arXiv:2611.28362v1 [cs.CE] 20 Nov 2026

通过可迁移表征学习高效优化机械超材料

AntonioVaragnolo[1†], Miaomiao Dai[2†], Bolei Deng[2*], Raphael Pestourie[1*]1*佐治亚理工学院计算科学与工程学院，美国佐治亚州亚特兰大，邮编30332。[2]佐治亚理工学院 Daniel Guggenheim航空航天工程学院

，

美国佐治亚州邮编30332。

通讯作者。电子邮箱：bolei.deng@gatech.edu；rpestourie3@gatech.edu；贡献作者：avaragnolo3@gatech.edu；mdai49@gatech.edu；

†作者贡献相同

摘要

设计用于非线性行为（如能量吸收）的机械超材料，受到高维参数化和昂贵仿真的限制。本文表明，可通过学习得到且可复用的搜索空间，将有关廉价线性属性的设计知识迁移至非线性、可能不可微的目标。TRAMA（面向超材料架构的可迁移表征学习）利用预训练视觉Transformer和物理信息增强的变分自编码器，对架构材料进行重新参数化。对于能量吸收效率，TRAMA 达到各个效率水平所需的评估次数，比最先进的无梯度优化减少 95% 至 98%，比基于梯度的优化少 3 至 5 次。其最佳设计与任何基线方法找到的最佳设计相当，同时评估次数至少减少 67%；在有限元仿真中，这相当于每次优化活动约节省 100 小时。制造出的设计与仿真的应力-应变曲线误差在 5.3% 以内。事后分析表明，该表征在哪些新属性上能够迁移，在哪些方面则不能。

关键词： 机械超材料、表示学习、逆向设计、架构超材料

1 引言

我们构建了一个数据驱动的搜索空间，将设计知识从线性杨氏模量迁移至非线性响应，例如能量吸收和应力-应变曲线，从而加速对评估成本高昂且可能不可微的属性进行优化。该潜在空间有效地

重新参数化架构材料，使计算上可行的设计与可制造设计之间实现有益的对齐，而通用的基于密度的搜索空间无法做到这一点。在最大化超材料的能量吸收效率时，与最先进的无梯度方法（CMA-ES）相比，所提出的方法平均以少 95% 至 98% 的评估次数达到各项效率；与基于梯度的方法（CCSA）相比，评估次数减少 66% 至 80%。考虑有限元仿真的成本，该方法从首次设计迭代起便可节省计算资源；与基于梯度的方法相比，节省的运行时间约为 100 小时，与无梯度方法相比则约为 1800 小时，而在缺少可微求解器时，无梯度方法是唯一可用的选择。此外，我们还对迁移至新属性的可迁移性进行事后分析，以形成关于这些方法适用范围和局限性的可解释认识。

机械超材料的有效属性主要源于几何结构，而非材料成分，因此可以通过逆向设计结构，而不是合成新材料，来实现强大的设计潜力。基于几何结构的设计已实现多稳态等先进功能 [1], 能量吸收 [2–4], 和可编程性 [5–9], 并应用于可拉伸电子产品 [10–12], 组织工程支架 [13], 和软体机器人 [14]。随着增材制造和制造技术的迅速发展，核心瓶颈日益从制造转向设计 [15–18]。

超材料的逆向设计仍然面临挑战，因为几何设计空间的维度极高，而结构-属性求解器的计算成本高昂。这种组合使穷举搜索的计算成本十分高昂。当梯度易于获取时，经典的逆向设计工具包括形状优化 [19] 和拓扑优化 [17, 20]。拓扑优化直接在像素空间中进行设计，能够生成丰富多样的结构，但也使性能与可制造性之间的权衡更加困难。例如，它通常会优化连续密度，从而产生灰度设计；这些设计在制造前必须转换为清晰的实体形状，且通常会因此牺牲性能，而参数化逆向设计从一开始就采用可制造的形状 [21]。此外，作为一种局部优化方法，它通常需要采用不同的初始值多次重新启动，才能摆脱较差的局部最优解。在无法获取梯度时，经典方法包括遗传算法和进化算法 [22, 23], ，这些方法已成功应用于超材料设计 [18, 24]。然而，随着设计变量增加，且对底层函数不作进一步假设时，无梯度方法的计算成本会呈指数级增长；由于维数灾难，为了探索设计空间，需要大幅增加评估次数 [25]。

深度学习已被广泛采用，通过快速的结构-属性正向预测 [26, 27]来加速设计性能的评估 [26, 28–35] 或使用生成模型 [36, 37]。然而，对可行空间的探索要实现加速，需要分摊生成数据集 [38]所需的数千次偏微分方程模拟。科学机器学习（SciML）物理信息机器学习（PIML）方法试图通过减少所需的物理评估次数或加速计算来降低基于模拟的设计成本。关键策略包括物理信息神经网络（PINNs） [39], 、求解器校正方法 [40, 41],和算子学习 [42–44]。然而，尽管取得了这些进展，大多数框架仍然优先进行从设计到属性的正向预测，而不是直接优化设计过程 [45]。此外，优化器可能会将设计推向代理模型不准确的区域，例如训练数据覆盖不足的区域，从而降低该方法的可靠性。

另一个限制是，深度学习模型难以在不同设计任务或材料属性之间迁移知识，因此优化新的属性（例如在多目标框架中）需要生成新的数据集，从而增加训练成本。为解决这一问题，迁移学习 [46] 已成功降低超材料中的数据集构建成本 [47–50], 例如用于多目标优化 [49]或用于不同聚合物复合材料之间 [48]。降低代理模型训练中数据集构建成本的其他途径，是采用多保真度 [51, 52] 或元学习方法[53–55]。最后，一个新兴方向是创建可在不同属性之间复用的学习得到的表征。这种方法已在自由空间光学等相邻领域展现出良好前景 [56–59], ，并且也已应用于机械超材料 [36, 60]。文献 [60] 利用变分自编码器（VAE）构建潜在空间，然后将潜在特征输入深度神经网络，以预测几何结构的属性。然而，如果将其用于学习物理响应，该策略会继承其他监督学习策略所面临的问题，例如分布外可靠性以及对新属性的可迁移性。相比之下，我们使用一种表征来发现高效的搜索空间。文献 [57] 学习了拓扑结构的表征，并将其迁移到相似的光学属性。相比之下，我们的表征对架构材料进行了重新参数化，使其能够在显著不同的线性到非线性属性之间进行迁移。文献 [36] 在像素网格上训练了一个变分自编码器，其针对刚度进行了优化，同时联合拟合刚度属性的回归器，使所得的低维潜在空间具有连续性，并按属性对微结构进行聚类，从而实现插值和用于形状相似性的距离度量。然而，该表征仅用于其拟合的线弹性属性，未迁移到不同的非线性或动态属性。此外，它高度依赖后处理工具来确保可制造性，例如通过阈值处理和移除孤立像素，或将非周期性微结构均质化为宏观尺度设计。相比之下，本工作关注参数化结构的可迁移性及其计算优势，这些结构能够更好地使可行空间与可制造空间相一致。

在此，我们介绍 TRAMA（超材料架构的可迁移表征学习），这是一种从线性到非线性的设计策略，它利用一个设计方案数据集——这些方案针对成本低廉的线性力学属性进行了高效优化——来学习超材料设计空间的紧凑且融入物理知识的表征。随后，我们将这一表征迁移至计算成本高昂的非线性逆向设计中。我们以旋转四边形超材料为例展示这一方法。这类架构材料以丰富的非线性行为著称，同时仍能确保可制造性 [32]。直接针对架构超材料开展工作，天然就使计算可行设计的空间与可制造设计的空间相一致。这种一致性避免了基于密度的优化所需的后处理步骤；这些步骤虽能提高可制造性，却通常会降低性能。首先，我们利用基于伴随法的优化，针对线性杨氏模量进行逆向设计。由于线性问题成本低且易于求导，我们得以高效生成一个大型几何形状数据集，涵盖广泛的弹性响应。该数据集提供了有关几何特征如何控制基本力学行为的知识。随后，我们训练一个变分自编码器，以学习这些优化设计的低维潜在空间。与原始参数化空间相比，该潜在空间作为重新参数化的设计域，潜变量数量减少了 40 倍，同时保留了具有物理意义的设计变化。我们利用所学习的潜在空间进行非线性逆向设计，目标是在大变形条件下实现完整的应力-应变响应。对于更复杂的非线性性质（例如断裂[61], ），可能无法获得可微分的非线性模拟器；因此，我们无需依赖此类模拟器，而是直接在紧凑的潜在空间中进行无梯度优化，例如贝叶斯优化。在涉及旋转、失稳和屈曲的非线性设计任务中，我们的方法显著优于在原始设计空间中进行的传统无梯度优化；在基于梯度的方法可用时，其性能与之相当，甚至更胜一筹。这些结果

表明，线性力学属性可作为高效的几何知识来源，用于非线性超材料设计，从而无需可微分的非线性仿真，即可实现复杂架构材料的可扩展逆向设计。为充分发挥这一方法的潜力，我们还提供了事后可迁移性分析，以揭示从线性到非线性的迁移可能取得的效果。

2 结果

我们以参数化超材料结构的高效设计为例，阐述我们的方法。在第 2.1 节中，我们展示所构建潜在空间的特征，并对构成我们流程的各个组件进行消融实验。第 2.2.3 节报告了该方法成功迁移至非线性能效最大化问题的结果，而在第 2.2.4 节中，我们进一步拓展到更具挑战性的非线性设计问题，使其匹配期望的应力–应变曲线。在这两个章节中，设计性能均与实验验证高度一致。第 2.3节对我们潜在空间在两个设计任务上的事后迁移能力进行了研究。

2.1 潜在空间的创建

通过对超材料的杨氏模量进行约100次短程逆向设计优化（每次40次迭代，远未达到收敛），生成了一个多样化的初始设计库。超材料的杨氏模量是一个线性、易于求导且计算成本低的属性。我们采用求解器的线性化版本，在1.6 ms内输出杨氏模量及其相对于设计参数的梯度（详细实现见补充说明 1 ）。在此应用中，选择杨氏模量似乎十分自然，因为这一属性简单且在力学领域具有广泛适用性，能够提供归纳偏置。我们采用了不同的基于梯度的优化器，并设置了不同的杨氏模量目标（最小化、最大化以及线性插值所得的数值）。具体而言，我们使用了以下三种优化器：结合增广拉格朗日处理硬约束的 Adam [62],保守凸可分离近似优化器（CCSA） [63], ，其代理模型具体采用移动渐近线方法中的形式；以及 SLSQP [64]。对于后两种算法，我们使用非线性优化库 `nlopt` [65]。为确保可制造性，我们对节点位移施加硬约束。具体而言，我们施加硬约束以避免节点交叉；对于旋转四边形，还约束最小边长以及相邻四边形边所形成的最小角度（详见补充说明 1 ）。随后，通过物理信息驱动的数据增强进一步扩充设计库。我们加入了镜像对称、所有节点的均匀平移，以及模拟周期模态的交替扰动（见图 5），以及上述操作的组合。平移和交替扰动采用六种幅值，$\{0.01, 0.02, 0.03, 0.05, 0.075, 0.10\}$ 单位为晶胞尺寸。由于更大的幅值几乎总会违反节点位移约束，因此未予采用。每个增强设计由一至三种操作组成，这些操作按随机顺序执行，且始终来自不同类型。链长度为一、二或三的概率分别为0.15、0.45和0.40。仅当候选设计满足与原始设计相同的节点位移约束，并且与其源设计及所有已保留变体在参数空间中的欧氏距离至少为0.05时，才会保留该候选设计。如图 5c所示，沿竖直方向的杨氏模量不受具有竖直对称轴的对称变换影响，而加入有利于旋转的交替扰动则会增强该模量。这些变换使我们能够将134次优化运行扩展为 ~ 30000 个最终训练实例（增加了 $\times 235$）。我们在这些设计上训练 ViT-VAE，将324维几何空间（144个水平位移参数和180个竖直位移参数）压缩为8维潜在表示，将维度降低约40倍。编码器所使用的ViT先在ImageNet-21k上进行预训练，再在ImageNet-1k上微调 [66]；其196个图块标记通过学习型注意力汇聚为一个768维特征

向量，然后进行潜在投影。解码器拆分为两个独立的多层感知机头（每个包含两个隐藏层，每层1024个单元）。使用预训练编码器可以降低数据需求，但代价是训练速度略慢：在训练初期，针对通用图像任务训练的Transformer难以区分这些设计图像，最初只会优化设计的均值。直到后续训练周期，模型才开始拟合更精细的特征。通过调度KL散度权重，在重构保真度与潜在空间结构之间取得平衡。图 1 展示了沿8维潜在空间中选定轴进行遍历的效果。在图形处理器上，VAE训练约需20分钟。有关训练的更多信息见补充说明2。我们通过比较留出集上的重构能力，评估数据增强和预训练视觉Transformer模型的有效性。我们将原始设计划分为3个数据折，并将每个变体分配到与其源设计相同的数据折中，从而确保同一原始设计的不同变体不会被分到训练集和测试集两侧。首先，在原始的、未经增强的设计上比较有无数据增强时的重构性能。我们观察到，在3个数据折中的全部3个折中，对于从头训练的模型，使用数据增强都会降低留出集原始设计上的均方误差，平均降幅为16.5%。相比之下，数据增强并不能改善预训练变分自编码器在原始设计上的表现。对于两种VAE训练情形，在留出集中纳入增强设计时，使用增强设计均能改善重构效果，这表明在增强数据上训练能够使VAE对扰动具有更稳定的鲁棒性（均方误差从0.033/0.041降至0.009/0.011）。随后，我们比较了从头训练ViT-VAE与使用预训练编码器权重进行热启动的训练方式。使用热启动时，留出集上的重构效果在3种情况中的3种均更好：使用数据增强时均方误差提高$15 \pm 1\%$，不使用数据增强时提高$52 \pm 7\%$。数据增强在一定程度上可以替代预训练，但它仍然是留出集重构性能提升的第二大驱动因素。利用预训练Transformer和数据增强构建的潜在空间，构成了后续所有优化实验的基础。

2.2 向非线性性质迁移

2.2.1 基线方法

作为基于梯度的基线，我们采用由 NLOPT库 [65] 实现的凸约束可分离近似（CCSA），这是当前受约束基于梯度的逆向设计和拓扑优化领域的最先进方法。请注意，得益于先前的工作，我们现在能够实施基于梯度的逆向设计 [32]；该工作针对这一特定机械超材料平台构建了一个完全可微的应力-应变曲线伴随模拟器（补充说明 1）。原则上，梯度可能无法轻易获得，因此我们还引入了无梯度基线。作为无梯度优化器，我们采用由 NLOPT [65] 实现的 CMA-ES，它直接作用于 324 维空间。比较两者的性能后，我们选择 CMA-ES 而非贝叶斯优化器（BO）作为全空间中的无梯度替代方案，因为其表现略好。对于 TRAMA 执行的贝叶斯优化，我们采用带有 RBF核 [67] 的高斯过程代理模型，并使用上置信界（UCB）采集函数 [68]。补充说明 3和补充表格 1 及2 阐述了基线方法优化超参数的选择依据。

2.2.2 迁移至应力-应变曲线匹配 ing

为评估我们构建的工程化搜索空间的可迁移性，我们首先将其用于非线性应力-应变曲线匹配，从计算成本低廉的线性任务拓展到计算成本高昂的非线性性质。应力-应变曲线描述了超材料在不同变形状态下的力学行为，因此是设计中的核心目标函数。曲线的不同特征直接对应不同功能：初始斜率决定

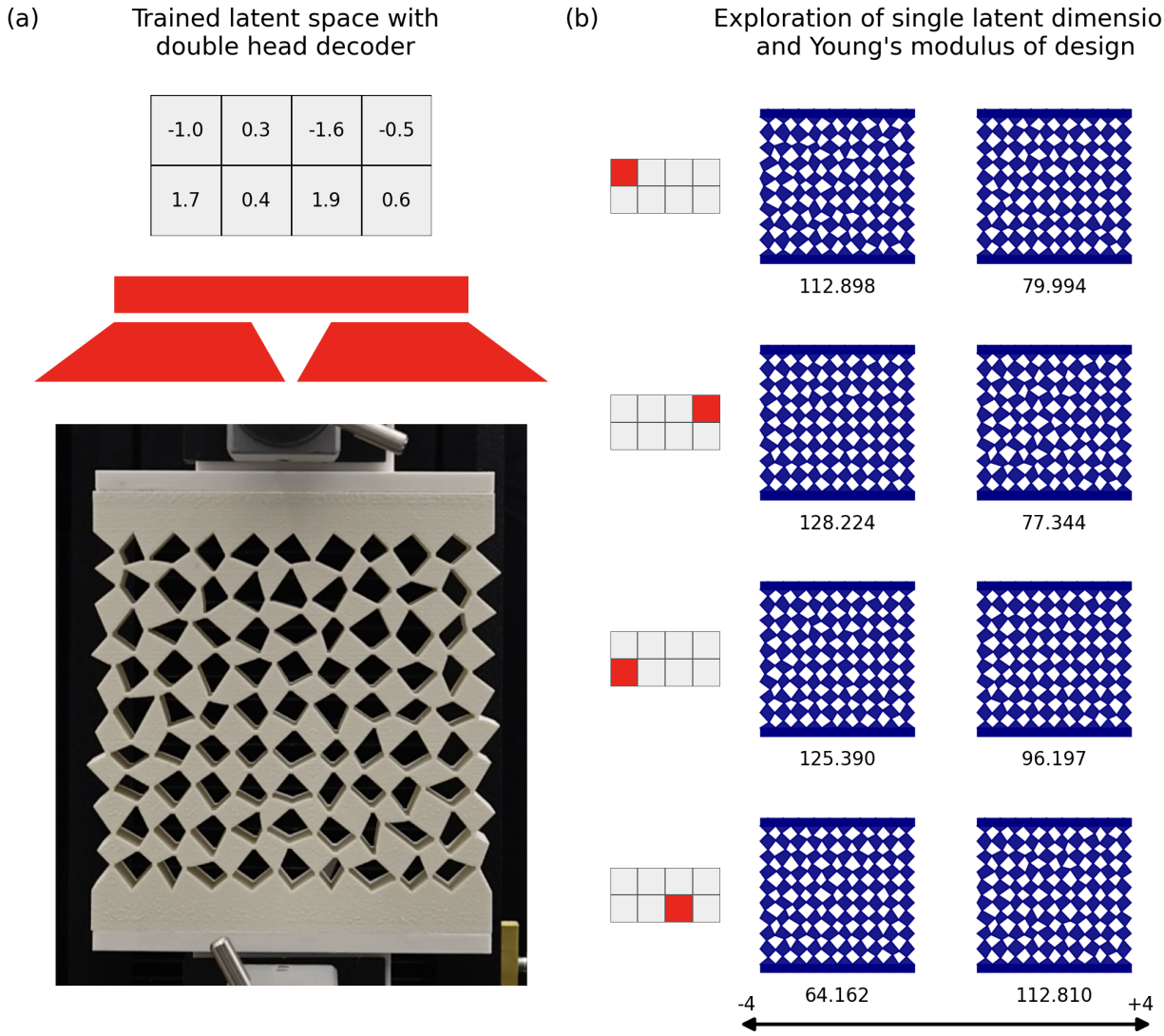


图 1：**TRAMA 的解码器和潜在空间探索**（a）双头解码器接收一个 8 维向量，并输出一种架构化超材料。这里展示了其中一个设计方案的已制造版本。（b）对单个潜在维度的探索，以及每个观测结构的杨氏模量。

刚度，曲线下面积决定吸收的能量，负斜率区域则会产生跳跃失稳和多稳态。因此，针对完整曲线进行优化，便能够设计出适用于多稳态、机器人驱动以及可编程或可重构响应等应用的材料 [5–9]。

杨氏模量在无穷小应变下定义，用于评估弹性响应，因此对应一个线性方程组。相比之下，应力-应变曲线取决于直到大变形的整个响应。在这些应变下，随着结构发生大

	线性			凹		
算法	Best	均值 ± 标准差	评估次数	Best	均值 ±标准差	评估次数
CCSA	3.12	3.68 ± 0.40	215 ± 40	**1.63**	**2.94 ± 2.79**	329 ± 200
CMA-ES	**0.98**	2.18 ± 0.68	1500	2.97	3.45 ± 0.55	1500
TRAMA	1.00	**1.93 ± 0.42**	**90 ± 28**	3.00	5.00 ± 0.79	**232 ± 51**

表 1：迁移到线性和凹应力-应变曲线匹配。误差是衡量指标 $e(f)$ (%)；评估次数将前向和反向传播各计为一次。所有优化器的均值 ± 标准差均基于 10 次运行。CMA-ES 在每次运行中均达到 1500 次评估的固定预算，但未能收敛。每列中的最佳值以粗体显示。

变形，四边形会发生大幅旋转，并开始出现屈曲等失稳现象，因此刚度本身也会随着变形而变化。

使用 TRAMA 和基线优化器，我们旨在找到一种能够匹配一组期望应力-应变曲线的设计。形式上，给定 N 个等间距点 ϵ_1，$...\epsilon_N = \epsilon_{\max}$，对于每条曲线，我们提供曲线的期望值 f_1，$...f_N$。目标是找到一种与曲线 f 匹配的超材料，最小化 $\sum_{i=1}^{N}(f(\epsilon_i) - f_i)^2$。我们针对难度逐渐增加的三条曲线进行实验：一条线性曲线、一条凹曲线，以及在第 2.2.4节讨论的一条突跳曲线（先线性增长，随后急剧下降）。我们在此讨论线性和凹曲线，因为它们是实现良好能量吸收效率的基本曲线（第 2.2.3节）；最后再讨论更具挑战性的突跳曲线。为便于直观展示误差，我们在图中报告以下指标，该指标类似于曲线分数误差：

$$e(f) = \frac{1}{n}\sum_{i=1}^{n}\frac{(f(\epsilon_i) - f_i)^2}{f_i^2} \cdot 100 \tag{1}$$

线性曲线和凹曲线的性能完整对比见表 1。我们注意到，性能差异在约 5% 的制造误差范围内。线性应力-应变曲线表示在整个变形范围内保持恒定刚度。不出所料，TRAMA 的表现最佳，其归纳偏置主要来源于杨氏模量。TRAMA 达到了与 CMA-ES 相当的误差（≈ 1.00%，而 CCSA 为 3.12%），所需评估次数比 CMA-ES 少约一个数量级，成本则与 CCSA 相当。（表 1）。

凹应力-应变曲线表现出渐进柔顺性（刚度逐渐降低，随着应变增加而松弛）。这一性质在几何形状空间中相对常见，并且在我们的训练集中得到了良好体现（见图 3）。这些曲线是构成良好能量吸收效率的基本要素之一，因此对冲击防护具有重要价值。由于这些几何形状已经包含在潜在空间中，我们预计其迁移效果良好。优化器达到的误差相近，范围为 1.63% 至 3.50%。TRAMA 的速度约比 CMA-ES 快一个数量级；在计入前向和反向求解后，其平均速度也快于 CCSA。

2.2.3 迁移至能量吸收效率最大化

随后，我们利用工程化搜索空间最大化能量吸收效率，其计算开销与曲线匹配相同。对于需要冲击缓解和冲击耗散的应用，能量吸收效率至关重要 [2, 3]。给定一条应力-应变曲线，能量效率

超材料的能量吸收能力，是以其所经历的最大应力与最大应变的乘积进行归一化后的能量吸收量。

$$\tau = \frac{\int_A f(\epsilon) d\epsilon}{\max_{\epsilon \in A} \epsilon \max_{\epsilon \in A} f(\epsilon)}$$

具有高能量吸收效率的应力-应变曲线，将刚性的初始线性阶段与较长的应力平台相结合。较长的平台会使应力保持在恒定水平，因此吸收的能量（即曲线下的面积）能够在整个可用应变范围内，以允许的最大应力持续累积；刚性的初始线性阶段则使结构几乎立即达到该平台应力，从而很少有可用应变被消耗在低应力水平下，而这些水平所做的功很小。理想能量吸收器是完美平坦的平台被立即达到这一极限情形。直观而言，迁移将会成功，因为这种力学响应与数据集中包含的物理规律有关，而我们已经在潜在空间中发现了这些规律，即由杨氏模量编码的初始响应以及凹形平台区域（见第 3 节）。在 5 个随机种子下，图 2a 展示了基线方法达到 TRAMA 的能量吸收效率所需的评估次数，按照每次迭代同时计算前向和反向传播的惯例进行计数。CMA-ES 需要约多 36 倍的评估次数，才能获得性能与 TRAMA 相当的设计，并且不同随机种子之间的差异更大。基于梯度的 CCSA 平均需要 631 ± 222 次评估，才能达到 TRAMA 通过 220 次前向评估实现的效果（约为计算开销的 $1/3$ ）。图 2c 展示了在给定评估次数下，各方法找到的最佳能量吸收效率的平均值和标准差。对于最高达 85% 的每个效率水平，TRAMA 所需的评估次数都比 CMA-ES 少 20–50 倍，对应于稳定的约 1.5 个数量级的加速。与 CCSA 相比，TRAMA 平均只需 3–5 倍更少的评估次数，即可达到给定的效率水平。作为局部优化器，CCSA 更依赖优化的起点，因此最终标准差更大。这意味着，尽管其平均表现相近，但其最佳设计比 TRAMA 的最佳设计高出 1 到 2 个百分点。图 2b 描绘了 TRAMA 的一个优化设计（并非总体最佳设计）及其在压缩下的制造版本。感兴趣的读者可参阅补充信息，了解有关制造（补充说明 1）和实验校准（补充说明 1）的信息。在图 2d 中，我们展示了制造并经过实验测试的设计的应力-应变曲线。仿真与现实差距定义为模拟应力-应变曲线与实测应力-应变曲线之间的逐点相对误差 $\|f_i^{\text{sim}} - f\text{exp}_i \quad \|_2 / \|f\text{exp}_i \quad \|_2$，在所有运行中的平均仿真与现实差距为 5.3%（TRAMA 为 5.9%，CCSA 为 4.8%）。对于两种算法，实测效率都低于模拟效率，平均下降 3.1 个百分点（TRAMA 下降 3.7 个百分点，CCSA 下降 2.5 个百分点）。重复我们在第 4 节中进行的示意性成本分析，并使用定制的动力学模拟器，我们发现，对于无梯度基线，在第一个设计上即可实现计算节省；对于基于梯度的基线，则在 6 个设计之后实现计算节省。在设计属性仍然相对廉价的情况下，重新利用潜在空间来适应新的属性，有助于摊销数据集创建和训练的固定成本。

2.2.4 迁移至非线性应力-应变突跳曲线匹配

我们展示了突跳曲线这一更难的案例，在该案例中迁移失效（图2e 和图2f）。在与其他实验相近的预算下，TRAMA表现不佳，最佳设计的误差为11。3%。相比之下，CCSA能够有效复现突跳，最佳误差达到5。9%。就误差而言，CMA-ES的表现与CCSA相近；但从定性角度看，CMA-ES的最佳设计比TRAMA的最佳设计更忠实地再现了应力的下降。正如我们在第2.3节的消融分析中所展示的，这一性能差距源于构造的潜在空间表达能力不足。我们用于创建潜在空间的最佳杨氏模量设计空间，使几何形状优化产生了与这种突跳行为不一致的偏差。

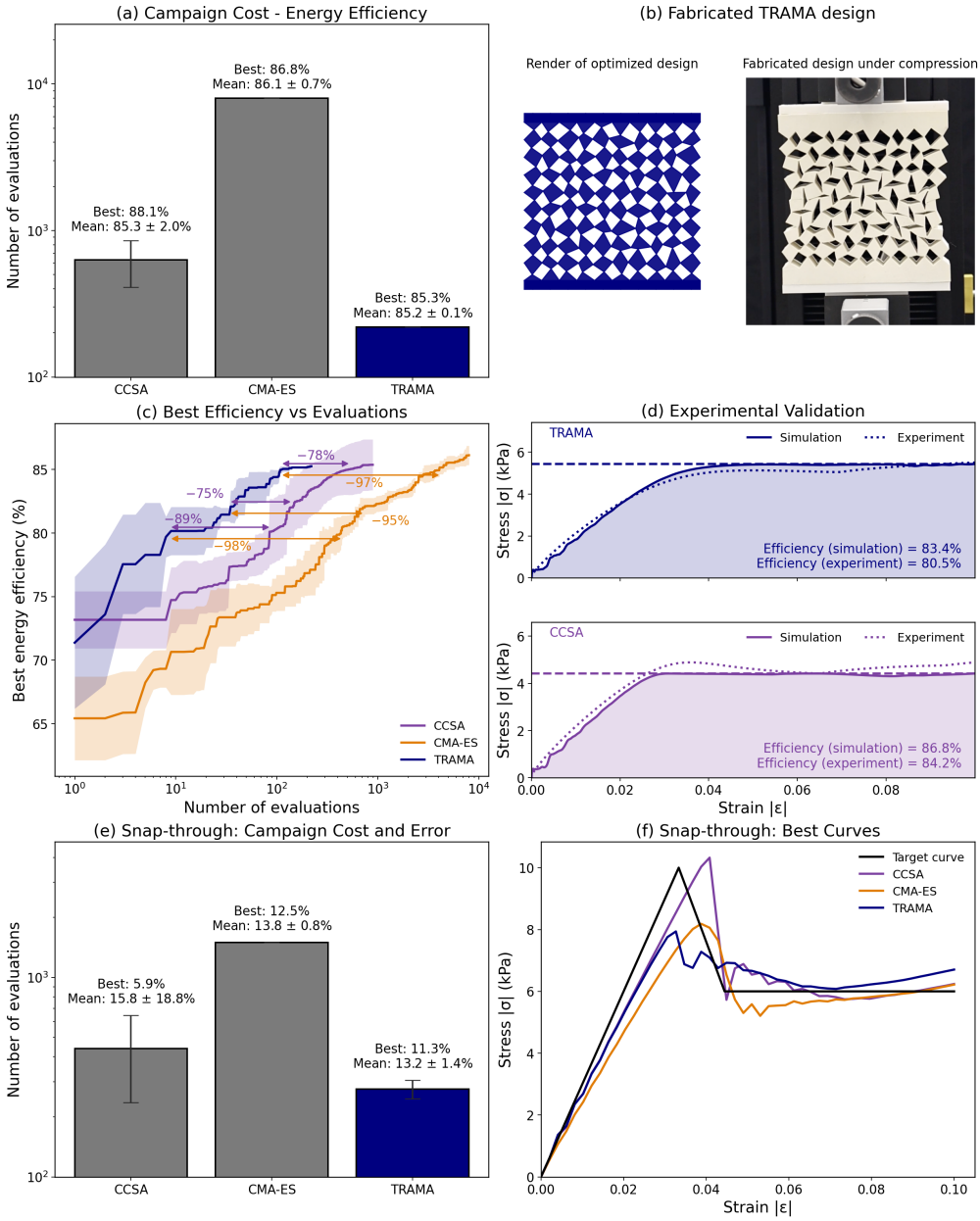


图 2：**能量吸收效率最大化设计与制造结果**（a）TRAMA 以 CCSA 预算的 35% 和 CMA-ES 预算的 3%，实现了相当的平均 85.2% 能量吸收效率。（b）TRAMA 优化的超材料几何结构及其在压缩下制造的设计方案。（c）对于每个评估次数，优化器找到的最佳设计（在 5 个随机种子下的均值 ± 标准差）。我们展示了达到 80%、82% 和 85% 目标时的平均预算节省。（d）仿真与实验应力-应变曲线。总体仿真与现实差距为 5.3%（TRAMA 为 5.9%，CCSA 为 4.8%）。（e-f）杨氏模量知识无法同样有效地迁移到突跳曲线设计中。CCSA 能够实现误差在 6% 以内的最佳设计。TRAMA 持续找到误差超过 10% 的设计，揭示了潜在空间表达能力限制。9

2.3 TRAMA 的可迁移性分析

我们分析了 TRAMA 从基于杨氏模量构建的潜在空间向两个目标属性——能量效率和 Snap 曲线匹配——的可迁移性。我们展示了评估向新属性迁移效果的三种可能方法。第 2.2.3 节表明，该空间成功迁移到了能量效率任务；而在恢复能够呈现突跳曲线的设计方面则失败了（第 2.2.4节）。

潜在空间是对数据集中存在的几何特征所学习到的表示。通过谨慎选择数据生成和增强方式，我们引导该表示学习原始设计任务中高性能设计的特征。当指导杨氏模量优化的相同底层几何特征对于新的目标属性仍然相关时，迁移便是成功的。相反，如果新问题依赖于训练数据中不存在的根本不同的几何基元，那么潜在表示很可能无法有效捕获或表达这些特征；更糟糕的是，它可能会使设计偏离高性能方案。从这个意义上说，我们进行的第一项可迁移性测试关注的是杨氏模量最佳设计空间与目标属性最佳设计空间之间的一致性。计算训练集内每个设计的目标属性后可以发现，能量效率分布（图 3a）中包含能量效率达到 79% 的具有竞争力的设计，仅比 CCSA 的最佳结果低 9.9%。相比之下，突跳曲线设计需要探索训练分布之外的区域，训练集中的最佳误差约比 CCSA 的最佳结果差 292%。事实上，在图 3b 中，我们展示了训练集中设计的应力-应变曲线。其中存在类似凹型和线性情形的曲线，但没有出现与目标具有相同幅度的突跳曲线。这项测试只需要对用于构建数据集的廉价属性进行提前停止优化，并对其中一部分设计评估目标属性。其成本主要由变分自编码器的训练决定，因此可以在投入构建潜在空间之前预先完成。

第二项测试考察训练后解码器对迁移属性的最优设计进行重构的能力。具体而言，我们检查训练后的自编码器能否重构 CCSA 找到的最优设计。如果最优设计不在训练集中，那么 CCSA 的解很可能无法在学习到的潜在空间中表示。合理地说，突跳曲线实验正是这种情况。如图3c 所示，5 个 CCSA 最优设计的平均重建误差约为最优杨氏模量设计（取自 ViT-VAE 测试集）的 4 到 5 倍。令人惊讶的是，5 个能量效率 CCSA 设计也呈现出类似结果。这表明，训练集中的优质设计在结构上不同于 CCSA 找到的设计，凸显了在数据集生成策略中使用多个优化器生成数据集的价值。图4a 定性展示了 CCSA 针对两个任务找到的最佳设计重构不佳所产生的影响（最左列）。中间列展示了自编码器重构后的两个设计。在这两个示例中，重构不佳导致目标函数分别比原始设计差.7% 和 1260.2%，对应的任务分别为能量效率和 Snap 曲线匹配。

最后，我们分析优化过程中如何探索解码器的潜在空间，并将其与 CCSA 运行结果进行比较。即使潜在空间中包含分布内解，且重构效果良好，贝叶斯优化器仍可能难以高效地找到这些解。为此，我们将 TRAMA 和 CCSA 探索的几何形状投影到设计空间的前两个主成分上（图 4b ），并通过点的透明度表示解的质量。有关主成分分析拟合的详细信息，请参阅补充说明 2。解的质量通过点的透明度表示（质量越好，透明度越低）。该空间具有高维度，而前几个主成分

仅涵盖了最优设计方案多样性的约 30%；尽管如此，主成分分析证实，TRAMA 完全跳过了具有优质 CCSA 设计方案的区域，因为相对于其训练集而言，这些方案属于分布外。第二项和第三项测试需要经过训练的变分自编码器，并在低维潜在空间中进行优化，因此属于事后分析：它们解释了迁移为何表现不佳，而非提前预测这一结果。

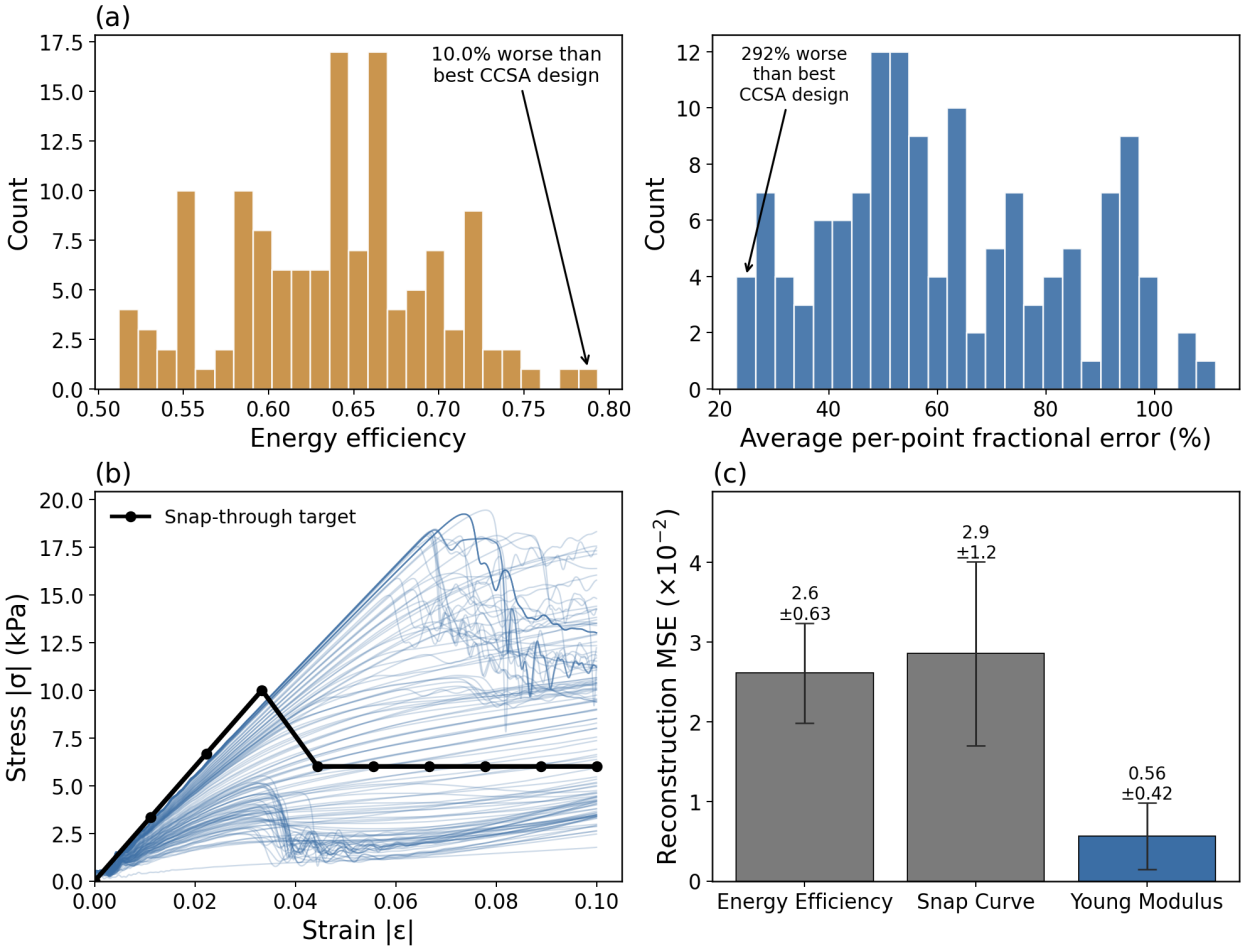


图 3：**TRAMA 训练集的可迁移性及定量分布外误差**(a) 训练集属性揭示了源设计空间与目标设计空间之间存在不匹配。左：能量效率分布显示，分布内设计具有竞争力，达到 79%，仅比 CCSA 的最佳结果低 9.9%。右：训练集中 Snap 曲线的曲线分数误差比 CCSA 高 292%，这表明目标最优结构位于学习到的分布之外。(b) 训练集中超材料的应力-应变曲线。(c) 对 CCSA 的最优设计进行重构时，两项属性的误差均为杨氏模量验证误差的 4 至 5 倍。对于 Snap 曲线匹配而言，这是训练集不匹配所致的预期结果。这表明，在实现能量吸收效率最大化时，基于梯度的优化器找到的设计在结构上不同于变分自编码器潜在空间中表示良好的设计。

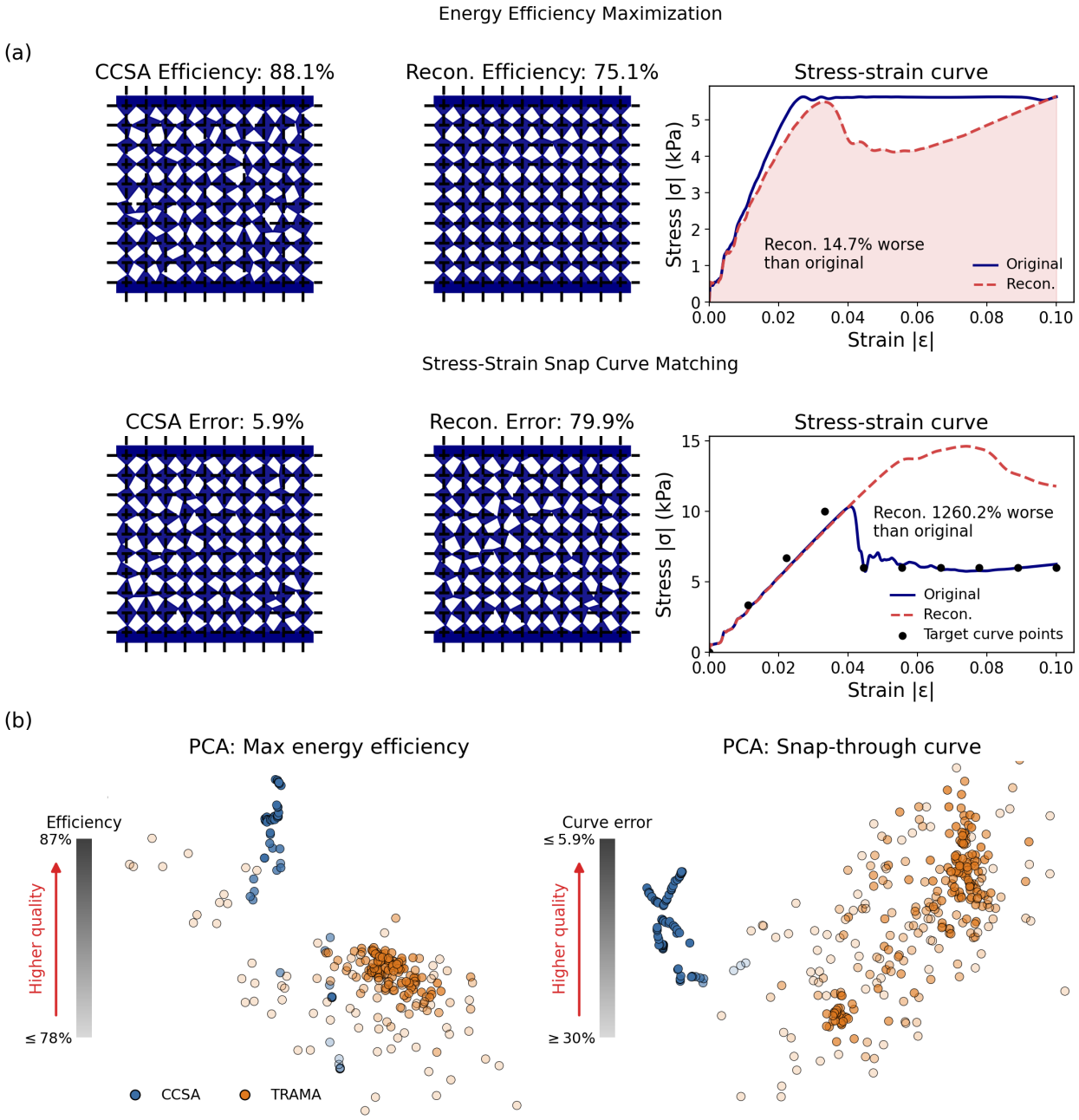


图4：**TRAMA 的定性分布外重构与优化轨迹**(a) VAE-ViT 对 CCSA 超材料进行较差重构的两个定性示例。左侧为 CCSA 针对两种迁移属性找到的设计方案。重构设计（中间）相较于原始结构，其指标分别差 14.7% 和 1260.2%。(b) 多次 CCSA 实验活动（蓝色）和最佳 TRAMA 实验活动（橙色）的 PCA 投影。感兴趣指标的质量越低，透明度越高。捕获的总方差分别为 32.5% 和 34.9%。

3 讨论

我们使用 TRAMA 提出了线性到非线性的可迁移性的实证证据。本研究补充了旨在实现此前无法企及的高成本设计方案的更广泛研究，并为仍有待探索的研究方向铺平了道路。TRAMA 可能扩展至力学领域的其他性质，以及偏微分方程约束优化的其他领域。针对我们的架构材料而特别定制的方法组件，应可直接应用于其他参数化方案。科学与工程中的许多设计任务，会将计算成本低且数学性质良好的物理过程，与计算成本高且可能不可微的物理过程结合起来。低成本物理过程为优良设计的几何特征提供信息，而由此得到的低维表示则使任意优化器——无论是否可微——都能够探索高成本问题。从这一角度看，学习得到的潜在空间是一项可复用资产，只需构建一次，便可在迁移至多个下游目标的过程中摊销其成本。相比之下，物理响应代理模型与单一目标绑定。只要存在低成本代理，这一逻辑便适用，例如从静态响应迁移到动态响应 [69], 、从单物理场迁移到多物理场 [70],，或从低保真模拟器迁移到高保真模拟器 [71, 72]。snap示例表明，静态潜在空间也存在局限性。未来，对潜在空间进行微调，或许能够更好地捕捉新性质对应的最优几何形状的多样性，而这种性质可能难以迁移。可以预先诊断可迁移性，以确定是否需要针对新性质进一步开展数据生成，从而避免失败的迁移尝试。我们的事后可迁移性分析通过提出一个多步骤诊断框架，朝着更好地理解我们的方法及其可迁移性的方向迈出了一步。我们选择从杨氏模量开始是出于便利性考虑，而向能量最大化的迁移则源于领域专业知识，并得到了实证结果的验证。然而，未来的研究可以更加系统地、以与物理无关的方式，为可迁移性质的课程安排提供依据。

另一项研究方向关注潜在空间结构与优化动态之间的相互作用。训练过程中 KL 散度的权重控制着潜在空间的平滑性和结构，但其与优化器行为及性能之间的关系仍有待进一步研究。将该方法扩展到完全三维参数化材料的应用，有望利用低维潜在空间获得更大的计算收益，因为参数化维度会呈指数级增长。仿真成本也会进一步拉大廉价代理与目标非线性属性之间的差距。

4 方法

此处将 TRAMA（用于超材料架构的可迁移表示学习）框架应用于参数化机械超材料设计知识从线性到非线性的迁移。这种混合方案将基于梯度的优化库与生成式变分自编码器（VAE）和贝叶斯优化器相结合。该框架分为三个阶段：首先，利用基于梯度的优化，针对杨氏模量等计算成本较低的线性属性，创建一个包含多种几何特征的设计方案库。尽管这些设计方案并非最优，但包含可迁移的几何基元，因此能为能量最大化和应变-应力曲线匹配等更复杂属性的优化提供有用的归纳偏置。随后，通过场域启发式方法扩充这些设计方案，无需额外模拟。其次，变分自编码器将这些设计方案压缩到紧凑的潜在空间中，使维度降低多个数量级，从而能够在这个新的低维潜在空间中使用无梯度优化器，而不是在高维参数化空间中进行优化。第三，采用现成的贝叶斯优化器，在这个低维潜在空间中高效搜索，以找到能优化计算成本更高或不可微属性的结构。

通过在逆向设计运行中设置提前停止，生成多样化的初始数据集

首先，我们选择一种计算成本较低且可微的属性来创建数据集。使用不同的基于梯度的优化器进行提前停止的逆向设计运行，与 [57]类似。我们的方法有两点显著不同。第一，使用不同的优化器能够提升最终设计方案的多样性（图5）。第二，迭代次数的选择依据是平均误差随迭代次数的变化情况，并在优化曲线的收益递减前停止（如图5所示）。

利用物理信息变换进行数据增强

无需额外模拟，为了丰富该数据集，我们通过应用一组物理规则来修改每个设计。这些规则在显著改变几何结构的同时，保持设计在其优化任务上的良好性能。这些变换及其多种组合使我们能够将数据集扩充约230倍，同时保留物理意义，并将有用的领域专家知识注入数据集。其中包括镜像对称（水平、垂直和180度旋转）、平移或特定于参数化的变换，如图5所示。如果某种变换使几何结构超出可制造空间（即满足特定于参数化的制造要求的设计集合），则会将该设计从数据集中筛除。

预训练图像模型

我们使用变分自编码器构建潜在空间。输入和重构输出取决于设计所采用的表示方式，主要包括设计本身的图像，或像架构材料情形中定义几何结构的参数。架构的选择取决于我们提供的高维输入：对于图像，通常优先采用卷积神经网络（CNNs）；而对于参数化向量，多层感知机（MLPs）则更为适用。当输入为图像时，编码器可受益于在大型通用图像数据集上预训练的模型。已有研究表明，这类预训练表示能够减少任务特定数据的需求量 [73, 74], ，这对于我们尤为有价值，因为我们关注的是整个流程的计算效率。编码器架构有多种选择，我们采用视觉变换器（ViT） [66], ，与文献 [57]类似。在这里，编码器将结构化晶格的渲染图作为单幅图像输入，并通过学习型注意力汇聚预训练ViT的图块标记，以获得潜在空间。相比标准图像自编码器，解码器并不重构图像：两个多层感知机分别输出晶格水平轴和垂直轴上的节点位移，因此，解码样本直接对应于几何结构，其可制造性可以得到精确评估。训练过程最小化重建误差，以促使潜在向量保留重建设计所需的信息；同时最小化与标准正态先验之间的Kullback–Leibler（KL）散度，从而促进形成平滑且结构良好的潜在空间，使相邻点解码为相似的设计。更多架构细节见补充说明2。

潜在—物理空间中的优化

训练好变分自编码器后，其解码器 $\mathcal{D}$ 被冻结，并用作设计生成器，将低维潜在空间映射到参数空间 $\mathcal{D}(\mathbf{z}) = \mathbf{p}$。由此，我们将原始参数空间中的约束设计问题重新表述为潜在空间中的低维无约束问题，

$$\min_{\mathbf{p}\in\mathbb{R}^{n_p}} J(\mathbf{u};\mathbf{p}) \quad \text{s.t. } \mathbf{F}(\mathbf{p}) \le \mathbf{0} \qquad \longrightarrow \qquad \min_{\mathbf{z}\in\mathcal{Z}} \tilde{J}(\mathbf{z}) = J\big(\mathbf{u};\mathcal{D}(\mathbf{z})\big) + \lambda\, G\big(\mathcal{D}(\mathbf{z})\big), \tag{2}$$

其中，$\mathbf{p}$ 是几何参数向量，J 是计算成本高且可能不可微的目标函数（例如能量吸收效率或应力-应变曲线失配），$\mathbf{u}$ 是通过模拟器获得的设计$\mathbf{p}$ 的力学响应，而$\mathbf{F}(\mathbf{p}) \leq \mathbf{0}$ 汇集了可制造性约束，这些约束仅取决于几何参数。在潜在空间中，约束作为惩罚项纳入目标函数：$G(\mathbf{p}) = \max_i F_i(\mathbf{p})$ 若违反程度超过容差；否则为 $G = 0$，并由系数 $\lambda \geq 0$加权。此处选用贝叶斯优化器（BO）作为优化器，如参考文献 [57] 所述；但也可采用任何现成方法，包括基于梯度的方法（前提是所优化的函数具有可微求解器）。

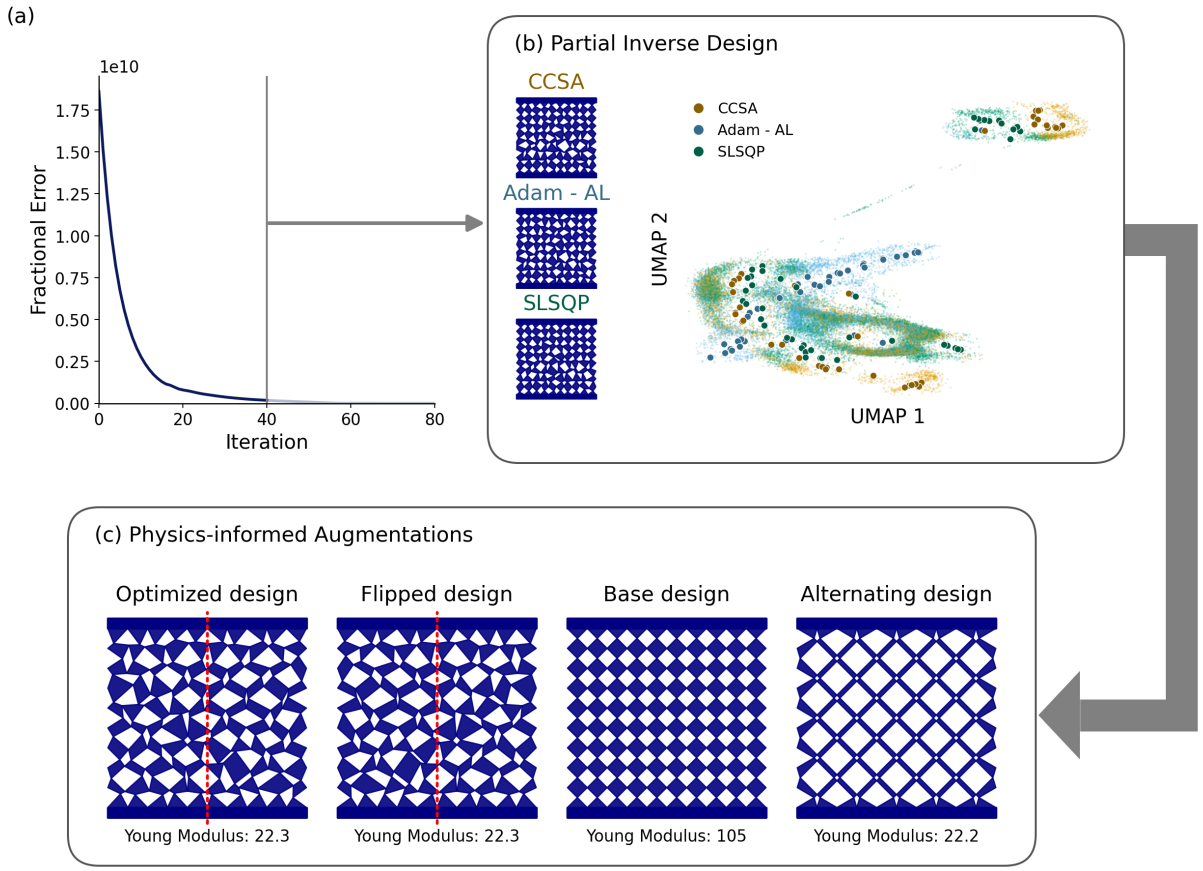


图 5：**TRAMA 的数据集生成流程**（a）数据集创建流程，包括使用一组多样化的优化器进行提前停止的逆向设计，以及物理信息增强。（b）生成的数据集的 UMAP 可视化，按所用优化器着色。较深的点对应原始数据集，较浅且较小的点对应这些设计方案的变体。ADAM-AL 与 CCSA 和 SLSQP 生成的设计方案可以区分开来，而后两者则较难区分。这是合理的，因为它们使用相似的代理模型。（c）不会改变或会降低设计方案杨氏模量的数据增强示例。这使数据集规模增加了 $\times 235$。

计算优势

下面给出一个示例，说明我们的方法能够节省多少计算资源。我们的方法成本分为固定的一次性部分（数据集构建和 VAE 训练）以及每轮优化的部分（在潜在空间中进行贝叶斯优化）。如果没有专门的动力学求解器，我们会使用 ABAQUS 上二维有限元模拟的成本来展示这种方法的计算优势。令 $c_N \approx 14$ 表示评估一次非线性属性（例如应力-应变曲线）的成本，令 $c_L \approx 1.6$ ms 表示评估廉价线性属性（例如杨氏模量）的成本。关于线性化实现的信息，请参见补充说明 1），则 $c_N/c_L \approx 525000$。我们针对杨氏模量优化 $N_0 = 130$ 个设计方案，并在 $K = 40$ 次迭代后停止，而不是进行完全收敛所需的 ~ 90 次迭代。数据集的成本为

$$C_{\text{data}} = N_0\, K\, c_L \approx 130 \cdot 40 \cdot 1.6 \times 10^{-3}\ \text{s} \approx 8\ \text{s}. \tag{3}$$

具有物理意义的数据增强可将数据集扩展至 $N_{\text{aug}} = 3 \times 10^4$ 个设计方案，而无需进行任何额外模拟。若通过优化生成相同的设计方案，使用低成本属性将耗时 $N_{\text{aug}} K c_L \approx 32$ 分钟。GPU加速的VAE训练耗时 $T_{\text{train}} \approx 20$ 分钟，因此我们策略的总固定成本为

$$C_{\text{fixed}} = C_{\text{data}} + T_{\text{train}} \approx 1.2 \times 10^3\ \text{s} \approx 20\ \text{min}, \tag{4}$$

潜在空间中的贝叶斯优化需要对高成本属性进行 $n_{\text{BO}} \approx 220$ 次评估，相比之下，高斯过程的拟合成本可以忽略不计：

$$C_{\text{BO}} = n_{\text{BO}}\, c_N \approx 50\ \text{h}. \tag{5}$$

替代方案包括在原始参数空间中完全收敛的基于梯度的优化，用于能量吸收效率评估的次数为 $n_{\text{GB}} \approx$630次；或者在原始空间中采用无梯度优化（CMA-ES），评估次数为 $n_{\text{GF}} \approx 8000$ 次：

$$C_{\text{GB}} = n_{\text{GB}}\, c_N \approx 147\ \text{h}, \qquad C_{\text{GF}} = n_{\text{GF}}\, c_N \approx 1867\ \text{h}. \tag{6}$$

潜在空间只需构建一次，即可针对每种新属性或目标重复使用，因此 M 使用我们的方法开展活动的成本为 $C_{\text{fixed}} + M\, C_{\text{BO}}$，而基线方法的成本分别为 $M\, C_{\text{GB}}$ 或 $M\, C_{\text{GF}}$ 。在这一现实条件下，我们的方法从第一个设计方案开始就更为便宜，并且可以复用于其他属性。贝叶斯优化循环的运行时间即使延长一倍，我们的方法仍然更便宜。在研究的其余部分，我们使用参考文献 [32] 中开发的可微动力学模拟器，并在补充说明1中进行了描述；对于关于可迁移性的实证结果，这不会造成一般性损失。请注意，2D FEM与所部署动力学模拟器之间的误差，小于2D FEM的仿真与现实差距。

数据可用性

本研究生成的数据集和训练后的模型权重已存储于 Zenodo [75]。同行评审期间，评审人员可以访问这些内容；论文发表后，这些内容将向公众开放。

代码可用性

用于生成我们结果的代码和可复现性说明已提供给审稿人，并将在论文发表后发布于 GitHub代码库 [76]。

参考文献

[1] Shan, S., Kang, S.H., Raney, J.R., Wang, P., Fang, L., Candido, F., Lewis, J.A., Bertoldi, K.：用于捕获弹性应变能的多稳态架构材料。《先进材料》**27**(29), 4296–4301 (2015)[2] Zhang, Y., Restrepo, D., Velay-Lizancos, M., Mankame, N.D., Zavattieri, P.D.：功能性二维相变胞元材料中的能量耗散。《科学报告》**9**(1), 12581 (2019)[3]Portela, C.M., Edwards, B.W., Veysset, D., Sun, Y., Nelson, K.A., Kochmann, D.M., Greer, J.R.：纳米架构碳的超音速冲击韧性。《自然·材料》**20**(11), 1491–1497 (2021)[4] Chen, Y., Jin, L.：利用宽超弹性柱回弹屈曲的可重复使用吸能架构材料。《先进功能材料》**31**(31), 2102113 (2021)[5]Haghpanah, B., Salari-Sharif, L., Pourrajab, P., Hopkins, J., Valdevit, L.：多稳态形状可重构架构材料。《先进材料》**28**(36), 7915–7920 (2016)[6] Florijn, B., Coulais, C., Van Hecke, M.：可编程机械超材料。《物理评论快报》**113**(17), 175503 (2014)[7]Jin, L., Forte, A.E., Deng, B., Rafsanjani, A., Bertoldi, K.：源自剪纸的可编程形状充气结构。《先进材料》**32**(33), 2001863 (2020)[8] Chen, T., Pauly,M.,Reis,P.M.：具有稳定记忆的可重新编程机械超材料。《自然》**589**(7842), 386–390 (2021)[9] Medina, E., Farrell, P.E., Bertoldi, K., Rycroft, C.H.：探索非线性机械超材料的设计空间以实现高级可编程性。《物理评论B》**101**(6), 064101 (2020)[10] Fan, J.A., Yeo, W.-H., Su, Y., Hattori, Y., Lee, W., Jung, S.-Y., Zhang, Y., Liu, Z., Cheng, H., Falgout, L., *et al.*：用于可拉伸电子产品的分形设计概念。《自然通讯》**5**(1), 3266 (2014)[11] Ma, Y., Feng, X., Rogers, J.A., Huang, Y., Zhang, Y.：可拉伸电子产品中"J形"应力–应变行为的设计与应用：综述。《芯片实验室》**17**(10), 1689–1704 (2017)[12]Herbert, R., Kim, J.-H., Kim, Y.S., Lee, H.M., Yeo, W.-H.：面向医学、医疗保健和人机界面的软材料赋能柔性混合电子器件。《材料》**11**(2), 187 (2018)[13] Chan, B., Leong, K.：组织工程中的支架构建：一般方法与组织特异性考量。《欧洲脊柱杂志》**17**(增刊4), 467–479 (2008)[14] Pal, A., Restrepo, V., Goswami, D., Martinez, R.V.：利用软体机器人中的机械失稳：控制、传感与驱动。《先进材料》**33**(19), 2006939 (2021)[15]Kumar, S., Tan, S., Zheng, L., Kochmann, D.M.：逆向设计的旋节超材料。《npj计算材料》**6**(1), 73 (2020)[16]Gu, G.X., Buehler, M.J.：通过控制多晶的织构来调节力学属性。

采用基于伴随法的梯度优化进行增材制造材料设计：Gx gu，mj buehler。《力学学报》**229**(10)，4033–4044（2018）[17] Zeng, Q., Duan, S., Zhao, Z., Wang, P., Lei, H.：通过拓扑优化对吸能超材料进行逆向设计。《先进科学》**10**(4)，2204977（2023）[18] Rodriguez Gomez, S.E., Das, R., Calius, E.P.：具有极端体积模量与剪切模量比的异质机械超材料：一种进化设计方法。载于：《材料力学行为国际会议》，第23–33页（2023）。Springer[19] Zol´esio, J.：形状优化问题与自由边界问题导论。《形状优化与自由边界》（加拿大魁北克省蒙特利尔，1990）**380**，397–457（1992）[20] Sigmund, O., Maute, K.：拓扑优化方法：比较性综述。《结构与多学科优化》**48**(6)，1031–1055（2013）[21] Hammond, A.M., Oskooi, A., Hammond, I.M., Chen, M., Ralph, S.E., Johnson, S.G.：通过子像素平滑投影统一并加速水平集与基于密度的拓扑优化。《光学快报》**33**(16)，33620–33642（2025）[22] Beyer, H.-G., Schwefel, H.-P.：进化策略——全面介绍。《自然计算》**1**(1)，3–52（2002）[23] Hansen, N., M¨uller, S.D., Koumoutsakos, P.：通过协方差矩阵自适应（cma-es）降低去随机化进化策略的时间复杂度。《进化计算》**11**(1)，1–18（2003）[24] Rodriguez Gomez, S.：以设计为核心的力学：多尺度机械超材料的进化设计优化。博士论文，RMIT大学（2025）[25] Hastie, T., Tibshirani, R., Friedman, J.H., Friedman, J.H.：《统计学习要素：数据挖掘、推断与预测》第2卷。Springer，???（2009）[26] Zheng, X., Zhang, X., Chen, T.-T., Watanabe, I.：机械超材料中的深度学习：从预测和生成到逆向设计。《先进材料》**35**(45)，2302530（2023）[27] Zong, Z., Wen, P., Chai, Z., Wu, D., Ding, H., Wu, Z.：深度学习辅助的机械超材料设计。《先进智能发现》**2**(4)，202500084（2026年）[28] Sun, X., Yue, L., Yu, L., Shao, H., Peng, X., Zhou, K., Demoly, F., Zhao, R., Qi, H.J.：机器学习—进化算法驱动的4D打印主动复合结构设计。《先进功能材料》**32**(10)，2109805（2022）[29] Gu, G.X., Chen, C.-T., Richmond, D.J., Buehler, M.J.：利用机器学习进行仿生分级复合材料设计：仿真、增材制造与实验。《材料视界》**5**(5)，939–945（2018）[30] Ha, C.S., Yao, D., Xu, Z., Liu, C., Liu, H., Elkins, D., Kile, M., Deshpande, V., Kong, Z., Bauchy, M.,

习。《自然通讯》**14**(1), 5765 (2023)[31] Liang, K., Wei, Z., Zhang, X.: 基于机器学习的变刚度机械超材料逆向设计。《国际机械科学杂志》，111265 (2026)[32] Deng, B., Zareei, A., Ding, X., Weaver, J.C., Rycroft, C.H., Bertoldi, K.: 通过神经网络加速进化策略实现具有目标非线性响应的机械超材料逆向设计。《先进材料》**34**(41), 2206238 (2022)[33] Yang, Z., Yu, C.-H., Buehler, M.J.: 用于预测分级复合材料中复杂应力应变场的深度学习模型。《科学进展》**7**(15), 7416 (2021)[34] Zhang, Z., Gu, G.X.: 面向数字材料的物理信息深度学习。《理论与应用力学快报》**11**(1), 100220 (2021)[35] Zhang, Z., Jin, Z., Gu, G.X.: 利用物理与数据驱动混合框架进行高效气动驱动建模。《细胞报告物理科学》**3**(4) (2022)[36] Wang,L., Chan, Y.-C., Ahmed, F., Liu, Z., Zhu, P., Chen, W.: 用于超材料系统机理学习与设计的深度生成建模。《计算机方法：应用力学与工程》**372**, 113377 (2020)[37] Zhang, X.-c., Song, Z.-y., Li, Y.-n., Xiao, L.-j., Xu, Z., Rao, L.-x., Ci, T.-j., Hui, X.-L.: 具有定制化应力应变响应的超材料生成式逆向设计。《国际机械科学杂志》，110875 (2025)[38] Pestourie, R.: 快速近似求解器：电磁学中的超材料设计。《IEEE天线与传播杂志》(2025)[39] Raissi, M., Perdikaris, P., Karniadakis, G.E.: 物理信息神经网络：一种用于求解涉及非线性偏微分方程的正向和逆向问题的深度学习框架。计算物理学杂志**378**, 686–707 (2019)[40] McBane, S., Choi, Y.: 分量式降阶模型晶格型结构设计。《计算机方法：应用力学与工程》**381**, 113813 (2021)[41] Zanardi, I., Diaz, A.N., Chung, S.W., Panesi, M., Choi, Y.: 面向动力系统的可扩展非线性流形降阶模型。arXiv预印本 arXiv:2412.00507 (2024)[42] Li, Z., Kovachki, N., Azizzadenesheli, K., Liu, B., Bhattacharya, K., Stuart, A., Anandkumar, A.: 用于参数化偏微分方程的傅里叶神经算子。arXiv预印本 arXiv:2010.08895 (2020)[43] Lu, L., Jin, P., Pang, G., Zhang, Z., Karniadakis, G.E.: 基于算子通用逼近定理，通过DeepONet学习非线性算子。《自然·机器智能》**3**(3), 218–229 (2021)[44] Kovachki, N., Li, Z., Liu, B., Azizzadenesheli, K., Bhattacharya, K., Stuart, A., Anandkumar, A.: 神经算子：学习函数空间之间的映射及其在偏微分方程中的应用。机器学习杂志

研究 **24**(89)，1–97（2023）

[45]Takeishi, N., Kalousis, A.: 用于稳健且可解释生成式建模的物理集成变分自编码器。《神经信息处理系统进展》**34**, 14809–14821 (2021)[46] Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., He, Q.: 迁移学习综合综述。IEEE学报**109**(1), 43–76 (2020)[47]Li, S., Liu, N., Becton, M., Zeng, X., Wang, X.: 利用迁移学习预测二维架构化胞状结构的力学性能。《微观力学与分子物理学杂志》**8**(01), 33–43 (2023)[48] Zhang, Z., Liu, Q., Wu, D.: 利用迁移学习预测应力–应变曲线：聚合物复合材料间的知识迁移。《材料与设计》**218**, 110700 (2022)[49] Apaza, G., Selva, D.: 超材料多目标生成式设计中的迁移学习。载于：国际设计工程技术会议暨计算机与信息工程会议，卷88360，第03–03006页 (2024)。美国机械工程师学会[50] Liu, X., Tang, Y., Zhao, H., Zhu, Y., Chen, C., Li, Y.: 通过迁移学习实现快速性能预测，推动可扩展且数据高效的超材料设计。《工程研究成果》，109362 (2026)[51] You, P., Chen, H., Bahmani, B., Espinosa, H.D.: 用于旋节线超材料力学分析的多保真贝叶斯神经算子。《npj计算材料》(2026)[52] Lu, L., Pestourie, R.,Johnson, S.G., Romano, G.: 多保真度深度神经算子：高效学习偏微分方程，并应用于纳米尺度热传输的快速逆向设计。《物理评论研究》**4**(2), 023210 (2022)[53] Finn, C., Abbeel, P., Levine, S.: 用于深度网络快速适应的模型无关元学习。载于：国际机器学习会议，第1126–1135页 (2017)。PMLR[54] Nichol, A., Achiam, J., Schulman, J.: 关于一阶元学习算法。arXiv预印本 arXiv:1803.02999 (2018)[55] Varagnolo, A., Pimonova, Y., Taylor, M.G., Pestourie, R., Lubbers, N.E.: 面向多目标化学搜索的可解释元学习。arXiv预印本 arXiv:2606.20497 (2026)[56] Ma, W., Cheng, F., Xu, Y., Wen, Q., Liu, Y.: 基于深度生成模型与半监督学习策略的超材料概率表示与逆向设计。《先进材料》**31**(35), 1901111 (2019)[57] Marzban, R., Abiri, H., Pestourie, R., Adibi, A.: Hilab：一种混合逆向设计框架。《小型方法》**9**(11), 00975 (2025)[58] Marzban, R., Adibi, A., Pestourie, R.: 通过表示学习实现纳米光子学逆向设计。《先进光学材料》**14**(1), 02062 (2026)

[59] Huang, X., Zhang, G., Lu, L., Pestourie, R.：神经算子支持的拓扑信息引导进化策略，用于偏微分方程约束优化。IEEE进化计算汇刊（2026年）[60]Ye, Z., Liu, X., Peng, B., Kan, C.：使用点云表示学习预测增材制造机械超材料的力学行为。工程计算与信息科学杂志 **24(6)**, 060901（2024年）[61] Shaikeea, A.J.D., Cui, H., O'Masta, M., Zheng, X.R., Deshpande, V.S.：机械超材料的韧性。《自然·材料》 **21(3)**, 297–304（2022年）[62] Conn, A.R., Gould, N.I.M., Toint, P.：一种用于带一般约束和简单边界的优化问题的全局收敛增广拉格朗日算法。SIAM数值分析杂志 **28**，545–572（1991年）https://doi.org/10.1137/0728030[63] Svanberg, K.：一种基于保守凸可分离近似的全局收敛优化方法。SIAM优化杂志 **12**，555–573（2002年）https://doi.org/10.1137/s1052623499362822[64] Kraft, D.：算法733：用于最优控制计算的TOMP–Fortran模块。ACM数学软件汇刊 **20**，262–281（1994年）https://doi.org/10.1145/192115.192124[65] Johnson, S.G.：NLopt非线性优化软件包。https://github.com/stevengj/nlopt （2007年）[66] Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S. 等：一张图像胜过16×16个词：用于大规模图像识别的变换器。国际学习表征会议（2021年）[67]Williams, C.K., Rasmussen, C.E.：用于机器学习的高斯过程，第2卷。麻省理工学院出版社，剑桥，马萨诸塞州，???（2006年）[68] Srinivas, N., Krause, A., Kakade, S.M., Seeger, M.：赌博机设定下的高斯过程优化：无遗憾与实验设计。arXiv预印本arXiv:0912.3995（2009年）[69] Gongora, A.E., Snapp, K.L., Pang, R., Tiano, T.M., Reyes, K.G., Whiting, E., Lawton, T.J., Morgan, E.F., Brown, K.A.：利用迁移学习设计用于冲击防护的晶格。Matter **5**（9），2829–2846（2022年）[70] Chakraborty, S.：基于迁移学习的多保真度物理信息深度神经网络。计算物理学杂志 **426**，109942（2021年）[71] Pestourie, R., Mroueh, Y., Rackauckas, C., Das, P., Johnson, S.G.：用于偏微分方程的物理增强深度代理模型。《自然·机器智能》**5**（12），1458–1465（2023年）[72] Varagnolo, A., Romano, G., Pestourie, R.：用于声子玻尔兹曼输运方程的物理增强深度代理模型。ASME传热传质杂志 **148**（8），082501（2026年）[73] Yosinski, J., Clune, J., Bengio, Y., Lipson, H.：深度神经网络中的特征具有多强的可迁移性？神经信息处理系统进展 **27**（2014年）

[74] Kornblith, S., Shlens, J., Le, Q.V.：更好的ImageNet模型是否具有更好的迁移效果？载于：IEEE/CVF计算机视觉与模式识别会议论文集，第2661–2671页（2019年）[75]Varagnolo, A., Dai, M., Deng, B., Pestourie, R.：基于可迁移表征学习的机械超材料高效优化数据。https://doi.org/10.5281/zenodo.24387483[76] Varagnolo, A., Dai, M., Deng, B., Pestourie, R.：基于可迁移表征学习的机械超材料高效优化代码仓库，年份 = 2026，发布形式 = https://github.com/tonioeltopoquegira/TRAMA/tree/release，备注 =GitHub代码库

资助

作者声明，本研究获得美国国家科学基金会的支持，[资助项目编号为 IIS-2435905]。

作者贡献

A.V. 和 M.D. 对本研究的贡献相同。M.D. 编写了线性化杨氏模量求解器的代码。A.V. 编写了机器学习模型、逆向设计实验和可迁移性分析的代码。M.D. 对制造设计进行了实验。B.D. 和 R.P. 提出了并监督该项目，并获得了经费。A.V. 和 M.D. 撰写了初稿；所有作者均审阅并编辑了稿件。

利益冲突

作者声明不存在利益冲突。

补充信息

补充说明 1：机械超材料仿真与实验平台

设计空间

该超材料由一个 $N_x \times N_y$ 四边形阵列组成。在此， a 表示未扰动正方形的对角线长度（补充图1(a)）。行和列分别从下到上、从左到右编号。对于每个四边形 $[i, j]$，局部索引 $p = 1$、2、3、4分别表示其右侧、上方、左侧和下方的顶点。

每个顶点都因扰动向量 $\mathbf{d}_p^{[i,j]} = [d_{p,x}^{[i,j]}, d_{p,y}^{[i,j]}]^\mathsf{T}$ 而偏离其原始位置。从未受扰动正方形的中心指向其四个受扰动顶点的位置向量为

$$\begin{aligned} \mathbf{r}_1^{[i,j]} &= \begin{bmatrix} a/2 \\ 0 \end{bmatrix} + \mathbf{d}_1^{[i,j]}, \qquad \mathbf{r}_2^{[i,j]} = \begin{bmatrix} 0 \\ a/2 \end{bmatrix} + \mathbf{d}_2^{[i,j]}, \\ \mathbf{r}_3^{[i,j]} &= \begin{bmatrix} -a/2 \\ 0 \end{bmatrix} + \mathbf{d}_3^{[i,j]}, \quad \mathbf{r}_4^{[i,j]} = \begin{bmatrix} 0 \\ -a/2 \end{bmatrix} + \mathbf{d}_4^{[i,j]}. \end{aligned} \tag{1.1}$$

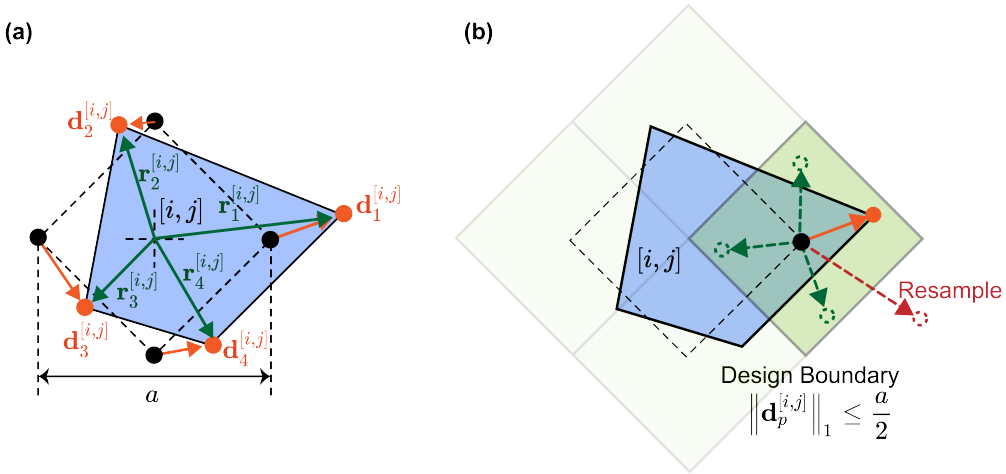
由于每个 顶点由两个相邻四边形共用，相应的扰动向量 rs 满足

$$\mathbf{d}_1^{[i,j]} = \mathbf{d}_3^{[i,j+1]}, \qquad \mathbf{d}_2^{[i,j]} = \mathbf{d}_4^{[i+1,j]}. \tag{1.2}$$

每个独立扰动向量都受到菱形设计区域的约束

$$\left\| \mathbf{d}_p^{[i,j]} \right\|_1 \leq \frac{a}{2}. \tag{1.3}$$

此区域之外的扰动将被拒绝并重新采样（补充图 1 (b)）。



补充图1：几何设计空间。(a) 四边形顶点的扰动向量$\mathbf{d}_p^{[i,j]}$和位置向量$\mathbf{r}_p^{[i,j]}$。(b) 每个扰动向量对应的菱形设计区域。

制造与实验

制造

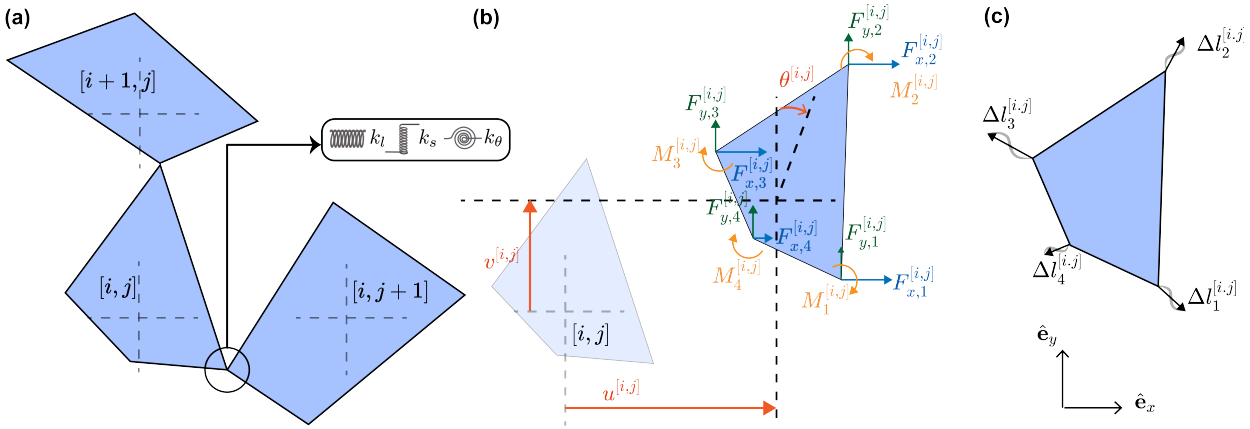
物理样品基于由未扰动正方形构成的 10×10 阵列，其对角线长度为$a = 10$ 毫米。未扰动结构的平面尺寸为 100 毫米 $\times$ 100 毫米，面外厚度为 20 毫米。连接韧带的面内厚度为 1 毫米。无序设计通过根据采样得到的扰动向量移动顶点而生成。

样品采用TEPUTM30A由Inkbit进行3D打印，该材料是一种软弹性体，邵氏A硬度为32，断裂伸长率为200%。选择这种材料是为了适应连接杆在压缩过程中的大幅局部变形。

压缩测试

压缩测试使用 Instron 万能试验机（68SC-1型，Instron，美国）进行。样品在位移控制下，以 3 mm/s 的恒定横梁速度压缩至 -10% 的压缩应变。根据顶行与底行四边形中心之间 90 mm 的初始距离，这对应于 9 mm 的位移。每个样品连续进行三次压缩循环。第一次循环用于机械预处理，不纳入分析。对随后两次循环的力-位移响应取平均值进行分析。

离散模型



补充图 2：离散力学模型。(a) 由韧带连接的刚性四边形，每条韧带均用纵向、剪切和扭转弹簧表示。(b) 四边形的面内平移和旋转，以及作用于其上的力和力矩。(c) 与四边形 $[i,j]$ 相连的四条连接杆的弹簧位移向量 $\Delta\mathbf{l}_p^{[i,j]}$。

四边形被视为刚体，并通过连接杆连接其顶点。每个四边形$[i,j]$ 具有三个自由度，汇总于自由度向量中

$$\mathbf{q}^{[i,j]} = \begin{bmatrix} u^{[i,j]} \\ v^{[i,j]} \\ \theta^{[i,j]} \end{bmatrix}, \tag{1.4}$$

其中， $u^{[i,j]}$ 和 $v^{[i,j]}$ 分别表示沿 x方向和 y方向的平移， $\theta^{[i,j]}$ 表示面内旋转角，并规定顺时针旋转为正（补充图2 （b））。采用“设计空间”部分中定义的顶点编号，每个连接处的相对旋转定义为四边形 $[i,j]$ 的旋转减去其相邻四边形的旋转：

$$\begin{bmatrix} \Delta\theta_1^{[i,j]} \\ \Delta\theta_2^{[i,j]} \\ \Delta\theta_3^{[i,j]} \\ \Delta\theta_4^{[i,j]} \end{bmatrix} = \begin{bmatrix} \theta^{[i,j]} - \theta^{[i,j+1]} \\ \theta^{[i,j]} - \theta^{[i+1,j]} \\ \theta^{[i,j]} - \theta^{[i,j-1]} \\ \theta^{[i,j]} - \theta^{[i-1,j]} \end{bmatrix}. \tag{1.5}$$

如补充图2 （a）所示，每根连接杆由一个刚度为 k_l的纵向弹簧、一个刚度为 k_s的剪切弹簧以及一个非线性扭转弹簧表示，后者施加的力矩为

$$M_{\theta,p}^{[i,j]} = -k_\theta \left[\Delta\theta_p^{[i,j]} + \gamma \left(\Delta\theta_p^{[i,j]} \right)^3 \right], \tag{1.6}$$

其中， k_θ 为扭转刚度， γ 为控制非线性旋转响应的无量纲参数。为简便起见，假设纵向弹簧和剪切弹簧始终分别平行于全局x轴或 y轴。所有连接杆均具有相同的刚度参数。

相邻四边形之间的接触采用额外的旋转弹簧进行建模。在每个连接处 p ，四边形 $[i,j]$的两对边可能发生接触。它们的角间隙分别记为 $\beta^{[i,j]}{}_{\mathrm{I}\,p}$ 和 $\beta^{[i,j]}{}_{\mathrm{II}\,p}$ （补充图 3）。随着相邻四边形相对旋转，一个间隙增大，而另一个间隙减小。当任一角间隙低于阈值 β_c时，接触被激活。

接触刚度定义为

$$K_{\mathrm{cont}}(\beta) = \begin{cases} k_{\mathrm{cont}}, & \beta < \beta_c, \\ 0, & \beta \geq \beta_c, \end{cases} \tag{1.7}$$

其中， k_{cont} 为激活状态下的接触刚度， β_c 为激活阈值。作用于顶点 p 处四边形 $[i,j]$ 上的接触力矩为

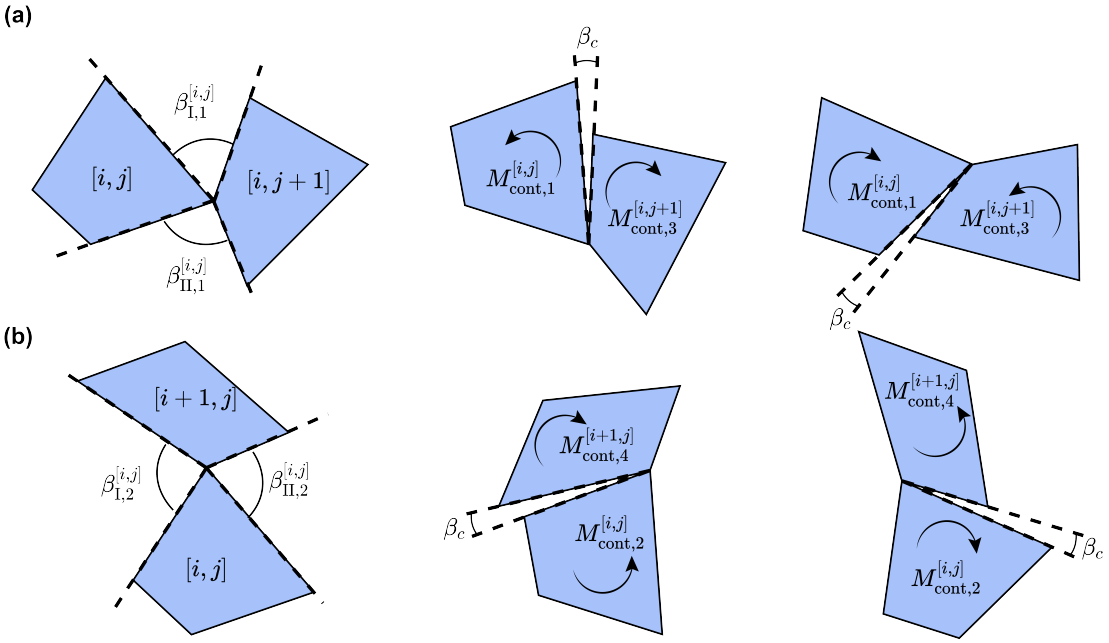
$$M_{\mathrm{cont},p}^{[i,j]} = \left(\beta_{\mathrm{I},p}^{[i,j]} - \beta_c \right) K_{\mathrm{cont}} \left(\beta_{\mathrm{I},p}^{[i,j]} \right) - \left(\beta_{\mathrm{II},p}^{[i,j]} - \beta_c \right) K_{\mathrm{cont}} \left(\beta_{\mathrm{II},p}^{[i,j]} \right). \tag{1.8}$$

根据设计空间部分定义的顶点编号，水平韧带连接四边形 $[i,j]$ 的顶点 $p=1$ 与四边形 $[i,j+1]$的顶点 $p=3$，而垂直韧带连接四边形 $[i,j]$ 的顶点 $p=2$ 与四边形 $[i+1,j]$ 的顶点 $p=4$（补充图 3）。作用于两个相邻四边形上的接触力矩大小相等、方向相反：

$$\begin{aligned} M_{\mathrm{cont},1}^{[i,j]} &= -M_{\mathrm{cont},3}^{[i,j+1]}, \\ M_{\mathrm{cont},2}^{[i,j]} &= -M_{\mathrm{cont},4}^{[i+1,j]}. \end{aligned} \tag{1.9}$$

每个四边形（补充图 2（b））的运动方程为

$$\begin{aligned} m\ddot{u}^{[i,j]} &= \sum_{p=1}^{4} F_{x,p}^{[i,j]}, \\ m\ddot{v}^{[i,j]} &= \sum_{p=1}^{4} F_{y,p}^{[i,j]}, \\ J\ddot{\theta}^{[i,j]} &= \sum_{p=1}^{4} M_p^{[i,j]}. \end{aligned} \tag{1.10}$$



补充图 3：水平相邻四边形 (a) 和垂直相邻四边形 (b) 的接触模型。当任一角间隙小于 β_c 时，接触弹簧会在相邻四边形上产生大小相等、方向相反的力矩。

其中， m 和 J 分别为每个四边形的质量和转动惯量。作用于与其第 p 个顶点相关联的四边形 $[i,j]$ 上的力分量和总力矩为

$$\begin{aligned} F_{x,p}^{[i,j]} &= \left(\mathbf{k}_p \Delta \mathbf{l}_p^{[i,j]}\right) \cdot \hat{\mathbf{e}}_x, \\ F_{y,p}^{[i,j]} &= \left(\mathbf{k}_p \Delta \mathbf{l}_p^{[i,j]}\right) \cdot \hat{\mathbf{e}}_y, \\ M_p^{[i,j]} &= -\left[\mathbf{R}\left(\theta^{[i,j]}\right) \mathbf{r}_p^{[i,j]} \times \left(\mathbf{k}_p \Delta \mathbf{l}_p^{[i,j]}\right)\right] \cdot \hat{\mathbf{e}}_z + M_{\theta,p}^{[i,j]} + M_{\mathrm{cont},p}^{[i,j]}. \end{aligned} \tag{1.11}$$

其中，$\hat{\mathbf{e}}_x$、$\hat{\mathbf{e}}_y$ 和$\hat{\mathbf{e}}_z$ 为笛卡尔单位向量，其中$\hat{\mathbf{e}}_z$ 指向平面外。刚度矩阵$\mathbf{k}_p$ 定义为

$$\mathbf{k}_p = \begin{cases} \begin{bmatrix} k_l & 0 \\ 0 & k_s \end{bmatrix}, & p = 1,3, \\ \begin{bmatrix} k_s & 0 \\ 0 & k_l \end{bmatrix}, & p = 2,4. \end{cases} \tag{1.12}$$

T描述每个四边形面内旋转的旋转矩阵$\mathbf{R}(\theta)$定义为

$$\mathbf{R}(\theta)=\begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}. \tag{1.13}$$

弹簧位移向量 $\Delta\mathbf{l}_p^{[i,j]}$描述了由连接杆 p （补充图2（c））连接的两个顶点之间的相对位移。根据设计空间部分中定义的位置向量，四个连接的位移向量为

$$\begin{aligned}
\Delta\mathbf{l}_1^{[i,j]} &= \begin{bmatrix}a\\0\end{bmatrix}+\begin{bmatrix}u^{[i,j+1]}-u^{[i,j]}\\ v^{[i,j+1]}-v^{[i,j]}\end{bmatrix}+\mathbf{R}\left(\theta^{[i,j+1]}\right)\mathbf{r}_3^{[i,j+1]}-\mathbf{R}\left(\theta^{[i,j]}\right)\mathbf{r}_1^{[i,j]},\\
\Delta\mathbf{l}_2^{[i,j]} &= \begin{bmatrix}0\\a\end{bmatrix}+\begin{bmatrix}u^{[i+1,j]}-u^{[i,j]}\\ v^{[i+1,j]}-v^{[i,j]}\end{bmatrix}+\mathbf{R}\left(\theta^{[i+1,j]}\right)\mathbf{r}_4^{[i+1,j]}-\mathbf{R}\left(\theta^{[i,j]}\right)\mathbf{r}_2^{[i,j]},\\
\Delta\mathbf{l}_3^{[i,j]} &= \begin{bmatrix}-a\\0\end{bmatrix}+\begin{bmatrix}u^{[i,j-1]}-u^{[i,j]}\\ v^{[i,j-1]}-v^{[i,j]}\end{bmatrix}+\mathbf{R}\left(\theta^{[i,j-1]}\right)\mathbf{r}_1^{[i,j-1]}-\mathbf{R}\left(\theta^{[i,j]}\right)\mathbf{r}_3^{[i,j]},\\
\Delta\mathbf{l}_4^{[i,j]} &= \begin{bmatrix}0\\-a\end{bmatrix}+\begin{bmatrix}u^{[i-1,j]}-u^{[i,j]}\\ v^{[i-1,j]}-v^{[i,j]}\end{bmatrix}+\mathbf{R}\left(\theta^{[i-1,j]}\right)\mathbf{r}_2^{[i-1,j]}-\mathbf{R}\left(\theta^{[i,j]}\right)\mathbf{r}_4^{[i,j]}.
\end{aligned} \tag{1.14}$$

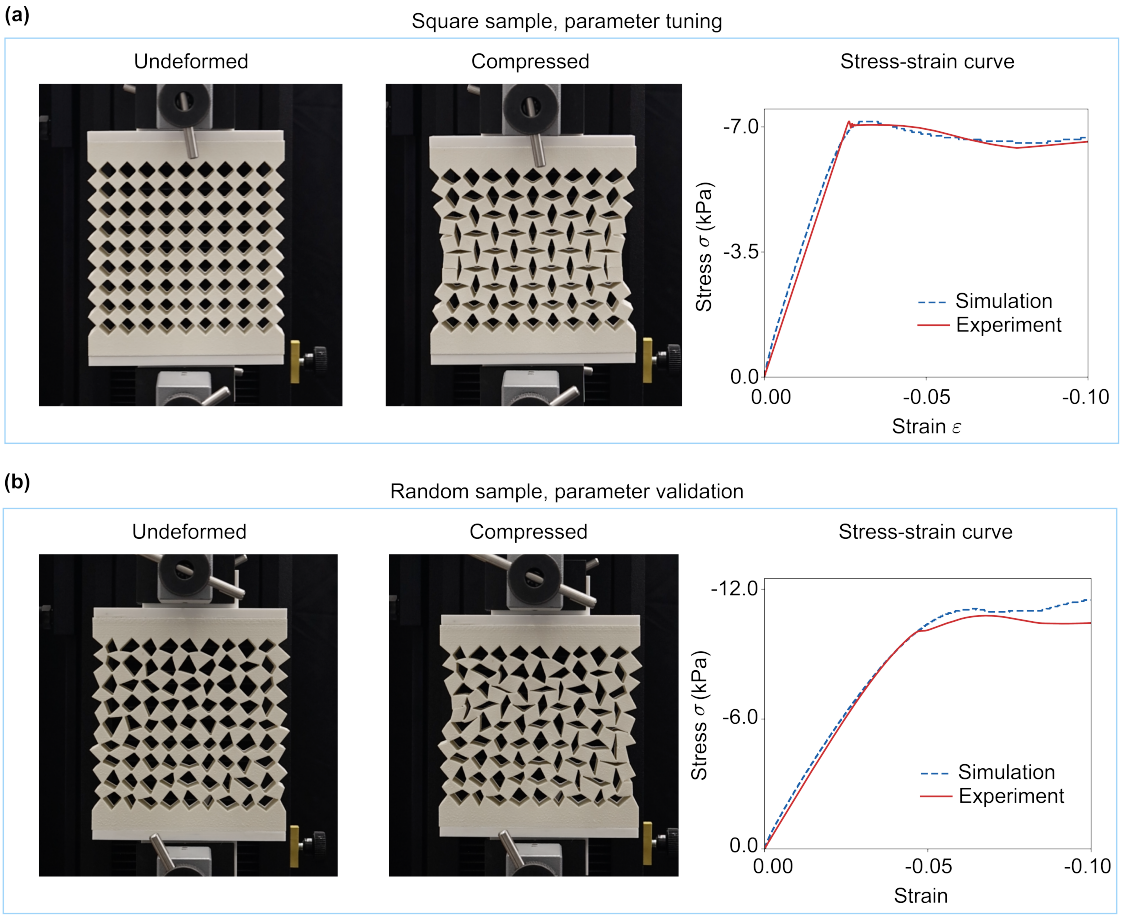
对于边界处的四边形，省略缺失相邻单元的贡献。

实验校准

离散模型根据未扰动 10×10正方形阵列的实验压缩响应进行校准，其对角线长度为 $a = 10$ 毫米（补充图4（a））。校准考虑了初始刚度、屈曲起始以及随后的屈曲后响应，使模型能够捕捉初始弹性阶段和非线性变形行为。模拟所用的参数集为

$$k_l = 5.5\ \mathrm{N/mm}, \qquad k_s = 1.65\ \mathrm{N/mm}, \qquad k_\theta = 1.375\ \mathrm{N\,mm/rad},$$
$$k_{\mathrm{cont}} = 5.5\ \mathrm{N\,mm/rad}, \qquad \gamma = -0.3.$$

为了检验校准后的参数能否预测另一种结构的响应，我们制作并测试了一个随机生成样品。随后，我们使用相同的参数模拟其压缩响应，未作进一步调整。预测的应力–应变曲线与实验结果高度吻合，仅在较大压缩应变下存在轻微差异（补充图4（b））。



补充图4：实验校准与验证。（a）用于校准模型的方形样品。（b）使用相同参数模拟、未经进一步调整的随机样品。每一行展示未变形样品和压缩样品，以及实验测得的（红色实线）和模拟得到的（蓝色虚线）应力–应变曲线。

线性化杨氏模量

有效的 超材料的杨氏模量根据其初始线性响应计算得出
离散模型。 使用为每个四边形定义的自由度向量 $\mathbf{q}^{[i,j]}$，见式（1.4），全局自由度向量

$\mathbf{Q}$ 由所有四边形的向量逐行堆叠而成，如下所示

$$\mathbf{Q} = \begin{bmatrix} \mathbf{q}^{[1,1]} \\ \vdots \\ \mathbf{q}^{[1,N_x]} \\ \mathbf{q}^{[2,1]} \\ \vdots \\ \mathbf{q}^{[2,N_x]} \\ \vdots \\ \mathbf{q}^{[N_y,1]} \\ \vdots \\ \mathbf{q}^{[N_y,N_x]} \end{bmatrix} \in \mathbb{R}^{3N_xN_y}. \tag{1.15}$$

同样，作用于每个四边形的内力和力矩按相同顺序组装，以形成全局内力向量

$$\mathbf{f}_{\text{int}}(\mathbf{Q}) = \begin{bmatrix} \sum_{p=1}^{4} F_{x,p}^{[1,1]} \\ \sum_{p=1}^{4} F_{y,p}^{[1,1]} \\ \sum_{p=1}^{4} M_{p}^{[1,1]} \\ \vdots \\ \sum_{p=1}^{4} F_{x,p}^{[N_y,N_x]} \\ \sum_{p=1}^{4} F_{y,p}^{[N_y,N_x]} \\ \sum_{p=1}^{4} M_{p}^{[N_y,N_x]} \end{bmatrix} \in \mathbb{R}^{3N_xN_y}. \tag{1.16}$$

为计算 初始线性响应，我们使用一阶泰勒展开对$\mathbf{f}_{\text{int}}(\mathbf{Q})$在 e
未变形状 态 $(\mathbf{Q} = \mathbf{0})$进行线
性化：

$$\mathbf{f}_{\text{int}}(\mathbf{Q}) \approx \mathbf{f}_{\text{int}}(\mathbf{0}) + \left.\frac{\partial \mathbf{f}_{\text{int}}}{\partial \mathbf{Q}}\right|_{\mathbf{Q}=\mathbf{0}} \mathbf{Q}. \tag{1.17}$$

由于未变形状态下无应力，$\mathbf{f}_{\text{int}}(\mathbf{0}) = \mathbf{0}$。初始刚度矩阵定义为$\mathbf{f}_{\text{int}}$ 关于$\mathbf{Q}$的负雅可比矩阵，并在未变形状态下求值：

$$\mathbf{K} = -\left.\frac{\partial \mathbf{f}_{\text{int}}}{\partial \mathbf{Q}}\right|_{\mathbf{Q}=\mathbf{0}}, \tag{1.18}$$

根据 $N_{\mathrm{dof}} = 3N_xN_y$，该矩阵可明确写为

$$\mathbf{K} = -\left[\begin{array}{ccc} \dfrac{\partial f_{\mathrm{int},1}}{\partial Q_1} & \cdots & \dfrac{\partial f_{\mathrm{int},1}}{\partial Q_{N_{\mathrm{dof}}}} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_{\mathrm{int},N_{\mathrm{dof}}}}{\partial Q_1} & \cdots & \dfrac{\partial f_{\mathrm{int},N_{\mathrm{dof}}}}{\partial Q_{N_{\mathrm{dof}}}} \end{array}\right]\Bigg|_{\mathbf{Q}=\mathbf{0}}. \tag{1.19}$$

刚度矩阵$\mathbf{K}$在未变形状态（$\mathbf{Q} = \mathbf{0}$）下求得，并在线性化模型中保持不变。假设处于静力平衡状态，外力和力矩与内力和力矩平衡，因此有

$$\mathbf{f}_{\mathrm{ext}} = -\mathbf{f}_{\mathrm{int}}(\mathbf{Q}) \approx \mathbf{K}\mathbf{Q}, \tag{1.20}$$

其中，$\mathbf{f}_{\mathrm{ext}}$ 是外力和力矩组成的全局向量。

我们将$\mathbf{Q}$ 划分为三个块：$\mathbf{Q}_b$、$\mathbf{Q}_f$和$\mathbf{Q}_t$，分别对应底部行、中间行中的四边形以及顶行。刚度矩阵$\mathbf{K}$ 也相应地进行划分。

底部行固定，而顶行的水平位移和旋转均为零。对于应变$\varepsilon < 0$ （压缩），每个顶行四边形的指定自由度向量为

$$\bar{\mathbf{q}}^{[N_y,j]} = \begin{bmatrix} 0 \\ \varepsilon(N_y - 1)a \\ 0 \end{bmatrix}, \qquad j = 1, \ldots, N_x. \tag{1.21}$$

¯将指定自由度向量组装到$\mathbf{Q}_t$ 中，得到

$$\bar{\mathbf{Q}}_t = \begin{bmatrix} \bar{\mathbf{q}}^{[N_y,1]} \\ \vdots \\ \bar{\mathbf{q}}^{[N_y,N_x]} \end{bmatrix} \in \mathbb{R}^{3N_x}. \tag{1.22}$$

由于没有外力或力矩直接作用于中间行的四边形，全局外力向量$\mathbf{f}_{\mathrm{ext}}$ 仅包含底部和顶部边界约束施加的力和力矩，分别记为$\mathbf{f}_b$ 和$\mathbf{f}_t$:

$$\mathbf{f}_{\mathrm{ext}} = \begin{bmatrix} \mathbf{f}_b \\ \mathbf{0} \\ \mathbf{f}_t \end{bmatrix}. \tag{1.23}$$

¯当 $\mathbf{Q}_b = \mathbf{0}$ 和 $\mathbf{Q}_t$ =$\mathbf{Q}_t$时，全局平衡方程（式（1.20））变为

$$\begin{bmatrix} \mathbf{K}_{bb} & \mathbf{K}_{bf} & \mathbf{K}_{bt} \\ \mathbf{K}_{fb} & \mathbf{K}_{ff} & \mathbf{K}_{ft} \\ \mathbf{K}_{tb} & \mathbf{K}_{tf} & \mathbf{K}_{tt} \end{bmatrix} \begin{bmatrix} \mathbf{0} \\ \mathbf{Q}_f \\ \bar{\mathbf{Q}}_t \end{bmatrix} = \begin{bmatrix} \mathbf{f}_b \\ \mathbf{0} \\ \mathbf{f}_t \end{bmatrix}. \tag{1.24}$$

展开三个块行可得

$$\begin{aligned}\mathbf{f}_b &= \mathbf{K}_{bf}\mathbf{Q}_f + \mathbf{K}_{bt}\bar{\mathbf{Q}}_t,\\ \mathbf{K}_{ff}\mathbf{Q}_f &= -\mathbf{K}_{ft}\bar{\mathbf{Q}}_t,\\ \mathbf{f}_t &= \mathbf{K}_{tf}\mathbf{Q}_f + \mathbf{K}_{tt}\bar{\mathbf{Q}}_t.\end{aligned} \tag{1.25}$$

求解方程（1.25）可得

$$\begin{aligned}\mathbf{f}_b &= \left(\mathbf{K}_{bt} - \mathbf{K}_{bf}\mathbf{K}_{ff}^{-1}\mathbf{K}_{ft}\right)\bar{\mathbf{Q}}_t,\\ \mathbf{f}_t &= \left(\mathbf{K}_{tt} - \mathbf{K}_{tf}\mathbf{K}_{ff}^{-1}\mathbf{K}_{ft}\right)\bar{\mathbf{Q}}_t.\end{aligned} \tag{1.26}$$

顶部边界约束所施加的力和力矩汇总如下

$$\mathbf{f}_t = \begin{bmatrix} f_u^{[N_y,1]} \\ f_v^{[N_y,1]} \\ f_\theta^{[N_y,1]} \\ \vdots \\ f_u^{[N_y,N_x]} \\ f_v^{[N_y,N_x]} \\ f_\theta^{[N_y,N_x]} \end{bmatrix} \in \mathbb{R}^{3N_x}, \tag{1.27}$$

其中， $f_u^{[N_y,j]}$、 $f_v^{[N_y,j]}$和 $f_\theta^{[N_y,j]}$分别是作用于四边形$[N_y, j]$上的水平力、竖直力和力矩。

顶行上的总竖直力为

$$F_t = \sum_{j=1}^{N_x} f_v^{[N_y,j]}. \tag{1.28}$$

相应的名义应力为

$$\sigma = \frac{F_t}{A} = \frac{1}{A}\sum_{j=1}^{N_x} f_v^{[N_y,j]}, \tag{1.29}$$

其中， A为样品加载面积。在压缩状态下， F_t 和 σ 均为负值。

在线性化模型中，应力与规定应变成正比。因此，有效杨氏模量为

$$E = \frac{\sigma}{\varepsilon} = \frac{1}{A\varepsilon}\sum_{j=1}^{N_x} f_v^{[N_y,j]}. \tag{1.30}$$

补充说明2：机器学习训练细节

数据集

该数据集由杨氏模量的部分逆向设计生成（目标值涵盖从最小化到最大化，每次运行最多迭代 50 次），采用 CCSA、带增广拉格朗日的 Adam 以及 SLSQP，共得到 134 个设计方案。通过对满足位移约束的设计方案施加镜像对称、均匀平移和交替扰动（单独使用或组合使用，每次最多三种），将样本扩增至 31，638 个（$\times 235$）。我们按源设计方案以 95/5 的比例划分数据集，因此训练集中不会出现任何验证设计方案的变体。

编码器和潜在后验

ViT-base 编码器的图块标记通过学习型注意力池化聚合为一个 768 维特征向量。随后，两个线性层将该向量映射为潜在后验的均值 $\boldsymbol{\mu} \in \mathbb{R}^8$ 和对数方差 $\log \boldsymbol{\sigma}^2 \in \mathbb{R}^8$ 。

解码器

通过重参数化技巧采样得到的潜在向量 $\mathbf{z} \sim \mathcal{N}(\boldsymbol{\mu},\ \mathrm{diag}(\boldsymbol{\sigma}^2))$ 会被输入两个相互独立的多层感知机（MLP），其中一个用于水平参数，另一个用于垂直参数。

训练

ViT-VAE 使用余弦学习率调度训练 20 个周期，学习率从最大值 10^{-4} 降至最小值 5×10^{-6}，批次大小为 64 个图像-参数对。KL散度权重 β 在整个训练过程中从 10^{-4} 线性增加至 10^{-2} 。

迁移分析

我们使用在该特定属性的优化运行过程中找到的所有设计方案来拟合主成分分析。能量实验的主成分分析将使用 TRAMA、CCSA 或 CMA-ES 针对能量效率问题找到的所有能量设计方案进行拟合。

补充说明 3：优化细节

能源效率最大化

我们扫描了停止 容差以及 CCSA 的最大评估次数和 CMA-ES 的最大评估次数。表1和2 展示了这些参数对 能效最大化的影响。

最终，我们为 CCSA 选择了容差 10^{-6}，为 CMA-ES 选择了容差 $n_{\mathrm{eval}} = 1200$ 。TRAMA 使用$\lambda = 10^{-4}$运行。

应力-应变曲线匹配

TRAMA 的运行参数为 $\lambda = 10^{-4}$，容差为 10^{-4}。CMA-ES 的最大迭代次数设为 1500。CCSA 的容差设为 10^{-4}。

补充表 1：在不同停止容差（目标函数的相对变化）下，CCSA 在能源效率任务上的表现。评估次数指前向计算和梯度评估次数。数值为 5 个随机种子的均值 ±标准差。正文中使用的配置以星号标记。

容差	能量效率（%）	评估次数
10^{-3}	82.1 ± 3.5	171 ± 87
10^{-4}	84.3 ± 3.0	432 ± 86
10^{-5}	84.6 ± 2.6	697 ± 35
10^{-6*}	85.3 ± 2.0	631 ± 222

补充表 2：不同评估预算下 CMA-ES 在能源效率任务上的表现。数值为五个随机种子的均值 ± 标准差；所有运行均使用完整预算。正文中使用的配置以星号标记。

评估预算	能量效率（%）
900	82.1 ± 0.6
1200	82.5 ± 0.8
1500	83.0 ± 0.4
2000	83.3 ± 0.5
3000	84.6 ± 0.7
8000*	86.1 ± 0.7