

Mechanistically informed data-driven modeling of cyclic plasticity via artificial neural networks

Daoping Liu^a, Hang Yang^a, K.I. Elkhodary^b, Shan Tang^{a,c,*}, Wing Kam Liu^{d,**}, Xu Guo^{a,c,*}

^a State Key Laboratory of Structural Analysis for Industrial Equipment, Department of Engineering Mechanics, Dalian University of Technology, Dalian, 116023, PR China

^b The Department of Mechanical Engineering, The American University in Cairo, 11835 New Cairo, Egypt

^c International Research Center for Computational Mechanics, Dalian University of Technology, 116023, PR China

^d Department of Mechanical Engineering, Northwestern University, 60208, USA

Received 1 November 2021; received in revised form 5 February 2022; accepted 14 February 2022

Available online 5 March 2022

Abstract

A mechanistically informed data-driven approach is proposed to simulate the complex plastic behavior of microstructured/homogenized solids subjected to cyclic loading, especially to simulate the Masing effect. Our proposed approach avoids the complicated mathematical construction of an appropriate yield surface, and does not require a large amount of data for training, by virtue of its mechanistic character, which couples the methods and tools of data science to the principles of mechanics. Specifically, a data-processing method is herein advanced to extract specific internal variables that characterize cyclic plastic behavior, which cannot be measured directly via physical experiments. A yield surface, represented by an artificial neural network (ANN), is then trained by stress–strain data and the extracted internal variables. Finally, the ANN is integrated into a finite element computational framework to solve different boundary value problems (BVPs). Results for demonstrative examples are presented, which illustrate the effectiveness and the reliability of the proposed approach for solids containing voids and particles in their microstructure. Compared with direct numerical simulation (DNS), our approach seems to predict the average levels of stress and plastic strain under cyclic loading more efficiently, as well as the regions of strain localization. In addition, results for a homogenized three-dimensional truss structure demonstrate that our approach can accurately describe the evolution of key internal variables. Our mechanistic approach requires much less data than the general pure data-driven methods, which shows a possible computational efficiency compared with the pure data-driven approach. Limitations of our proposed approach are also discussed.

© 2022 Elsevier B.V. All rights reserved.

Keywords: Data-driven; Constitutive law; Cyclic plasticity; Mechanistic; Artificial neural network (ANN)

1. Introduction

Load-bearing members made of elastoplastic materials, such as metals and alloys, are widely used in many engineering structures, which usually operate under cyclic loading conditions. How to describe the complex plastic

* Corresponding authors at: State Key Laboratory of Structural Analysis for Industrial Equipment, Department of Engineering Mechanics, Dalian University of Technology, Dalian, 116023, PR China.

** Corresponding author at: Department of Mechanical Engineering, Northwestern University, 60208, USA.

E-mail addresses: shantang@dlut.edu.cn (S. Tang), w-liu@northwestern.edu (W.K. Liu), guoxu@dlut.edu.cn (X. Guo).

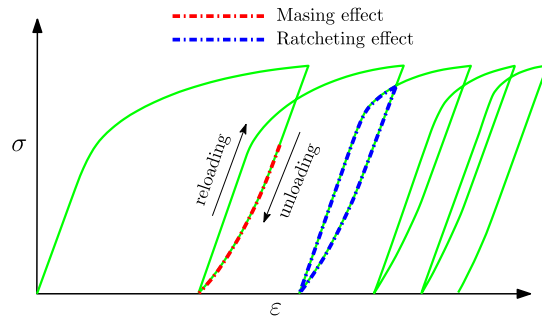


Fig. 1. Two specific effects observed in the experiments under the uniaxial cyclic loadings for many materials: Masing effect and ratcheting effect.

behavior of these materials has always caught the attention of researchers [1]. Fig. 1 shows a typical stress–strain curve for cyclic loading. The curve exhibits an incomplete elastic relationship when unloading, which is called the Masing effect [2]. A related ratcheting effect can also be noted from the curve [3,4]. To capture the Masing effect under cyclic loads, various constitutive models [5–14] have been proposed, which are based on seminal works on kinematic hardening [15,16]. In particular, to describe the Masing effect the yielding surface of the proposed models plays a central role, which can be used to judge the states of loading and unloading, and to characterize plastic evolution under cyclic loads. For instance, the subloading surface model, as a typical example [13,14], includes a normal-yield surface, and a subloading surface, which together control back-stress evolution and plastic deformation flow. Specifically, plastic behavior, under both loading and unloading, can be described by the translation, expansion, and contraction of the subloading surface, within the normal-yield surface.

Modern day materials are most frequently microstructured (with controlled voids or particles), purposely developed for and used in diverse engineering applications. To model cyclic plasticity for such complex materials, both the matrix and the microstructure are defined, and their mechanical behavior under cyclic loading is described. Thus, the average stress–strain response can be obtained through established homogenization methods, e.g. [17–21]. It is evident that the matrix and the microstructure may be described by differing yield surfaces and subyield surfaces. Their evolution during plastic deformation may also exhibit great differences, leading to a strongly nonlinear coupling between matrix and microstructure. Choosing the appropriate yield surface and subyield surface is therefore very important to correctly model cyclic plastic behavior (average behavior) for such microstructured solids. Nevertheless, it remains very difficult to accurately describe the microstructured solid by adding or subtracting yield surfaces, and/or by interpolating model parameters, leaving the matter open to active research.

With the growth of data science in recent years, numerous data-driven approaches have been proposed to replace traditional constitutive models and their associated difficulties [22]. Various neural networks were trained to represent a constitutive law, completely from stress–strain data [23–33]. In addition to using neural networks, Liu et al. proposed self-consistent clustering analysis (SCA) with the help of clustering technique in machine learning [34,35]. Based on SCA, FEM-cluster based analysis (FCA) [36,37] and virtual clustering analysis (VCA) [38] were further developed. In the same vein, without resorting to any constitutive function, Ortiz et al. proposed a pure (model-free) data-driven approach to solve boundary value problems (BVPs), directly respecting compatibility and equilibrium constraints simultaneously [39,40]. Specifically, for cyclic plasticity, many pure data-driven models have been proposed [41–44]. For instance, Mozaffar et al. [43] used data of 15,000 different cyclic loading/unloading paths to train a recurrent neural network (RNN) on loading-path dependent plasticity. Theoretically, artificial neural networks can be used to approximate arbitrarily complex functions, based on the universal approximation theorems [45]. For example, Shen et al. [46] gave the minimum number of neural network hidden layers and neurons to fit any continuous d-dimensional function. In other words, any constitutive model can always be replaced by a neural network trained with enough data. However, as the plastic behavior becomes more complex, the amount of data required to train a constitutive law, through a purely data-driven method, is expected to increase significantly, as the plastic behavior is a long-term, path (history) dependent phenomenon.

On the other hand, physical invariants and symmetries of mechanical quantities, hidden in the data, can be harnessed to reduce the required amount of training data. For instance, Tang et al. [47,48] harnessed the coaxial

characteristic of deviatoric strain and deviatoric stress to expand one-dimensional data sets to capture three-dimensional plastic behaviors. Likewise, Yang et al. [49,50] constructed a unified functional form for the yield condition, through an artificial neural network (ANN) trained by data from 22 loading paths, which they found sufficient to fully describe cyclic plastic behavior. Thus, the methods and tools coupling data science with mechanics principles, termed mechanistic data science, have very recently arisen, charting a new and promising methodology to solve engineering problems [22]. Mechanistic data science can effectively reduce data dimensionality when learning complex hidden relationships, by extracting mechanistic features from input data. Indeed to characterize the Masing effect via a pure data-driven approach would be quite challenging in terms of the volume of stress–strain data required for training, although this effect is a widely observed phenomenon in cyclic loading experiments [4,51–53].

In this paper, a mechanistic data-driven approach is therefore proposed to describe cyclic plasticity in microstructured solids, by leveraging existing constitutive model structures. This approach can accurately describe the incomplete elastic relationship between stress and strain during unloading. Theoretically, the proposed approach is possible to be applied to any existing constitutive model of cyclic plasticity. The subloading surface model [54] will be herein selected for demonstration purposes. Key physical quantities (that is, data on the mechanical features of the Masing effect) required by the subloading surface, such as the active stress and hardening function, cannot be measured directly in physical experiments. They are instead extracted from the measured stress–strain data sets. Then, this extracted data is used to train the ANN to obtain the desired yield function. Finally, the learned yield function is integrated into finite element computations. Compared with traditional function-based material models, the proposed data-driven approach avoids the calibration of material parameters involved and the construction of mathematical forms of the yield function. In addition, the proposed approach does not require a large amount of data, or a complex neural network structure, to obtain the yield function, as compared with the other pure data-driven methods discussed above.

The structure of this paper is as follows. In Section 2, data generation, data processing and ANN training are described. In Section 3, the effectiveness of the proposed approach is illustrated through finite element computations on some representative structures subjected to cyclic loading/unloading. Some concluding remarks are presented in Section 4. Note that all the tensors in this paper are recorded as vectors, using the Voigt notation.

2. Methodology

The mechanistic data-driven approach that we propose to describe cyclic plasticity in microstructured solids, can be illustrated as shown in Fig. 2. The data-driven approach is essentially divided into three parts: data generation, data processing (mechanistic feature extraction), and ANN training and learning. Data generation is herein treated by numerical experiments (as surrogates of physical experiments). First, however, we identify the key physical quantities that enable a successful characterization of cyclic plasticity, to define the mechanistic features to be extracted by our data-driven approach. Next, an RVE that contains the representative microstructure is created, to generate averaged stress–strain data. Then, the elastic material parameters of interest, such as Young's modulus and Poisson's ratio, can be identified from the RVE stress–strain data. Furthermore, to describe the cyclic plastic behavior, which includes the Masing and ratcheting effects, data for the subloading surface (active stress and hardening function) needs also to be extracted from the same stress–strain data. As such, a new data set is formed (extracted feature). Finally, a yield function expressed by ANN is obtained, as trained with the new data set via machine learning. Solutions to BVPs can finally be driven by this learned yield function, along with a loading/unloading criterion.

2.1. Mechanistic features of cyclic plasticity

To extract correct characteristic variables for cyclic plasticity, it is necessary to understand the important role of internal variables in the constitutive modeling of cyclic plasticity. According to our investigation, most elastoplastic material models in commercial software may fail in describing the Masing effect. Conversely, Hashiguchi's subloading surface model [54] succeeds, by proposing three hypersurfaces in stress space: a normal-yield surface, a subloading surface, and an elastic-core surface (see Fig. 3). It will therefore be implemented and employed in our study.

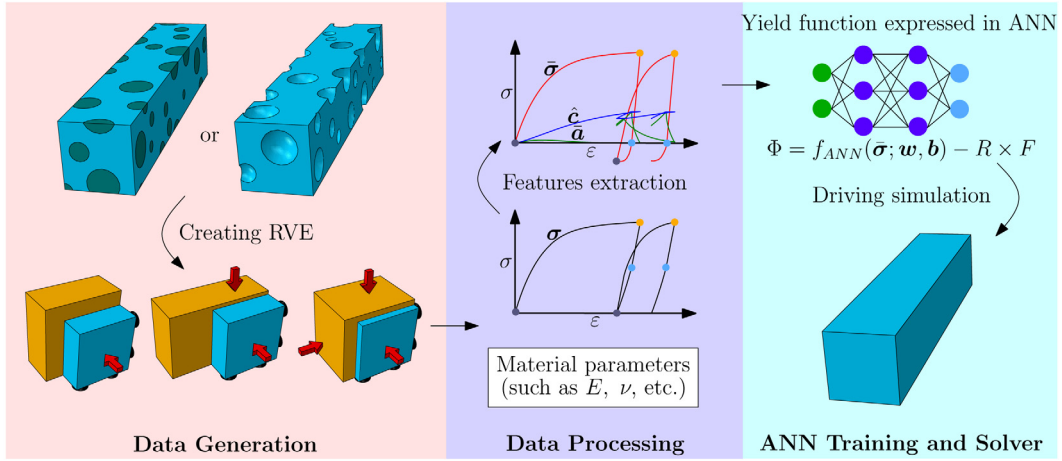


Fig. 2. A summary of the proposed data-driven approach to solve the cyclic plasticity problem of microstructured materials.

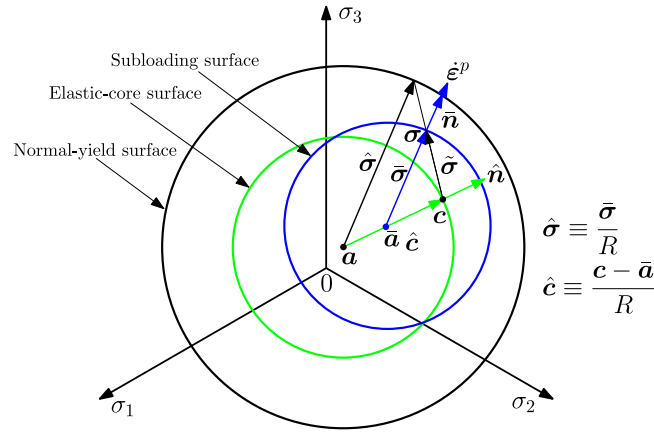


Fig. 3. Normal-yield, subloading and elastic-core surfaces of the subloading surface model.

The normal-yield surface is selected with isotropic and kinematic hardening, as adopted in many cyclic plasticity models (e.g., in ABAQUS), which may be given as,

$$f(\hat{\sigma}) - F = 0 \quad (1)$$

where $\hat{\sigma} = \sigma - a$. Here σ is the Cauchy stress, and a is the back stress (kinematic hardening variable), first proposed by Prager [15] to describe the anisotropy caused by plastic hardening. $f(\hat{\sigma})$ is a homogeneous function of stress $\hat{\sigma}$. $F(\geq 0)$ is a function of the accumulated plastic strain, introduced to describe isotropic hardening or softening, i.e. the expansion or contraction of the yielding surface. F usually takes the following form [54],

$$F = F_0 [1 + h_1 (1 - e^{-h_2 H})] \quad (2)$$

where F_0 is the initial value for isotropic hardening, and h_1 and h_2 are material constants. $H = \sqrt{\frac{2}{3}} \int \|\dot{\epsilon}^p\| dt$ is the accumulated plastic strain, in which $\dot{\epsilon}^p$ is the plastic strain rate.

To capture the Masing effect, a subloading surface that lies within the normal-yield surface in stress-space is introduced (see the blue circle in Fig. 3). This subloading surface is defined as,

$$f(\bar{\sigma}) - R \times F = 0 \quad (3)$$

where $\bar{\sigma} = \sigma - \bar{a}$ is the active stress, and $\bar{a}$ is the geometric center of the subloading surface. R ($0 \leq R \leq 1$) defines the ratio of the subloading surface size to the normal-yield surface size. R serves as a measure of the extent by

which the subloading surface approaches the normal-yield surface. The symbol ' $\times$ ' represents the scalar product, emphasizing that R and F represent two physical quantities, respectively. It should be noted that the subloading surface is always contained within the normal-yield surface, and maintains geometric similarity to it. The plastic strain rate is then derived based on this subloading surface.

To help establish a connection between back stresses $\mathbf{a}$ and $\bar{\mathbf{a}}$, an elastic-core surface is further introduced, as shown by the green circle in Fig. 3. The geometric relationships between the normal-yield surface ($\bar{\sigma}$), the subloading surface ($\hat{\sigma}$), and the elastic-core surface ($\hat{c}$), lead to the following expressions for the associated stresses,

$$\begin{cases} \hat{\sigma} \equiv \bar{\sigma}/R \\ \hat{c} \equiv \mathbf{c} - \mathbf{a} = (\mathbf{c} - \bar{\mathbf{a}})/R \end{cases} \quad (4)$$

where $\mathbf{c}$ designates any similarity-center that arises between the normal-yield surface and the subloading surface. $\mathbf{c}$ has been physically interpreted as any stress-point where deformation is most elastic, and has thus been termed the elastic-core, or the elastic-center. Correspondingly, an elastic-core surface that passes through each elastic-core $\mathbf{c}$, while maintaining a similar shape and orientation to the normal-yield surface with the evolution of the kinematic-hardening variable $\mathbf{a}$, may be defined as,

$$f(\hat{\mathbf{c}}) - \Re \times F = 0 \quad (5)$$

where $\Re$ ($0 \leq \Re \leq \chi$) is the ratio of the size of the elastic-core surface to that of the normal-yield surface, and χ (< 1) is a material constant.

When the material experiences plastic loading, the plastic strain rate $\dot{\epsilon}^p$ is not equal to $\mathbf{0}$. The evolution equations of R , $\mathbf{a}$, $\mathbf{c}$ and $\bar{\mathbf{a}}$ are as follows [55],

$$\begin{cases} \dot{R} = r_1 e^{r_2 \Re(\bar{\mathbf{n}}:\dot{\mathbf{n}})} \ln(R) \|\dot{\epsilon}^p\| \\ \dot{\mathbf{a}} = C_a \left(\bar{\mathbf{n}} - \sqrt{\frac{2}{3}} \frac{\mathbf{a}}{r_a F} \right) \|\dot{\epsilon}^p\| \\ \dot{\mathbf{c}} = C_c \left(\bar{\mathbf{n}} - \frac{\Re}{\chi} \hat{\mathbf{n}} \right) \|\dot{\epsilon}^p\| + \sqrt{\frac{2}{3}} \frac{dF/dH}{F} \hat{\mathbf{c}} \|\dot{\epsilon}^p\| \\ \dot{\bar{\mathbf{a}}} = (1 - R)\dot{\mathbf{c}} - \mathbf{c}\dot{R} + R\dot{\mathbf{a}} + \mathbf{a}\dot{R} \end{cases} \quad (6)$$

where $\bar{\mathbf{n}} = \bar{\sigma}'/\|\bar{\sigma}'\|$, and $\bar{\sigma}'$ is the deviatoric part of $(\bar{\sigma} - \mathbf{c})$. $\bar{\mathbf{n}} = \frac{\partial f/\partial \bar{\sigma}}{\|\partial f/\partial \bar{\sigma}\|}$ and $\hat{\mathbf{n}} = \frac{\partial f/\partial \hat{c}}{\|\partial f/\partial \hat{c}\|}$ are the normal tensors of the subloading surface and elastic-core surface, respectively. r_1 , r_2 , C_a , r_a , C_c are material constants. The ratio of the size of the elastic-core surface to that of the normal-yield surface $\Re$ can be updated by Eq. (5). A practical method to compute the plastic strain rate is detailed in Section 2.5.

Conversely, when the material experiences elastic loading, elastic unloading, or neutral loading, the plastic strain rate $\dot{\epsilon}^p$ is equal to $\mathbf{0}$. The isotropic hardening variable F , the back stress $\mathbf{a}$ and the geometric center of the subloading surface $\bar{\mathbf{a}}$ remain unchanged. R , $\hat{\mathbf{c}}$ and $\Re$ can be updated according to Eqs. (3), (4), (5),

$$\begin{cases} R = \frac{f(\bar{\sigma})}{F} \\ \mathbf{c} = \frac{\bar{\mathbf{a}} - R\mathbf{a}}{1 - R} \\ \Re = \frac{f(\mathbf{c} - \mathbf{a})}{F} \end{cases} \quad (7)$$

In summary, the subloading surface is key to describing the complex behavior of materials under cyclic loading. Its expansion reflects plastic loading, and the flow of plastic deformation. It should be emphasized that the internal variables, such as R , F and $\bar{\sigma}$, are very important to describe this complex behavior, though they are difficult to measure directly in physical experiments.

2.2. Data generation

An RVE with microstructure (voids or particles) is adopted to generate the required stress-strain data by numerical experiments. To account for the Masing effect by the data generated, the subloading surface model is employed as the material law of the constituents of the microstructured RVE. As shown in Fig. 4, an RVE can

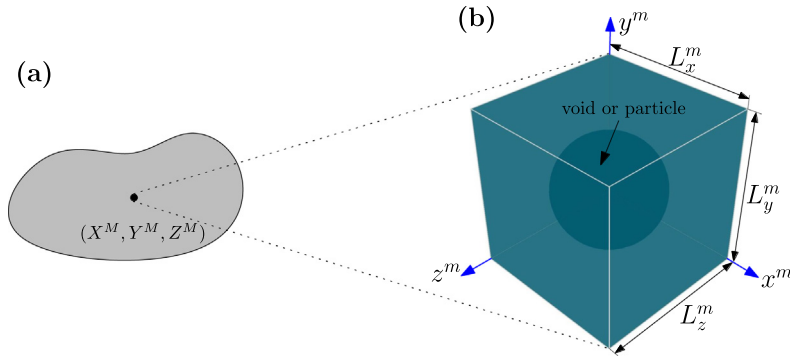


Fig. 4. (a) A material point (X^M, Y^M, Z^M) with microstructure in the deformable body (initial configuration). (b) The RVE with microstructure (void or particle) for data generation.

be thought of as being attached to a macro-scale material point (X^M, Y^M, Z^M) of a continuum. Here, and in the following, superscripts “ m ” and “ M ” represent quantities at the meso-scale (RVE) and the macro-scale (parent domain) respectively.

The RVE is herein taken as a cubic cell, which we associate with a Cartesian coordinate system and principal axes x^m , y^m and z^m . The base unit vectors are $\{\mathbf{e}_1^m, \mathbf{e}_2^m, \mathbf{e}_3^m\}$ respectively, and the RVE lengths are L_x^m , L_y^m and L_z^m , measured along the x^m , y^m and z^m axes respectively.

The boundary conditions on the RVE are given by

$$\begin{cases} u_1^m = 0 & \text{on } x^m = 0 \\ u_2^m = 0 & \text{on } y^m = 0 \\ u_3^m = 0 & \text{on } z^m = 0 \\ u_1^m = (\lambda_1 - 1)L_x^m & \text{on } x^m = L_x^m \\ u_2^m = (\lambda_2 - 1)L_y^m & \text{on } y^m = L_y^m \\ u_3^m = (\lambda_3 - 1)L_z^m & \text{on } z^m = L_z^m \end{cases} \quad (8)$$

where u_1^m , u_2^m and u_3^m are the components of the displacement vector $\mathbf{u}^m$ in the principal space, $\mathbf{u}^m = u_1^m \mathbf{e}_1^m + u_2^m \mathbf{e}_2^m + u_3^m \mathbf{e}_3^m$. λ_1 , λ_2 and λ_3 are the principal stretches in the x^m , y^m and z^m directions respectively, which are chosen as the loading parameters.

As clarified in our previous work [50], the data along 22 loading paths is sufficient to train the yield surface in stress space. The principal stretches on the RVE (the loading parameters) along the 22 loading paths are given by,

$$\{\lambda_1, \lambda_2, \lambda_3\}^{[\alpha, \beta]} = 1 + \eta \mathbf{P}_\alpha \frac{\beta}{N_S} \quad (\alpha = 1, \dots, N_L; \beta = 1, \dots, N_S) \quad (9)$$

where α and β represent the loading path number and the loading step respectively; N_L and N_S (taken as 22 and 1000) designate the maximum values of α and β , respectively. η (taken as 10%) is a coefficient that defines the range of strain for the given data set, and $\mathbf{P}_\alpha$ (as shown in Table 1) represents the directions of loading paths α (described in principal stress space). Based on homogenization theory [17,18], the averaged principal strain of an RVE can be obtained as

$$\{\varepsilon_1^M, \varepsilon_2^M, \varepsilon_3^M\}^{[\alpha, \beta]} = \{\lambda_1 - 1, \lambda_2 - 1, \lambda_3 - 1\}^{[\alpha, \beta]} \quad (10)$$

The corresponding homogenized (averaged) stress $\boldsymbol{\sigma}^M$ can be expressed by,

$$\boldsymbol{\sigma}^M = \sigma_1^M \mathbf{e}_1^m \otimes \mathbf{e}_1^m + \sigma_2^M \mathbf{e}_2^m \otimes \mathbf{e}_2^m + \sigma_3^M \mathbf{e}_3^m \otimes \mathbf{e}_3^m, \quad (11)$$

Table 1The loading directions $\mathbf{P}_\alpha$ used in data generation.

path ID(α)	$\mathbf{P}_\alpha$	path ID(α)	$\mathbf{P}_\alpha$	path ID(α)	$\mathbf{P}_\alpha$
1	{-1.00, 0.00, 0.00}	9	{-0.75, -0.5, 0.43}	17	{-0.43, 0.50, 0.75}
2	{-0.87, -0.50, 0.00}	10	{-0.75, -0.43, 0.50}	18	{-0.25, 0.43, 0.87}
3	{-0.87, -0.43, -0.25}	11	{-0.75, 0.43, 0.50}	19	{0.00, 0.00, 1.00}
4	{-0.87, -0.43, 0.25}	12	{-0.50, -0.43, 0.75}	20	{0.00, 0.50, 0.75}
5	{-0.87, -0.25, 0.43}	13	{-0.50, 0.00, 0.87}	21	{0.25, 0.43, 0.87}
6	{-0.87, 0.00, 0.50}	14	{-0.50, 0.43, 0.75}	22	{0.43, 0.50, 0.75}
7	{-0.87, 0.25, 0.43}	15	{-0.43, -0.25, 0.87}		
8	{-0.75, -0.50, -0.43}	16	{-0.43, 0.25, 0.87}		

Table 2

The material parameters of the matrix.

E	ν	F_0	h_1	h_2	r_1	r_2	C_a	r_a	C_c	χ
200.0	0.3	1.0	0.2	20.0	100.0	1.0	5.0	0.5	15.0	0.6

where $\otimes$ is dyadic symbol for vectors. The principal components (σ_1^M , σ_2^M and σ_3^M) of the average stress $\boldsymbol{\sigma}^M$ can be computed from,

$$\begin{cases} \sigma_1^M = \frac{1}{L_y^m L_z^m} \int_0^{L_y^m} \int_0^{L_z^m} t_1^m dy^m dz^m \\ \sigma_2^M = \frac{1}{L_x^m L_z^m} \int_0^{L_x^m} \int_0^{L_z^m} t_2^m dx^m dz^m \\ \sigma_3^M = \frac{1}{L_x^m L_y^m} \int_0^{L_x^m} \int_0^{L_y^m} t_3^m dx^m dy^m, \end{cases} \quad (12)$$

which is based on the traction force acting on the outer boundary, where t_1^m , t_2^m , and t_3^m are the components of the traction force $\mathbf{t}^m$, written in principal space (the component form of the traction force can thus be expressed as $\mathbf{t}^m = t_1^m \mathbf{e}_1^m + t_2^m \mathbf{e}_2^m + t_3^m \mathbf{e}_3^m$). Based on this data generation method, the average principal stress and the average principal strain for any RVE are obtained, at each loading step, and along all loading paths (Table 1), using Eqs. (10) and (12). That is, data sets for average principal strain of the form $\{\varepsilon_1^M, \varepsilon_2^M, \varepsilon_3^M\}^{[\alpha, \beta]}$, and for average principal stress of the form $\{\sigma_1^M, \sigma_2^M, \sigma_3^M\}^{[\alpha, \beta]}$, are created for further processing.

Fig. 5 shows the stress-strain data generated along three loading paths ($\alpha = 1, 6$ or 18) as an example. In this example, an RVE with a void at the center is illustrated. The void volume fraction in the RVE is 15.45%. The matrix is assumed to be an isotropic elastoplastic material, as described by the subloading surface model with material parameters as listed in Table 2. To investigate the effect of the void in the RVE, Fig. 5 compares the stress-strain predictions for a homogeneous bulk matrix RVE (without void), under the same conditions. As can be seen, the RVE without a void predicts very different results, especially under triaxial loading (Loading path ID=18). This finding suggests that one cannot approximate the behavior of the RVE with a void directly from a homogeneous RVE without void (described by the matrix material only).

From hereon, unless otherwise stated, all the data that concerns physical quantities is homogenized for ANN training. As such, the superscript “ M ” of these physical quantities is omitted, to avoid cluttering notation. Note that only the stress-strain data for the loading process is herein generated. This feature of our work compares favorably with previous works, which require to generate stress-strain data along approximately 15,000 loading and unloading paths (instead of our mere 22 loading paths), to train the material law for cyclic plasticity, e.g., [43].

2.3. Feature extraction for cyclic effects

Clearly, the generated principal stress-strain data sets, $\{\varepsilon_1, \varepsilon_2, \varepsilon_3\}^{[\alpha, \beta]}$ and $\{\sigma_1, \sigma_2, \sigma_3\}^{[\alpha, \beta]}$, are not adequate to properly train an elastoplastic constitutive model for cyclic plasticity, since the internal variables associated with these data sets need to be also identified. Since the average elastoplastic behavior of the microstructured solid can

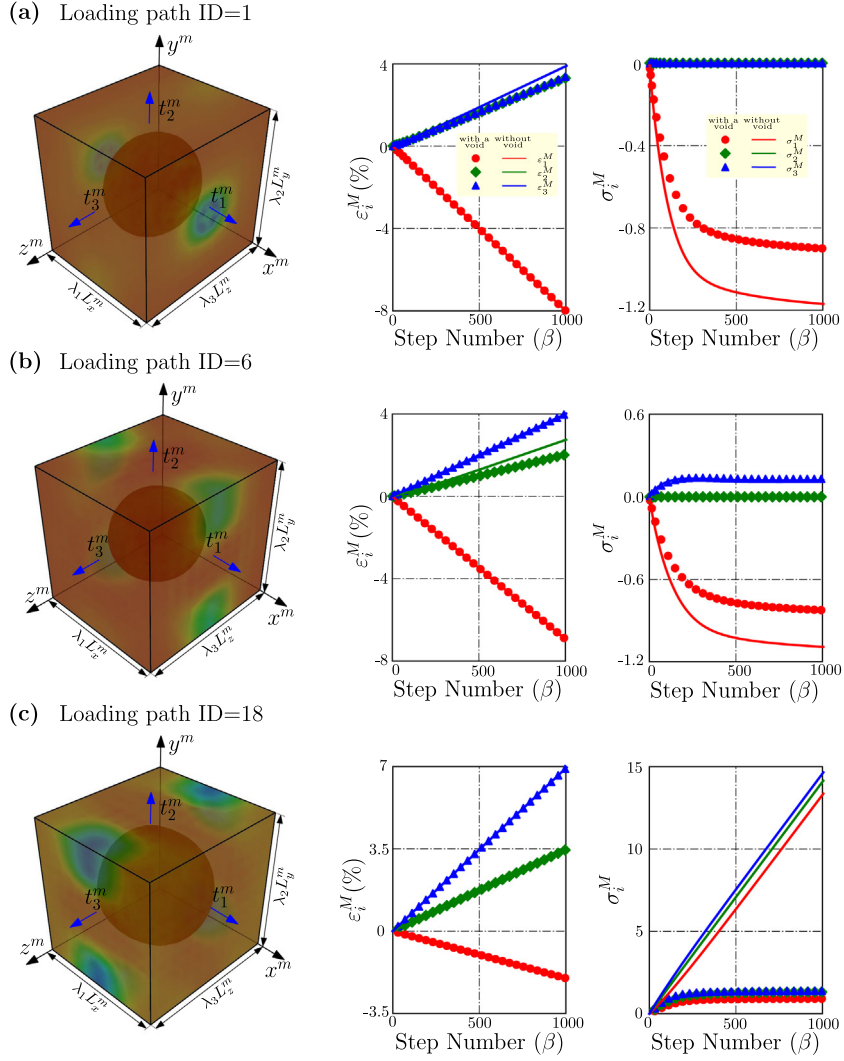


Fig. 5. The visualization of data generation of average principal strain $\{\varepsilon_1^M, \varepsilon_2^M, \varepsilon_3^M\}$ and average principal stress $\{\sigma_1^M, \sigma_2^M, \sigma_3^M\}$ through the RVE for (a) loading path ID 1; (b) loading path ID 6; (c) loading path ID 18 in Table 1. The material parameters of the matrix are shown in Table 2. The response of the homogeneous bulk matrix RVE under the same conditions is plotted by solid lines.

also be fairly described by a certain material model, its associated internal variables can be extracted by mapping to this model the available principal stress–strain data. Basically, this introduction of a specific material model for macro-scale cyclic plasticity, into our data-driven approach, is an especially expedient feature. On one hand it helps extract a set of internal variables (data for mechanistic features) required to characterize cyclic plasticity. On the other hand, the amount of data required for subsequent ANN training (on long-term material behavior) is vastly reduced, as noted earlier.

We will herein consider for the microstructured solid the same subloading surface model that we did for each constituent, for its adequacy. Thus, the set $\{\bar{\sigma}, R, F\}$ describes the evolution of the subloading surface for the microstructure, which is important to describe its cyclic plastic behavior. Again, these variables are not directly obtainable like the stress and the strain are. Hence, in the following sub-section, a two-step approach is proposed to generate new data sets that are based on the available stress–strain data sets, with the help of the subloading surface model. These new data sets will be used to train the ANN-based elastoplastic model for cyclic loading.

2.3.1. Step I: Identifying macro-scale material parameters

The Young's modulus and Poisson's ratio are identified first. The elastic stress-strain relationship (Hooke's law) written in principal directions for an isotropic elastic material is,

$$\check{\sigma}_i = D_{ij} \varepsilon_j \quad (i, j = 1, 2, 3) \quad (13)$$

where D_{ij} is the elastic stiffness matrix:

$$D_{ij} = \frac{E}{(1+\nu)(1-2\nu)} \begin{bmatrix} 1-\nu & \nu & \nu \\ \nu & 1-\nu & \nu \\ \nu & \nu & 1-\nu \end{bmatrix} \quad (14)$$

where E and ν are the Young's modulus and Poisson's ratio respectively, which can be obtained by a least squares fit (to ensure the data is in the elastic stage, only the data of steps $\beta = 1, 2$ are taken),

$$\min_{E, \nu} \sum_{\alpha=1}^{N_L} \sum_{\beta=1}^2 \sum_{i=1}^3 \left(\check{\sigma}_i^{[\alpha, \beta]} - \sigma_i^{[\alpha, \beta]} \right)^2 \quad (15)$$

For complex elastic behaviors, Young's modulus E and Poisson's ratio ν can be calculated by adding a few steps of measured data (that is, $\beta > 2$). Such as the nonlinear elastic problem, this has been carried out in our previous works [49].

The strain increment $\Delta \varepsilon_i^{[\alpha, \beta]}$ and stress increment $\Delta \sigma_i^{[\alpha, \beta]}$ for any step β ($\beta \geq 2$) can be computed using the data from steps $\beta - 1$ and β ,

$$\begin{cases} \Delta \varepsilon_i^{[\alpha, \beta]} = \varepsilon_i^{[\alpha, \beta]} - \varepsilon_i^{[\alpha, \beta-1]} \\ \Delta \sigma_i^{[\alpha, \beta]} = \sigma_i^{[\alpha, \beta]} - \sigma_i^{[\alpha, \beta-1]} \end{cases} \quad (16)$$

Combined with Hooke's law, the plastic strain increment $\Delta \varepsilon_i^{p[\alpha, \beta]}$ can be expressed as

$$\Delta \varepsilon_i^{p[\alpha, \beta]} = \Delta \varepsilon_i^{[\alpha, \beta]} - D_{ij}^{-1} \Delta \sigma_j^{[\alpha, \beta]} \quad (17)$$

The accumulated plastic strain $H^{[\alpha, \beta]}$ and its increment $\Delta \bar{\varepsilon}^{p[\alpha, \beta]}$ can be computed by

$$\begin{cases} \Delta \bar{\varepsilon}^{p[\alpha, \beta]} = \sqrt{\frac{2}{3} \sum_{j=1}^3 \left(\Delta \varepsilon_j^{p[\alpha, \beta]} - \frac{1}{3} \sum_{i=1}^3 \Delta \varepsilon_i^{p[\alpha, \beta]} \right)^2} \\ H^{[\alpha, \beta]} = \sum_{i=1}^{\beta} \Delta \bar{\varepsilon}^{p[\alpha, i]} \end{cases} \quad (18)$$

We can define the data set $\mathbb{X}$ describing the elastic stage according to the accumulated plastic strain,

$$\mathbb{X} := \{[\alpha, \beta] | H^{[\alpha, \beta]} \leq 0\} \quad (19)$$

According to Eq. (2), F_0 and $F_0 + h_1$ take on their values when the accumulated plastic strain H takes on its minimum and maximum values within hardening function F , respectively. They can be obtained in the elastic stage and when $\beta = N_s$. Therefore, the values of F_0 and h_1 can be obtained through an optimization process as,

$$\begin{cases} \min_{F_0} \sum_{\alpha=1}^{N_L} \left(\max_{[\alpha, \beta] \in \mathbb{X}} \sigma_e^{[\alpha, \beta]} - F_0 \right)^2 \\ \min_{h_1} \sum_{\alpha=1}^{N_L} \left(\sigma_e^{[\alpha, N_s]} - F_0 - h_1 \right)^2 \end{cases} \quad (20)$$

where $\sigma_e^{[\alpha, \beta]} = \sqrt{\frac{3}{2} \sum_{j=1}^3 \left(\sigma_j^{[\alpha, \beta]} - \frac{1}{3} \sum_{i=1}^3 \sigma_i^{[\alpha, \beta]} \right)^2}$ is the equivalent stress, and $\max_{[\alpha, \beta] \in \mathbb{X}} \sigma_e^{[\alpha, \beta]}$ represents the maximum value of σ_e under loading path α in the data set $\mathbb{X}$.

As an important and welcome simplification, we note that all other material parameters, i.e., $r_1, r_2, C_a, r_a, C_c, h_2$ and χ , in the subloading surface model, are not required for our method to be accurate. This claim is verified by our

numerical examples shown in [Appendix A](#). This observation implies that we may use the material parameters of the matrix for our microstructured solid, so long as the microstructure's behavior follows similar trends in evolution, as expected for typical alloys.

2.3.2. Step II: Extracting macro-scale mechanical features

The initial values of internal variables are assumed to be,

$$\begin{cases} \Delta \boldsymbol{\varepsilon}^{p[\alpha,1]} = [1, 0, 0]^T \\ \bar{\mathbf{n}}^{[\alpha,1]} = \hat{\mathbf{n}}^{[\alpha,1]} = [1, 0, 0]^T \\ \mathbf{a}^{[\alpha,1]} = \mathbf{c}^{[\alpha,1]} = \bar{\mathbf{a}}^{[\alpha,1]} = [0, 0, 0]^T \\ \bar{\boldsymbol{\sigma}}^{[\alpha,1]} = [\sigma_1^{[\alpha,1]}, \sigma_2^{[\alpha,1]}, \sigma_3^{[\alpha,1]}]^T \\ R^{[\alpha,1]} = \mathfrak{N}^{[\alpha,1]} = 10^{-16} \end{cases} \quad (21)$$

All internal variables are functions of the plastic strain increment $\Delta \boldsymbol{\varepsilon}^{p[\alpha,\beta]}$. It is also assumed that an associative plastic flow rule is followed. Therefore, in combination with Eqs. (2), (6), (18), the internal variables can be computed by,

$$\begin{cases} H^{[\alpha,\beta]} = H^{[\alpha,\beta-1]} + \Delta \bar{\varepsilon}^{p[\alpha,\beta]} \\ F^{[\alpha,\beta]} = F_0 \left\{ 1 + h_1 \left[1 - e^{-h_2 H^{[\alpha,\beta]}} \right] \right\} \\ F'^{[\alpha,\beta]} = F_0 h_1 h_2 e^{-h_2 H^{[\alpha,\beta]}} \\ \bar{\mathbf{n}}^{[\alpha,\beta]} = \frac{\Delta \boldsymbol{\varepsilon}^{p[\alpha,\beta]}}{\Delta \bar{\varepsilon}^{p[\alpha,\beta]}} \\ \mathbf{a}^{[\alpha,\beta]} = \mathbf{a}^{[\alpha,\beta-1]} + C_a \left(\bar{\mathbf{n}}^{[\alpha,\beta]} - \sqrt{\frac{2}{3}} \frac{\mathbf{a}^{[\alpha,\beta-1]}}{r_a F^{[\alpha,\beta]}} \right) \Delta \bar{\varepsilon}^{p[\alpha,\beta]} \\ \mathbf{c}^{[\alpha,\beta]} = \mathbf{c}^{[\alpha,\beta-1]} + C_c \left(\bar{\mathbf{n}}^{[\alpha,\beta]} - \frac{\mathfrak{N}^{[\alpha,\beta]} \hat{\mathbf{n}}^{[\alpha,\beta-1]}}{\chi} \right) \Delta \bar{\varepsilon}^{p[\alpha,\beta]} + \Delta \mathbf{a}^{[\alpha,\beta]} + \sqrt{\frac{2}{3}} \frac{F'^{[\alpha,\beta]}}{F^{[\alpha,\beta]}} (\mathbf{c}^{[\alpha,\beta-1]} - \mathbf{a}^{[\alpha,\beta]}) \Delta \bar{\varepsilon}^{p[\alpha,\beta]} \\ \hat{\mathbf{n}}^{[\alpha,\beta]} = \frac{\Delta \mathbf{c}^{[\alpha,\beta]}}{\sqrt{\frac{3}{2} \text{dev}(\Delta \mathbf{c}^{[\alpha,\beta]}) : \text{dev}(\Delta \mathbf{c}^{[\alpha,\beta]})}} \\ \mathfrak{N}^{[\alpha,\beta]} = \frac{\|\mathbf{c}^{[\alpha,\beta]} - \mathbf{a}^{[\alpha,\beta]}\|}{\|\boldsymbol{\sigma}^{[\alpha,\beta]} - \bar{\mathbf{a}}^{[\alpha,\beta]}\|} \\ R^{[\alpha,\beta]} = R^{[\alpha,\beta-1]} - r_1 e^{\left[r_2 \mathfrak{N}^{[\alpha,\beta]} \left(\hat{\mathbf{n}}^{[\alpha,\beta]} : \frac{\bar{\boldsymbol{\sigma}}'}{\|\bar{\boldsymbol{\sigma}}'\|} \right) \right]} \ln([\alpha,\beta-1]) \Delta \bar{\varepsilon}^{p[\alpha,\beta]} \\ \bar{\mathbf{a}}^{[\alpha,\beta]} = \bar{\mathbf{a}}^{[\alpha,\beta-1]} + (1 - R^{[\alpha,\beta]}) \Delta \mathbf{c}^{[\alpha,\beta]} - \mathbf{c}^{[\alpha,\beta]} \Delta R^{[\alpha,\beta]} + R^{[\alpha,\beta]} \Delta \mathbf{a}^{[\alpha,\beta]} + \mathbf{a}^{[\alpha,\beta]} \Delta R^{[\alpha,\beta]} \\ \bar{\boldsymbol{\sigma}}^{[\alpha,\beta]} = \boldsymbol{\sigma}^{[\alpha,\beta]} - \bar{\mathbf{a}}^{[\alpha,\beta]} \end{cases} \quad (22)$$

where $\Delta \boldsymbol{\varepsilon}^{p[\alpha,\beta]} = [\Delta \varepsilon_1^{p[\alpha,\beta]}, \Delta \varepsilon_2^{p[\alpha,\beta]}, \Delta \varepsilon_3^{p[\alpha,\beta]}]^T$, $\boldsymbol{\sigma}^{[\alpha,\beta]} = [\sigma_1^{[\alpha,\beta]}, \sigma_2^{[\alpha,\beta]}, \sigma_3^{[\alpha,\beta]}]^T$, and $\text{dev}(\Delta \mathbf{c}^{[\alpha,\beta]})$ is the deviatoric part of $\Delta \mathbf{c}^{[\alpha,\beta]}$. In this manner, the new data sets $\{\bar{\boldsymbol{\sigma}}^{[\alpha,\beta]}, R^{[\alpha,\beta]}, F^{[\alpha,\beta]}\}$ are obtained, which are used to train the subloading surface. It should be noted here that this data processing strategy is very similar to the update procedures of stress and of the internal variables in a typical finite element computation. However, the plastic strain increment at each time step is in our strategy known, unlike for typical incremental plasticity in finite elements, which solves for the plastic increment at each time step by a nonlinear analysis. Thus, in our case it is very straightforward and fast to obtain the new data sets using Eq. (22).

2.4. ANN training and learning

Since the subloading surface plays the key role in describing the Masing effect in cyclic loading, the ANN is herein designed to describe it mathematically. The ANN is then trained by the processed data sets, $\{\bar{\boldsymbol{\sigma}}^{[\alpha,\beta]}, R^{[\alpha,\beta]}, F^{[\alpha,\beta]}\}$. With the trained subloading surface, the normal-yield surface and elastic-core surface can then be described mathematically to completely replace the overall subloading surface model, and to drive subsequent online finite element computations.

According to Eq. (3), the subloading surface can be rewritten as a function $\hat{\Phi}$ of the following form,

$$\hat{\Phi} = f(\bar{\boldsymbol{\sigma}}_m, \bar{J}_2, \bar{J}_3) - R \times F \quad (23)$$

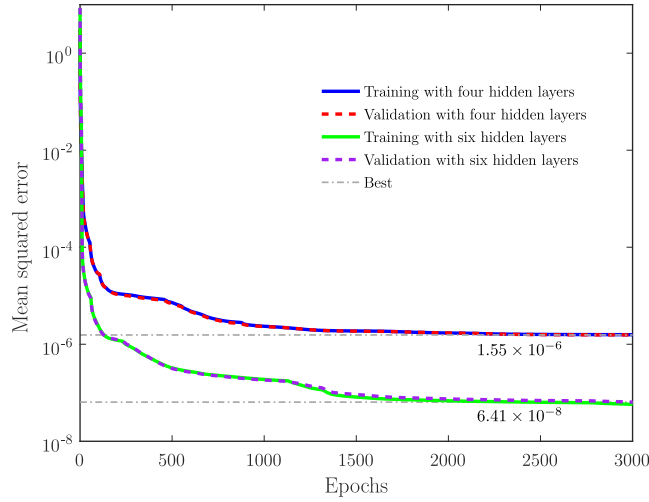


Fig. 6. The effect of the number of hidden layers on training performance of ANN. The variation of training and validation loss with four and six hidden layers are plotted. The best validation performance of the ANN is also marked.

where $\bar{\sigma}_m$, $\bar{J}_2$ and $\bar{J}_3$ are three invariants of $\bar{\sigma}$ defined as,

$$\begin{cases} \bar{\sigma}_m = \bar{\sigma}_{kk}/3 \\ \bar{J}_2 = \sqrt{\bar{\sigma}'_{ij}\bar{\sigma}'_{ij}} \\ \bar{J}_3 = \sqrt[3]{\bar{\sigma}'_{ik}\bar{\sigma}'_{kj}\bar{\sigma}'_{ij}} \end{cases} \quad (24)$$

where $\bar{\sigma}'_{ij}$ is a component of the deviatoric part of $\bar{\sigma}$. Practically, we only need data for which the accumulated plastic strain is greater than zero (that is, $F > F_0$) when training the ANN. Therefore, a corresponding data set $\mathbb{Y}$ in the plastic range may be defined as,

$$\mathbb{Y} := \{[\alpha, \beta] | F^{[\alpha, \beta]} > F_0\} \quad (25)$$

The next step is to obtain a new function $\bar{\Phi}$ determined by the ANN, which can approximate $\hat{\Phi}$. To compare the impact of the hidden layers on ANN performance, the ANN training performance with four and six hidden layers is investigated. 70%, 15%, and 15% of the data sets are randomly selected as training, validation, and test sets, respectively. We take the loss function of the artificial neural network as the mean square error. The mean square error for the ANN with four hidden layers and six hidden layers vs. their training epochs is shown in Fig. 6. After 3000 epochs, the mean square errors are approximately 1.55×10^{-6} with four hidden layers and 6.41×10^{-8} with six hidden layers. Therefore, as the number of hidden layers increases, the value of the loss function for the trained ANN is smaller.

Therefore, we create an ANN that includes an input layer, three hidden layers, and an output layer, as shown in Fig. 7. The input layer and the output layer have three neurons and a single neuron respectively. The number of neurons in each hidden layer is six. The new function $\bar{\Phi}$ with inputs $(\bar{\sigma}_m, \bar{J}_2, \bar{J}_3)$, outputs $(R \times F)$, weights $(\mathbf{w})$ and biases $(\mathbf{b})$ represented by ANN is expressed as follows,

$$\bar{\Phi} = f_{ANN}(\bar{\sigma}_m, \bar{J}_2, \bar{J}_3; \mathbf{w}, \mathbf{b}) - R \times F \quad (26)$$

where f_{ANN} is the function to be trained and can be written:

$$f_{ANN} = \tanh(\tanh(\tanh([\bar{\sigma}_m, \bar{J}_2, \bar{J}_3]\mathbf{w}^2 + \mathbf{b}^2)\mathbf{w}^3 + \mathbf{b}^3)\mathbf{w}^4 + \mathbf{b}^4)\mathbf{w}^5 + \mathbf{b}^5 \quad (27)$$

here, $\mathbf{w}^i$ and $\mathbf{b}^i$ ($i = 2, 3, 4, 5$) are the weights and biases for the link between $(i - 1)$ th layer and i th layer, respectively. The hyperbolic tangent function $\tanh(\cdot)$ is used as the activation function. The weights and biases of

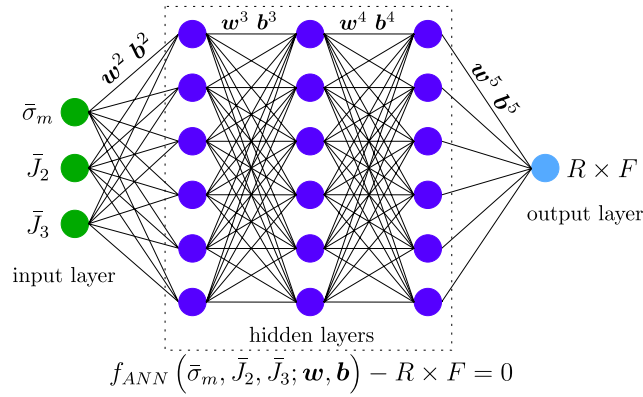


Fig. 7. The ANN used in the proposed approach. The inputs, output, weights and biases are all marked.

the ANN can be obtained by solving the following minimization problem,

$$\min_{\mathbf{w}, \mathbf{b}} \sum_{[\alpha, \beta] \in \mathbb{Y}} \left[f_{ANN}(\bar{\sigma}_m^{[\alpha, \beta]}, \bar{J}_2^{[\alpha, \beta]}, \bar{J}_3^{[\alpha, \beta]}; \mathbf{w}, \mathbf{b}) - R^{[\alpha, \beta]} \times F^{[\alpha, \beta]} \right]^2 \quad (28)$$

where $\bar{\sigma}_m^{[\alpha, \beta]}$, $\bar{J}_2^{[\alpha, \beta]}$ and $\bar{J}_3^{[\alpha, \beta]}$ are the data set of the three invariants of stress generated from the corresponding $\bar{\sigma}^{[\alpha, \beta]}$, as defined in Eq. (24). The training is carried out using MATLAB's Neural Fitting Toolbox (nftool). When the training is complete, the new yield function $\bar{\Phi}$ is obtained.

2.5. Macro-scale stress and internal variables update

The trained function $\bar{\Phi}$ can be used to determine whether the current state at each step of a material point in the finite element computation is loaded or unloaded. When the current state is known, a stress update can be performed using the classic return-mapping algorithm. It is assumed that the stress $\boldsymbol{\sigma}^n$, strain $\boldsymbol{\epsilon}^n$ and the other internal variables are known for the previous time step n . The strain increment $\Delta \boldsymbol{\epsilon}$ at step $n + 1$ is also given. The trial stress and the trial active stress are introduced as,

$$\begin{cases} \boldsymbol{\sigma}^{tr} = \boldsymbol{\sigma}^n + \mathbf{D} : \Delta \boldsymbol{\epsilon} \\ \bar{\boldsymbol{\sigma}}^{tr} = \boldsymbol{\sigma}^{tr} - \bar{\mathbf{a}}^n \end{cases} \quad (29)$$

where $\mathbf{D}$ is fourth-order elastic tensor, which is the generalized form of D_{ij} in Eq. (14).

Loading/unloading criterion. Cyclic loading can render a conventional loading/unloading criterion unpractical. For example, during the unloading process, the deformation of the material is not purely elastic, but also includes plasticity. This is the reason for the observed Masing effect indicated in Fig. 1. Furthermore, under the subsequent loading process, even when the equivalent stress is less in magnitude than the maximum stress incurred over the deformation history, plastic deformation may nevertheless take place. As a result of these difficulties, the following loading/unloading criterion, based on the subloading surface, can be instead used [55],

$$\begin{cases} \bar{\Phi}(\bar{\sigma}_m^{tr}, \bar{J}_2^{tr}, \bar{J}_3^{tr}, R^n, F^n; \mathbf{w}, \mathbf{b}) \leq 0 \text{ or } \frac{\partial \bar{\Phi}}{\partial \bar{\boldsymbol{\sigma}}^{tr}} : \Delta \bar{\boldsymbol{\sigma}}^{tr} \leq 0 & \implies \Delta \boldsymbol{\epsilon}^p = \mathbf{0} \\ \text{Otherwise} & \implies \Delta \boldsymbol{\epsilon}^p \neq \mathbf{0} \end{cases} \quad (30)$$

Here, $\bar{\sigma}_m^{tr}$, $\bar{J}_2^{tr}$ and $\bar{J}_3^{tr}$ are three invariants of the trial active stress $\bar{\boldsymbol{\sigma}}^{tr}$, computed by Eq. (24), $\Delta \bar{\boldsymbol{\sigma}}^{tr} = \bar{\boldsymbol{\sigma}}^{tr} - \bar{\boldsymbol{\sigma}}^n$ is an increment of the trial active stress, and $\Delta \boldsymbol{\epsilon}^p$ is an increment of the plastic strain (based on the plastic strain rate $\dot{\boldsymbol{\epsilon}}^p$ outlined in the section on the subloading surface model).

Stress update. When the plastic strain increment $\Delta \boldsymbol{\epsilon}^p$, or plastic strain rate $\dot{\boldsymbol{\epsilon}}^p$, is equal to $\mathbf{0}$, F and $\mathbf{c}$ remain unchanged ($F^{n+1} = F^n$, $\mathbf{c}^{n+1} = \mathbf{c}^n$), and the stress at step $n + 1$ is set equal to the trial stress ($\boldsymbol{\sigma}^{n+1} = \boldsymbol{\sigma}^{tr}$). R^{n+1} ,

$\bar{\mathbf{a}}^{n+1}$, $\mathbf{a}^{n+1}$, $\mathfrak{R}^{n+1}$ and $\bar{\boldsymbol{\sigma}}^{n+1}$ can be then updated as follows,

$$\begin{cases} \bar{\mathbf{a}}^{n+1} = \mathbf{c}^{n+1} - R^{n+1}(\mathbf{c}^{n+1} - \mathbf{a}^n) \\ f(\boldsymbol{\sigma}^{n+1} - \bar{\mathbf{a}}^{n+1}) = R^{n+1} \times F^{n+1} \\ \mathbf{a}^{n+1} = \mathbf{c}^{n+1} - (\mathbf{c}^{n+1} - \bar{\mathbf{a}}^{n+1})/R^{n+1} \\ f(\mathbf{c}^{n+1} - \mathbf{a}^{n+1}) = \mathfrak{R}^{n+1} \times F^{n+1} \\ \bar{\boldsymbol{\sigma}}^{n+1} = \boldsymbol{\sigma}^{n+1} - \bar{\mathbf{a}}^{n+1} \end{cases} \quad (31)$$

When the plastic strain increment $\Delta \boldsymbol{\varepsilon}^p$, or plastic strain rate $\dot{\boldsymbol{\varepsilon}}^p$, is not equal to $\mathbf{0}$, the internal variables and the trial stress must then be updated, to restore the function Φ to zero. The stresses can be updated as follows,

$$\boldsymbol{\sigma}^{tr} = \boldsymbol{\sigma}^n + \mathbf{D} : \left(\Delta \boldsymbol{\varepsilon} - \Delta \lambda \frac{\partial \Phi / \partial \bar{\boldsymbol{\sigma}}^n}{\|\partial \Phi / \partial \bar{\boldsymbol{\sigma}}^n\|} \right) \quad (32)$$

where $\Delta \lambda$ is the flow factor, and $\Delta \boldsymbol{\varepsilon}^p = \Delta \lambda \frac{\partial \Phi / \partial \bar{\boldsymbol{\sigma}}^n}{\|\partial \Phi / \partial \bar{\boldsymbol{\sigma}}^n\|}$. We note that $\partial \Phi / \partial \bar{\boldsymbol{\sigma}}^n$ can be expressed by the neural network through a back-propagation algorithm [50]. We also note that all these internal variables are exclusively functions of $\Delta \lambda$. As such, we can rewrite Eq. (26) by substituting Eqs. (2), (6), (32),

$$\Phi^*(\Delta \lambda) = f_{ANN}(\bar{\boldsymbol{\sigma}}_m^{tr}(\Delta \lambda), \bar{J}_2^{tr}(\Delta \lambda), \bar{J}_3^{tr}(\Delta \lambda); \mathbf{w}, \mathbf{b}) - R^{tr}(\Delta \lambda) \times F^{tr}(\Delta \lambda) \quad (33)$$

where $R^{tr}(\Delta \lambda)$ and $F^{tr}(\Delta \lambda)$ are trial values of R and F , respectively. Newton's iteration method can be used to solve the equation $\Phi^*(\Delta \lambda) = 0$, with $\Delta \lambda$ updated during the iteration process by,

$$\Delta \lambda^{i+1} = \Delta \lambda^i - \frac{\Phi^*(\Delta \lambda^i)}{\partial \Phi^*(\Delta \lambda^i) / \partial \Delta \lambda^i} \quad (34)$$

where $\Delta \lambda^i$ is the value at the i -th iteration of $\Delta \lambda$, with the initial value of $\Delta \lambda$ set to zero. Once $\Delta \lambda$ is obtained, all the variables including the plastic strain increment $\Delta \boldsymbol{\varepsilon}^p$ can be updated by Eqs. (18) and (22).

Fig. 8 summarizes the entire process of building the data-driven constitutive model for cyclic plasticity, starting from original stress-strain data sets, to carrying out subsequent finite element simulations. For the implicit finite element computation, the tangent modulus $\mathbf{L}$ is also required, which can be derived, as shown in Appendices B and C. Our ANN trained material model was implemented as a user material subroutine in a commercial or in-house finite element solver, i.e., UMAT in ABAQUS.

3. Numerical examples

In this section, we employ the proposed approach to predict the cyclic plastic behavior of engineering structures composed of microstructured solids. The numerical examples show that our proposed approach can predict the Masing effect of engineering structures accurately and effectively, saving orders of magnitude of computational time, compared with DNS. In all our examples, a consistent unit of measurement is adopted. The unit of length is mm; of force is N; of time is s; of stress, pressure, and the elastic modulus is MPa.

3.1. Cantilever beam with voids

In this example, a short beam containing 20 spherical voids of the same size is considered (volume fraction is 15.45%), as shown in Fig. 9(a). The matrix material of the beam is assumed to be isotropic and elastoplastic, as described by the subloading surface model, with material parameters as given in Table 2, and with $f(\bar{\boldsymbol{\sigma}}) = \bar{\sigma}_e$. While our previous work [50] has demonstrated the capability of ANNs to learn complex yield functions (such as the hyperbolic Drucker-Prager and Mohr-Coulomb models), our present work, however, focuses on how to describe cyclic plasticity by means of ANNs, and does not explore complex yield functions or subloading surfaces. One end of the beam is fixed while the other end is coupled to a 'reference point'. Cyclic loading, totaling seven cycles, as shown in Fig. 9(b), is applied to this reference point. In each loading cycle, the forces in the x , y , and z directions increase linearly from zero to 0.4, 0.5 and 0.6, respectively, and then decrease linearly back to zero.

This example is thus solved by two approaches. The first approach is direct numerical simulation (DNS), using the finite element method (FEM) with three dimensional, six-node elements (C3D6). The second approach is our proposed approach, where an RVE composed of a homogeneous matrix with a single void, as shown in Fig. 9(c), is first established for data generation. The generated stress-strain data from the RVE is then processed for ANN

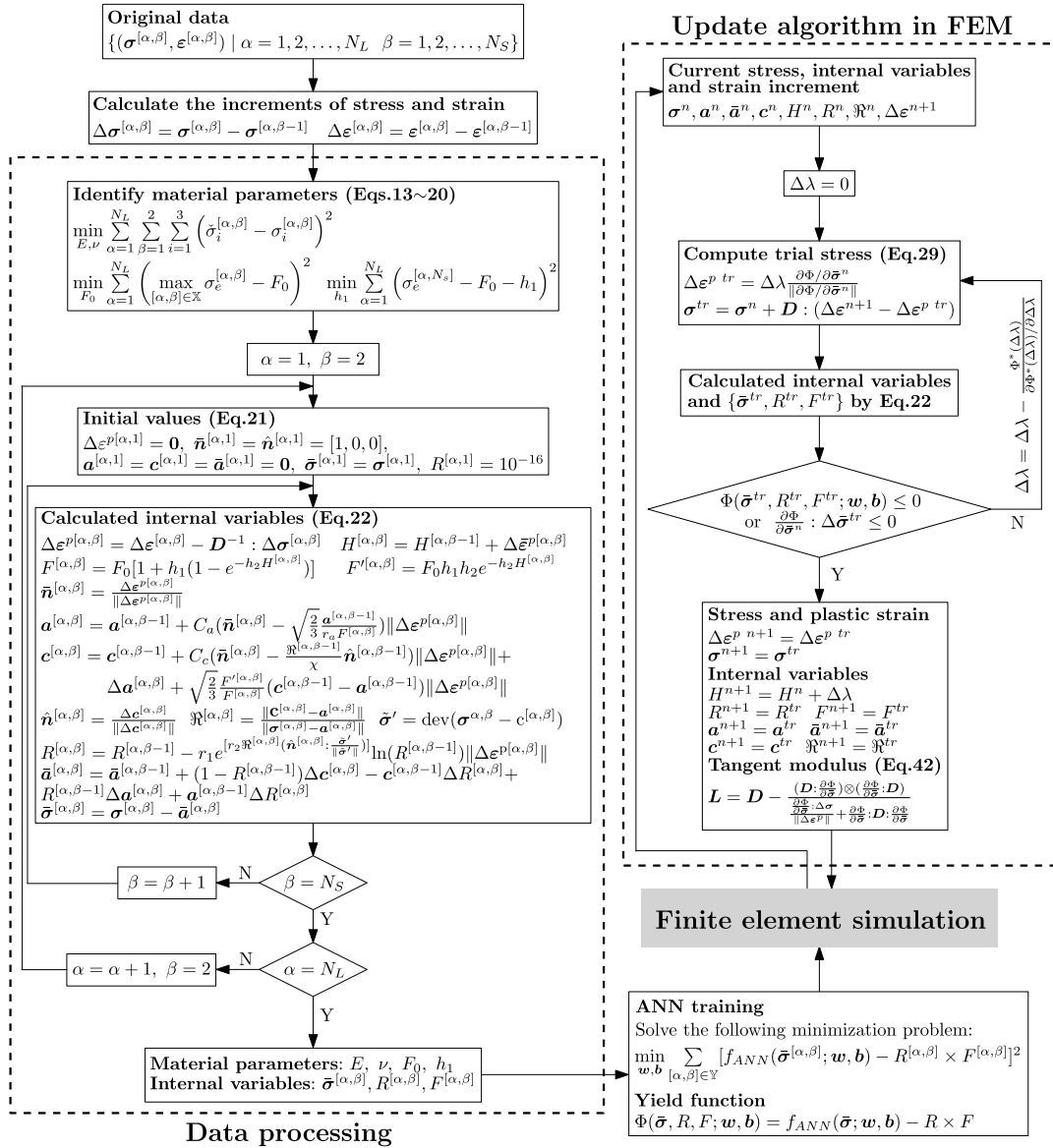


Fig. 8. A detailed flow chart of data-driven constitutive law driving finite element simulation. The dotted box on the left is the process of identifying material parameters and computing the internal variables from the original data through the subloading surface model. The dotted box on the right is the update process of variables in finite element simulation.

train of the constitutive law for the homogenized material (microstructure at the macro-scale). The ANN consists of three hidden layers, each with six neurons. The material parameters E, ν, F_0 and h_1 can be identified by Eqs. (15) and (20), while the other material parameters $h_2, r_1, r_2, C_a, r_a, C_c$ and χ only need to be assumed, based on the material parameters of the matrix used with DNS (see Appendix A). The material parameters for the trained constitutive law are listed in Table 3. Our proposed approach can thus smear out the micro-voids across the beam by homogenization, as shown in Fig. 9(d). The finite element mesh for the homogenized problem includes 3146 nodes and 2500 elements; much less than that of DNS, with 22,426 nodes and 113,895 elements. The element type used for the homogenized finite element analysis is the three dimensional, eight-node element (C3D8), permitted by the simpler topology.

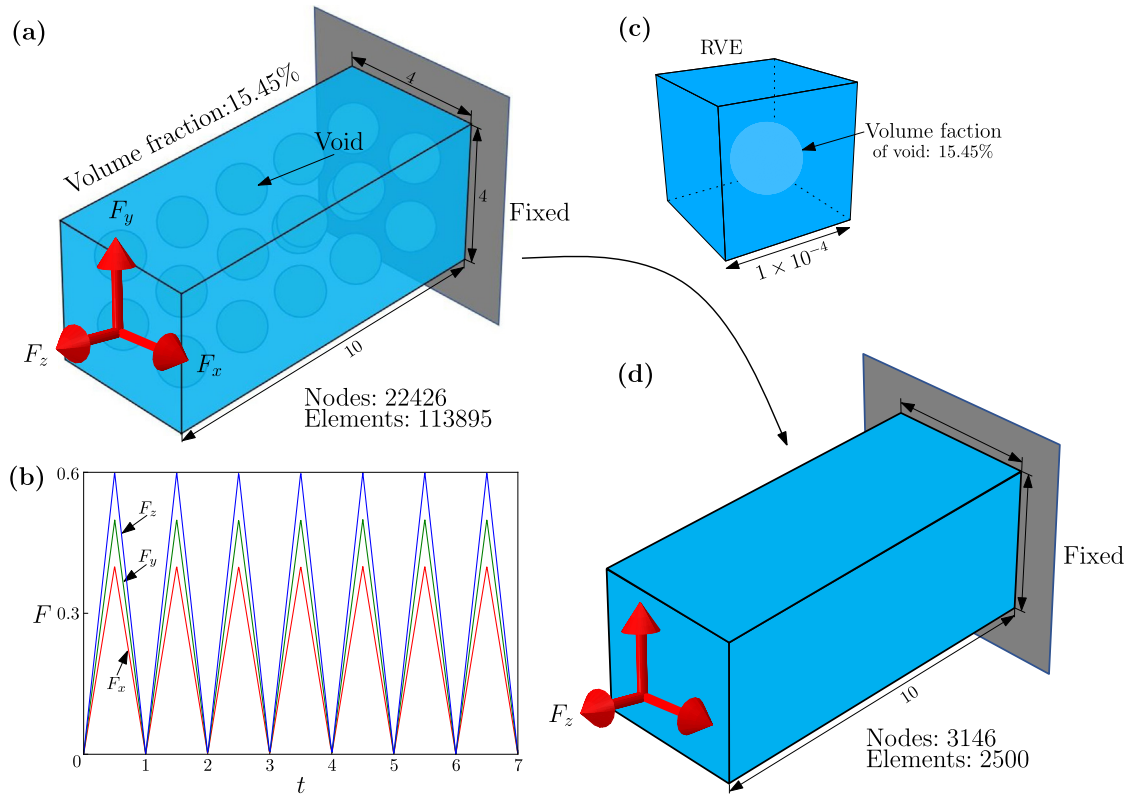


Fig. 9. The finite element analysis on a cantilever beam with spherical voids. The number of nodes and elements used in finite element analysis is listed. (a) The geometry and boundary conditions of the problem. The void volume fraction of the beam is listed. (b) Cyclic loading applied to the reference point. Red, green and blue represent the x , y and z directions, respectively. (c) The RVE used to generate data. (d) The homogenized finite element model of the proposed approach.

Table 3

The material parameters of the trained constitutive law.

Parameter	Identified				Assumed						
	E	ν	F_0	h_1	h_2	r_1	r_2	C_a	r_a	C_c	χ
Value	153.68	0.28	0.87	0.23	15	90	0.8	6	0.6	12	0.8

Fig. 10(a) demonstrates the relationship obtained between the force and the reaction displacement in the x , y , and z directions for the reference point, respectively. As can be seen, the results predicted by our proposed approach and by DNS are both plotted, using solid lines and dotted lines respectively. The results of both approaches, in terms of force and reaction displacement, are practically the same. As the cycles of loading increase, the stress–strain hysteresis loop becomes wider. This may be due to the material’s cyclic softening [14], rendering more evident the Masing effect. The engineering strain may be defined as U/L , where U can be the displacement of the reference point in the x , y and z directions respectively, and L is the dimension of the beam in the corresponding direction. The engineering strain vs. loading cycle is shown in Fig. 10(b). As can be seen, as the cycles increase the average engineering strain gradually increases (at a constant rate). Similar trends were confirmed experimentally for low-carbon steels when subjected to uniaxial cyclic loadings (Okorokov et al. [56]). Therefore, although our proposed approach bypasses the need for, and complexity of, the direct modeling of voids, it successfully predicted the cyclic elastoplastic behavior of a voided beam.

To analyze the effect of microstructure (particles or voids) on the response of bulk materials, the homogenized cantilever beam is solved using the subloading surface model and our ANN model, respectively. The subloading surface model adopts the material parameters of the matrix (listed in Table 2), since it is difficult to calibrate them

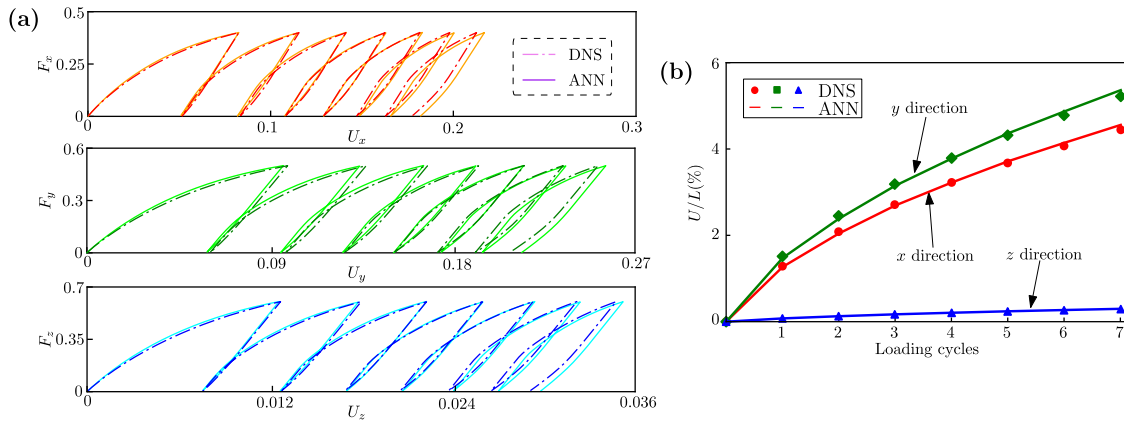


Fig. 10. The results predicted by the proposed approach and DNS for the problem shown in Fig. 9. (a) The force versus the predicted displacement of the reference point. The solid lines and dash dotted lines represent the results computed by the proposed approach and DNS, respectively. (b) The average engineering strain U/L versus loading cycles. The lines and symbols represent the results computed by the proposed approach and DNS, respectively. Red, green and blue lines/symbols represent quantities in the x , y and z directions.

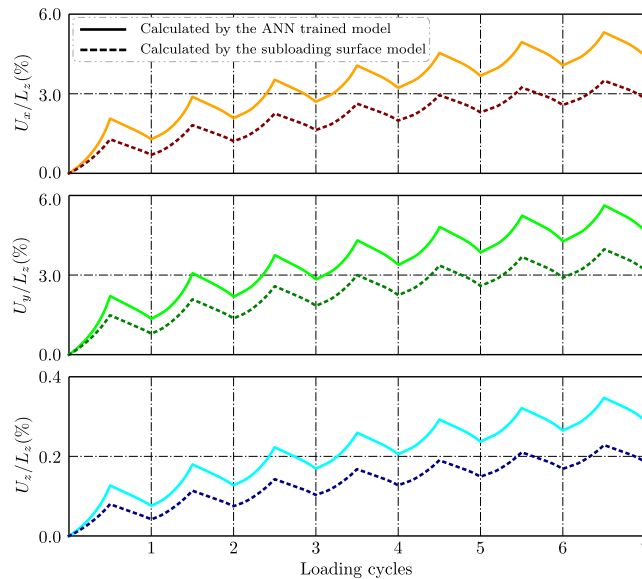


Fig. 11. The average engineering strain U/L of homogeneous beam versus loading cycles. The solid lines and dotted lines represent the results of homogeneous beam described by the ANN trained model and subloading surface model, respectively.

to the case with microstructure without clear physical meaning and experiments. In contrast, the parameters of the ANN model are obtained by training the data of RVE with a single void. Fig. 11 shows the average engineering strain vs. the loading cycles of the homogeneous beam computed by the two models. It can be seen that the difference between the two models is noticeable, which is again consistent with the conclusion in Fig. 5.

Fig. 12(a) plots the contours for equivalent stress, as predicted by the two approaches, for times $t = 0.5, 1.0$ and 1.5 . Only the elements with x coordinates between 0 and 3 are shown in the figure, to help focus our comparison. As can be seen, our approach (though homogenized) obtains a stress distribution quite similar to DNS. As shown in Fig. 12(b), the difference between the equivalent stresses predicted by the two approaches is analyzed for Surf-Y1 (marked in Fig. 12(a)) at the same times. The maximum relative differences at the three moments ($t = 0.5, 1.0$ and 1.5) are less than 6.25%, which is deemed negligible.

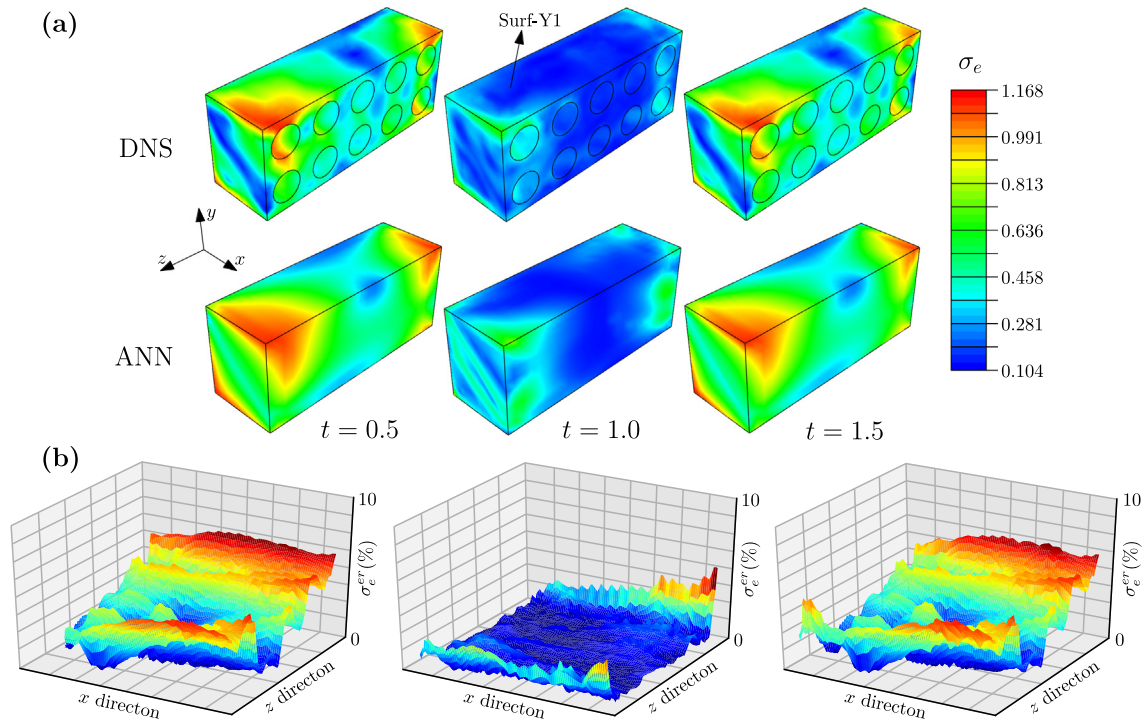


Fig. 12. The difference of equivalent stress between the results obtained by the proposed approach and DNS at time moment $t = 0.5$, 1.0 and 1.5. (a) The contour plots of equivalent stress predicted by DNS (σ_e^D) and the proposed approach (σ_e^A). The surface Surf-Y1 is marked. (b) The relative difference σ_e^{er} ($=|\sigma_e^D - \sigma_e^A|/|\sigma_e^D|$) between the equivalent stress predicted by the proposed approach and DNS on the surface Surf-Y1.

Likewise, the contour plots for the accumulated plastic strain are shown in Fig. 13(a), and the relative difference of the accumulated plastic strain H for Surf-Y1 is shown in Fig. 13(b). As the loading cycles increase, the relative difference grows a little larger. Nevertheless, overall the relative difference stays within a small range (almost all less than 5%), again being deemed negligible. Therefore, it can be seen that our proposed approach can successfully predict the distribution of both the equivalent stress and the accumulated plastic strain for the voided beam problem.

Finally, the computational time required to solve this voided beam problem was 283.15 min for DNS and 30.61 min for our data-driven approach. Counting the time cost for data generation (594.27 min), feature extraction (0.87 min), and ANN training (2.03 min), the total computational time of the proposed approach will be more than twice that of the DNS. However, these steps only need to be performed in the offline stage. Our proposed approach thus exhibits much greater computational efficiency than DNS, obtained by drastically reducing the degrees of freedom required for a convergent finite element analysis. In addition, for other loading modes such as torsion, the proposed approach can also obtain almost the same results as DNS at the macroscopic level, and the results are given in Appendix D.

3.2. Cantilever bending beam, with randomly distributed particles

In this example, a beam with randomly distributed particles is studied under cyclic bending, as shown in Fig. 14(a). 74 particles (volume fraction is 13.80%), with diameters of 5 and 3, are randomly distributed across the matrix. For DNS, both the matrix and the particles are assumed to be isotropic elastoplastic materials, as described by the subloading surface model. The corresponding material parameters are shown in Table 2. As indicated in Table 4, different subloading surface functions are employed to distinguish the matrix from the particles. Fig. 14(c) shows the stress-strain response for the matrix and the particle materials when subjected to the uniaxial cyclic loading shown in 14(b).

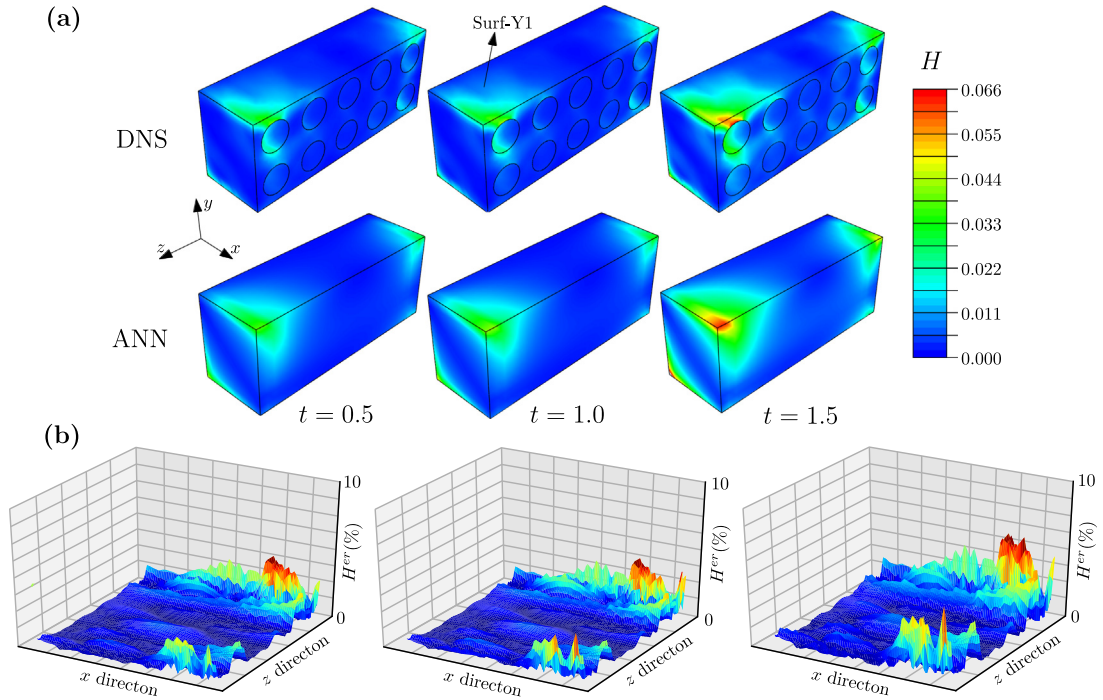


Fig. 13. The difference of accumulated plastic strain H between the results obtained by the two approaches at time moment $t = 0.5$, 1.0 and 1.5. (a) The contour plots of accumulated plastic strain predicted by DNS (H^D) and the proposed approach (H^A). The marked surface Surf-Y1 is same as Fig. 12(a). (b) The relative difference $H^{er} (=|H^D - H^A|/|H^D|)$ between the accumulated plastic strain predicted by the proposed approach and DNS on the surface Surf-Y1.

Table 4

The subloading surface functions of the matrix and particles.

Part	Subloading surface
Matrix	$f(\bar{\sigma}) = \bar{\sigma}_e$
Particles	$f(\bar{\sigma}) = \bar{\sigma}_e + \ell \bar{\sigma}_m$ ($\ell = 0.1$)

One end of the beam is held fixed, while the other end (marked by force F_x , F_y and F_z) is coupled to a reference point. As shown in Fig. 14(d), a total of three loading cycles, with progressively decreasing amplitude, are applied to the reference point. In the first loading cycle ($t = 0 \sim 1$), the forces in the x , y , and z directions increase linearly from zero to 6, 5, and 4, respectively, and then decrease linearly to back zero. The forces of the second cycle ($t = 1 \sim 2$) and the third cycle ($t = 2 \sim 3$) vary similarly as the first cycle, but with amplitudes that are 80% and 60% of the first cycle, respectively.

Similar to the voided beam problem, this example is solved by finite elements, using DNS with three dimensional, six-node elements (C3D6), and using our approach with three dimensional, eight-node elements (C3D8). For our proposed approach, as shown in Fig. 14(e), the RVE is established first according to the average particle volume fraction of the beam. Then, the trained constitutive law is obtained, where the ANN for training consists of three hidden layers, again with six neurons. The material parameters of the trained constitutive law are as follows: $E=202.52$, $\nu=0.302$, $F_0=1.0$, $h_1=0.2$, $h_2=15.0$, $r_1=90.0$, $r_2=0.8$, $C_a=6.0$, $r_a=0.6$, $C_c=12.0$, $\chi=0.8$. Finally, the problem is solved using the trained ANN model, where the particles are smeared out. Compared with the finite element mesh for DNS (16,285 nodes and 84,966 elements), our proposed approach drastically reduces the mesh size to only 1908 nodes and 1335 elements.

The relationship between the force and the reaction displacement in the x , y , and z directions for the reference point is shown in Fig. 15, where the solid lines and the dash dotted lines represent the results predicted by our proposed approach and by DNS, respectively. Clearly, there is practically no difference between the results

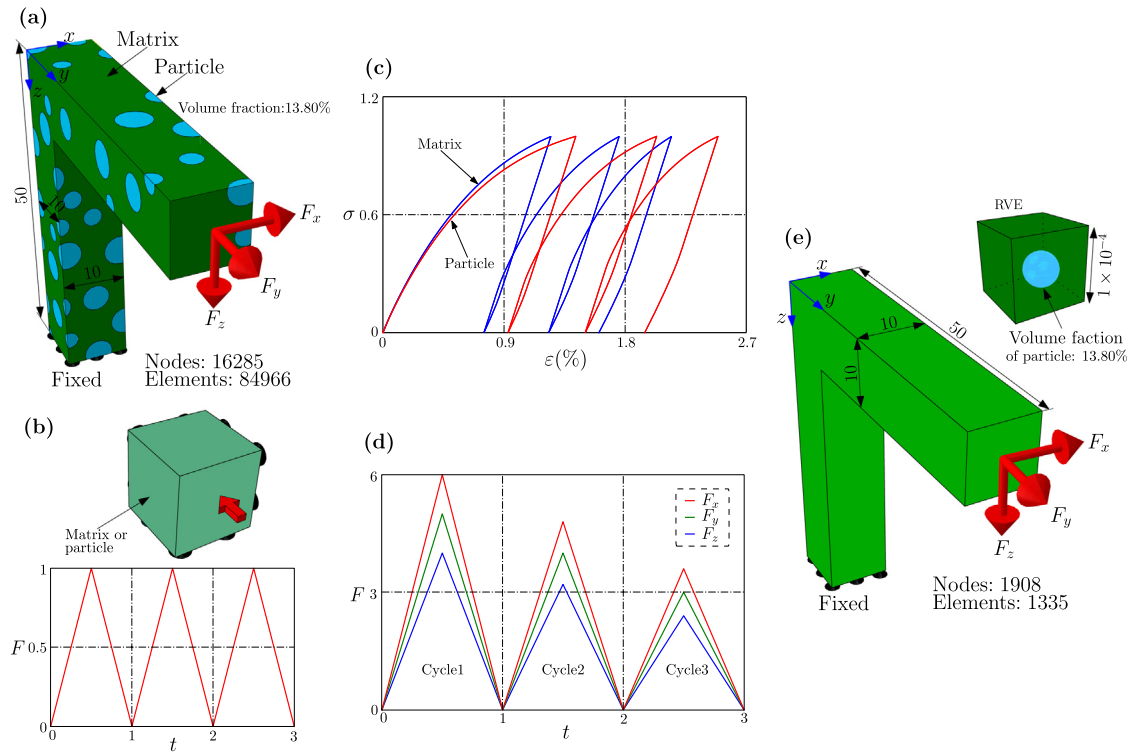


Fig. 14. The finite element analysis on a beam contained with 74 randomly distributed particles. The number of nodes and elements used in finite element analysis is listed. (a) The geometry and boundary conditions of the problem. (b) Cyclic loadings and boundary conditions of the matrix or particle materials under uniaxial conditions. (c) The stress–strain curves of the matrix and particle materials under uniaxial cyclic loadings. (d) Cyclic loading applied to the reference point. Red, green and blue represent the x , y and z directions, respectively. (e) The RVE used to generate data and the homogenized finite element model of the proposed approach.

predicted by the two approaches. It is worth noting that although the mechanical properties (material parameters and subloading surface function) of the matrix and the particles are different, our proposed approach can still accurately describe the cyclic elastoplastic behavior of the beam structure. In addition, although the amplitudes of the cyclic loading in the second and third cycles are smaller than for the first cycle, the enhancement of plastic deformation is nonetheless observed with increasing number of cycles. This finding is consistent with experimental observations, e.g., [4,57]. The reason is revealed by investigating the subloading surface model [54,58]. In it, the occurrence of plastic deformation only depends on whether the subloading surface expands. This subloading surface also expands under stress levels that are lower than the maximum recorded over a given deformation history.

The contour plots of the accumulated plastic strain H , the ratio of the subloading surface to normal-yield surface sizes R , and the equivalent (Cauchy) stress σ_e obtained by the two approaches at time $t = 0.5$ are summarized in Fig. 16. As can be seen, the results of the two approaches differ locally (especially for a region close to the particles). Nevertheless, the difference in terms of the overall field distribution (H , R or σ_e) is negligible. In fact, the region of localized plastic deformation for the beam has also been identified by our proposed approach, in very good agreement with that predicted by DNS.

The computational time for our proposed approach with this beam bending problem is 9.35 min; far less than the 139.65 min required by DNS. We can conclude, therefore, from this example and the one before it, that when describing the cyclic plastic behavior of materials with microstructure (particles or voids), our approach demonstrates its advantages over DNS in terms of computational efficiency, having leveraged the accurate subloading surface function.

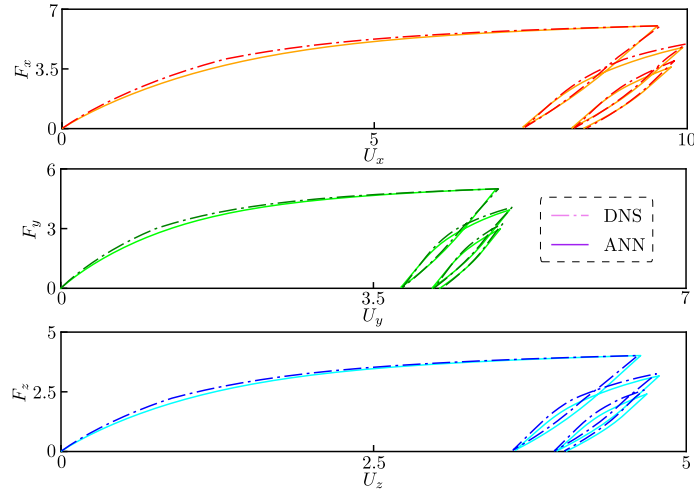


Fig. 15. The imposed force versus predicted displacement of DNS (by dash dotted lines) and the proposed data-driven approach (by solid lines). Red, green and blue represent quantities in the x , y and z directions, respectively.

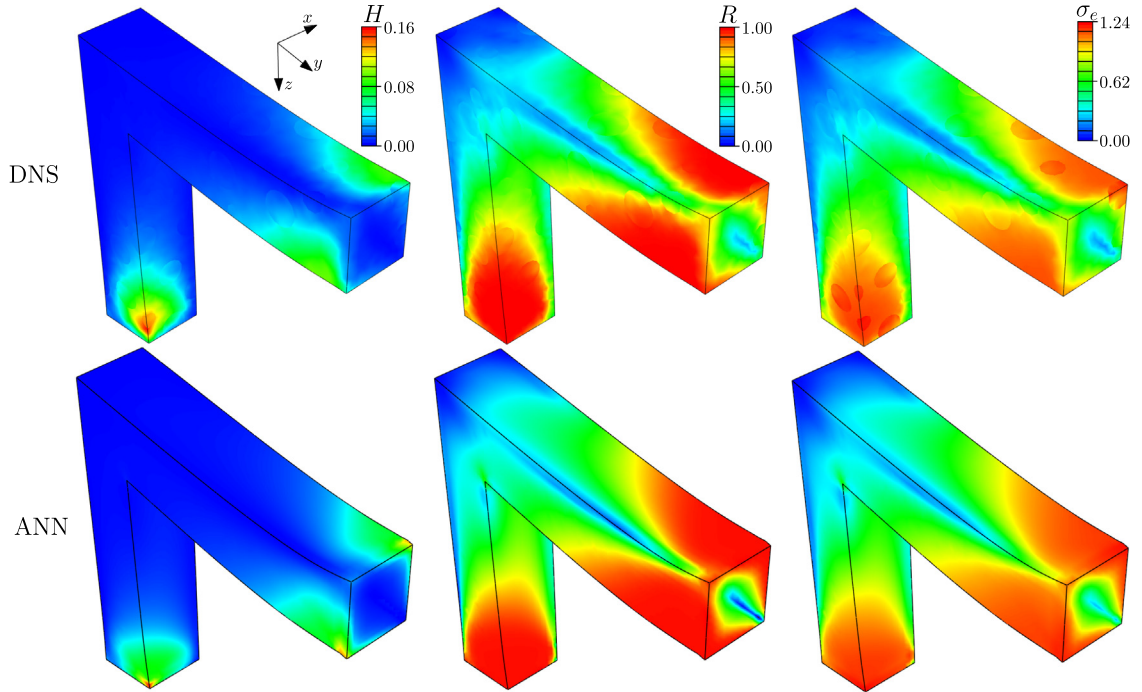


Fig. 16. The contour plots of accumulated plastic strain H , the ratio of subloading surface to normal-yield surface R and equivalent stress σ_e obtained by DNS and the proposed approach at the time moment $t = 0.5$.

3.3. Three-dimensional truss structure

In this example, as shown in Fig. 17(a), the cyclic behavior of a three-dimensional truss structure, widely used in engineering, is studied. As in our first example, the truss structure is assumed to be isotropic and homogeneous elastoplastic material, as described by the subloading surface model (the material parameters are the same as those of DNS model in the first example) with $f(\bar{\sigma}) = \bar{\sigma}_e$. The material parameters of the ANN trained model are: $E=199.31$, $\nu=0.298$, $F_0=1.0$, $h_1=0.2$, $h_2=15.0$, $r_1=90.0$, $r_2=0.8$, $C_a=6.0$, $r_a=0.6$, $C_c=12.0$, $\chi=0.8$. The bottom ends

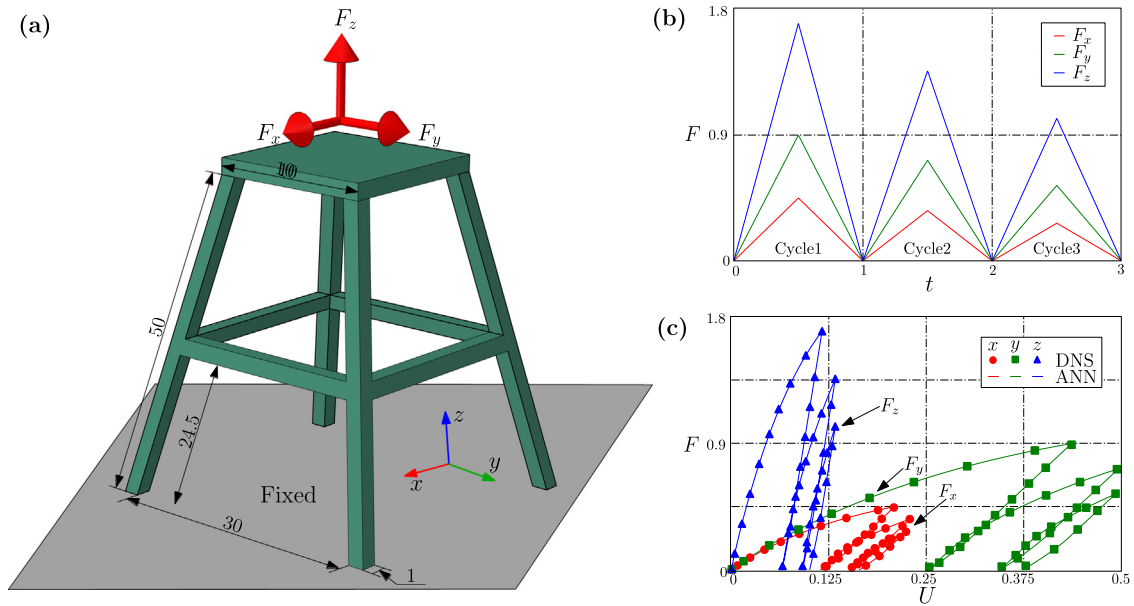


Fig. 17. The finite element analysis on a three-dimensional truss structure. (a) The geometry and boundary conditions of the problem. (b) Cyclic loading applied to the reference point. The colors of line have the same meanings as those in Fig. 14(d). (c) The force vs. displacement of the reference point predicted by the proposed data-driven approach (by lines) and DNS (by symbols). The colors of line or symbol have the same meanings as those in Fig. 10(b).

of the truss are held fixed, while the top plane is coupled to a reference point. As in the second example, three cycles of loading, as shown in Fig. 17(b), were applied to the reference point. The maximum loadings in the x , y , and z directions for the first cycle ($t = 0 \sim 1$) are 1.7, 0.9 and 0.45, respectively. The maximum loads of the second cycle ($t = 1 \sim 2$) and the third cycle ($t = 2 \sim 3$) are again 80% and 60% of the first cycle, respectively.

In this example, three dimensional, six-node elements (C3D6) are employed to mesh the truss structure, for finite element analysis by both DNS, and our approach. The force vs. the reaction displacement curve at the reference point, as predicted by DNS (markers) and our proposed approach (lines), is plotted in Fig. 17(c). It can again be clearly seen that our approach applies well to the cyclic plastic behavior of the three-dimensional truss structure.

The contour plots for the equivalent active stress $\bar{\sigma}_e$, as obtained by both approaches for the unloading phase in the first cycle to the unloading phase in the second cycle ($0.5 \sim 1.5$), are shown in Fig. 18(a) and (b), respectively. It is found that the stress distributions are nearly identical, although some negligible differences may be discerned. Though there is no microstructure, these small differences result from the different internal variables obtained from our data processing in comparison with those used by DNS.

Fig. 18(c) shows the evolution of some equivalent variables in a representative element from the mesh (element number 5357) for the ANN trained model, where the value of the variable is the average value of all integration points of the element. The equivalent (Cauchy) stress σ_e vs. the equivalent strain ε_e predicted by our proposed approach is shown in Fig. 18(c-1). It can be seen that the unloading process of the first cycle can be divided into a completely elastic part ($t_1 \sim t_2$) and an elastoplastic part ($t_2 \sim t_3$). However, from the perspective of equivalent active stress $\bar{\sigma}_e$ shown in Fig. 18(c-2), the elastoplastic part ($t_2 \sim t_3$) is a reverse loading process. Also considering the accumulated plastic strain H in Fig. 18(c-3), there is no plastic strain generated from t_1 to t_2 , being purely elastic; but plastic strain is correctly generated from t_2 to t_3 . In other words, the Cauchy stress σ should not be used to assess cyclic plasticity; the active stress $\bar{\sigma}$ on the subloading surface should instead be used to judge its onset. These results thus confirm that our trained ANN model correctly describes the evolution of the internal variables (such as $\bar{\sigma}$), in alignment with the requirements of the subloading surface model.

4. Conclusions

In this study, a mechanistic data-driven approach is proposed to predict the complex elastoplastic behavior of microstructured/homogenized solids subjected to cyclic loadings. Our proposed approach is capable of performing

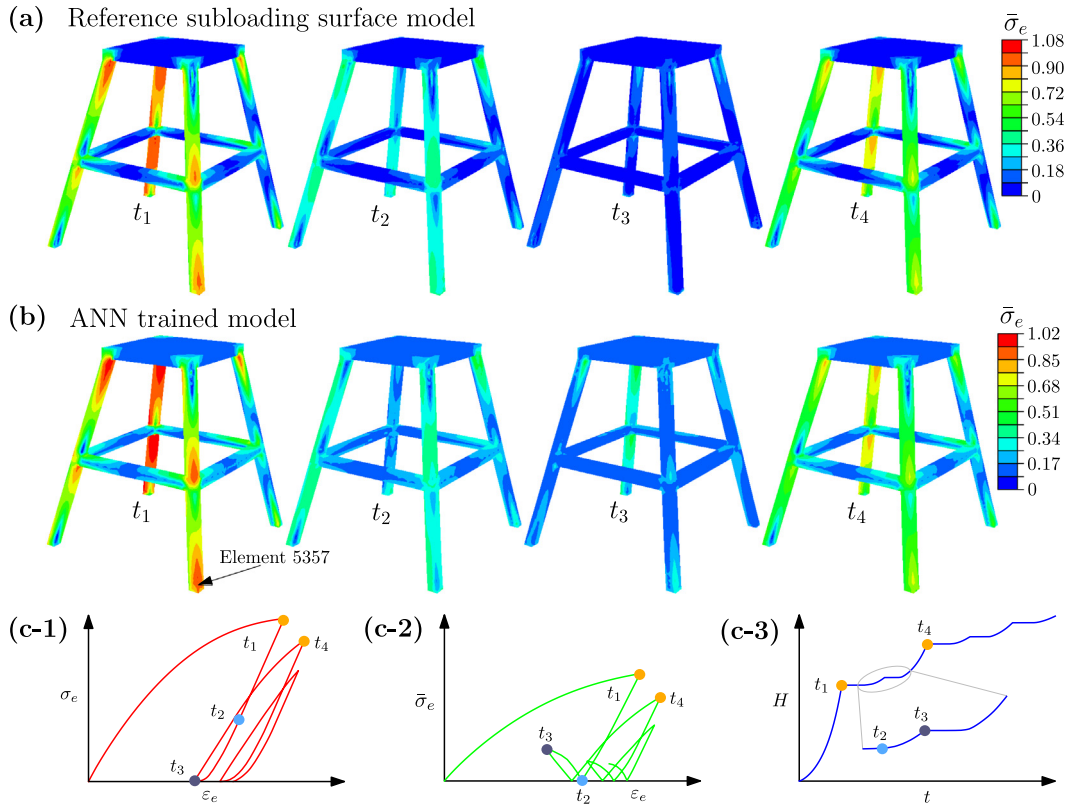


Fig. 18. The results predicted by the proposed data-driven approach and DNS at time 0.2~0.4. Contour plots of equivalent active stress $\bar{\sigma}_e$ predicted by (a) the reference subloading surface model and (b) the ANN trained model at time moment t_1 , t_3 , t_3 , and t_4 . Here, t_1 , t_3 , and t_4 represent 0.5, 1.0, and 1.5, respectively, and t_2 represents the time moment when plastic strain occurs at the unloading of first cycle. (c) The evolution of equivalent stress σ_e , equivalent active stress $\bar{\sigma}_e$ and accumulated plastic strain H of the element marked in (b).

numerical simulations on elastoplastic materials, without explicitly knowing the mathematical functional form of the yield function. The approach is illustrated through three examples that involve microstructured/homogenized materials under complex cyclic loadings. Compared with the results obtained from the DNS for traditional cyclic elastoplastic models (the subloading surface model in the present work), we found that our approach successfully predicts the elastoplastic behavior of materials rather well, and can also predict the distribution (and localization) of some key physical quantities, such as the equivalent stress, and the accumulated plastic strain, across an engineering structure. Our proposed approach can thus significantly strengthen the capability of conventional mathematical function-based plasticity models, for cyclic loading.

How to obtain a well-calibrated constitutive model for microstructured materials that are subjected complex cyclic loadings, in a manner that is quick and efficient, is the issue alleviated by our proposed approach. The present approach combines tools of data science with the existing material model built through the mechanics principle. Specifically, by taking the subloading surface model, as an example, internal variables that characterize the Masing and ratcheting effects as mechanistic features are extracted from raw stress–strain data, as they cannot be measured directly during physical experiments (or their numerical surrogates), rendering a pure data-driven approach very expensive to train. In our combined method, a subloading surface is represented in the trained ANN, by means of these extracted internal variables, and requiring much less training data.

Our proposed approach has not yet considered the rotation of the principal stress axes to capture the effect of anisotropy (often associated with microstructures), nor has it considered non-associative flow rules. On the other hand, dealing with noisy data is considered in many applications of data science [22]. Therefore, the influence of noise contained in experimental data should be taken into account along the line of the proposed approach. These unexplored directions leave much room for further development of our proposed model.

Table 5

Four cases with different material parameters to verify the robustness of the proposed approach.

Case	h_2	r_1	r_2	C_a	r_a	C_c	χ
Reference subloading surface model	20	100	1	5	0.5	15	0.6
ANN trained model I	19.47	142.93	0.85	3.48	0.38	16.74	0.83
ANN trained model II	14.80	90.18	0.58	3.70	0.31	10.26	0.64
ANN trained model III	23.19	59.67	1.32	6.59	0.61	9.75	0.71

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.

Acknowledgments

X.Guo acknowledges the support from National Key Research and Development Plan 2020YFB1709401. The support from the NSF of China (Project No. 11872139, 11821202, 11732004), and Dalian University of Technology Fundamental Research Fund (Project No. DUT20RC(5)008) is also acknowledged. W.K.L. acknowledges the support of the United States National Science Foundation (NSF) under Grant No. MOMS/CMMI-1762035.

Appendix A

In the data processing shown in Section 2.3, it has been shown that the material parameters h_2 , r_1 , r_2 , C_a , r_a , C_c and χ have to be set in prior to extract the internal variables hidden in the data. Then the related issue is if the different choice of these material parameters will affect the trained ANN model and its prediction. In order to illustrate the insensitivity to the material parameters h_2 , r_1 , r_2 , C_a , r_a , C_c and χ , three cases are designed, in which the material parameters h_2 , r_1 , r_2 , C_a , r_a , C_c and χ are set as different values shown in Table 5. These material parameters are also different from the reference subloading surface model for generating the data in the numerical experiments, varying in a range (50% ~ 150% of those in the reference subloading surface model. χ is less than 1). Three ANN trained models are obtained by the proposed approach using the data generated by the reference subloading surface model, where the ANN for training consists of three hidden layers with six neurons. Other involved material parameters are set with the same value: $E=200$, $\nu=0.3$, $F_0=1.0$, $h_1=0.2$.

The trained ANN model with different material parameters is then used to predict the stress–strain response under the two loading paths ($\alpha = 2$ and 15, refer to Table 1), the relative difference between the principal stress predicted by the three ANN trained models and the reference subloading surface model at different loading step are shown in Fig. 19. The relative difference is defined as $\sigma_i^{er} = |\sigma_i^H - \sigma_i^A|/|\sigma_i^H|$, where σ_i^H and σ_i^A are the principal stress predicted by the reference subloading surface model and ANN trained model, respectively. i (=1, 2 or 3) represents the principal direction of stress. It can be seen from Fig. 19 that the relative differences of the three ANN trained models in most cases are less than 3%, except for the loading path ID=15 where the relative difference is close to 6%. The results under other loading paths are not shown to save the space. It is also observed that the difference between the results of the three ANN trained models and the reference subloading surface model is very small.

The principal stress of two loading paths ($\alpha = 2$ and 15) vs. the step number for the reference subloading surface model and the ANN trained model III are further plotted in Fig. 20. The symbols and lines represent the stresses predicted by the reference subloading surface model and the ANN trained model III, respectively. It can be seen that although the relative difference of the ANN training model III under the loading path ($\alpha = 15$) shown in Fig. 19 is a little large, the actual difference of the principal stress is not obvious.

In summary, the proposed approach does not require the high accuracy of material parameters h_2 , r_1 , r_2 , C_a , r_a , C_c and χ . They can be chosen referring to the matrix material for microstructured solids with a little arbitrary. This is the advantage of the proposed approach.

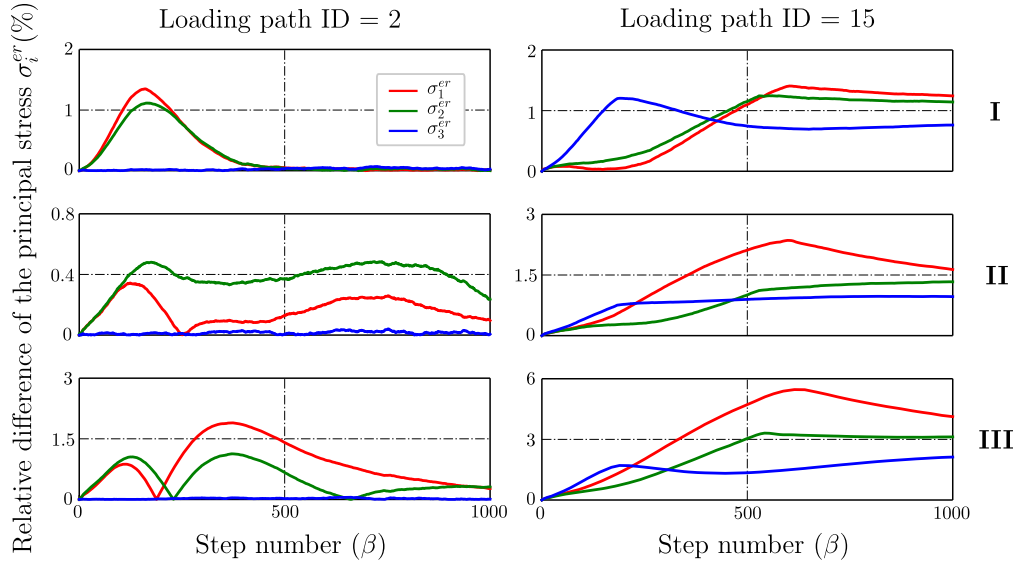


Fig. 19. The relative difference (σ_i^{er}) of the principal stress between the reference subloading surface model and the ANN trained model. I, II and III represent three ANN trained models listed in Table 5, respectively. The red, green and blue lines represent the relative difference of σ_1 , σ_2 and σ_3 , respectively.

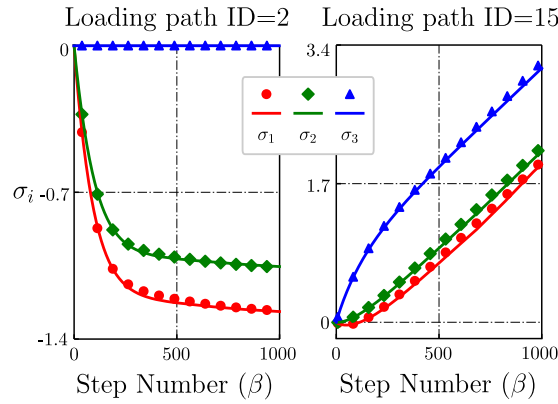


Fig. 20. The difference of principal stress between the reference subloading surface model (by symbols) and the ANN trained model III (by lines). Red, green and blue respectively represent the three principal stress σ_1 , σ_2 and σ_3 , respectively.

Appendix B

The tangent modulus required in the finite element computation is derived in this appendix. It is observed from Eq. (3) that,

$$f(\bar{\sigma}) = \frac{\partial f(\bar{\sigma})}{\partial \bar{\sigma}} : \bar{\sigma} = R \times F, \quad (35)$$

by means of Euler's theorem applied to the homogeneous function $f(\bar{\sigma})$. Upon substitution with the consistency condition,

$$\frac{\partial f(\bar{\sigma})}{\partial \bar{\sigma}} : d\bar{\sigma} - R dF - F dR = 0, \quad (36)$$

and then integrating the definition of the normal of the subloading surface $\bar{n}$ into Eq. (36) yields,

$$\bar{n} : d\sigma - \bar{M}^p \|d\epsilon^p\| = 0, \quad (37)$$

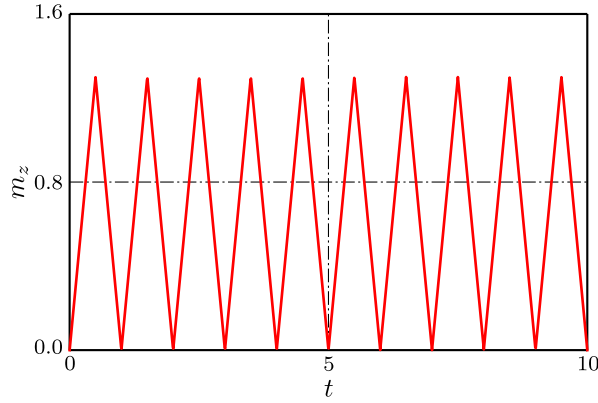


Fig. 21. The torque m_z twisted around the z direction.

where

$$\bar{M}^p = \bar{n} : \left[\frac{d\bar{a}}{\|d\epsilon^p\|} + \frac{\sigma - \bar{a}}{\|d\epsilon^p\|} \left(\frac{dF}{F} + \frac{dR}{R} \right) \right]. \quad (38)$$

Then, by substituting

$$d\sigma = D : (d\epsilon - \bar{n} \|d\epsilon^p\|) \quad (39)$$

into Eq. (37) we obtain,

$$\|d\epsilon^p\| = \frac{\bar{n} : D : d\epsilon}{\bar{M}^p + \bar{n} : D : \bar{n}}. \quad (40)$$

As such, the relation between $d\sigma$ and $d\epsilon$ is obtained by means of Eqs. (39), (40),

$$d\sigma = \left[D - \frac{(D : \bar{n}) \otimes (\bar{n} : D)}{\bar{M}^p + \bar{n} : D : \bar{n}} \right] : d\epsilon. \quad (41)$$

Finally, the elastic–plastic tangent modulus L can be denoted as

$$L = D - \frac{(D : \bar{n}) \otimes (\bar{n} : D)}{\bar{M}^p + \bar{n} : D : \bar{n}}, \quad (42)$$

where the normal of the subloading surface $\bar{n}$ can be computed by,

$$\bar{n} = \frac{\partial f(\bar{\sigma}) / \partial \bar{\sigma}}{\|\partial f(\bar{\sigma}) / \partial \bar{\sigma}\|}, \quad (43)$$

and the quantity $\partial f(\bar{\sigma}) / \partial \bar{\sigma}$ can be derived based on the trained ANN (that is, $\partial f(\bar{\sigma}) / \partial \bar{\sigma} = \partial f_{ANN}(\bar{\sigma}) / \partial \bar{\sigma}$), as shown in Appendix C.

Appendix C

For stress update in the finite element computation, the derivatives of the active stress invariants defined in Eq. (24) with respect to active stress are required, which are derived as follows:

$$\begin{cases} \frac{\partial \bar{\sigma}_m}{\partial \bar{\sigma}_{ij}} = \frac{1}{3} \delta_{ij} \\ \frac{\partial \bar{J}_2}{\partial \bar{\sigma}_{ij}} = \frac{\bar{\sigma}'_{ij}}{\bar{J}_2} \\ \frac{\partial \bar{J}_3}{\partial \bar{\sigma}_{ij}} = \bar{J}_3^{-2} \left(\bar{\sigma}'_{im} \bar{\sigma}'_{jm} - \frac{1}{3} \bar{J}_2^2 \delta_{ij} \right) \end{cases} \quad (44)$$

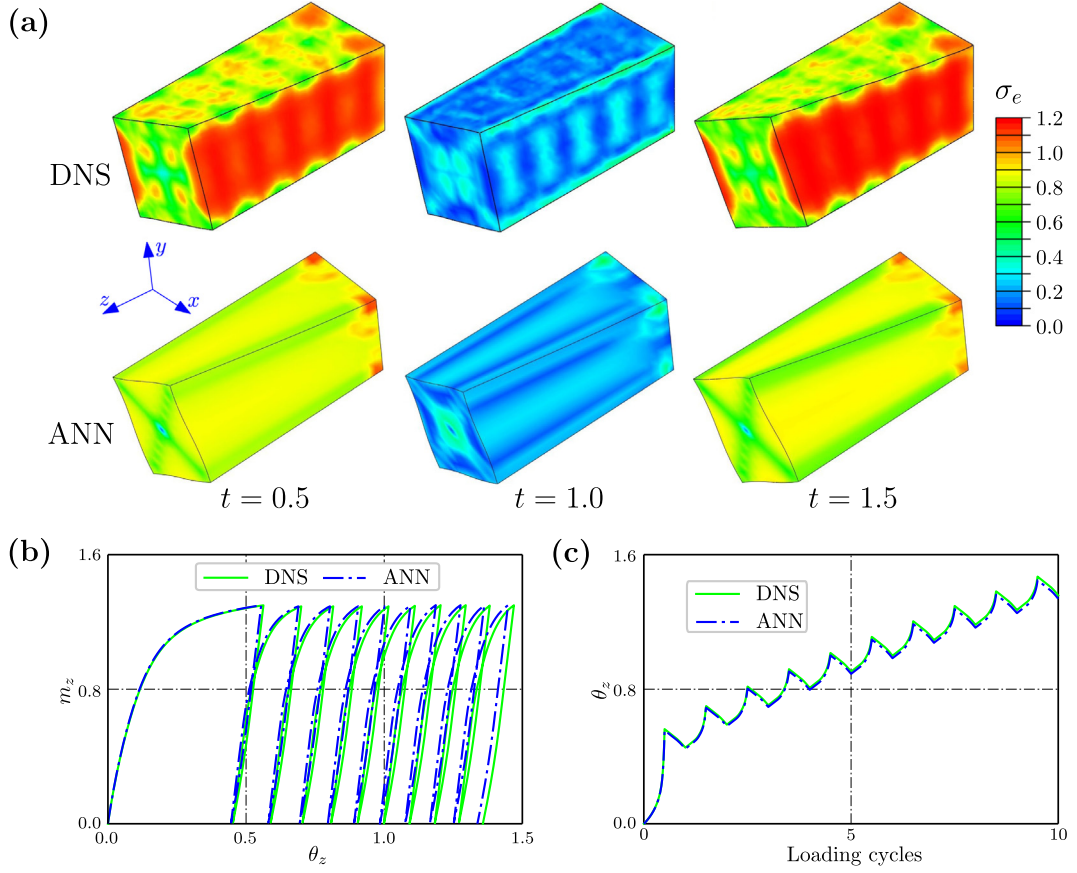


Fig. 22. The results predicted by the two approaches. (a) The contour plots of equivalent stress predicted by DNS and the proposed approach at the time moment $t = 0.5, 1.0$ and 1.5 . (b) The torsion m_z versus the predicted rotation angle θ_z of the reference point. (c) The predicted torsion angle θ_z versus loading cycles. The green solid lines and blue dash dotted lines represent the results computed by DNS and the proposed approach, respectively.

where δ_{ij} is the Kronecker symbol, and $\bar{\sigma}'_{ij}$ is the deviatoric part of $\bar{\sigma}_{ij}$. According to Eq. (26) and the chain rule, the quantity of $\partial f_{ANN} / \partial \bar{\sigma}_{ij}$ can be obtained:

$$\frac{\partial f_{ANN}}{\partial \bar{\sigma}_{ij}} = \frac{\partial f_{ANN}}{\partial \bar{\sigma}_m} \frac{\partial \bar{\sigma}_m}{\partial \bar{\sigma}_{ij}} + \frac{\partial f_{ANN}}{\partial \bar{J}_2} \frac{\partial \bar{J}_2}{\partial \bar{\sigma}_{ij}} + \frac{\partial f_{ANN}}{\partial \bar{J}_3} \frac{\partial \bar{J}_3}{\partial \bar{\sigma}_{ij}} \quad (45)$$

in which $\frac{\partial f_{ANN}}{\partial \bar{\sigma}_m}$, $\frac{\partial f_{ANN}}{\partial \bar{J}_2}$ and $\frac{\partial f_{ANN}}{\partial \bar{J}_3}$ can be computed by the back propagation algorithm [50].

Appendix D

Torsion is one of the typical load models on engineering structures. So, it is necessary to test the prediction capability of our proposed ANN approach under cyclic torsional loads. The geometry and boundary conditions are the same as those in Fig. 6(d) for our first example. The cyclic torsional load m_z applied vs. time at a reference point are shown in Fig. 21.

The stress contours and the twist angle θ_z vs. load cycles as predicted by DNS and by our proposed ANN approach are shown in Fig. 22. Due to the existence of voids, the local stress (equivalent stress σ_e) in DNS is significantly higher than that for ANN. Nevertheless, the twist angle θ_z at the reference point predicted by the two approaches under cyclic torsional loads is essentially identical.

References

- [1] J.L. Chaboche, A review of some plasticity and viscoplasticity constitutive theories, *Int. J. Plast.* 24 (10) (2008) 1642–1693.
- [2] C.J. Smithells, C.S. Smith, *Metals Reference Book*, Physics Today, New York, 1949.
- [3] G. Johansson, M. Ekh, K. Runesson, Computational modeling of inelastic large ratcheting strains, *Int. J. Plast.* 21 (5) (2005) 955–980.
- [4] Y. Jiang, J. Zhang, Benchmark experiments and characteristic cyclic plasticity deformation, *Int. J. Plast.* 24 (9) (2008) 1481–1515.
- [5] H. Burlet, G. Cailletaud, Numerical techniques for cyclic plasticity at variable temperature, *Eng. Comput.* 3 (2) (1986) 143–152.
- [6] N. Ohno, J.-D. Wang, Kinematic hardening rules with critical state of dynamic recovery, part I: formulation and basic features for ratchetting behavior, *Int. J. Plast.* 9 (3) (1993) 375–390.
- [7] N. Ohno, J.-D. Wang, Kinematic hardening rules with critical state of dynamic recovery, part II: Application to experiments of ratchetting behavior, *Int. J. Plast.* 9 (3) (1993) 391–403.
- [8] D. McDowell, Stress state dependence of cyclic ratchetting behavior of two rail steels, *Int. J. Plast.* 11 (4) (1995) 397–421.
- [9] J.L. Chaboche, P. Kanoute, F. Azzouz, Cyclic inelastic constitutive equations and their impact on the fatigue life predictions, *Int. J. Plast.* 35 (2012) 44–66.
- [10] Z. Mróz, H.P. Shrivastava, R.N. Dubey, A non-linear hardening model and its application to cyclic loading, *Acta Mech.* 25 (1) (1976) 51–61.
- [11] Y.F. Dafalias, E.P. Popov, A model of nonlinearly hardening materials for complex loading, *Acta Mech.* 21 (3) (1975) 173–192.
- [12] Y.F. Dafalias, E.P. Popov, Plastic internal variables formalism of cyclic plasticity, *J. Appl. Mech.* 43 (4) (1976) 645–651.
- [13] K. Hashiguchi, Subloading surface model in unconventional plasticity, *Int. J. Solids Struct.* 25 (8) (1989) 917–945.
- [14] K. Hashiguchi, M. Ueno, Elastoplastic constitutive equation of metals under cyclic loading, *Internat. J. Engrg. Sci.* 111 (2017) 86–112.
- [15] W. Prager, A New method of analyzing stresses and strains in work - hardening plastic solids, *J. Appl. Mech.* 23 (1956) 493–496.
- [16] C.O. Frederick, P. Armstrong, A mathematical representation of the multiaxial bauschinger effect, *Mater. High Temp.* 24 (1) (2007) 1–26.
- [17] R. Hill, J. Rice, Constitutive analysis of elastic-plastic crystals at arbitrary strain, *J. Mech. Phys. Solids* 20 (6) (1972) 401–413.
- [18] A. Bensoussan, J.L. Lions, G. Papanicolaou, *Asymptotic Analysis for Periodic Structures*, North-Holland, Amsterdam, 1978.
- [19] E. Sanchez-Palencia, A. Zaoui, *Homogenization Techniques for Composite Media*, Springer, Berlin, 1987.
- [20] M. Hori, S. Nemat-Nasser, On two micromechanics theories for determining micro-macro relations in heterogeneous solids, *Mech. Mater.* 31 (10) (1999) 667–682.
- [21] J. Yvonnet, D. Gonzalez, Q.-C. He, Numerically explicit potentials for the homogenization of nonlinear elastic heterogeneous materials, *Comput. Methods Appl. Mech. Engrg.* 198 (33) (2009) 2723–2737.
- [22] W.K. Liu, M. Fleming, Z. Gan, *Mechanistic Data Science for STEM Education and Applications*, Springer, Cham, 2021.
- [23] N. Huber, C. Tsakmakis, Determination of constitutive properties from spherical indentation data using neural networks. Part II: plasticity with nonlinear isotropic and kinematic hardening, *J. Mech. Phys. Solids* 47 (7) (1999) 1589–1607.
- [24] S. Jung, J. Ghaboussi, Neural network constitutive model for rate-dependent materials, *Comput. Struct.* 84 (15) (2006) 955–963.
- [25] M. Oeser, S. Freitag, Fractional derivatives and recurrent neural networks in rheological modelling - part I: theory, *Int. J. Pavement Eng.* 17 (2) (2016) 87–102.
- [26] A. Yamanaka, R. Kamijyo, K. Koenuma, I. Watanabe, T. Kuwabara, Deep neural network approach to estimate biaxial stress-strain curves of sheet metals, *Mater. Des.* 195 (2020) 108970.
- [27] M.A. Bessa, R. Bostanabad, Z. Liu, A. Hu, D.W. Apley, C. Brinson, W. Chen, W.K. Liu, A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality, *Comput. Methods Appl. Mech. Engrg.* 320 (2017) 633–667.
- [28] K. Wang, W. Sun, Meta-modeling game for deriving theory-consistent, microstructure-based traction-separation laws via deep reinforcement learning, *Comput. Methods Appl. Mech. Engrg.* 346 (2019) 216–241.
- [29] Y. Heider, K. Wang, W. Sun, SO(3)-invariance of informed-graph-based deep neural network for anisotropic elastoplastic materials, *Comput. Methods Appl. Mech. Engrg.* 363 (2020) 112875.
- [30] S. Im, H. Kim, W. Kim, M. Cho, Neural network constitutive model for crystal structures, *Comput. Mech.* 67 (1) (2021) 185–206.
- [31] L. Zhang, L. Cheng, H. Li, J. Gao, C. Yu, R. Domel, Y. Yang, S. Tang, W.K. Liu, Hierarchical deep-learning neural networks: finite elements and beyond, *Comput. Mech.* 67 (1) (2021) 207–230.
- [32] L. Zhang, Y. Lu, S. Tang, W.K. Liu, HiDeNN-PGD: reduced-order hierarchical deep learning neural networks, 2021, *CoRR*, arXiv:2105.06363.
- [33] S. Saha, Z. Gan, L. Cheng, J. Gao, O.L. Kafka, X. Xie, H. Li, M. Tajdari, H.A. Kim, W.K. Liu, Hierarchical deep learning neural network (HiDeNN): An artificial intelligence (AI) framework for computational science and engineering, *Comput. Methods Appl. Mech. Engrg.* 373 (2021) 113452.
- [34] Z. Liu, M. Bessa, W.K. Liu, Self-consistent clustering analysis: An efficient multi-scale scheme for inelastic heterogeneous materials, *Comput. Methods Appl. Mech. Engrg.* 306 (2016) 319–341.
- [35] Z. Liu, M. Fleming, W.K. Liu, Microstructural material database for self-consistent clustering analysis of elastoplastic strain softening materials, *Comput. Methods Appl. Mech. Engrg.* 330 (2018) 547–577.
- [36] G. Cheng, X. Li, Y. Nie, H. Li, FEM-cluster based reduction method for efficient numerical prediction of effective properties of heterogeneous material in nonlinear range, *Comput. Methods Appl. Mech. Engrg.* 348 (2019) 157–184.
- [37] Y. Nie, G. Cheng, X. Li, L. Xu, K. Li, Principle of cluster minimum complementary energy of FEM-cluster-based reduced order method: fast updating the interaction matrix and predicting effective nonlinear properties of heterogeneous material, *Comput. Mech.* 64 (2) (2019) 323–349.
- [38] S. Tang, L. Zhang, W.K. Liu, From virtual clustering analysis to self-consistent clustering analysis: a mathematical study, *Comput. Mech.* 62 (6) (2018) 1443–1460.

- [39] T. Kirchdoerfer, M. Ortiz, Data-driven computational mechanics, *Comput. Methods Appl. Mech. Engrg.* 304 (2016) 81–101.
- [40] K. Karapiperis, L. Stainier, M. Ortiz, J. Andrade, Data-driven multiscale modeling in mechanics, *J. Mech. Phys. Solids* 147 (2021) 104239.
- [41] T. Furukawa, G. Yagawa, Implicit constitutive modelling for viscoplasticity using neural networks, *Internat. J. Numer. Methods Engrg.* 43 (2) (1998) 195–219.
- [42] J.-S. Lee, E.-C. Lee, T. Furukawa, Formulation of the neural network for implicit constitutive model (II) : Application to inelastic constitutive equations, *Int. J. Fuzzy Logic Intell. Syst.* 8 (4) (2008) 264–269.
- [43] M. Mozaffar, R. Bostanabad, W. Chen, K. Ehmann, J. Cao, M.A. Bessa, Deep learning predicts path-dependent plasticity, *Proc. Natl. Acad. Sci. USA* 116 (52) (2019) 26414–26420.
- [44] G. Chen, Recurrent neural networks (RNNs) learn the constitutive law of viscoelasticity, *Comput. Mech.* 67 (3) (2021) 1–11.
- [45] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neural Netw.* (ISSN: 0893-6080) 2 (5) (1989) 359–366.
- [46] Z. Shen, H. Yang, S. Zhang, Deep network approximation: Achieving arbitrary accuracy with fixed number of neurons, 2021, CoRR, arXiv:2107.02397.
- [47] S. Tang, G. Zhang, H. Yang, Y. Li, W.K. Liu, X. Guo, MAP123: A Data-driven approach to use 1D data for 3D nonlinear elastic materials modeling, *Comput. Methods Appl. Mech. Engrg.* 357 (2019) 112587.
- [48] S. Tang, Y. Li, H. Qiu, H. Yang, S. Saha, S. Mojumder, W.K. Liu, X. Guo, MAP123-EP: A mechanistic-based data-driven approach for numerical elastoplastic analysis, *Comput. Methods Appl. Mech. Engrg.* 364 (2020) 112955.
- [49] H. Yang, X. Guo, S. Tang, W.K. Liu, Derivation of heterogeneous material laws via data-driven principal component expansions, *Comput. Mech.* 64 (2) (2019) 365–379.
- [50] H. Yang, H. Qiu, S. Tang, Q. Xiang, X. Guo, Exploring elastoplastic constitutive law of microstructured materials through artificial neural network (ANN) - a mechanistic-based data-driven approach, *J. Appl. Mech.* 87 (9) (2020) 1–16.
- [51] P. Verma, N.S. Srinivas, S. Singh, V. Singh, Low cycle fatigue behavior of modified 9Cr–1Mo steel at room temperature, *Mater. Sci. Eng. A Struct. Mater. Prop. Microstruct. Process.* 652 (2016) 30–41.
- [52] B. Das, A. Singh, Understanding strain controlled low cycle fatigue response of P91 steel through experiment and cyclic plasticity modeling, *Fusion Eng. Des.* 138 (2019) 125–137.
- [53] D. Lu, K. Zhang, G. Hu, Y. Lan, Y. Chang, Investigation of yield surfaces evolution for polycrystalline aluminum after pre-cyclic loading by experiment and crystal plasticity simulation, *Materials* 13 (14) (2020) 3069.
- [54] K. Hashiguchi, *Foundations of Elastoplasticity: Subloading Surface Model*, Springer, Tokyo, 2017.
- [55] K. Hashiguchi, Exact formulation of subloading surface model: Unified constitutive law for irreversible mechanical phenomena in solids, *Arch. Comput. Methods Eng.* 23 (3) (2016) 417–447.
- [56] V. Okorokov, Y. Gorash, D. Mackenzie, R. van Rijswijk, New formulation of nonlinear kinematic hardening model, Part II: Cyclic hardening/softening and ratcheting, *Int. J. Plast.* 122 (2019) 244–267.
- [57] K. Kolasangiani, M. Shariati, K. Farhangdoost, A. Varvani-Farahani, Ratcheting progress at notch root of 1045 steel samples over asymmetric loading cycles: Experiments and analyses, *Fatigue Fract. Eng. Mater. Struct.* 41 (9) (2018) 1870–1883.
- [58] K. Hashiguchi, Multiplicative hyperelastic-based plasticity for finite elastoplastic deformation/sliding: A comprehensive review, *Arch. Comput. Methods Eng.* 26 (3) (2019) 597–637.