



Words to Matter: *De novo* Architected Materials Design Using Transformer Neural Networks

Zhenze Yang^{1,2} and Markus J. Buehler^{1,3*}

¹Laboratory for Atomistic and Molecular Mechanics (LAMM), Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, United States, ²Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA, United States, ³Center for Computational Science and Engineering, Schwarzman College of Computing, Massachusetts Institute of Technology, Cambridge, MA, United States

Transformer neural networks have become widely used in a variety of AI applications, enabling significant advances in Natural Language Processing (NLP) and computer vision. Here we demonstrate the use of transformer neural networks in the *de novo* design of architected materials using a unique approach based on text input that enables the design to be directed by descriptive text, such as “a regular lattice of steel”. Since transformer neural nets enable the conversion of data from distinct forms into one another, including text into images, such methods have the potential to be used as a natural-language-driven tool to develop complex materials designs. In this study we use the Contrastive Language-Image Pre-Training (CLIP) and VQGAN neural networks in an iterative process to generate images that reflect text prompt driven materials designs. We then use the resulting images to generate three-dimensional models that can be realized using additive manufacturing, resulting in physical samples of these text-based materials. We present several such word-to-matter examples, and analyze 3D printed material specimen through associated additional finite element analysis, especially focused on mechanical properties including mechanism design. As an emerging new field, such language-based design approaches can have profound impact, including the use of transformer neural nets to generate machine code for 3D printing, optimization of processing conditions, and other end-to-end design environments that intersect directly with human language.

Keywords: transformer neural nets, deep learning, mechanics, architected materials, materiomics, *de novo* design, hierarchical, additive manufacturing

INTRODUCTION

Materials design has been an endless Frontier for science and engineering, especially translating insights and geometries from biology into engineering (Buehler, 2010; Qin et al., 2014; Wegst et al., 2015; Palkovic et al., 2016; Buehler and Misra, 2019), and interacting with human design input that often drives conceptualization and ideation. Indeed, *de novo* materials design can be challenging, and complex geometric aspects as evolved in nature are difficult to realize in engineering using rigorous methods, including the interface with human descriptions. With the advent of deep learning, and advanced network architectures such as GANs, transformer/attention models, a new era is beginning that is changing the way we model, characterize, design and manufacture materials (Guo et al., 2021).

OPEN ACCESS

Edited by:

Davide Ruffoni,
University of Liège, Belgium

Reviewed by:

Emilio Turco,
University of Sassari, Italy
Domenico De Tommasi,
Politecnico di Bari, Italy
Marco Agostino Deriu,
Politecnico di Torino, Italy

*Correspondence:

Markus J. Buehler
mbuehler@mit.edu

Specialty section:

This article was submitted to
Mechanics of Materials,
a section of the journal
Frontiers in Materials

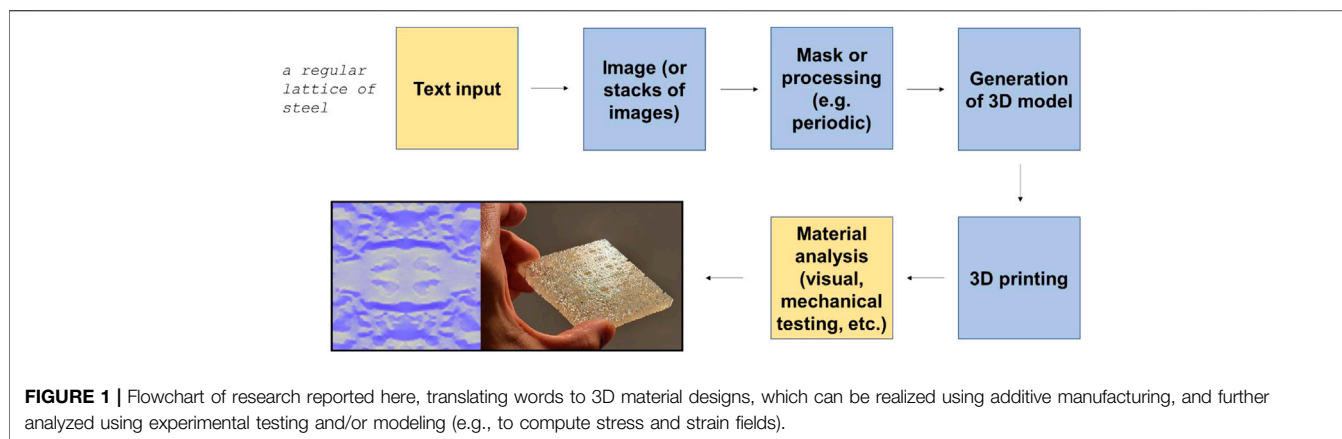
Received: 13 July 2021

Accepted: 20 September 2021

Published: 06 October 2021

Citation:

Yang Z and Buehler MJ (2021) Words
to Matter: *De novo* Architected
Materials Design Using Transformer
Neural Networks.
Front. Mater. 8:740754.
doi: 10.3389/fmats.2021.740754



Indeed, researchers have long sought novel approaches to develop designs that can build on natural evolution and ideation (Cranford Buehler and Markus, 2012). One such method is to use bioinspired design; however, such translation can be challenging. Recently, neural nets have been used to develop novel nature-inspired materials, such as materials developed from music or fire (Giesa et al., 2011; Yu et al., 2019a; Milazzo and Buehler, 2021). In these approaches, neural nets provide a powerful and systematic approach to translate information across manifestations, offering a systematic and consistent approach for such traversions. This can complement conventional approaches, either based on traditional bio-inspired mimicking or category theoretic approaches (Giesa et al., 2011), (Spivak et al., 2011; Brommer et al., 2016; Milazzo et al., 2019). Neural networks have also been used to predict complex stress and strain fields in hierarchical composite materials, enabling the direct association of mechanical fields from geometric or microstructural input (Yang et al., 2021a), (Yang et al., 2021b). However, many if not all of these examples rely on either an algorithmic input or a direct mimicking approach to learn from nature, and is devoid of providing human-readable natural language input into the design process. Such human input remains, to date, largely in the realm of the natural creative process, and interfacing with rigorous computational algorithms is not yet well developed. We provide an advance towards that goal in this paper.

Transformer Neural Nets

Transformer neural networks have become widely used in various AI applications, enabling significant advances in Natural Language Processing (NLP) and vision developments (Guo et al., 2021). In fact, in recent years, inspired by the concept that humans tends to pay greater attention to certain factors when processing information, a mechanism referred to as “attention” has been developed and applied that directs neural networks to focus on important parts of the input data, as opposed to consider all data equally (Bahdanau et al., 2015), (Chaudhari et al., 2019). Based on this idea, Google first built a model known as the “transformer” neural net based on such an attention mechanisms, completely dispensing with recurrence and convolution methods (Vaswani et al., 2017). It has been found that such transformer neural nets can achieve great

success in various NLP settings including language translation. With the capacity of facilitating greater parallelization during training, transformers enable the development of extremely large models including “Bidirectional Encoder Representations from Transformers” (BERT) (Devlin et al., 2019) and “Generative Pre-trained Transformer” (GPT) (Brown et al., 2020). Given these early successes of transformers in NLP, the networks can serve as powerful tools for materials design problems, which under many circumstances, are sequential tasks such as topology optimization, electronic circuit design, or 3D printing toolpath generation (Olivetti et al., 2020), (Liu et al., 2017). We hypothesize that the use of transformers has the potential to parallelize these tasks and make them computationally tractable, which opens the door for many complex engineering challenges that cannot yet be solved today using computational methods.

Apart from NLP applications, recent research showed that attention mechanisms and transformers can also be applied to computer vision (CV), serving as an alternative to convolutional approaches (Wang and Tax, 2016). The well-known pretrained transformers for images include “DEtection TRansformer” (DETR) (Carion et al., 2020) for object detection and “Vision Transformer” (ViT) (Dosovitskiy et al., 2020) for image classification.

In terms of current applications of transformer neural nets in materials research, the “Molecular Transformer” model accurately predicts the outcomes of organic chemical reactions trained with about millions of data point (Schwaller et al., 2019). In addition, Grechishnikova *et al.* considered target-specific drug design as a translational problem and applied transformers to generate *de novo* drugs (Grechishnikova, 2021). However, because transformer models are relatively data greedy and expensive to train, transfer learning has been shown to be a useful tool to build off of trained models. For example, Pesciullesi et al. (2020) used transfer learning to adapt the original Molecular Transformer to predict carbohydrate reactions with a small set of specialized data.

General Concepts in Materials Design

Design of materials has been conducted by humans through eras of civilization, often conceptualizing an idea, then formulating it

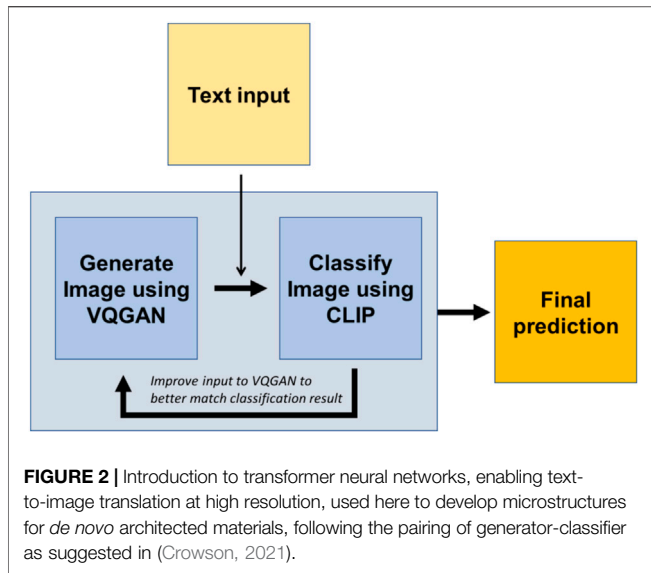


FIGURE 2 | Introduction to transformer neural networks, enabling text-to-image translation at high resolution, used here to develop microstructures for *de novo* architected materials, following the pairing of generator-classifier as suggested in (Crowson, 2021).

either as a drawing or creating prototypes, or initiating the process based on written text or collaborative artistic or artisan work. The key to this human experience of materials design is the intersection of language, and thought, and materialization of these concepts. All material concepts are typically developed through this text-to-material paradigm. However, we have not yet had the opportunity to realize such concepts using conventional computational approaches, even using neural nets. Computational methods force us to formulate our idea into a model, or to express it as a set of numbers, constraints, and ultimately into an algorithmic representation.

Here we propose that transformer neural nets can address this gap and for the first time enable a text-to-material translation, and open a very wide array of opportunities to develop human-AI partnerships, and material design concepts. Some earlier work has used NLP methods in various settings, including the mining of literature to extract optimal processing steps for material design (Jensen et al., 2019). However, this approach did not use direct text input to drive the design process; rather, it mined papers written by humans to extract recipe-type information to create materials.

As transformer neural nets enable the conversion of text into other data formats and representations including images, such methods have the potential to be used as a novel interactive tool to demonstrate the development of materials designs, potentially in the future, relying on voice input and the design of novel computer aided design software. In this work we use the CLIP (Radford et al., 2021) and VQGAN (Esser et al., 2020) neural networks to generate images that reflect materials designs, based on an integration of CLIP and VQGAN (Crowson, 2021) similar to the Big Sleep method (details see Materials and Methods). There are wide ranges of artistic uses of such methods (Crowson, 2021), but they have not been employed in engineering design, nor in the creation of physical objects or materials.

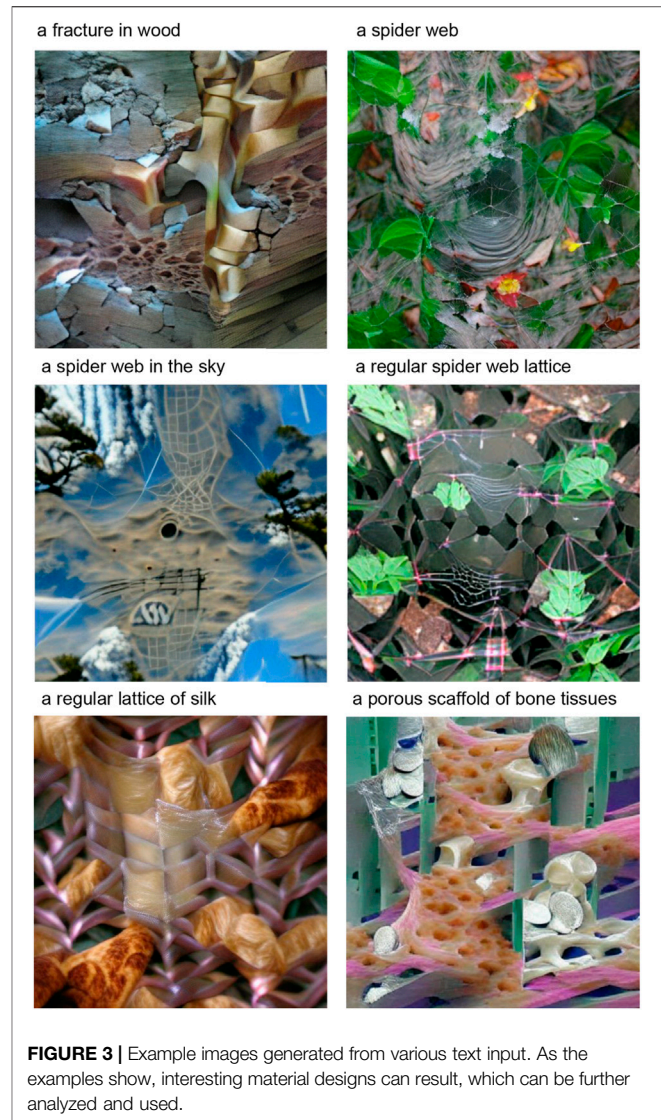


FIGURE 3 | Example images generated from various text input. As the examples show, interesting material designs can result, which can be further analyzed and used.

Outline of This Paper

In this study we demonstrate the use of a transformer neural network in the design of materials. **Figure 1** depicts a flowchart of approach reported here, translating words—human readable descriptive text—towards 3D physical material designs. We proceed with various examples of the approach, describe how we convert the predicted images into 3D models, use additive manufacturing to print them. We then report experimental and computational analysis of mechanical properties to assess the viability of the designs generated.

RESULTS

Figure 2 shows a flowchart that illustrates how transformer neural networks are used to facilitate text-to-image translation at high resolution. This is achieved through the integrated use of the VQGAN transformer neural net, employed here as the

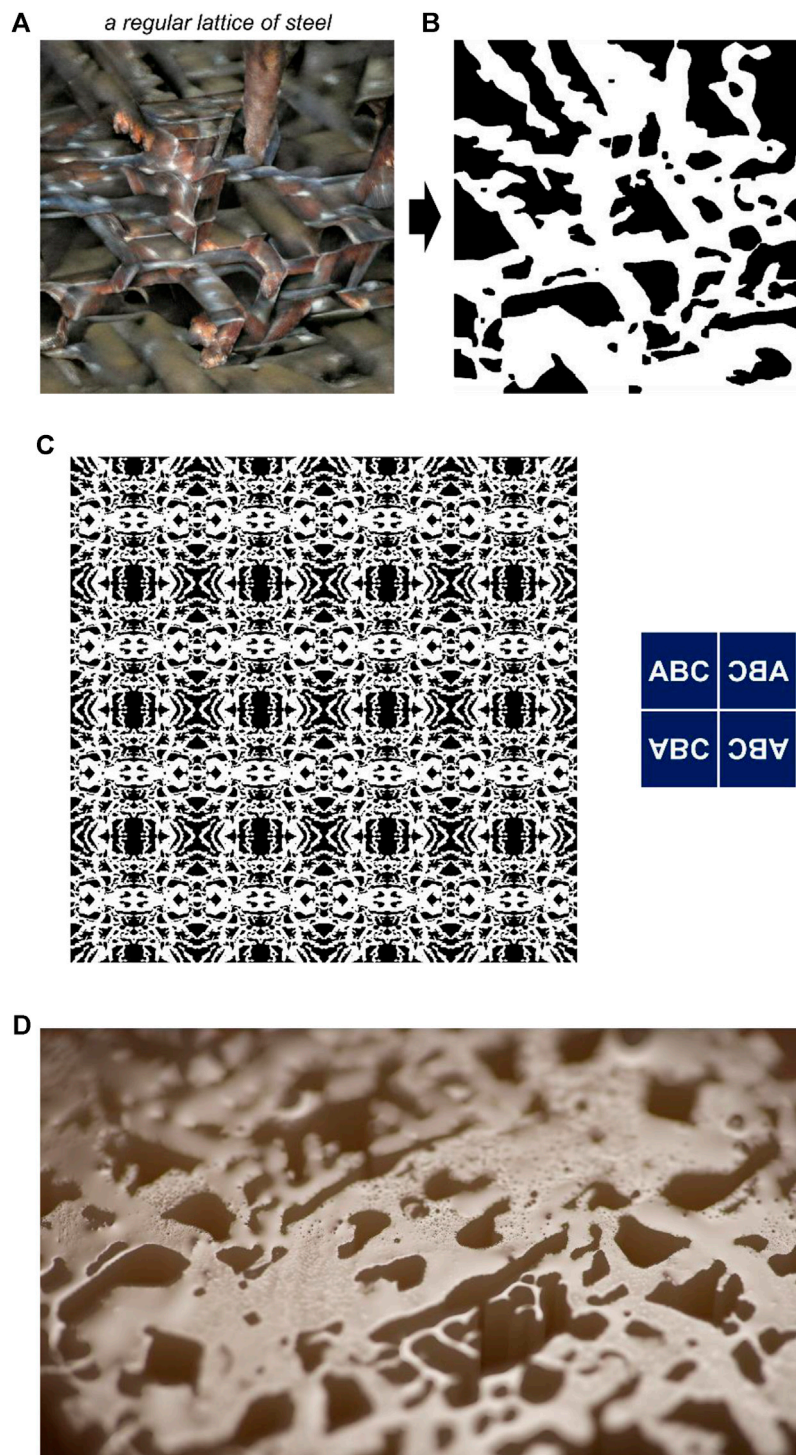
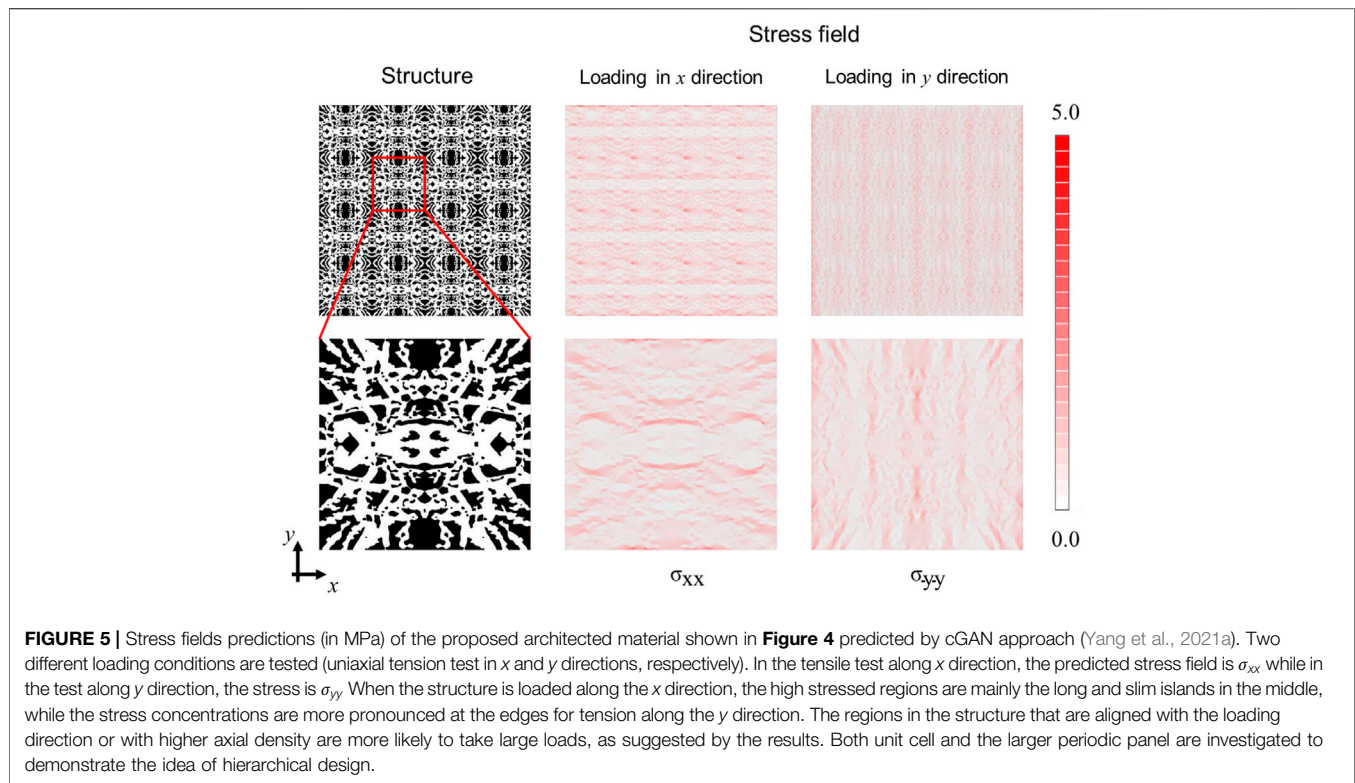


FIGURE 4 | Text-based image generation (A), processing into a black-and-white mask using smoothing operations (B), and generation of periodic material design using geometric operations (first: horizontal flip, then vertical flip of the original and initially flipped design, yielding a periodic structure—see inset on the right with letters “ABC”) (C). Panel (D) shows a close-up view of the 3D printed material. More 3D printed structures are shown in **Figure 6**.

image generator, and CLIP as an image classifier. These two models work in tandem over iterations to generate images that successively match a given text prompt as the solution

converges. Both CLIP and VQGAN models are pretrained on huge datasets, thus providing great capacity of generalization and high robustness. Therefore, we are able



to use those models directly to translate texts to images for materials design.

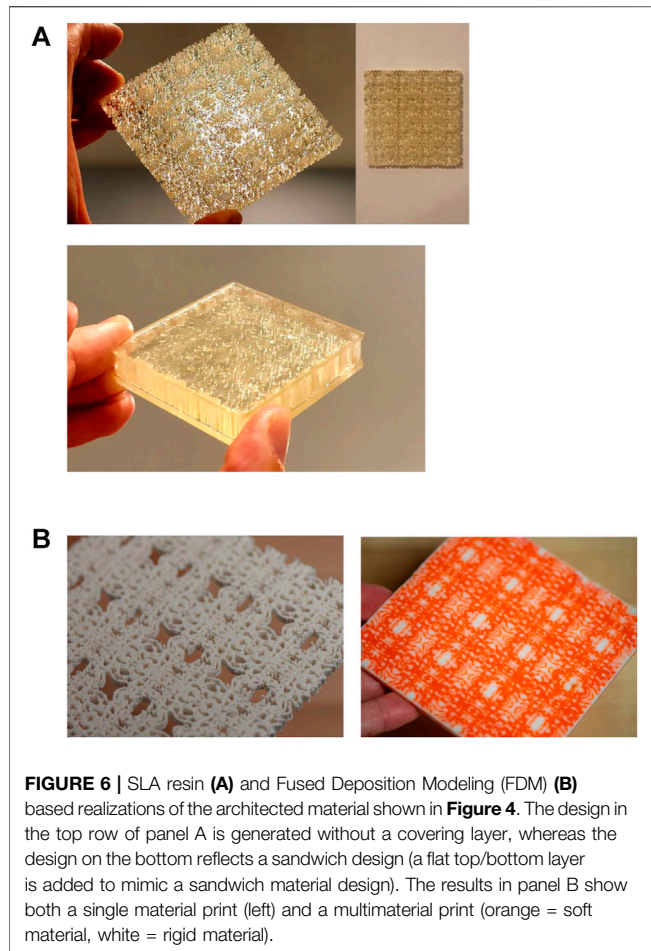
We proceed to demonstrate the use of the CLIP-VQGAN model employed here in realizing material designs. The design process starts from text prompts that describe the desired material design. **Figure 3** shows various example images generated from various text input. As the examples illustrate, interesting material designs can result from given text prompts. In the following sections we will systematically explore these and propose ways to develop physical material samples from these “written words” that act as drivers for the design process. The examples related to spider web and silk indicate that the approach can be used to design structures that resemble biomaterials. Furthermore, when the text input is “a porous scaffold of bone tissues,” the integrated approach provides an image with combined features of bones, porosity and matrix-like shape. The outcome implies potential applications in the field of tissue engineering. As these examples show, the method opens new creative approaches to identify radical new material concepts.

As can be seen from the examples in **Figure 3**, while the material designs generated by the algorithm provide a wealth of structural details, they typically require further processing and selection, for the purpose of physically manufacturing a material or selecting specific functions. We use tools from image processing to achieve this goal. **Figure 4** summarizes results of image generation (**Figure 4A**) based on the text prompt “a regular lattice of steel”, processing the resulting image into a black-and-white mask using smoothing operations (**Figure 4B**), and generation of periodic material designs using geometric

operations (first: horizontal flip, then vertical flip of the original and initially flipped design, yielding a periodic structure—as seen in the inset on the right exemplified with letters “ABC”) (**Figure 4C**). **Figure 4D** shows a close-up view of the resulting SLA 3D printed material, where one can see the capacity of the additive method to reproduce fine details of the design.

The resulting materials can be further analyzed using other computational methods, for instance to extract stress-strain distribution or to predict certain functional properties. **Figure 5** shows stress field predictions based on a cGAN method developed by the authors that enables us to predict mechanical tensor field data directly from microstructural designs (Yang et al., 2021a). Uniaxial tensile tests in different directions are tested considering the anisotropy of the structure. In the tensile test along the x direction, the long and slim islands in the middle show relatively high stress concentration as the local material density is low and the islands are aligned with the loading direction. By contrast, when the structure is stretched along the y direction, the high stress region is witnessed more at the left and right edge. Such an integrated use of neural nets to generate material designs, and using another model to assess mechanical features, can form the basis for systematic optimization and adaptation of designs over multiple iterations.

Figure 6 depicts realizations of the architected material shown in **Figure 4** (panel A: resin based printing using SLA, panel B, Fused Deposition Modeling (FDM) based printing). The design in the top row of **Figure 4A** (is generated without a covering layer, whereas the design on the bottom reflects a sandwich design (a



flat top/bottom layer is added). These examples demonstrate how the design concepts can be adapted further to generate meaningful engineering structures.

We move on to other examples, especially focused on architected materials that feature mechanical mechanisms. Figure 7A shows an analysis of the effect of slight variations of input text, yielding small changes in the material design. The use of the phrase “black and white” in the text input yields a grayscale image, whereas the inclusion of terms such as “periodic,” “lattice” and “high contrast” generates designs that have clear distinct phases that are important for further processing. To demonstrate that process, Figure 7B shows how the design can be used to extract small sections of the large image that show interesting material features. These can be utilized in further analysis.

Figure 8 shows an analysis in which we exploit one of the material designs shown in Figure 7 and focus specifically on the extraction of a structure with a continuous geometry and material islands (i.e., microstructural features that are not connected to others) removed. The final design is then printed using polyurethane, and used for mechanical testing. Notably, the resulting design represents a simple topological gripper mechanism (e.g., for soft robotics applications), as shown in Figure 9 from both a computational and experimental perspective.

Figures 9A,C show a series of Finite Element simulations to assess the mechanisms, including a direct comparison with experiment as shown in Figures 9B,D, providing a quantitative analysis of the deformation. We find that when one handle is twisted clockwise and the other counter-clockwise, the upper domain of the structure opens up and the lower domain closes down shown in Figures 9A,B. For the diagonal compression in Figures 9C,D, the loading not only leads to shrinkage of the main body along the diagonal, but also a twisting or rotation of the two handles, showing an interesting deformation mechanism. Movies M1-M2 show the complete process of the deformation given two different loading conditions based on the FEM results, and Movies M3-4 show results for the same loading condition as observed in experiment. There is good agreement between the computational prediction and the experiment result. Supplementary Figure S1 shows snapshots of various other material architectures extracted from the source images shown in Figure 7B, offering insights into the diversity of designs one can achieve.

Figure 10 shows stress field predictions. Similarly, as reported in Figure 5, two different loading conditions are investigated for the anisotropic structure. As the stress fields show, in the uniaxial tension test along x direction, the two symmetric horizontal lines feature higher stress levels. By contrast, when the loading is along y direction, the vertical X-shape regions bear the most load. Similar as in Figure 5, the stress concentration mechanism suggests that the continuous regions aligned with the loading direction are most essential parts in resisting deformation.

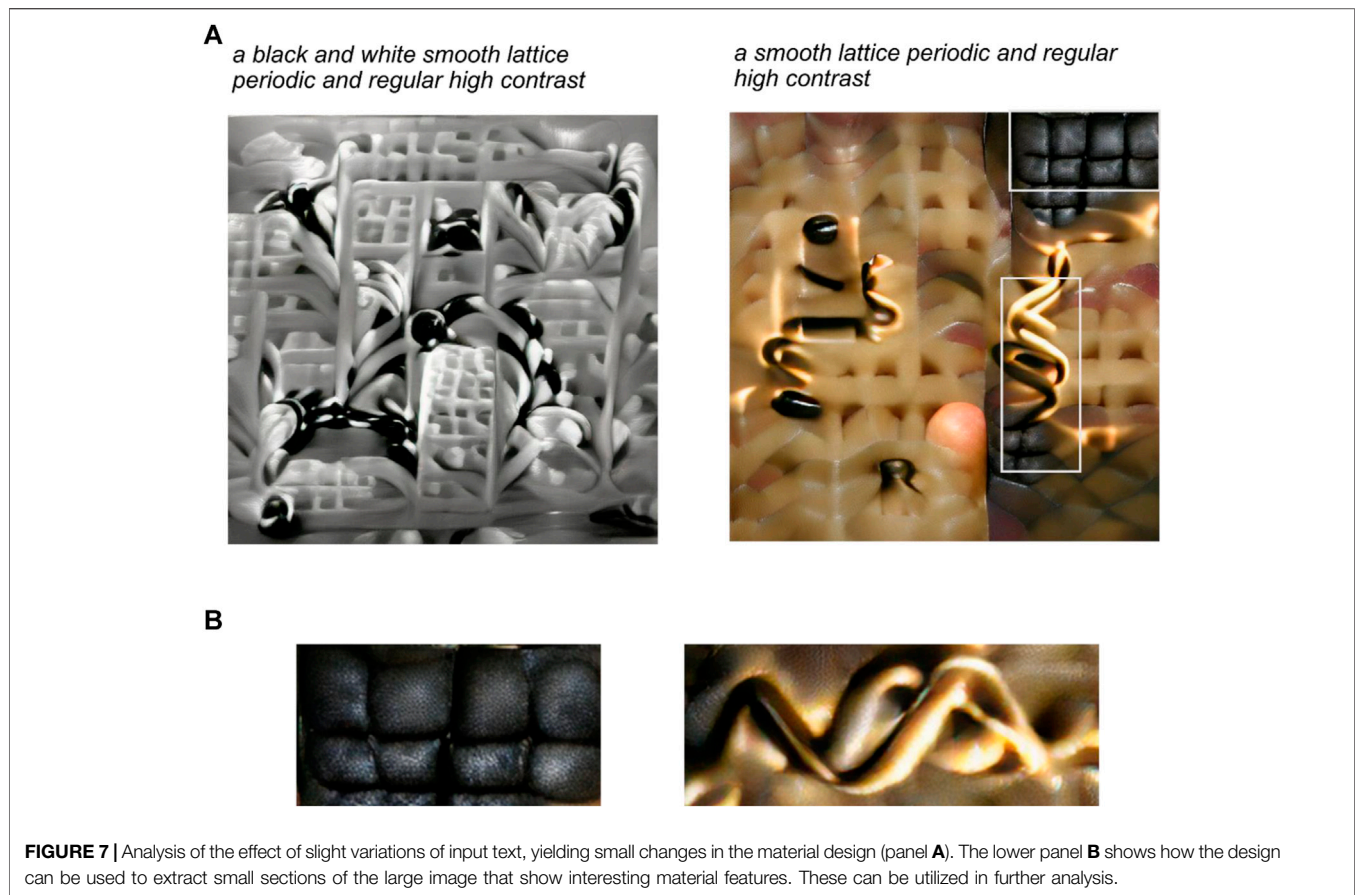
Supplementary Figure S1 shows the results of SLA based resin 3D printing process (slicing, top; printing, bottom). Supplementary Figure S2 shows the results of FDM based 3D printing process, showing that the materials can be constructed using a variety of methods.

DISCUSSION AND CONCLUSION

In this paper we demonstrated the use of transformer neural nets to enable the conversion of text input to “describe” a material into models and 3D printed physical specimen. Such methods have the potential to be used as a natural-language-driven tool to develop complex materials designs, with broad implications as a novel design approach.

Indeed, future research may build on this work in numerous ways. For instance, the capabilities of vision transformer neural nets to generate 3D image renderings could be used to output realistic geometry (e.g., STL) files directly, based on perspective changes already realized for conventional imagery, and even offer automatic rendering into 2D images or for virtual reality applications. Future work could direct text input into other forms of representations, such as 3D geometries (e.g., volume meshes, or surface meshes (Brommer et al., 2015)), or even directly into 3D printing instructions. The diverse transformations that can be achieved by transformer neural nets opens many pathways for future developments.

Another important development will be to develop novel training sets, perhaps based on transfer learning in case of



data scarcity, to offer a directed way to generate novel microstructures, more specifically targeting microstructural material aspects and sources for structural information such as naturally occurring patterns. Transfer learning may reduce the need for significant datasets, and could be a viable approach to realize successful material renderings. However, there remain significant research challenges such as the need for large and adequate datasets, or the development of proper transfer learning methods.

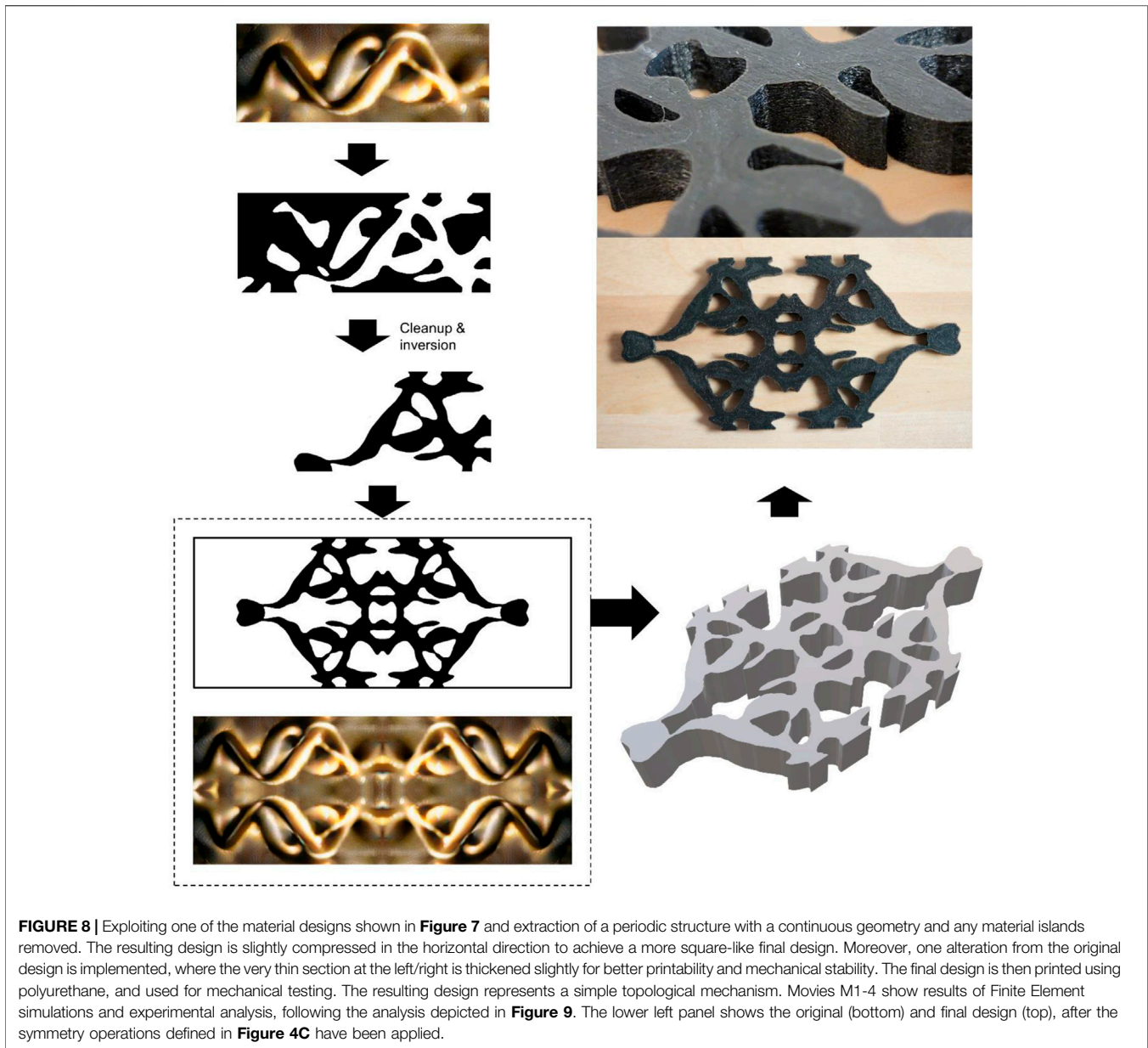
The work reported here explored the use of text-driven architected material design as an explorative tool, without deliberate optimization. Future work can address that limitation of the present work. Indeed, building on the work reported here, the use of adequate objective functions can be a powerful way to direct design to meet certain design demands. It can either be complemented to the text input (i.e., the generative process will be informed by additional constraints), or use a systematic variation of the text input itself. The use of genetic algorithms (Yu et al., 2019b), for instance, can be a good way to offer directed evolution realized in a simulated environment, and allow the rapid exploration of a wide design space, including natural language prompts and variations.

In this work we relied solely on existing trained models. In spite of the use of an already trained model, the results show that intriguing materials designs can be generated using this approach, and that the generative tool is capable of producing abstract

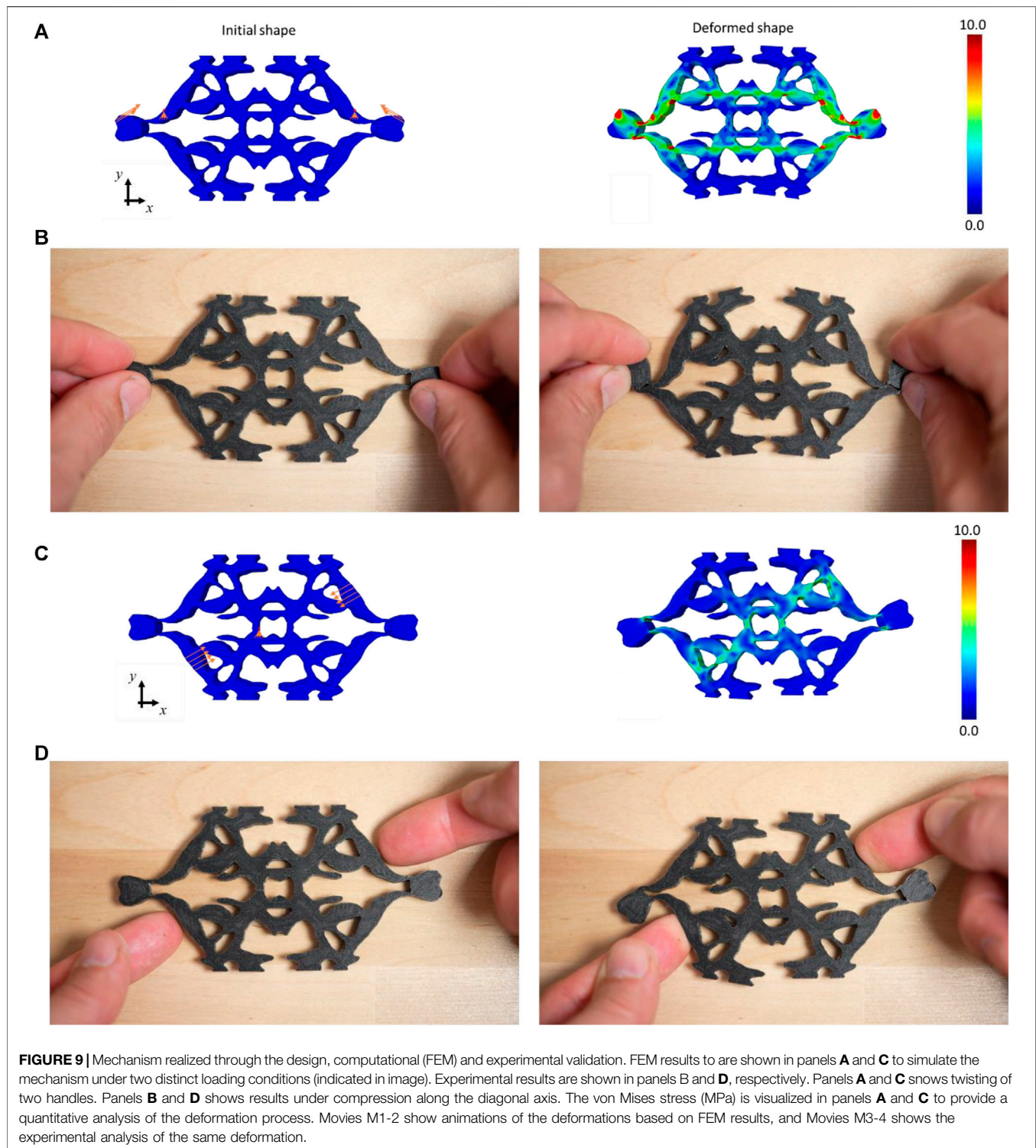
microstructural designs that can form either regular architected lattice-like materials (see **Figures 4–6**, for instance), or yield mechanisms (see **Figures 7–9**). The various demonstrations developed within this study shed light onto the fact that the learned word associations, and realizations into material designs, can form the basis for a generic framework that works with pre-trained models that can be further improved, and adapted, towards specific material applications and constraints. However, we still need to be careful when using those large pretrained base model since many academic terminologies are not common words or the design may not be easy to describe. In other words, the existing trained models may lack knowledge for more delicate materials design purposes which need complex texts input or rare words input. In this case, we could also use transfer learning techniques to further tune the pretrained model with additional data that include necessary information about materials or designs. Given a well-trained model like the CLIP or VQGAN model, a small amount of data should readily enable us to handle more specific design problems.

The contributions of this work to materials research, and future opportunities, are:

- The work enables a conversion from text input to microstructure images directly, which can be leveraged to design materials and further manufacture the generated design using 3D printing.

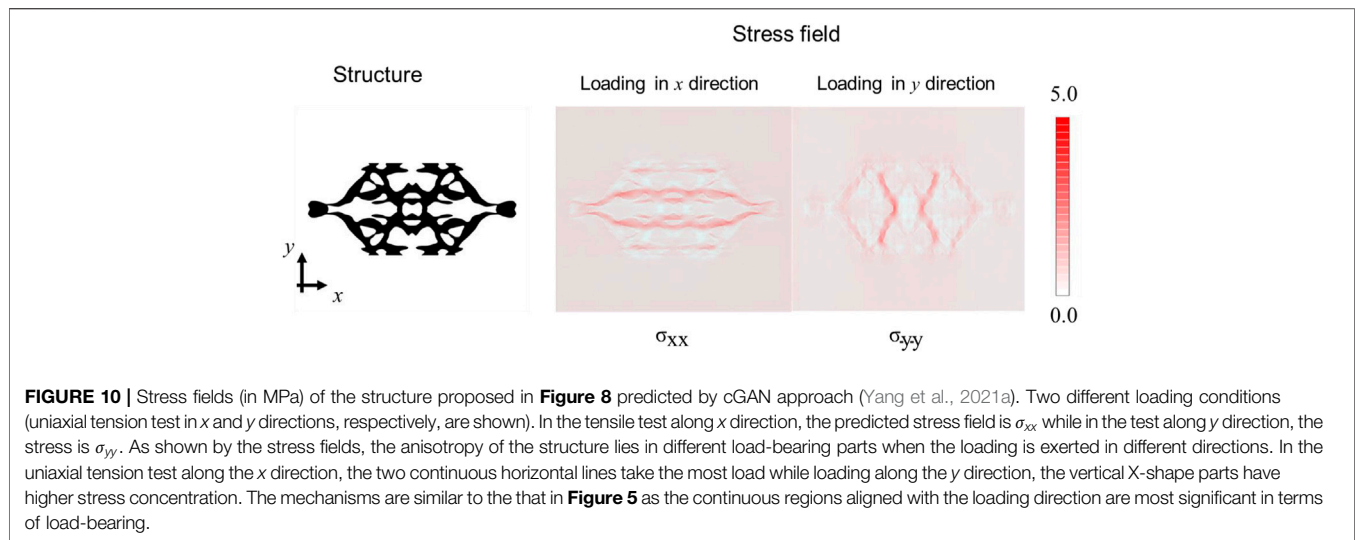


- In future work, the input text can be tuned to contain design protocol such as constraints (e.g., component material) and design objectives (e.g., desirable properties). As a result, designing materials with desirable properties given certain constraints can be realized, including through the use of genetic algorithms, or Bayesian optimization (Bock et al., 2019).
 - More generally, the text input can be any sequential data and the image output can be more than generated structures. For example, materials represented by sequential data such as SMILES strings and protein sequences can be inputted to the neural networks for predicting 2D energy landscapes or physical properties (e.g., charge density) that will then be shown as images or stacks of images (for volume data).
 - In addition, the sequential input can not only be the material itself, but also the design process of materials (such as machine specific CNC code, e.g., RS-274/G-code) for 3D printing and layout strategy of a material manufacturing process. The corresponding output include yield geometries or physical fields under certain loading conditions (e.g., strain and stress).
 - Apart from the several examples mentioned above, there are more potential applications of the idea converting sequences to images with transformer neural networks and additive methods.
- In conclusion, this study opened a first perspective into the novel approach of designing materials via text prompts, and engaging into a genuine human-computer partnership.



Such a promise pushes us beyond existing uses of AI primarily for information gathering and interacting with computers (e.g., Alexa, Cortana, or Google Assistant) into materialization of text. It may realize a new “magical” Frontier where spoken words are materialized through the

integrated use of language, AI, and novel manufacturing techniques. This partnership between human intelligence, creativity and knowledge with deep learning tools can form the basis for exciting new discoveries in science, technology, and artistic expression.



MATERIALS AND METHODS

The methods used in this paper include various deep neural nets, additive manufacturing, and image-based analysis. We also perform mechanical analyses using both cGAN approach and finite element method (FEM), to realize the overall flowchart depicted in **Figure 1**.

Contrastive Language-Image Pre-Training Model

CLIP (Radford et al., 2021) (Contrastive Language-Image Pre-Training) is a neural network trained on a variety of (image, text) pairing, to carry out complex image classification tasks. The CLIP model serves as a classifier to assess whether or not a produced image meets the target text prompt provided by the user. The model is designed to perform a great variety of classification benchmarks with similar “zero-shot” capacity like GPT models. In other words, there is no need for direct optimization given the model’s generalization performance and robustness. The model consists of two encoders to encode images and texts separately and further predict the correct pairing of the batches. To train the model, 400 million (image, text) pairs are collected from the internet in the original work (Radford et al., 2021).

Being tested on more than 30 different benchmark datasets, the model shows the power of transferring non-trivially to those tasks and even being competitive compared to existing baseline models. Thanks to the universality and robustness of the approach, we here use pretrained CLIP model to evaluate the generated images from VQGAN given the text inputs, providing better match between the images and texts. More specifically, the pretrained CLIP model used in this work is based on one specific version of ViT named “ViT-B/32” (Dosovitskiy et al., 2020) which contains 12 neural layers with hidden size of 768. The size of the multilayer perceptron (MLP) in the model is 3,072 and the number of attention heads is 12. The total number of parameters is around 86 million. The “32” in the name of the

model refers to the input patch size and the “B” is in short of “base” given other variants such as “ViT-L” (“L” for “Large”) or “ViT-H” (“H” for “Huge”).

VQGAN Model

CNNs can efficiently extract local correlation in an image while transformers are designed to obtain long-range interactions on sequential data. The VQGAN model (Esser et al., 2020) aims at combining the effectiveness of CNNs with the expressivity of transformers to generate high-resolution images. To link two different model architectures (CNNs and transformers), the VQGAN model uses a convolutional encoder to learn a codebook (similar as the latent vectors) of context-rich visual parts which is then input to the transformer to extract long-range interaction within the compositions. The codebook is further passed to the CNN decoder which serves as the generator in the GAN to generate the desired image. Like any other GAN models, a discriminator is finally used to compare the generated image to the original image. The VQGAN model used in this study has been trained on the ImageNet (Deng et al., 2010) dataset and used for generating image candidates given the text description. In terms of the hyperparameters, the pretrained VQGAN used in this work has a codebook size of 16×16 and the factor of first-stage encoding is 16.

Integrated Contrastive Language-Image Pre-Training and VQGAN Model

The integrated model as reported in (Crowson, 2021) is based on the BigSleep (Komatsuzaki, 2021), (lucidrains/big-sleep, 2021) algorithm that takes a text prompt and visualizes an image to fit the text provided. Here, VQGAN is used as an image generator guided by the CLIP model. It works as follows, whereby VQGAN is used as the image generator method, and CLIP used as the scoring method to classify the images produced. Over the iterations, the algorithm searches for images that match the provided text prompt increasingly well by tweaking the input

provided to the VQGAN algorithm. A schematic of the method is depicted in **Figure 2**. The input image size is 480×480 pixels with the number of cutouts at 64. The integrated model is trained around 4,000 iterations for each text input to obtain the relevant image.

Image Processing and Preparation for Additive Manufacturing

The primary objective for image processing is to transform a complex image with multi-channel (or grayscale) data into a printable geometry. To that end we use colormap operations, smoothing functions and image processing methods to remove small image parcels, and/or islands, to generate printable mechanically functional designs. The target objective is to yield image data with distinct colors at each pixel that refer to the material type (e.g., white = void/no material, black = material, or white = soft material, black = rigid material in Multimaterial 3D printing). All image operations are done using the OpenCV computer vision package implemented in Python (Bradski, 2000). Once the basic parameters are set, the model makes automatic predictions which are generally printable using suitable methods.

Mechanical Analysis Using conditional GAN

To predict the stress field, we use a conditional GAN (cGAN) approach reported in earlier work (Yang et al., 2021a). As the original work is aimed at dealing with composites that contain more than one material, we fill the designs proposed in this work with an extremely soft materials to serve as models for voids.

This setup enables us to simulate the mechanical behaviors of those complex and sometimes discontinuous structures which could be a nightmare for normal FEM simulations. To simplify, both the stiff (black color) and soft material (white color which is the void) are modeled as linear elastic material since the deformation in the mechanical test is not large and no hysteresis is taken into consideration. To model the voids, we set the Young's modulus of soft material (1 MPa) to be 1% of the stiff material (100 MPa). The data generation and training follow the same strategy proposed in the author's previous work (Yang et al., 2021a). The stress fields under both uniaxial tensile tests can be predicted by 1 ML model through rotating the geometries by 90° .

Mechanical Analysis Using Finite Element Method

We use FEM to reproduce and prove the deformation mechanism shown in the experiment. The simulations are performed using a commercial Abaqus/Standard code (Dassault Systems Simulia Corp., 2010). To keep consistence with the ML predictions on the mechanical fields, a linear elastic material with Young's modulus being 100 MPa is used for the material. The design geometry in STL format is first transferred to SAT format which converts a mesh body to a solid body using SolidWorks (Dassault Systems). The transferred SAT file is then imported as a part to the Abaqus for simulation.

Once the material properties and the geometry are defined, the structure is meshed with tetrahedral elements (C3D10) using the

default algorithm of Abaqus for 3D stress analysis. In the loading condition shown in **Figure 9A**, the displacement amplitudes are 10 and 4 mm along the x direction and y direction, respectively. In the loading condition shown in **Figure 9C**, the displacement amplitudes are 3 and 1.2 mm along the x direction and y direction respectively. The von Mises stresses are plotted using Abaqus Visualization. The contour style used is "CONTINUOUS" and the "FEATURE" option is applied to visible edges. Any legend, title, state, annotation, compass and reference point are turned off to exclude useless information but turned on in the **Supplementary Movies**.

Additive Manufacturing

We employ additive manufacturing to generate 3D models of the materials designed from words. 3D files are sliced using Cura 4.9.1 and printed using a Ultimaker S3 multi-material printer and a QIDI X-Pro printer. Material samples are printed using Ultimaker TPU and Ultimaker PLA filaments, as well as generic PLA and TPU filaments in the QIDI X-Pro printer. SLA resin printing is performed using a Anycubic Photon X printer, whereby STL files are sliced using Chitobox, and printed using black and transparent resins. The use of transparent resins, as shown in **Figure 6**, enables one to assess the internal microstructure of the architected material. **Supplementary Figure S2** shows snapshots of the SLA resin printing approach. **Supplementary Figure S3** shows snapshots of the FDM printing approach, including multimaterial printing. Some complex geometries may require multimaterial printing for support material, such as PVA.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author.

AUTHOR CONTRIBUTIONS

MJB conceived the idea and developed the research. MJB performed the neural network calculations, developed the image processing, and performed the 3D printing and experimental tests. ZY performed FE and cGAN modeling, and analyzed the results. MB and ZY wrote the paper.

ACKNOWLEDGMENTS

The authors acknowledge support the MIT-IBM AI Lab, as well as ONR (N000141912375) and ARO (W911NF1920098).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmats.2021.740754/full#supplementary-material>

REFERENCES

- Bahdanau, D., Cho, K., and Bengio, Y. (2015). *Neural Machine Translation by Jointly Learning to Align and Translate*, 1–15. arXiv [Preprint]. Available at: <https://arxiv.org/abs/1409.0473>
- Bock, F. E., Aydin, R. C., Cyron, C. J., Huber, N., Kalidindi, S. R., and Klusemann, B. (2019). A Review of the Application of Machine Learning and Data Mining Approaches in Continuum Materials Mechanics. *Front. Mater.* 6, 110. doi:10.3389/fmats.2019.00110
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's J. Softw. Tools*.
- Brommer, D. B., Giesa, T., Spivak, D. I., and Buehler, M. J. (2015). Categorical Prototyping: Incorporating Molecular Mechanisms into 3D Printing. *Nanotechnology* 27 (2), 024002. doi:10.1088/0957-4484/27/2/024002
- Brommer, D. B., Giesa, T., Spivak, D. I., and Buehler, M. J. (2016). Categorical Prototyping: Incorporating Molecular Mechanisms into 3D Printing. *Nanotechnology* 27 (2), 024002. doi:10.1088/0957-4484/27/2/024002
- Brown, T. B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., et al. (2020). *Language Models are Few-Shot Learners*. in *Advances in Neural Information Processing Systems, 1877–1901*. Available at: <https://proceedings.neurips.cc/paper/2020/file/1457c0d6bfc4967418bfb8ac142f64a-Paper.pdf>.
- Buehler, M. J., and Misra, A. (2019). Mechanical Behavior of Nanocomposites. *MRS Bull.* 44 (1), 19. doi:10.1557/mrs.2018.323
- Buehler, M. J. (2010). Tu(r)ning Weakness to Strength. *Nano Today* 5, 379. doi:10.1016/j.nantod.2010.08.001
- Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., and Zagoruyko, S. (2020). “End-to-End Object Detection with Transformers,” in *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 12346 LNCS, (Glasgow, United Kingdom, August 23–August 28, 2020), 213–229. doi:10.1007/978-3-030-58452-8_13
- Chaudhari, S., Mithal, V., Polatkan, G., and Ramanath, R. (2019). An Attentive Survey of Attention Models. arXiv [Preprint]. Available at: <http://arxiv.org/abs/1904.02874>.
- Cranford Buehler, S. W., and Markus, J. (2012). *Biomateriomics*. Dordrecht, Netherlands: Springer Netherlands.
- Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, Li., et al. (2010). *A Large-Scale Hierarchical Image Database*, in *Institute of Electrical and Electronics Engineers (IEEE)*, 248–248. doi:10.1109/cvpr.2009.5206848
- Devlin, J., Chang, M. W., Lee, K., and Toutanova, K. (2019). “BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding,” in *NAACL HLT 2019 - 2019 Conf. North Am. Chapter Assoc. Comput. Linguist. Hum. Lang. Technol. - Proc. Conf.*, vol. 1, no. Mlm, (Minneapolis, Minnesota, June 2–June 7, 2019), 4171–4186.
- Dosovitskiy, A., et al. (2020). *An Image Is Worth 16x16 Words: Transformers for Image Recognition at Scale*. arXiv [Preprint]. Available at: <http://arxiv.org/abs/2010.11929> (Accessed May 31, 2021).
- Esser, P., Rombach, R., and Ommer, B. (2020). *Taming Transformers for High-Resolution Image Synthesis*. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 12873–12883.
- Giesa, T., Spivak, D. I., and Buehler, M. J. (2011). Reoccurring Patterns in Hierarchical Protein Materials and Music: The Power of Analogies. *Bionanoscience* 1 (4), 153. doi:10.1007/s12668-011-0022-5
- Grechishnikova, D. (2021). Transformer Neural Network for Protein-specific De Novo Drug Generation as a Machine Translation Problem. *Sci. Rep.* 11 (1), 1–13. doi:10.1038/s41598-020-79682-4
- Guo, K., Yang, Z., Yu, C.-H., and Buehler, M. J. (2021). Artificial Intelligence and Machine Learning in Design of Mechanical Materials. *Mater. Horiz.* 8 (4), 1153–1172. doi:10.1039/d0mh01451f
- Jensen, Z., Kim, E., Kwon, S., Gani, T. Z. H., Román-Leshkov, Y., Moliner, M., et al. (2019). A Machine Learning Approach to Zeolite Synthesis Enabled by Automatic Literature Data Extraction. *ACS Cent. Sci.* 5, 892. doi:10.1021/acscentsci.9b00193
- Komatsuzaki, A. (2021). When You Generate Images with VQGAN + CLIP, the Image Quality Dramatically Improves if You Add “unreal Engine” to Your prompt, Twitter. [Online]. Available: <https://twitter.com/arankomatsuzaki/status/1399471244760649729> (Accessed May 31, 2021).
- Liu, Y., Zhao, T., Ju, W., Shi, S., Shi, S., and Shi, S. (2017). Materials Discovery and Design Using Machine Learning. *J. Materiomics* 3 (3), 159–177. doi:10.1016/j.jmat.2017.08.002
- lucidrains/big-sleep (2021). Lucidrains/Big-Sleep: A Simple Command Line Tool for Text to Image Generation, Using OpenAI’s CLIP and a BigGAN. Technique was originally created by <https://twitter.com/advadnoun>.
- Milazzo, M., Negrini, N. C., Scialla, S., Marelli, B., Farè, S., Danti, S., et al. (2019). Additive Manufacturing Approaches for Hydroxyapatite-Reinforced Composites. *Adv. Funct. Mater.* 29 (35), 1903055. doi:10.1002/adfm.201903055
- Milazzo, M., and Buehler, M. J. (2021). Designing and Fabricating Materials from Fire Using Sonification and Deep Learning. *iScience* 24 (8), 102873. doi:10.1016/j.isci.2021.102873
- Olivetti, E. A., Cole, J. M., Kim, E., Kononova, O., Ceder, G., Han, T. Y. G., et al. (2020). Data-driven Materials Research Enabled by Natural Language Processing and Information Extraction. *Appl. Phys. Rev.* 7 (4), 041317. doi:10.1063/5.0021106
- Palkovic, S. D., Brommer, D. B., Kupwade-Patil, K., Masic, A., Buehler, M. J., and Büyükköztürk, O. (2016). Roadmap across the Mesoscale for Durable and Sustainable Cement Paste - A Bioinspired Approach. *Constr. Build. Mater.* 115, 13. doi:10.1016/j.conbuildmat.2016.04.020
- Pesciullesi, G., Schwaller, P., Laino, T., and Reymond, J. L. (2020). Transfer Learning Enables the Molecular Transformer to Predict Regio- and Stereoselective Reactions on Carbohydrates. *Nat. Commun.* 11 (1), 4874–4878. doi:10.1038/s41467-020-18671-7
- Qin, Z., Dimas, L., Adler, D., Bratzel, G., and Buehler, M. J. (2014). Biological Materials by Design. *J. Phys. Condens. Matter* 26 (7), 073101. doi:10.1088/0953-8984/26/7/073101
- Radford, A., Kim, J. W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., et al. (2021). *Learning Transferable Visual Models From Natural Language Supervision*. arXiv [Preprint]. Available at: <https://arxiv.org/abs/2103.00020>
- Schwaller, P., Laino, T., Gaudin, T., Bolgar, P., Hunter, C. A., Bekas, C., et al. (2019). Molecular Transformer: A Model for Uncertainty-Calibrated Chemical Reaction Prediction. *ACS Cent. Sci.* 5 (9), 1572–1583. doi:10.1021/acscentsci.9b00576
- Spivak, D. I., Giesa, T., Wood, E., and Buehler, M. J. (2011). Category Theoretic Analysis of Hierarchical Protein Materials and Social Networks. *PLoS One* 6 (9), e23911. doi:10.1371/journal.pone.0023911
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention Is All You Need. *Adv. Neural Inf. Process. Syst.* 2017-Decem (Nips), 5999–6009.
- CrowsonK. (2021). *VQGAN+CLIP*. Available at: <https://colab.research.google.com/github/chigozienri/VQGAN-CLIP-animations/blob/main/VQGAN-CLIP-animations.ipynb>
- Wang, F., and Tax, D. M. J. (2016). *Survey on the Attention Based RNN Model and its Applications in Computer Vision*. arXiv [Preprint]. Available at: <http://arxiv.org/abs/1601.06823>
- Wegst, U. G. K., Bai, H., Saiz, E., Tomsia, A. P., and Ritchie, R. O. (2015). Bioinspired Structural Materials. *Nat. Mater.* 14 (1), 23–36. doi:10.1038/nmat4089
- Yang, Z., Yu, C.-H., and Buehler, M. J. (2021). Deep Learning Model to Predict Complex Stress and Strain fields in Hierarchical Composites. *Sci. Adv.* 7 (15), eabd7416. doi:10.1126/sciadv.abd7416
- Yang, Z., Yu, C.-H., Guo, K., and Buehler, M. J. (2021). End-to-end Deep Learning Method to Predict Complete Strain and Stress Tensors for Complex Hierarchical Composite Microstructures. *J. Mech. Phys. Sol.* 154, 104506. doi:10.1016/j.jmps.2021.104506
- Yu, C.-H., Qin, Z., and Buehler, M. J. (2019). Artificial Intelligence Design Algorithm for Nanocomposites Optimized for Shear Crack Resistance. *Nano Futur.* 3 (3), 035001. doi:10.1088/2399-1984/ab36f0
- Yu, C.-H., Qin, Z., Martin-Martinez, F. J., and Buehler, M. J. (2019). A Self-Consistent Sonification Method to Translate Amino Acid Sequences into Musical Compositions and Application in Protein Design Using

Artificial Intelligence. *ACS Nano* 13, 7471–7482. doi:10.1021/acsnano.9b02180

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in

this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Yang and Buehler. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.