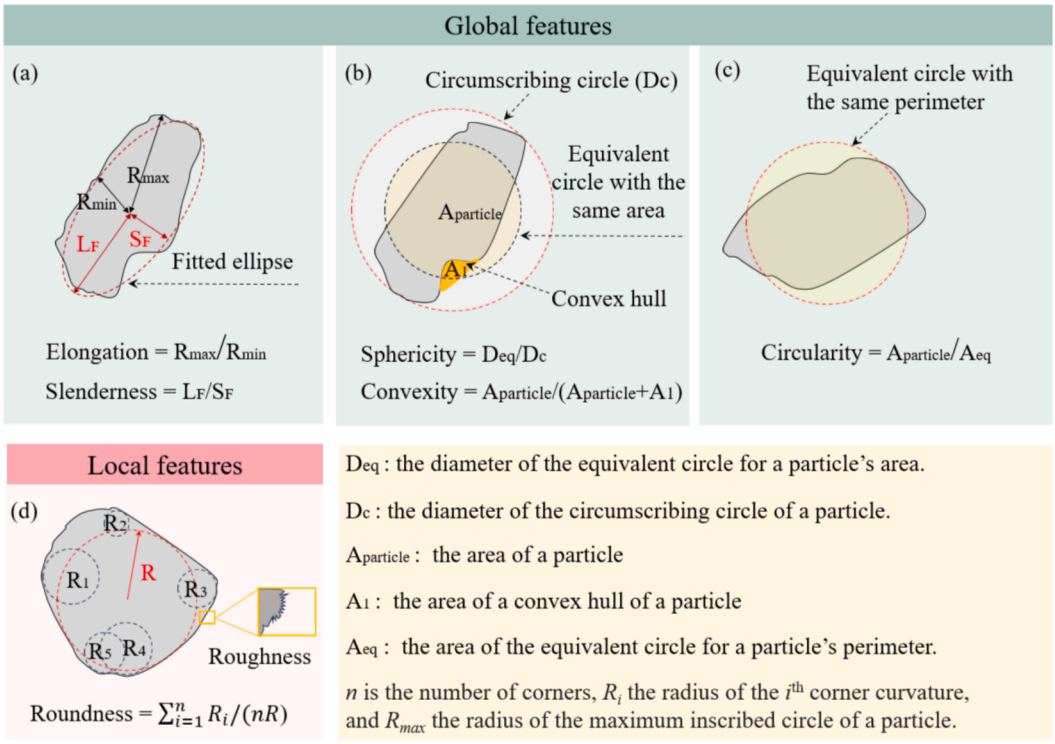


Fig. 14. Typical descriptors for characterizing the morphology of a particle.

Table 3
A summary of ML-based grain shape and size identification.

Reference	ML Models	Datasets	Annotation	Input	Output	Notes
(Zheng and Hryciw 2018)	AdaBoost	85,000 images with a pixel size of 32 × 32	N/A	RGB images	Roundness and sphericity	N/A
(Liang et al., 2019)	U-Net	540 images after data augmentation	N/A	RGB images	Elongation, orientation index, root of form factor	N/A
(Yang et al., 2021)	Mask R-CNN	30 images; resolution: 3024 × 4032 pixels	VIA (Dutta et al., 2016)	RGB images	Elongation, angularity, roughness,	Fine-tuned from Detectron2
(Zhang et al., 2022c)	3D-CNN	2D X-RCT slices for 100 ballast and 100 Fujian sands	VIA	2D binary images	Roundness, sphericity, convexity, aspect ratio	Considering 3D
(Kim et al., 2022)	Inception-ResNet-v2	Over 7000 sand images from six different sands	N/A	Grayscale or Binary	Roundness, sphericity, slenderness, circularity	N/A
(Zhou et al., 2021)	U-Net	150 images of excavated muck from TBM	Manual	RGB images	Elongation, circularity, equivalent diameter	N/A
(Sun et al., 2022)	NSNet: encoder–decoder	430 raw SEM/TEM images; 4300 augmented images; Resolution: 512 × 512 pixels	LabelMe	Grayscale	Blaschke shape factor; size distribution	Nano particles
(Zhang et al., 2023c)	Mask R-CNN	50 raw images, Data augmentation to 1000; Resolution: 2048 × 2048 pixels	LabelMe	RGB images	Equivalent particle size	A shaker-photography system was designed to consider 3D particle shape
(Gong et al., 2024a)	GAN and SOlO v2	800 raw images with a resolution of 4480 × 4480	Manual	RGB images	Size distribution, sphericity, roundness	Super-resolution
(Zhou et al., 2025a)	AlexNet and YOLO-v3	3,800 individual particle images	K-means clustering	RGB images	Particle shape class	N/A

Note: [i] R-CNN: Region Convolutional Neural Network; VIA: VGG Image Annotator; [ii] Detectron2: a platform for object detection, segmentation and other visual recognition tasks. [iii] Blaschke shape factor $= 32A/(\pi P)^2$ (Sun et al., 2022), root of form factor $= 2\sqrt{\pi A}/P$ (Liang et al., 2019), with A and P being the projected area and perimeter of each particle, respectively. The particle orientation index is the angle between the particle's longest axis and the positive *x*-axis, measured clockwise.

contrast to the autoencoder (AE) model, VAE encodes the input into a probabilistic distribution, characterized by a mean and variance. This allows for probabilistic sampling from the latent space and variability in the generated outputs (Fig. 16b). VAE has been used to learn from around 30,000 X-ray micro-computed tomography (μCT) scanned sand particles and generate unseen 3D sand grains (Shi et al., 2021a). Their research found that the generated sand particles exhibit statistical similarities to the natural sand particles. Through the use of VAE, it is found

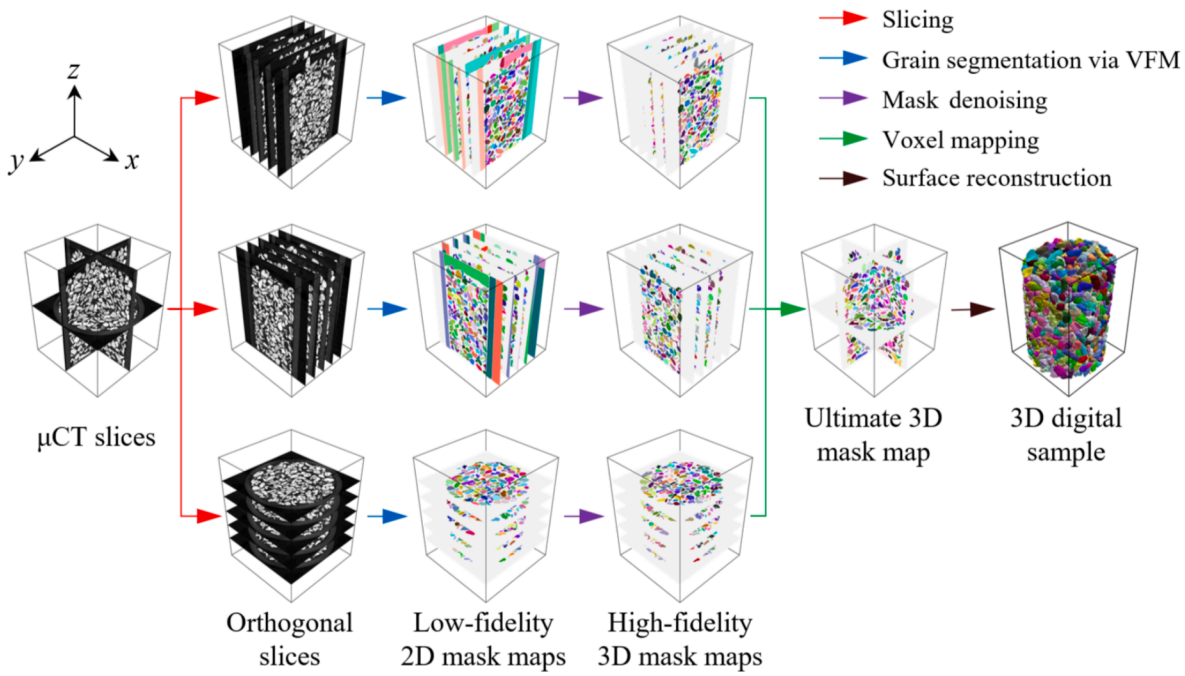


Fig. 15. A computational procedure for VFM-assisted 3D grain reconstruction method. Image adapted from (Li et al., 2024a).

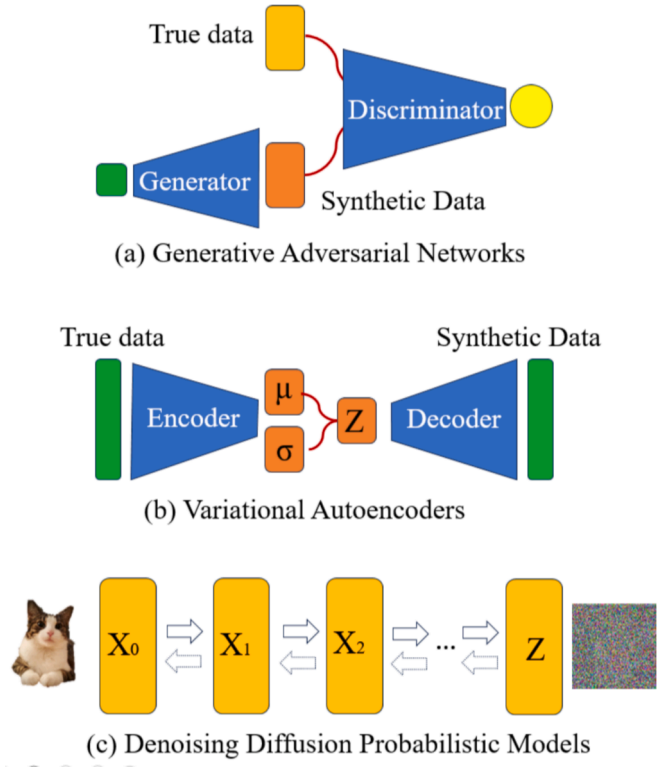


Fig. 16. Three representative generative AI models.

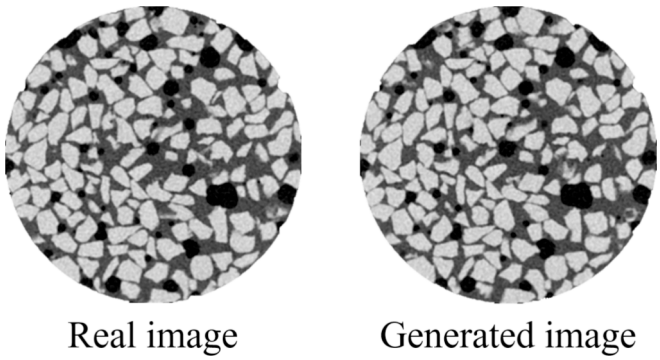


Fig. 17. Reconstructed specimens via generative models (Argilaga 2023).

that 20 descriptors in the latent space can uniquely define the morphology of a particle (Macedo et al., 2023). In addition, a metaball-based VAE is developed to generate realistic sand shapes. The idea is that VAE uses 3D metaball descriptors, instead of a 3D digital matrix of particle shape, as inputs and outputs. The particle shape is computed based on the generated metaball descriptors (Zhao et al., 2023b).

DDPM is a type of generative models that utilize diffusion processes to iteratively remove noise from corrupted data (Fig. 16c), allowing for the generation of high-quality samples that closely resemble the original data distribution (Ho et al., 2020). Recently, DDPMs have been employed to generate particles that have never been observed in reality (Vlassis et al., 2023). As shown in Fig. 18, a real particle is represented as a point cloud $p_i=(x_i, y_i, z_i)$ in three-dimensional Euclidean space, denoted as $\boldsymbol{P}=[p_1, p_2, p_3,\ldots, p_i,\ldots, p_N]$. An autoencoder architecture is utilized to discover a latent representation of the point clouds of sand grains. Subsequently, DDPMs are trained through a diffusion process that progressively adds noise to the lower-dimensional embeddings of these point clouds. This is followed by a reverse diffusion process that constructs lower-dimensional embeddings of the generated sand grains through gradual denoising. The trained decoder then utilizes these generated embeddings to produce synthetic sand that closely resembles the original data distribution. The rationality of generated particle shapes is assessed by comparing the statistical consistency of key particle shape indicators with the original distribution of natural particle assemblies.

4.4. ML-based contact force predictions

Granular media resist external loads through interparticle contacts, forming highly inhomogeneous contact force chains that evolve

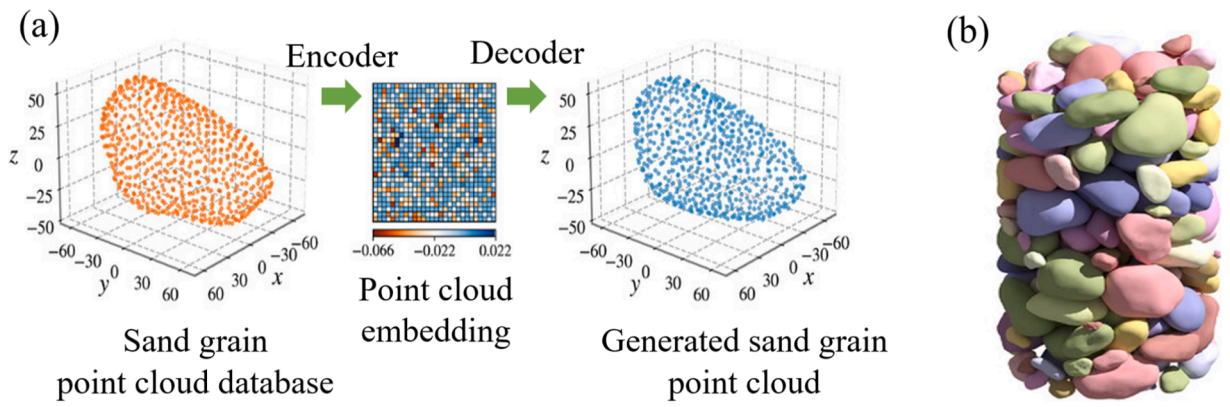


Fig. 18. DDPM-based sand grain generation: (a) process of generating sand grain and (b) an example of generated samples, adapted from (Vlassis et al., 2023).

dynamically within the system (Shi et al., 2021b, Gong et al., 2024b). The evolution of the force-chain network is partially responsible for the complex collective behaviour observed in sheared granular media (Peng and Yin, 2023). In laboratory experiments, particle positions and velocities can be measured, but direct measurement of forces between particles is almost impossible. This becomes an important motivation for using ML to estimate contact force distributions in granular media using kinematic data from μCT (Cheng et al., 2023) and photoelastic experiments (Lin, 2024).

Although MLP is found effective at predicting contact forces (Wu and Wang, 2022b), GNN has gained more attention for modelling contact force chains (Cheng and Wang, 2022). Under specific loading conditions, GNNs can predict the location of force chains in granular materials from the initially undeformed structure (Mandal et al., 2022, Li et al., 2023c). Recent studies indicate that GNNs can scale to predict larger and more complex systems that were not seen during training (Aminimajd et al., 2024). Instead of developing a surrogate model to predict contact force distributions directly, SHapley Additive exPlanations (SHAP) value analysis reveals that the coordination number and local clustering coefficient are keys for forming these force chains (Cheng and Lin, 2024). Furthermore, ML is used to reconstruct forces acting between particles in photoelastic granular material. Due to the scarcity of experimental data, a CNN is first pre-trained on a synthetic dataset derived from analytical solutions of photoelastic theory and then fine-tuned on a smaller experimental dataset, showing satisfactory results (Sergazinov and Kramár, 2021).

4.5. ML-assisted descriptions of material behaviour

The use of AI to predict the macroscopic behaviour of granular media has a long history (Antony et al., 2006). One basic motivation is that some physical behaviours or processes are too complex to be mathematically formulated or numerically simulated, such as the evolution of PSD due to particle attrition in long-term fluidization processes (Farizhandi et al., 2016). ML can serve as a surrogate model to establish the physical relationships between relevant variables based on available experimental data. For example, the mechanical behaviour of granular biomass materials was learned using ML models based on laboratory data from cyclic axial compression and ring-shear tests (Li et al., 2025).

Stiffness and strength are essential properties of granular media. To analyze stiffness characteristics, methods include using 3D CNNs to extract the fabric of granular soil for predicting its constrained modulus (Zhang et al., 2024d) and applying MLPs to forecast the small-strain shear modulus of sand and sand-fines mixtures (Khodkari et al., 2024). Regarding strength behaviour, various ML models are used to predict the uniaxial compressive strength (UCS) of the Voronoi-based universal distinct element code (UDEC) model based on contact-scale cohesion and friction angle (Fathipour-Azar, 2022).

Another strategy for harnessing AI in the description of material behaviour is linking particle-scale physics with free parameters in phenomenological material models. For instance, MLPs are used to capture the correlation between PSDs and two free parameters in the Duncan–Chang hyperbolic model (Gonzalez Tejada and Antolin, 2022), i.e. the tangent elastic modulus and the ultimate deviatoric stress (Duncan and Chang, 1970). Furthermore, building an AI constitutive model to describe how a material deforms under specific external loads is a crucial aspect of this topic. Given the resemblance between data-driven constitutive modelling and multiscale modelling, further details can refer to Section 3.1.

4.6. ML-based pattern recognition from granular dynamics

ML is a useful tool for identifying inherent patterns in granular scenarios. In process engineering, ML is used for grain classification by detecting kinematic patterns to sort different granular media (Laudari et al., 2022), for differentiating small from large grains in the observable dynamics of disordered granular flows (Laudari et al., 2023). In some experimental settings, a dilated convolution UNet++ model is employed to recognize the quasi-static region resulting from sphere impacts on granular beds, replacing the traditionally used Particle Image Velocimetry (PIV) measure technique (Chen et al., 2021).

In addition, Mask R-CNN is utilized to track the position and orientation of dilute rod-like particles across two-view videos. The identified 2D coordinates from both views are converted into 3D trajectories by minimizing reprojection error, enabling automated 3D statistical analysis of particle trajectories with the aid of ML (Puzyrev et al., 2020).

4.7. Discovery of granular physics via machine learning

Many physical laws are hidden in complex patterns and environments. These underlying physics are typically hard to explain, describe, and predict. Some efforts have been made to employ ML to gain granular insights that might not have been possible with traditional methods alone. Discovering ordinary differential equations (ODEs) and partial differential equations (PDEs) to describe physical systems and processes is a fundamental task in physics. Representative data-driven methods for finding governing equations include: (i) sparse regression, (ii) symbolic regression, and (iii) partial dependence analysis.

Sparse regression is a technique that uses regularization to identify a subset of features and reduce model complexity in regression analysis. A notable example is using sparse regression to identify the governing equation of granular temperature during homogeneous cooling (Zhao et al., 2023a). In this case, domain expertise is leveraged to choose a candidate variable library associated with granular temperature, followed by linear regression to model its evolution using DEM simulation data. In contrast, symbolic regression is an ML approach that uses

algorithms like genetic programming to uncover underlying mathematical relationships in data. This technique has been used to discover elastoplasticity constitutive models in conjunction with the neural polynomial method (Bahmani et al., 2024). In addition to sparse and symbolic regression, partial dependence analysis can reveal how a specific feature influences models' predictions while keeping other features unchanged (Parr and Wilson, 2021). This technique has been employed to develop explicit permeability formulas in granular materials, resulting in a permeability prediction equation that demonstrates robust applicability across various datasets (Zhang et al., 2024c).

When the relationships among physical variables are too complex to be mathematically formulated, using a data-driven surrogate model to capture these relationships is a viable option. For example, Gradient Boosting Decision Trees (GBDT) can map short- and medium-range structural features to identify crystalline phase precursors in monodisperse packings (Zhang et al., 2022e). Extreme Gradient Boosting (XGBoost) links micro slips (indicated by squared nonaffine displacements) to slip avalanches (reflected in macroscopic stress fluctuations) in sheared granular gouges (Ma et al., 2022b). A 3D CNN connects voxel matrix representations of microscopic dynamics to macroscopic stress fluctuations (Mei et al., 2023). Graph Convolutional Neural Networks (GCNNs) serve to identify particles with high plasticity susceptibility based on solely initial particle positions (Mei et al., 2024). GNNs can capture particle crushing strength using connected fragment graphs instead of traditional morphology descriptors (Zheng et al., 2024) and examine the relationship between microscopic plasticity and macroscopic stress (Mei et al., 2022). A causal link is established between particle-level structural indicators (such as coordination number and local packing fraction) and the plastic instability of shearing granular materials (Zou et al., 2023). CNNs are employed to identify the structural characteristics of granular materials (Cui et al., 2023). MLPs explore shock compaction-induced heating mechanisms in energetic granular media by linking local density to local temperature and intramolecular strain energy (Hamilton and Germann, 2023). In addition, MLP and random forest are leveraged to predict the number and sizes of crushed ballast particles (Aela et al., 2022).

ML has been used to investigate the heterogeneous dynamics of disordered granular flows (Cubuk et al., 2015, Cubuk et al., 2017). A random forest classifier reveals that force and acceleration in collisional zones of bidisperse silo flows can differentiate small from large grains, while other indicators like velocity, are ineffective, particularly in dense zones of the silo flow (Laudari et al., 2023). Although the predictions are not entirely satisfactory, ML has been reported to uncover meaningful physics underlying the clogging process in 2D granular hoppers (Hanlan et al., 2024). Beyond solid mechanics, ML is applied to correlate pore networks, particle contact networks, and permeability in porous granular media (Yasuda et al., 2021).

ML has also been used to explore geological physics associated with granular media (Rouet-Leduc et al., 2017, Jaza et al., 2021). For instance, an XGBoost model predicts instantaneous global friction coefficients for sheared granular systems, which are designed to simulate the frictional behaviour of geological faults (Ren et al., 2019). Through tentatively adding particle-scale information to ML models, researchers found that velocity signals from individual particles contain details about intermittent frictional stick–slip dynamics. Similarly, the Light Gradient Boosting Machine (LightGBM) analysis revealed that plate motion signals in a sheared granular fault during the initial slip stage contain precursor information about the duration of slip in laboratory earthquakes (Wei and Gao, 2024). Notably, the knowledge gained from particle-scale simulation data can be transferred to predictions of laboratory fault friction via transfer learning (Wang et al., 2021a).

Based on well-trained ML models, SHAP values are widely used to explain individual predictions by computing the contribution of each feature. The application of SHAP analysis can be found in examining the plastic instability of disordered granular media, identifying important particle shape descriptors related to crushing strength (Wang et al., 2021b), discovering key variables associated with the shear modulus of gap-graded granular mixtures (Liu et al., 2024), identifying crucial variables affecting hydraulic fracture behaviour in conglomerate rock (Shentu et al., 2024), and exploring the structure–property relationship of granular materials (Zhang et al., 2022f). However, SHAP analysis relies on specific models and data. Using identical datasets with different models, or identical models with different datasets, can yield different explanations. This variability is particularly pronounced when the data does not cover all possible scenarios or contains significant noise, potentially leading to biased results. In addition, while SHAP analysis provides insights into the contribution of each variable, it does not fully explain the underlying physics.

5. Machine learning for inverse problems in computational granular mechanics

Inverse problems involve using observed data or outcomes of a system to deduce underlying causes or parameters. Solving inverse problems is essential in engineering and science applications, but it can be challenging due to the possibility of non-unique or infinite solutions. Typical inverse problems in granular mechanics include parameter calibration in DEM and simulation-based optimization tasks. In contrast to conventional methods, ML brings new solutions for tackling these longstanding inverse problems.

5.1. Parameter calibration in DEM

Realistic simulations require accurate input parameters. However, particle-scale parameters are often difficult to measure directly for most granular materials. In addition, DEM models simplify the complexities of real physical systems, including shape and size distributions of grains and contact behaviour. Thus the fundamental philosophy of parameter calibration is to (1) acknowledge the simplifications made in DEM and (2) adjust particle-scale parameters to match physical observations as closely as possible (Qu et al., 2020a).

5.1.1. Supervised machine learning for parameter calibration

ML is often used to build a macro-microscopic parameter relation (Benvenuti et al., 2016). Basically, there are two main strategies. One is the direct method which uses macroscopic parameters as input while particle-scale parameters as outputs (Zhou et al., 2022, Wang et al., 2023b, Liu et al., 2025b). The other one is the surrogate inversion method which builds an ML-based surrogate model to evaluate macro parameters based on micro properties (Shentu and Lin, 2023) and then leverages optimization methods, such as orthogonal search algorithm (Ye et al., 2019) and genetic algorithm (GA) (Gu et al., 2023), to search for satisfactory particle-scale parameters to reproduce the target macro properties.

In terms of the direct method, multi-fidelity residual neural network (MFRNN) utilizes existing empirical or analytical macro–micro relations as low-fidelity data while DEM simulations as high-fidelity data to reduce the costs of training a reliable parameter calibration model in bonded granular materials (Zhou et al., 2024). A potential risk is that there might be many particle-scale parameter combinations that will yield similar macroscopic results. Thus, the direct prediction from macroscopic properties to estimate potential microscopic parameters might suffer from problems. To mitigate this issue, Long et al., (Long et al., 2023) leveraged the entire stress–strain sequences as inputs to estimate particle-scale parameters, based on the consideration that the macroscopic stress–strain sequences resulting from particle-scale DEM simulations are highly associated with the parameters used. The results show this scheme can yield satisfactory interpolated results as shown in Fig. 19.

In the surrogate inversion method, ML serves to replace time-consuming DEM simulations with an efficient data-driven forward simulator (Irazábal et al., 2023), thereby accelerating the calibration

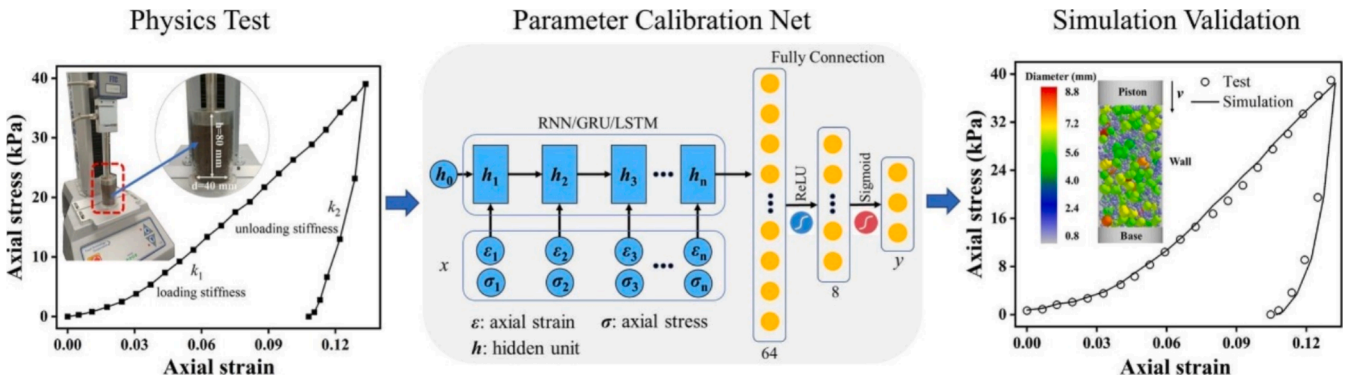


Fig. 19. A RNN approach for particle-scale parameter calibration (Long et al., 2023).

process. This approach typically requires a certain number of iterative simulations (Qu et al., 2020b). In contrast to the direct method, the surrogate inversion method is found to be more accurate (Pan et al., 2023).

Supervised learning algorithms, including the direct method and surrogate inversion method, face challenges in incorporating particle size, shape distributions, and microstructures of granular systems, factors that impact the macroscopic behaviour of granular media. As the number of parameters to calibrate grows, the parameter space expands significantly, necessitating numerous DEM simulations for training data generation. Consequently, despite showcasing strong calibration performance, supervised learning methods frequently encounter difficulties in generalizing scenarios with diverse packings.

5.1.2. Reinforcement learning (RL)

RL is a branch of ML in which an agent learns to make optimal decisions by interacting with an environment. This approach is inspired by how humans naturally acquire knowledge through interacting with their surroundings. RL consists of four basic components: (1) agent, (2) environment, (3) state, and (4) action. Through numerous interactions with its environment, RL enables virtual agents to determine optimal actions for any given state, making it particularly effective for combinatorial optimization problems. A reinforcement learning framework is proposed to calibrate particle-scale parameters (Westbrink et al., 2021), as shown in Fig. 20. Calibration is conducted based on both the static angle of repose (AoR) using lifting cylinder tests and the dynamic AoR with rotating drum tests. The agent's states are defined by the particle-scale parameters that require calibration, while actions involve adjusting these material parameters to maximize rewards. Rewards are determined by the difference between experimentally measured macroscopic properties and their simulated counterparts in DEM. The closer the numerical results are to the experimental targets, the higher the reward.

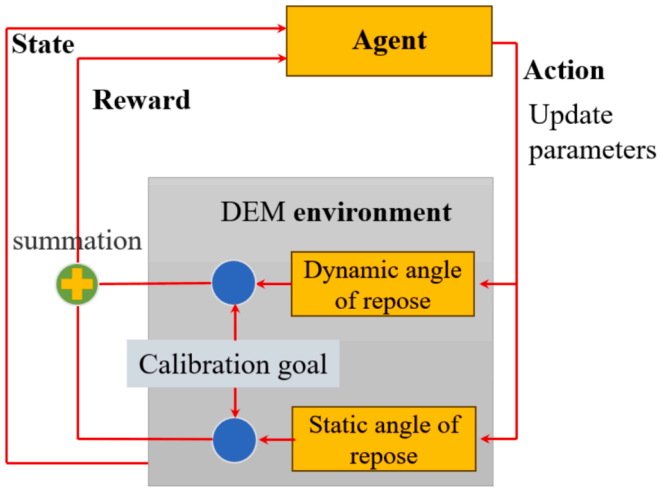


Fig. 20. Structure of Actor-Critic RL algorithm for parameter calibration in DEM, adapted from (Westbrink et al., 2021).

5.2. Optimisation problems in granular mechanics

Optimisation is a key focus in industry and engineering. The goal of optimization is to find the best solutions from a set of feasible solutions to maximize or minimize an objective function under specified constraints. ML-based surrogate models excel at solving optimisation problems for several reasons: (1) some surrogate models are differentiable so that gradient-based optimisation methods can offer iterative directions for a prescribed goal. (2) ML surrogate models provide high computational efficiency for simulating granular processes, allowing for numerous iterative computations in a reasonable time. (3) ML can uncover low-dimensional latent representations of high-dimensional parameter spaces, thereby lowering the cost of finding feasible solutions. Overall, the potential of ML to optimize various problems related to granular materials is being actively researched.

5.2.1. Gradient-based optimization

The primary technique for training deep learning models largely relies on automatic differentiation (AD) and gradient-based optimisations. Through establishing a differentiable surrogate model, the gradient information can be leveraged for optimization (Allen et al., 2022). Recent advancements include the GNS-based inverse design method for granular flows. Automatic inverse design is an optimization problem that involves defining an objective function (L_θ) and iteratively adjusting the parameters $\boldsymbol{\theta}$ to be optimized. Gradient-based iteration is a common choice in optimization and there are two main strategies for computing the gradient of the objective function against the parameters being optimized: one uses perturbation or the finite difference method to approximate the gradient of the objective function L_θ with respect to the parameters $\boldsymbol{\theta}$, i.e. $\nabla_\theta L_\theta$ (Jiang et al., 2024). For example, if suppose $\boldsymbol{\theta}=(\theta_1,\theta_2)$, the gradient $\nabla_\theta L_\theta$ can be approximated using Eq. (14).

$$\nabla_\theta L_\theta \approx \left(\frac{L_\theta(\theta_1+\epsilon_1,\theta_2)-L_\theta(\theta_1,\theta_2)}{\epsilon_1}, \frac{L_\theta(\theta_1,\theta_2+\epsilon_2)-L_\theta(\theta_1,\theta_2)}{\epsilon_2}\right) \tag{14}$$

where θ_1 and θ_2 represent two parameters to be optimized (more parameters can be treated in the same way); ϵ_1 and ϵ_2 are small numbers representing a perturbation to θ_1 and θ_2, respectively.

The alternative strategy is to leverage the AD of a GNS (Kumar and Choi, 2023):

$$\nabla_\theta L_\theta = \frac{\partial L_\theta}{\partial \mathbf{S}^k}\frac{\partial \mathbf{S}^k}{\partial \theta} \tag{15}$$

where $\mathbf{S}^k$ represents the GNS state at the k^{th} time step. With $\mathbf{S}_0$ being the initial state, the following states can be computed via:

$$\mathbf{S}^1 = \mathrm{GNS}(\mathbf{S}^0;\theta)$$

$$\mathbf{S}^2 = \mathrm{GNS}(\mathbf{S}^1;\theta)$$

…

$$\mathbf{S}^k = \mathrm{GNS}(\mathbf{S}^{k-1};\theta) \tag{16}$$

In addition, $\frac{\partial \mathbf{S}^{k-1}}{\partial \theta}$ can be computed recursively by applying the chain rule to the entire simulation trajectory:

$$\frac{\partial \mathbf{S}^k}{\partial \theta} = \frac{\partial \mathbf{S}^k}{\partial \mathbf{S}^{k-1}} \cdots \frac{\partial \mathbf{S}^1}{\partial \mathbf{S}^0} \frac{\partial \mathbf{S}^0}{\partial \theta} \tag{17}$$

Eqs. (15) and (17) indicate that accurately calculating ∇L_θ requires propagating gradients from the final state $\mathbf{S}^k$ back to the initial state $\mathbf{S}^0$, utilizing all intermediate states $\mathbf{S}^1, \mathbf{S}^2, \cdots, \mathbf{S}^{k-1}$. Since the GNS consists of multiple MLPs with numerous parameters, storing all intermediate states for long trajectories can be memory-intensive. Research showed that GNS simulations of granular flows with approximately 3,000 particles can only handle 3 to 4 timesteps on GPUs with 40 GB of memory (Choi and Kumar, 2024b). However, they found that the gradient checkpointing technique (Chen et al., 2016) can reduce memory consumption effectively and enable the computation of a 3D granular flow problem over hundreds of timesteps.

With gradient information available, gradient-based optimization methods can iteratively minimize the difference between the GNS simulation results and the specified goals. For example, in the gradient descent method, the parameter θ can be updated as follows:

$$\theta := \theta - \eta \nabla_\theta L_\theta \tag{18}$$

where η is a hyperparameter called learning rate. Other first-order gradient-based optimisation methods, such as AdaGrad (Duchi et al., 2011) and Adam (Kingma and Ba 2014), can be used seamlessly for the optimisation purpose. An example of the GNS-based optimization framework can be found in Fig. 21.

5.2.2. Gradient-free optimisation

In the absence of gradient information, ML can provide an efficient surrogate simulator for gradient-free optimization, which typically requires many simulations. One such example is to control the final packing shape of granular materials by optimising the accelerations of a blade in rigid body-granular media interactions. A GNN-based surrogate simulator is integrated with Differential Dynamic Programming (DDP) to optimize a rigid body-driven granular system by iteratively refining control policies (Aoyama et al., 2024). Another gradient-free optimisation problem is searching which kinds of particle shapes can pack the densest granular matter in a certain specimen-making procedure. The solution is conducted in the following steps: First, a vector $\boldsymbol{\alpha}_\mathbf{p}$ is defined to represent the particle shapes. Next, a numerical or ML model is used to determine the packing density ϕ_p for a given $\boldsymbol{\alpha}_\mathbf{p}$, following a specific sampling-making protocol. Finally, gradient-free methods, such as NelderMead, Simulated Annealing (SA), and Differential Evolution (DE) can be used to minimize $\phi_p(\boldsymbol{\alpha}_\mathbf{p})$ (Baule et al., 2023).

$$\boldsymbol{\alpha}_p^* = \mathrm{argmin}\, \phi_p(\boldsymbol{\alpha}_p) \tag{19}$$

5.2.3. Dimension reduction in optimisation

Traditional optimization methods struggle with high-dimensional parameter spaces and implicit features. Taking Eq. (21) as an example, in the case that the combinatorial space of $\boldsymbol{\alpha}_\mathbf{p}$ is too large, PCA can be used to reduce the dimensionality of the parameter space, thereby making the optimization more tractable. One more example is that in the inverse design of microstructure for Li-Ion Batteries (Xu et al., 2021a, Ge and Liu, 2024), the features to be optimised are not well-defined. Autoencoders offer a novel solution by transforming intricate inputs like microstructure images into optimizable latent variables, which can then be decoded back into optimal designs. Then conventional optimization methods, such as Bayesian optimization, can search for optimal solutions in the low-dimensional latent space (Jung et al., 2020).

5.3. Physics-informed neural network (PINN) for granular mechanics

PINNs integrate PDEs or ODEs into the loss function of neural networks as soft constraints (Raissi et al., 2019). During training, the network minimizes the residuals of these PDEs or ODEs using automatic differentiation. The differentiability of PINNs enables their use for solving inverse problems via backpropagation. For instance, these methods can infer unknown parameters that best fit the observed data by treating the parameters as trainable variables adjusted during optimization. PDEs have been employed to characterize the evolution of granular systems. For example, population balance equations (PBEs) are often used to describe particle-based processes like particle breakage analysis (Gupta and Mishra, 2024) and powder mixing in chemical engineering. PINNs are used to solve PBEs and identify unknown parameters such as aggregation and breakage rate constants (Chen et al., 2021), demonstrating the potential of PINNs in solving forward and inverse PDE problems. However, the identified parameters may not be unique and PDEs are often unavailable or oversimplified for many granular behaviours.

PINNs represent a special type of unsupervised learning, where the loss function is formulated using analytical equations, rather than relying solely on labelled data. A similar idea is being increasingly integrated into the solution process of FEM (Wang et al., 2024c), where an NN is used to predict nodal displacements by using nodal coordinates as inputs. The loss function can be formulated using energy functional (Wang et al., 2025c) and boundary/initial conditions (Zhang et al., 2025a). Together with other standard FEM computational procedures,

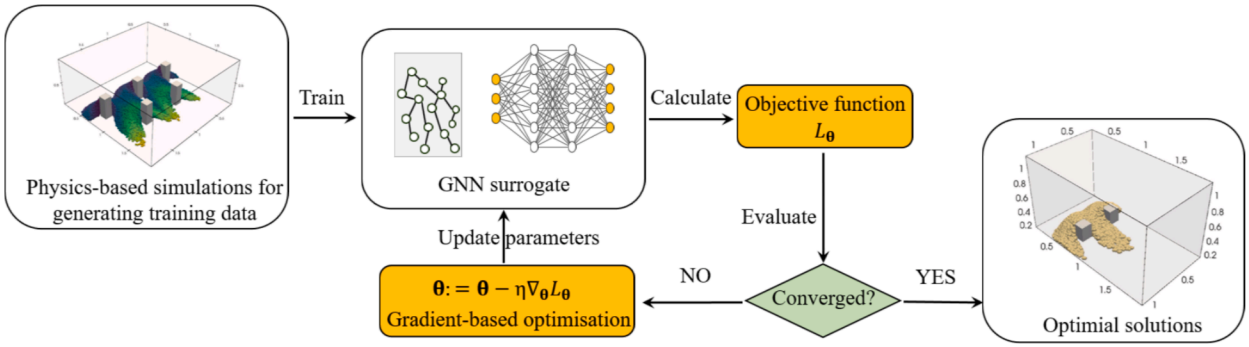


Fig. 21. A GNS-based optimization simulation procedure (adapted from (Choi and Kumar 2024b, Jiang et al., 2024)).

the NN can be trained to minimize the global residuals, yielding nodal predictions that satisfy governing equations in FEM. Alternatively, optimizers in NN training, e.g. Adam, and LBFGS, have been explored to directly minimize the global residual of the governing equations after Galerkin discretization in FEM (Wang et al., 2025b). These attempts enable FEM models to be differentiable. This, in turn, provides greater flexibility in tackling inverse problems.

6. Discussion

6.1. Challenge in data

The effectiveness of data-driven models is inherently limited by the quality and representativeness of the data on which they are built. To extract unbiased insights, it is essential to ensure representative sampling; however, sampled datasets often fail to cover the full range of scenarios the model is intended to address. This challenge is complicated by the fact that data is often collected independently, limiting ML experts' control over the sampling process. To minimize sampling biases, data collectors should understand the problem context, define the model's intended scope, and ensure adequate sampling coverage.

When active sampling is infeasible, particularly in laboratory or engineering scenarios, synthetic data generation methods can be used to produce extensive datasets that are difficult or costly to obtain through physical monitoring. These methods include high-fidelity physics-based simulators (e.g. DEM) (Zhang and Yin, 2021), analytical formulation-based models (Zhang et al., 2020), and even Generative AI techniques (Argilaga 2023, Vlassis et al., 2023). Note that the lack of data under extreme conditions is a key challenge that affects many AI applications. In granular scenarios under extreme conditions, such as high strain rates and particle breakage, numerical simulations can provide a critical source of data for mitigating data scarcity (Zheng et al., 2024). These synthetic datasets can serve as low-fidelity data, to alleviate data scarcity by leveraging similarities between related tasks through cross-task learning techniques such as transfer learning (Qu et al., 2023b) or multi-fidelity learning (Zhang et al., 2022b).

However, establishing explicit guidelines or standards for data synthesis is essential to ensure the reproducibility of the data generation process. For instance, parameter calibration in numerical and analytical models is the prerequisite to minimize the discrepancy between synthetic and physical data. Key information such as model parameters, simulation settings, random seeds, and software or code details should be documented as metadata to ensure the transparency and reproducibility of computational simulations.

Recognizing the importance of data and the challenges associated with its collection and management, we have launched a data management platform called "Clear Data Bay" (https://www.cleardatabay.com). This platform seeks to integrate all available data resources related to granular media and its associated industrial and engineering fields, including experimental, analytical, and numerical simulations. It aims to promote efficient data circulation and maximize the value and usage of data within the community. Unlike most data management systems, "Clear Data Bay" is a lightweight, centralized platform that features a substantial amount of metadata, providing information on existing open-source raw datasets and their corresponding source addresses.

6.2. Challenge in models

A significant challenge for current ML models is their reliability. Developing explainable and trustworthy ML models relies not only on comprehensive data coverage but also on advancements in computational algorithms. Several key areas are suggested below:

(1) **Uncertainty Quantification**. As data-driven models are typically not underpinned by a solid theoretical foundation, an AI model that can assess its uncertainty or confidence in each prediction can effectively enhance trustworthiness, particularly for out-of-distribution (OOD) predictions. Furthermore, the use of uncertainty estimators allows for an interactive AI training strategy, where predictions with high uncertainty can trigger specified needs for numerical computations or laboratory experiments to further enrich the training dataset (Qu et al., 2023a).

(2) **Physics-involved Data-driven Models**. Well-established physical knowledge, such as thermodynamics and frame-indifference, should be explored for integration into the training process of models, such as Physics-informed loss function (soft constraints), Physics-guided architecture (hard constraints) and even Physics-based parameter initialization. The development of principled physics-involved ML models could facilitate the development of high-quality models using limited data.

(3) **Explainable ML models**. AI models are often criticized as "black box" due to the lack of clarity on how outputs are generated (Chan et al., 2022). In certain cases, interpretability may not be necessary for applications like generative structural design, because domain experts can leverage mechanics-based models or knowledge to assess the quality of ML outputs. However, for AI-informed decisions that directly or immediately affect high-consequence actions, research in interpretability can effectively enhance confidence in the predictions made by AI models. Most current interpretability research in CGM relies on post hoc explanations, such as global techniques (e.g. feature importance), and local methods (e.g. SHAP). Further development in explainable ML models will enhance the reliability and trustworthiness of AI in accelerating granular simulations, knowledge discovery and other scientific applications in CGM.

6.3. Prospects in AI-driven CGM

AI is transforming numerous fields, including computational granular materials. While ML is a potent fitting tool, it often struggles to make accurate predictions for underrepresented scenarios during the training phase. The true value of AI will become evident when it successfully addresses previously unresolved challenges in the CGM community. For instance, ML can be used to create scalable simulators that model physical objects over larger lengths and time scales that are inaccessible to traditional physics-based computational tools or develop highly efficient data-centric virtual twins that instantly interact with physical assets within the digital twin framework.

High-fidelity simulations remain essential in the age of AI. On the one hand, Physics-based simulators can generate abundant data efficiently at a low cost for training AI models. On the other hand, AI models can be integrated into physics-based numerical simulators, acting as a type of non-intrusive reduced-order model to accelerate computations or empower differentiability by modifying certain components in the original computational framework. Combining physics-based simulations and AI techniques offers a complementary approach that can leverage the strengths of both to advance the development of CGM.

While AI has been extensively explored in various aspects of CGM, the potential of generative AI to drive innovations in this field has not yet been fully realized. There is a growing trend towards fine-tuning, general-purpose LLMs for specialized usage. However, integrating diverse LLM agents into targeted application scenarios is equally crucial. Current standard uses of LLMs include generating codes (Kim et al., 2025), debugging, brainstorming, and refining writing. Yet, further exploration of AI-human collaboration across broader scenarios is not just a trend but necessary. Additionally, despite some progress, the application of multi-modal LLMs or LVMs in CGM is still in its early stages. It can be expected that the integration of these generative AI techniques could unlock more frontiers in CGM.

7. Summary

This review summarizes the state-of-the-art machine learning applications in modelling granular materials and related industrial/

engineering problems. The capabilities of AI in transforming computational granular media can be categorized into three main areas:

(1) **AI-accelerated simulations**: This includes multiscale and constitutive modelling of granular soils, MLP-based contact laws in DEM, CNN-based intrusive reduced-order modelling of DEM, RNN and GNN-based non-intrusive reduced-order modelling of DEM, and ML-assisted CFD-DEM modelling. While AI cannot replace CGM, it serves as an important complementary tool that enhances CGM. AI-driven CGM is expected to provide better accuracy-speed trade-offs compared to conventional, physics-based granular simulators.
(2) **AI-enabled pattern recognition and physics discovery**: This encompasses the recognition of shape and gradation in granular media, 3D reconstruction of granular specimens, automatic generation of granular specimens, predictions of contact force network, characterization of macroscopic property, and data-driven physics discovery. AI can enhance human capabilities by identifying complex patterns that might otherwise go unnoticed, leading to more effective simulation-assisted knowledge discovery and improved representation of granular materials in simulations.
(3) **AI-assisted inverse modelling**. This includes automatic calibration of particle-scale parameters, scenario-specific optimization of simulations, and the use of PINNs for granular modelling. Differentiable and lightweight AI-based surrogate simulators offer new solutions for addressing long-standing inverse problems in CGM.

To further promote the development of data-driven CGM, a specialized data platform featuring comprehensive metadata has been established. This platform aims to accelerate the circulation of CGM data and unlock its value for advancing the field. Although AI has been applied in various granular simulations, numerous challenges remain in terms of data and model/algorithm development. It is important to note that AI is not a panacea; AI systems often struggle to generalize beyond their training data, and produce poorly interpretable predictions. However, as AI continues to evolve and more principled methods that integrate domain knowledge with AI models are developed, AI-driven CGM approaches hold promises for improving the efficiency and accuracy of granular simulations, while extracting deeper insights from the simulation results.

CRediT authorship contribution statement

Tongming Qu: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization. **Jidong Zhao:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Y.T. Feng:** Writing – review & editing, Supervision, Project administration, Conceptualization.

Declaration of competing interest

The authors 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 study was financially supported by Research Grants Council of Hong Kong (under GRF 16208224, GRF 16203123, and TBS T22-607/24-N) and National Natural Science Foundation of China (NSFC) (52439001).

Data availability

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

References

Aela, P., Wang, J., Yousefian, K., Fu, H., Yin, Z.-Y., Jing, G., 2022. Prediction of crushed numbers and sizes of ballast particles after breakage using machine learning techniques. Constr. Build. Mater. 337, 127469.
Allen, K.R., Lopez-Guevara, T., Stachenfeld, K., Sanchez-Gonzalez, A., Battaglia, P., Hamrick, J., Pfaff, T., 2022. Physical design using differentiable learned simulators. arXiv preprint arXiv:2202.00728.
Aminimajd, A., Maia, J., Singh, A., 2024. Scalability of graph neural network in accurate prediction of force chain network in suspensions. arXiv preprint arXiv:2409.13160.
Antony, S., Zhou, C., Wang, X., 2006. An integrated mechanistic-neural network modelling for granular systems. Appl. Math. Modell. 30 (1), 116–128.
Aoyama, Y., Haeri, A., Theodorou, E.A., 2024. Optimal control of granular material. 2024 IEEE International Conference on Robotics and Automation (ICRA), IEEE.
Argilaga, A., 2023. Fractal Informed Generative Adversarial Networks (FI-GAN): Application to the generation of X-ray CT images of a self-similar partially saturated sand. Computers and Geotechnics 158, 105384.
Arifuzzaman, S., Dong, K., Yu, A., 2022. Process model of vibrating screen based on DEM and physics-informed machine learning. Powder Technol. 410, 117869.
Ashwin, N.R., Cao, Z., Muralidhar, N., Tafti, D., Karpatne, A., 2022. Deep learning methods for predicting fluid forces in dense particle suspensions. Powder Technol. 401, 117303.
Bahmani, B., Suh, H.S., Sun, W., 2024. Discovering interpretable elastoplasticity models via the neural polynomial method enabled symbolic regressions. Comput. Methods in Appl. Mech. Eng. 422, 116827.
Bai, F., Fan, M., Yang, H., Dong, L., 2021. Image segmentation method for coal particle size distribution analysis. Particuology 56, 163–170.
Baule, A., Kurban, E., Liu, K., Makse, H.A., 2023. Machine learning approaches for the optimization of packing densities in granular matter. Soft Matter 19 (36), 6875–6884.
Benvenuti, L., Kloss, C., Pirker, S., 2016. Identification of DEM simulation parameters by Artificial Neural Networks and bulk experiments. Powder Technol. 291, 456–465.
Cevallos, S.B., Jerves, A.X., Mital, U., Medina, D.A., Quinteros, V.S., Mulas, M., Torgersrud, Ø., 2023. Towards a more accurate characterization of granular media 2.0: Involving AI in the process. Comput. Geotech. 160, 105510.
Chan, K.H.R., Yu, Y., You, C., Qi, H., Wright, J., Ma, Y., 2022. ReduNet: A white-box deep network from the principle of maximizing rate reduction. J. Mach. Learn. Res. 23 (114), 1–103.
Chen, B., Xiang, J., Latham, J.-P., Bakker, R.R., 2020. Grain-scale failure mechanism of porous sandstone: an experimental and numerical FDEM study of the Brazilian Tensile Strength test using CT-Scan microstructure. Int. J. Rock Mech. Mining Sci. 132, 104348.
Chen, G., Yamashita, H., Ruan, Y., Jayakumar, P., Gorsich, D., Knap, J., Leiter, K.W., Yang, X., Sugiyama, H., 2023. Hierarchical MPM-ANN multiscale terrain model for high-fidelity off-road mobility simulations: a coupled MBD-FE-MPM-ANN approach. J. Comput. Nonlinear Dynamics 18 (7), 071001.
Chen, T., Xu, B., Zhang, C., Guestrin, C., 2016. Training deep nets with sublinear memory cost. arXiv preprint arXiv:1604.06174.
Chen, W., Qiu, T., 2012. Numerical simulations for large deformation of granular materials using smoothed particle hydrodynamics method. Int. J. Geomech. 12 (2), 127–135.
Chen, X., Wang, L.G., Meng, F., Luo, Z.-H., 2021. Physics-informed deep learning for modelling particle aggregation and breakage processes. Chem. Eng. J. 426, 131220.
Cheng, C.-H., Lin, C.-C., 2024. Prediction of force chains for dense granular flows using machine learning approach. Phys. Fluids 36 (8).
Cheng, Z., Wang, J., 2022. Estimation of contact forces of granular materials under uniaxial compression based on a machine learning model. Granul. Matter 24, 1–14.
Cheng, Z., Wang, J., Xiong, W., 2023. A machine learning-based strategy for experimentally estimating force chains of granular materials using X-ray micro-tomography. Géotechnique 1–13.
Cho, K., 2014. Learning phrase representations using RNN encoder-decoder for statistical machine translation. arXiv preprint arXiv:1406.1078(2014).
Choi, Y., Kumar, K., 2024a. Graph neural network-based surrogate model for granular flows. Comput. Geotech. 166, 106015.
Choi, Y., Kumar, K., 2024b. Inverse analysis of granular flows using differentiable graph neural network simulator. Comput. Geotech. 171, 106374.
Chow, J.K., Li, Z., Su, Z., Wang, Y.-H., 2022. Characterization of particle orientation of kaolinite samples using the deep learning-based technique. Acta Geotech. 1–14.
Cubuk, E.D., Ivancic, R., Schoenholz, S.S., Strickland, D., Basu, A., Davidson, Z., Fontaine, J., Hor, J.L., Huang, Y.-R., Jiang, Y., 2017. Structure-property relationships from universal signatures of plasticity in disordered solids. Science 358 (6366), 1033–1037.
Cubuk, E.D., Schoenholz, S.S., Rieser, J.M., Malone, B.D., Rottler, J., Durian, D.J., Kaxiras, E., Liu, A.J., 2015. Identifying structural flow defects in disordered solids using machine-learning methods. Phys. Rev. Lett. 114 (10), 108001.
Cui, J., Li, Y., Zhao, C., Zheng, W., 2023. Classification and structural characteristics of amorphous materials based on interpretable deep learning. Chin. Phys. B 32 (9), 096101.
Cundall, P.A., Strack, O.D., 1979. A discrete numerical model for granular assemblies. Geotechnique 29 (1), 47–65.

Davydzenka, T., Tahmasebi, P., 2022. High-resolution fluid–particle interactions: a machine learning approach. J. Fluid Mech. 938, A20.
de Bono, J., McDowell, G., 2016. Particle breakage criteria in discrete-element modelling. Géotechnique 66 (12), 1014–1027.
Duchi, J., Hazan, E., Singer, Y., 2011. Adaptive subgradient methods for online learning and stochastic optimization. J. Mach. Learn. Res. 12 (7).
Duncan, J.M., Chang, C.-Y., 1970. Nonlinear analysis of stress and strain in soils. J. Soil Mech. Found. Div. 96 (5), 1629–1653.
Dutta, A., Gupta, A., Zissermann, A., 2016. VGG image annotator (VIA). URL: http://www. robots. ox. ac. uk/∼ vgg/software/via.
Eghbalian, M., Pouragha, M., Wan, R., 2023. A physics-informed deep neural network for surrogate modeling in classical elasto-plasticity. Comput. Geotech. 159, 105472.
Ellis, G., Yao, C., Zhao, R., Penumadu, D., 1995. Stress-strain modeling of sands using artificial neural networks. J. Geotech. Eng. 121 (5), 429–435.
Faessel, M., Jeulin, D., 2010. Segmentation of 3D microtomographic images of granular materials with the stochastic watershed. J. Microscopy 239 (1), 17–31.
Fan, H., Tian, Z., Xu, X., Sun, X., Ma, Y., Liu, H., Lu, H., 2022. Rockfill material segmentation and gradation calculation based on deep learning. Case Stud. Constr. Mater. 17, e01216.
Farizhandi, A.A.K., Zhao, H., Lau, R., 2016. Modeling the change in particle size distribution in a gas-solid fluidized bed due to particle attrition using a hybrid artificial neural network-genetic algorithm approach. Chem. Eng. Sci. 155, 210–220.
Fathipour-Azar, H., 2022. Machine learning-assisted distinct element model calibration: ANFIS, SVM, GPR, and MARS approaches. Acta Geotech. 17 (4), 1207–1217.
Fazily, P., Yoon, J.W., 2023. Machine learning-driven stress integration method for anisotropic plasticity in sheet metal forming. Int. J. Plasticity 166, 103642.
Feng, Y., 2023. Thirty years of developments in contact modelling of non-spherical particles in DEM: a selective review. Acta Mech. Sin. 39 (1), 722343.
Feng, Y., Han, K., Owen, D., 2007. Coupled lattice Boltzmann method and discrete element modelling of particle transport in turbulent fluid flows: Computational issues. Int. J. Numer. Methods Eng. 72 (9), 1111–1134.
Gao, L., Wang, D., Miao, Y., 2024. A review of two-dimensional image-based technologies for size and shape characterization of coarse-grained granular soils. Powder Technol. 120115.
Ge, R., Liu, Q., 2024. Digital design and manufacturing of microstructural granular materials. Digital Eng., 100008
Ghaboussi, J., Garrett Jr, J., Wu, X., 1991. Knowledge-based modeling of material behavior with neural networks. J. Eng. Mech. 117 (1), 132–153.
Ghaboussi, J., Sidarta, D., 1998. New nested adaptive neural networks (NANN) for constitutive modeling. Comput. Geotech. 22 (1), 29–52.
Giannis, K., Thon, C., Yang, G., Kwade, A., Schilde, C., 2024. Predicting 3D particles shapes based on 2D images by using convolutional neural network. Powder Technol. 432, 119122.
Gong, J., Liu, Z., Nie, J., Cui, Y., Jiang, J., Ou, X., 2025. Study on the automated characterization of particle size and shape of stacked gravelly soils via deep learning. Acta Geotech.
Gong, J., Liu, Z., Zhao, K., Xu, H., Zheng, Y., Jiang, J., Ou, X., 2024a. Quantification of particle size and shape of sands based on the combination of GAN and CNN. Powder Technol., 120122.
Gong, J., Pang, X., Tang, Y., Liu, M., Jiang, J., Ou, X., 2024b. Effects of particle shape, physical properties and particle size distribution on the small-strain stiffness of granular materials: A DEM study. Comput. Geotech. 165, 105903.
Gonzalez Tejada, I., Antolin, P., 2022. Use of machine learning for unraveling hidden correlations between particle size distributions and the mechanical behavior of granular materials. Acta Geotech. 17 (4), 1443–1461.
Gorgogianni, A., Karapiperis, K., Stainier, L., Ortiz, M., Andrade, J.E., 2023. Adaptive goal-oriented data sampling in data-driven computational mechanics. Comput. Methods Appl. Mech. Eng. 409, 115949.
Gu, Z., Ding, X., Hou, X., Ye, J., 2023. A genetic evolved machine learning approach for 3D DEM modelling of anisotropic materials with full consideration of particulate interactions. Compos. Part B: Eng. 250, 110432.
Guan, Q., Yang, Z., Guo, N., Chen, L., 2024. Deep learning-accelerated multiscale approach for granular material modeling. Int. J. Numer. Anal. Methods Geomech. 48 (5), 1372–1389.
Guan, S., Feng, Y., Ma, G., Qu, T., Wang, M., Zhou, W., 2023a. An explicit FEM-NN framework and the analysis of error caused by NN-predicted stress. Acta Geotech., 1-20.
Guan, S., Qu, T., Feng, Y., Ma, G., Zhou, W., 2023. A machine learning-based multi-scale computational framework for granular materials. Acta Geotech. 18 (4), 1699–1720.
Guo, N., Yang, Z., Yuan, W., Zhao, J., 2021. A coupled SPFEM/DEM approach for multiscale modeling of large-deformation geomechanical problems. Int. J. Numer. Anal. Methods Geomech. 45 (5), 648–667.
Guo, N., Zhao, J., 2014. A coupled FEM/DEM approach for hierarchical multiscale modelling of granular media. Int. J. Numer. Meth. Eng. 99 (11), 789–818.
Guo, N., Zhao, J., 2016a. Multiscale insights into classical geomechanics problems. Int. J. Numer. Anal. Meth. Geomech. 40 (3), 367–390.
Guo, N., Zhao, J., 2016b. Parallel hierarchical multiscale modelling of hydro-mechanical problems for saturated granular soils. Comput. Methods Appl. Mech. Eng. 305, 37–61.
Gupta, A., Mishra, B., 2024. Neuroevolving monotonic PINNs for particle breakage analysis. 2024 IEEE Conference on Artificial Intelligence (CAI), IEEE.
Hadi, A., Moradi, M., Pang, Y., Schott, D., 2024. Adaptive AI-based surrogate modelling via transfer learning for DEM simulation of multi-component segregation. Sci. Rep. 14 (1), 27003.
Haeri, A., Holz, D., Skonieczny, K., 2024. Subspace graph networks for real-time granular flow simulation with applications to machine-terrain interactions. Eng. Appl. Artificial Intelligence 135, 108765.
Hager, A., Kloss, C., Pirker, S., Goniva, C., 2014. Parallel resolved open source CFD-DEM: method, validation and application. J. Comput. Multiphase Flows 6 (1), 13–27.
Hajisharifi, A., Romanò, F., Girfoglio, M., Beccari, A., Bonanni, D., Rozza, G., 2023. A non-intrusive data-driven reduced order model for parametrized CFD-DEM numerical simulations. J. Comput. Phys. 491, 112355.
Hamilton, B.W., Germann, T.C., 2023. Using limited neural networks to assess relative mechanistic influence on shock heating in granular solids. Phys. Rev. Mater. 7 (8), 085601.
Hanlan, J.M., Dillavou, S., Liu, A.J., Durian, D.J., 2024. Cornerstones are the key stones: Using interpretable machine learning to probe the clogging process in 2D granular hoppers. arXiv preprint arXiv:2407.05491.
He, C., Sadeghpour, H., Shi, Y., Mishra, B., Roshankhah, S., 2024. Mapping distribution of fractures and minerals in rock samples using Res-VGG-UNet and threshold segmentation methods. Comput. Geotech. 175, 106675.
He, G.-F., Zhang, P., Yin, Z.-Y., Jin, Y.-F., Yang, Y., 2023. Multi-fidelity data-driven modelling of rate-dependent behaviour of soft clays. Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards 17(1), 64-76.
He, L., Tafti, D.K., 2019. A supervised machine learning approach for predicting variable drag forces on spherical particles in suspension. Powder Technol. 345, 379–389.
Heider, Y., Wang, K., Sun, W., 2020. SO (3)-invariance of informed-graph-based deep neural network for anisotropic elastoplastic materials. Comput. Methods Appl. Mech. Eng. 363, 112875.
Hesse, R., Krull, F., Antonyuk, S., 2021. Prediction of random packing density and flowability for non-spherical particles by deep convolutional neural networks and Discrete Element Method simulations. Powder Technol. 393, 559–581.
Ho, J., Jain, A., Abbeel, P., 2020. Denoising diffusion probabilistic models. Adv. Neural Inf. Proces. Syst. 33, 6840–6851.
Hu, G., Zhou, B., Zheng, W., Li, C., Wang, H., 2024. A machine learning-based drag model for sand particles in transition flow aided by spherical harmonic analysis and resolved CFD-DEM. Acta Geotech.
Hu, H.H., Patankar, N.A., Zhu, M., 2001. Direct numerical simulations of fluid–solid systems using the arbitrary Lagrangian–Eulerian technique. J. Comput. Phys. 169 (2), 427–462.
Huang, D., Fuhg, J.N., Weißenfels, C., Wriggers, P., 2020. A machine learning based plasticity model using proper orthogonal decomposition. Comput. Methods Appl. Mech. Eng. 365, 113008.
Huang, S., Wang, P., Lai, Z., Yin, Z.-Y., Huang, L., Xu, C., 2024. Machine-learning-enabled discrete element method: The extension to three dimensions and computational issues. Comput. Meth. Appl. Mech. Eng. 432, 117445.
Hwang, S., Pan, J., Fan, L.-S., 2021. A machine learning-based interaction force model for non-spherical and irregular particles in low Reynolds number incompressible flows. Powder Technol. 392, 632–638.
Hwang, S., Pan, J., Fan, L.-S., 2024. Neural-network-based drag force model for polydisperse assemblies of irregular-shaped particles. Powder Technol. 440, 119783.
Hwang, S., Pan, J., Sunny, A.A., Fan, L.-S., 2022. A machine learning-based particle-particle collision model for non-spherical particles with arbitrary shape. Chem. Eng. Sci. 251, 117439.
Irazábal, J., Salazar, F., Vicente, D.J., 2023. A methodology for calibrating parameters in discrete element models based on machine learning surrogates. Comput. Particle Mech. 10 (5), 1031–1047.
Jaza, R., Mollon, G., Descartes, S., Paquet, A., Berthier, Y., 2021. Lessons learned using machine learning to link third body particles morphology to interface rheology. Tribol. Int. 153, 106630.
Jiang, Y., Byrne, E., Glassey, J., Chen, X., 2024. Integrating graph neural network-based surrogate modeling with inverse design for granular flows. Ind. Eng. Chem. Res.
Jung, J., Yoon, J.I., Park, H.K., Jo, H., Kim, H.S., 2020. Microstructure design using machine learning generated low dimensional and continuous design space. Materialia 11, 100690.
Karapiperis, K., Stainier, L., Ortiz, M., Andrade, J., 2021. Data-driven multiscale modeling in mechanics. J. Mech. Phys. Solids 147, 104239.
Kawamoto, R., Andò, E., Viggiani, G., Andrade, J.E., 2016. Level set discrete element method for three-dimensional computations with triaxial case study. J. Mech. Phys. Solids 91, 1–13.
Kawamoto, R., Andò, E., Viggiani, G., Andrade, J.E., 2018. All you need is shape: Predicting shear banding in sand with LS-DEM. J. Mech. Phys. Solids 111, 375–392.
Khodkari, N., Hamidian, P., Khodkari, H., Payan, M., Behnood, A., 2024. Predicting the small strain shear modulus of sands and sand-fines binary mixtures using machine learning algorithms. Transp. Geotech. 44, 101172.
Kim, T., Yun, T.S., Suh, H.S., 2025. Can ChatGPT implement finite element models for geotechnical engineering applications? Int. J. Numer. Anal. Meth. Geomech.
Kim, Y., Ma, J., Lim, S.Y., Song, J.Y., Yun, T.S., 2022. Determination of shape parameters of sands: a deep learning approach. Acta Geotech. 17 (4), 1521–1531.
Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
Kirchdoerfer, T., Ortiz, M., 2016. Data-driven computational mechanics. Comput. Methods Appl. Mech. Eng. 304, 81–101.
Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., 2023. Segment anything. Proceedings of the IEEE/CVF International Conference on Computer Vision.
Kishida, N., Nakamura, H., Ohsaki, S., Watano, S., 2023. Development of ultra-fast computing method for powder mixing process. Chem. Eng. J. 475, 146166.

Kishida, N., Nakamura, H., Ohsaki, S., Watano, S., 2024. Optimizing data-sampling period in a machine learning-based surrogate model for powder mixing simulations. Powder Technol., 120584

Kumar, K., Choi, Y., 2023. Accelerating particle and fluid simulations with differentiable graph networks for solving forward and inverse problems. Proceedings of the SC'23 Workshops of The International Conference on High Performance Computing, Network, Storage, and Analysis.

Kumar, R., Patel, C.M., Jana, A.K., Gopireddy, S.R., 2018. Prediction of hopper discharge rate using combined discrete element method and artificial neural network. Adv. Powder Technol. 29 (11), 2822–2834.

Lai, W., Hu, F., Kong, X., Yan, P., Bian, K., Dai, X., 2022a. The study of coal gangue segmentation for location and shape predicts based on multispectral and improved Mask R-CNN. Powder Technol. 407, 117655.

Lai, Z., Chen, Q., 2019. Reconstructing granular particles from X-ray computed tomography using the TWS machine learning tool and the level set method. Acta Geotech. 14, 1–18.

Lai, Z., Chen, Q., Huang, L., 2022. Machine-learning-enabled discrete element method: Contact detection and resolution of irregular-shaped particles. Int. J. Numer. Anal. Meth. Geomech. 46 (1), 113–140.

Laudari, S., Marks, B., Rognon, P., 2022. Classifying grains using behaviour-informed machine learning. Sci. Rep. 12 (1), 13915.

Laudari, S., Marks, B., Rognon, P., 2023. Insights on the internal dynamics of bi-disperse granular flows from machine learning. Granul. Matter 25 (4), 73.

Li, K.-Q., Yin, Z.-Y., Zhang, N., Liu, Y., 2023a. A data-driven method to model stress-strain behaviour of frozen soil considering uncertainty. Cold Regions Sci. Technol. 213, 103906.

Li, R., Yin, Z.-Y., He, S., 2024a. 3D reconstruction of arbitrary granular media utilizing vision foundation model. Appl. Soft Comput. 112599.

Li, R., Zhang, P., Yin, Z.Y., Sheil, B., 2024. Enhanced hybrid algorithms for segmentation and reconstruction of granular grains from X-ray micro computed-tomography images. Int. J. Numer. Anal. Meth. Geomech.

Li, S., Duan, G., Sakai, M., 2023b. On POD-based modal analysis in simulations of granular flows. Powder Technol. 413, 118058.

Li, S., Sakai, M., 2024. Advanced graph neural network-based surrogate model for granular flows in arbitrarily shaped domains. Chem. Eng. J., 157349

Li, X., Jin, W., Klinger, J., Saha, N., Lajnef, N., 2025. Data-driven mechanical behavior modeling of granular biomass materials. Comput. Geotech. 177, 106907.

Li, X., Roth, C.C., Mohr, D., 2019. Machine-learning based temperature-and rate-dependent plasticity model: Application to analysis of fracture experiments on DP steel. Int. J. Plast 118, 320–344.

Li, Z., Li, X., Zhang, H., Huang, D., Zhang, L., 2023. The prediction of contact force networks in granular materials based on graph neural networks. J. Chem. Phys. 158 (5).

Liang, W., Zhao, J., 2019. Multiscale modeling of large deformation in geomechanics. Int. J. Numer. Anal. Meth. Geomech. 43 (5), 1080–1114.

Liang, Z., Nie, Z., An, A., Gong, J., Wang, X., 2019. A particle shape extraction and evaluation method using a deep convolutional neural network and digital image processing. Powder Technol. 353, 156–170.

Liao, Z., Yang, Y., Sun, C., Wu, R., Duan, Z., Wang, Y., Li, X., Xu, J., 2021. Image-based prediction of granular flow behaviors in a wedge-shaped hopper by combing DEM and deep learning methods. Powder Technol. 383, 159–166.

Lin, C.-C., 2024. Application of artificial neural networks to predict the particle-scale contact force of photoelastic disks. Adv. Powder Technol. 35 (1), 104284.

Lin, H., Gao, Q., 2024. A calibration-informed deep learning model for three-dimensional particle reconstruction of volumetric particle image velocimetry. Phys. Fluids 36 (11).

Ling, J., Jones, R., Templeton, J., 2016. Machine learning strategies for systems with invariance properties. J. Comput. Phys. 318, 22–35.

Liu, H., Su, H., Sheil, B., 2025. Automatic determination of 3D particle morphology from multiview images using uncertainty-evaluated deep learning. Comput.-Aided Civil Infrastructure Eng.

Liu, L., Wang, T., Long, X., Ji, S., 2025b. Deep-learning-based parametric calibration of discrete element method for the breakage simulation of sea ice. Comput. Particle Mech. 1–18.

Liu, M., Liu, G., 2010. Smoothed particle hydrodynamics (SPH): an overview and recent developments. Arch. Comput. Meth. Eng. 17, 25–76.

Liu, X., Yang, J., Zou, D., Li, Z., Chen, Y., Cao, X., 2024. Utilizing DEM and interpretable ML algorithms to examine particle size distribution's role in small-strain shear modulus of gap-graded granular mixtures. Constr. Build. Mater. 428, 136232.

Liu, Y., Zhang, J., Zhao, T., Wang, Z., Wang, Z., 2023. Reconstruction of the meso-scale concrete model using a deep convolutional generative adversarial network (DCGAN). Constr. Build. Mater. 370, 130704.

Long, S., Xu, S., Zhang, Y., Li, B., Sun, L., Wang, Y., Wang, J., 2023. Method of soil-elastoplastic dem parameter calibration based on recurrent neural network. Powder Technol. 416, 118222.

Lu, L., Gao, X., Dietiker, J.-F., Shahnam, M., Rogers, W.A., 2021. Machine learning accelerated discrete element modeling of granular flows. Chem. Eng. Sci. 245, 116832.

Ma, G., Guan, S., Wang, Q., Feng, Y., Zhou, W., 2022. A predictive deep learning framework for path-dependent mechanical behavior of granular materials. Acta Geotech. 17, 3463–3478.

Ma, G., Mei, J., Gao, K., Zhao, J., Zhou, W., Wang, D., 2022b. Machine learning bridges microslips and slip avalanches of sheared granular gouges. Earth Planet. Sci. Lett. 579, 117366.

Macedo, R.B.D., Monfared, S., Karapiperis, K., Andrade, J.E., 2023. What is shape? Characterizing particle morphology with genetic algorithms and deep generative models. Granular Matter 25 (1), 2.

Mahdi, F.M., Holdich, R.G., 2017. Using statistical and artificial neural networks to predict the permeability of loosely packed granular materials. Sep. Sci. Technol. 52 (1), 1–12.

Mandal, R., Casert, C., Sollich, P., 2022. Robust prediction of force chains in jammed solids using graph neural networks. Nat. Commun. 13 (1), 4424.

Masi, F., Stefanou, I., 2023. Evolution TANN and the identification of internal variables and evolution equations in solid mechanics. J. Mech. Phys. Solids 174, 105245.

Masi, F., Stefanou, I., Vannucci, P., Maffi-Berthier, V., 2021. Thermodynamics-based Artificial Neural Networks for constitutive modeling. J. Mech. Phys. Solids 147, 104277.

Mayr, A., Lehner, S., Mayrhofer, A., Kloss, C., Hochreiter, S., Brandstetter, J., 2023. Boundary graph neural networks for 3d simulations. Proceedings of the AAAI Conference on Artificial Intelligence.

Mei, J., Ma, G., Cao, W., Wu, T., Zhou, W., 2024. Graph neural network unveils the spatiotemporal evolution of structural defects in sheared granular materials. Int. J. Plast 104218.

Mei, J., Ma, G., Tang, L., Gao, K., Cao, W., Zhou, W., 2023. Spatial clustering of microscopic dynamics governs the slip avalanche of sheared granular materials. Int. J. Plasticity 163, 103570.

Mei, J., Ma, G., Wang, Q., Wu, T., Zhou, W., 2022. Micro-and macroscopic aspects of the intermittent behaviors of granular materials related by graph neural network. Int. J. Solids Struct. 251, 111763.

Mital, U., Andrade, J.E., 2022. Bridging length scales in granular materials using convolutional neural networks. Comput. Particle Mech. 9, 221–235.

Munjiza, A., Owen, D., Bicanic, N., 1995. A combined finite-discrete element method in transient dynamics of fracturing solids. Eng. Comput. 12 (2), 145–174.

Muralidhar, N., Bu, J., Cao, Z., He, L., Ramakrishnan, N., Tafti, D., Karpatne, A., 2020. Physics-guided deep learning for drag force prediction in dense fluid-particulate systems. Big Data 8 (5), 431–449.

Nasato, D.S., Goniva, C., Pirker, S., Kloss, C., 2015. Coarse graining for large-scale DEM simulations of particle flow–an investigation on contact and cohesion models. Procedia Eng. 102, 1484–1490.

Nascimento, A., Roongta, S., Diehl, M., Beyerlein, I.J., 2023. A machine learning model to predict yield surfaces from crystal plasticity simulations. Int. J. Plasticity 161, 103507.

Oñate, E., Idelsohn, S.R., Del Pin, F., Aubry, R., 2004. The particle finite element method—an overview. Int. J. Comput. Methods 1 (02), 267–307.

Pan, X., Niu, Y., Zhao, Y., Huang, P., Wu, Y., 2023. Parameter calibration method of clustered-particle logic concrete DEM model using BP neural network-particle swarm optimisation algorithm (BP-PSO) inversion method. Eng. Fract. Mech. 292, 109659.

Parr, T., Wilson, J.D., 2021. Partial dependence through stratification. Mach. Learning Appl. 6, 100146.

Peng, Y., Yin, Z.-Y., 2023. Micromechanical analysis of suction pile-granular soil interaction under inclined pulling load of mooring line: Mooring depth effect. Mar. Struct. 92, 103499.

Puzyrev, D., Harth, K., Trittel, T., Stannarius, R., 2020. Machine learning for 3D particle tracking in granular gases. Microgravity Sci. Technol. 32 (5), 897–906.

Qu, S., Zhang, W., You, C., 2022a. Modeling of dynamic characteristic of particle in transient gas–solid flow via a machine learning approach. Powder Technol. 412, 117939.

Qu, T., Di, S., Feng, Y., Wang, M., Zhao, T., 2021a. Towards data-driven constitutive modelling for granular materials via micromechanics-informed deep learning. Int. J. Plasticity 144, 103046.

Qu, T., Di, S., Feng, Y.T., Wang, M., Zhao, T., Wang, M., 2021b. Deep learning predicts stress–strain relations of granular materials based on triaxial testing data. Comput. Model. Eng. Sci. 128 (1), 129–144.

Qu, T., Feng, Y., Wang, M., 2021c. An adaptive granular representative volume element model with an evolutionary periodic boundary for hierarchical multiscale analysis. Int. J. Numer. Meth. Eng. 122 (9), 2239–2253.

Qu, T., Feng, Y., Wang, M., Jiang, S., 2020a. Calibration of parallel bond parameters in bonded particle models via physics-informed adaptive moment optimisation. Powder Technol. 366, 527–536.

Qu, T., Feng, Y., Zhao, J., 2022b. A strain energy-based elastic parameter calibration method for lattice/bonded particle modelling of solid materials. Powder Technol. 410, 117860.

Qu, T., Guan, S., Feng, Y., Ma, G., Zhou, W., Zhao, J., 2023a. Deep active learning for constitutive modelling of granular materials: From representative volume elements to implicit finite element modelling. Int. J. Plasticity 164, 103576.

Qu, T., Zhao, J., Guan, S., Feng, Y., 2023b. Data-driven multiscale modelling of granular materials via knowledge transfer and sharing. Int. J. Plasticity 171, 103786.

Qu, T.M., Feng, Y.T., Zhao, T.T., Wang, M., 2020b. A hybrid calibration approach to Hertz-type contact parameters for discrete element models. Int. J. Numer. Anal. Meth. Geomech. 44 (9), 1281–1300.

Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. J. Comput. Phys. 378, 686–707.

Rangel, R.L., Franci, A., Oñate, E., Gimenez, J.M., 2024a. Multiscale data-driven modeling of the thermomechanical behavior of granular media with thermal expansion effects. Comput. Geotech. 176, 106789.

Rangel, R.L., Gimenez, J.M., Oñate, E., Franci, A., 2024b. A continuum–discrete multiscale methodology using machine learning for thermal analysis of granular media. Comput. Geotech. 168, 106118.

Ren, C.X., Dorostkar, O., Rouet-Leduc, B., Hulbert, C., Strebel, D., Guyer, R.A., Johnson, P.A., Carmeliet, J., 2019. Machine learning reveals the state of intermittent frictional dynamics in a sheared granular fault. Geophys. Res. Lett. 46 (13), 7395–7403.
Rouet-Leduc, B., Hulbert, C., Lubbers, N., Barros, K., Humphreys, C.J., Johnson, P.A., 2017. Machine learning predicts laboratory earthquakes. Geophys. Res. Lett. 44 (18), 9276–9282.
Sanchez-Gonzalez, A., Godwin, J., Pfaff, T., Ying, R., Leskovec, J., Battaglia, P., 2020. Learning to simulate complex physics with graph networks. International conference on machine learning. PMLR.
Santos, D.M., Neto, A.G., 2023. Evaluating the influence of loss function on performance of a neural net-work for particle 3D shape reconstruction from 2D projections. XLIV Ibero-Latin American Congress on Computational Methods Engineering.
Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G., 2008. The graph neural network model. IEEE Trans. Neural Netw. 20 (1), 61–80.
Sergazinov, R., Kramár, M., 2021. Machine learning approach to force reconstruction in photoelastic materials. Mach. Learn.: Sci. Technol. 2 (4), 045030.
Shentu, J., Lin, B., 2023. A novel machine learning framework for efficient calibration of complex dem model: A case study of a conglomerate sample. Eng. Fract. Mech. 279, 109044.
Shentu, J., Lin, B., Jin, Y., Yoon, J.S., 2024. Investigation of hydraulic fracture propagation in conglomerate rock using discrete element method and explainable machine learning framework. Acta Geotech. 19 (6), 3837–3862.
Shi, J.-J., Zhang, W., Wang, W., Sun, Y.-H., Xu, C.-Y., Zhu, H.-H., Sun, Z.-X., 2021a. Randomly generating three-dimensional realistic schistous sand particles using deep learning: Variational autoencoder implementation. Eng. Geol. 291, 106235.
Shi, X., Liu, K., Yin, J., 2021. Effect of initial density, particle shape, and confining stress on the critical state behavior of weathered gap-graded granular soils. J. Geotech. Geoenviron. Eng. 147 (2), 04020160.
Silling, S.A., 2000. Reformulation of elasticity theory for discontinuities and long-range forces. J. Mech. Phys. Solids 48 (1), 175–209.
Su, D., Yan, W., 2018. 3D characterization of general-shape sand particles using microfocus X-ray computed tomography and spherical harmonic functions, and particle regeneration using multivariate random vector. Powder Technol. 323, 8–23.
Su, M., Guo, N., Yang, Z., 2023. A multifidelity neural network (MFNN) for constitutive modeling of complex soil behaviors. Int. J. Numer. Anal. Meth. Geomech.
Su, M., Yu, Y., Chen, T., Guo, N., Yang, Z., 2024. A thermodynamics-informed neural network for elastoplastic constitutive modeling of granular materials. Comput. Methods Appl. Mech. Eng. 430, 117246.
Sulsky, D., Chen, Z., Schreyer, H.L., 1994. A particle method for history-dependent materials. Comput. Methods Appl. Mech. Eng. 118 (1–2), 179–196.
Sun, Q., Zheng, Y., Li, B., Zheng, J., Wang, Z., 2019. Three-dimensional particle size and shape characterisation using structural light. Géotech. Lett. 9 (1), 72–78.
Sun, Z., Shi, J., Wang, J., Jiang, M., Wang, Z., Bai, X., Wang, X., 2022. A deep learning-based framework for automatic analysis of the nanoparticle morphology in SEM/TEM images. Nanoscale 14 (30), 10761–10772.
Ulloa, J., Gorgogianni, A., Karapiperis, K., Ortiz, M., Andrade, J.E., 2023. Data-driven breakage mechanics: Predicting the evolution of particle-size distribution in granular media. J. Mech. Phys. Solids 178, 105328.
Ummenhofer, B., Prantl, L., Thuerey, N., Koltun, V., 2019. Lagrangian fluid simulation with continuous convolutions. International Conference on Learning Representations.
Vargas, W.L., McCarthy, J.J., 2007. Thermal expansion effects and heat conduction in granular materials. Phys. Rev. E—Stat., Nonlinear, and Soft Matter Phys. 76(4), 041301.
Vidyapati, V., Subramaniam, S., 2012. Granular rheology and phase transition: DEM simulations and order-parameter based constitutive model. Chem. Eng. Sci. 72, 20–34.
Vlassis, N.N., Sun, W., 2021. Sobolev training of thermodynamic-informed neural networks for interpretable elasto-plasticity models with level set hardening. Comput. Methods Appl. Mech. Eng. 377, 113695.
Vlassis, N.N., Sun, W., 2022. Component-based machine learning paradigm for discovering rate-dependent and pressure-sensitive level-set plasticity models. J. Appl. Mech. 89 (2), 021003.
Vlassis, N.N., Sun, W., Alshibli, K.A., Regueiro, R.A., 2023. Synthesizing realistic sand assemblies with denoising diffusion in latent space. Int. J. Numer. Anal. Meth. Geomech.
Wang, J., Ma, G., Qu, T., Guan, S., Zhou, W., Chang, X., 2025. GPM-PeNN: A generalized plasticity model-based data-driven constitutive modeling framework using physics-encoded neural network. Comput. Methods Appl. Mech. Eng. 436, 117694.
Wang, K., Johnson, C.W., Bennett, K.C., Johnson, P.A., 2021a. Predicting fault slip via transfer learning. Nat. Commun. 12 (1), 7319.
Wang, K., Sun, W., 2018. A multiscale multi-permeability poroplasticity model linked by recursive homogenizations and deep learning. Comput. Methods Appl. Mech. Eng. 334, 337–380.
Wang, K., Sun, W., 2019. Meta-modeling game for deriving theory-consistent, microstructure-based traction–separation laws via deep reinforcement learning. Comput. Methods Appl. Mech. Eng. 346, 216–241.
Wang, M., Feng, Y., Guan, S., Qu, T., 2024a. Multi-layer perceptron-based data-driven multiscale modelling of granular materials with a novel Frobenius norm-based internal variable. J. Rock Mech. Geotech. Eng.
Wang, M., Kumar, K., Feng, Y., Qu, T., Wang, M., 2024b. Machine learning aided modeling of granular materials: a review. Arch. Comput. Meth. Eng. 1–38.
Wang, M., Qu, T., Guan, S., Zhao, T., Liu, B., Feng, Y., 2022. Data-driven strain–stress modelling of granular materials via temporal convolution neural network. Comput. Geotech. 152, 105049.
Wang, X., Tian, K., Su, D., Zhao, J., 2019a. Superellipsoid-based study on reproducing 3D particle geometry from 2D projections. Comput. Geotech. 114, 103131.
Wang, X., Wu, W., Zhu, H.-H., 2024c. Solving fluid flow in discontinuous heterogeneous porous media and multi-layer strata with interpretable physics-encoded finite element network. J. Rock Mech. Geotech. Eng.
Wang, X., Yin, Z.-Y., Wu, W., Zhu, H.-H., 2025b. Differentiable finite element method with Galerkin discretization for fast and accurate inverse analysis of multidimensional heterogeneous engineering structures. Comput. Methods Appl. Mech. Eng. 437, 117755.
Wang, X., Yin, Z.-Y., Wu, W., Zhu, H.-H., 2025c. Neural network-augmented differentiable finite element method for boundary value problems. Int. J. Mech. Sci. 285, 109783.
Wang, X., Zhang, H., Yin, Z.-Y., Su, D., Liu, Z., 2023. Deep-learning-enhanced model reconstruction of realistic 3D rock particles by intelligent video tracking of 2D random particle projections. Acta Geotech. 18 (3), 1407–1430.
Wang, Y., Ma, G., Mei, J., Zou, Y., Zhang, D., Zhou, W., Cao, X., 2021b. Machine learning reveals the influences of grain morphology on grain crushing strength. Acta Geotech. 16 (11), 3617–3630.
Wang, Y., Wang, G., Wu, S., Liu, Z., Fang, Y., 2023b. A calibration method for ore bonded particle model based on deep learning neural network. Powder Technol. 420, 118417.
Wang, Z., Teng, Y., Liu, M., 2019. A semi-resolved CFD–DEM approach for particulate flows with kernel based approximation and Hilbert curve based searching strategy. J. Comput. Phys. 384, 151–169.
Wei, M., Gao, K., 2024. Machine learning predicts the slip duration and friction drop of laboratory earthquakes in sheared granular fault. J. Geophys. Res.: Mach. Learn. Comput. 1 (4) e2024JH000398.
Wen, J., Zou, Q., Wei, Y., 2021. Physics-driven machine learning model on temperature and time-dependent deformation in lithium metal and its finite element implementation. J. Mech. Phys. Solids 153, 104481.
Westbrink, F., Elbel, A., Schwung, A., Ding, S.X., 2021. Optimization of DEM parameters using multi-objective reinforcement learning. Powder Technol. 379, 602–616.
Wu, M., Wang, J., 2022a. Constitutive modelling of natural sands using a deep learning approach accounting for particle shape effects. Powder Technol. 404, 117439.
Wu, M., Wang, J., 2022b. Prediction of 3D contact force chains using artificial neural networks. Eng. Geol. 296, 106444.
Wu, W., Chen, K., Tsotsas, E., 2024. Prediction of rod-like particle mixing in rotary drums by three machine learning methods based on DEM simulation data. Powder Technol. 120307.
Xi, J., Gao, L., Zheng, J., Wang, D., Wang, G., Guan, Z., Zheng, J., 2025. Gravel particle shape classification from half-particle point clouds using a dynamic graph edge convolution neural network. Comput. Geotech. 179, 107015.
Xiang, H., Cheng, C., Zhang, P., Jiang, G., 2024. Machine learning and numerical investigation on drag coefficient of arbitrary polygonal particles. Int. J. Mech. Syst. Dynamics.
Xiao, Y., Peng, Y., Wang, M., Ning, Y., Zhou, Y., Kong, K., Long, Y., 2024. A novel method for predicting coarse aggregate particle size distribution based on segment anything model and machine learning. Constr. Build. Mater. 429, 136429.
Xie, H., Yao, H., Sun, X., Zhou, S., Zhang, S., 2019. Pix2vox: Context-aware 3d reconstruction from single and multi-view images. Proceedings of the IEEE/CVF international conference on computer vision.
Xie, Z., Gu, X., Shen, Y., 2022. A machine learning study of predicting mixing and segregation behaviors in a bidisperse solid–liquid fluidized bed. Ind. Eng. Chem. Res. 61 (24), 8551–8565.
Xiong, W., Wang, J., 2021. Gene mutation of particle morphology through spherical harmonic-based principal component analysis. Powder Technol. 386, 176–192.
Xiong, W., Wang, J., 2024. Deep transfer learning-aided constitutive modelling of granular soils considering out-of-range particle morphology. Tunnelling and Underground Space Technology 144, 105547.
Xiong, W., Wang, J., Wu, M., 2023. Data-driven constitutive modelling of granular soils considering multiscale particle morphology. Comput. Geotech. 162, 105699.
Xu, D., Shen, Y., 2022. An improved machine learning approach for predicting granular flows. Chem. Eng. J. 450, 138036.
Xu, H., Zhu, J., Finegan, D.P., Zhao, H., Lu, X., Li, W., Hoffman, N., Bertei, A., Shearing, P., Bazant, M.Z., 2021. Guiding the design of heterogeneous electrode microstructures for Li-ion batteries: microscopic imaging, predictive modeling, and machine learning. Adv. Energy Mater. 11 (19), 2003908.
Xu, K., Huang, D.Z., Darve, E., 2021b. Learning constitutive relations using symmetric positive definite neural networks. J. Comput. Phys. 428, 110072.
Yang, D., Wang, X., Zhang, H., Yin, Z.-Y., Su, D., Xu, J., 2021. A Mask R-CNN based particle identification for quantitative shape evaluation of granular materials. Powder Technol. 392, 296–305.
Yasuda, T., Ookawara, S., Yoshikawa, S., Matsumoto, H., 2021. Machine learning and data-driven characterization framework for porous materials: Permeability prediction and channeling defect detection. Chem. Eng. J. 420, 130069.
Ye, F., Wheeler, C., Chen, B., Hu, J., Chen, K., Chen, W., 2019. Calibration and verification of DEM parameters for dynamic particle flow conditions using a backpropagation neural network. Adv. Powder Technol. 30 (2), 292–301.
Zeng, R., Wang, S., Zhang, Y., Qu, T., 2025. CFD-DEM modeling of seepage in foam-conditioned soil. Comput. Geotech. 177, 106818.
Zhan, X., Pan, X., Dai, B., Liu, Z., Lin, D., Loy, C.C., 2020. Self-supervised scene de-occlusion. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition.
Zhang, H., Yin, Z.-Y., Zhang, N., Wang, X., 2024a. A rapid segmentation and occlusion completion method for morphology analysis of packed granular particles considering uncertainty. Can. Geotech. J.(ja).

Zhang, H., Yin, Z.-Y., Zhang, N., Wang, X., Ding, Z., 2024b. A scale-adaptive Mask R-CNN strategy for foreground particle segmentation and geometrical analysis of granular aggregates. Appl. Soft Comput. 164, 111931.
Zhang, J., Ma, G., Yang, Z., Mei, J., Zhang, D., Zhou, W., Chang, X., 2024c. Knowledge extraction via machine learning guides a topology-based permeability prediction model. Water Resour. Res. 60 (7) e2024WR037124.
Zhang, L., Chen, J., Zhang, H., Huang, D., 2023a. The prediction of dynamical quantities in granular avalanches based on graph neural networks. J. Chem. Phys. 159 (21).
Zhang, N., Xu, K., Yin, Z.Y., Li, K.-Q., Jin, Y.-F., 2025a. Finite element-integrated neural network framework for elastic and elastoplastic solids. Comput. Methods Appl. Mech. Eng. 433, 117474.
Zhang, N., Zhou, A., Jin, Y.-F., Yin, Z.-Y., Shen, S.-L., 2023b. An enhanced deep learning method for accurate and robust modelling of soil stress–strain response. Acta Geotech. 18 (8), 4405–4427.
Zhang, P., Karapiperis, K., Weeger, O., 2025. t-PiNet: A thermodynamics-informed hierarchical learning for discovering constitutive relations of geomaterials. J. Mech. Phys. Solids, 106049.
Zhang, P., Yin, Z.-Y., 2021. A novel deep learning-based modelling strategy from image of particles to mechanical properties for granular materials with CNN and BiLSTM. Comput. Meth. Appl. Mech. Eng. 382.
Zhang, P., Yin, Z.-Y., Chen, Q., 2022a. Image-based 3D reconstruction of granular grains via hybrid algorithm and level set with convolution kernel. J. Geotech. Geoenviron. Eng. 148 (5), 04022021.
Zhang, P., Yin, Z.-Y., Chen, W.-B., Jin, Y.-F., 2021. CNN-based intelligent method for identifying GSD of granular soils. Int. J. Geomech. 21 (12), 04021229.
Zhang, P., Yin, Z.-Y., Jin, Y.-F., Yang, J., Sheil, B., 2022b. Physics-informed multifidelity residual neural networks for hydromechanical modeling of granular soils and foundation considering internal erosion. J. Eng. Mech. 148 (4), 04022015.
Zhang, P., Yin, Z.Y., Jin, Y.F., 2022c. Three-dimensional quantitative analysis on granular particle shape using convolutional neural network. Int. J. Numer. Anal. Meth. Geomech. 46 (1), 187–204.
Zhang, P., Yin, Z.Y., Jin, Y.F., Sheil, B., 2022d. Physics-constrained hierarchical data-driven modelling framework for complex path-dependent behaviour of soils. Int. J. Numer. Anal. Meth. Geomech. 46 (10), 1831–1850.
Zhang, P., Yin, Z.Y., Jin, Y.F., Ye, G.L., 2020. An AI-based model for describing cyclic characteristics of granular materials. Int. J. Numer. Anal. Meth. Geomech. 44 (9), 1315–1335.
Zhang, R., Li, K., Yu, F., Zhang, H., Gao, Z., Huang, Y., 2023c. Aggregate particle identification and gradation analysis method based on the deep learning network of Mask R-CNN. Mater. Today Commun. 35, 106269.
Zhang, T., Li, S., Yang, H., Zhang, F., 2024c. Prediction of constrained modulus for granular soil using 3D discrete element method and convolutional neural networks. J. Rock Mech. Geotech. Eng.
Zhang, X., Zhao, T., Liu, Y., Chen, Q., Wang, Z., Wang, Z., 2023d. A data-driven model for predicting the mixed-mode stress intensity factors of a crack in composites. Eng. Fract. Mech. 288, 109385.
Zhang, Y., Ma, G., Tang, L., Zhou, W., 2022e. Predicting the crystalline phase generation effectively in monosized granular matter using machine learning. Granul. Matter 24, 1–13.
Zhang, Y., Zhou, W., Ma, G., Cheng, R., Chang, X., 2022f. The structure–property relationship of granular materials with different friction coefficients: Insight from machine learning. Extreme Mech. Lett. 54, 101759.
Zhao, B., He, M., Wang, J., 2023a. Data-driven discovery of the governing equation of granular flow in the homogeneous cooling state using sparse regression. Phys. Fluids 35 (1).
Zhao, J., Shan, T., 2013. Coupled CFD–DEM simulation of fluid–particle interaction in geomechanics. Powder Technol. 239, 248–258.
Zhao, J., Xie, H., Li, C., Liu, Y., 2024. Deep learning-based reconstruction of 3D morphology of geomaterial particles from single-view 2D images. Materials 17 (20), 5100.
Zhao, S., Chen, H., Zhao, J., 2025. A physical-information-flow-constrained temporal graph neural network-based simulator for granular materials. Comput. Methods Appl. Mech. Eng. 433, 117536.
Zhao, S., Zhao, J., Liang, W., 2021. A thread-block-wise computational framework for large-scale hierarchical continuum-discrete modeling of granular media. Int. J. Numer. Meth. Eng. 122 (2), 579–608.
Zhao, Y., Gao, X., Zhang, P., Lei, L., Li, S.Z., Galindo-Torres, S., 2023b. Reconstruction and generation of 3D realistic soil particles with metaball descriptor. Comput. Geotech. 161.
Zhao, Y., Liu, J., Gao, X., Torres, S., Li, S.Z., 2023c. Scalable Particle Generation for Granular Shape Study. NeurIPS 2023 AI for Science Workshop.
Zheng, J., Hryciw, R.D., 2018. Identification and characterization of particle shapes from images of sand assemblies using pattern recognition. J. Comput. Civ. Eng. 32 (3), 04018016.
Zheng, J., Zhang, Z., Li, C., Li, Z., Gao, L., 2022. Laboratory-on-a-smartphone for estimating angularity of granular soils. Acta Geotech. 1–24.
Zheng, T., T. Zhang, Q. Guan, W. Huang, Z. Feng, M. Song and C. Chen. Graph Neural Networks-based hybrid framework for predicting particle crushing strength. Expert Systems with Applications 255(2024). 124691.
Zhong, J., Zheng, J., Gao, L., Wu, Q., Guan, Z., Li, S., Wang, D., 2024. Study on micromechanical behavior and energy evolution of granular material generated by latent diffusion model under rotation of principal stresses. Particuology.
Zhong, X., Sun, W., 2018. An adaptive reduced-dimensional discrete element model for dynamic responses of granular materials with high-frequency noises. Int. J. Multiscale Comput. Eng. 16 (4).
Zhou, B., Li, C., Andrade, J.E., 2025a. Autonomous particle-shape-based classification and identification of calcareous soils through machine learning. Géotechnique 1–40.
Zhou, X., Q. Gong, Y. Liu and L. Yin. Automatic segmentation of TBM muck images via a deep-learning approach to estimate the size and shape of rock chips. Automation in Construction 126(2021). 103685.
Zhou, Y., Wang, H., Wu, B., Wang, L., Chen, X., 2025b. Physics informed neural network model for multi-particle interaction forces. Particuology 96, 126–138.
Zhou, Y., H. Zhou, T. Chen, C. Hu, Z. Liang, C. Zhao and F. Wang. Detection of rockfill gradation based on video image recognition. Automation in Construction 154(2023). 104965.
Zhou, Z., Bai, S., Chu, K., Li, J., Sun, J., Wang, M., Sun, Y., Li, M., Liu, Y., 2022. Calibration of DEM macro and micro parameters via XGBoost method. Granul. Matter 24 (4), 106.
Zhou, Z., Y. Zhen-Yu, G.-F. He, P. Zhang and M. Jiang. The potential of a multi-fidelity residual neural network based optimizer to calibrate DEM parameters of rock-like bonded granular materials. Computers and Geotechnics 168(2024). 106137.
Zienkiewicz, O.C., Taylor, R.L., Nithiarasu, P., Zhu, J., 1977. The finite element method. Elsevier.
Zou, Y., Ma, G., Zhang, Y., Zhou, W., Wang, Q., Chang, X., 2023. Identifying strain-dependent structural defects in granular materials from the hidden structure-plasticity relationship. Int. J. Solids and Struct. 276, 112332.