



Topological dimensionality reduction-based machine learning for efficient gradient-free 3D topology optimization

Zhaoyou Sun^a, Yaguang Wang^a, Pai Liu^a, Yangjun Luo^{b,*}

^aState Key Laboratory of Structural Analysis for Industrial Equipment, Dalian University of Technology, Dalian 116024, China

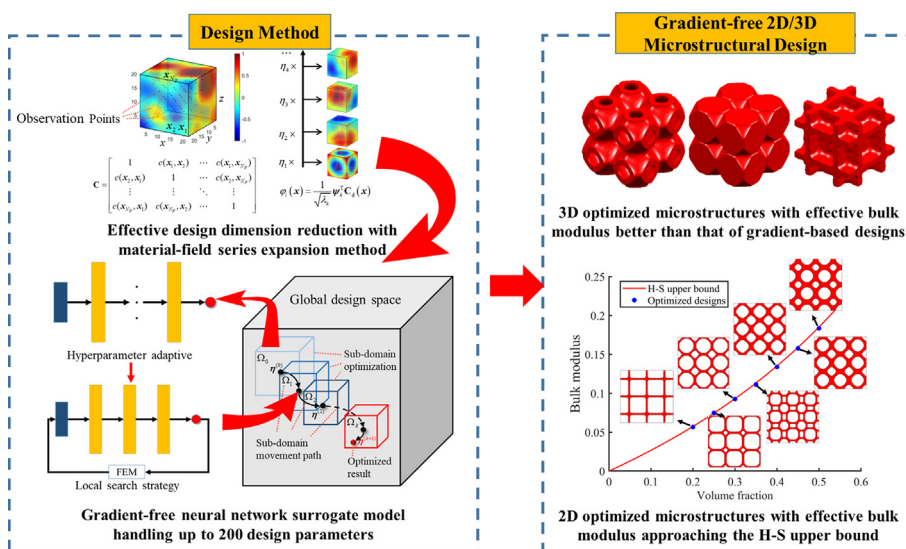
^bSchool of Science, Harbin Institute of Technology, Shenzhen 518055, China

HIGHLIGHTS

- A novel topology optimization method combining self-adaptive neural network and material-field series expansion is proposed.
- The self-adaptive neural network model handles up to 200 design parameters under topological dimensionality reduction.
- This method reduces the computational time by 1–4 orders of magnitude compared with the coarse-mesh-based gradient-free methods.
- The optimized cellular material effective bulk modulus by the proposed method achieves 98% of the theory bound.

GRAPHICAL ABSTRACT

Illustration: A gradient-free topological design method integrating the material-field series expansion topological parameterization and the deep neural networks is proposed. This optimization method reduces the computational time by 1–4 orders of magnitude compared with the coarse-mesh-based gradient-free methods and can handle up to 200 design parameters in design dimension reduction model. Several designs with high performances are obtained with the present method.



ARTICLE INFO

Article history:

Received 14 March 2022

Revised 18 June 2022

Accepted 21 June 2022

Available online 24 June 2022

Keywords:

Gradient-free

Topology optimization

ABSTRACT

Powerful gradient-free topology optimization methods are needed for structural design concerning complex responses. In this paper, a novel gradient-free optimization method is proposed by integrating the material-field series expansion topological parameterization and the deep neural networks, providing two-fold advances: firstly, it generally reduces the massive topological design variables to fewer than 200, while keeps the capability to represent relative complex 3D topologies and clear boundaries; secondly, by constructing a sequential neural network surrogate model, it sufficiently explores the reduced design space and is capable of handling multi-peak and discontinuous optimization problems. The effectiveness of this method is illustrated via several design problems, among which the optimized material

* Corresponding author.

E-mail address: yangjunluo@hit.edu.cn (Y. Luo).

Deep neural network
Material-field series expansion

effective bulk modulus achieves 98% of the H-S bound and the highly-nonlinear peak weld stress in a phone dropping process is decreased by 16.59%. This method reduces the computational time by 1–4 orders of magnitude compared with the coarse-mesh-based gradient-free methods, and it is the first time to successfully conduct gradient-free 3D topology optimization with thousands of finite elements. The method's ease of implementation and compatibility with various simulation software, brings topology optimization into complex industrial applications and proves that gradient-free technology represents an effective optimization benchmark for improving structural performance.

© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

Topology optimization, which aims to determine the optimal material distribution within a given design domain for the best performance of structures, has attracted widespread attention within academia and engineering contexts over the past two decades [1,2]. From a fundamental perspective, a topological design is a highly non-convex 0–1 integer programming problem in an infinite-dimensional design space, posing significant challenges for achieving the optimal solutions [3]. Unlike size or shape optimization related to continuous and finite-dimensional design space, the topology optimization procedure typically experiences an artificial transition from discrete to continuous variables, thereby giving rise to three main steps, i.e., topological representation, material interpolation modelling, and the optimization algorithm. In the topological representation step, typically, the density-based method [4] represents the structural topology by thousands of elemental relative densities, while the evolutionary structural optimization method [5] represents the structural topology directly through the use of finite elements. By describing the structural topology with radial basis functions or beam component functions, the parameter level-set method [6,7] and the moving morphable component method [8] were shown to reduce the design dimensionality into several hundred parameters. However, in most existing topology representations, only gradient-based optimization algorithms, such as optimality criteria methods [9], sequential linear programming methods [10], the method of moving asymptotes [11], or the sequential approximate integer programming [12], can be used to effectively solve the corresponding topology optimization model, thereby requiring a complicated sensitivity analysis and easily falling to a local solution. Material interpolation modelling is also critical in the topology design of continua, giving rise to, e.g., the interpolation of solid isotropic material with penalization (SIMP) [13], the rational approximation of material properties [14], and others [15].

There are two significant advantages to gradient-free topology optimization compared to gradient optimization methods as follows [16]: (i) Because the sensitivity information is not mandatory and only the evaluation of the objective function is required, gradient-free topology optimization is easy to implement for multi-physical problems; (ii) in complex optimization problems with multiple-peak objective functions or non-connected design spaces, gradient-free topology optimization has better global search capabilities. Existing works on gradient-free methodologies, e.g., those addressing the genetic algorithm (GA), the generalized simulated annealing (GSA) and the self-directed online learning optimization (SOLO), showed that they would be useful for selected complex problems in which standard gradient methods fail [17–19]. However, these methods have also been regarded as hopelessly inefficient with poor performance for density-based topology optimization problems [20] due to a large number of design variables that will inevitably lead to a dimensional explosion in gradient-free algorithms. Based on the genetic algorithm and a specific generative algorithm to coarsely represent topology,

Lohan et al. [21] studied the topological design of steady-state heat-conduction problems. Yoshimura et al. [22] investigated the flow-channel topology design with the genetic algorithm assisted by the Kriging surrogate model. After simplifying the structural topology by using several beam component functions, Raponi et al. [23] developed a Kriging-assisted level-set method for the design of two-dimensional (2D) crash structures. However, the exploitation of the gradient-free topology optimization with a complete representation of structural topology to handle complex three-dimensional (3D) topology optimization problems that are likely to be unsolvable using gradient methods, remains elusive.

Recently, a new topology representation model, referred to as material-field series expansion (MFSE) [24], was demonstrated as effectively and significantly reducing the number of design variables in topology representation while providing an accurate and clear description of the structural boundaries. The proposed MFSE-based topology representation with reduced dimensionality is fundamentally different from the conventional density-based method. Accordingly, MFSE can be combined with gradient-free optimization algorithms to achieve the optimized topology of any structure. By combining MFSE and a Kriging-based optimization algorithm, Kriging-based MFSE (KG-MFSE) [25] has been successfully applied for dealing with non-smooth and multiple local minima optimization problems without requiring sensitivity analysis, e.g., photonic or phononic crystal bandgap problems [26–28], material microstructure design [29], and freely vibrating design [30]. For 3D topology optimization problems, the number of MFSE design variables used for the complete representation of structural topology is typically above 100; in such a case, the Kriging-based algorithm will not perform well because of its extremely inaccurate surrogate in high-dimensional space. Therefore, how to create a more effective gradient-free solution strategy is an open problem still to be completed.

As a burgeoning field in artificial intelligence research, neural networks have had interesting impacts on topology optimization. Existing work on topology optimization with the aid of neural networks aims to create a large solution data set that is based on traditional topology optimization, as well as establish the mapping between pixels, e.g., mapping from the intermediate grayscale structure [31–33] or the initial design domain and boundary conditions [34–36] to the final structural topology. This method requires a vast supporting data set; furthermore, it is difficult to impose complex constraint conditions. Hence, large amounts of calculation, a weak generalization ability, and the violation of constraints are inherent defects of such an approach. In response, some scholars have put forward corresponding improvements to compensate for these shortcomings through the use of a larger integrated neural network [37].

In this study, a new gradient-free topology optimization methodology based on the deep neural network and MFSE (DNN-MFSE) model is proposed. Topological representation using a bounded material field with spatial dependency, which can use a reduced series expansion to greatly decrease the topology optimization design variables, as well as a new optimization algorithm

that uses a self-adapted DNN process, is presented. Finally, we provide detailed numerical examples by way of a compliance design problem, a microstructure bulk modulus problem, and the dropped phone problem to illustrate the method's validity and applicability.

2. Topology optimization model

Continuum topology optimization can be explained as finding the optimal distribution of materials in a structural design domain. The general mathematical formula of topology optimization problems can be summarized as follows:

$$\begin{aligned} \min \quad & f(\chi, \mathbf{u}) \\ \text{s.t.} \quad & G(\chi, \mathbf{u}) = 0 \\ & g_k(\chi, \mathbf{u}) \leq 0, \quad k = 1, 2, \dots \end{aligned} \quad (1)$$

where χ denotes the material distribution function, $f(\chi, \mathbf{u})$ is the performance function to be optimized, which can be given in the form of compliance, structural stress, band gap, etc. $\mathbf{u}$ is the structural response obtained by solving the equilibrium function $G(\chi, \mathbf{u}) = 0$, and g_k is the inequality constraint in topology optimization, which typically contains a volume or a mass constraint.

2.1. Topological description method based on material field

The issue of how to characterize the material distribution in the structural design domain is the main difference between various topology optimization methods. Among these techniques, the MFSE topological description model [24] plays an important role in the gradient-free topology optimization strategy proposed in this study. It introduces a bounded material field function, i.e., $\varphi(\mathbf{x}) \in [-1, 1]$, to define the 2D or 3D design domain topology of the structure $\mathbf{x} \in \Omega_{\text{des}}$. For continuum topology optimization, the distribution of materials often has a specific spatial consistency; as such, the material field function can naturally be endowed with a correlation definition. Herein, it is assumed that the correlation is expressed as an exponential distance function as follows:

$$c(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / l_c^2) \quad (2)$$

where $c(\mathbf{x}_i, \mathbf{x}_j)$ represents the correlation of the material field function value of the two points $\mathbf{x}_i$ and $\mathbf{x}_j$, and l_c represents the correlation length, which can control the topological complexity. If l_c is large, the spatial correlation of the material field is strong, which means the spatial fluctuation degree of the material field is relatively gentle, and vice versa. It is worth noting that, based on the definition of the above formula, there are obviously $c(\mathbf{x}_i, \mathbf{x}_i) = 1$ and $c(\mathbf{x}_i, \mathbf{x}_j) \in [0, 1]$. In practical engineering problems, the correlation length is typically set to 30% of the size of the design domain, i.e., $l_c = 0.3 \cdot \min(a, b, c)$, where a , b , and c are the dimensions of the 3D design domain.

The realization of each material-field function represents a topological configuration of the structure. Because the material field function is continuous, it must be projected to get an approximation of discrete material distribution χ at any coordinate point $\mathbf{x}$ as follows:

$$\chi(\mathbf{x}) = \begin{cases} \varphi(\mathbf{x})e^{-\beta} + 1 - e^{-\beta\varphi(\mathbf{x})}, & \text{if } \varphi(\mathbf{x}) \geq 0 \\ \varphi(\mathbf{x})e^{-\beta} - 1 + e^{-\beta\varphi(\mathbf{x})}, & \text{if } \varphi(\mathbf{x}) < 0 \end{cases} \quad (3)$$

where β is the smoothing parameter, which controls the smoothness of the projection. When $\beta = 0$, $\chi(\mathbf{x}) = \varphi(\mathbf{x})$, and as β becomes larger, the projection for establishing $\chi(\mathbf{x})$ tends toward 1 and -1 . Note that $0 \leq \chi(\mathbf{x}) < 1$ refers to the solid material while $-1 \leq \chi(\mathbf{x}) < 0$ refers to void in single-material topology optimization. In the practical optimization process, the continuation approach for gradually increasing β from 0 to 40 is used.

2.2. Material-field series expansion

When the bounded material field is endowed with a specific spatial correlation, the material field function $\varphi(\mathbf{x})$ can be expressed using a non-probabilistic series expansion method that inherently maintains the spatial correlation of the original material field. By arranging N_p observation points evenly in the structural design domain, a correlation matrix $\mathbf{C}$ can be constructed as follows:

$$\mathbf{C} = \begin{bmatrix} 1 & c(\mathbf{x}_1, \mathbf{x}_2) & \cdots & c(\mathbf{x}_1, \mathbf{x}_{N_p}) \\ c(\mathbf{x}_2, \mathbf{x}_1) & 1 & \cdots & c(\mathbf{x}_2, \mathbf{x}_{N_p}) \\ \vdots & \vdots & \ddots & \vdots \\ c(\mathbf{x}_{N_p}, \mathbf{x}_1) & c(\mathbf{x}_{N_p}, \mathbf{x}_2) & \cdots & 1 \end{bmatrix} \quad (4)$$

where upon the eigenvalue problem becomes

$$\mathbf{C} = \{\psi_1 \ \psi_2 \ \cdots \ \psi_{N_p}\}^T \begin{Bmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \\ & & & \lambda_{N_p} \end{Bmatrix} \{\psi_1 \ \psi_2 \ \cdots \ \psi_{N_p}\} \quad (5)$$

where λ_k and ψ_k are the eigenvalue and eigenvector of the correlation matrix in a descending order, respectively.

Through the correlation matrix, the material field can be expressed by a linear series expansion [38] as

$$\varphi(\mathbf{x}) = \sum_{k=1}^{N_p} \eta_k \frac{1}{\sqrt{\lambda_k}} \psi_k^T \mathbf{C}_d(\mathbf{x}), \quad \mathbf{x} \in \Omega_{\text{des}} \quad (6)$$

where η_k is the MFSE coefficient to be determined, $\mathbf{C}_d(\mathbf{x})$ is the division of the correlation matrix, and the specific expression is $\mathbf{C}_d(\mathbf{x}) = [c(\mathbf{x}, \mathbf{x}_1), c(\mathbf{x}, \mathbf{x}_2), \dots, c(\mathbf{x}, \mathbf{x}_{N_p})]$.

To better understand the expansion shown in Eq. (6), the first nine orders of eigenvalues λ_k and the corresponding eigenvectors ψ_k of an illustrative correlation matrix for a cube design domain are given in Fig. 1. The different orders of eigenvectors show different basis functions; furthermore, the eigenvalues decrease rapidly with increased order, thus indicating that the major topology configurations can be approximately represented by the first several orders of eigenvectors.

Based on the above observation, it is convenient to ignore the expansion items with smaller eigenvalues. The number of truncated terms not only affects the approximate degree of the structural topology, but also affects the optimization efficiency of gradient-free methods. In this study, the number of truncation terms, i.e., M , is selected by the formulation

$$\varepsilon = \sum_{i=1}^M \lambda_i / \sum_{j=1}^{N_p} \lambda_j \geq 99.9\% \quad (7)$$

and the truncation series expression of the material field is

$$\varphi(\mathbf{x}) \approx \sum_{k=1}^M \eta_k \frac{1}{\sqrt{\lambda_k}} \psi_k^T \mathbf{C}_d(\mathbf{x}) = \boldsymbol{\eta}^T \boldsymbol{\Lambda}^{-1/2} \boldsymbol{\Phi}^T \mathbf{C}_d(\mathbf{x}) \quad (8)$$

where $\boldsymbol{\eta} = \{\eta_1, \eta_2, \dots, \eta_M\}^T$ is the vector of coefficients to be determined, $\boldsymbol{\Lambda} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_M)$ is an $M \times M$ diagonal matrix, and $\boldsymbol{\Phi} = \{\psi_1, \psi_2, \dots, \psi_M\}$ is the $N_p \times M$ matrix.

To facilitate the construction of the mapping of the material field function to the structural topology, it is necessary to ensure that the material field is bounded in $[-1, 1]$; accordingly, it follows that

$$-1 \leq \boldsymbol{\eta}^T \boldsymbol{\Lambda}^{-1/2} \boldsymbol{\Phi}^T \mathbf{C}_d(\mathbf{x}_i) \leq 1, \quad (i = 1, 2, \dots, N_p) \quad (9)$$

By squaring the Eq. (9), we get

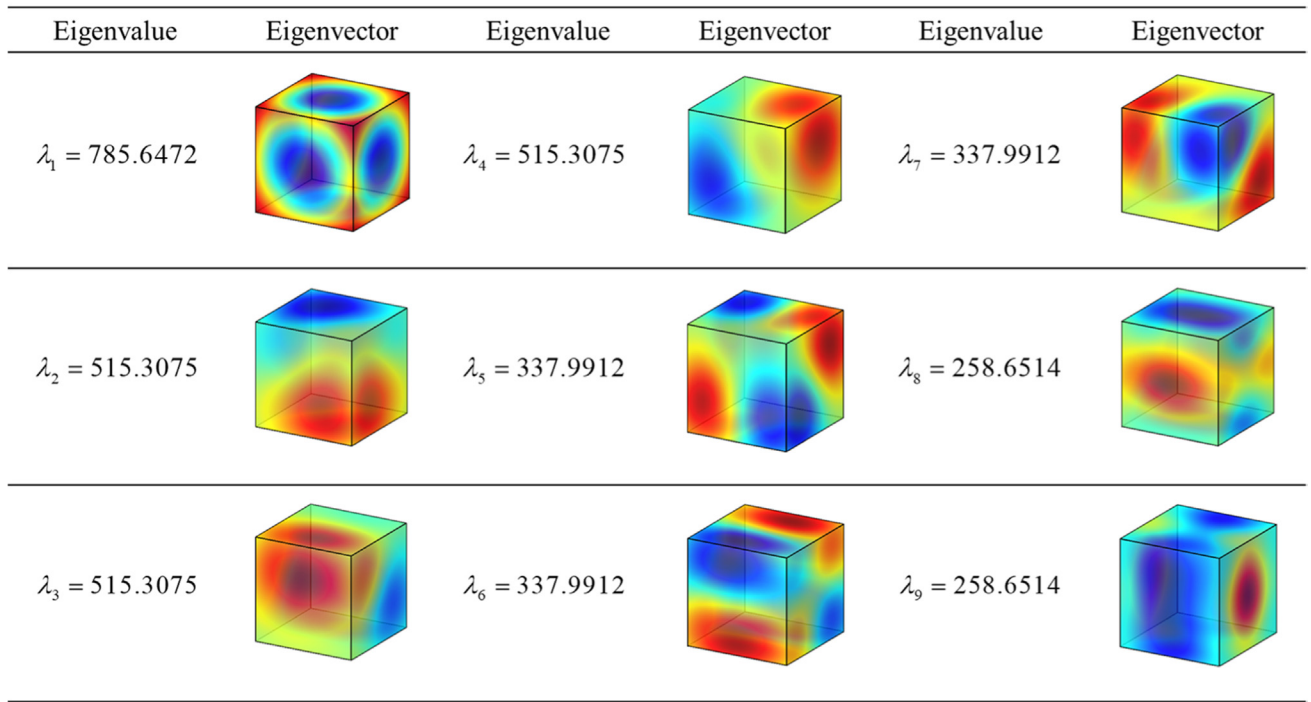


Fig. 1. The first nine orders of eigenvalues and eigenvectors of the correlation matrix, based on the cube material field.

$$\boldsymbol{\eta}^T \boldsymbol{\Lambda}^{-1/2} \boldsymbol{\Phi}^T \mathbf{C}_d(\mathbf{x}_i) \mathbf{C}_d(\mathbf{x}_i)^T \boldsymbol{\Phi} \boldsymbol{\Lambda}^{-1/2} \boldsymbol{\eta} \leq 1, \quad (i = 1, 2, \dots, N_p) \quad (10)$$

Let's introduce the matrix $\mathbf{W}_i = \boldsymbol{\Lambda}^{-1/2} \boldsymbol{\Phi}^T \mathbf{C}_d(\mathbf{x}_i) \mathbf{C}_d(\mathbf{x}_i)^T \boldsymbol{\Phi} \boldsymbol{\Lambda}^{-1/2}$, Eq. (10) could be given as follows:

$$\boldsymbol{\eta}^T \mathbf{W}_i \boldsymbol{\eta} \leq 1, \quad (i = 1, 2, \dots, N_p) \quad (11)$$

2.3. MFSE-based topology optimization model

At each finite element, Young's modulus E for the isotropic material is expressed using a power-law interpolation function, i.e.,

$$E = E_{\min} + \left(\frac{1 + \chi(\mathbf{x}_e)}{2} \right)^p (E_0 - E_{\min}), \quad (e = 1, 2, \dots, N_e) \quad (12)$$

where E_0 and $\mathbf{x}_e$ denote the Young's modulus for the solid material and the center coordinate point of the e -th element, respectively. A smaller value, $E_{\min} = 10^{-9} E_0$, is taken to avoid the singularity of the stiffness matrix, and $p = 3$ is the penalty coefficient.

Through the material field interpolation and the finite element analysis, the MFSE-based topology optimization model in $\boldsymbol{\eta}$ -space can be expressed as follows:

$$\begin{aligned} \min_{\boldsymbol{\eta}} \quad & f(\mathbf{u}, E(\boldsymbol{\eta})) \\ \text{s.t.} \quad & \begin{cases} G(\mathbf{u}) = 0 \\ g_k(\mathbf{u}, E(\boldsymbol{\eta})) \leq 0 \\ \boldsymbol{\eta}^T \mathbf{W}_i \boldsymbol{\eta} \leq 1 \end{cases} \end{aligned} \quad (13)$$

where the MFSE coefficients $\boldsymbol{\eta} = \{\eta_1, \eta_2, \dots, \eta_M\}^T$ are taken as the design variables, $f(\mathbf{u}, E(\boldsymbol{\eta}))$ is the performance function, and $E(\boldsymbol{\eta})$ is the isotropic material's Young's modulus as determined by Eq. (12).

Following the truncation of the material field series, the topological configuration of the structure can be approximated to the series with a small number of design variables. When working with 2D problems, the design space dimension of topology opti-

mization can typically be reduced to 50–100; in the case of 3D problems, it can generally be reduced to fewer than 200. Compared with the density-based topology optimization method, the MFSE method can reduce the design variables by several orders of magnitude, which also provides the possibility for the practical application of the gradient-free topology optimization method for solving complex nonlinear problems.

3. DNN-MFSE optimization method

This section presents a combination of the neural network and the MFSE model. First, the background knowledge of neural networks is reviewed. Next, three major strategies for solving the MFSE topology optimization model, combined with the DNN-based surrogate optimization algorithm are presented. To improve the accuracy of the surrogate, new neural network architecture and a method for adaptively selecting hyperparameters are presented.

3.1. An overview of the neural network

The surrogate model plays an important role in dealing with the computationally expensive analysis process [39]. In terms of efficiency, compared with the original optimization problem, optimization based on the surrogate model can obtain a high-quality optimized design within a short time. The surrogate model is actively employed in various fields of academia and industry [40–44] and includes the commonly used Kriging, artificial neural network (ANN), and support vector machine (SVM) models. In particular, ANNs have recently gained considerable attention due to their flexibility and universal approximation abilities. However, only a small amount of work applied neural networks as surrogate optimization models for topology optimization. Generally, ANNs can be divided into shallow neural networks and DNNs, based on their number of layers. A shallow neural network is a model that includes only one hidden layer. In complicated problems, e.g., topology optimization design, the input and output usually show a high degree of nonlinear dependence. It is difficult, even when

working with the compliance-minimization problem, to fully describe the mapping relationship using a shallow neural network, which presents the need for using a DNN method.

Compared with the traditional shallow neural network, a significant improvement delivered by a DNN is an increase in the number of hidden layers as shown in Fig. 2, which strengthens the complexity of the model and its capability for managing problems related to high nonlinearity [45]. According to the general approximation theorem [46], a neural network can approximate any continuous function with arbitrary precision. Several DNNs with a considerable number of hidden layers were proposed to address real problems [47–49].

In a forward neural network, the mathematical formulation between the output on the j -th neuron z_j and the input x_i to the next layer can be expressed as

$$z_j = f\left(\sum_i (w_{ij}x_i + b_j)\right) \quad (14)$$

where $f(\cdot)$ denotes the nonlinear activation function, while w_{ij} and b_j denote the weight vector and bias term, respectively.

For each layer of the neural network, the parameters are independent and the multiple layers enable the mapping of complex behaviors between inputs and outputs. Through a layer-by-layer feedforward method, the output of the neural network with n_l layers can be obtained as

$$M(\mathbf{x}, \theta) = f^{(n_l-1)}(f^{(\dots)}(f^{(2)}(f^{(1)}(\mathbf{x})))) \quad (15)$$

where M denotes the DNN output, θ denotes the parameters of the network containing w_{ij} and b_j , $f^{(i)}$ denotes the activation function enacted by the i -th layer neural network, and n_l is the number of layers.

3.2. Optimization strategies in the DNN-MFSE model

To efficiently obtain a solution of the MFSE topology optimization model in Eq. (13), a DNN-based model was used as a surrogate for the topology optimization problem. First, the topology optimization model needs to be transformed into an unconstrained form as follows:

$$\min_{\boldsymbol{\eta}} f_{\text{obj}}(\boldsymbol{\eta}) = f(\mathbf{u}, E(\boldsymbol{\eta})) + p_0 \max(g_k, \boldsymbol{\eta}^T \mathbf{W}_i \boldsymbol{\eta} - 1, 0) \quad (16)$$

where p_0 is the penalty multiplier and is set as

$$p_0 = 10^{\text{floor}(1 + \log_{10} |f(\mathbf{u}, \rho(\boldsymbol{\eta}))|)} \quad (17)$$

Due to the samples being distributed uniformly, the Latin hypercube sampling was chosen as the sampling technology in this paper to get the various MFSE design variables $\boldsymbol{\eta}$. Then, MFSE design variables $\boldsymbol{\eta}$ is taken as a neural network input, and the unconstrained objective $f_{\text{obj}}(\boldsymbol{\eta})$ is set as the output; in this way, mapping throughout the neural network can be achieved. The neural network learns through the use of the corresponding input and output via back-propagation in the training phase, i.e., deep learning. In this paper, the task of the neural network is to establish regression mapping from the design variables to the penalty objective function, so that the mean square error between the predicted penalty objective function $M(\boldsymbol{\eta}, \theta)$ and the real penalty objective function $f_{\text{obj}}(\boldsymbol{\eta})$ is selected as the loss function in the DNN model, which can be written as the following optimization problem:

$$\min_{\theta} : F = \sum_{i=1}^N \|M(\boldsymbol{\eta}_i, \theta) - f_{\text{obj}}(\boldsymbol{\eta}_i)\|_2 \quad (18)$$

where $f_{\text{obj}}(\boldsymbol{\eta}_i)$ is the unconstrained objective value corresponding to $\boldsymbol{\eta}_i$, operator $\|\cdot\|_2$ represents 2-norm, and N is the total number of input samples.

To increase the robustness of the algorithm in the optimization method, based on DNN and MFSE, this study introduces three major strategies, i.e., a sub-domain self-movement strategy, a local adding-point strategy, and a sub-domain adaptive reduction strategy.

3.2.1. The sub-domain self-movement strategy

When solving topology optimization problems using surrogate models, it is inefficient to directly use design variables in the entire domain as samples for the objective function. Not only does this require enough sample points to improve the accuracy of the neural network-based surrogate model, but there may also be many sample points with poor performance that make the surrogate optimization process difficult. Forrester et al. [50] pointed out that the only solution for complex function surrogates was to use a local strategy so that the performance of the objective function would be relatively simple, and fewer samples could be used to better surrogate the objective function. Accordingly, we adopted

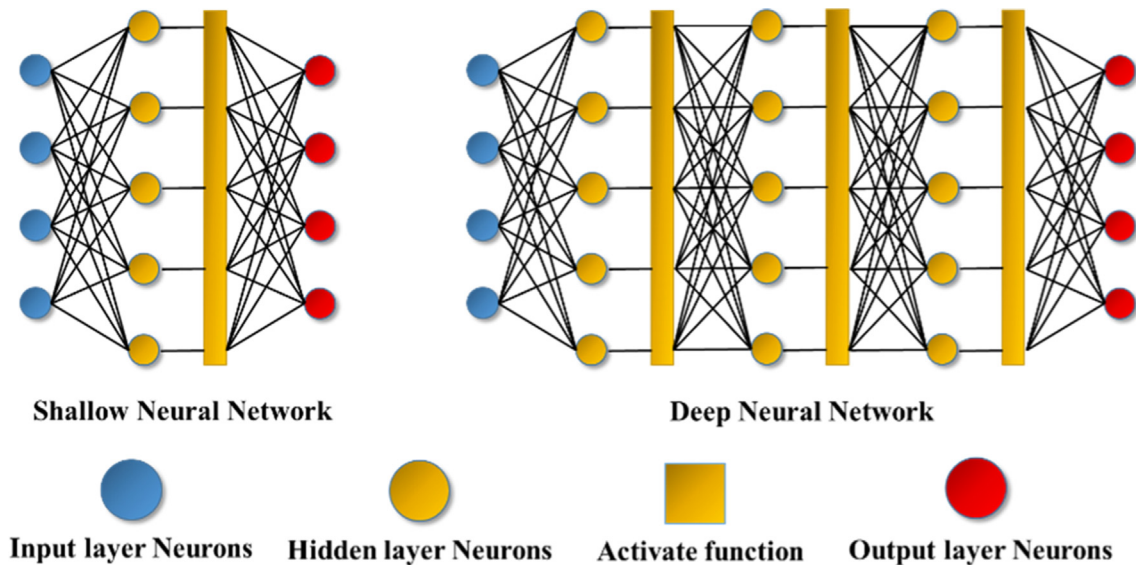


Fig. 2. Shallow neural network and deep neural network.

a sub-domain self-movement strategy for the DNN-surrogate model.

3.2.2. The local adding-point strategy

Typically, it is uneconomical for the DNN surrogate model to yield very accurate predictions at sample points far removed from the optimal point. High accuracy is required in a potentially promising region for the DNN-surrogate model. Accordingly, the local adding-point strategy can be used to locally increase the sampling points in the sub-domain where the objective function is better, so that the neural network has a higher approximation accuracy in the current sub-domain and can more promisingly find an optimized point.

3.2.3. The sub-domain adaptive reduction strategy

In topology optimization, it will be difficult for the DNN surrogate model to accurately represent the performance function by a small sample point if a larger search domain is considered. Furthermore, structural disconnection probability in sample points will be increased, giving rise to difficulties within the surrogate model. In this regard, the range of the search sub-domain should be reduced in the iterative process, which is in line with the method using a “rough search” first, followed by a “fine search”.

According to the above three strategies, this study gives the following empirical formulation for adjusting the design domain:

$$\Omega_k = \{ \|\boldsymbol{\eta} - \boldsymbol{\eta}^{(k)}\|_{\infty} \leq r_k \} \quad (k = 0, 1, 2, \dots) \quad (19)$$

$$\boldsymbol{\eta}_{k+1} = \{ \boldsymbol{\eta}^* \mid \min_{\boldsymbol{\eta} \in \Omega_k} f_{\text{obj}}(\boldsymbol{\eta}) \} r_{k+1} = 0.9r_k$$

where k represents the number linked to the sub-optimization problem, $\|\cdot\|_{\infty}$ represents the infinite norm, and $\boldsymbol{\eta}^{(k)}$ is the central point of the k -th sub-problem.

Furthermore, the respective criteria for stopping each sub-domain optimization iteration and the entire optimization are

$$M(\boldsymbol{\eta}^{(n+1)}, \boldsymbol{\theta}) - M(\boldsymbol{\eta}^{(n)}, \boldsymbol{\theta}) \leq 0.01 \cdot y_{\min} \quad \text{or} \quad N_k \leq 50 \quad (20)$$

and

$$\|\boldsymbol{\eta}^{(k+1)} - \boldsymbol{\eta}^{(k)}\|_{\infty} \leq 0.01, \quad (21)$$

where $M(\boldsymbol{\eta}^{(n)}, \boldsymbol{\theta})$ is the n -th optimized prediction of the DNN, $y_{\min}$ is the current minimum of f_{obj} , and N_k is the number of resamples in the k -th sub-domain. The first stop criterion in Eq. (20) indicates that the sub-optimization problem changes when reaching the maximum number of resamples or, alternatively, when finding the optimized solution via a neural network. The second criterion in Eq. (21) indicates the final optimized solution obtained when the two sub-problem solutions show relatively little change.

3.3. Neural network hyperparameter selection

In each sub-domain, in order to save finite element costs, the number of Latin hypercube samples was small for the neural network. To avoid overfitting in the neural network, a regularization strategy was required. This study used the BatchNorm and Dropout layers for regularization. To reduce the computational cost and improve the surrogate performance of the model, the PRelu function, rather than the commonly used Sigmoid function, was used as the activation function, and the layers including the Linear activation layer, BatchNorm layer, PRelu function, and Dropout layer were set as a basic code block in the hidden layer. Then, the number of hidden layer blocks was set as a hyperparameter of the neural network structure, as shown in Fig. 3.

Next, the remaining neural network hyperparameter strategies were introduced. The parameters representing the network's predefined architecture, e.g., the number of hidden layers, were labeled the hyperparameters. As noted above, these hyperparameters

had a significant influence on the performance of the neural network, and the value of the hyperparameters could be discrete or continuous. This represents a typical non-convex combinatorial optimization problem that is difficult to solve by traditional optimization methods. In the process of topology optimization, because this study adopted a movement strategy of the design domain, the behavior of the local function of each generation will change significantly. Therefore, it is particularly important to develop an adaptive network hyperparameter selection strategy to improve surrogate accuracy.

Herein, based on the self-adaptive reduction and movement strategy of design domain, a self-adaptive selection algorithm of network hyperparameters is proposed and its key ideas are outlined as follows:

- (1) Given the selection range of the hyperparameters, one of the hyperparameters was randomly selected as the initial value of the iterative process.
- (2) According to the different selection ranges of the various hyperparameters, given a search radius and the construction of a sub-domain hypercube, the hyperparameters in this domain were randomly sampled to calculate the total error of the test sample (note that the test sample was not visible to the neural network) until the maximum number of samples was reached.
- (3) The total error of each hyperparameter was compared; the best hyperparameter selection in the sub-domain was obtained and used as the center point for moving the sub-domain, as well as to reduce the search radius until the hyperparameter selection reached convergence or the maximum computational cost.

In short, this study used the number of code blocks, the learning rate, the dropout rate, and the number of neurons in each layer as hyperparameters, and adopted the above self-adaptive selection algorithm. The ranges of these hyperparameters are shown in Table 1.

Theoretically, the DNN surrogate model predicts accuracy will be greater as the number of samples increases. A qualitative test was carried out to verify the influence of the number of samples on the self-adaptive neural network performance. The problem of minimizing the compliance of the MBB beam structure is taken as an example, and the numerical conclusion is shown in Fig. 4.

It can be seen from Fig. 4 that the performance of the self-adaptive DNN model increases rapidly as the number of samples increases when the number of samples is less than 300, but it becomes slower when the number of samples further increases. Thus, it is reasonable to control the number of samples per-Latin hypercube sampling with 300 in this research.

To show the influence of the search domain area on the surrogate accuracy of the DNN-based model and verify the effectiveness of the proposed hyperparameter self-adaptive method, this paper takes the former cases in Fig. 4 for verification. We express the design domain as follows:

$$\|\boldsymbol{\eta} - \boldsymbol{\eta}^{(0)}\|_{\infty} \leq r \quad (22)$$

where $\boldsymbol{\eta}^{(0)}$ is the central point of the sub-problem design domain, and r is the parameters that control the size of the sub-problem design domain. Then, 300 sample points are taken to calculate the structural performance, of which 90% are used to train the DNN model, and 10% are used to test the DNN model.

In surrogate-based optimization, the average error affects the overall performance of the optimization, and the standard deviation affects the position of the sample point in the local adding-point strategy; as such, comprehensive consideration of the above

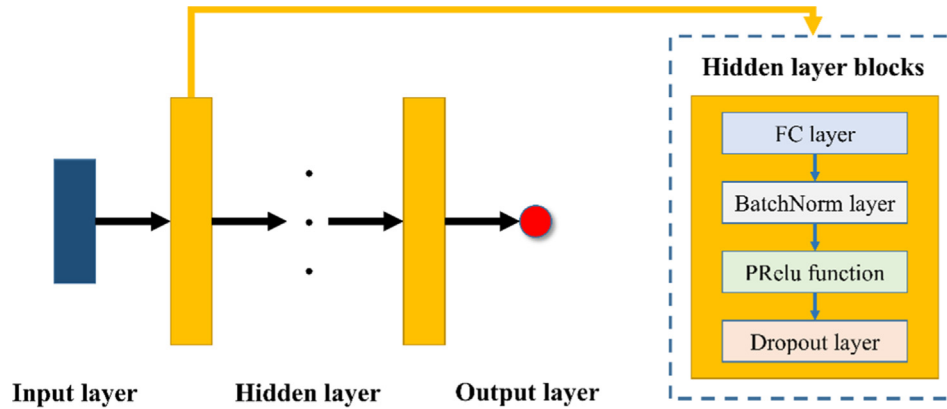


Fig. 3. Basic blocks in deep neural network.

Table 1

The hyperparameter ranges that were used in this study.

Hyperparameter	Ranges	Parameter Type	Hyperparameter	Ranges	Parameter Type
Blocks numbers	[1, 4]	Discrete	Number of nodes	[4, 128]	Discrete
Learning rate	$[10^{-5}, 10^{-1}]$	Continuous	Dropout rate	[0.2, 0.5]	Continuous

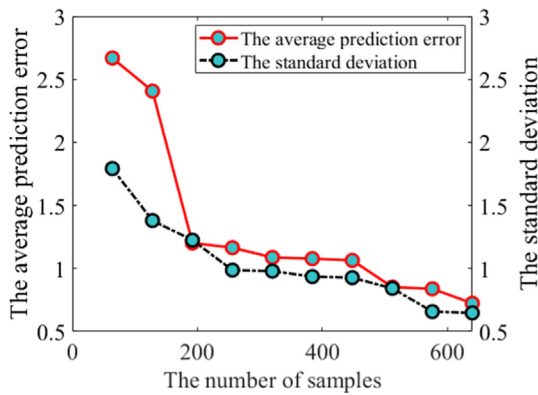


Fig. 4. The average prediction error and the standard deviation of the self-adaptive deep neural network under the different number of samples.

two indicators is needed to evaluate the surrogate performance. The average prediction error A_{error} and the mean standard deviation M_{std} between the objective $\hat{y}_i$ and the predicted y_i is calculated as

$$A_{\text{error}} = \frac{1}{30} \sum_{i=1}^{30} (y_i - \hat{y}_i)^2$$

$$M_{\text{std}} = \frac{1}{30} \sum_{i=1}^{30} ((y_i - \hat{y}_i)^2 - A_{\text{error}})^2 \quad (23)$$

The relationship between the average prediction error A_{error} and mean standard deviation M_{std} of the surrogate model with the domain area and design space dimensions are plotted in Fig. 5. The figure shows that the average error and mean standard deviation of the DNN model decrease as the range of the search area is reduced. In the MBB beam case, the search range of the surrogate model, which is controlled in $\|\eta - \eta_0\|_{\infty} \leq 0.4$, is ideal. Moreover, Fig. 5 shows that the performance of the optimized neural network has been greatly improved, whereby the average and standard deviations of its error have decreased to some extent under different conditions, thereby verifying the importance of adopting the hyperparameter self-adaptive selection strategy.

In summary, the DNN optimization approach in the sub-domain proposed in this study is shown in Fig. 6. It is noted that the number of sample points for each Latin hypercube sampling is within a few hundred, which is a small sample numbers for deep learning; thus, the time required for training a neural network only is very short, and the hyperparameter adaptive change process is only performed when the projection parameter increases. Accordingly, the proportion of the total time required to implement the above algorithm within the entire iteration process can be ignored in numerical practice compared with finite element evaluations.

The topology optimization method of DNN-MFSE can be summarized as follows. First, the given design domain is expanded using the material-field series expansion method and undergoes a reduction in the number of design variables. Second, the center position and the search area of the initial design sub-domain are given, and Latin hypercube sampling is performed for the initial design domain. Then, sample points are divided into training and testing datasets, and the network hyperparameters are automatically selected when the projection parameter is increased. Third, using an optimized neural network as a surrogate penalty objective function, and based on a local adding-point strategy and surrogate optimization, the best sub-domain point was found and selected as the center point of the next self-adaptive sub-domain until the topology optimization reached convergence. Fig. 7 shows the 3-step workflow program that was used to construct the DNN-MFSE as a means to solve topology optimization problems.

3.4. Comparison between the DNN-MFSE and the KG-MFSE

In [25], a KG-MFSE gradient-free topology optimization method was developed. This method included advantages related to problems when fewer design variables were involved. However, it is difficult to deal with the problem with a large number of design variables (e.g., $M \geq 100$). To quantitatively compare the accuracy of the two surrogate methods (Kriging and DNN models), the effects of the search-field area and the design variable dimensions on the prediction accuracy and the mean standard deviation are compared in Fig. 8, based on the dataset presented in Section 3.3.

Fig. 8 shows that, compared with the Kriging model, the self-adapted DNN model is less sensitive to an increase in dimensional-

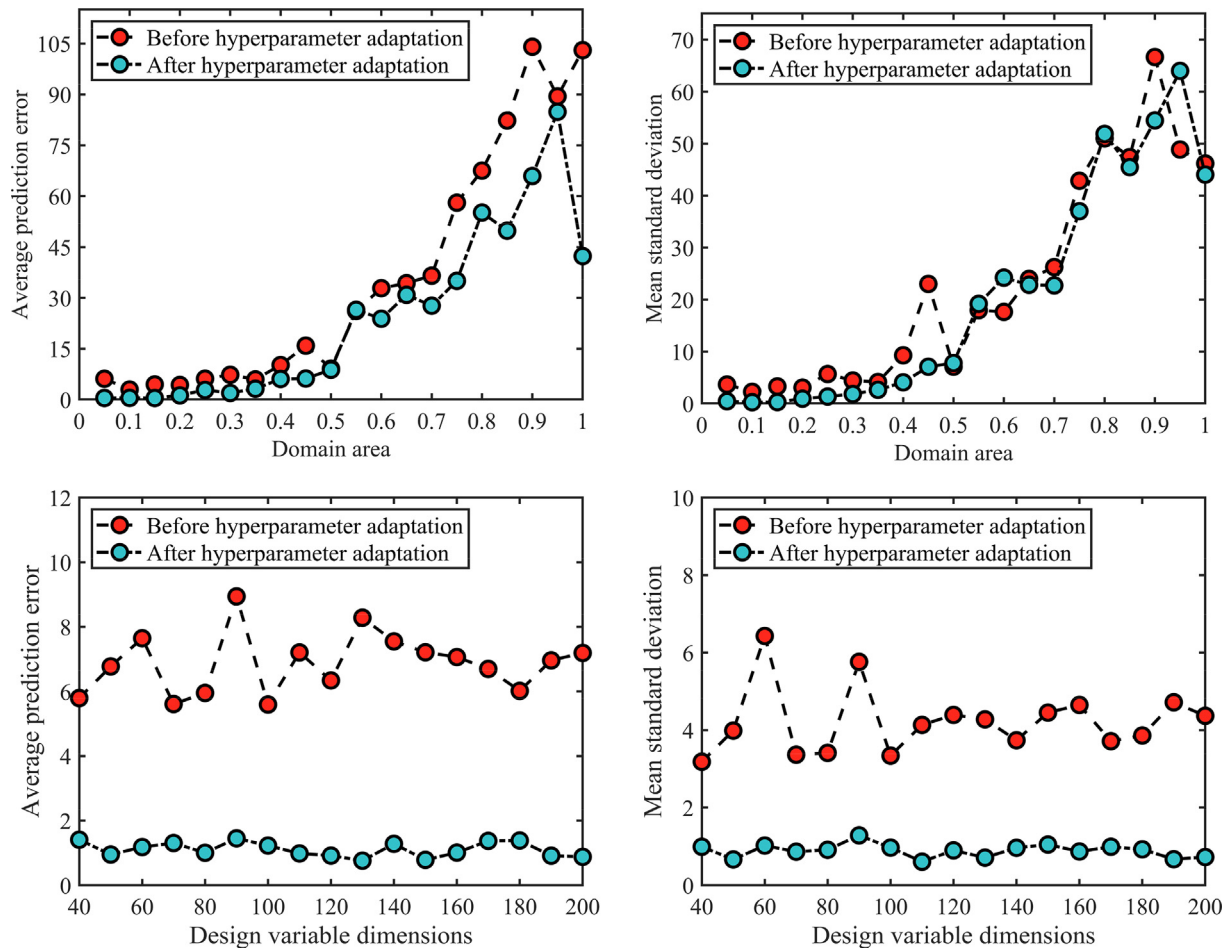


Fig. 5. The average prediction error and the standard deviation of deep neural network before and after hyperparameter adaptation.

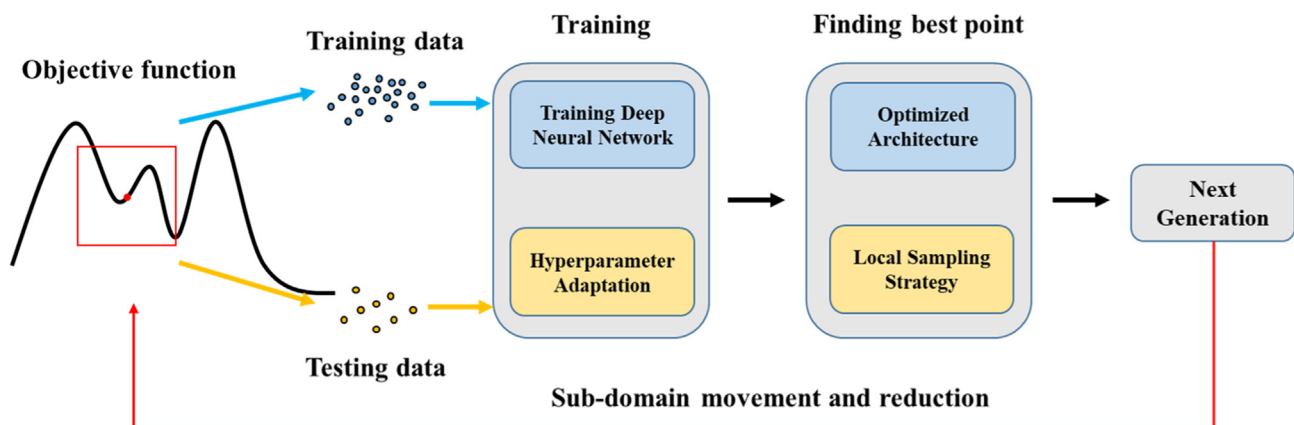


Fig. 6. The optimization algorithm in the sub-domain.

ity and, as the domain area expands, the decrease in surrogate performance is also within the acceptable range. Regarding the above observation, for high-dimensional MFSE-based topology optimization problems, the introduction of the self-adapted DNN model

may solve problems that are difficult to solve using the Kriging model. The subsequent numerical examples showed that the method proposed in this study was very effective, even for application to complex topology optimization problems.

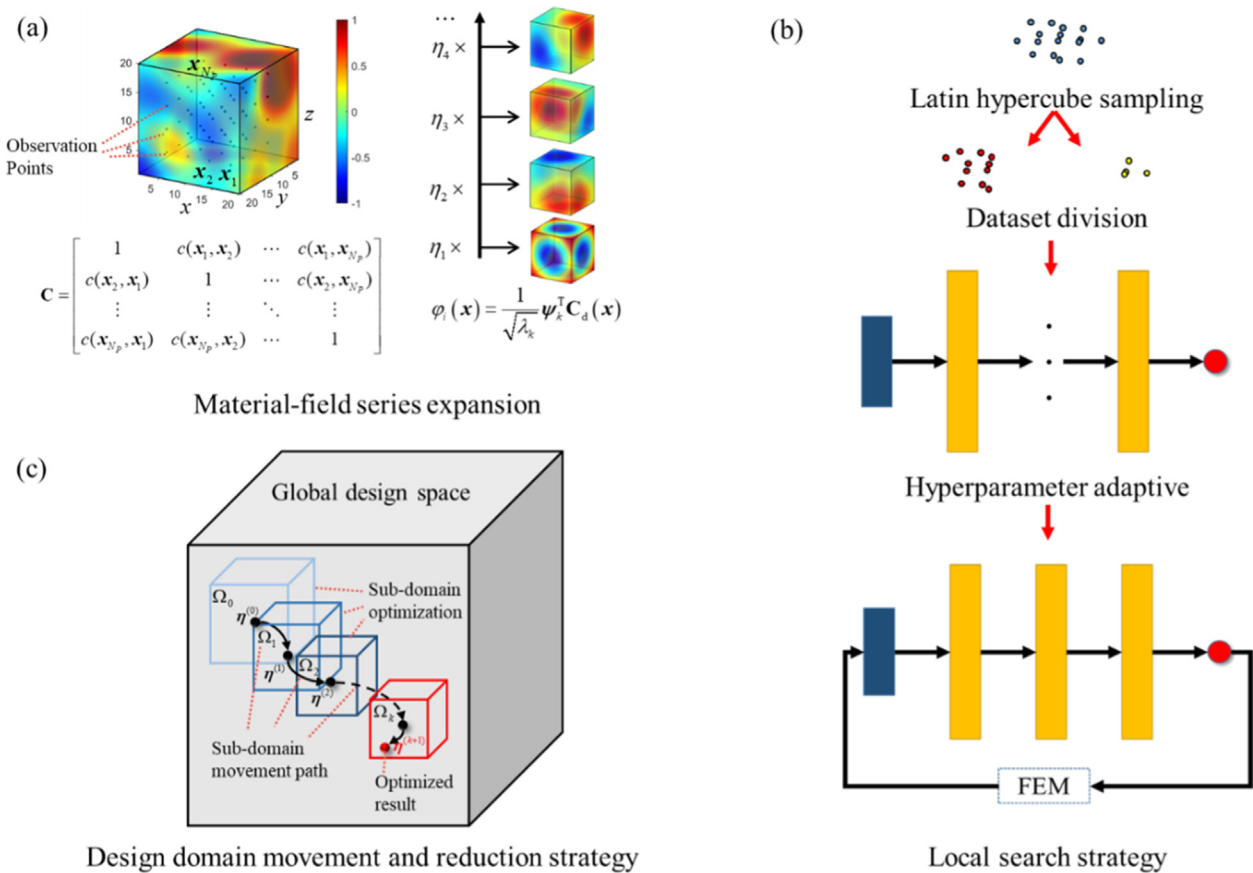


Fig. 7. The workflow of the DNN-MFSE for solving topology optimization problems. (a) The pre-processing operation based on MFSE; (b) the neural network adaptation and local search strategy; (c) the design domain movement and reduction strategy.

4. Numerical examples

In this section, the proposed DNN-MFSE method is used to solve several typical topology optimization problems, including 2D and 3D compliance-minimization problems, the maximum volume modulus design of microstructures, and the stress minimization of a dropped mobile phone problem. By comparing the optimization results with the gradient method and the Kriging-based gradient-free method, the versatility and robustness of the proposed method were underscored.

4.1. Compliance-minimization 2D topology optimization problems

In order to facilitate a comparison with other gradient-free topology optimization techniques, two typical compliance-minimization cases were considered using the same optimization settings, i.e., a short cantilever beam and an MBB beam.

Young's modulus of the solid material was set to $E_0 = 200$ GPa, and Poisson's ratio was set to $\nu = 0.3$. The objective was to minimize the compliance with a volume ratio of 50%. As shown in Fig. 9, the right middle of the short cantilever beam is subjected to a point force of $F = 1000$ N, while the left side is fully fixed. The size of the design domain is 80×80 mm, uniformly distributed in 40×40 observation points, and discretized into 80×80 four-node plane-stress elements. When the truncation accuracy was 99.9%, the number of design variables in this example was 50. Table 2 shows the optimization results obtained by using the proposed method for two random experiments. The results obtained by this method were compared with those obtained by the

gradient-based MFSE method [24]. It can be seen from Table 2 that the topologies were essentially the same, and the difference between the objective functions was negligible, which indicated that the proposed DNN-MFSE method could effectively solve the problem without requiring the sensitivity of the objective function.

Fig. 10 shows the optimization history of the penalty objective function. The reason for fluctuations in the objective function was due to the Latin hypercube sampling in each sub-optimization. The iterative process among sub-optimizations was robust. The number of finite element (FE) evaluations required by the proposed method was 1941 with 14 Latin hypercube sampling processes.

It is noted that the similar topology optimization problem with a coarse mesh (having only $11 \times 11 = 121$ FEs) has also been performed by other non-gradient methods, such as the density-based GSA and SOLO methods [19]. The GSA performed 4×10^6 FE evaluations while the SOLO converged after 10,243 FE function calculations. Compared with directly using GSA and SOLO, the proposed DNN-MFSE can handle very fine mesh, and simultaneously reduces the number of FE evaluations to 1/2000 and 1/5, respectively.

Further, to exhibit the changes of hyperparameters during the optimization process, the hyperparameters function curves are plotted in Fig. 11. It can be seen that the number of blocks for the self-adaptive neural network is kept at 3–4, which is consistent with the aforementioned description: shallow neural networks cannot accurately surrogate the topology optimization problem. The dropout rate in the first sub-problem remains high due to the sub-domain being relatively large in the early stages of optimization, and it becomes smaller during the last several sub-

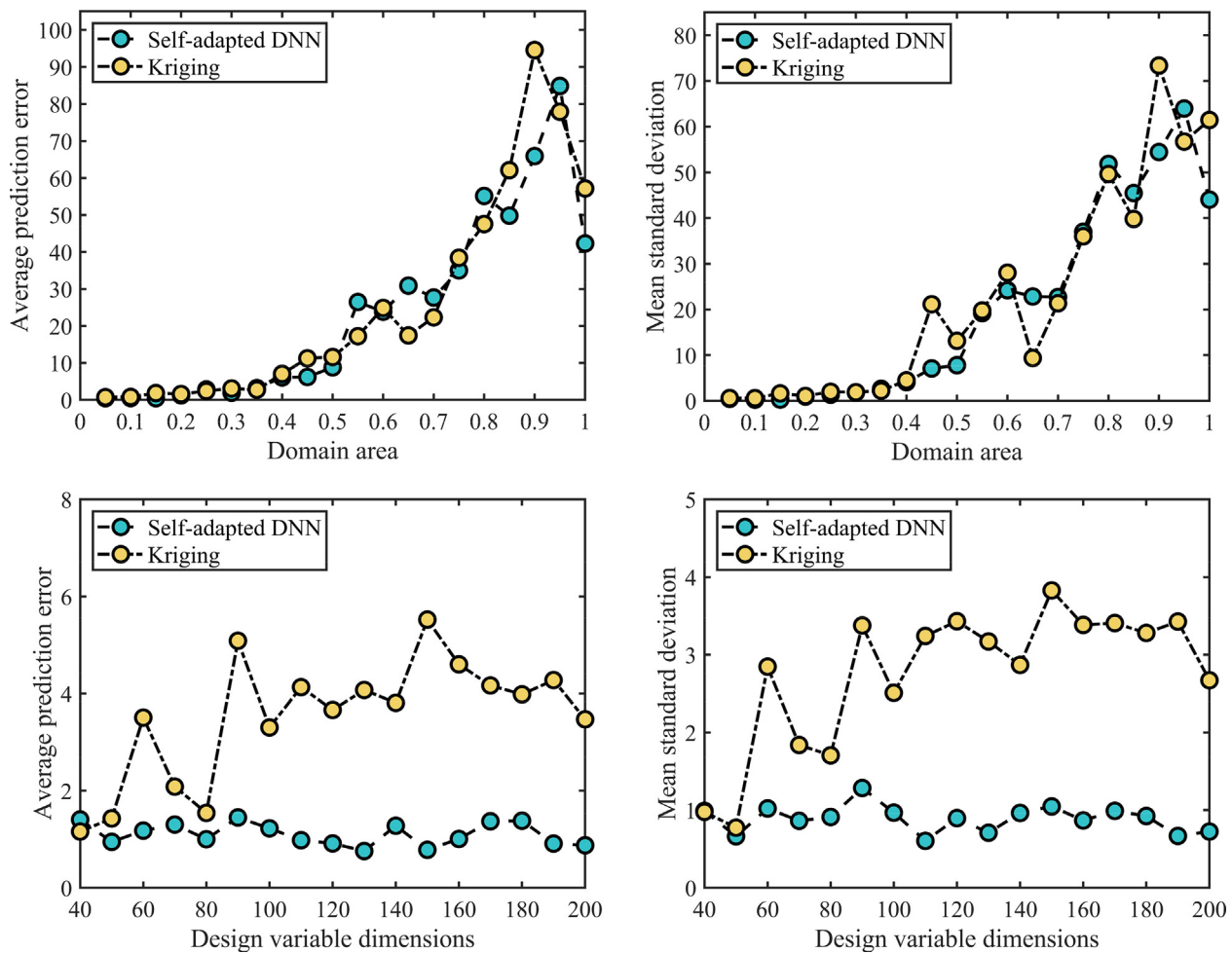


Fig. 8. Comparison of Kriging surrogate model and self-adapted deep neural network for the average prediction error and the standard deviation.

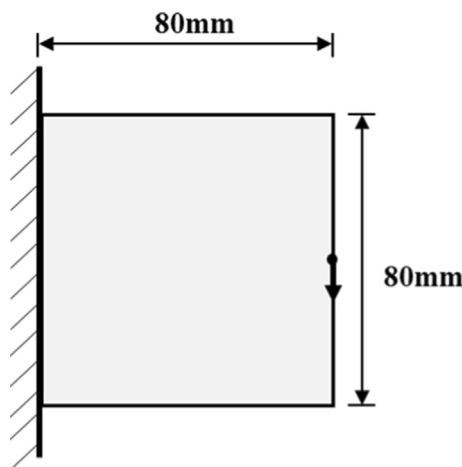


Fig. 9. Cantilever beam design domain.

problems, proving the necessity of a hyperparameter self-adaptive selection strategy.

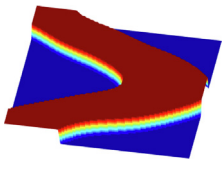
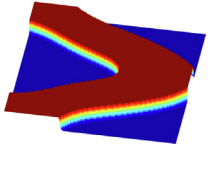
The second numerical example is an MBB beam as shown in Fig. 12. The elastic modulus and Poisson's ratio were the same as in the cantilever problem; the design domain length of the structure was 180 mm, and the width was 30 mm. A concentrated force

of 1000 N was applied at the center of the top edge, and the volume ratio was 50%. Only half of the design domain, with 90×30 uniformly distributed observation points and 180×60 plane stress four-node elements, were included in the analysis. After truncation of the material-field series, the number of MFSE design variables was 64.

The optimization processes of the DNN-MFSE were tested twice randomly and the results are shown in Fig. 13. Because of the random sampling, the results of each run were slightly different; however, it can be seen that a clear structural topology and load transfer path were obtained for each optimization, and the difference between the target compliances was slight. The iterative history of the penalty objective function and the corresponding topologies are given in Fig. 14, which shows that the optimization history of the objective function was stable during 3398 finite element analyses and 17 sub-domain searches. Although the number of finite element calls for this problem was higher than the KG-MFSE method, it should be noted that this method was still more effective than the non-gradient topology optimization algorithm based on the genetic algorithm or the gradient-based method with sensitivity difference approximation. It is noted that the proposed DNN-MFSE method was able to simultaneously employ computing in parallel with accelerating the finite element calculation process in the Latin hypercube sampling, and this process could easily be applied to complex topology optimization problems.

To verify the stability of the proposed method, we ran the optimization procedure of the cantilever beam and the MBB beam 10

Table 2
Optimized result of different methods.

Method	Compliance (Nm)	Topology configuration	Material field distribution
DNN-MFSE (test-1)	0.7562		
DNN-MFSE (test-2)	0.7538		
Gradient-based MFSE in [24]	0.7472		

times, respectively, calculated their statistical characteristics, and compared them with those of the KG-MFSE method. The results in Table 3 show that for each problem, the mean and standard deviation of compliance using the proposed DNN-MFSE optimization method is significantly smaller, and show a more outstanding and stable performance compared with the KG-MFSE.

4.2. Compliance-minimization of 3D topology optimization problems

In this example, to further prove the DNN-MFSE algorithm's abilities, a 3D long cantilever beam with the design domain of $48 \text{ mm} \times 16 \text{ mm} \times 12 \text{ mm}$ (Fig. 15) is selected for optimization. A uniform distributed load, i.e., $q = 100 \text{ N/mm}$, is applied on the bottom of the free end, and the allowable volume ratio is set to 15%. This example is to show that our method can handle more design variables than the KG-MFSE, which makes the 3D topology optimization problem can be solved efficiently by the gradient-free method became realistic.

Given the symmetry in the Z-direction, one-half of the design domain was considered. First, a coarse discretization was established for the material field representation and the finite element analysis, in which $24 \times 8 \times 3$ observation points were uniformly distributed in the design domain, and the mesh was divided into $48 \times 16 \times 6$ eight-node solid elements. After the truncation of the material-field series with $l_c = 13 \text{ mm}$, a total of 108 design variables were included in this optimization. After 2198 cycles of finite element analysis, convergence was reached. The optimization history with corresponding evolutionary topologies is shown

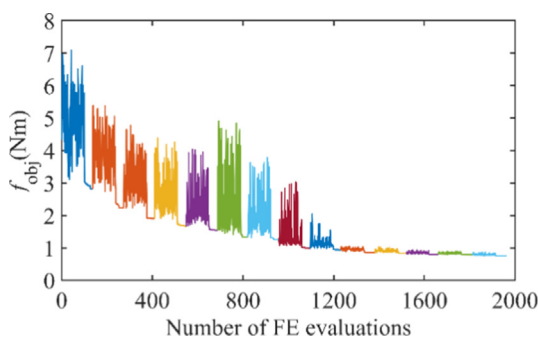


Fig. 10. Optimization history for cantilever beam by DNN-MFSE.

in Fig. 16 and the compliance value of the final design, shown in Fig. 17(a), is 4.41 Nm, which is only 1.35% greater than the value (4.35 Nm) obtained using the gradient-based MFSE method, thus indicating the ability and accuracy of the method for solving 3D problems. To the author's knowledge, this is the first time that gradient-free algorithms were successfully applied to 3D topology optimization examples.

In addition, we set the correlation length to be $l_c = 9.5 \text{ mm}$ and, accordingly, the number of design variables was 200 according to Equation (8), which could not be effectively solved by the KG-MFSE method. In this case, $48 \times 16 \times 6$ observation points were uniformly distributed in the design domain and the mesh was divided into $96 \times 32 \times 12$ eight-node solid elements. The number of finite element calls required was 3424. The final topological configuration is shown in Fig. 17(b) and the corresponding compliance value is 3.80 Nm, which is only 3.23% greater than the gradient-based solution (3.68 Nm). The result further indicated the ability of the proposed method to solve topology optimization problems with relatively large design variable dimensions.

4.3. Bulk modulus maximization for the microstructure design

Cellular materials with periodic microstructures have received extensive attention due to their various advanced physical, mechanical, and thermal properties compared with solid materials, and their material properties can be obtained by numerical homogenization. It is well-established that the microstructure design is a typical problem with multiple local solutions. When using a gradient-based algorithm to solve this problem, a solution strongly depends on the initial guess of the structure [51].

Here, we want to show in this example the superior performance of DNN-MFSE compared to the gradient method for multi-local solution problems such as the maximizing microstructures bulk modulus. To ensure the orthotropy of the microstructure, the unit cell design domain was established considering geometry symmetry. The MFSE topology optimization model is given as follows:

$$\begin{aligned}
 \min_{\boldsymbol{\eta}} \quad & -\sum_i \sum_j \mathbf{D}_{ij} \\
 \text{s.t.} \quad & \begin{cases} \mathbf{K}\boldsymbol{\chi}^i = \mathbf{f}^i \\ \sum_e v_e \leq V^* \\ \boldsymbol{\eta}^T \mathbf{W}_i \boldsymbol{\eta} \leq 1 \end{cases}
 \end{aligned} \quad (24)$$

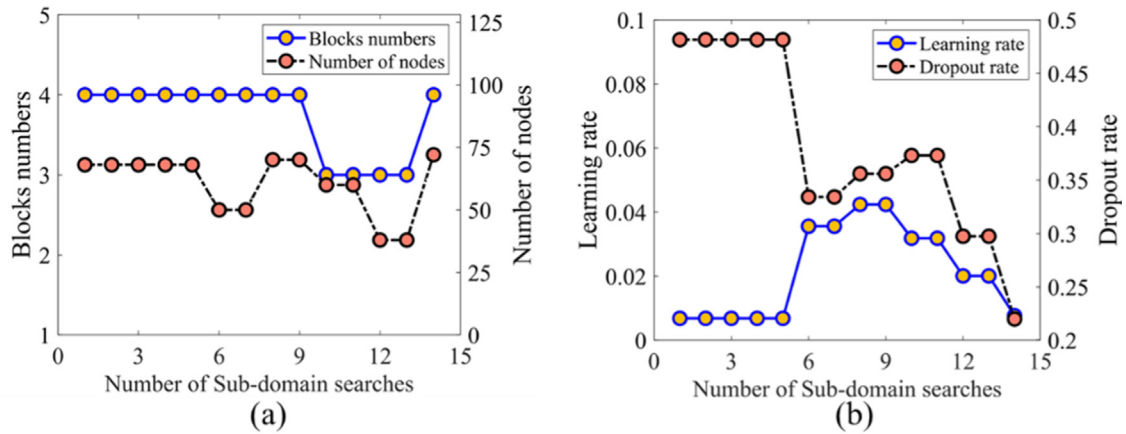


Fig. 11. The hyperparameter self-adaptive history in cantilever beam case. (a) Discrete hyperparameters. (b) Continuous hyperparameters.

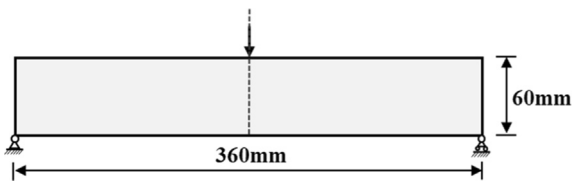


Fig. 12. Messerschmitt-Bölkow-Blohm (MBB) beam design domain.

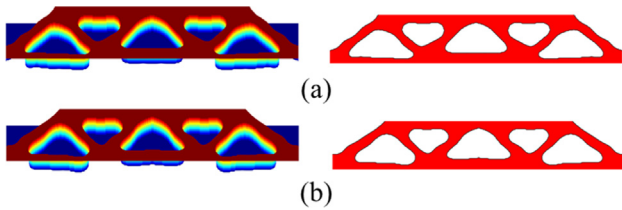


Fig. 13. The results of MBB beams by DNN-MFSE with two random tests (left, material field distribution; right, structure topology). (a) Test 1, compliance is 10.2370; (b) Test 2, compliance is 10.2742.

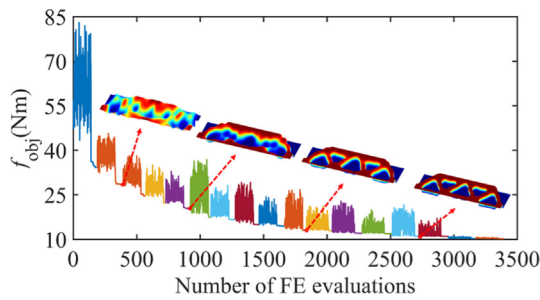


Fig. 14. Optimization history for MBB beam by DNN-MFSE.

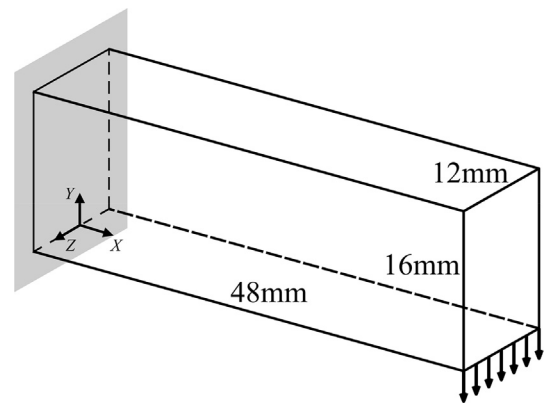


Fig. 15. The 3D cantilever beam design domain.

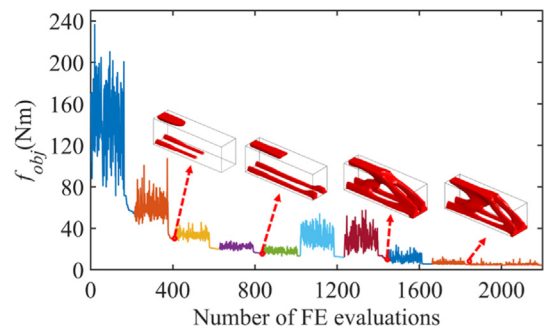


Fig. 16. Optimization iterative history of the 3D cantilever beam.

where V^* is the allowable material volume, $i, j = 2$ in the 2D case, and $i, j = 3$ in the 3D case. Young's modulus and Poisson's ratio of the base material are 1 and 0.3, respectively.

Table 3

The performance of different methods for compliance-minimization topology optimization problems.

Example	Gradient-based MFSE (Nm)	DNN-MFSE		KG-MFSE [25]	
		Mean of compliance (Nm)	Standard deviation of compliance (Nm)	Mean of compliance (Nm)	Standard deviation of compliance (Nm)
Cantilever beam	0.7472	0.7564	0.0022	0.7565	0.0251
MBB beam	9.8625	10.2221	0.0882	10.3315	0.6423

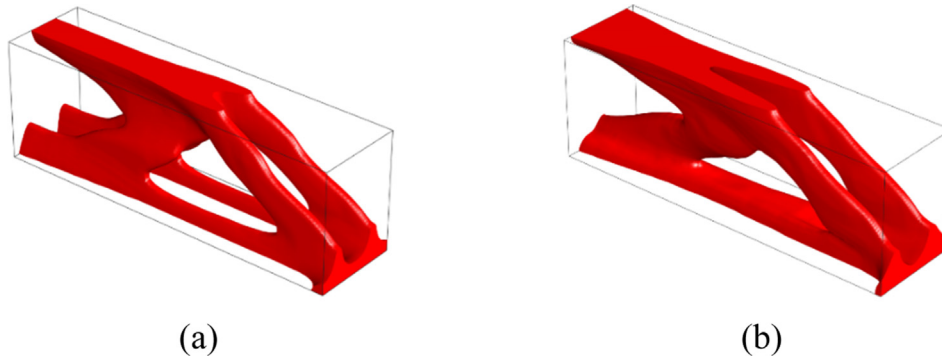


Fig. 17. Topology optimization solutions of 3D cantilever beams with different correlation lengths. (a) $l_c = 13$ mm; (b) $l_c = 9.5$ mm.

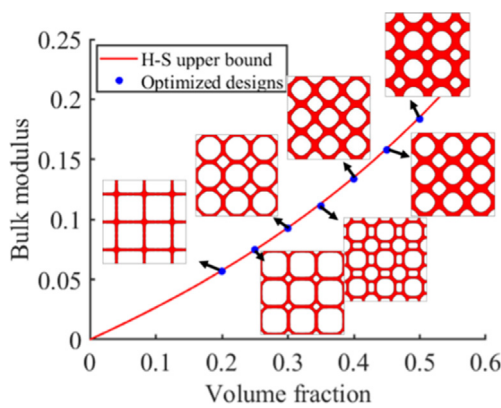


Fig. 18. Comparison of the bulk modulus of the optimized designs with the H-S upper bound.

The 2D maximum bulk modulus material design was conducted first. The unit cell was discretized with 100×100 bi-linear quadrilateral finite elements, the correlation length was set as 15% of the unit cell length, and the number of design variables in this example was 99. Starting from an initial homogeneous base material distribution, the optimized solutions with different volume ratios are shown in Fig. 18, and all volume fraction constraints are satisfied. The required number of finite element calls for all cases was fewer than 2600. We compared the bulk modulus with the H-S upper bound for all cases plotted in Fig. 19. The conclusion was that in all the cases, deviation from the H-S upper bound was less than 2%, thereby indicating that the better performance of the proposed method without the need for experience-based guessing regarding the initial layout.

Second, the proposed DNN-MFSE method was applied for maximizing the bulk modulus of the 3D microstructures. The unit cell was discretized with $40 \times 40 \times 40$ eight-node solid finite elements, and the allowable volume ratios were 30%, 40%, and 50%, respectively. After implementing the proposed method without the initial microstructural geometry and the sensitivity information, optimized results of three cases, including the optimized topologies of the unit cell, the $2 \times 2 \times 2$ unit cells, the corresponding homogenized constitutive matrix, and the bulk modulus, were shown in Fig. 20. When the volume ratio was 30%, the optimization problem converged after successively searching in 17 sub-domains with 2674 finite element calls, and the bulk modulus was 0.93. The optimized bulk modulus value was the same as in the results [52] but showed a different topological configuration. It was also clearly observed that the optimized configurations of the unit cell with

50% material usage, which were obtained by the proposed method, were consistent with the results [52]. Moreover, the results of all the numerical examples were slightly better than the results [52]. Hence, the effectiveness of the proposed method for 3D microstructural material design was validated.

4.4. Minimization of the maximum weld stress in the dropped mobile phone problem

Smartphones are present in thousands of households due to their portability and usability. When using a mobile phone, it is very common for it to accidentally be dropped onto a hard surface, which may cause physical damage to its internal parts or to the external casing, particularly the solder joints, which are important for maintaining functionality. In this example, a dynamics problem is applied to confirm the ability of the proposed method to solve complex nonlinear problems, i.e., minimizing the weld stress of a dropped mobile phone (as shown in Fig. 21), in which gradient-based methods are difficult to address. Lim et al. [53] pointed out that horizontal drop orientations generally produced the largest impact responses. A smartphone free-drop from a pocket is one of the most common situations in which this can occur. Therefore, assuming that the height of the mobile phone free-drop is 1 m, this paper converts this value into an equivalent initial velocity, i.e., the velocity of first contact with the rigid ground surface is 4.427 m/s. In summary, this study assumes that the phone drops horizontally onto a hard surface with an initial velocity of $v_0 = 4.427$ m/s, and the optimization objective is to find the optimal topology of the shielding frame for minimizing the maximum principal stress on the solder joints during the impact process.

The MFSE-based optimization model for this problem can be summarized as follows:

$$\begin{aligned} \min_{\eta} \quad & \max_t S_t(\eta) \\ \text{s.t.} \quad & \begin{cases} \mathbf{M}\mathbf{a} + \mathbf{C}\mathbf{v} + \mathbf{F} = \mathbf{R} \\ \eta^T \mathbf{W}_i \eta \leq 1 \end{cases} \end{aligned} \quad (25)$$

where t represents the time in the collision analysis, S_t is the maximum principal stress on the solder joints at time t , and $\mathbf{M}\mathbf{a} + \mathbf{C}\mathbf{v} + \mathbf{F} = \mathbf{R}$ represents the dynamic finite element equation that the collision obeys. In the finite element formula, $\mathbf{M}$, $\mathbf{C}$, $\mathbf{F}$, and $\mathbf{R}$ are the global mass matrix, the damping matrix, the internal force vector, and the load vector, respectively. Response vector $\mathbf{a}$ and $\mathbf{v}$ are the acceleration vector and the velocity vector, respectively. Note that the collision problem is a typical nonlinear transient-state dynamic problem, and it is difficult to derive the design sensitivity information of the objective function, which is indispensable to gradient-based topology optimization methods.

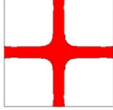
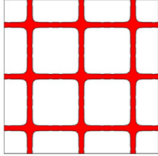
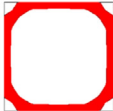
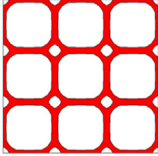
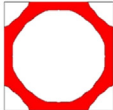
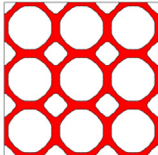
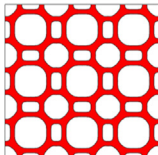
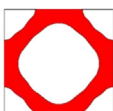
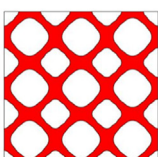
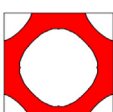
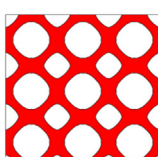
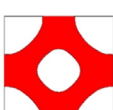
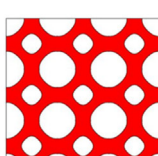
Base cell	3 by 3 base cells	Homogenized constitutive matrix and the bulk modulus
		$\mathbf{D}^H = \begin{bmatrix} 0.1081 & 0.0055 & 0.0000 \\ 0.0055 & 0.1081 & 0.0000 \\ 0.0000 & 0.0000 & 0.0000 \end{bmatrix}$ $K_{eff} = 0.0568$
(a)		
		$\mathbf{D}^H = \begin{bmatrix} 0.1021 & 0.0452 & 0.0000 \\ 0.0452 & 0.1021 & 0.0000 \\ 0.0000 & 0.0000 & 0.0024 \end{bmatrix}$ $K_{eff} = 0.0737$
(b)		
		$\mathbf{D}^H = \begin{bmatrix} 0.1056 & 0.0796 & 0.0000 \\ 0.0796 & 0.1056 & 0.0000 \\ 0.0000 & 0.0000 & 0.0088 \end{bmatrix}$ $K_{eff} = 0.0926$
(c)		
		$\mathbf{D}^H = \begin{bmatrix} 0.1293 & 0.0932 & 0.0000 \\ 0.0932 & 0.1293 & 0.0000 \\ 0.0000 & 0.0000 & 0.0139 \end{bmatrix}$ $K_{eff} = 0.1113$
(d)		
		$\mathbf{D}^H = \begin{bmatrix} 0.1492 & 0.1174 & 0.0000 \\ 0.1174 & 0.1492 & 0.0000 \\ 0.0000 & 0.0000 & 0.0602 \end{bmatrix}$ $K_{eff} = 0.1333$
(e)		
		$\mathbf{D}^H = \begin{bmatrix} 0.1851 & 0.1305 & 0.0000 \\ 0.1305 & 0.1821 & 0.0000 \\ 0.0000 & 0.0000 & 0.0627 \end{bmatrix}$ $K_{eff} = 0.1533$
(f)		
		$\mathbf{D}^H = \begin{bmatrix} 0.2347 & 0.1322 & 0.0000 \\ 0.1322 & 0.2347 & 0.0000 \\ 0.0000 & 0.0000 & 0.0594 \end{bmatrix}$ $K_{eff} = 0.1835$
(g)		

Fig. 19. Optimized 2D microstructures and their effective material parameters with (a) 20%, (b) 25%, (c) 30%, (d) 35%, (e) 40%, (f) 45%, and (g) 50% base material usages.

Base cell	2 by 2 by 2 base cells	Homogenized constitutive matrix and the bulk modulus
		$\mathbf{D}^H = \begin{bmatrix} 0.15 & 0.08 & 0.08 & 0 & 0 & 0 \\ 0.08 & 0.15 & 0.08 & 0 & 0 & 0 \\ 0.08 & 0.08 & 0.15 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.06 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.06 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.06 \end{bmatrix},$ $K_{eff} = 0.93$
		$\mathbf{D}^H = \begin{bmatrix} 0.29 & 0.10 & 0.10 & 0 & 0 & 0 \\ 0.10 & 0.29 & 0.10 & 0 & 0 & 0 \\ 0.10 & 0.10 & 0.29 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.08 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.08 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.08 \end{bmatrix},$ $K_{eff} = 1.47$
		$\mathbf{D}^H = \begin{bmatrix} 0.40 & 0.14 & 0.14 & 0 & 0 & 0 \\ 0.14 & 0.40 & 0.14 & 0 & 0 & 0 \\ 0.14 & 0.14 & 0.40 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.11 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.11 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.11 \end{bmatrix},$ $K_{eff} = 2.04$

Fig. 20. Optimized 3D microstructures and their effective material parameters with (a) 30%, (b) 40%, and (c) 50% base material usages.

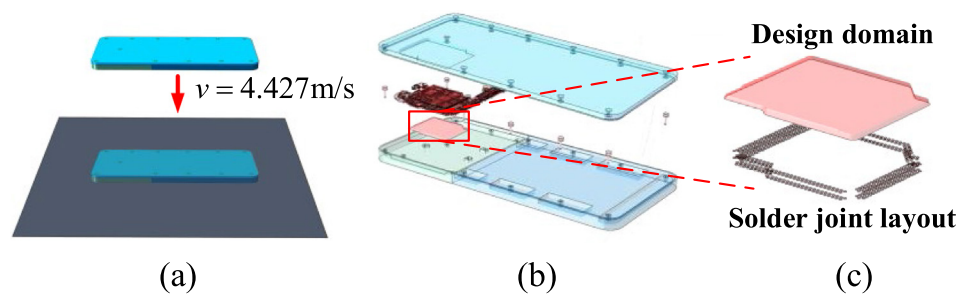


Fig. 21. A mobile phone design problem. (a) A schematic diagram of the dropped mobile phone model; (b) an exploded view of the dropped mobile phone; (c) the shielding frame topology optimization design domain and the layout of the solder joints.

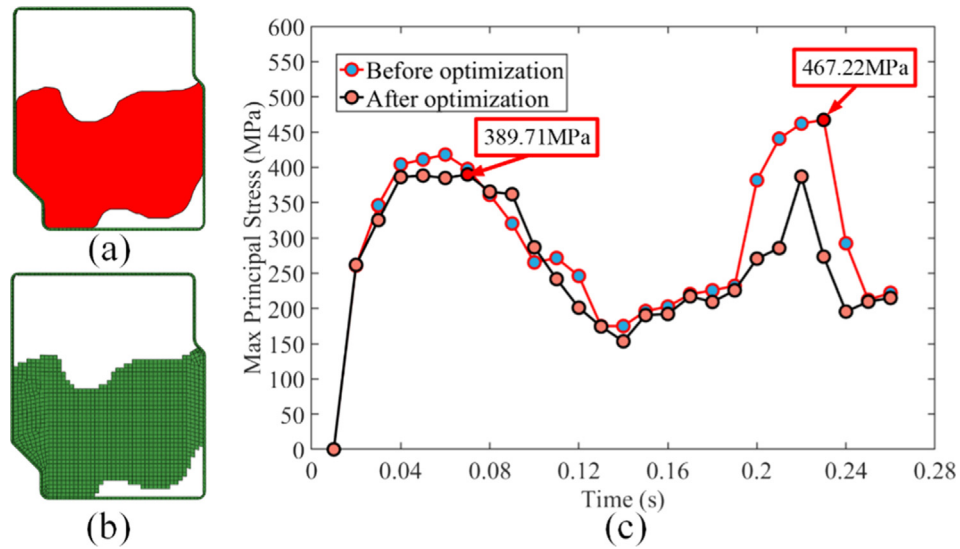


Fig. 22. Optimized results of the mobile phone shielding frame. (a) optimized topology; (b) the finite element model; (c) the maximal principal stress curve of solder joints.

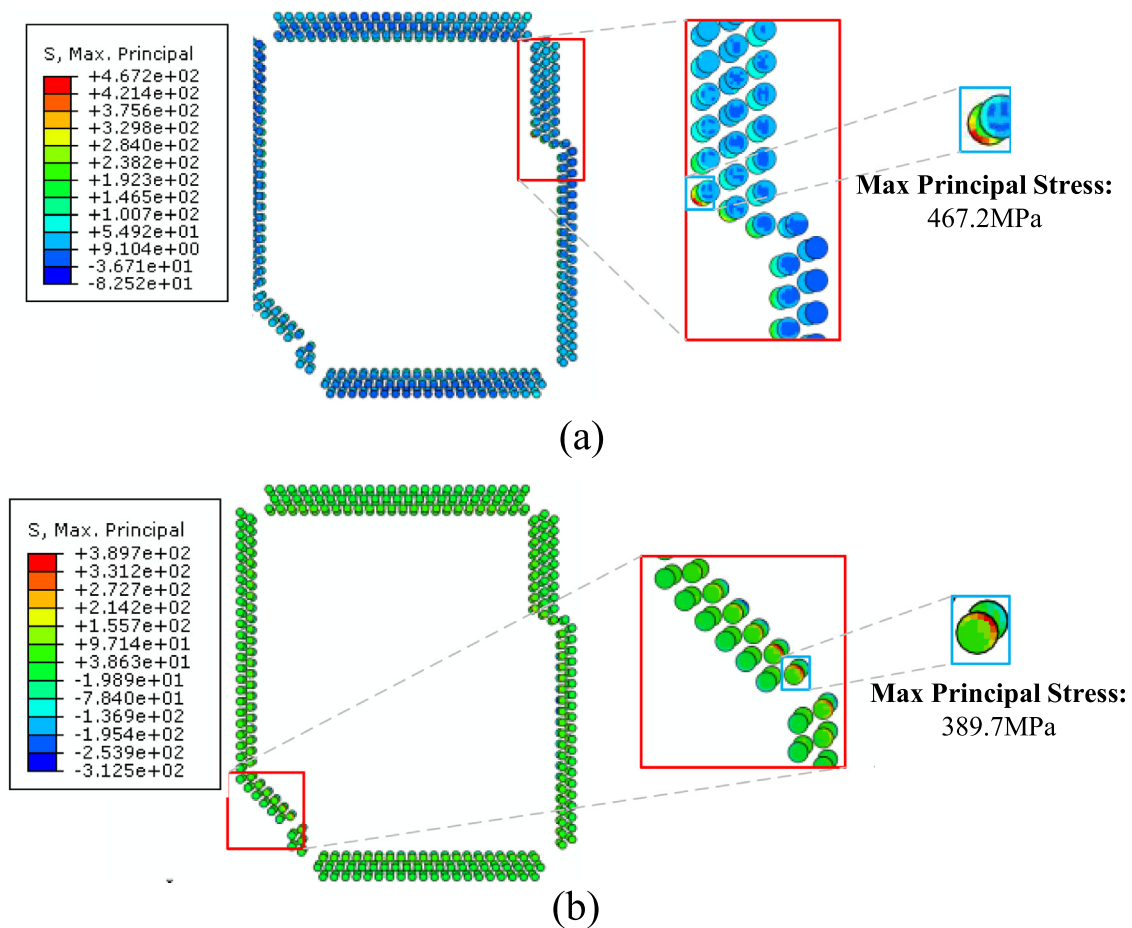


Fig. 23. The maximal principal stress distribution of solder joints. (a) the stress distribution before optimization (at 0.23 s); (b) the stress distribution after optimization (at 0.07 s).

To address non-linear problems such as the impact situation, an explicit dynamics finite element method can deliver better convergence and high accuracy and efficiency. Accordingly, the Abaqus/Explicit finite element analysis is adopted to obtain the time curve of the maximum principal stress in this example, which uses a central difference system to solve the equilibrium equation in Eq. (25), i.e.,

$$\mathbf{v}_t = \frac{1}{2\Delta t}(\mathbf{s}_{t+1} - \mathbf{s}_{t-1}), \mathbf{a}_t = \frac{1}{\Delta t^2}(\mathbf{s}_{t+1} - 2\mathbf{s}_t + \mathbf{s}_{t-1}) \quad (26)$$

where $\mathbf{v}_t$, $\mathbf{s}_t$ and $\mathbf{a}_t$ are velocity, displacement, and acceleration at time t , respectively, and Δt is the time increment. In this example, the time increment is set as $\Delta t = 0.01$ s, and the total analysis time is 0.26 s.

The finite element model for the entire phone model comprised up more than 390,000 elements. The ground type was assumed to be a rigid body that collided with a free-dropped mobile phone. The upper shielding frame of the mobile phone was set as the design domain in the proposed DNN-MFSE optimization procedure. The results of the optimized shielding frame are shown in Fig. 22, where the maximal stress of solder joints during the impact process is 389.72 MPa, indicating a reduction of 16.59% compared to the maximum value (467.22 MPa) of the stress spectrum before optimization. The stress peak of solder joints shifted from one peak to another and was then significantly reduced. The maximum stress distribution before and after the optimization is given in Fig. 23, which shows that the maximum stress appears at 0.23 s before optimization and switched to 0.07 s after optimization. Furthermore, the position where the maximum stress appears is transformed from the lower layer to the upper layer of solder joints. This numerical example demonstrates the powerful ability of the method proposed herein to solve complex nonlinear dynamic topology optimization problems.

5. Conclusions

Topology optimization has broad application prospects in many engineering disciplines, but gradient-based optimization methods require sensitivity information and easily fall into local solutions, thus necessitating the demand for gradient-free topology optimization approaches for complex engineering problems. In this study, a systematic DNN-MFSE method was proposed that indicated better performance than other gradient-free approaches. Using the presented algorithm, 3D topology optimization problems, including compliance-minimization problems, maximum bulk modulus problems, and the dropped phone problem, were successfully solved. The results of the numerical examples confirmed the robustness and effectiveness of the DNN-MFSE method. Since this method is a surrogate-based optimization method, it can easily be implemented in parallel computing, and there is no need to obtain sensitivity information. Moreover, this method presents significant convenience for solving complex structures and can potentially be extended to multi-physics coupling microstructural material design problems. That means this framework active in a promising way in complex nonlinear or multifunctional material design and will further investigation in our future work.

6. Data availability statement

Data will be made available on request.

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

This work was supported financially by the National Natural Science Foundation of China (11972104) and the LiaoNing Revitalization Talents Program (XLYC1807187).

References

- [1] J.D. Deaton, R.V. Grandhi, A survey of structural and multidisciplinary continuum topology optimization: post 2000, *Struct. Multidiscip. Optim.* 49 (1) (2014) 1–38, <https://doi.org/10.1007/s00158-013-0956-z>.
- [2] C. Wang, Z. Zhao, M. Zhou, O. Sigmund, X.S. Zhang, A comprehensive review of educational articles on structural and multidisciplinary optimization, *Struct. Multidiscip. Optim.* 64 (5) (2021) 2827–2880, <https://doi.org/10.1007/s00158-021-03050-7>.
- [3] M.P. Bendsoe, Optimal shape design as a material distribution problem, *Struct. Optim.* 1 (4) (1989) 193–202, <https://doi.org/10.1007/BF01650949>.
- [4] M. Zhou, G.I.N. Rozvany, The COC algorithm, Part II: Topological, geometrical and generalized shape optimization, *Comput. Methods Appl. Mech. Eng.* 89 (1) (1991) 309–336, [https://doi.org/10.1016/0045-7825\(91\)90046-9](https://doi.org/10.1016/0045-7825(91)90046-9).
- [5] Y.M. Xie, G.P. Steven, A simple evolutionary procedure for structural optimization, *Comput. Struct.* 49 (5) (1993) 885–896, [https://doi.org/10.1016/0045-7949\(93\)90035-C](https://doi.org/10.1016/0045-7949(93)90035-C).
- [6] M.Y. Wang, X. Wang, D. Guo, A level set method for structural topology optimization, *Comput. Methods Appl. Mech. Eng.* 192 (1) (2003) 227–246, [https://doi.org/10.1016/S0045-7825\(02\)00559-5](https://doi.org/10.1016/S0045-7825(02)00559-5).
- [7] G. Allaire, F. Jouve, A.-M. Toader, Structural optimization using sensitivity analysis and a level-set method, *J. Comput. Phys.* 194 (1) (2004) 363–393, <https://doi.org/10.1016/j.jcp.2003.09.032>.
- [8] X. Guo, W. Zhang, W. Zhong, Doing topology optimization explicitly and geometrically—a new moving morphable components based framework, *J. Appl. Mech.* 81 (8) (2014), <https://doi.org/10.1115/1.4027609>.
- [9] O. Sigmund, A 99 line topology optimization code written in Matlab, *Struct. Multidiscip. Optim.* 21 (2) (2001) 120–127, <https://doi.org/10.1007/s001580050176>.
- [10] J. Zhang, N.-H. Kim, L. Lasdon, An Improved Successive Linear Programming Algorithm 31(10) (1985) 1312–1331, doi:10.1287/mnsc.31.10.1312.
- [11] K. Svanberg, The method of moving asymptotes—a new method for structural optimization 24(2) (1987) 359–373, doi:10.1002/nme.1620240207.
- [12] Y. Liang, G. Cheng, Topology optimization via sequential integer programming and Canonical relaxation algorithm, *Comput. Methods Appl. Mech. Eng.* 348 (2019) 64–96, <https://doi.org/10.1016/j.cma.2018.10.050>.
- [13] A. Rietz, Sufficiency of a finite exponent in SIMP (power law) methods, *Struct. Multidiscip. Optim.* 21 (2) (2001) 159–163, <https://doi.org/10.1007/s001580050180>.
- [14] M. Stolpe, K. Svanberg, An alternative interpolation scheme for minimum compliance topology optimization, *Struct. Multidiscip. Optim.* 22 (2) (2001) 116–124, <https://doi.org/10.1007/s001580100129>.
- [15] J.S. Jensen, O. Sigmund, Topology optimization of photonic crystal structures: a high-bandwidth low-loss T-junction waveguide, *J. Opt. Soc. Am. B* 22 (6) (2005) 1191–1198, <https://doi.org/10.1364/JOSAB.22.001191>.
- [16] W. Hare, J. Nutini, S. Tesfamariam, A survey of non-gradient optimization methods in structural engineering, *Adv. Eng. Software* 59 (2013) 19–28, <https://doi.org/10.1016/j.advengsoft.2013.03.001>.
- [17] S.Y. Wang, K. Tai, Structural topology design optimization using Genetic Algorithms with a bit-array representation, *Comput. Methods Appl. Mech. Eng.* 194 (36) (2005) 3749–3770, <https://doi.org/10.1016/j.cma.2004.09.003>.
- [18] L. Shen, Z. Ye, S. He, Design of two-dimensional photonic crystals with large absolute band gaps using a genetic algorithm, *Phys. Rev. B* 68 (3) (2003) 035109, <https://doi.org/10.1103/PhysRevB.68.035109>.
- [19] C. Deng, Y. Wang, C. Qin, Y. Fu, W. Lu, Self-directed online machine learning for topology optimization, *Nat. Commun.* 13 (1) (2022) 388, <https://doi.org/10.1038/s41467-021-27713-7>.
- [20] O. Sigmund, On the usefulness of non-gradient approaches in topology optimization, *Struct. Multidiscip. Optim.* 43 (5) (2011) 589–596, <https://doi.org/10.1007/s00158-011-0638-7>.
- [21] D.J. Lohan, E.M. Dede, J.T. Allison, Topology optimization for heat conduction using generative design algorithms, *Struct. Multidiscip. Optim.* 55 (3) (2017) 1063–1077, <https://doi.org/10.1007/s00158-016-1563-6>.
- [22] M. Yoshimura, K. Shimoyama, T. Misaka, S. Obayashi, Topology optimization of fluid problems using genetic algorithm assisted by the Kriging model 109(4) (2017) 514–532, doi:10.1002/nme.5295.
- [23] E. Raponi, M. Bujny, M. Olhofer, N. Aulig, S. Boria, F. Duddeck, Kriging-assisted topology optimization of crash structures, *Comput. Methods Appl. Mech. Eng.* 348 (2019) 730–752, <https://doi.org/10.1016/j.cma.2019.02.002>.

- [24] Y. Luo, J. Bao, A material-field series-expansion method for topology optimization of continuum structures, *Comput. Struct.* 225 (2019) 106122, <https://doi.org/10.1016/j.compstruc.2019.106122>.
- [25] Y. Luo, J. Xing, Z. Kang, Topology optimization using material-field series expansion and Kriging-based algorithm: an effective non-gradient method, *Comput. Methods Appl. Mech. Eng.* 364 (2020) 112966, <https://doi.org/10.1016/j.cma.2020.112966>.
- [26] X. Zhang, J. Xing, P. Liu, Y. Luo, Z. Kang, Realization of full and directional band gap design by non-gradient topology optimization in acoustic metamaterials, *Extreme Mech. Lett.* 42 (2021) 101126, <https://doi.org/10.1016/j.eml.2020.101126>.
- [27] X. Zhang, Y. Luo, Y. Yan, P. Liu, Z. Kang, Photonic Band Gap Material Topological Design at Specified Target Frequency 4(10) (2021) 2100125. doi:10.1002/adts.202100125.
- [28] X. Zhang, Z. Jia, Y. Luo, Y. Wang, P. Liu, Z. Kang, A Precisely-Controlled Multichannel Phononic Crystal Resonant Cavity (*Adv. Theory Simul.* 12(2021) 4(12) (2021) 2170029. doi:10.1002/adts.202170029.
- [29] P. Liu, Y. Yan, X. Zhang, Y. Luo, Z. Kang, Topological design of microstructures using periodic material-field series-expansion and gradient-free optimization algorithm, *Mater. Des.* 199 (2021) 109437, <https://doi.org/10.1016/j.matdes.2020.109437>.
- [30] P. Liu, X. Zhang, Y. Luo, Topological design of freely vibrating bi-material structures to achieve the maximum band gap centering at a specified frequency, *J. Appl. Mech.* 88 (8) (2021), <https://doi.org/10.1115/1.4050621>.
- [31] I. Sosnovik, I. Oseledets, Neural networks for topology optimization %J Russ. J. Numer. Anal. Math. Model. 34(4) (2019) 215–223. doi:10.1515/rnam-2019-0018.
- [32] Q. Lin, J. Hong, Z. Liu, B. Li, J. Wang, Investigation into the topology optimization for conductive heat transfer based on deep learning approach, *Int. Commun. Heat Mass Transfer* 97 (2018) 103–109, <https://doi.org/10.1016/j.icheatmasstransfer.2018.07.001>.
- [33] Y. Yu, T. Hur, J. Jung, I.G. Jang, Deep learning for determining a near-optimal topological design without any iteration, *Struct. Multidiscip. Optim.* 59 (3) (2019) 787–799, <https://doi.org/10.1007/s00158-018-2101-5>.
- [34] B. Li, C. Huang, X. Li, S. Zheng, J. Hong, Non-iterative structural topology optimization using deep learning, *Comput. Aid. Des.* 115 (2019) 172–180, <https://doi.org/10.1016/j.cad.2019.05.038>.
- [35] D.W. Abueidda, S. Koric, N.A. Sobh, Topology optimization of 2D structures with nonlinearities using deep learning, *Comput. Struct.* 237 (2020) 106283, <https://doi.org/10.1016/j.compstruc.2020.106283>.
- [36] D. Wang, C. Xiang, Y. Pan, A. Chen, X. Zhou, Y. Zhang, A deep convolutional neural network for topology optimization with perceptible generalization ability, *Eng. Optim.* (2021) 1–16, <https://doi.org/10.1080/0305215X.2021.1902998>.
- [37] A. Halle, L.F. Campanile, A. Hasse, An Artificial Intelligence-Assisted Design Method for Topology Optimization without Pre-Optimized Training Data 11 (19) (2021) 9041.
- [38] C.C. Li, A.D. Kiureghian, Optimal Discretization of Random Fields, 119(6) (1993) 1136–1154 10.1061/(ASCE)0733-9399(1993)119:6(1136).
- [39] A. Bhosekar, M. Ierapetritou, Advances in surrogate based modeling, feasibility analysis, and optimization: a review, *Comput. Chem. Eng.* 108 (2018) 250–267, <https://doi.org/10.1016/j.compchemeng.2017.09.017>.
- [40] C. Qian, W. Ye, Accelerating gradient-based topology optimization design with dual-model artificial neural networks, *Struct. Multidiscip. Optim.* 63 (4) (2021) 1687–1707, <https://doi.org/10.1007/s00158-020-02770-6>.
- [41] Y. Su, L. Xu, E.D. Goodman, Hybrid surrogate-based constrained optimization with a new constraint-handling method, *IEEE Trans. Cybern.* (2020) 1–14, <https://doi.org/10.1109/TCYB.2020.3031620>.
- [42] Y. Seong, C. Park, J. Choi, I. Jang, Surrogate Model with a Deep Neural Network to Evaluate Gas–Liquid Flow in a Horizontal Pipe 13(4) (2020) 968.
- [43] S. Barmada, N. Fontana, A. Formisano, D. Thomopoulos, M. Tucci, A deep learning surrogate model for topology optimization, *IEEE Trans. Magn.* 57 (6) (2021) 1–4, <https://doi.org/10.1109/TMAG.2021.3063470>.
- [44] L. Zheng, S. Kumar, D.M. Kochmann, Data-driven topology optimization of spinodoid metamaterials with seamlessly tunable anisotropy, *Comput. Methods Appl. Mech. Eng.* 383 (2021) 113894, <https://doi.org/10.1016/j.cma.2021.113894>.
- [45] N. Calik, M.A. Belen, P. Mahouti, S. Koziel, Accurate modeling of frequency selective surfaces using fully-connected regression model with automated architecture determination and parameter selection based on Bayesian optimization, *IEEE Access* 9 (2021) 38396–38410, <https://doi.org/10.1109/ACCESS.2021.3063523>.
- [46] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, *Neu. Netw.* 2 (5) (1989) 359–366, [https://doi.org/10.1016/0893-6080\(89\)90020-8](https://doi.org/10.1016/0893-6080(89)90020-8).
- [47] A. Chandrasekhar, K. Suresh, TOuNN: Topology Optimization using Neural Networks, *Struct. Multidiscip. Optim.* 63 (3) (2021) 1135–1149, <https://doi.org/10.1007/s00158-020-02748-4>.
- [48] M.O. Elingaard, N. Aage, J.A. Bærentzen, O. Sigmund, De-homogenization using convolutional neural networks, *Comput. Methods Appl. Mech. Eng.* 388 (2022) 114197, <https://doi.org/10.1016/j.cma.2021.114197>.
- [49] F. Liu, X. Jiang, X. Wang, L. Wang, Machine learning-based design and optimization of curved beams for multistable structures and metamaterials, *Extr. Mech. Lett.* 41 (2020) 101002, <https://doi.org/10.1016/j.eml.2020.101002>.
- [50] Exploring and Exploiting a Surrogate, *Engineering Design via Surrogate Modelling* (2008) 77–107.
- [51] Y. Zhang, M. Xiao, H. Li, L. Gao, Topology optimization of material microstructures using energy-based homogenization method under specified initial material layout, *J. Mech. Sci. Technol.* 33 (2) (2019) 677–693, <https://doi.org/10.1007/s12206-019-0123-6>.
- [52] X. Huang, A. Radman, Y.M. Xie, Topological design of microstructures of cellular materials for maximum bulk or shear modulus, *Comput. Mater. Sci.* 50 (6) (2011) 1861–1870, <https://doi.org/10.1016/j.commatsci.2011.01.030>.
- [53] C.T. Lim, C.W. Ang, L.B. Tan, S.K.W. Seah, E.H. Wong, Drop impact survey of portable electronic products, in: *53rd Electronic Components and Technology Conference Proceedings*, 2003, pp. 113–120.