Modeling Atomistic Dynamic Fracture Mechanisms Using a Progressive Transformer Diffusion Model

Citation: Buehler, Markus J. 2022. "Modeling Atomistic Dynamic Fracture Mechanisms Using a Progressive Transformer Diffusion Model." Journal of Applied Mechanics, 89 (12).

As Published: 10.1115/1.4055730

Publisher: ASME International

Persistent URL: https://hdl.handle.net/1721.1/148574

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Modeling Atomistic Dynamic Fracture Mechanisms Using a Progressive Transformer Diffusion Model

Markus J. Buehler
Laboratory for Atomistic and Molecular
Mechanics (LAMM);
Center for Computational Science and
Engineering,
Schwarzman College of Computing,
Massachusetts Institute of Technology,
77 Massachusetts Avenue,
Cambridge, MA 02139
e-mail: mbuehler@mit.edu

Dynamic fracture is an important area of materials analysis, assessing the atomic-level mechanisms by which materials fail over time. Here, we focus on brittle materials failure and show that an atomistically derived progressive transformer diffusion machine learning model can effectively describe the dynamics of fracture, capturing important aspects such as crack dynamics, instabilities, and initiation mechanisms. Trained on a small dataset of atomistic simulations, the model generalizes well and offers a rapid assessment of dynamic fracture mechanisms for complex geometries, expanding well beyond the original set of atomistic simulation results. Various validation cases, progressively more distinct from the data used for training, are presented and analyzed. The validation cases feature distinct geometric details, including microstructures generated by a generative neural network used here to identify novel bio-inspired material designs for mechanical performance. For all cases, the model performs well and captures key aspects of material failure. [DOI: 10.1115/1.4055730]

Keywords: fracture, dynamics, mechanics, deep learning, attention models, progressive diffusion models, transformer, language models, computational mechanics, constitutive modeling of materials, flow and fracture, mechanical properties of materials, micromechanics, structures

1 Introduction

The dynamics of fracture [1–10] is an intriguing challenge in mechanics and materials science. In brittle materials, the focus of this study, cracks initiate upon a critical loading level and propagate rapidly throughout the material. Depending on the details of the scenario, such as the orientation of the crack relative to the applied load, the crystal orientation, critical temporal aspects such as the loading rate, the precise mechanisms, and temporal evolution, of fracture, changes drastically. A method commonly applied to model the dynamics of brittle fracture is the use of molecular dynamics (MD) [10,11]. However, such models can be challenging in terms of computational needs and quickly reach their limits especially when combined with optimization algorithms that seek to identify a material's optimal micro- or nano-structure to meet certain design targets—such as preventing fractures from originating or directing fractures to move in a certain geometric fashion. New modeling tools are hence required to expand the MD results in a multiscale scheme, to effectively scale a set of MD results toward broader applicability (Fig. 1).

In this paper, we propose a method to use MD simulation results to train a deep neural network to directly and efficiently predict dynamical information of crack spreading. In recent work, AI systems have been shown to generate complex image solutions to text prompts [12–15]. In the context of modeling a physical system, here we replace the text prompt with an expression that uniquely characterizes the input microstructure of a material—for instance, used to characterize the presence of voids or cracks, as outlined in Fig. 2. We then train a model to not only generate a single image but rather to produce a series of simulation frames that predict the *dynamics* of defect evolution from the input.

Various deep learning methods have been used to generate field data such as images and, in some cases, video. Earlier attempts used autoencoder architectures [16,17] and generative adversarial neural networks (GANs) [18–20]. More recently, transformer-based diffusion model architectures have emerged as a broadly applicable approach for many data modalities including state of the art text-to-image generation (e.g., VQGAN-CLIP [21,22], GLIDE [14], DALL-E 2 [12,23], Imagen [13], or Latent/Stable Diffusion [24]. Other recent work has focused on using deep neural networks, including GAN and transformer architectures for video generation [25–27]. However, in GANs, solutions are generated from noise, in diffusion models a neural network is trained to translate, through a series of progressive *reverse* diffusion steps [28], noise into a proper solution, via a Markov chain. This denoising strategy can be conditioned with additional information, such as text prompts or as done in this study, an embedding of the material microstructure.

The purpose of this study is to explore the use of diffusion models for modeling dynamical physical phenomena and specifically the dynamics of fracture. Our model is trained against MD simulation results, here using a simple 2D Lennard–Jones material [8] that is known to be a brittle material. It is noted that we anticipate that the framework can be explored and possibly extended to other dynamical phenomena in materials science, chemistry, physics, and engineering. Naturally, instead of using simulated data, the model could also be trained against other dynamical data (e.g., video or image sequences) from experimental studies, or use data from various mechanical engineering devices. Since the model can convert relatively low-dimension input (text, a single slice of a microstructure, etc.) into a higher-order output—here, frames of a time sequence of fracture states—it may also be possible to use such models to predict folding patterns in origami or other 2D-to-3D transformations [29,30].

1.1 Opportunities and Challenges.

One of the challenging aspects of applying deep learning to mechanics is to expand

Contributed by the Applied Mechanics Division of ASME for publication in the Journal of Applied Mechanics. Manuscript received August 23, 2022; final manuscript received September 20, 2022; published online October 6, 2022. Tech. Editor: Pradeep Sharma.

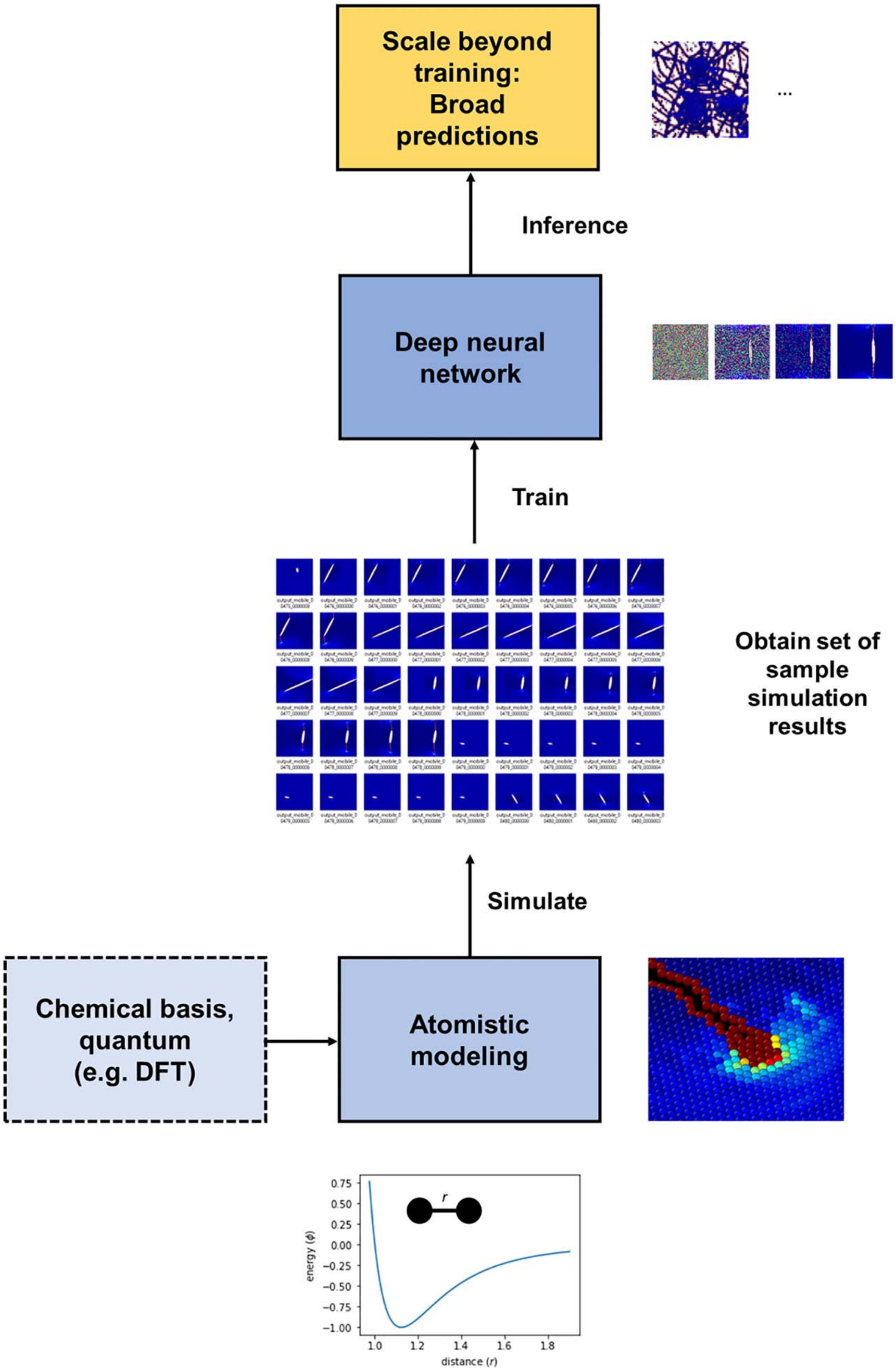

Fig. 1 Overall workflow of the results reported in the paper, using a progressive transformer diffusion model (details on the model in Fig. 3(*a*), Materials and Methods, and Supplementary Information). We conduct a series of atomistic simulations based on an empirical potential (albeit, in principle, these simulations could be obtained based on more complex interatomic potentials or even quantum-level simulations; albeit not done here) to train a deep neural network. Having been trained on a relatively small set of sample simulation results, the deep neural network is then applied to generalize predictions toward a wider range of geometries, beyond those included in the training set.

models beyond relatively simple interpolation solutions or "curve fitting" to make predictions for scenarios far from the data set that was used to train the model. Developing a better understanding of the predictive capacity of physics-based deep learning has been a subject of various investigations [31,32] and is an important frontier in the field. In this paper, we focus on using deep learning for dynamic fracture modeling and investigate the use of a small data set generated from lower-level atomistic-level simulation results and the resulting generalization capacity in the context of other geometries, to bridge scales [33–37]. Modeling field predictions has been reported using earlier deep learning models developed for solid mechanics

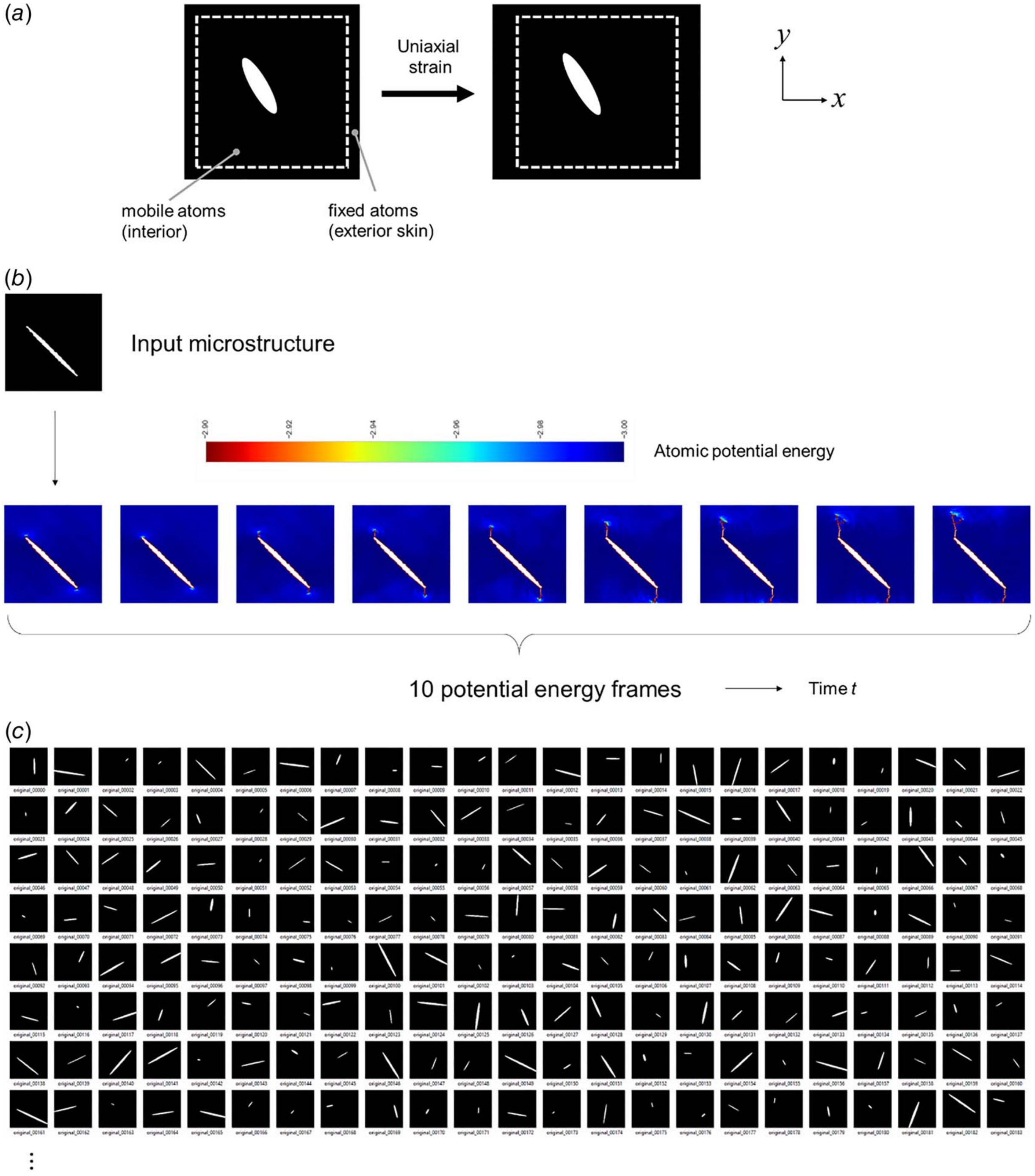

Fig. 2 Molecular dynamics (MD) simulation setup (*a*) and sample result from MD simulation, used to build the training set (*b*) and (*c*). The MD simulation setup consists of a 2D hexagonal lattice with LJ interactions; exposed to tension in the *x*-direction. The small area outside of the area defined by the dashed line is fixed, whereas all atoms within are free to move in *x* and *y*. We extract the potential energy fields in the region with mobile atoms for the construction of the data set. (*b*) Example input microstructure and predicted 10 frames that capture the evolution of the potential energy fields, thereby offering a detailed visualization of defect dynamics. Crack surfaces and dislocations are visualized in dark shade/red color; a perfect hexagonal lattice in light shade/blue color. (*c*) Samples of the type of microstructures used in the training set. All input microstructures feature a single crack, varied randomly in size, shape, and orientation. All cracks are elliptical.

problems including elasticity and fracture cases [38–42] but has not yet been applied to predict high-resolution full-field failure sequences.

1.2 Scope of This Paper and Outline. It is emphasized that this paper is meant to provide the first application of diffusion models to describe dynamic materials failure, with numerous research directions and open questions left to future work. The model is tasked to learn to construct the evolution of the field solution, here focused on the potential energy field so that we can closely examine the spreading of cracks and other atomistic defects. We first report details on the data set generation and training, followed by a series of testing and validation steps. We successively expand the challenge of the test conducted to investigate the predictive limit of the model. We conclude the paper with a discussion on limitations and possible next steps.

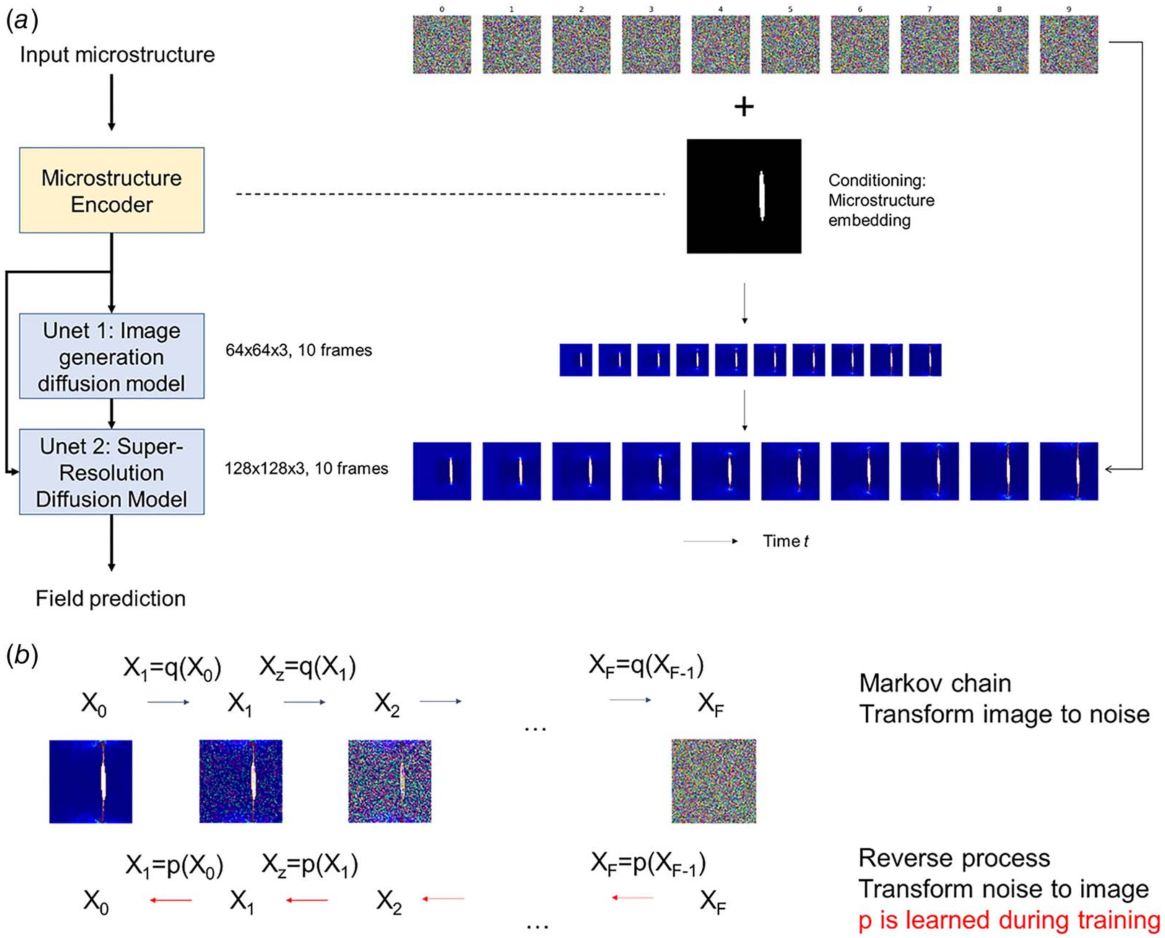

Fig. 3 Summary of the neural network architecture used in this study, whereby a microstructure encoder is fed to two successive U-Nets (Unet 1 and Unet 2; simple visual of a U-Net shown in Figure S4 available in the Supplemental Materials on the ASME Digital Collection). Each U-Net produces a series of 10 frames, scaling up to a final resolution of 128 × 128 field predictions. (*a*) Summary of the model architecture and condition of the noise input based on the microstructure and (*b*) illustration of the learned process *p* during training, denoising data to yield proper field solution, as described in Eq. (4). The schematic in *b* only shows one frame, whereas the model is trained to predict a successive number of frames to capture dynamical information.

2 Results and Discussion

To start, Fig. 2(*a*) shows the molecular dynamics (MD) simulation setup and sample result from the MD simulation. The results from a large number of MD simulations similar to that one are used to build the training set. The MD simulation setup consists of a two-dimensional (2D) hexagonal lattice with Lennard–Jones (LJ) interactions (as described in Materials and Methods). The crystal is exposed to tension in the *x*-direction. Figure 2(*b*) depicts an example input microstructure and 10 frames predicted from MD simulation. These 10 frames capture the evolution of the potential energy fields, thereby offering a detailed visualization of defect dynamics (high-energy atoms are displayed in red color, representing surfaces or other defects that deviate from the perfect crystal structure that is visualized in blue). Figure 2(*c*) shows a larger number of samples of the type of microstructures used in the training set. All input microstructures feature a single crack, varied randomly in size, shape, and orientation, and all cracks are elliptical.

Figure 3(*a*) shows a summary of the neural network architecture used in this study, whereby a microstructure encoder is fed to two successive U-Nets (Unet 1 and Unet 2). Each U-Net produces a series of 10 frames (selected here for a demonstration of the method, but can be increased in principle), scaling up to a final resolution of 128 × 128 field predictions. To provide additional details to the formulation of the diffusion model, Fig. 3(*b*) depicts the denoising process that is learned. The top defines a Markov chain operator q that successively adds Gaussian noise (according to a defined noise schedule that defines how much noise ε_i is added at each step i), translating the original data X_0 (left) into pure noise (right), X_F via: whereby

$$X_{i+1} = q(X_i) \tag{1}$$

In a diffusion model, a deep neural network is trained to *reverse* this process, identifying a reverse operator p that maximizes the likelihood of the training data, thereby offering a means to translate noise to solutions

$$X_{i-1} = p(X_i) \tag{2}$$

The deep neural network trained results in the reverse operator p, which can then be used to systematically convert Gaussian noise into a proper field solution and, in our case, a set of dynamical frames that reflect the temporal process of cracking. Since all X_i have the same dimensionality, the input and output of p must also be of the same dimension and hence the use of a U-Net architecture. There are additional details in the exact formulation of the model, including the encoding of the time-step in the neural network (since the amount of noise changes at every denoising step, the model needs to understand where in the process it is situated in order to predict the correct amount of noise removal). The reader is referred to these papers for additional mathematical details [43–46].

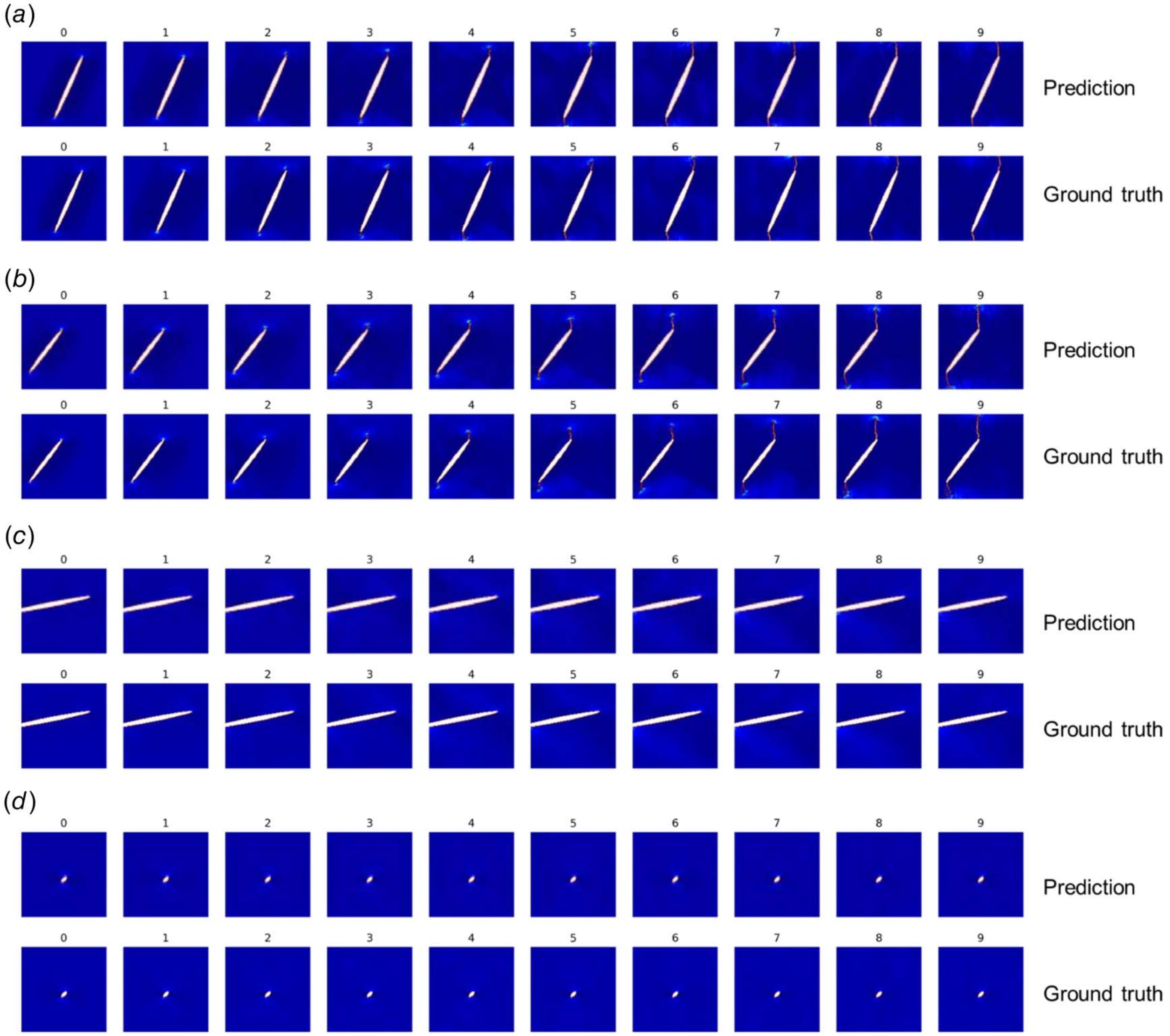

Fig. 4 Sample neural network prediction results and comparison with ground truth obtained from MD simulation, for validation data. As can be seen, the model accurately predicts whether a crack not only nucleates but also captures the dynamics of fracture. (*a*) and (*b*) Cases where fractures initiate, due to the presence of a large crack exposed to loading orthogonal to the long defect axis. (c) and (*d*) Cases where cracks are oriented largely in the direction of applied load (c) or too small to nucleate further damage (*d*).

Figure 4 shows sample neural network prediction results, and a comparison with ground truth obtained from MD simulation, for validation data. As can be seen, the model accurately predicts whether a crack not only nucleates but also captures the dynamics of fracture. Figures 4(*a*) and 4(*b*) show cases where fractures initiate, due to the presence of a large crack exposed to loading orthogonal to the long defect axis. Figures 4(*c*) and 4(*d*) represent cases where cracks are oriented predominantly in the direction of stress or too small to nucleate further damage.

Figure 5 provides a detailed analysis of testing the predictive capacity of the model for cases with multiple cracks, not seen in the training process. Figure 4(*a*) shows a multi-crack case where nucleation of fracture does not occur. Conversely, Fig. 4(*b*) shows a case where cracks nucleate. Figure 4(*c*) shows a direct comparison of the dynamics of fracture, for the prediction and ground truth. While some differences emerge, the overall damage spreading mechanisms and dynamics are predicted well.

For further analysis, we analyze the statistics of the potential energy field for all validation cases, with results shown in Fig. 6. Figure 6(*a*) shows predicted versus ground truth standard deviation and Fig. 6(*b*) predicted versus ground truth variance. Excellent agreement is found, with R2 values of ~0.99 for both cases.

Next, Fig. 7 depicts an analysis of even more complex validation cases. Here, the analysis is only shown for the final frame. The input microstructures considered here are vastly distinct from the types of cracks in the training set. Still, the model makes good predictions and good agreement with respect to dynamics of fracture, and initiation location of the damage. Another important observation is that the model has learned to identify surface atoms; this can be seen, for instance, in Fig. 7(*d*) where the interior of the heart-shaped defect is clearly surrounded by a row of dark shaded/red-colored atoms indicating the presence of a material surface. It is further noted that Fig. 7(*f*) shows a complex spider web-like microstructure (more details on that case below, including how this microstructure was generated using a deep neural network). A first important observation about this case is that both prediction and ground truth show that no damage is predicted to occur. We believe that this is likely due to a flaw-tolerance size effect since the filaments in the architected material are only a few atomic spacings large. The column on the right shows, for each case, a normalized histogram of the potential energy field.

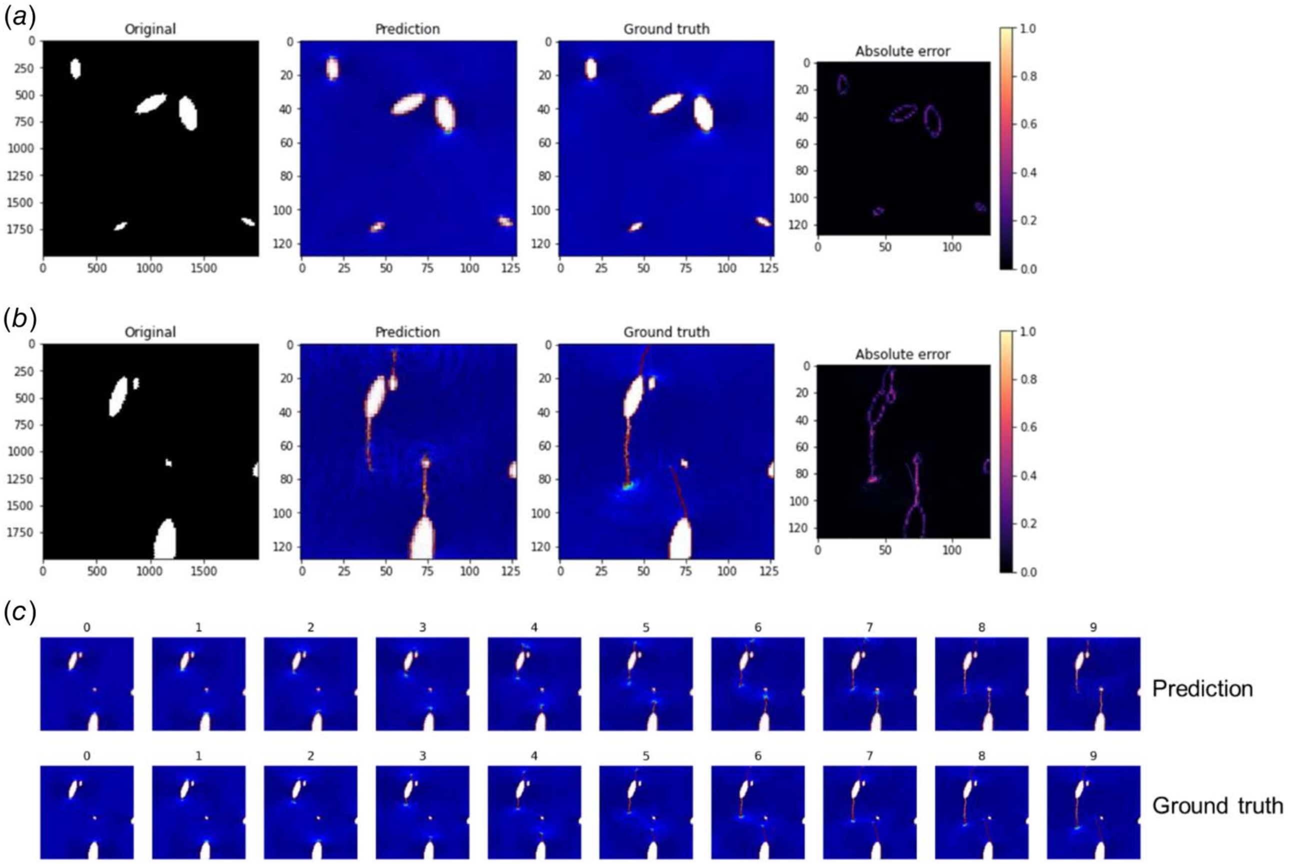

Fig. 5 Testing the predictive capacity of the model for cases with multiple cracks, not seen in the training process. Panel (*a*) shows a multi-crack case where nucleation of fracture does not occur. Conversely, panel (b) shows a case where cracks nucleate. Both panels (*a*) and (*b*) show the final predicted result (frame 10). (*c*) A direct comparison of the dynamics of fracture, for the prediction and ground truth. Certain differences can be observed, such as slightly different locations of crack nucleation and stochastic variations of the precise crack path. This could be a subject of further study and investigation, especially to explore the types of variations seen in the MD simulation itself (which show stochastic results themselves).

We now go into a bit more depth and analyze two cases that are included in Fig. 7. First, Fig. 8 shows a detailed analysis of the dynamics of fracture for one of the validation cases (Fig. 6(*c*)). The model has captured the details of the spreading mechanisms and dynamics well. Next, Fig. 9 presents a detailed analysis of the final potential energy field for the case presented in Fig. 7(*f*). Figure 9 shows the process by which the input microstructure was generated from a text prompt using the Latent Diffusion model, a text-to-image translation algorithm [24,47]. The produced image in “sketch” style, used to obtain an input that is largely bimodal in terms of colors, is processed by thickening the lines and increasing the contrast, yielding a

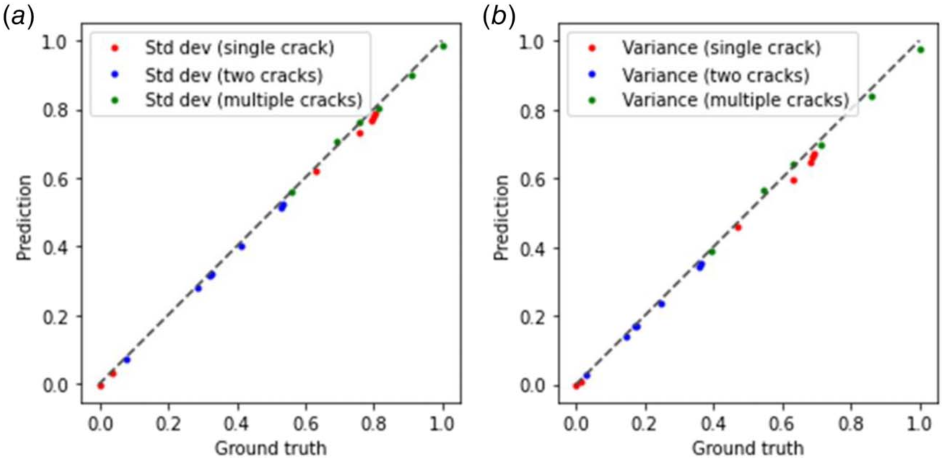

Fig. 6 As further analysis, we analyze the statistics of the potential energy field for all validation cases. (*a*) Predicted versus ground truth standard deviation. (*b*) Predicted versus ground truth variance. Excellent agreement is found, with *R*2 values of ∼0.99 for both cases.

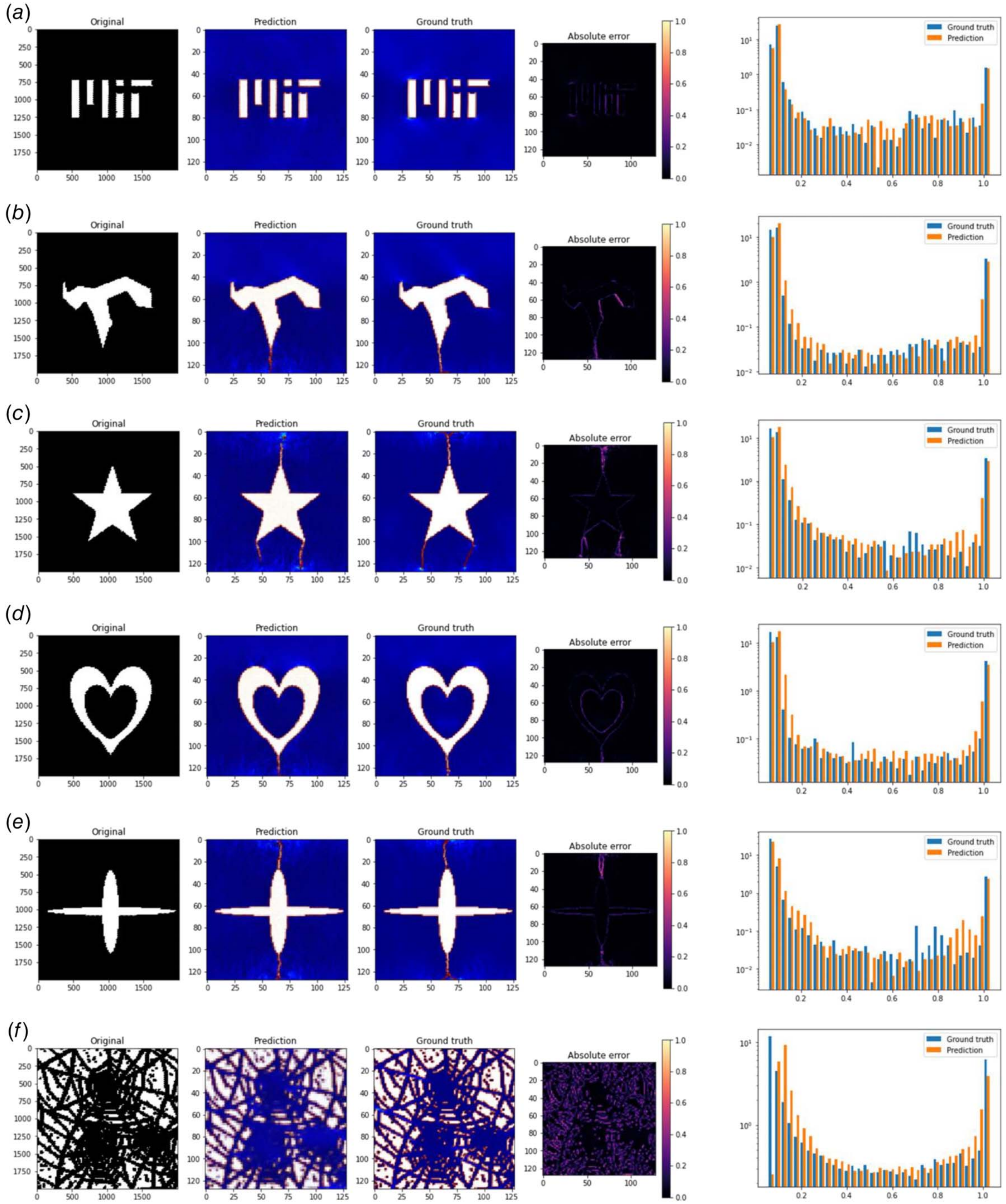

Fig. 7 Analysis of more complex validation cases; analysis shown for the final frame. The input microstructures considered here are vastly distinct from the types of cracks in the training set. Still, the model makes good predictions and good agreement with respect to dynamics of fracture, and initiation location of the damage. Note, panel (*f*) shows a complex spider web-like microstructure. Both prediction and ground truth show that no damage is predicted to occur likely due to a flaw-tolerance size effect since the filaments in the architected material are only a few atomic spacings large. The column on the right shows, for each case, a normalized histogram of the potential energy field.

black-and-white image (black = material, white = no material; consistent with the notation used for the other cases considered, e.g., shown in Fig. 2). This simple example, which could be expanded upon in future work, shows how a dynamic diffusion model as used here could work hand-in-hand with a generative text-to-image model to produce interesting new designs, which can then be assessed against a particular material function or objective function.

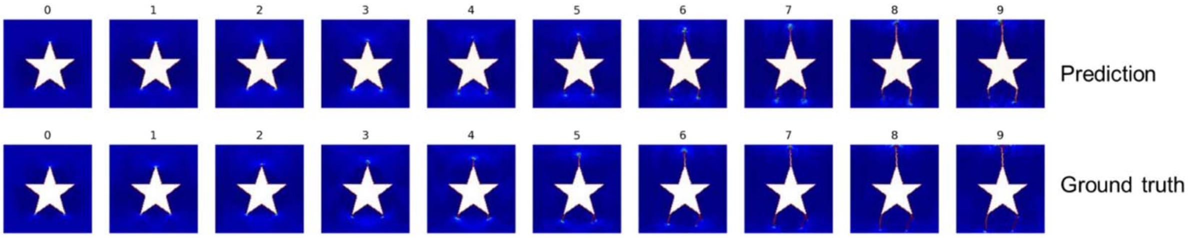

Fig. 8 Detailed analysis of the dynamics of fracture for one of the validation cases (Fig. 7(c))

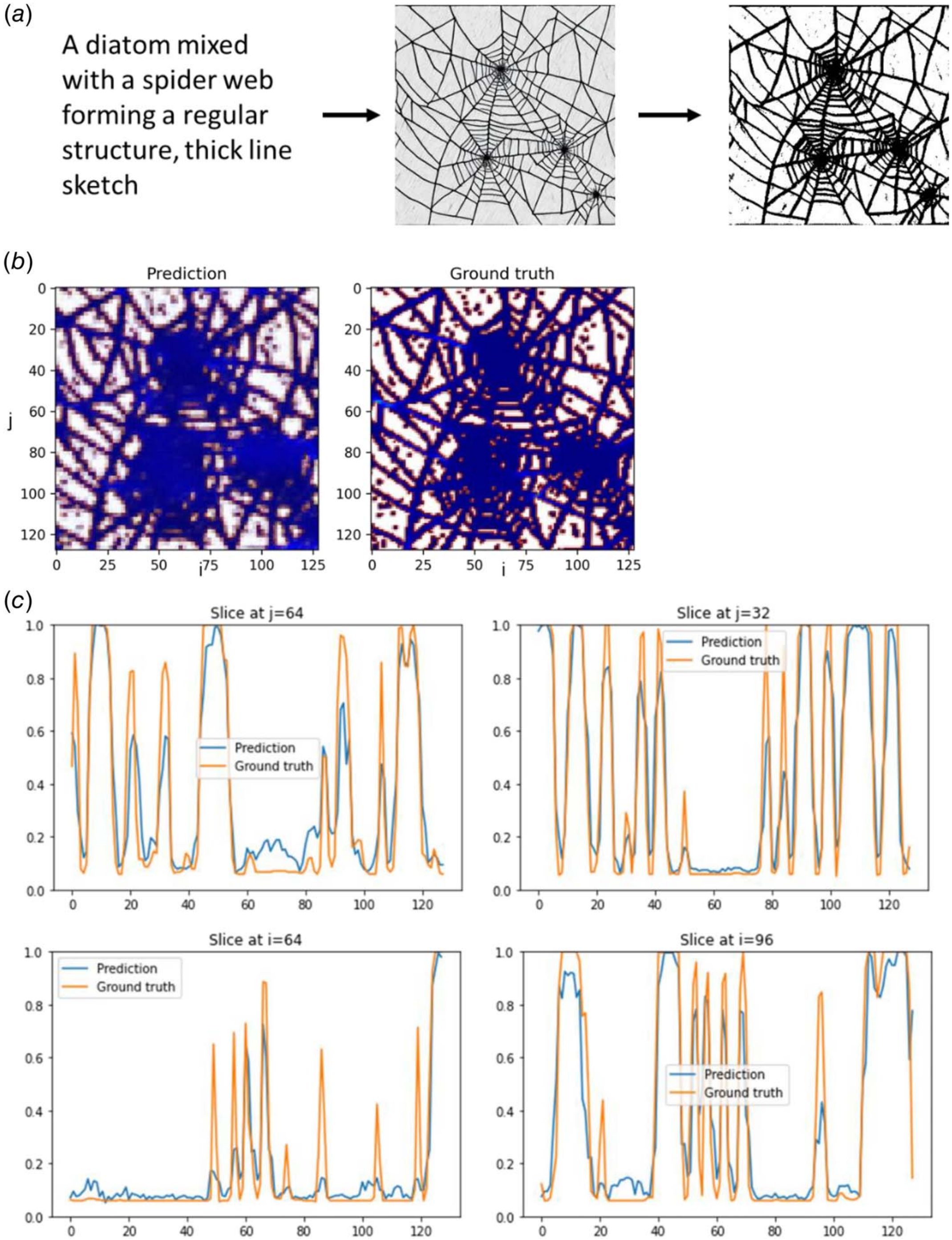

Fig. 9 Detailed analysis of the final potential energy field for the case presented in Fig. 7(*f*). (*a*) Generation of the input microstructure from a text prompt using the latent diffusion model [24,47]. The produced image in "sketch" type is processed by thickening the lines and increasing the contrast, yielding a black-and-white image (black = material, white = no material; consistent with the notation used for the other cases considered, e.g., shown in Fig. 2).

3 Conclusion

Dynamic fracture has been an area of intense interest in the materials modeling community. Here, we have shown that the aggregate complexity of the results of MD simulations can be effectively learned and upscaled using a diffusion model. This model is able to predict a series of frames that reflect the dynamics of microstructural change of a material, here examined for potential energy fields used to characterize the presence and geometry of cracks and other damage.

The model was shown to generalize well. In spite of being trained on a small data set with simple single cracks (Figs. 2(*b*) and 2(*c*)), the model has a capacity to make predictions for microstructures vastly different from the training set. This has been shown in Figs. 5–9, for instance. We analyzed not only the spreading dynamics of cracks, but also considered the effective statistical properties of the potential energy fields produced by the model. As expected in MD simulations, there are stochastic variations of the precise path a crack takes. This can be easily validated by studying the error plots included in the various figures. More work could be done to explore the nature of these statistical variations, especially when cracking is unstable (e.g., deviating from a straight line), as is often the case.

The most interesting performance assessment of the model is seen for the cases considered in Fig. 7, where we predicted damage evolution for microstructures that were vastly different from the training set. Cracks were no longer ellipsoidal, but took on various shapes including hand-drawn ones, text-generated microstructures, and furthermore, featuring straight lines that the model had never seen during training. While some differences are clearly visible, the results are remarkably good considering the small data set used to train. We anticipate that using a larger more diverse data set can offer significantly improved performance, and even consider a variety of other prompts beyond simply the microstructure, but also including boundary conditions. This could easily be realized by providing multiple input layers beyond a single microstructural image, but also including different material types, chemical details, concentration information, or locations of displacement, force, and other boundary conditions.

Finally, we explored one case where the input microstructure was not constructed mathematically but instead produced by a generative diffusion model (Fig. 9). This resulted in a significantly complex material design but the model was still able to yield good results. The method opens a variety of new directions in mechanics, especially converting design data into dynamical or higher-dimensional information. The formulation of a diffusion model, rooted in statistical physics (Fig. 3(*b*)), may offer added benefits toward use in engineering.

There are many ways by which the model could be improved. The scope of the study reported in this paper was deliberately focused on a very small data set to explore how predictive a model could be under this constraint, but increasing the dataset for training—or exploring the use of fine-tuning or transfer learning to predict distinct dynamical processes—could be a useful future research direction. Other potential steps to be taken include an increase in the number of denoising steps. In our study, we used 10, but better results are expected with ∼64 steps, which as shown in Ref. [46] is still significantly reduced from 1,000 steps or more as done in the original paper [43]. One may also explore to deepen the U-Net architecture and increase the size of the neural network by adding additional attention modules or more attention heads. These and other avenues are left to future work.

4 Materials and Methods

4.1 Molecular Dynamics Simulations. We use a simple fracture model to generate the dataset. The data set consists of an input microstructure and a series of images that capture the evolution of the potential energy field [10,48,49] (the simulation setup is similar as reported in, with a higher strain applied to drive crack initiation [39]). We consider the geometry shown in Figs. 1(*a*) and 1(*b*), depicting a simple triangular lattice with a 12:6 LJ interatomic interactions [50,51] under uniaxial loading, where interatomic interactions as a function of interatomic distance r_{ij} are defined by

$$\phi(r_{ij}) = 4\,\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] \tag{3}$$

while the LJ potential is a simplistic representation of atomic bonding in materials, its use in a 2D hexagonal geometry is known to yield a brittle material [10,51] and hence a good choice for the proof of concept study done here.

The input geometry features a series of randomly located and oriented cracks of different sizes and aspect ratios, as shown in Fig. 1(*b*), showing also the fixed boundaries in which forces and velocities of atoms are set to zero to allow for the application of strain for mechanical loading. The system is confined to 2D atomic motions in the *x*- and *y*-directions. The LJ interatomic potential parameters used in the generation of the data set, as defined in Eq. (2) (see Ref. [39] for further details) are $\varepsilon = 1$ and $\sigma = 1$ with a cutoff radius of $r_{\text{cut}} = 1.2$ (all simulations carried out in nondimensional units).

The MD simulations are carried out using the LAMMPS simulation package [52]. All samples are exposed to a homogeneous uniaxial tensile strain of 2% in the *x*-direction, applied homogeneously using the LAMMPS "deform erate" command.

4.2 Data Set Generation. The resulting field data obtained from MD simulations are visualized using matplotlib. We save images of both the input microstructure and the resulting potential energy field to generate the data sets (only atoms in the mobile region are saved into images). We use a "jet" colormap, with a range between potential energy values of −3 (light shade/blue) to −2.9 (dark shade/red). The time-step used in the MD simulation is $\Delta t = 0.0035$, and frames are output every 1000 simulation steps after loading is applied. The simulation is run for a total of 10,000 steps, resulting in 10 frames.

Input and output images are stored in CSV files and then used by the data reader for processing in the machine learning code for training, testing, and validation or general inference.

The training set used to train the model includes *only* cases with a larger number of randomly situated cracks (see Supplementary Information). Validation cases include both very small/single crack cases or scenarios with two and a much larger number of defects and with a different crack geometry than what was used in training.

The final training set used for the neural network training is based on ∼630 images total for the computational experiments, 90% of which is used for training and the rest for testing. A sample data set can be accessed online.[1]

We further create a set of validation examples consisting of single cracks, two cracks, and a larger number of cracks. We also consider a variety of cases with very different crack geometries than those used in the training set.

4.3 Progressive Diffusion Model. The dynamical diffusion model is based on the Imagen architecture that introduced a multistep progressive approach to translation and upscaling of input to output [13]. Instead of conditioning the input on text prompts, we condition the predictions on microstructure embeddings. The model consists of two U-Net architectures composed of convolutional and transformer layers with skip connections (Fig. S2 available in the Supplemental Materials on the ASME Digital Collection for details of all layers used in the neural network): Unet 1 (constructed based on [45]) and Unet 2 (constructed as Efficient U-Nets as proposed in [53]). U-Nets are a type of neural network that features the same input dimension as the output

[1] https://www.dropbox.com/s/6i301iwnv9ljnw9/dynamic_fracture_small.zip?dl=0

dimension, commonly used in problems such as image segmentation. The U-Nets implemented here feature a more complex architecture as described below and in Ref. [13], including the use of ResNet blocks, attention modules, and skip connections.

The U-Nets are used to translate the input microstructure coding into the final field output, over successive stages of translation and upscaling, ultimately reaching a resolution of 128 × 128 with three color channels, and producing a series of simulation frames. The loss is the L2 distance between the actual, added noise ε_i and the predicted added noise ε'_i. That way, the trained model can predict the added noise. Knowing this quantity then allows us to realize a numerical solution to the problem stated in Eq. (2), which is used to generate the next iteration of the denoised image

$$X_{i-1} = X_i - \varepsilon'_i \tag{4}$$

In Eq. (4), the image X_i at step i is transformed by removing the noise ε'_i. This process is performed iteratively; whereas the neural network predicts, given the current state X_i, the noise to be removed (Fig. S4 available in the Supplemental Materials on the ASME Digital Collection).

We use the improved noise schedule, sampling, and training processes proposed in Ref. [46] since it provides us with enhanced and computationally efficient predictions, specifically obtaining results within just 10 denoising steps for each of the U-Nets. Table 1 provides details about the model architecture parameters, and Figure S2 (available in the Supplemental Materials on the ASME Digital Collection) shows a detailed PyTorch model readout to reveal the entire architecture. The implementation is based on the code published in Ref. [54].

Table 1 Parameters used in the progressive transformer diffusion model (parameters for the two U-Nets, the integrated architecture, and additional parameters are provided)

Neural network component	Parameter	Value
Unet 1	Dimension	64
	Dimension multipliers	1, 2, 4, 8
	Resnet blocks	1
	Attention heads	8
	Feedforward multiplier	2.0
Unet 2	Dimension	64
	Dimension multipliers	1, 2, 4, 8
	Resnet blocks	1
	Attention heads	8
	Feedforward multiplier	2.0
Integrated architecture (comprised of Unet 1 and Unet 2)	Image sizes (each frame)	64 × 64 128 × 128 (each with 3 color channels)
	Frames predicted	10
	Cond. drop probability	0.1
	Sample steps (for Unet 1, Unet 2	10, 10
	$\sigma_{\min}$	0.002
	$\sigma_{\max}$ (for Unet 1, Unet 2)	80, 160
	σ_{data}	0.5
	ρ	7
	P_{mean}	−1.2
	P_{std}	1.2
	P_{churn}	80
	$S_{\text{t,min}}$	0.05
	$S_{\text{t,max}}$	50.
	S_{noise}	1.003
Additional parameters	Optimizer and parameters	Adam learning rate = 1E-4, epsilon = 1e-8, betas = (0.9,0.99)
	Batch size	5

Note: See also Figure S2 (available in the Supplemental Materials on the ASME Digital Collection).

The microstructure encoder scales the pixel data values to be between −1…1 and feeds each of the three-color channels to the neural networks in the embedding dimension. Since the input microstructures solely consist of white (no material) and black (material) building blocks, the input embeddings consist solely of (1,1,1) and (−1,−1,−1) tensors for each building block that makes up the material microstructure. In some sense, this is viewed as a language prompt that provides a series of letters defined at every pixel of the input geometry—"W" for void, "B" for material, to the model. Future work can expand on this easily and provide either more material choices or gradations or text prompts that describe certain types of boundary conditions (see also discussion in the main text).

The total number of parameters in the model is 101,506,434. Unet1 features 41,507,444 trainable parameters, and Unet2 59,998,990 trainable parameters.

The model is trained successfully whereby each of the U-Nets is trained individually (first Unet 1, then Unet 2). Training performances are shown in Fig. S3 (available in the Supplemental Materials on the ASME Digital Collection). Unet 1 is trained for 5.5 K steps, and Unet 2 is trained for 12 K steps.

The microstructure shown in Figs. 7(*f*) and 9 are generated from a text prompt using a latent diffusion model, as reported in [24,47]. Weights were used as reported in Ref. [55]. The text prompt used is provided in the main text.

4.4 Data Analysis.

For the discussions in this paper, we limit the exploration to the potential energy fields. Slices of the potential energy field plots are obtained by converting the field data with three color channels into a grayscale image, by multiplying the unscaled data in the three color channels with [0.299, 0.5870, 0.114] and then adding them up. This yields scalar field data of the potential energy field and microstructure, respectively, that measures pixel-level intensity (the operation is done consistently for ground truth and prediction data).

All potential energy values are normalized such that the smallest value overall is 0, and the largest 1, for simpler comparison in the field plots and the histogram analysis.

5 Software Versions and Hardware

We use PYTHON 3.8.12, PyTorch 1.10 with CUDA (CUDA version 11.6), and a NVIDIA RTX A6000 with 48 GB VRAM for training and inference.

Acknowledgment

We acknowledge support from ARO (Grant No.W911NF1920098), NIH (Grant Nos. U01EB014976 and 1R01AR077793), and ONR (Grant Nos. N00014-19-1-2375 and N00014-20-1-2189). Additional support is provided by USDA (Grant No. 2021-69012-35978). Support from the MIT-IBM Watson AI Lab is acknowledged.

Conflict of Interest

The author declares no conflict of interest.

Data Availability

A sample data set is included as a link to a repository in Materials and Methods. Other data and codes are either available on GitHub or will be provided by the author based on reasonable request.

Author Contributions

MJB conceived the study, developed, and trained the neural network and associated data analysis, including fracture models, and wrote the paper.

References

[1] Tong, Q., and Li, S., 2020, "A Concurrent Multiscale Study of Dynamic Fracture," Comput. Methods Appl. Mech. Eng., **366**.

[2] Anderson, T. L., 2005, *Fracture Mechanics: Fundamentals and Applications*, Taylor & Francis, London.

[3] Jung, G., Qin, Z., and Buehler, M. J., 2015, "Molecular Mechanics of Polycrystalline Graphene With Enhanced Fracture Toughness," Extreme Mech. Lett., **2**, pp. 52–59.

[4] Buehler, M. J., Tang, H., van Duin, A. C. T., and Goddard, W. A., 2007, "Threshold Crack Speed Controls Dynamical Fracture of Silicon Single Crystals," Phys. Rev. Lett., **99**(16).

[5] Gao, H., Buehler, M. J., and Abraham, F. F., 2005, "Hyperelasticity in Dynamic Fracture: The Characteristic Energy Length Scale," Proceedings of the 11th International Conference on Fracture 2005, Turino, Italy.

[6] Freund, L. B., 1990, *Dynamic Fracture Mechanics*, Cambridge University Press.

[7] Buehler, M. J., Tang, H., van Duin, A. C. T., and Goddard, W. A., 2006, "Threshold Crack Speed in Dynamic Fracture of Silicon," Materials Research Society Symposium Proceedings, Boston, MA.

[8] Buehler, M. J., and Gao, H., 2006, "Dynamical Fracture Instabilities due to Local Hyperelasticity at Crack Tips," Nature, **439**(7074), pp. 307–310

[9] Jung, G. S., Wang, S., Qin, Z., Zhou, S., Danaie, M., Kirkland, A. I., Buehler, M. J., and Warner, J. H., 2019, "Anisotropic Fracture Dynamics Due to Local Lattice Distortions," ACS Nano, **13**(5), pp. 5693–5702

[10] Buehler, M. J., 2008, "*Atomistic Modeling of Materials Failure*," *Atomistic Modeling of Materials Failure*, Springer US, Boston, MA.

[11] Buehler, M. J., 2006, "Atomistic and Continuum Modeling of Mechanical Properties of Collagen: Elasticity, Fracture, and Self-assembly," J. Mater. Res., **21**(8), pp. 1947–1961.

[12] Ramesh, A., Dhariwal, P., Nichol, A., Chu, C., and Chen, M., 2022, "Hierarchical Text-Conditional Image Generation with CLIP Latents," arXiv preprint

[13] Saharia, C., Chan, W., Saxena, S., Li, L., Whang, J., Denton, E., Ghasemipour, S. K. S., et al., 2022, "Photorealistic Text-to-Image Diffusion Models With Deep Language Understanding," arXiv preprint

[14] Nichol, A., Dhariwal, P., Ramesh, A., Shyam, P., Mishkin, P., McGrew, B., Sutskever, I., and Chen, M., 2021, "GLIDE: Towards Photorealistic Image Generation and Editing With Text-Guided Diffusion Models," arXiv preprint

[15] Yang, Z., and Buehler, M. J., 2021, "Words to Matter: De Novo Architected Materials Design Using Transformer Neural Networks," Front. Mater., **8**, p. 740754.

[16] Hinton, G., and Zemel, R., 1994, "Autoencoders, Minimum Description Length and Helmholtz Free Energy," Adv. Neural Inf. Process. Syst., **6**(3–10).

[17] Dong, G., Liao, G., Liu, H., and Kuang, G., 2018, "A Review of the Autoencoder and Its Variants: A Comparative Perspective From Target Recognition in Synthetic-Aperture Radar Images," IEEE Geosci. Remote Sens. Mag., **6**(3), pp. 44–68.

[18] Goodfellow, I. J., et al., 2014, "Generative Adversarial Networks," arXiv preprint

[19] Lebese, T., Mellado, B., and Ruan, X., 2021, "The Use of Generative Adversarial Networks to Characterise New Physics in Multi-Lepton Final States at the LHC," Int. J. Mod. Phys. A

[20] Makoś, M. Z., Verma, N., Larson, E. C., Freindorf, M., and Kraka, E., 2021, "Generative Adversarial Networks for Transition State Geometry Prediction," J. Chem. Phys., **155**(2), p. 024116.

[21] Crowson, K., Biderman, S., Kornis, D., Stander, D., Hallahan, E., Castricato, L., and Raff, E., 2022, "VQGAN-CLIP: Open Domain Image Generation and Editing With Natural Language Guidance," arXiv preprint.

[22] 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, Seattle, WA, pp. 12873–12883.

[23] Marcus, G., Davis, E., and Aaronson, S., 2022, "A Very Preliminary Analysis of DALL-E 2," arXiv preprint arXiv:2204.13807.

[24] Rombach, R., Blattmann, A., Lorenz, D., Esser, P., and Ommer, B., 2022, "High-Resolution Image Synthesis With Latent Diffusion Models," Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, pp. 10684–10695.

[25] Haim, N., Feinstein, B., Granot, N., Shocher, A., Bagon, S., Dekel, T., and Irani, M., 2022, "Diverse Video Generation From a Single Video," arXiv preprint

[26] Yan, W., Zhang, Y., Abbeel, P., and Srinivas, A., 2021, "VideoGPT: Video Generation using VQ-VAE and Transformers," arXiv preprint arXiv:2104.10157.

[27] Aldausari, N., Sowmya, A., Marcus, N., and Mohammadi, G., 2022, "Video Generative Adversarial Networks: A Review," ACM Comput. Surv., **55**(2), pp. 1–25.

[28] Anderson, B. D. O., 1982, "Reverse-Time Diffusion Equation Models," Stoch Process Their Appl., **12**(3), pp. 313–326.

[29] Lim, S., Luan, H., Zhao, S., Lee, Y., Zhang, Y., Huang, Y., Rogers, J. A., and Ahn, J. H., 2020, "Assembly of Foldable 3D Microstructures Using Graphene Hinges," Adv. Mater., **32**(28), p. 2001303.

[30] Zhai, Z., Wang, Y., Lin, K., Wu, L., and Jiang, H., 2020, "In Situ Stiffness Manipulation Using Elegant Curved Origami," Sci. Adv., **6**(47).

[31] Lejeune, E., 2020, "Mechanical MNIST: A Benchmark Dataset for Mechanical Metamodels," Extreme Mech. Lett., **36**, p. 100659.

[32] Yuan, L., Park, H. S., and Lejeune, E., 2022, "Towards Out of Distribution Generalization for Problems in Mechanics," Comput. Methods Appl. Mech. Eng., **400**, p. 115569.

[33] Qin, H., 2020, "Machine Learning and Serving of Discrete Field Theories," Sci. Rep., **10**(1), pp. 1–15.

[34] Tang, J., Geng, X., Li, D., Shi, Y., Tong, J., Xiao, H., and Peng, F., "Machine Learning-Based Microstructure Prediction During Laser Sintering of Alumina," Sci. Rep., **11**(1), pp. 1–11.

[35] Kauwe, S. K., Graser, J., Murdock, R., and Sparks, T. D., 2020, "Can Machine Learning Find Extraordinary Materials?," Comput. Mater. Sci., **174**, p. 109498.

[36] Schmidt, J., Marques, M. R. G., Botti, S., and Marques, M. A. L., 2019, "Recent Advances and Applications of Machine Learning in Solid-State Materials Science," npj Comput. Mater., **5**(1), pp. 1–36.

[37] Reyes, K. G., and Maruyama, B., 2019, "The Machine Learning Revolution in Materials?," MRS Bull., **44**(7), pp. 530–537.

[38] Buehler, M. J., 2022, "FieldPerceiver: Domain Agnostic Transformer Model to Predict Multiscale Physical Fields and Nonlinear Material Properties Through Neural Ologs," Mater. Today, **57**, pp. 9–25.

[39] Buehler, E. L., and Buehler, M. J., 2022, "End-to-End Prediction of Multimaterial Stress Fields and Fracture Patterns Using Cycle-Consistent Adversarial and Transformer Neural Networks," Adv. Biomed. Eng., **4**, p. 100038.

[40] 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).

[41] Hsu, Y. C., Yu, C. H., and Buehler, M. J., 2020, "Using Deep Learning to Predict Fracture Patterns in Crystalline Solids," Matter, **3**(1), pp. 197–211.

[42] Lew, A. J., and Buehler, M. J., 2021, "A Deep Learning Augmented Genetic Algorithm Approach to Polycrystalline 2D Material Fracture Discovery and Design," Appl. Phys. Rev., **8**(4), p. 041414.

[43] Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N., and Ganguli, S., 2015, "Deep Unsupervised Learning Using Nonequilibrium Thermodynamics," Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, Vol. 3, pp. 2246–2255.

[44] Ho, J., Jain, A., and Abbeel, P., 2020, "Denoising Diffusion Probabilistic Models," Adv. Neural Inf. Process. Syst., **33**, pp. 6840–6851.

[45] Nichol, A., and Dhariwal, P., 2021, "Improved Denoising Diffusion Probabilistic Models," arXiv preprint

[46] Karras, T., Aittala, M., Aila, T., and Laine, S., 2022, "Elucidating the Design Space of Diffusion-Based Generative Models," arXiv preprint arXiv:2206.00364.

[47] "CompVis/Latent-Diffusion: High-Resolution Image Synthesis With Latent Diffusion Models," arXiv preprint

[48] Wang, X., Tabarraei, A., and Spearot, D. E., 2015, "Fracture Mechanics of Monolayer Molybdenum Disulfide," Nanotechnology, **26**(17), p. 175703.

[49] Buehler, M. J., van Duin, A. C. T., and Goddard W. A., III2006, "Multiparadigm Modeling of Dynamical Crack Propagation in Silicon Using a Reactive Force Field," Phys. Rev. Lett., **96**(9), p. 095505.

[50] Xu, Z., and Buehler, M. J., 2010, "Interface Structure and Mechanics Between Graphene and Metal Substrates: A First-Principles Study," J. Phys. Conden. Matter, **22**(48), p. 485301.

[51] Allen, M. P., and Tildesley, D. J., 1987, *Computer Simulation of Liquids*, Clarendon, Oxford.

[52] Thompson, A. P., Aktulga, H. M., Berger, R., Bolintineanu, D. S., Brown, W. M., Crozier, P. S., in't Veld, P. J., et al., 2022, "LAMMPS—A Flexible Simulation Tool for Particle-Based Materials Modeling at the Atomic, Meso, and Continuum Scales," Comput. Phys. Commun., **271**, p. 108171.

[53] Saharia, C., Chan, W., Chang, H., Lee, C., Ho, J., Salimans, T., Fleet, D., and Norouzi, M., 2021, "Palette: Image-to-Image Diffusion Models," arXiv preprint

[54] Lucidrains/imagen-Pytorch: Implementation of Imagen, Google's Text-to-Image Neural Network, in Pytorch, https://github.com/lucidrains/imagen-pytorch

[55] Model weights used as published at: https://ommer-lab.com/files/latent-diffusion/nitro/txt2img-f8-large/