ARTICLE   OPEN

Check for updates

Deep learning model to predict fracture mechanisms of graphene

Andrew J. Lew [1,2], Chi-Hua Yu [1,3], Yu-Chuan Hsu[4] and Markus J. Buehler [1,5,6 ✉]

Understanding fracture is critical to the design of resilient nanomaterials. Molecular dynamics offers a way to study fracture at an atomistic level, but is computationally expensive with limitations of scalability. In this work, we build upon machine-learning approaches for predicting nanoscopic fracture mechanisms including crack instabilities and branching as a function of crystal orientation. We focus on a particular technologically relevant material system, graphene, and apply a deep learning method to the study of such nanomaterials and explore the parameter space necessary for calibrating machine-learning predictions to meaningful results. Our results validate the ability of deep learning methods to quantitatively capture graphene fracture behavior, including its fractal dimension as a function of crystal orientation, and provide promise toward the wider application of deep learning to materials design, opening the potential for other 2D materials.

INTRODUCTION

One of the most fundamental challenges in technological applications of nanomaterials is materials fracture[1]. However, the scientific foundations by which materials break is still poorly understood, especially 2D materials such as graphene. One way to model fracture is using molecular modeling, which can complement experimental methods such as HRTEM in an efficient way[2–4]. However, MD simulations are expensive, especially when a large number of simulations is required such as for materials design applications. Here, alternate methods to enable multiscaling of material behavior across levels of material representation is a critical frontier[5,6]. Indeed, machine learning (ML)—especially deep learning[7]—has developed into a powerful tool to capture complex physical phenomena[8] and offer a complementary approach to achieve multiscaling across spatial and temporal domains, in various areas of materials science including soft material design[9], proteins[10–15], and nanomaterials[16–18]. ML methods have begun application toward fracture, with predictions of how defects in graphene affect fracture stress[19].

Recently, a deep learning method for predicting fracture patterns in crystalline Lennard-Jones material has been developed[20]. This method uses a Convolutional Long Short-term Memory (ConvLSTM)-based model to tease out the spatiotemporal relationships underpinning fracture propagation from datasets generated by physics-based molecular dynamics (MD) simulations. The excellent predictive power of the model across various loading conditions and material orientations is promising for reducing the large computational burden of fracture prediction from traditional methods. In this paper, we apply this ML model to the fracture of a real, technologically relevant material—graphene, with significantly higher chemical complexity and covalent bond breaking at the core of its fracture mechanics[21–26]. We investigate the parameter calibration process necessary in obtaining meaningful fracture predictions, achieve predicted fractures with good agreement to MD, and further demonstrate the potential power of

an ML approach to material systems of increasing complexity and scale.

RESULTS AND DISCUSSION

MD studies of graphene fracture

We consider a graphene system, as depicted in Fig. 1a, 32×24 nm with sheet orientation rotated about the center by θ degrees. For MD simulations, boundaries are periodic in tensile direction and non-periodic in fracture and depth directions. During tensile tests, a 3-Å thick border is restricted to only allow motion in tensile direction. As has been shown in the literature, graphene fracture behavior is significantly dependent on its orientation[27], with documented differences in mechanical properties when loading along the armchair vs. zigzag directions[24]. To cover this chirality-dependent range of fracture behavior, loading orientations from purely armchair direction (0°) to purely zigzag direction (30°) in 10° increments are simulated, as shown in Fig. 1b.

As each fracture event is heavily affected by random molecular vibrations, 11 MD fracture simulations are run for each orientation, shown in Fig. 1c. These fracture paths are averaged over each orientation in order to illustrate a representative likely fracture envelope. We can clearly see a preference for the material to fracture along the zigzag edge in Fig. 1d, with the path of most likely fracture deflected downward as graphene orientation is rotated clockwise. This predominance of zigzag fracture edges is consistent with previous reporting in the literature[25]. Crack branching becomes more prevalent as the sample approaches 30°, at which point fracture paths going up or down at a 30° angle are equally likely due to the symmetry of both being zigzag edges. A comparison of stress–strain curves between our data and the example from the literature[27] in Fig. 1e shows our MD results behave as expected, as they are nearly identical. Details of the MD simulation setup and visualization of graphene fracture from the raw MD data are explained in "Methods."

[1]Laboratory for Atomistic and Molecular Mechanics (LAMM), Massachusetts Institute of Technology, Cambridge, MA, USA. [2]Department of Chemistry, Massachusetts Institute of Technology, Cambridge, MA, USA. [3]Department of Engineering Science, National Cheng Kung University, Tainan City, Taiwan. [4]Department of Civil Engineering, National Taiwan University, Taipei, Taiwan. [5]Center for Computational Science and Engineering, Schwarzman College of Computing, Massachusetts Institute of Technology, Cambridge, MA, USA. [6]Center for Materials Science and Engineering, Cambridge, MA, USA. ✉email: mbuehler@mit.edu
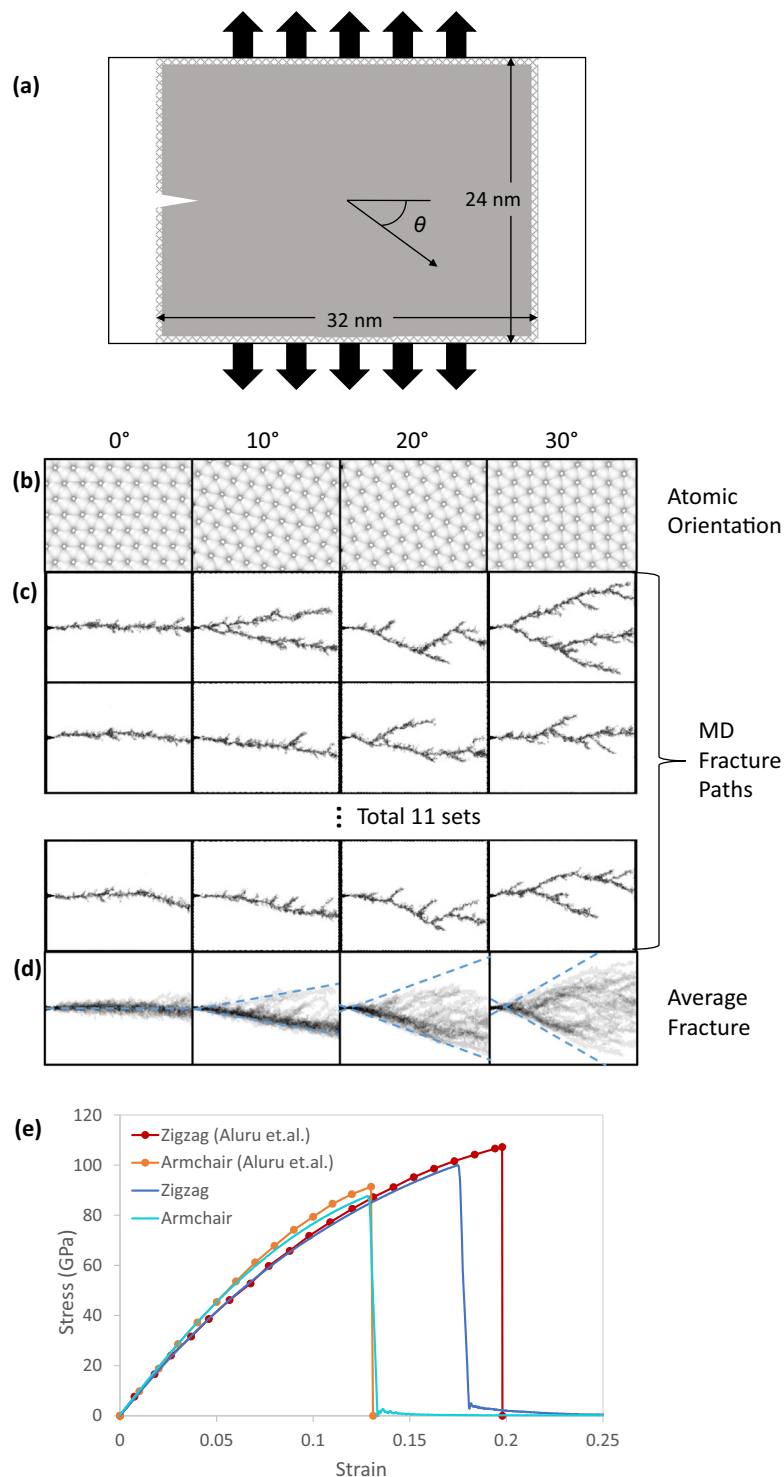
Fig. 1 **MD studies of graphene fracture. a** Graphene sample dimensions with periodic boundaries in only the tensile direction. The border atoms are restricted to only allow motion in the tensile direction. The fracture behavior of graphene was modeled across **b** four sheet orientations from armchair loading to zigzag loading. Each orientation was **c** fractured 11 separate times for a total for 44 samples. Averaging over each orientation illustrates fracture paths **d** fall within an envelope determined exactly by the rotation angle, biased toward fracture along zigzag edges. Stress–strain curves of **e** graphene fracture at both zigzag and armchair orientations show our MD results are nearly identical to previously published graphene crack simulations from Zhao et.al.[27] specifically focusing on the orientation-dependent elastic properties of graphene under uniaxial tension.

ML parameter optimization

The 44 graphene fracture MD simulations are subsequently used to train the ML model in capturing the complex orientation-dependent fracture behavior of graphene. The structure of the ML model is given in Fig. 2a and further elaborated in "Methods." Each fracture path image is split into segments of sequential input $X[t]$ and output $y[t]$ matrix pairs in order to focus on the spatiotemporal relationship behind each step of crack propagation. By
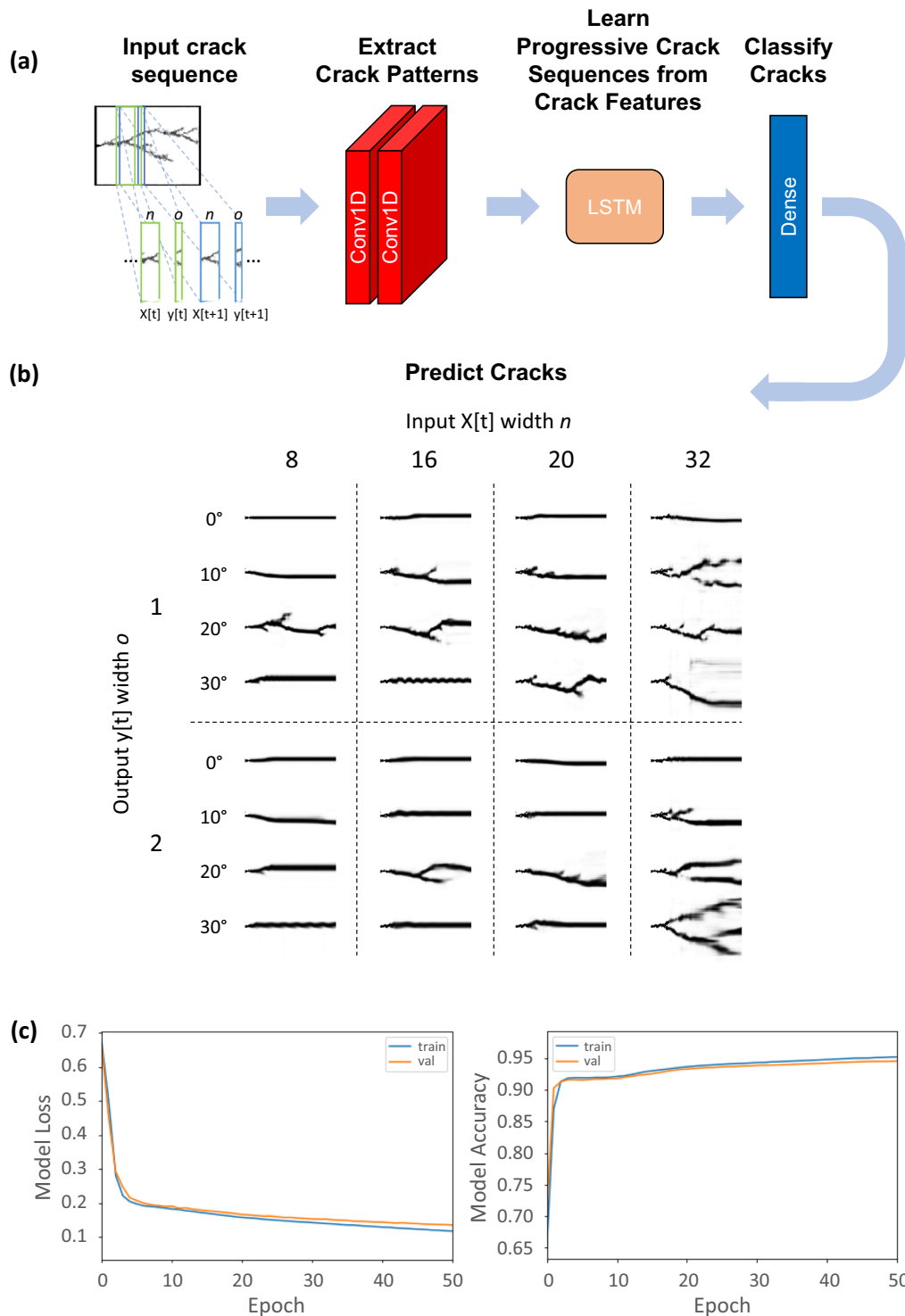
Fig. 2 ML parameter optimization. The ML model consists of multiple layers including **a** two convolutional layers to learn geometric features of crack slices, an LSTM layer to learn sequential relations between them, and a dense layer to classify the results. Each MD graphene fracture image is sliced into ML training data by **b** eight different combinations of input and output widths to yield eight sets of ML-predicted fracture paths. The set of parameters yielding predictions closest to MD fracture results is input width 32 output width 2. Comparing the **c** training and validation loss of the model shows close agreement without overfitting and a model accuracy near 95%.

varying the pixel width of the input and output matrices, dependencies at different length scales are probed. A systematic exploration of how various input and output widths affect the predicted fracture paths is provided in Fig. 2b and details of the slicing process are explained in "Methods."

Predicted fracture paths tend to become more complex and in closer alignment to MD results as the input width increases to 32, signifying graphene crack propagation has a large dependence on its history. Specifically, not accounting for enough fracture history leads to spurious predictions of straight horizontal line paths with

little branching. The accurate value of 32 is larger than the amount of history originally used for the Lennard-Jones model[20]. Subsequently, using an output width of two columns at a time rather than one provides better predictions closer to the average MD paths. Interestingly, this result has a structural link, as a two-column width corresponds to the graphene armchair unit cell length. Though increasing the width of the input and output slices decreases the number of training sets obtainable from each MD image, the quality of predictions greatly increases. This indicates that the quality of training pairs is much more important than the quantity of training pairs. The training history of the model using 32 column wide input segments and 2 column wide output segments for a total of 5544 training pairs has nearly 95% validation accuracy and close agreement between training and validation losses, as shown in Fig. 2c, indicating the training dataset is sufficient.

Evaluating predictive power of ML model

Comparing the calibrated ML-predicted fracture paths with averaged MD results side-by-side as in Fig. 3a illustrates how well the ML model has managed to learn the training cases. However, it is essential that the ML model be able to predict fracture paths of graphene orientations not used for training, else the model be overfit to the specific training cases and useless as a general tool. Importantly, this calibrated model can predict fracture behavior of orientations it never been exposed to, as in Fig. 3b, with good

qualitative agreement to MD in terms of both envelope angle and branching behavior.

As a quantitative measure of prediction efficacy, the fractal dimensions of average MD and predicted ML fracture paths are compared in Fig. 3c. The fractal dimension is a measure of complexity as a ratio of the change in detail to the change in scale[28], and while the ML paths trend toward underestimating this value compared to the real MD results, the underestimated discrepancy of <0.1 is consistent between most of the trained (0°, 10°, 20°) and test (5°, 15°, 25°) orientations. The existence of this 0.1 discrepancy in fractal dimension is visually apparent in how the ML-predicted paths are somewhat simplified compared to the average MD. This issue may be ameliorated by increasing the resolution limit (and correspondingly number of nodes) of the ML model. Regardless, the consistency of the discrepancy quantitatively signifies that the ML model is able to predict test orientations just as well as orientations it has trained on, and has indeed gathered a working general understanding of graphene fracture behavior.

Further quantitative comparisons in crack length and crack energy are shown in Fig. 3d. The overestimation in crack length is consistent with the fractal dimension discussion of how ML-predicted paths are somewhat visually simplified with heavier line weights—and thus more cracked pixels—than the more nuanced MD paths. This overestimation of crack length manifests as a slight overestimation of crack energy values following Eq. (1) where γ is the surface energy from MD, a is the aforementioned crack length,
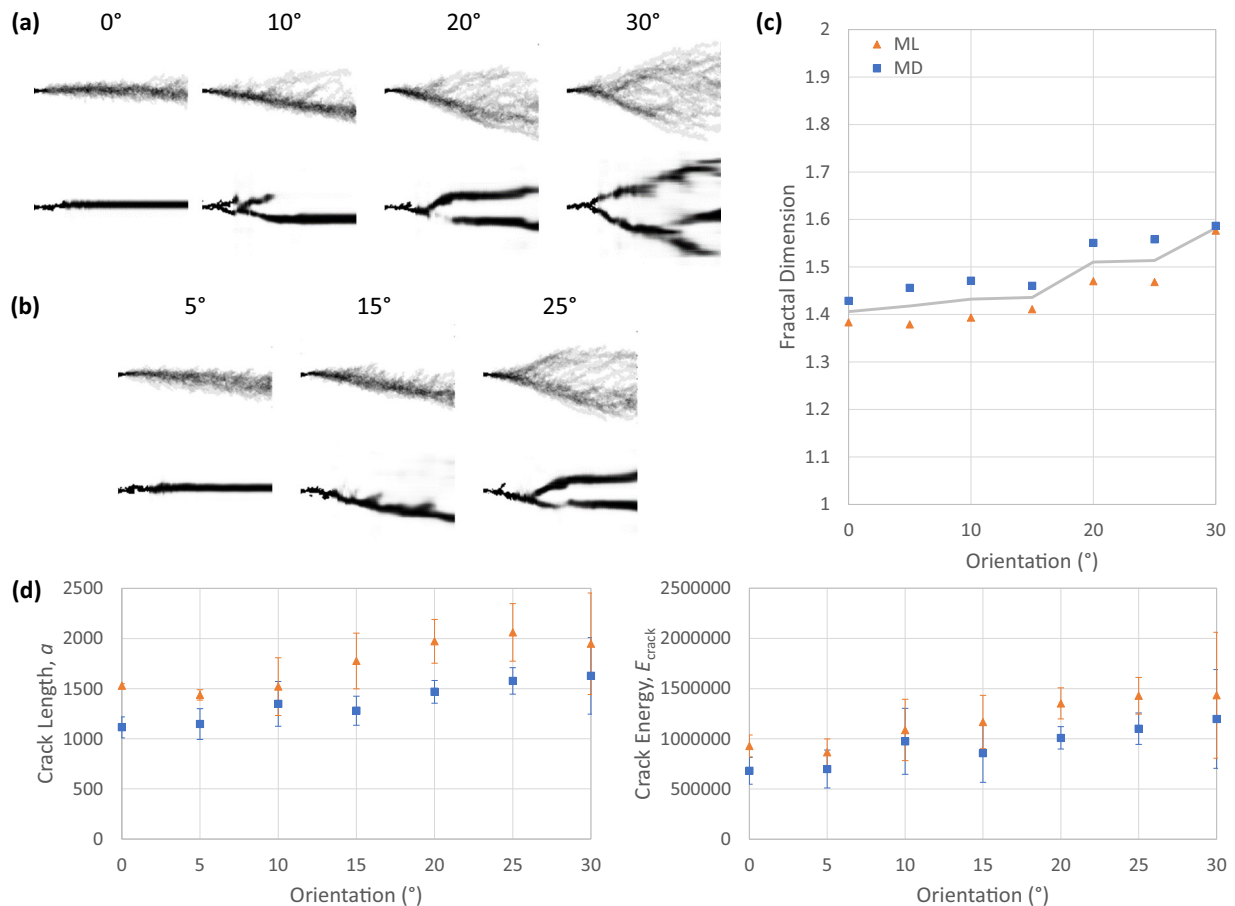

Fig. 3 Evaluating predictive power of ML model. Calibrated ML predictions line up not just with **a** graphene orientations used for the training data, but also **b** orientations the model has never seen before. The **c** fractal dimensions of ML-predicted fracture paths between the training and test orientations are of comparable accuracy, indicating that the model has correctly learned fracture behavior without overfitting to the specific training orientations. Quantitatively, **d** the fracture toughness values ML predicts only slightly overestimate crack energy values. Data points and error bars represent the mean and standard deviation over 11 sets of fracture paths.

and B is the thickness (where $B = 1$ in 2D materials)

$$E_{crack} = 2\gamma aB \qquad (1)$$

Generalized fracture prediction

The usefulness and generality of the model can be seen in its ability to reasonably predict crack nucleation of a pristine surface, as in Fig. 4a, despite being trained on purely pre-notched graphene samples. Furthermore, more complex graphene systems beyond single crystal sheets can be easily treated with this ML approach. Graphene bicrystal fracture is simulated with MD as a baseline comparison for the ML predictions shown in Fig. 4b. The left side of each graphene sample is kept at 0°, while the right side is rotated in 5° increments to form six total misorientation cases. Each misorientation is run 11 times as before, and averaged to obtain the general fracture behavior. Though the ML model is not trained on any bicrystal samples, both MD and ML show crack branching beginning at 15° misorientation. Within the MD results, the difference between 20°, 25°, and 30° misorientations is subtle, and the ML model correctly predicts very similar fracture paths for these three. The ability to predict bicrystal fracture indicates that a large part of fracture behavior near a boundary may be understood just by the orientations on either side. Beyond bicrystal cases, Fig. 4c illustrates how ML predictions can be applied to samples of longer length scales containing a variety of arbitrary graphene orientations. While not perfect, the model approaches a qualitatively reasonable prediction—crack deflection near a grain boundary.

In terms of how other defects affect fracture, we next focus on point defects. It has been reported in the literature[29] that nanocrystalline graphene exhibits a degree of flaw tolerance, able to retain its theoretical pristine strength even with small hole defects. We treat the effect of small hole defects on ML-predicted fracture paths by pre-labeling a square region of size $L \times L$ in the center of the fracture prediction images as "cracked" prior to the start of fracture prediction, as shown in Fig. 5a. Our ML model predicts that small hole defects barely perturb the crack path of single crystal graphene, but significant deviation occurs in both 0° and 30° orientations as the flaw size increases to 3.2 nm, as shown in Fig. 5b. At small flaw sizes, fracture strength of nanocrystalline graphene is reported to saturate at a plateau close to the theoretical pristine value of 26.86 GPa[29], unaffected by the presence of defects until they reach a critical size. Larger flaws then contribute to a decrease in strength consistent with the Griffith model. Intriguingly, the reported size for hole defects to begin deteriorating the strength of nanocrystalline graphene is just at this point where our ML predictions indicate a change in behavior, diameters of 3.2 nm[29] and above. The predicted change in fracture behavior of 0° and 30° oriented graphene, corresponding to both pure armchair and pure zigzag orientations, may be the mechanism by which nanocrystalline graphene loses its fracture strength. These results show the predictive power of our model to capture intricate nanomechanical effects previously only seen in physics-based simulations.

DISCUSSION

In summary, by tuning the binning sizes used to prepare training data for the ML model, predicted fracture paths can vary widely. It is important to include enough historical fracture data as input for the ML model in order for it to accurately determine where the crack will next go. It appears that more complex systems, such as our current graphene case, require more such historical information for accurate predictions than simpler systems, such as the Lennard-Jones lattices the model was originally developed on. The output crack size for each propagation step also impacts the overall form of predicted cracks, and it may be important to select a distance structurally relevant to the material under study. Once
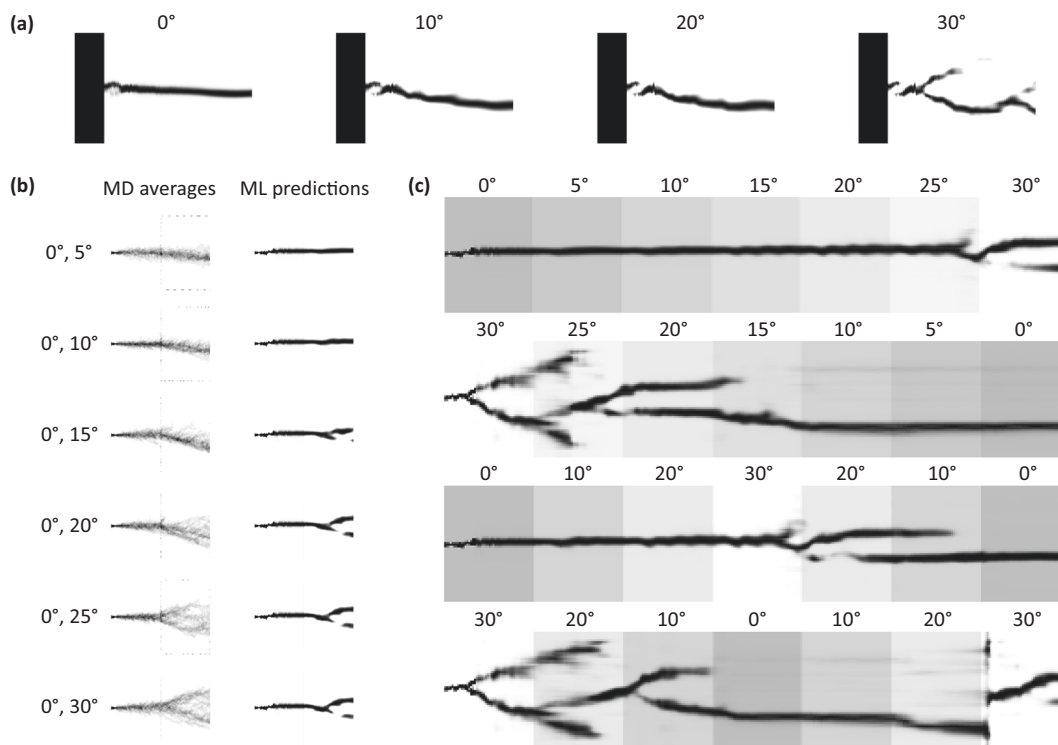

Fig. 4 Generalized fracture prediction. Despite being trained on exclusively pre-notched graphene samples, the model can **a** reasonably predict crack nucleation given a pristine free surface. Furthermore, the ML model can quickly and easily predict fracture in increasingly complex graphene systems including **b** bicrystal graphene of various misorientation angles and **c** linear and sinusoidal gradients of orientation within a larger scale sample.
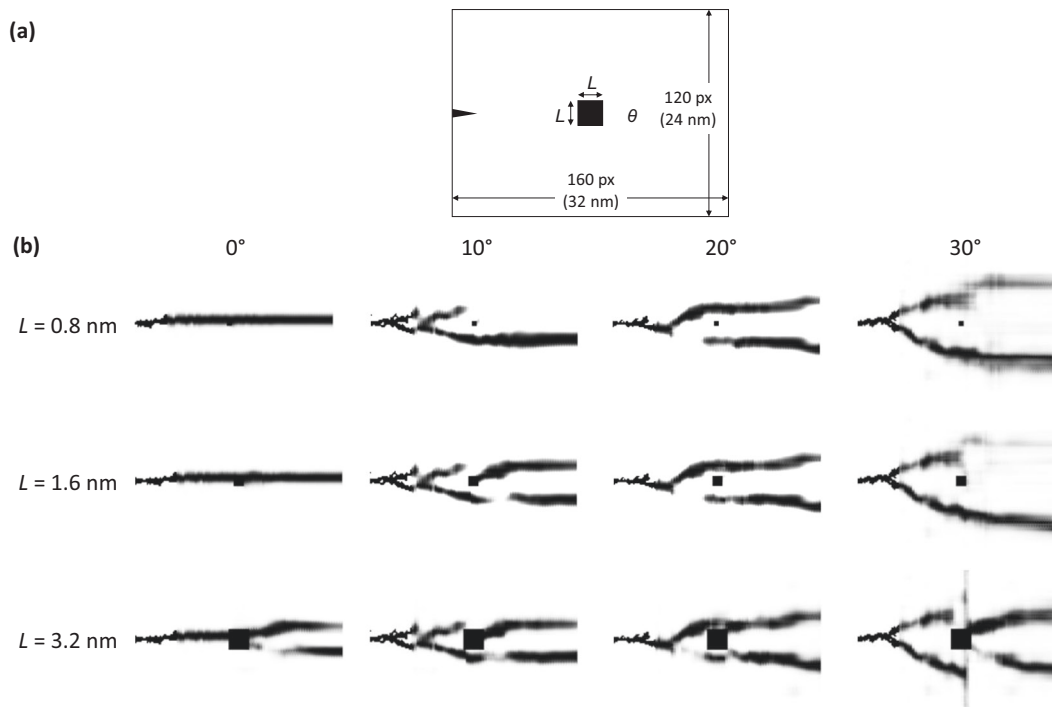
(a)

(b)

Fig. 5 Prediction of flaw tolerance. ML predictions show a degree of flaw tolerance, assessing fracture mechanics due to a pre-existing crack. The ML model predicts fracture paths given **a** 160×120 pixel images populated with values θ, representing sheets of graphene 32×24 nm with orientation angle θ. To introduce hole defects to the sample, a square region in the center of the image of size $L \times L$ is pre-labeled as "cracked" with the value -1 prior to starting fracture predictions. As the images show, **b** small hole defects placed in the center of graphene sheets only slightly affect the predicted fracture paths from the pristine samples. However, as flaw size increases to 3.2 nm, significant deviation occurs in the form of additional crack branching. This change in behavior at 3.2 nm is consistent with the threshold reported in the literature[29] for nanocrystalline graphene flaw tolerance—and may be the mechanism by which fracture strength is reportedly unaffected by flaw sizes below 3.2 nm but decreased with larger flaw sizes.

proper calibration of these parameters is complete without overfitting to specific training cases, ML shows itself to be an effective tool in predicting general fracture behaviors of real materials. The predicted results are closely aligned to results in the literature and much more computationally intensive MD simulations, even in cases incorporating differing surface, grain boundary, and point defects very different from the original training data. Furthermore, the ML method can be expanded to cases of increasing scale and complexity that are less feasible for MD to treat directly.

We believe that the predictions could be further improved in future work, as there are some discrepancies in fractal dimension between ML-predicted and MD-simulated fracture paths. Further work may also go toward increasing the fidelity of predicted fracture paths. Incorporating grain boundaries directly into the training data may result in increasingly accurate predictions for polycrystal samples. And there is undoubtedly a vast space for exploring how this deep learning method for predicting fracture patterns can apply to other crystalline materials, and treat materials with various defects. Rapid fracture prediction in this way allows us to calculate fracture toughness of structures without running atomic MD simulations every time, and increase our ability to design for and reverse engineer desired fracture behaviors. In addition, as this ML method works on image-based matrices scrubbed of any discrete atomic information, predictions are not limited to particular material scales and may aid as an important tool for multiscale design. In short, the landscape is vast and this work serves as a step toward validating the power of applying a deep learning approach toward mechanical design problems.

METHODS

MD simulations

Graphene sheets are generated by simply tiling the atomic coordinates of a single graphene unit cell to a region spanning 32×24 nm^2. Different sheet orientations from armchair to zigzag are obtained by rotating the atomic coordinates clockwise at the chosen angle about the center of the sample. Bicrystal graphene samples are obtained by rotating the atomic coordinates of half of the sheet independently from the other half. An initial crack of 2 nm is inserted at the left middle edge of the sample.

All MD simulations are performed using LAMMPS[30] and the AIREBO[31] potential. Though the AIREBO potential has trouble predicting the edge energy of armchair and zigzag edges compared with ab initio calculations[32], AIREBO can well characterize the elastic mechanics of carbon allotropes[26] and has often been used in the literature to model graphene[27] and provide a more realistic system than a toy L-J model. Samples are equilibrated with a Nosé–Hoover thermostat[33] at a constant temperature of 300 K and barostat with a constant pressure of 0 bar in the tensile direction for 50 ps. Periodic boundary conditions are applied across the tensile direction. Fracture is conducted at a strain rate of 0.0001% fs^{-1} with a timestep of 1 fs.

Data preprocessing and dataset preparation

To represent the MD simulation results in a matrix form amenable to the ML model, each sample is mapped into an image with a width of 120 pixels and a length of 160 pixels. Visualization of fracture paths is done using OVITO[34]. First, all atoms are rendered as squares sufficiently large enough to fill the empty space between them. Then, atoms with fewer neighbors i.e., those along the fracture path, are colored darker using a black-and-white color map. All atomic positions are frozen to their initial position. Binning atoms into square pixels may not accurately represent fine graphene structure if the image resolution is too low, biasing the ML predictions at that scale. However, as our fractal dimension analysis shows, the MD images used here have greater detail than our ML-predicted paths. Thus, ML predictions are not in the regime of guessing at structure in between

MD square pixel binning, and there is sufficient atomistic MD data at the current image resolution. Averaged MD images are obtained by overlaying fracture paths, summing the black intensity at each pixel location, and dividing by the number of image paths considered. The darker a pixel in the averaged image, the more frequently fracture paths went through that location for simulations of that orientation. These steps simplify the MD information from tens of thousands of atoms into a continuous image containing whether a given location in the sample has cracked or not.

These fracture path images are then split into batches of input and output matrices to capture the temporal cause-and-effect relationships behind crack propagation. The input "geometric" matrices are of height m, the height in pixels of the fracture path image, and width n, the number of columns of pixels under consideration for affecting the next step of crack propagation. These inputs are populated with the geometric orientation angle of the graphene sheet in non-cracked (white) locations and the value -1 at cracked (black) locations. Output "crack" matrices are of width o, comprising the next o columns to the right of each input matrix, and represents the crack formed after each propagation step from the input crack. Values are populated with 0 for non-cracked (white) locations and 1 at cracked (black) locations. This dataset of input/output matrix pairs is then used in training the ML model to predict the fracture propagation behavior of graphene materials. In this way the model was trained to use the last n pixel wide segment to predict the next o pixel wide segment.

ML Predictions

We utilize the ConvLSTM-based model reported in ref. [20] as a basis for the work here, applied to learn the spatiotemporal relations of crack propagation from physics-based MD simulations. In brief, the input data flow through two 1D convolutional layers to extract geometric features[7,35], each composed with 64 filters and kernel sizes of 60 and 61. Following the convolutional layers, a LSTM layer with 512 LSTM units is adopted to learn sequential relations[36] along crack propagation, and one dense layer with 120 neurons as the output layer. Training is done using an Adam optimizer with a learning rate of 0.0001 with decay of 0.001. Implementation is in Python with the Tensorflow package. Seventy-five percent of the input–output pairs are immediately used for training, with the remaining 25% reserved for validation.

An iterative process is used to generate a predicted fracture pattern. For each prediction step, the predicted crack matrix is converted to the geometric matrix format and used to update the input for the next prediction. Specifically, the oldest o columns of the input matrix are removed while the converted output matrix is appended to create the input for the next prediction step. The complete fracture prediction is comprised of each predicted crack step appended to the initial input.

Fractal analysis of fracture predictions

To quantitatively provide a measure of the fracture behavior and act as a point of comparison for ML performance, the Hausdorff fractal dimension of MD-simulated and ML-predicted fracture paths are calculated according to the traditional "box counting" method[37]. A MATLAB tool implementing this "box counting" method[38] provides a calculation of the fractal dimension from the black-and-white fracture path images processed by OVITO and output by the ML model.

DATA AVAILABILITY

The data that support the findings of this paper are available from the corresponding author upon reasonable request.

CODE AVAILABILITY

The code used to generate the findings of this paper is available from the corresponding author upon reasonable request.

REFERENCES

1. Buehler, M. J. *Atomistic Modeling of Materials Failure*, https://doi.org/10.1007/978-0-387-76426-9 (Springer, 2008).
2. Jung, G. S. et al. Interlocking friction governs the mechanical fracture of bilayer MoS2. *ACS Nano* **12**, 3600–3608 (2018).
3. Jung, G. S. et al. Anisotropic fracture dynamics due to local lattice distortions. *ACS Nano* **13**, 5693–5702 (2019).
4. Wang, S. S. et al. Atomically sharp crack tips in monolayer MoS2 and their enhanced toughness by vacancy defects. *ACS Nano* **10**, 9831–9839 (2016).
5. Peng, G. C. Y. et al. Multiscale modeling meets machine learning: what can we learn? *Arch. Comput. Methods Eng.* **1**, 3 (2020).
6. Tong, Q. & Li, S. A concurrent multiscale study of dynamic fracture. *Comput. Methods Appl. Mech. Eng.* **366**, 113075 (2020).
7. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444 (2015).
8. Udrescu, S.-M. & Tegmark, M. Symbolic pregression: discovering physical laws from raw distorted video. *arXiv* 2005.11212 (2020).
9. Jackson, N. E., Webb, M. A. & de Pablo, J. J. Recent advances in machine learning towards multiscale soft materials design. *Curr. Opin. Chem. Eng.* **23**, 106–114 (2019).
10. Xu, J. Distance-based protein folding powered by deep learning. *bioRxiv* 465955, https://doi.org/10.1101/465955 (2018).
11. AlQuraishi, M. End-to-end differentiable learning of protein structure. *Cell Syst.* **8**, 292–301 (2019).
12. Yu, C.-H. & Buehler, M. J. Sonification based de novo protein design using artificial intelligence, structure prediction, and analysis using molecular modeling. *APL Bioeng.* **4**, 016108 (2020).
13. Wang, J., Cao, H., Zhang, J. Z. H. & Qi, Y. Computational protein design with deep learning neural networks. *Sci. Rep.* **8**, 6349 (2018).
14. Yu, C.-H., Qin, Z., Martin-Martinez, F. J. & Buehler, M. J. 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 (2019).
15. Qin, Z. et al. Artificial intelligence method to design and fold alpha-helical structural proteins from the primary amino acid sequence. *Extrem. Mech. Lett.* **36**, 100652 (2020).
16. Gu, G. X., Chen, C.-T., Richmond, D. J. & Buehler, M. J. Bioinspired hierarchical composite design using machine learning: simulation, additive manufacturing, and experiment. *Mater. Horiz.* **5**, 939–945 (2018).
17. Schütt, K. T., Arbabzadah, F., Chmiela, S., Müller, K. R. & Tkatchenko, A. Quantum-chemical insights from deep tensor neural networks. *Nat. Commun.* **8**, 6–13 (2017).
18. Yu, C. H., Qin, Z., Martin-Martinez, F. & Buehler, M. J. A self-consistent sonification method to translate amino acid sequences into musical compositions and application in protein design using AI. *ACS Nano* **13**, 7471–7482 (2019).
19. Dewapriya, M. A. N., Rajapakse, R. K. N. D. & Dias, W. P. S. Characterizing fracture stress of defective graphene samples using shallow and deep artificial neural networks. *Carbon* **163**, 425–440 (2020).
20. Hsu, Y. C., Yu, C. H. & Buehler, M. J. Using deep learning to predict fracture patterns in crystalline solids. *Matter* **3**, 197–211 (2020).
21. Sen, D. & Buehler, M. J. Chemical complexity in mechanical deformation of metals. *Int. J. Multiscale Comput. Eng.* **5**, 181–202 (2007).
22. Qin, Z. & Buehler, M. J. Nonlinear viscous water at nanoporous two-dimensional interfaces resists high-speed flow through cooperativity. *Nano Lett.* **15**, 3939–3944 (2015).
23. Cranford, S. W. & Buehler, M. J. Mechanical properties of graphyne. *Carbon* **49**, 4111–4121 (2011).
24. Liu, F., Ming, P. & Li, J. Ab initio calculation of ideal strength and phonon instability of graphene under tension. *Phys. Rev. B - Condens. Matter Mater. Phys.* **76**, 064120 (2007).
25. Chu, Y., Ragab, T. & Basaran, C. The size effect in mechanical properties of finite-sized graphene nanoribbon. *Comput. Mater. Sci.* **81**, 269–274 (2014).
26. Jung, G. S., Qin, Z. & Buehler, M. J. Molecular mechanics of polycrystalline graphene with enhanced fracture toughness. *Extrem. Mech. Lett.* **2**, 52–59 (2015).
27. Zhao, H., Min, K. & Aluru, N. R. Size and chirality dependent elastic properties of graphene nanoribbons under uniaxial tension. *Nano Lett.* **9**, 3012–3015 (2009).
28. Mandelbrot, B. B. *The Fractal Geometry of Nature* (W. H. Freeman, 1983).
29. Zhang, T., Li, X., Kadkhodaei, S. & Gao, H. Flaw insensitive fracture in nanocrystalline graphene. *Nano Lett.* **12**, 4605–4610 (2012).
30. Plimpton, S. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.* **117**, 1–19 (1995).
31. Stuart, S. J., Tutein, A. B. & Harrison, J. A. A reactive potential for hydrocarbons with intermolecular interactions. *J. Chem. Phys.* **112**, 6472–6486 (2000).
32. Liu, Y., Dobrinsky, A. & Yakobson, B. I. Graphene edge from armchair to zigzag: the origins of nanotube chirality? *Phys. Rev. Lett.* https://doi.org/10.1103/PhysRevLett.105.235502 (2010).
33. Nosé, S. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268 (1984).

34. Stukowski, A. Visualization and analysis of atomistic simulation data with OVITO-the Open Visualization Tool. *Model. Simul. Mater. Sci. Eng.* **18**, 015012 (2010).

35. LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2323 (1998).

36. Oore, S., Simon, I., Dieleman, S., Eck, D. & Simonyan, K. This time with feeling: learning expressive musical performance. *Neural Comput. Appl.* **32**, 955–967 (2020).

37. Falconer, K. J. The Hausdorff dimension of self-affine fractals. *Math. Proc. Camb. Philos. Soc.* **103**, 339–350 (1988).

38. Costa, A. Hausdorff (box-counting) fractal dimension. MATLAB Central File Exchange, http://mathworks.com/matlabcentral/fileexchange/30329-hausdorff--box-counting--fractal-dimension (2013).

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

M.J.B. conceived the idea. A.J.L. carried out the molecular dynamics simulations, curated the training data, performed the parameter optimization and subsequent analyses. C.-H.Y. and Y.-C.H. worked on providing accessibility to the machine-learning model and participated in discussions. M.J.B. supervised the project. A.J.L. wrote the manuscript with M.J.B.

COMPETING INTERESTS

The authors declare no competing interests.

ADDITIONAL INFORMATION

Correspondence and requests for materials should be addressed to M.J.B.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.