Advanced Engineering Informatics 67 (2025) 103524

Contents lists available at ScienceDirect

Advanced Engineering Informatics

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

Full length article

# Integrating large models with topology optimization for conceptual design realization

Zelong Liang [a,b], Yuan-Fang Zhang [c], Yingjun Wang [a,b],*, Weihua Li [a],**

[a] School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou, 510641, Guangdong, China
[b] National Engineering Research Center of Novel Equipment for Polymer Processing, The Key Laboratory of Polymer Processing Engineering of the Ministry of Education, Guangdong Provincial Key Laboratory of Technique and Equipment for Macromolecular Advanced Manufacturing, Guangzhou, 510641, Guangdong, China
[c] Shien-Ming Wu School of Intelligent Engineering, South China University of Technology, Guangzhou, 511442, Guangdong, China

ARTICLE INFO

*Keywords:*
Topology optimization
Generative structural design
Deep learning
Conceptual design

ABSTRACT

Topology optimization (TO) has been extensively applied in various domains, including robotics, equipment manufacturing, household products, and civil engineering, enhancing the performance of structural designs. However, TO-designed structures often lack adaptability to human preference despite high physical performance. Trained extensively on human knowledge, large visual–language models (LVLM) exhibit a strong ability to understand human intent and generate satisfactory designs efficiently. In this paper, a large visual–language model-guided topology optimization (LMTO) approach is proposed to automatically generate and edit efficient structural designs according to concepts. By integrating the TO into the large model knowledge space through a UDF-Weighting block, LMTO can optimize performance in the direction of human preference. Experimental results show that, despite significant variations in appearance, the performance of the designs remains comparable or superior to those obtained by the BESO method, indicating the effectiveness of our approach in exploring the joint space. Our method can yield diverse designs from the same prompt and is well-adapted to 2D and 3D cases, highlighting its effectiveness and practicality.

## 1. Introduction

How to build a chair shaped like a penguin or an avocado? What about a Gothic-style bridge? Great designs typically begin with a simple concept and evolve into tangible objects. Computer-aided design (CAD) software packages, such as SolidWorks, Catia, and Creo, are traditionally employed to address geometric shape modeling. In addition, structural performance analysis is handled by computer-aided engineering (CAE) software packages, namely ANSYS, Abaqus, and Nastran. However, translating abstract concepts into feasible designs remains a significant challenge for artists and engineers [1]. This challenge is multifaceted: First of all, designers require extensive time to accumulate experience and inspiration to produce creative designs. In the second place, due to the different mathematical representations in CAD and CAE software packages, designers must coordinate design intent with physical performance across platforms. A data-driven design approach, aligning physical performance with human preferences, should be proposed to overcome these difficulties.

Topology optimization (TO) is a numerical method used to determine the optimal material layout within a given design domain [2], extensively proven to create lightweight, high-performance structures in various fields, such as aircraft engineering [3,4], architectural design [5,6], and furniture manufacturing [7,8]. There are quite a few TO techniques, the most prevalent of which are the Solid Isotropic Material with Penalization (SIMP) [9,10], the level set (LSM) [11,12], and the Evolutionary Structural Optimization (ESO) [13,14] method. Initially, ESO allowed only material removal, requiring an oversized initial design domain and often leading to local optima. To address this issue, Quein et al. introduced a bi-directional ESO (BESO) approach, enabling both material removal and addition [14]. Huang et al. developed a soft element removal method to enhance convergence in stiffness optimization problems and further discussed various BESO applications [15]. Early studies mainly focused on optimizing physical performance, while, in recent years, complex design requirements and comprehensive knowledge have been introduced into TO designs. For example, specific reference textures or geometrical patterns have been employed as guidance to modify TO design styles [16–18]. Moreover, additional constraints related to the design domain [19,20], material density distribution [21], and sensitivity [22,23] have been

* Corresponding author at: School of Mechanical and Automotive Engineering, South China University of Technology, Guangzhou, 510641, Guangdong, China.
** Corresponding author.
*E-mail addresses:* wangyj84@scut.edu.cn (Y. Wang), whlee@scut.edu.cn (W. Li).

https://doi.org/10.1016/j.aei.2025.103524
Received 18 February 2025; Received in revised form 24 May 2025; Accepted 28 May 2025

incorporated into BESO within a generative structural design, thereby expanding the design space available to designers.

The rapid advancement of artificial intelligence has profoundly transformed the engineering fields [24,25]. In particular, the introduction of generative models [26–28] and reinforcement learning [29] into TO has enabled the direct fabrication of structures generated through generative and conceptual designs [30,31]. The diffusion model is one of the most promising generative models, and Topodiff [32] explores the beam design space with a conditional diffusion model. However, the application of these advanced models is often constrained by the limited scope of available datasets, typically restricting their use to specific categories, such as wheel and MBB beam designs. Deep learning features are also utilized to modulate the outcomes of TO designs. For instance, Vulimiri and colleagues have attempted to merge a reference image with an optimized structure using a pre-trained neural network, to enhance the design's aesthetic [33]. Similarly, Zhang et al. developed a machine learning-assisted TO method (MLATO) tailored to architectural designs incorporating artistic elements [34]. Impressive results of image-influenced structures have also been proposed [17,18,34]. However, due to the limited information within a single image, realizing substantial designs that fulfill the extensive demands of designers is still challenging.

In recent years, large models such as ChatGPT [35] have successfully understood human intents and generated high-quality answers, profoundly impacting human–machine interaction. Large visual–language models such as OpenAI's DALL-E 2 [36], Google Brain's Imagen [37], and StabilityAI's Stable Diffusion [38] began to approach the quality of real photographs and introduce novel methods for human–computer interaction via text guidance. These methodologies are also being extended to 3D datasets, enabling users to create 3D representations of tables and chairs in mesh and point cloud formats with simple prompts [39–41]. These developments highlight the potential of AI-driven methodologies in 3D modeling and structural design.

Leveraging the implicit knowledge embedded in large-scale models to guide structural optimization has emerged as a feasible strategy for enhancing the diversity of structural designs. Picard et al. employed GPTv4 to scrutinize TO designs rendered through CAD, underscoring the significant constraint posed by the absence of explicit accuracy in such endeavors [42]. Zhong et al. leveraged the text encoder CLIP to generate aesthetically pleasing beam structures, elucidating the challenges of attaining specific volume fractions within these designs [43]. Bastek and Jan-Hendrik utilized video diffusion models integrated with finite element simulations (FEM) for inverse structural design, generating periodic structural materials with nonlinear deformation and stress responses under compression in the large-strain regime [44]. Although these models leverage knowledge extracted from large models, they cannot achieve local minima in performance owing to constraints imposed by the statistically-based inference mechanisms. Wei Zhang et al. proposed a multi-stage optimization method that integrates the TO with latent diffusion models, enhancing the performance and computational efficiency of the TO [45]. As several TO steps are employed in the second stage to enhance performance, the final design escapes from local minima in performance.

In this study, we introduce LMTO, a novel hybrid model that integrates conventional TO with LVLM, aiming to automatically generate efficient structural designs that align with human preferences. The knowledge priors embedded in the large model enrich the diversity of final designs. LMTO facilitates the manipulation of style and shape in final designs, making it an innovative tool for conceptual design tasks. Within LMTO, TO procedures are guided by an LVLM incorporating Unsigned Distance Field Weighting (UDF-Weighting), enabling the design to evolve from preference-optimal toward performance-optimal configurations. (Fig. 1) Our approach is suitable for generative and editing tasks across 2D and 3D scenarios. By bridging LVLM and TO domains, our methodology enables inexpert designers to produce aesthetically pleasing, high-performance structural designs, ranging from coarse conceptualizations to refined iterations.

The remainder of this paper is organized as follows. Sections 2 and 3 present the theoretical foundations of the Soft-Kill BESO method and the Latent Diffusion Model, respectively. Next, Section 4 describes the implementation of the proposed LMTO framework. Specifically, we first introduce the overall workflow of the LMTO methodology, followed by individual explanations of the LVLM block and the UDF-Weighting module. Afterwards, Section 5 provides experimental validation for the LMTO approach. We begin with a Gothic bridge case study to illustrate the general effectiveness of our method in achieving both design diversity and high performance. To further assess its scalability and adaptability, two distinct design tasks are investigated: a star tracker's bracket, which emphasizes structural performance, and a chair design, which prioritizes aesthetic diversity and user preference. Additionally, compression experiments are conducted to evaluate the manufacturability of the resulting designs. Then, Section 6 shows the limitations. Finally, Section 7 concludes the paper and outlines potential directions for future research.

## 2. Soft-Kill BESO method

Soft-Kill BESO is an advanced TO method designed to improve material distribution in structural design by iteratively removing and adding material based on sensitivity analysis. As opposed to, traditional hard-kill approaches that abruptly remove material elements, the Soft-Kill strategy gradually adjusts material properties, ensuring a more stable and efficient convergence process. This method is particularly effective in handling complex optimization problems where a smooth transition in material distribution is crucial for achieving high-performance structures. Regarding Soft-Kill BESO [15,22], the formulation of optimization problem considering maximum stiffness subject to a volume constraint is expressed as follows:

$$\begin{aligned} \min: \quad & C = \frac{1}{2}\boldsymbol{U}^T\boldsymbol{K}\boldsymbol{U} \\ \text{s.t.}: \quad & V^* = \sum_{i=1}^{N} v_i \rho_i \\ & \rho_i = \rho_{\min} \text{ or } 1 \end{aligned}$$

where $C$, $\boldsymbol{K}$, $\boldsymbol{U}$, $V^*$, $v_i$ and $V$ denote the total compliance of the design (serving as an inverse measure of structural stiffness), the global stiffness matrix, the displacement vector, the volume of the $i$th element, and the total volume of design space, respectively. A smaller volume ratio $V^*/V$ implies a more efficient use of material in achieving the desired structural functionality. The design domain is discretized into $N$ elements, each of which has a design variable $\rho_i$ that specifies whether the element is solid ($\rho_i = 1$) or void ($\rho_i = \rho_{min} = 0.001$).

To ensure that each element closely approximates a solid or void state, the material model for structural elements is defined as:

$$E(\rho_i)^k = \rho_i^p E_0$$

where $E(\rho_i)^k$, $E_0$ and $p$ donate Young's modulus of $i$th structural element at the $k$th iteration, the basic Young's modulus of structural elements, and the penalty exponent which is considered $p = 3$ in the present work, respectively. Additionally, the TO progresses based on the relative ranking of the sensitivity numbers $dc_i$ which is typically defined as follows:

$$dc_i = -\frac{1}{p}\frac{\partial C}{\partial x_i} = \begin{cases} \frac{1}{2}\boldsymbol{u}_i^T\boldsymbol{k}_i\boldsymbol{u}_i, & \text{when } \rho_i = 1 \\ \frac{\rho_{\min}^{p-1}}{2}\boldsymbol{u}_i^T\boldsymbol{k}_i\boldsymbol{u}_i, & \text{when } \rho_i = \rho_{\min} \end{cases} \tag{1}$$

where $\boldsymbol{u}_i$ and $\boldsymbol{k}_i$ denote the elemental displacement vector and the elemental stiffness matrix, respectively.

To mitigate checkerboarding and mesh dependency issues, a convolution filter is applied to the sensitivity numbers $dc_i$ as follows:

$$\tilde{dc_i} = \frac{\sum_{j=1}^{N}(r_{min} - r_{ij})dc_i}{\sum_{j=1}^{N}(r_{min} - r_{ij})} \tag{2}$$

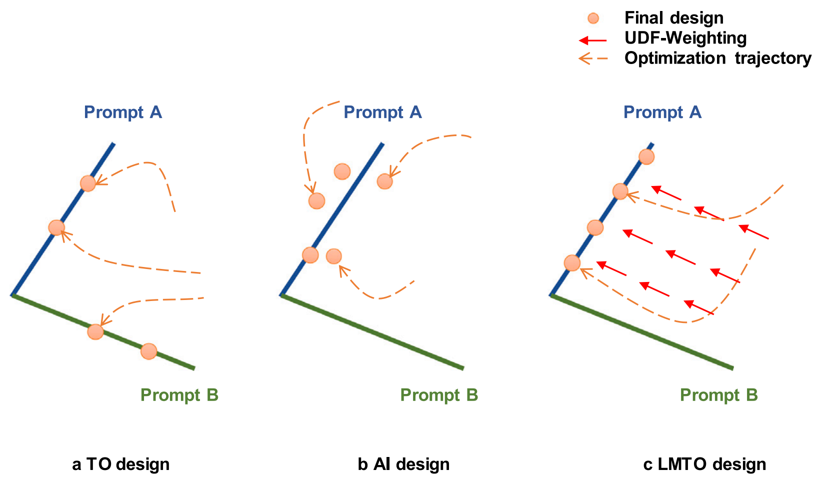


**Fig. 1.** Optimization trajectory of TO design, AI design, and LMTO design. TO is a physical performance optimization process, and the final designs are usually located at the boundary of the design space. AI design is a resampling process, and the final designs are located near the boundary. LMTO can generate designs with preference and performance by the guidance of UDF-Weighting.

where $\tilde{dc}_i$, $r_{min}$ and $r_{ij}$ denote the filtered sensitivity number, the filter radius, and the distance between the centers of elements $i$ and $j$.

To enhance the convergence of the BESO technique, a moving average of historical sensitivity numbers is employed as follows:

$$\hat{dc}_i = \frac{\tilde{dc}_i^{k} + \tilde{dc}_i^{k+1}}{2}$$

where $\hat{dc}_i$ is the averaged sensitivity number. Structural elements are dynamically added or removed based on the sensitivity threshold in each iteration to align the current volume $V^k$ closely with the target volume $V^*$. Then, the next iteration target volume $V^{k+1}$ needs to be updated before elements are removed from or added to the current design:

$$V^{k+1} = V^k(1 \pm ER)$$

where $V^k$ and $ER$ denote the current volume and the evolutionary volume ratio, respectively. Finally, the bisection method is used to add or remove elements based on $\hat{dc}_i$, and the iteration loop stops once the following convergence criterion is satisfied.

$$\frac{\left| \sum_{m=1}^{M} (C_{k-m+1} - C_{k-m-m-1}) \right|}{\sum_{m=1}^{M} C_{k-m+1}} \le \tau$$

The sliding window length is $M = 5$, and the allowable convergence error is $\tau = 0.001$.

## 3. Latent diffusion model

The Latent Diffusion Model (LDM) [38] is a generative model that synthesizes images by progressively denoising representations in a lower-dimensional latent space. In contrast to conventional diffusion models that perform denoising directly in the high-dimensional pixel space, LDMs leverage a lower-dimensional latent space, significantly reducing computational complexity while maintaining high-quality outputs.

To learn a probabilistic generative process in a lower-dimensional latent space, LDM first trains an encoder and a decoder, which transform the high-dimensional data into a compact latent representation. Subsequently, a diffusion model [46] is trained to approximate the data distribution within this latent space.

A diffusion model consists of both forward and backward diffusion processes. The forward diffusion process is a Markov process that gradually adds noise to the data in a series of $T$ steps. Starting with an initial data sample $x_0$, the process progressively transforms data into pure noise. Using a series of conditional diffusion steps, the forward diffusion process can be defined as follows:

$$x_t = \sqrt{1-\beta_t}x_{t-1} + \beta_t \epsilon_t$$

where $\beta_t$ and $\epsilon_t$ denote the variance schedule (controlling the amount of noise added at each time step) and a Gaussian distribution with $\mathcal{N}(0,1)$.

The backward process aims to reverse the noise corruption process, in which a neural network learns noise $\epsilon_\theta$. It starts from pure noise $x_T \in \mathcal{N}(0,1)$ and recovers the data $x_0$ through a series of $T$ steps. The backward process is defined as follows:

$$x_{t-1} = \frac{1}{\sqrt{\alpha_t}}(x_t - \frac{1-\alpha_t}{\sqrt{1-\bar{\alpha}_t}}\epsilon_\theta(x_t, t)) + \sigma_t \epsilon$$
$$\alpha_t = 1 - \beta_t$$
$$\bar{\alpha}_t = \prod_{i=1}^{t} \alpha_i$$

where $\epsilon$ and $\sigma_t$ denote a normal distribution (the mean and variance of which, in turn, equal 0 and 1) and standard deviation of the noise (added at time step $t$), respectively.

LDM [38] is a large-scale visual–language model trained on LAION-5B, demonstrating a remarkable ability to comprehend human design intentions, particularly architectural designs and furniture products.

## 4. Proposed method

### 4.1. Overview of LMTO

Ensuring manufacturability while preserving human preferences in structural design is a significant challenge arising from two primary factors. First, human preferences often conflict with high-performance designs, and accommodating these preferences may come at the expense of substantial performance. Second, optimization algorithms typically struggle to distinguish between preference- and performance-related features, which can result in the unintended removal of critical preference features. Consequently, the final design may deviate from the intended-design objectives, undermining its alignment with human expectations. Previous studies focus on aligning the preference prompt with structural features and embedding them into the same latent space [43]. Their methods impose a high requirement on the quality of the latent space. It is important to note that once the boundary conditions are defined, most manufacturable structural designs incline to be closely aligned with the optimized TO design. Accordingly, the final design can be decomposed into a combination of the designs

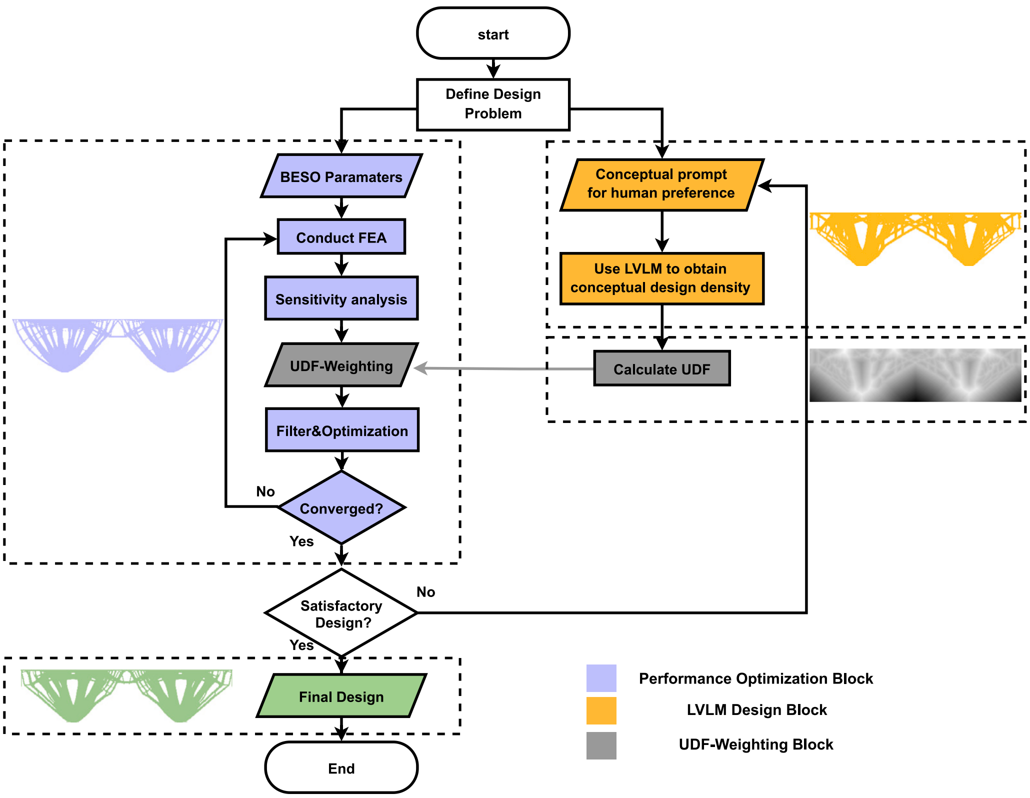


**Fig. 2.** Overview of LMTO framework. The LVLM Design Block is used to generate a human-preference design according to the conceptual prompt. The Performance Optimization Block is used to optimize the physical performance of the whole design. A satisfactory design can be generated by balancing human preference and performance with UDF-Weighting.

based on the TO and the LVLM. This relationship can be formulated as follows:

$$\boldsymbol{x} = \bar{\boldsymbol{x}} + \Sigma_{i=1}^{N}(\alpha_i(\tilde{\boldsymbol{x}}(C_i) - \bar{\boldsymbol{x}})) \tag{3}$$

Here, $\bar{\boldsymbol{x}}$ and $\tilde{\boldsymbol{x}}(C_i)$ represent the optimized TO design and LVLM design under human preference condition $C_i$, respectively. By integrating LVLM design feature sets into TO designs, LMTO achieves high-performance design while ensuring alignment with the design-intended objectives.

As illustrated in Fig. 2, the LMTO consists of three main modules. In the LVLM design block, a coarse conceptual structure design is generated according to users' prompts. Given that these LVLMs encompass extensive human knowledge, users can utilize the rich design space to support structural design and improvement. It is worth noting that the prompt can be an entity or even a concept. For instance, Figs. 9 and 5 show the entity prompt 'penguin chair' and the concept prompt 'Gothic bridge', respectively. In this manner, LMTO can control shape and style using a unified approach.

After the LVLM design block, the coarse conceptual structural design is transformed into an unsigned distance field (UDF) and will subsequently participate in the UDF-Weighting operation. By introducing UDF-Weighting to the TO process, the satisfactory design is as varied as the LVLM designs and as efficient as conventional TO designs. For detailed formulations on performance optimization, the reader is referred to SP-BESO [22]. It should be noted that the LVLM block and the UDF-Weighting block are executed only once a process, their computational cost for inference is negligible. In addition, the complexity of the Performance Optimization block is approximately equivalent to that of the BESO algorithm. Therefore, the overall computational complexity of the proposed method is slightly higher than that of BESO.

By leveraging prompts and UDF-Weighting, satisfactory designs can be generated within the confluence of the expansive design space of the large visual–language model and the physical high-performance TO space. In this context, although distinct prompts contribute to increased diversity and aesthetics of designs, UDF-Weighting governs the performance characteristics.

### *4.2. LVLM block for optimized structure augmentation*

The key idea of the LVLM block is to generate human preference features that guide the subsequent performance optimization block. In this regard, output images produced by image generative models can be treated as a combination of high-level features. To this end, it is initially assumed that TO inherently lacks human preferences. Therefore, the preference information related to the prompt can be represented as $\tilde{\boldsymbol{x}}(C_i) - \bar{\boldsymbol{x}}$, where $\tilde{\boldsymbol{x}}(C_i)$ and $\bar{\boldsymbol{x}}$ denote LVLM raw design under condition in $C_i$ and the TO design. This assumption is plausible because the results produced by TO are completely physical-principle-driven and free from prior knowledge or human-imposed biases.

The TO design is augmented using the general knowledge in LVLM. Indeed, to ensure that the general model's semantic understanding capability in LVLM is not compromised, ControlNet [47] is employed here to generate preference features related to the TO design (see the LVLM architecture in Fig. 3). It should be noted that the fine-tuned dataset contains 2000 BESO design results with different boundary conditions, and AdamW is employed as the optimizer. The feature conversion process can be formulated as follows:

$$\boldsymbol{x}_{n+1} = F(\boldsymbol{x}_n; \boldsymbol{\theta}) + Z(F(\boldsymbol{x}_n + Z(\boldsymbol{\rho}; \boldsymbol{\theta}_{z1}); \boldsymbol{\theta}_\rho); \boldsymbol{\theta}_{z2}) \tag{4}$$

where, the parameter $F$ represents a neural network block that transforms input feature maps from $\boldsymbol{x}_n$ layer to $\boldsymbol{x}_{n+1}$ layer. The parameters $\boldsymbol{\theta}$ are pre-trained and remain fixed during training, whereas the parameters $\boldsymbol{\theta}_\rho$ are not frozen and continue to be updated. The parameter $Z$ refers to $1 \times 1$ convolution layer called "zero convolution", which is employed to fine-tune the large visual–language model. The parameter $\boldsymbol{\rho}$ is the density distribution from TO. By introducing the density

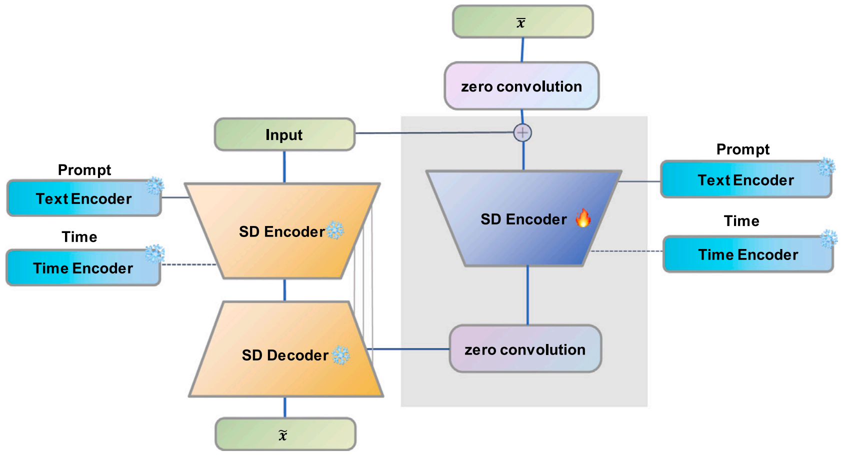


**Fig. 3.** Architecture of LVLM block framework. A pre-trained LVLM is on the left, and a fine-tuned ControlNet is employed to map the prompt to the TO design space according to the general knowledge in the LVLM.

distribution to a large visual–language model, our approach can explore designs within the coupled space of high-performance and general designs.

After the coupled space is constructed, the prompt is mapped to the latent space through CLIP. The coarse conceptualization design, generated through the diffusion process, is expressed as follows:

$$\boldsymbol{x}_{t-1} = \sqrt{\alpha_{t-1}}\left(\frac{\boldsymbol{x}_t - \sqrt{1-\alpha_t}\epsilon_\theta^t(\boldsymbol{x}_t, \boldsymbol{p})}{\sqrt{\alpha_t}}\right) + \sqrt{1-\alpha_{t-1}}\epsilon_\theta^t(\boldsymbol{x}_t, \boldsymbol{p}), \quad t = T, \dots, 1 \tag{5}$$

$\boldsymbol{x}_t$ and $\boldsymbol{x}_{t+1}$ are samples in different diffusion time steps. $\alpha_t$ and $\alpha_{t-1}$ control the noise scaling during the diffusion process. Also, $\epsilon_\theta^t$ and $\boldsymbol{p}$ denote the noise predicted by a neural network at the time step $t$ and the embedding features based on the prompt, respectively. Fig. 5b shows the density distribution of coarse conceptual designs in which the density distribution of materials changes as $\boldsymbol{x}_0$ varies with the random seed, $R$. The designs demonstrate the potential of the LVLM to generate various designs, each of which exhibits aesthetic characteristics that align with the semantics of the input prompt.

### 4.3. UDF-weighting block for high-performance structure generation

The UDF-Weighting block generates a weighted compliance sensitivity $\tilde{dc}$ based on Eq. (1), according to the coarse conceptual design produced by the LVLM block. Then, the performance optimization block utilizes an iterative mechanism inside BESO to identify and select features with the least performance degradation from the feature set generated by the LVLM.

Unsigned Distance Field (UDF) is widely used in computer vision for shape registration and model fitting [48]. In this paper, UDF measures the minimum distance from each point in the design space to the boundary of the coarse conceptualization design $\tilde{x}$, as depicted in Fig. 4. Assume the position of point $i$ in the design area is $\boldsymbol{f}_i$ and the position of point $j$ on the boundary of coarse conceptualization design is $\boldsymbol{g}_j$. The minimum distance to the boundary of the conceptualization design for each point $i$ can be represented as follows:

$$mindis_i = \min_{j\in\mathbb{R}} \|\boldsymbol{f}_i - \boldsymbol{g}_j\| \tag{6}$$

The UDF value for point $i$ equals to the following statement:

$$UDF_i = 1/(mindis_i + 1) \tag{7}$$

Next, we introduce $alpha$ — the scalar factor used in UDF-Weighting — which serves to characterize the similarity between LMTO and TO designs. UDF-Weighting is formulated as follows:

$$w_{UDF} = \max(10^{-alpha} - 10^{-alpha_{min}}, 0) \tag{8}$$

$$\check{dc}_i = (1 - w_{UDF}) \times dc_i + w_{UDF} \times UDF_i \tag{9}$$

To ensure a low compliance of the satisfactory design, the effect of UDF should be small. As a result, $w_{UDF}$ is defined to rescale the UDF value. It should be noted that $\check{dc}_i$ is the weighted $dc_i$ containing balanced information between the TO and the LVLM designs. Additionally, $alpha$ lies between 0 and $alpha_{max}$, $[0, alpha_{max}]$, and it is sensible that the bigger the alpha, the more similar the satisfactory design to the TO design. With this in mind, we consider 9 for $alpha_{max}$ in this study. Hence, considering 9 for $alpha$, the satisfactory design degenerates into a Soft-Kill BESO design. Due to the utilization of UDF-Weighting, a connection has been established between the stiffness of the TO design and the aesthetic features of LVLM design. This connection enables the development of a satisfactory structural design through a balance between stiffness and aesthetics. Then, Eq. (2) is applied to $\check{dc}$ to smooth the optimization process and improve the quality of the final designs. This step corresponds to the "Filter&Optimization" module shown in Fig. 2.

## 5. Results and discussion

### 5.1. 2D Gothic bridge generation

LVLM plays a crucial role in ensuring the diversity of the final designs in LMTO processes. It enables the TO design space to move beyond the constraints of domain-specific datasets and transition toward a generalized representation of design patterns. In this study, "the Gothic style bridge" is studied as an example of abstract conceptual structural design. To demonstrate the adaptability of LMTO to different boundary conditions, two types of boundary conditions, namely boundary condition 1 (BC1) and boundary condition 2 (BC2), are employed, as illustrated in Fig. 5a and b, respectively. The design region is discretized using 400 × 100 linear quadrilateral finite elements. The design volume is $V^* = 0.5V$, (where $V$ denotes the volume of the design domain), and the filter radius is 1.5 times as long as the element size. The experiments are implemented on a Dell Workstation (Processor: Intel(R) Xeon(R) Gold 5218 CPU @ 2.30 GHz, GPU: NVIDIA GeForce RTX 3090). Fig. 5c and d show the diverse bridge designs generated through LVLM based on BC1 and BC2, respectively. Here, LVLM refers to a pre-trained ControlNet [47], where the grayscale intensity of the generated images is analogous to the density distribution in structural design. Because ControlNet is trained on a broad dataset, it circumvents the issue of excessive dependence on the TO datasets observed in deep learning approaches. It is worth noting that although these designs are creative and conceptually similar to the TO designs, they exhibit comparatively lower physical performance.

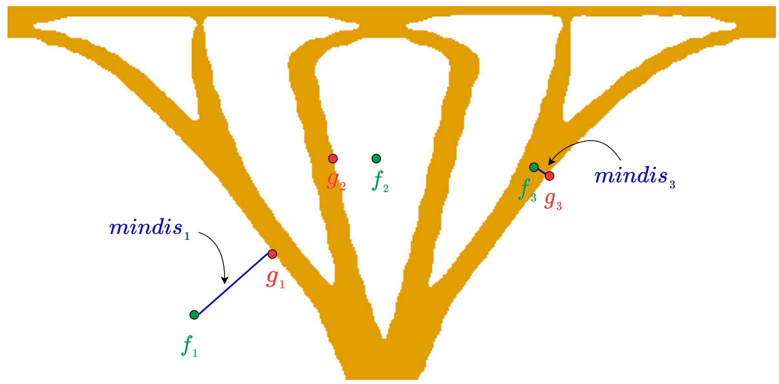


**Fig. 4.** UDF sketch map. Typical positions of $f_i$ and $g_j$ in coarse conceptualization design. The blue line is the distance of $mindis_i$.

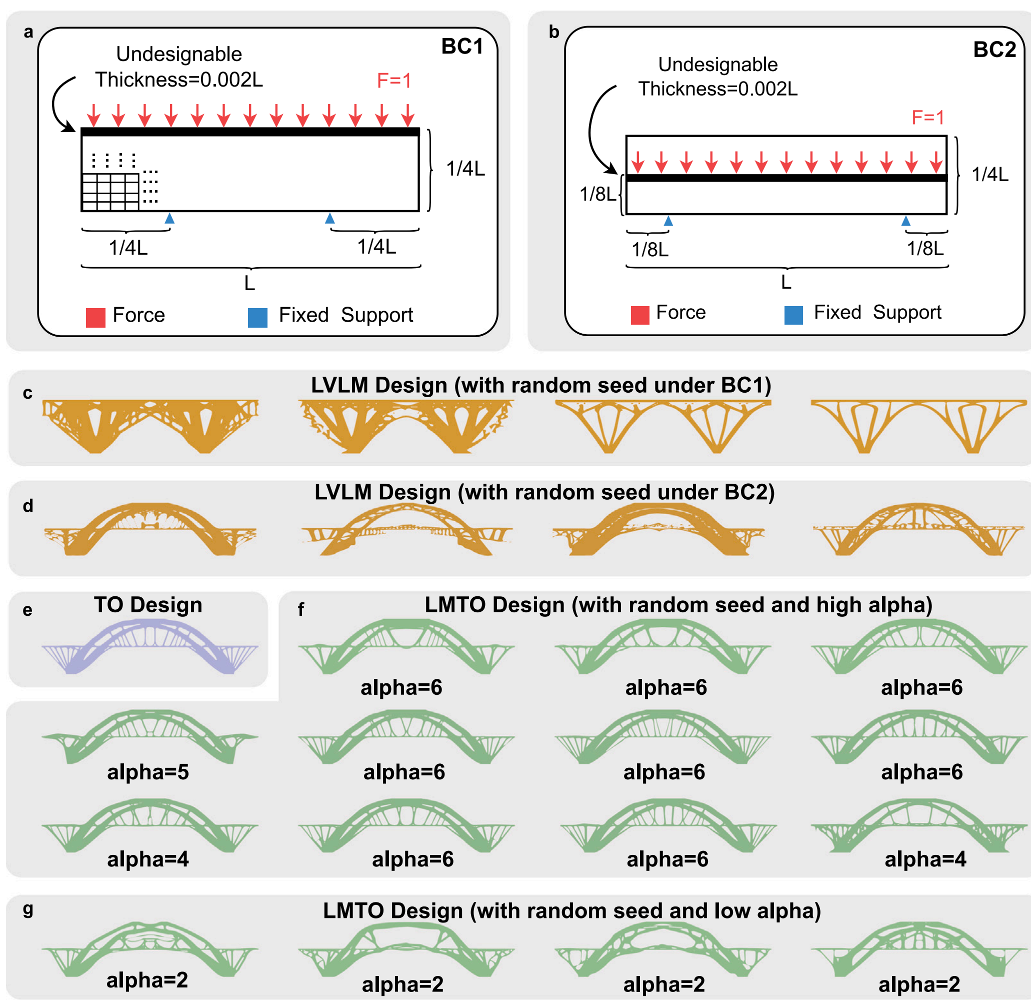


**Fig. 5.** Diverse bridge designs with LMTO and LVLM. (a) Boundary condition 1 (BC1) for bridge structural design. (b) Boundary condition 2 (BC2) for bridge structural design. (c) The LVLM bridge designs under BC1 using different random seeds *R*. (d) The LVLM designs under BC2 with different *R*. Although these LVLM designs indicate a high human preference, they exhibit low physical performance and are difficult to manufacture for spatial discontinuity. (e) The TO design under BC2, which is high physical performance. (f) The LMTO design under high-performance requirements. (g) The LMTO designs without high-performance requirements. The LMTO designs combine high performance with design diversity.

The LMTO method enables structural designs to be adjusted according to human preferences while maintaining excellent performance. Taking the designs under BC2 as an example, Fig. 5f shows the design results obtained via LMTO. In contrast to the TO designs in Fig. 5e, which focus solely on performance, the LMTO designs incorporate elements from LVLM designs in Fig. 5d, leading to greater diversity.

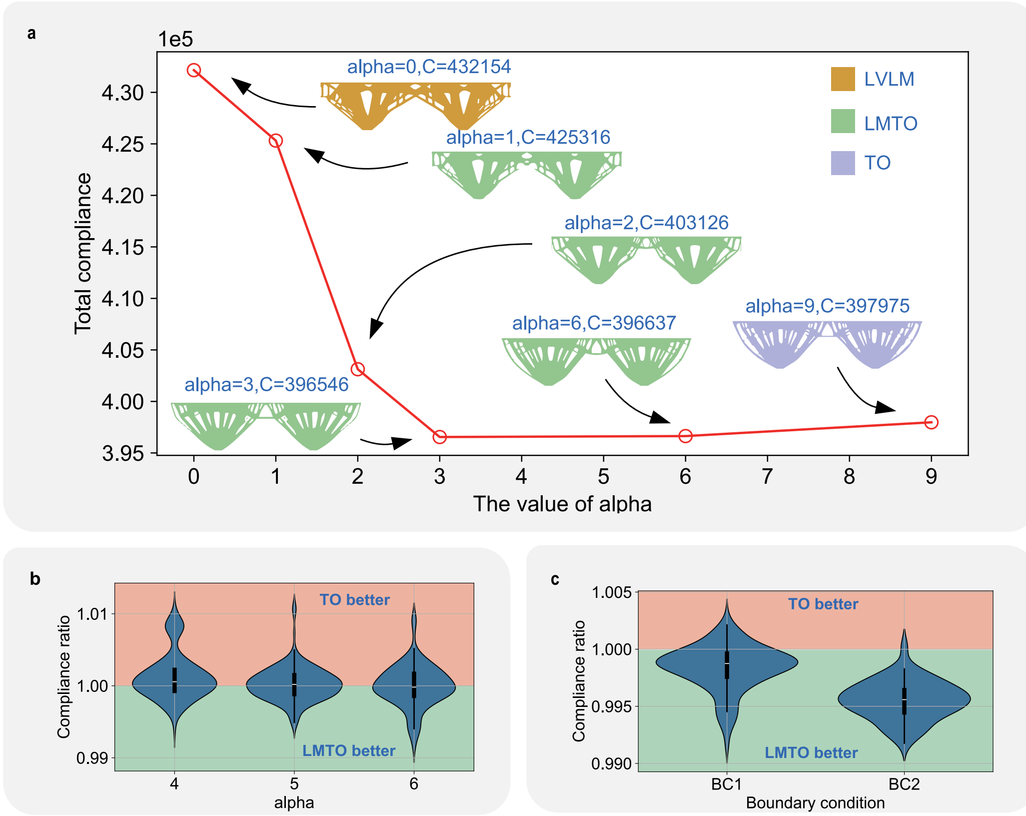


**Fig. 6.** Physical performance of the LMTO approach compared with TO approach. (a) The relationship between $alpha$ and design compliance under BC1 in the LMTO. LMTO designs vary from high preference to high performance and can reach a balance point when $alpha = 3$. (b) The compliance ratio distribution corresponds to $alpha$ among the LMTO designs under BC2. As $alpha$ grows, the median of the compliance ratio shifts below 1, indicating that designs with better performance than the TO designs become increasingly easier to find. (c) The distribution of the optimal compliance ratio as $R$ varies under different fixed boundary conditions. Regarding BC1 and BC2, the Q3 is less than 1, indicating that by selecting an appropriate $alpha$, 75% of the LMTO designs perform better than the BESO designs.

Furthermore, by adjusting the parameter $alpha$, users can enhance the emphasis on human preferences. For instance, the designs in Fig. 5g exhibit more innovation than those in Fig. 5f. Notably, despite the increased innovation, the designs in Fig. 5f show almost no performance loss relative to Fig. 5e, achieving a balance between human preferences and high physical performance in structural design.

In the UDF-Weighting Block, $alpha$ is employed to balance physical performance and human preference, and the relative relationship is shown in Fig. 6a. When $alpha$ equals 0, the LMTO design degenerates into an LVLM design, characterized by low performance but with an appearance similar to that shown in Fig. 5c. When $alpha$ equals 9, the LMTO design degenerates into a TO (Soft-Kill BESO) design, resulting in optimal performance but lacking human preference information. As $alpha$ increases, the physical performance gradually improves while human preferences diminish. $alpha$ approaches to 3 exhibits a balance between performance and preference.

It is noteworthy that certain results exhibit superior performance compared to Soft-Kill BESO. To investigate why LMTO designs outperform Soft-Kill BESO designs, we statistically analyzed 1158 randomly generated designs by varying the parameters, including $alpha$ and random seed($R$) under BC1 and BC2. In Fig. 6b and c, violin plots present the performance distribution of LMTO designs, using the compliance ratio (CR) to evaluate the performance of the LMTO relative to the TO. The CR is defined as the ratio of the compliance of the LMTO design to that of the TO design under identical boundary conditions. If most of the CR distribution is below 1, it indicates that the LMTO design outperforms the TO design. In the violin plots of Fig. 6, the black box's upper and lower edges represent the third quartile (Q3) and the first quartile (Q1), corresponding to 75% and 25%, respectively. The short white line within the black box indicates the median or 50%. The median below 1 signifies that in half of the random experiments, the LMTO design outperforms the TO design. The compliance ratio distribution corresponding to $alpha$ under BC2 is shown in Fig. 6b. It is crystal clear that as $alpha$ increases, the median of the distribution moves towards 1, aligning with the observation that higher alpha values correlate with better performance.

The compliance initially decreases and then increases as $alpha$ increases in Fig. 6a. Therefore, investigating the distribution of the minimum compliance with varying $alpha$ can illustrate the algorithm's capability to find optimal solutions within the design space. In Fig. 6c, each sample point represents the minimum relative compliance achieved by varying alpha alone (under certain constant conditions), indicating the maximum performance gain achievable through modifications to alpha. The distribution shown in the violin plot demonstrates that the third quartile (Q3) is below 1 under both BC1 and BC2. This distribution indicates that varying alpha yields a high probability of achieving values surpassing the Soft-Kill BESO performance, suggesting the effectiveness of the LMTO in exploring the design space.

In addition to analyzing the design space, Table 1 summarizes the optimal designs in performance produced by several mainstream methods. Design methods based on artificial intelligence generally rely on the TO datasets, making it difficult to surpass the performance of conventional TO approaches. In contrast, LMTO enhances the search around local optima, increasing the likelihood of finding designs that simultaneously satisfy both performance requirements and human preferences.

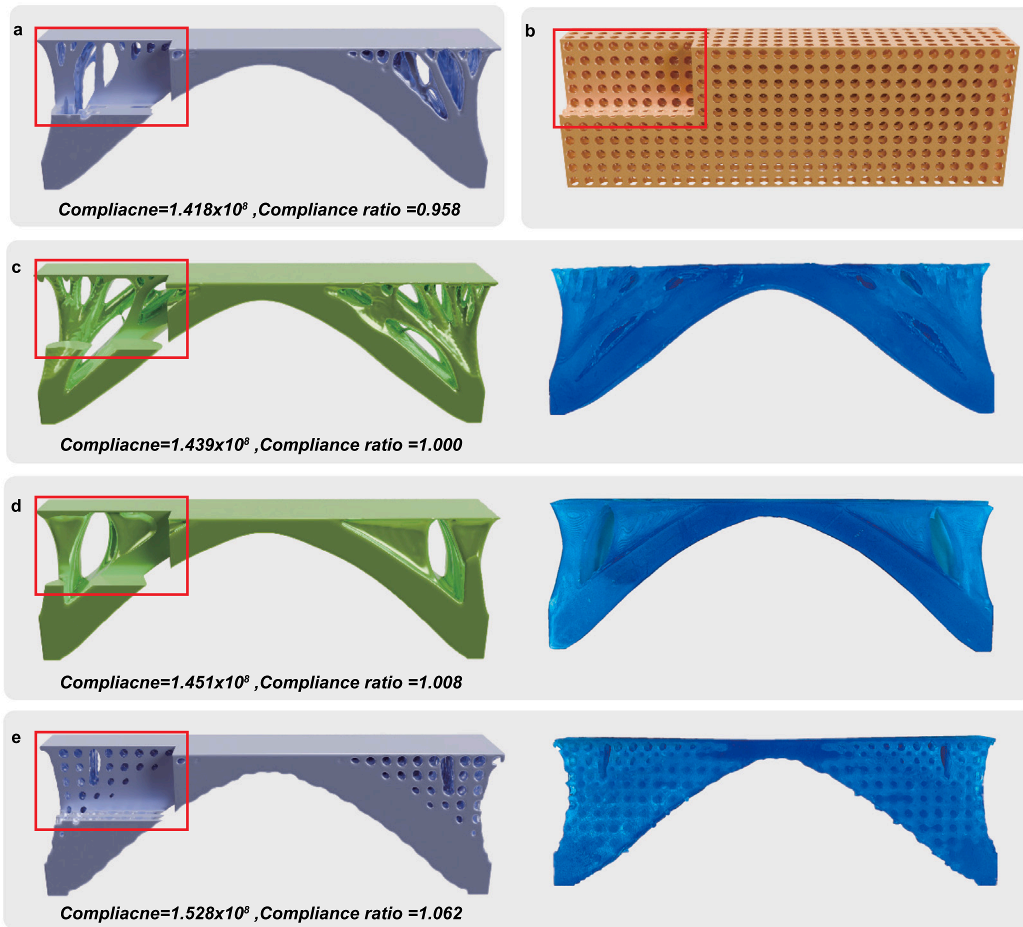


**Fig. 7.** LMTO 3D bridge designs and 3D-printed objects. (b) The yellow model represents manually edited UDF features employed to generate designs using the LMTO approach, which resemble a cheese-like structure with a uniform matrix of spherical voids. In the left column, the purple models illustrate structures refined using the LMTO optimization technique, while the green models depict structures optimized through the TO. (a) LMTO design with $alpha = 4$. (c) Soft-Kill BESO design. (d) SIMP design. (e) LMTO design with $alpha = 3$. The compliance of the structures is ranked in ascending order from top to bottom. To provide a clearer view of the internal structure, a rectangular section has been excised from the upper left corner, indicated by a red box. In (c), (d), and (e), designs in the right column are the 3d-printed results of those in the left column.

**Table 1**
2D bridge performance between different methods under BC2.

| | Volfrac | E | nu | Compliance | Compliance ratio |
|---|---|---|---|---|---|
| BESO | 0.25 | 1.0 | 0.3 | 364 029 | 1.000 |
| LMTO | 0.25 | 1.0 | 0.3 | 357 702 | 0.98 |
| MLATO | 0.25 | 1.0 | 0.3 | 388 421 | 1.067 |
| Topodiff | 0.25 | 1.0 | 0.3 | 372 457 | 1.023 |

### *5.2. 3D architecture design generation*

The LMTO formulations can be seamlessly applied to 3D problems by merely replacing the 2D model with its 3D counterpart. For the 3D scenario, the choice of LVLM is Shap-E [41], thereby naturally transforming 2D distance (from pixel to edge) into 3D distance (from point to surface) in UDF-Weighting. Similarly, 2D Soft-Kill BESO is replaced by 3D Soft-Kill BESO. As the 3D form of UDF-Weighting can traverse all positions in space and provide weighting for the TO model, the LMTO is a fully 3D voxel-based model other than treating 3D space as 2D slices for layer-by-layer optimization, which is the foundation for constructing 3D semantics.

Although most of bridge design research concentrates on 2D or pseudo-three-dimensional (pseudo-3D) systems, the realization of certain complex deformations and functionalities necessitates using 3D space. To verify the control ability of LMTO over 3D structures and compare its performance with SIMP and Soft-Kill BESO methods, we manually modified UDF-Weighting to fill the structure with uniformly sized spherical voids in Fig. 7b. Based on this modification, we generated an LMTO bridge design in a discretized area of 80 × 480 × 160 cells, with boundary conditions consistent with those of BC1.

The porous bridge design results from different methods are illustrated in the left column of Fig. 7, where each design displays distinct characteristics from one another. Specifically, the compliance of the structures is ranked in ascending order from top to bottom. The LMTO design in Fig. 7e retains the most human preferences compared with the SIMP design in Fig. 7d and with the Soft-Kill BESO design in Fig. 7c. It is crystal clear that the LMTO design in Fig. 7e is filled with regularly arranged spherical voids from the cross-sectional view within the red box in the upper left corner. From a performance perspective, it achieves a compliance ratio of 1.061 compared to Soft-Kill BESO, which slightly outperforms SIMP design, the compliance ratio of which equals 1.008. As previously mentioned, alpha can impact the performance of LMTO, so that the higher the $alpha$ in UDF-Weighting, the higher the preference in LMTO designs. Accordingly, setting $alpha = 4$ results in a compliance ratio of 0.985 (see Fig. 7a) and the LMTO design that slightly outperforms Soft-Kill BESO. Relative 3D printed models of these design structures are listed in the right column of Fig. 7, which serves as evidence of the manufacturability of LMTO designs.

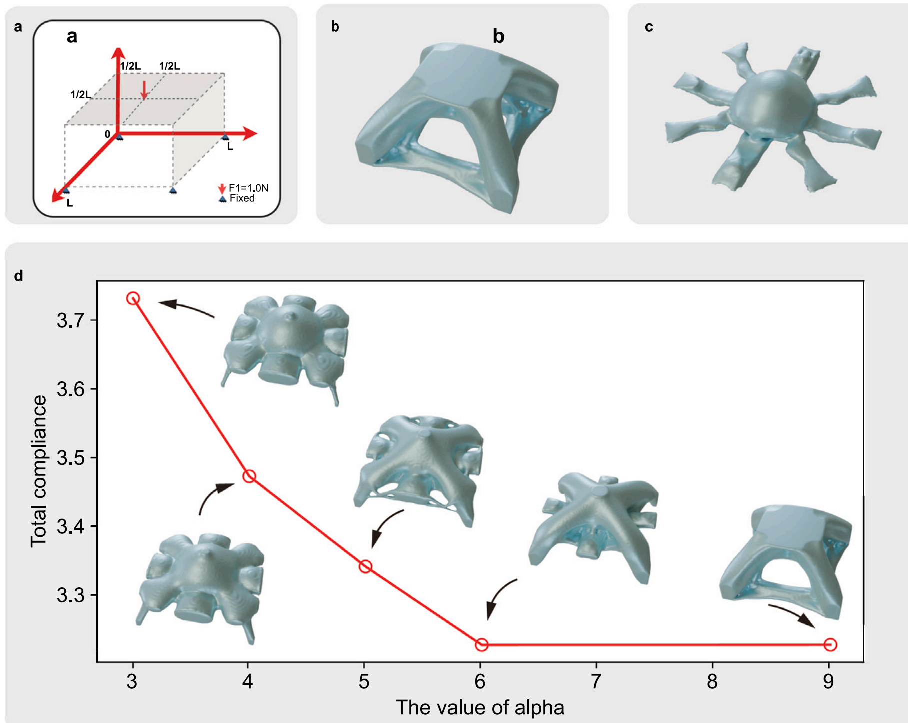


**Fig. 8.** Performance of $alpha$ on 3D Star Tracker Bracket Designs. (a) The boundary condition for star tracker's design. (b) The TO design with high performance. It is identical in shape to the LMTO design when $alpha = 9$. (c) The LVLM design with prompt "octopus spaceship". This design represents human preference and is graphically identical to the LMTO design when $alpha = 0$. (d) The relationship between $alpha$ and total compliance in the LMTO design. As $alpha$ increases, the LMTO design becomes less similar to the LVLM design while improving performance.

### 5.3. 3D spacecraft equipment design generation

Another 3D example is the appearance design of the star tracker. A star tracker is a device used to determine a spacecraft's orientation by measuring the stars' positions. In this device, the star tracker's bracket is the core component that connects the satellite deck and the star tracker. Due to the high rigidity requirements of the star tracker bracket, the industry typically employs the TO methods for structural design. Traditionally, limitations in design techniques have made it a challenging task to alter the shape of the design. By integrating artificial intelligence with structural design, the LMTO can generate numerous designs that align with human preferences. In the current study, we present the structural design of a star tracker in the shape of an "octopus spaceship".

The design domain, whose boundary condition is provided in Fig. 8a, is divided into a grid of $200 \times 200 \times 100$ cells. The filter radius and the design volume are set $r_{min} = 5$ and $V^* = 0.3V$, respectively. Fig. 8d illustrates the curve of total compliance with respect to the parameter $alpha$, whose growth leads to enhancing the performance of the LMTO design. Conversely, the LMTO design progressively resembles the preferred "octopus spaceship" configuration as alpha decreases. These results illustrate the LMTO algorithm's capability to achieve an optimal balance between aesthetic preferences and performance metrics in the design process.

### 5.4. 3D furniture design generation

Due to its ability to balance human preferences with mechanical performance, our approach is also particularly suitable for the structural design of furniture, such as a penguin chair designed for aquariums. This 3D problem space is discretized by $80 \times 80 \times 80$ linear hexahedral finite element mesh. The filter radius and the design volume are set $r_{min} = 5$ and $V^* = 0.2V$, respectively. Different design results are shown in Fig. 9, and 'a penguin' is the given prompt for the LVLM. The outcomes show that LMTO fully preserves human preferences and generates a chair-functional design in the shape of a penguin. The modulation of performance by $alpha$ remains consistent with the 2D Gothic bridge. As a result, the penguin-like appearance has been guaranteed, but the compliance ratio of performance is only 1.003. This result indicates that in a 3D context, the LMTO can construct 3D semantic information while maintaining excellent performance. More details are provided in Table 2.

A detailed comparison of 3D chair designs between the LMTO approach and a conventional method is presented in Table 2. The first column represents the design method. The second column describes the prompt used for the design. The Young's modulus(E) and Poisson's ratio(nu) are set to 1 and 0.3, respectively. If a local modification is made, the note 'part' should be appended to the prompt. When $alpha$ is large, LMTO design can give conceptual designs with a little compliance increase compared with the TO designs. If $alpha$ decreases, the compliance grows giving a design with high human preference. This conclusion is suitable for both design generation and design editing.

The ability of the LMTO to reconcile structural performance with user-driven semantic preferences is rooted in its inherent mechanism of introducing void regions during the optimization process. This fact enables a spatially efficient allocation of material density within the design domain. (see Fig. 9c and d in which the cross-section reveals the presence of cavity features).

By manipulating UDF-Weighting, our approach can edit designs based on the LVLM designs or the TO designs. In cases where the designs are to be edited, manipulating the UDF-Weighting enables localized modifications to the final design outcomes (the details of editing the design of furniture are provided in Appendix C). As the modification of structural designs can be executed semantically, this

**Table 2**
3D chair design performance among different design methods.

| Category | Name | $alpha$ | Volfrac | E | nu | Compliance | Compliance ratio |
|---|---|---|---|---|---|---|---|
| TO (Soft-Kill BESO) | Conventional chair | 9 | 0.2 | 1.0 | 0.3 | $1.783 \times 10^5$ | 1.000 |
| LMTO (generation) | Penguin chair | 2 | 0.2 | 1.0 | 0.3 | $2.091 \times 10^5$ | 1.173 |
| LMTO (generation) | Penguin chair | 5 | 0.2 | 1.0 | 0.3 | $1.789 \times 10^5$ | 1.003 |
| LMTO (generation) | Avocado chair | 5 | 0.2 | 1.0 | 0.3 | $1.789 \times 10^5$ | 1.003 |
| LMTO (editing) | Avocado chair (part) | 5 | 0.2 | 1.0 | 0.3 | $1.798 \times 10^5$ | 1.008 |

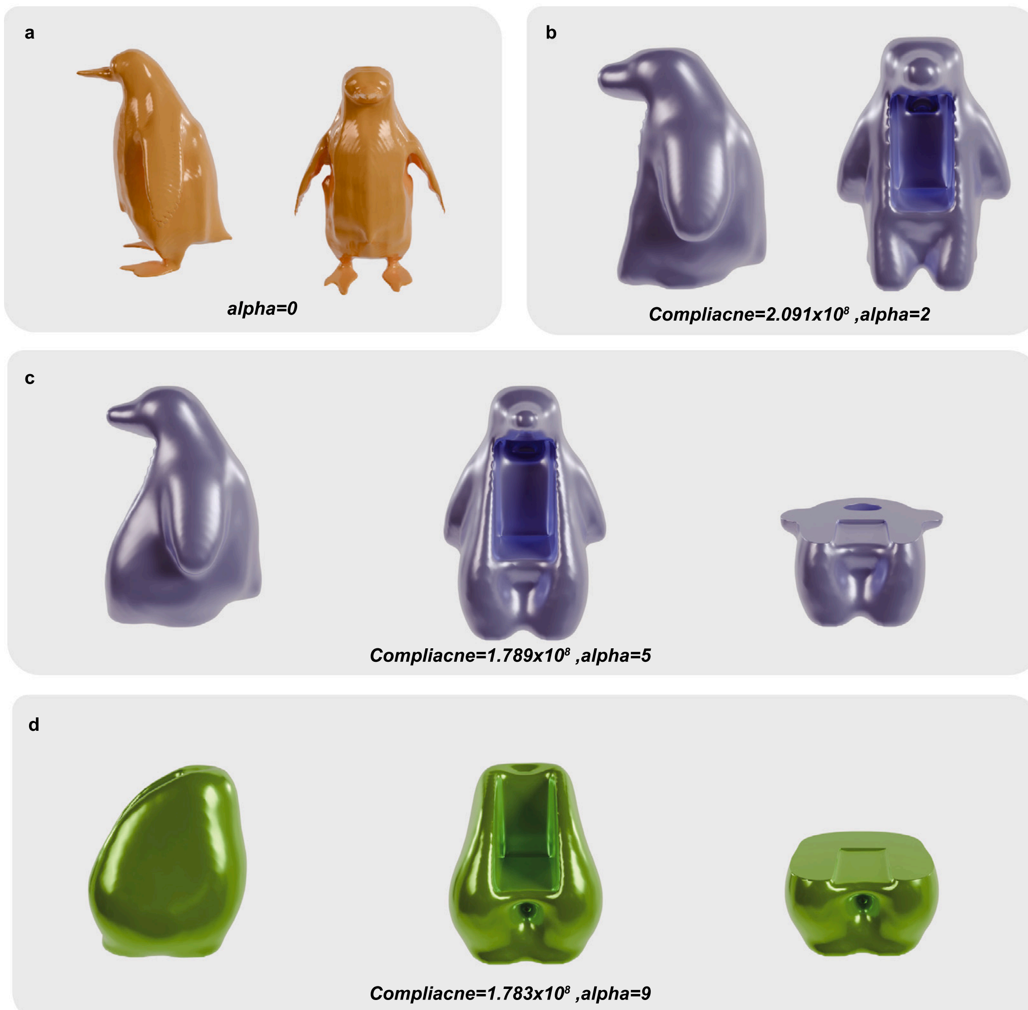


**Fig. 9.** Concrete design example of penguin chair. (a) Penguins generated by LVLM (Shap-E) with the prompt "a penguin". (b) The penguin chair design generated by LMTO with $alpha = 2$. (c) The penguin chair design generated by LMTO with $alpha = 5$. (d) The chair design generated by the TO (Soft-Kill BESO) method. As $alpha$ increases, the compliance gradually decreases. The smaller the $alpha$ is, the more similar the LMTO and the LVLM designs are. The cavity structure inside the design contributes to low compliance in LMTO design, which can be seen from their cross-sections in (c) and (d).

approach significantly expands the dimensions of design exploration. To summarize, a satisfactory design can be constructed by combining the attributes from different sources while keeping compliance within acceptable limits.

### 5.5. CAE simulation and mechanical experiments

To validate the difference between the LMTO and Soft-Kill BESO in terms of physical performance, we 3D-printed the Gothic bridge design as shown in Fig. 10c. For the convenience of 3D printing, the design volume, the filter radius, and the boundary condition are set to $V^* = 0.5V$, $r_{min} = 5$, and BC1, respectively. Regarding the parameter $alpha$, $alpha = 5$ and $alpha = 9$ are assigned to the high-performance LMTO design and the TO (Soft-Kill BESO) design, respectively. The results of the LMTO design are presented in Fig. 10a.

In the current model, simulation parameters such as Young's modulus and Poisson ratio are set to 1 MPa and 0.3, respectively. The bottom of the model is bounded to a fixed plate, whose Young's modulus and Poisson's ratio are set to $2 \times 10^5$ MPa and 0.3, respectively. The upper surface of the model undergoes a gradual parallel downward displacement of 5 mm. As illustrated in Fig. 10b, although the two designs have significantly different appearances, their energy distribution is roughly similar after applying the force. Fig. 10f shows the numerical representation of displacement as a function of applied force, demonstrating that the performances of the two designs are very similar from a simulation perspective.

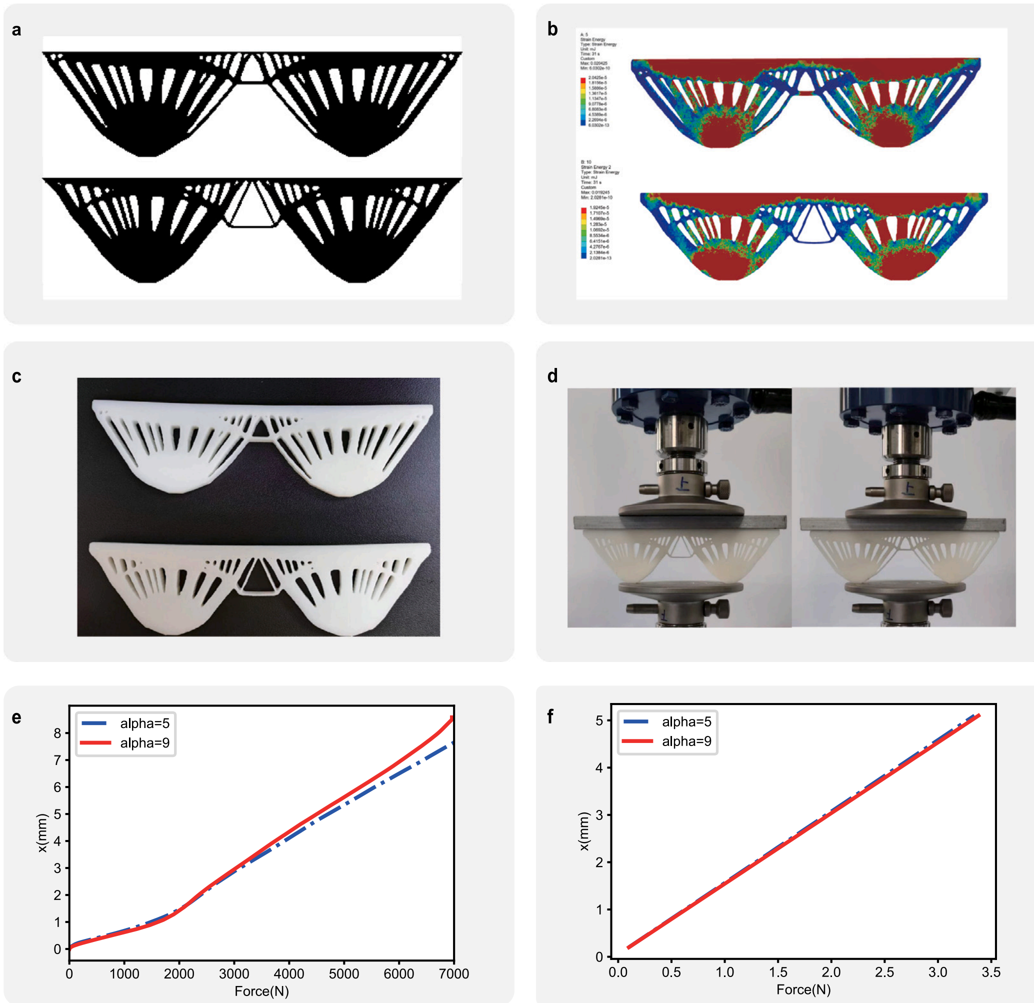


**Fig. 10.** Simulation and compression test results of LMTO bridge design and TO bridge design. (a) The results of the LMTO and the TO designs where the alpha value is set to 5 (top) and 9 (bottom), respectively. (b) ANSYS simulation results of strain energy. (c) 3D printed models. (d) Compression test (A metal block is placed on the top of the 3D printed model within the machine to ensure the uniform force distribution on the upper surface during testing). (e) Numerical results of the compression test. (f) Numerical results of Ansys simulation. The alignment of red and blue lines suggests little performance difference between the LMTO and the TO (Soft-Kill BESO) designs.

During the physical testing phase, a steel bar was placed on top of the printed bridge to ensure that the load was evenly distributed across bridge's surface (see Fig. 10d). Experimental results are shown in Fig. 10e, and empirical measurements indicate that, under the conditions of elastic deformation, the two design structures exhibit similar performance. The lines in Fig. 10e are not perfectly straight, which may be attributed to inconsistencies in the 3D printing process. Comparing Fig. 10e and f, mechanical experiment trends are consistent with the simulation results. Constrained by the precision of the 3D printer, the parameters of the bridge design are narrowed in CAE simulations and mechanical experiments, resulting in validation outcomes slightly trailing those obtained through topology optimization. Nonetheless, the LMTO and BESO designs attain closely approximate physical performance while simultaneously notably altering the aesthetic aspect, thus underscoring the method's reliability in terms of physical functionality.

## 6. Limitations

It should be pointed out that the generative capability and morphological diversity of the LMTO method are inherently constrained by the expressiveness and generalization capacity of the underlying large vision–language model. If the large vision–language model fails to accurately interpret users' intent or exhibits domain-specific limitations, the resulting design outcomes may not meet expectations. Therefore, careful selection and evaluation of the LVLM are essential to ensure semantic fidelity and design feasibility. As research on large vision–language models continues to progress, current limitations, such as insufficient coverage or inadequate semantic resolution, are expected to be progressively mitigated, further enhancing the applicability and robustness of the large vision–language model framework.

Although LMTO can achieve designs that satisfy both performance requirements and human preferences, its computational efficiency could be further improved. For example, artificial intelligence techniques could be employed to accelerate the finite element analysis process. Additionally, further research is needed to enhance the fine-grained control of large-scale and complex models. These challenges will be addressed in future research to facilitate a fully integrated process from conceptual generation to final product manufacturing.

## 7. Conclusion

In the present work, a novel design approach integrated with a large visual–language model with the TO is proposed for conceptual design. By semantically decomposing structural designs within large models, the LMTO approach efficiently explores semantic features in the joint space of human preference and performance. A bridge is also

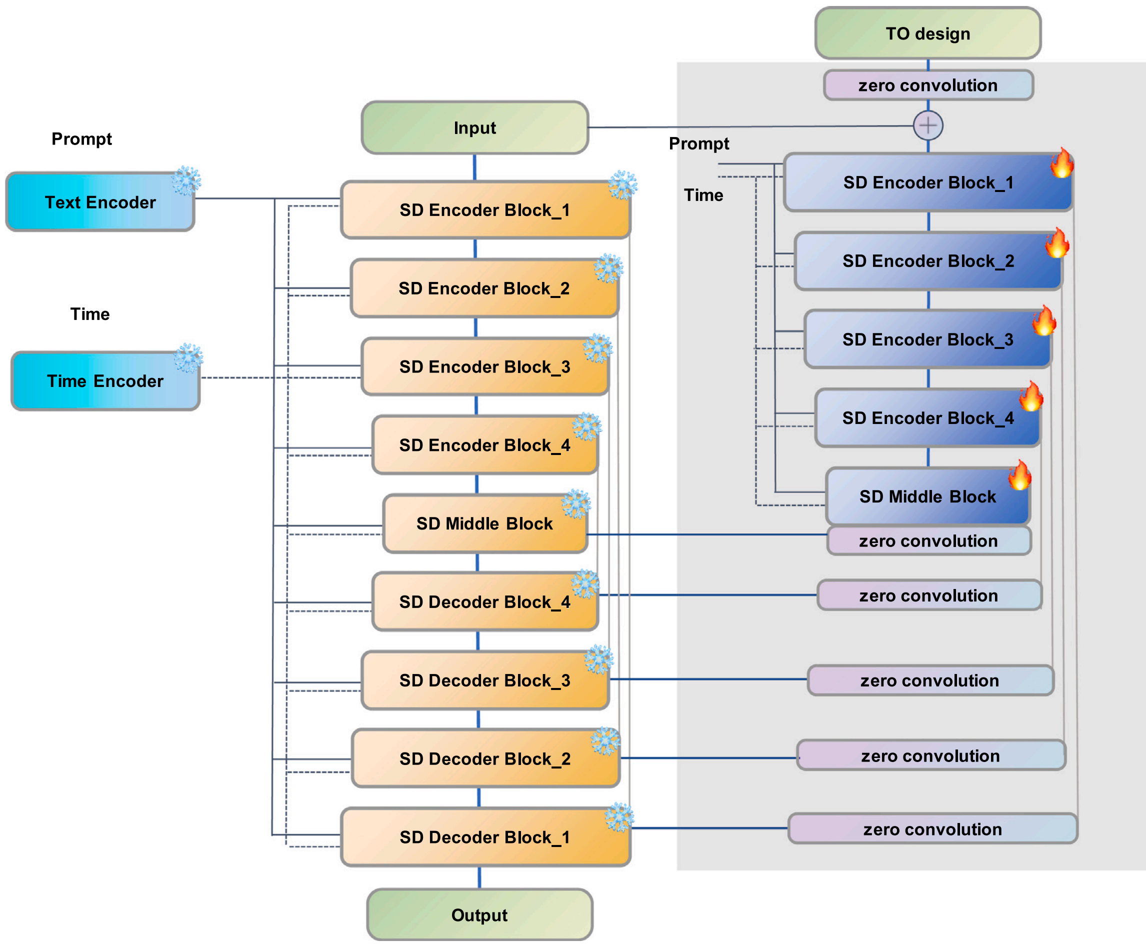


**Fig. A.11.** The network structure of ControlNet. (a) Stable Diffusion network trained on a large 2D dataset. (b) The additional ControlNet to modify Stable diffusion.

built using the TO to large models through UDF-Weighting, allowing for design outcomes that fully satisfy design boundary conditions. Simultaneously, the method offers a diverse appearance based on human preferences, thus making the materialization of conceptual designs a reality.

## CRediT authorship contribution statement

**Zelong Liang:** Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Yuan-Fang Zhang:** Writing – review & editing, Supervision, Software, Resources, Project administration. **Yingjun Wang:** Writing – review & editing, Supervision, Software, Resources, Project administration. **Weihua Li:** Writing – review & editing, Supervision, Software, Resources, Project administration, Funding acquisition.

## 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 has been supported by the National Natural Science Foundation of China [No. 52075184]. This support is gratefully acknowledged. Finally, the authors would like to thank the editors and reviewers for their valuable comments and constructive suggestions.

## Appendix A. ControlNet network structure

ControlNet is an additional network that improves correlation between Stable Diffusion and the given condition. Stable Diffusion is a latent diffusion model trained on a large dataset containing enormous image information. Due to the high cost of fine-tuning all of the Stable Diffusion's parameters, lightweight fine-tuning networks, such as ControlNet, have been proposed to reduce the conditioning control cost of the network. ControlNet structure is shown in Fig. A.11. To reduce computing costs and preserve the universal information in the original Stable Diffusion, the parameters in the Stable Diffusion part are frozen. ControlNet part parameters are fine-tuned with the additional dataset. As zero convolution exists, the whole network degenerates to Stable Diffusion if no further information is added as a condition. The outputs show a high correlation with input conditions only if additional information is given as a condition.

In this paper, the version of Stable Diffusion and ControlNet are v1.5 and v1.1. The prompt is 'Gothic bridge', and taking TO design as the condition yields the best quality for LVLM design. It is necessary to rescale the density to $[0, 255]$ and repeat its result to generate an RGB image for two reasons: First of all, the TO design has a single channel density on a scale of $[0, 1]$, and secondly, the input condition is a three-channel RGB image. To improve the resolution of the output image, the input condition is padding to the shape of $512 \times 512 \times 3$. Finally, after ControlNet, the output image is required to undergo padding removal, averaging of the third dimension, and rescaling to a range of [0,1].

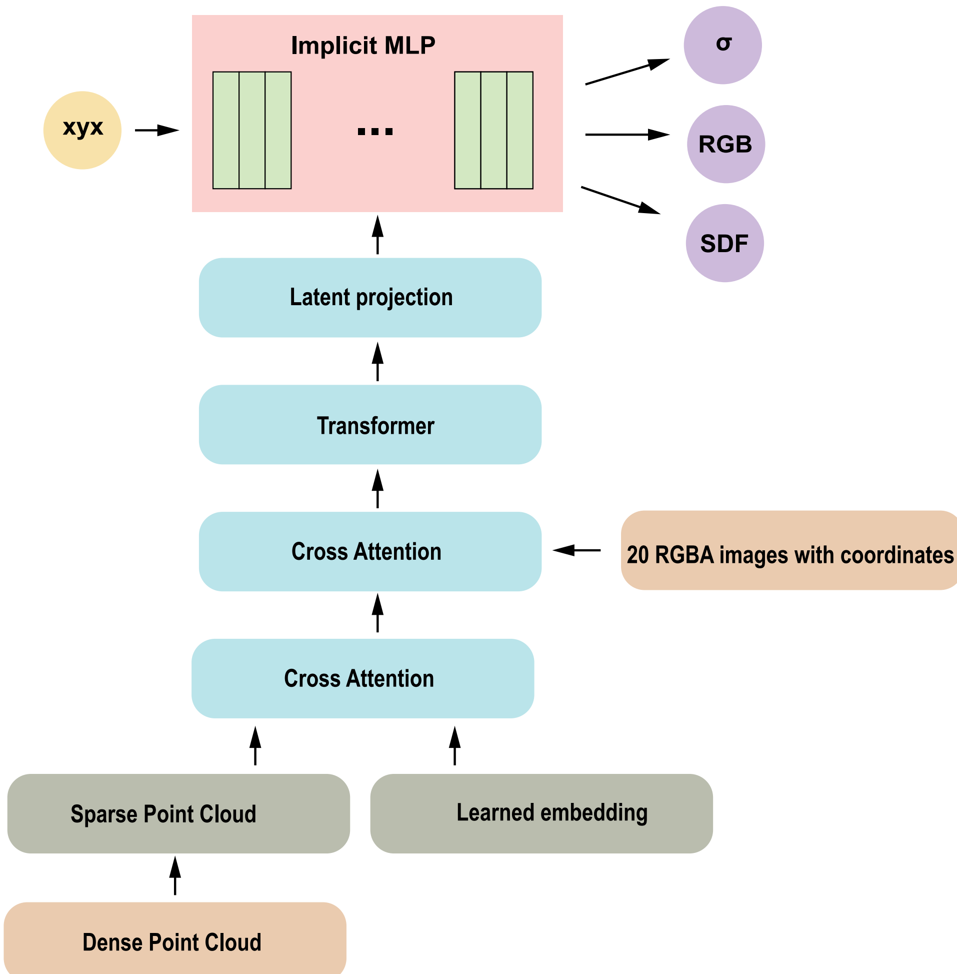


**Fig. B.12.** Shap-E network structure. Shap-E can transform the position of a space point (xyz) to density value ($\sigma$), point color (RGB), and its signed texture field (SDF) according to the information trained in the implicit MLP. Mesh can be reconstructed with this information by Marching Cubes. The implicit MLP is trained with 16k point cloud and 20 RGBA images.

After these processes, the output image is transformed into the LVLM density design.

## Appendix B. Shap-E network structure

Shap-E is an implicit neural representation (INR) trained on a dataset containing several million 3D assets. This INR can represent both meshes and neural radiance field (NeRF). The network of Shap-E is shown in Fig. B.12. By pre-training 16k point cloud and 20 different view RGBA images in Shap-E, high-dimensional information is encoded into latent projection. The latent projection is then embedded in an implicit MLP. Given a point position (xyz) in the space, implicit MLP outputs the point density value ($\sigma$), the point color (RGB), and its signed texture field (SDF). An object mesh can be reconstructed with this information through Marching Cubes.

In the 3D conceptualization design ‘penguin chair’ and ‘avocado chair’ generation, the prompts are ‘penguin’ and ‘avocado chair’, respectively. The prompt affects the resulting mesh through the attention block. After Marching Cubes, smooth meshes are produced. Then, the mesh is rescaled and translated to the center of the design space. In this step, we strive to align the symmetry axis of the design space as closely as possible with the symmetry axis of the mesh to ensure the aesthetic appeal of the final generated design.

## Appendix C. LMTO for 3D avocado chair feature editing

The LMTO algorithm can generate holistic designs that align with human preferences based on UDF-Weighting. By performing local semantic edits on the UDF-Weighting, localized modifications within the overall design is possible.

The results of localized modifications and the Boundary conditions are shown in Figs. C.13 and C.13a, respectively. Under the given prompt “avocado chair”, Fig. C.13e represents the TO design, optimized for maximum performance. At the same time, Fig. C.13f represents the LVLM design, which reflects human aesthetic preferences. Additionally, Fig. C.13g represents the LMTO design, effectively balancing performance with human preferences, and furthermore, Fig. C.13c illustrates their UDF-Weighting. If the upper half of the UDF-Weighting in Fig. C.13c is extracted and the bottom part is cut off, only semantic features from the top half will be converted to the final design from the LVLM design in Fig. C.13h compared to Fig. C.13g. To be exact, the round arm is similar to the LVLM design, while the bottom half part is the same as the TO design. In terms of performance, it has been observed that the compliance of both designs is nearly the same, while there are significant changes in their appearance (see Table 2). This result demonstrates the LMTO’s capability for localized 3D semantic modifications in structural design.

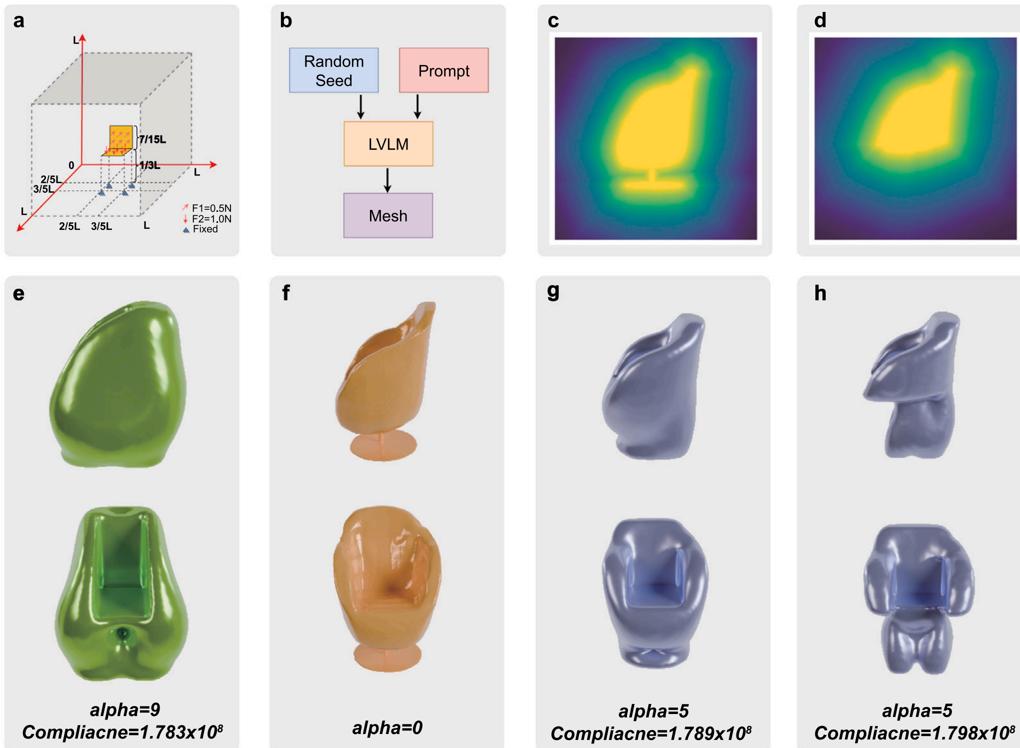


**Fig. C.13.** Results for LMTO structure editing. (a) The boundary condition for 3D chair design. (e) TO chair designs under the boundary condition (a). (b) The process for 3D chair design generation in LVLM. (f) Chair designs generated by Shap-E with the prompt 'avocado chair'. (c) UDF-Weighting generated according to (f). (g) The avocado chair design generated by LMTO according to UDF-Weighting in (c). (d) A new UDF-Weighting constructed in relation to (c) by removing the lower part of the features. (h) The avocado chair design generated by LMTO corresponding to UDF-Weighting in (d). From (g) to (h) the local spatial semantics were modified through the manipulation of UDF-Weighting. Despite the significant differences in appearance, the performance of both, in terms of compliance, remains nearly identical.

## Data availability

Data will be made available on request.

## References

[1] Y. Wang, M. Xiao, Z. Xia, P. Li, L. Gao, From computer-aided design (CAD) toward human-aided design (HAD): an isogeometric topology optimization approach, Eng. 22 (2023) 94–105.
[2] Y. Wang, X. Li, K. Long, P. Wei, Open-source codes of topology optimization: A summary for beginners to start their research, Comput. Model. Eng. Sci. 137 (2023) 1–34.
[3] A. LarsKrog, Application of Topology, Sizing and Shape Optimization Methods To Optimal Design of Aircraft Components, Altair product design, 2011.
[4] J.-H. Zhu, W.-H. Zhang, L. Xia, Topology optimization in aircraft and aerospace structures design, Arch. Comput. Methods Eng. 23 (2016) 595–622.
[5] Y. Xie, Z. Zuo, X. Huang, J. Tang, B. Zhao, P. Felicetti, Architecture and urban design through evolutionary structural optimisation algorithms, in: Proc. Int. Symp. Algorithmic Design for Architecture and Urban Design, Vol. 22, 2011.
[6] T.-U. Lee, Y.M. Xie, Simultaneously optimizing supports and topology in structural design, Finite Elem. Anal. Des. 197 (2021) 103633.
[7] J. Ma, Z. Li, Z.-L. Zhao, Y.M. Xie, Creating novel furniture through topology optimization and advanced manufacturing, Rapid Prototyp. J. 27 (9) (2021) 1749–1758.
[8] V. Bhooshan, M. Fuchs, S. Bhooshan, 3D-printing, topology optimization and statistical learning: a case study, in: Proceedings of the Symposium on Simulation for Architecture and Urban Design, 2017, pp. 107–114.
[9] M.P. Bendsoe, O. Sigmund, Topology Optimization: Theory, Methods, and Applications, Springer, 2013.
[10] M.P. Bendsoe, Optimal shape design as a material distribution problem, Struct. Optim. 1 (1989) 193–202.
[11] M.Y. Wang, X. Wang, D. Guo, A level set method for structural topology optimization, Comput. Methods Appl. Mech. Engrg. 192 (1–2) (2003) 227–246.
[12] 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.
[13] Y.M. Xie, G.P. Steven, A simple evolutionary procedure for structural optimization, Comput. Struct. 49 (5) (1993) 885–896.
[14] O.M. Querin, G.P. Steven, Y.M. Xie, Evolutionary structural optimisation (ESO) using a bidirectional algorithm, Eng. Comput. 15 (8) (1998) 1031–1048.
[15] X. Huang, M. Xie, Evolutionary Topology Optimization of Continuum Structures: Methods and Applications, Wiley, 2010.
[16] J. Martínez, J. Dumas, S. Lefebvre, L.-Y. Wei, Structure and appearance optimization for controllable shape design, ACM Trans. Graph. 34 (6) (2015) 1–11.
[17] J. Hu, M. Li, S. Gao, Texture-guided generative structural designs under local control, Comput.-Aided Des. 108 (2019) 1–11.
[18] T. Navez, M.-P. Schmidt, O. Sigmund, C.B. Pedersen, Topology optimization guided by a geometrical pattern library, Struct. Multidiscip. Optim. 65 (4) (2022) 108.
[19] Y. Li, Y. Lai, G. Lu, F. Yan, P. Wei, Y.M. Xie, Innovative design of long-span steel–concrete composite bridge using multi-material topology optimization, Eng. Struct. 269 (2022) 114838.
[20] Y. Rong, Z.-L. Zhao, X.-Q. Feng, Y.M. Xie, Structural topology optimization with an adaptive design domain, Comput. Methods Appl. Mech. Engrg. 389 (2022) 114382.
[21] X. Yan, D. Bao, Y. Zhou, Y. Xie, T. Cui, Detail control strategies for topology optimization in architectural design and development, Front. Archit. Res. 11 (2) (2022) 340–356.
[22] Z. Li, T.-U. Lee, Y.M. Xie, Interactive structural topology optimization with subjective scoring and drawing systems, Comput.-Aided Des. 160 (2023) 103532.
[23] W. Zhang, X. Zhuang, X. Guo, S.-K. Youn, Human-augmented topology optimization design with multi-framework intervention, Eng. Comput. (2025) 1–18.
[24] A.A. Khan, R.K. Madendran, U. Thirunavukkarasu, M. Faheem, D2pam: Epileptic seizures prediction using adversarial deep dual patch attention mechanism, CAAI Trans. Intell. Technol. 8 (3) (2023) 755–769.
[25] A.A. Khan, R.K. Mahendran, K. Perumal, M. Faheem, Dual-3dm 3 ad: mixed transformer based semantic segmentation and triplet pre-processing for early multi-class Alzheimer's diagnosis, IEEE Trans. Neural Syst. Rehabil. Eng. 32 (2024) 696–707.
[26] N. Yüksel, H.R. Börklü, H.K. Sezer, O.E. Canyurt, Review of artificial intelligence applications in engineering design perspective, Eng. Appl. Artif. Intell. 118 (2023) 105697.
[27] S. Shin, D. Shin, N. Kang, Topology optimization via machine learning and deep learning: A review, J. Comput. Des. Eng. 10 (4) (2023) 1736–1766.

[28] T. Zang, M. Yang, Y. Liu, P. Jiang, Text2shape: Intelligent computational design of car outer contour shapes based on improved conditional Wasserstein generative adversarial network, Adv. Eng. Inform. 62 (2024) 102892.

[29] S. Jang, S. Yoo, N. Kang, Generative design by reinforcement learning: enhancing the diversity of topology optimization designs, Comput.-Aided Des. 146 (2022) 103225.

[30] X. Li, Y. Wang, Z. Sha, Deep learning methods of cross-modal tasks for conceptual design of product shapes: A review, J. Mech. Des. 145 (4) (2023).

[31] R. Tsumoto, K. Yaji, Y. Nomaguchi, K. Fujita, Deep concept identification for generative design, Adv. Eng. Inform. 65 (2025) 103354.

[32] F. Mazé, F. Ahmed, Diffusion models beat GANs on topology optimization, 2022, arXiv preprint arXiv:2208.09591.

[33] P.S. Vulimiri, H. Deng, F. Dugast, X. Zhang, A.C. To, Integrating geometric data into topology optimization via neural style transfer, Mater. 14 (16) (2021) 4551.

[34] W. Zhang, Y. Wang, Z. Du, C. Liu, S.-K. Youn, X. Guo, Machine-learning assisted topology optimization for architectural design with artistic flavor, Comput. Methods Appl. Mech. Engrg. 413 (2023) 116041.

[35] Y. Liu, T. Han, S. Ma, J. Zhang, Y. Yang, J. Tian, H. He, A. Li, M. He, Z. Liu, et al., Summary of ChatGPT-related research and perspective towards the future of large language models, Meta-Radiol. (2023) 100017.

[36] A. Ramesh, P. Dhariwal, A. Nichol, C. Chu, M. Chen, Hierarchical text-conditional image generation with CLIP latents, 2022, arXiv preprint arXiv:2204.06125.

[37] C. Saharia, W. Chan, S. Saxena, L. Li, J. Whang, E.L. Denton, K. Ghasemipour, R. Gontijo Lopes, B. Karagol Ayan, T. Salimans, et al., Photorealistic text-to-image diffusion models with deep language understanding, Adv. Neural Inf. Process. Syst. 35 (2022) 36479–36494.

[38] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, B. Ommer, High-resolution image synthesis with latent diffusion models, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022, pp. 10684–10695.

[39] G. Metzer, E. Richardson, O. Patashnik, R. Giryes, D. Cohen-Or, Latent-NeRF for shape-guided generation of 3D shapes and textures, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 12663–12673.

[40] J. Xu, X. Wang, W. Cheng, Y.-P. Cao, Y. Shan, X. Qie, S. Gao, Dream3D: Zero-shot text-to-3D synthesis using 3D shape prior and text-to-image diffusion models, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023, pp. 20908–20918.

[41] H. Jun, A. Nichol, Shape-E: Generating conditional 3D implicit functions, 2023, arXiv preprint arXiv:2305.02463.

[42] C. Picard, K.M. Edwards, A.C. Doris, B. Man, G. Giannone, M.F. Alam, F. Ahmed, From concept to manufacturing: Evaluating vision-language models for engineering design, 2023, arXiv preprint arXiv:2311.12668.

[43] S. Zhong, P. Punpongsanon, D. Iwai, K. Sato, Topology optimization with text-guided stylization, Struct. Multidiscip. Optim. 66 (12) (2023) 256.

[44] J.-H. Bastek, Inverse-design of nonlinear mechanical metamaterials via video denoising diffusion models: dataset and model checkpoints, 2023.

[45] W. Zhang, G. Zhao, L. Su, Research on multi-stage topology optimization method based on latent diffusion model, Adv. Eng. Inform. 63 (2025) 102966.

[46] J. Ho, A. Jain, P. Abbeel, Denoising diffusion probabilistic models, Adv. Neural Inf. Process. Syst. 33 (2020) 6840–6851.

[47] L. Zhang, A. Rao, M. Agrawala, Adding conditional control to text-to-image diffusion models, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 3836–3847.

[48] J. Chibane, G. Pons-Moll, et al., Neural unsigned distance fields for implicit function learning, Adv. Neural Inf. Process. Syst. 33 (2020) 21638–21652.