npj | computational materials

Article

Published in partnership with the Shanghai Institute of Ceramics of the Chinese Academy of Sciences

https://doi.org/10.1038/s41524-025-01741-z

An interleaved physics-based deep-learning framework as a new cycle jumping approach for microstructurally small fatigue crack growth simulations

Vignesh Babu Rao & Ashley D. Spear ✉

Conventional fracture mechanics asserts that the relevant physics governing small crack growth occurs near the crack front. However, for fatigue, computing these physics for each crack-growth increment over the entire microstructurally small crack regime is computationally intractable. Properly trained deep-learning surrogate models can massively accelerate fatigue crack-growth predictions by virtually propagating an initial crack using micromechanical fields corresponding to just the initially cracked microstructure. As the predicted crack front advances, however, the fields no longer reflect relevant near-crack-front physics, leading to error and uncertainty accumulation. To address this, we present an interleaved physics-based deep-learning (PBDL) framework, where updates to the crack representation in the physics-based model are triggered intermittently using model uncertainty, thereby updating micromechanical fields passed to the deep-learning model. We show that this framework, representing a novel cycle-jumping approach, effectively limits error accumulation in history-dependent fatigue crack evolution and forms a template for other time-series applications in materials.

Given that the early stages of fatigue-crack evolution can consume the majority of total structural life for many technologically relevant applications[1,2], being able to predict the growth of microstructurally small fatigue cracks (MSCs) is essential for developing next-generation fatigue crack-resistant materials and realizing concepts like the airframe digital twin[3–5]. To make such predictions accurately requires careful treatment of the relevant microstructural features that influence crack evolution and a sufficiently high-fidelity representation of the evolving crack surface.

Simulation frameworks that leverage crystal plasticity constitutive models are typically adopted for MSC modeling tasks, as they accurately capture material deformation mechanisms and resolve the resulting micromechanical fields at the scale of MSCs. For instance, Castelluccio et al.[6] employed a crystal plasticity-based finite element (FE) model to simulate crack evolution using a fatigue indicator parameter while implicitly representing cracks by systematically reducing the elastic stiffness of elements in the FE mesh. Rovinelli et al.[7] used a fast Fourier transform-based crystal plasticity model, representing static cracks in the voxelized microstructure as voxels with void properties. Proudhon et al.[8] used crystal plasticity FE models to simulate MSC propagation in experimentally observed microstructures, representing cracks explicitly in the mesh. Phung et al.[9] employed a voxel-based remeshing framework to explicitly represent cracks in the FE mesh and employed crystal plasticity modeling to assess various crack growth criteria for predicting experimentally observed crack paths. While the aforementioned works represent progressive advancements in the fidelity with which MSCs can be represented, simulating MSC growth with high fidelity over realistic cycle counts remains computationally intractable. To illustrate, under high-cycle fatigue conditions—where cracks can spend up to 90% of their lifetime in the MSC regime—such effort could involve simulating MSC growth over cycle counts on the order of 10^4.

Cycle jumping[10–12] is a popular technique in continuum damage mechanics for accelerating fatigue simulations involving a large number of loading cycles. This approach begins by simulating a few initial fatigue cycles to establish the evolution of internal state variables, followed by strategically skipping a specified number of cycles while extrapolating key variables of interest (such as plastic strain or damage evolution) to account for the skipped cycles[13]. While cycle jumping has been shown to be a computationally efficient approach to effectively accelerate fatigue predictions at the continuum scale, its applicability to mesoscale models that explicitly simulate crack growth (such as MSC growth simulations) remains unexplored. Unlike continuum damage mechanics models, where damage evolves smoothly,

Department of Mechanical Engineering, University of Utah, Salt Lake City, UT, USA. ✉e-mail: ashley.spear@utah.edu

MSC simulations require predicting crack front behavior within localized fields that emerge from the interactions of the crack front with the microstructural features[14]. These interactions lead to significant variability in local stress and strain fields and thus make simple extrapolations challenging.

One promising avenue is the use of deep learning, which has demonstrated remarkable success in extracting complex patterns from large datasets across various domains, including natural language processing[15,16], image recognition[17], and bioinformatics[18]. Its application in crack growth prediction has gained significant attention in recent years. For instance, Pierson et al.[19,20] proposed using convolutional neural networks (CNNs) and XGBoost on data extracted from the micromechanical fields of uncracked polycrystals to predict crack paths. Rovinelli et al.[21] utilized Bayesian networks to derive analytical equations for small crack propagation. Hsu et al.[22] used ConvLSTM, combining CNN and long short-term memory (LSTM), to predict fracture patterns from 2D atomistic simulations, with input data encoded with crystal orientation and initial crack location. In a previous study[23], the authors demonstrated the ability of bi-directional long short-term memory (BiLSTM) networks to predict crack growth rates and propagation directions using data from high-fidelity physics-based simulations. Despite the impressive predictive power of BiLSTM models, the reliability over a large number of successive crack propagation increments remains a concern due to the potential for error propagation as the predicted crack front progresses farther away from the actual crack front (hence, deeper into the uncracked domain of the microstructure).

In this work, we propose an interleaved, physics-based deep-learning (PBDL) framework that leverages the rapid predictive capabilities of well-trained deep-learning models and the high accuracy of physics-based models within a tightly coupled loop (see Fig. 1). Uncertainty quantification (UQ) estimates from deep-learning ensemble predictions are used to inform the instant at which the physics-based model must be updated to reflect both the evolved crack surface and the corresponding micromechanical state prior to resuming deep-learning predictions. The UQ-informed interleaved PBDL framework marks a significant advancement in the field of fatigue modeling and also serves as a template for other time-series applications.

Results

In the previous study[23], it was determined that trained BiLSTM models effectively learned to map MSC growth parameters (viz., local crack extension, Δa, and kink angle, Φ^{crit}) at virtual crack front points (CFPs) to input sequences containing microstructural and micromechanical features extracted from the uncracked local neighborhood of the CFPs. Because the microstructure does not evolve with crack propagation in our physics-based simulation, the key variables driving the accuracy of BiLSTM models are the micromechanical features.

Interleaved physics-based deep-learning framework

Figure 1 illustrates the prediction process for the proposed interleaved PBDL framework. The process begins by evaluating micromechanical fields using the physics-based framework with an assumed initial crack explicitly represented in the FE mesh, and subsequently collecting input sequences at CFPs along the crack front. Trained ensembles of BiLSTM models predict Δa and Φ^{crit} point-wise at each CFP using these input sequences. Besides predicting the MSC growth parameters, the ensemble models also provide associated uncertainty estimates. The predictions are accepted if the model uncertainty estimate is below a user-specified threshold for either model. Accepted predictions of Δa and Φ^{crit} are combined to virtually propagate the crack front from $k = 0$ to $k = 1$. New input sequences are then collected from the virtually updated crack front at $k = 1$ (shown as white dots in Fig. 1) to predict the next crack growth increment. This way, the crack continues to propagate virtually using only the deep-learning predictions. However, as the virtual crack propagates farther away from the initial crack front, the micromechanical fields extracted in the uncracked region of the $k = 0$ simulation become less relevant for the deep-learning model's prediction. With each subsequent step relying on the previous one, model error and uncertainty accumulate over time.

Once the uncertainty accumulation reaches a user-specified threshold, the MSC propagation process switches from deep-learning predictions to physics-based simulation. At this point, the virtually propagated crack surface is explicitly represented in the FE mesh, and a physics-based simulation is conducted to update the micromechanical fields, thereby better representing the near-crack-front driving mechanisms. After updating the micromechanical fields, the process resumes virtual crack propagation using the deep-learning framework, now with features informed by the updated micromechanical fields. This interleaved approach of the PBDL framework leverages the strengths of both deep-learning and physics-based simulations, ensuring accurate and efficient predictions.

Demonstration of interleaved PBDL framework

In this section, we demonstrate the application of the interleaved PBDL framework using one of the microstructural instantiations in the test dataset.

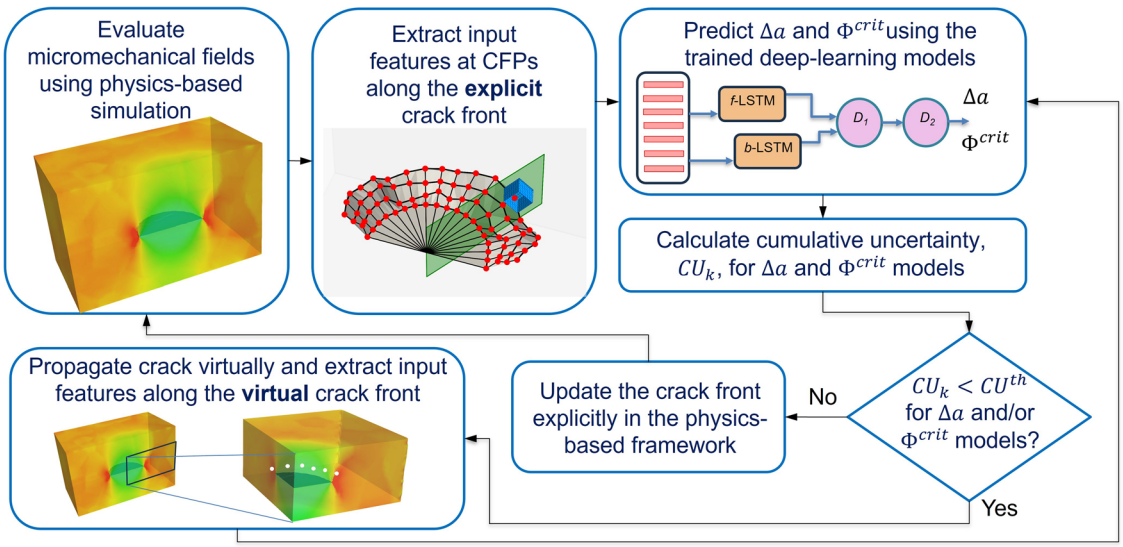

Fig. 1 | Workflow of interleaved physics-based deep-learning (PBDL) framework for predicting history-dependent propagation of microstructurally small fatigue cracks. The framework combines a physics-based crystal plasticity model, which evaluates micromechanical fields in a cracked microstructure, with a deep-learning surrogate that predicts a sequence of crack-growth increments using fields from the current state of the physics-based model. Uncertainty metrics are used to trigger intermittent updates of the explicit crack representation in the physics-based model to maintain relevant micromechanical fields as input for the deep-learning predictions.

We evaluate and compare the error and computational expense associated with propagating a crack from $k = 0$ to $k = 7$ across three frameworks: the interleaved PBDL framework, traditional physics-based simulation, and a deep-learning-only approach. Additionally, because the decision to update the micromechanical fields is governed by the cumulative uncertainty in crack extension (refer to Section "Uncertainty quantification and calibration procedure"), we present results from the PBDL framework under different user-specified thresholds for cumulative uncertainty, which effectively controls the frequency of updates to the physics-based model.

Figure 2 illustrates the evolution of the crack surface at discrete increments as it propagates from $k = 0$ to $k = 7$, along with the micromechanical fields used for predicting the next crack growth increment. Three cases are presented for comparison in Fig. 2: (1) sequence of forward predictions using only the deep-learning model with no updates to the physics-based model after the $k = 0$ simulation (i.e., without representing stress redistribution associated with formation of new traction-free surface), (2) interleaved PBDL predictions corresponding to a cumulative uncertainty threshold of $T_1 = 0.8$ μm, and (3) interleaved PBDL predictions corresponding to a cumulative uncertainty threshold of $T_2 = 0.4$ μm. The thresholds T_1 and T_2 are selected arbitrarily to result in updates to the physics-based model every four and two increments of crack growth, respectively. Of note, the threshold values are smaller than the voxel resolution of 1 μm used in the physics-based simulations and therefore remain within the intrinsic resolution error. In Fig. 2a–c, a cut-plane view shows the predicted crack trajectory from the PBDL framework (in white) overlaid with the corresponding reference trajectory from physics-based simulations (in black) for visual comparison. In Fig. 2b, predictions from $k = 0$ to $k = 4$ rely on the micromechanical fields computed at $k = 0$, but once the cumulative uncertainty reaches the predefined threshold at $k = 4$, the explicit crack surface is updated and the micromechanical fields are reevaluated at $k = 4$, and subsequent predictions from $k = 5$ to $k = 7$ rely on the micromechanical fields corresponding to the crack state at $k = 4$. Figure 2c illustrates an even more frequent update strategy, where the micromechanical fields are updated every two crack growth increments. Predictions from $k = 0$ to $k = 2$ are based on the initial fields at $k = 0$; predictions from $k = 3$ to $k = 4$ use fields updated at $k = 2$; predictions from $k = 5$ to $k = 6$ are based on fields updated at $k = 4$, and so on. As expected, more frequent updates of the micromechanical fields lead to better alignment between the predicted and reference crack trajectories.

It is important to recognize that although uncertainty is reset at each physics-based update in the interleaved PBDL framework, error continues to accumulate with crack propagation. Unlike long cracks, MSCs evolve as highly tortuous 3D surfaces, and therefore, we adopt the difference in crack surface area as the primary error metric. This choice is particularly meaningful because the predicted crack surfaces in our deep-learning framework are constructed based on both crack extension and kink angle predictions, and thus, the surface area error naturally integrates the error arising from both components. However, to facilitate interpretation, we define an equivalent crack radius r_{eq} by equating the error in crack surface area to that of an idealized flat semicircular crack. This scalar measure can then be compared to the average crack size, thus providing a basis for comparing the error.

Figure 3 shows the absolute error in total crack surface area relative to the corresponding high-fidelity physics-based simulation (ground truth), plotted as a function of average crack radius. Results are shown for the deep-learning-only framework and the interleaved PBDL framework with the two different uncertainty thresholds, T_1 and T_2. To assess long-term behavior, the error trends are extrapolated to a crack size of 1650 μm, which corresponds to the minimum reliably detectable crack size via ultrasonic inspection[24]. Extrapolation is performed using polynomial regression, with the polynomial order determined via trial and error. The interleaved PBDL framework's error closely followed a second-degree polynomial, whereas the deep-learning-only framework followed a third-degree polynomial. As shown in the inset of Fig. 3, the absolute error increases with crack growth in both approaches; however, the rate of growth is more pronounced in the deep-learning-only case. Moreover, the error associated with the PBDL framework using the tighter threshold T_2 is consistently smaller than that using a looser threshold T_1, highlighting the role of update frequency on overall predictive accuracy.

Although the difference in absolute error between the two methods may appear modest initially, the difference becomes increasingly significant with propagation. In the deep-learning-only framework, the error can eventually lead to highly unreliable predictions (error in equivalent crack radius, r_{eq}, translates to ≈ 318% of the actual average radius). In contrast, the interleaved PBDL framework, through periodic physics-based updates, limits error propagation. The quadratic error growth observed in the PBDL framework translates to a more reasonable deviation, with errors limited to ≈ 41% and ≈ 29% of the average crack radius for thresholds T_1 and T_2, respectively.

Figure 4 shows the projected computational costs incurred when predicting the evolution of a microstructurally small crack using the interleaved PBDL framework compared to purely physics-based simulation and deep-learning only approach (following the initial state, $k = 0$). The inset shows data points from a specific microstructural instantiation, which are used to extrapolate to a final crack size that would be detectable by ultrasonic testing. The results indicate that while the interleaved PBDL framework is more computationally expensive than the deep-learning-only approach, it is significantly less expensive than the traditional physics-based framework, reducing computational cost by a factor of four using threshold T_1 and by a factor of three using threshold T_2. As shown in Fig. 4, the computational savings of the interleaved PBDL framework depends on the user-specified tolerance of cumulative uncertainty.

Discussion

The interleaved PBDL framework proposed in this study exemplifies a synergistic approach whereby physics-based and deep-learning models complement each other to enhance predictive performance[25]. This strategy leverages the strengths of both methodologies, addressing their respective limitations. In the proposed framework, deep-learning models help physics-based models by rapidly predicting the local crack growth parameters, which is a crucial advantage given the high computational cost associated with physics-based simulations. Conversely, physics-based models contribute by providing high-fidelity micromechanical fields that are essential for accurate predictions. The intermittent updating of micromechanical fields ensures that the deep-learning model's input remains relevant and reasonably accurate, thus mitigating the degradation of prediction quality over time.

Results from Figs. 2 and 3 show that, for this history-dependent application, updating the physics-based model using deep-learning predictions that contain slight errors provides better overall accuracy than making no updates at all (i.e., using only a sequence of deep-learning predictions). Specifically, as shown in Fig. 3, the interleaved PBDL framework with uncertainty thresholds T_1 and T_2, which trigger micromechanical field updates at different frequencies, offers a consistent decrease in prediction error compared to deep-learning-only models (albeit at a slightly higher computational cost). The trend suggests that the error difference between the interleaved PBDL framework and the deep-learning-only approach will become more pronounced as the crack grows in size. We expect the benefits of incorporating updated micromechanical fields to outweigh the additional computational expense as the error accumulation in deep-learning-only models becomes more pronounced. Moreover, the ability to adjust the cumulative uncertainty threshold provides a flexible mechanism to balance accuracy and computational expense. By setting a tighter threshold, the framework can ensure more frequent updates to micromechanical fields, leading to more accurate predictions (but with additional computational expense). Therefore, the interleaved PBDL framework allows users to tailor the trade-off between accuracy and computational cost based on specific needs and constraints.

In conclusion, the interleaved PBDL approach effectively combines the rapid prediction capabilities of deep learning and the high fidelity of the

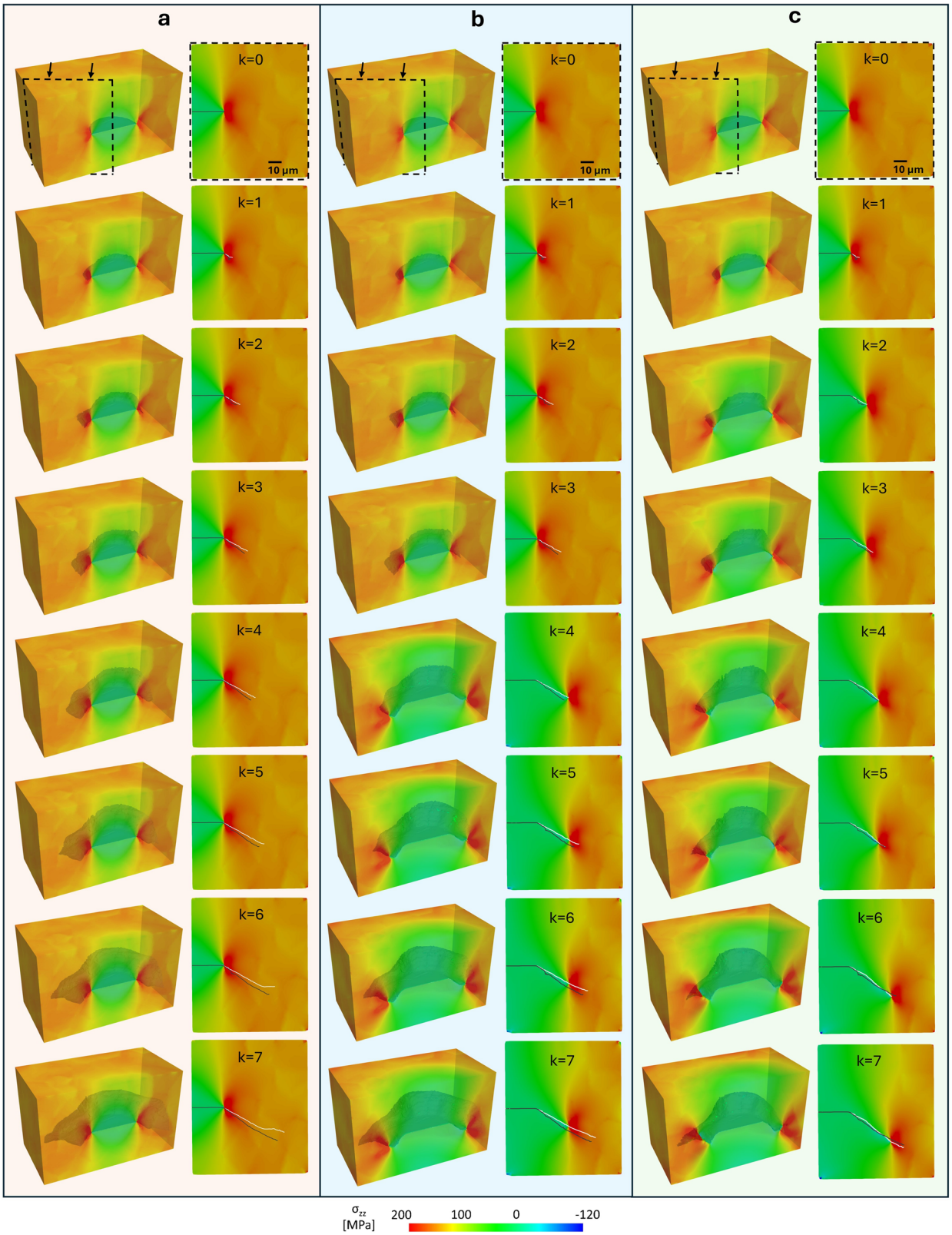

Fig. 2 | Influence of cumulative uncertainty threshold on the interleaved PBDL predictions. Illustration of interleaved PBDL framework predictions of microstructurally small crack growth over eight increments (k) within a 3D polycrystalline microstructure. A representative micromechanical field used for making forward predictions is overlaid at each increment. **a** Fields not updated after $k = 0$ (i.e., deep-learning-only case), **b** threshold T_1 resulting in micromechanical field update at $k = 4$ (refer to Supplementary Movie 1 for an animated version), and **c** threshold T_2 resulting in micromechanical field updates at $k = 2$, $k = 4$, and $k = 6$. In each cut-plane view, the predicted crack trajectory from the PBDL framework (in white) is overlaid with the corresponding reference trajectory from physics-based simulations (in black) for visual comparison.

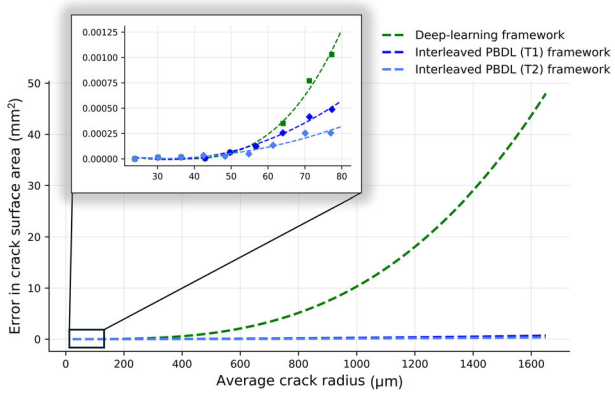

Fig. 3 | Comparison of error in predicted crack surface area. Projected absolute error in total crack surface area relative to a corresponding physics-based simulation (ground truth), plotted as a function of nominal crack size. Results are shown for deep-learning-only framework and the interleaved PBDL framework with two different cumulative uncertainty thresholds, T_1 and T_2. The inset shows actual data points from a specific microstructural instantiation, which are used to extrapolate to a final crack size that would be detectable by ultrasonic testing.

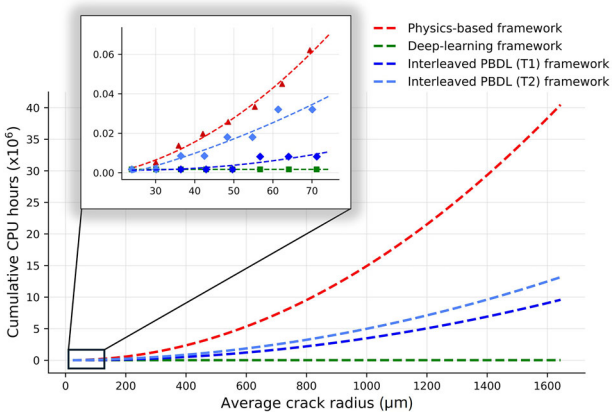

Fig. 4 | Comparison of computational expense. Projected computational costs for predicting the evolution of microstructurally small crack using the interleaved PBDL framework (with two different cumulative uncertainty thresholds, T_1 and T_2), in comparison to purely physics-based simulations and a deep-learning-only approach. The inset shows actual data points from a specific microstructural instantiation, which are used to extrapolate to a final crack size that would be detectable by ultrasonic testing.

physics-based simulations to predict MSC growth. By leveraging calibrated uncertainty estimates, the framework provides a reliable way to trade-off between speed and accuracy of predictions. The interleaved PBDL framework demonstrated here not only advances MSC growth modeling but also holds potential for broader applications in computational simulations across various time-series domains.

Methods

MSC growth dataset

The MSC growth data comprise two distinct datasets for developing predictive models for Δa and Φ^{crit}, respectively. Each dataset contains 18,000 data points collected from physics-based MSC growth simulations conducted in 40 statistically similar yet unique microstructures. A representative microstructure is illustrated in Fig. 5a. The physics-based simulations employ a high-fidelity FE framework to explicitly model 3D crack propagation at the microscale. For comprehensive details regarding the high-fidelity framework and the 40 simulations, the reader is referred to our previous work[14]. A summary is provided here for completeness. The high-fidelity framework integrates: (1) a voxel-based remeshing code that facilitates explicit representation of the evolving crack surface within the microstructure[26]; (2) an in-house meshing tool, Crackmesher[27], for adaptive mesh refinement along the evolving crack front; and (3) a crystal plasticity constitutive model, implemented as a user-defined subroutine in Abaqus®, to capture the anisotropic plastic behavior in face-centered cubic materials. The crystal plasticity model governs the plastic response through shear deformations on 12 octahedral slip systems, with the slip rate on each slip system, α, evolving as a function of the resolved shear stress τ and slip resistance g. Detailed constitutive equations are provided in the Supplementary Information.

Each microstructural instantiation is initialized with a flat, semicircular crack of radius 24 μm, oriented for mode-I loading, which propagates incrementally under prescribed cyclic loading conditions. At each crack growth increment, micromechanical quantities such as the cyclic crack tip displacement range (ΔCTD) and the direction of localization of the D_5 fatigue indicator parameter[28] are evaluated at each node along the crack front, which is then used to determine the local crack extension and the kink angle, respectively. The functional relationships used for these evaluations are described in the Supplementary Information under MSC Growth Criteria.

Five crack growth increments are simulated in each microstructural instantiation, resulting in 200 crack growth increments across all 40 microstructural instantiations. As illustrated in Fig. 5b, the crack front at each crack growth increment is discretized uniformly into 90 CFPs (although only 19 points are shown for each crack front in Fig. 5b for illustration). From each CFP, input sequences and corresponding ground truth labels for both Δa and Φ^{crit} models are extracted. These input sequences contain microstructural (e.g., maximum Schmid factor) and

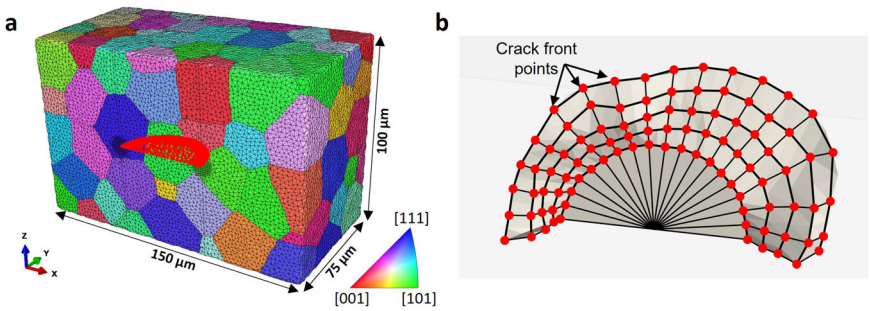

Fig. 5 | Meshed microstructure and crack front discretization strategy. a A representative microstructural instantiation used in physics-based simulations of crack growth, shown after conformal and adaptive meshing. The inverse pole figure (IPF) colors indicate the crystallographic orientations of grains in the microstructure. The highlighted crack surface nodes represent the initial semicircular crack explicitly introduced within the mesh. **b** A schematic illustration of crack front evolution from $k = 0$ to $k = 4$, along with the radial discretization strategy used to generate crack front points (CFPs). These CFPs serve as the basis for extracting input sequences used to train the deep-learning models.

micromechanical features (e.g., equivalent plastic strain) extracted both locally from the CFP location and from nonlocal points in the neighborhood, effectively encoding the microstructural dependence of MSC growth. It is worth noting that the micromechanical features used for training the deep-learning model are extracted solely from the crystal-plasticity FE simulation containing the initial crack (i.e., $k = 0$). While the process of generating input sequence is identical for both Δa and Φ^{crit} datasets, they primarily differ in the type of features collected: the Δa dataset includes magnitude-dependent features, whereas the Φ^{crit} dataset includes direction-dependent features. A list of all features used for training the Δa and Φ^{crit} models is provided in Tables 1 and 2, respectively, and is detailed further in the authors' previous work[14,23].

The 18,000 data sequences from 40 MSC growth simulations are split instantiation-wise into training (12,600 sequences from 28 microstructural instantiations), validation (2700 sequences from 6 microstructural instantiations), and testing (2700 sequences from 6 microstructural instantiations) datasets. The sequences in the training datasets of Δa and Φ^{crit} are used to train two separate BiLSTM models for predicting Δa and Φ^{crit}, respectively.

Table 1 | A list of features used to train BiLSTM model for predicting Δa[23]

Feature	Description
m_{max}	Maximum of the Schmid factors evaluated across 12 slip systems
$\sum_1^5 m$	Sum of the five largest Schmid factors out of those evaluated across 12 slip systems
d_{fs}^x, d_{fs}^y	Nearest distance to the free surface in X and Y directions
a	Half-crack length
$\lvert\gamma\rvert$	Absolute angular position of CFP
ω_{avg}	Disorientation angle
$E_{[001]}$	Elastic modulus along the loading direction [001]
Γ	A binary variable indicating whether a CFP is at the intersection of a grain boundary
Δa^{k-1}	Crack extension at crack growth increment $k-1$
D_5^{avg}	Fatigue indicator parameter extracted using the micromechanical fields at $k = 0$
ϵ_{eq}^{avg}	Equivalent plastic strain extracted using the micromechanical fields at $k = 0$
σ_{triax}^{avg}	Stress triaxiality extracted using the micromechanical fields at $k = 0$
ϵ_{33}^{avg}	Strain along Z-direction extracted using the micromechanical fields at $k = 0$
$\nabla\epsilon_{33}^{avg}$	Strain gradient along Z-direction extracted using the micromechanical fields at $k = 0$

Leveraging the BiLSTM network's ability to capture dependencies in sequential data, the models are trained to learn microstructure-sensitive behavior encoded in the input sequences extracted from spatially ordered CFPs. The network architecture consists of forward and backward LSTM layers that process the sequence in both directions, enabling the model to learn context from the entire neighborhood of a given CFP. Dropout is applied to both layers to reduce overfitting, followed by a time-distributed dense layer for dimensionality reduction, and a final dense layer that outputs the prediction.

Model training is performed using TensorFlow's Keras API, with the Adam optimizer minimizing the mean squared error between predictions and ground truth values. Key hyperparameters, including the number of LSTM units, learning rate, and dropout rate, are tuned using the Hyperband algorithm in the Keras Tuner library. The optimized parameters are listed in Table 3. The training and validation loss curves for the BiLSTM models predicting Δa and Φ^{crit}, trained using their respective tuned hyperparameters, are shown in Fig. 6. Once trained, the models can rapidly predict Δa and Φ^{crit} at any given CFP, given its corresponding input sequence. Additionally, each model is trained ten times with different random initializations, resulting in an ensemble of BiLSTM models. The arithmetic mean of the ensemble predictions provides the final predicted output. This ensemble approach not only enhances model accuracy but also provides meaningful uncertainty estimates. Performance metrics, including R^2, root mean square error (RMSE), and median absolute error (MDAE), are reported in Table 4.

While MSC growth is simulated for five crack growth increments (from $k = 0$ to $k = 4$) in 40 instantiations, for one of the microstructural instantiations in the test dataset, MSC is propagated for three additional increments (up to $k = 7$). These extended simulations provide ground truth data to assess the accuracy of the interleaved PBDL framework.

Uncertainty quantification and calibration procedure

The BiLSTM ensemble predictions are used not only to estimate the crack growth parameters but also to quantify the associated uncertainty. We assume a Gaussian posterior predictive distribution for each model prediction, with the ensemble's mean and standard deviation parameterizing

Table 3 | Tuned hyperparameters of the BiLSTM models that predict Δa and Φ^{crit}[23]

Hyperparameter	Δa BiLSTM model	Φ^{crit} BiLSTM model
Learning rate	0.006	0.0009
No. of units in the hidden state of LSTM cell	21	13
Dropout rate	0.01	0.3

Table 2 | A list of features used to train BiLSTM model for predicting Φ^{crit}[23]

Feature	Description
$\lambda_1, \lambda_2, \lambda_3, \lambda_4$	Angles formed by the trace of the slip planes 1 to 4 with $\vec{r}$ on the radial plane
$\Omega_{max}, \Omega_{min}$	Maximum and minimum tilt angles of the four slip planes relative to the crack orientation
Γ	A binary variable indicating whether a CFP is at the intersection of a grain boundary
$\lambda_{max(\Omega)}, \lambda_{min(\Omega)}$	Slip plane angle having the maximum and minimum tilt angles
$\lvert\gamma\rvert$	Absolute angular position of CFP
n_{grains}	Number of grains intersecting the radial probe
$\lambda_{max1(m)}, \lambda_{max2(m)}, \lambda_{max3(m)}, \lambda_{max4(m)}$	Slip plane angles sorted based on the maximum Schmid factor of the respective slip plane
Φ^{k-1}	Kink angle at crack growth increment $k-1$
$\omega_1, \omega_2, \omega_3, \omega_4$	Disorientation angle between the grain containing the CFP and the grain containing a probe point that intersects with the trace of the slip planes 1 to 4
$\Phi_{max(\epsilon_{eq}^p)}$	Angle along the radial probe having the maximum equivalent plastic strain extracted using the micromechanical fields at $k = 0$
$\Phi_{max(M_{micro})}$	Angle along the radial probe having the maximum micromechanical Taylor factor extracted using the micromechanical fields at $k = 0$

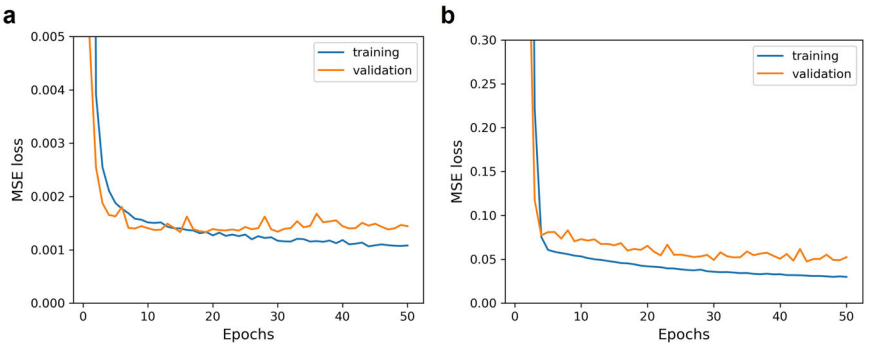

Fig. 6 | Loss curves of BiLSTM models predicting Δa and Φ^{crit}. Mean squared error (MSE) loss for both training and validation datasets plotted as a function of training epochs for the BiLSTM models predicting **a** crack extension (Δa) and **b** kink angle (Φ^{crit}).

Table 4 | Performance metrics of the ensembled BiLSTM models predicting crack extension (Δa) and kink angle (Φ^{crit})

Metric	Δa BiLSTM model	Φ^{crit} BiLSTM model
R^2	0.72	0.75
RMSE	0.31	0.51
MDAE	0.14	0.22

Root mean square error (RMSE) and median absolute error (MDAE) are reported in micrometers for the Δa model and in radians for the Φ^{crit} model.

the Gaussian distribution. Using the quantile function of this Gaussian distribution, we evaluate the confidence interval (CI) for each prediction, which serves as the measure of uncertainty. However, ensemble methods, while providing a way to estimate confidence levels, often result in inaccurate predictive intervals. For instance, a 90% confidence interval should ideally contain the true outcomes 90% of the time, but this is not always the case. To address this, we apply a calibration procedure proposed by Kuleshov et al.[29] to improve the quality of the uncertainty estimates. The calibration is performed on the validation dataset to develop a calibration curve, which is then used to correct the uncertainties in the test dataset. The calibration procedure involves determining the cumulative distribution function (CDF) of the ground truth relative to the model's predictive distribution for a particular output (predicted CDF) and comparing it with the CDF of that output relative to all other ground truths in the validation dataset (empirical CDF). A calibration curve is obtained by fitting the predicted and empirical CDFs of all validation data points using isotonic regression. The calibration curve is then used to adjust the quantiles for the predicted confidence intervals in the test dataset. While we use isotonic regression as a post hoc calibration method to improve predictive coverage of BiLSTM ensemble outputs after training, we note for the reader's reference that other calibration techniques exist with more intrinsic uncertainty quantification approaches, which are directly tied to the model's training objective. For instance, quantile regression has been widely used not only to quantify uncertainty by directly estimating prediction intervals[30,31], but also as a post-processing calibration tool to adjust predictive coverage[32,33]. Similarly, conformal prediction[34] is another technique that provides valid uncertainty bounds with minimal assumptions about the underlying data distribution. It can also be combined with quantile regression[35] to yield calibrated and sharp uncertainty bounds.

The calibrated uncertainty estimates based on BiLSTM ensemble predictions are recorded for each CFP at each crack growth step. As the crack grows in increments, the cumulative uncertainty (CU) at any given crack growth increment, k, on the radial plane is evaluated using the following equation:

$$\mathrm{CU}_k = \sqrt{\sum_{i}^{k} CI_i^2}, \quad (1)$$

where i is the crack growth increment relative to the most recent physics-based representation of the crack.

Data availability

The microstructural dataset used to generate the physics-based simulations for training data are reported in our previous work[23] and are available from the corresponding author upon request.

Code availability

All non-commercial codes to reproduce the findings of this study are available from the corresponding author upon reasonable request.

Received: 14 April 2025; Accepted: 17 July 2025;
Published online: 05 August 2025

References

1. Suresh, S. *Fatigue of Materials* 2nd edn (Cambridge University Press, 1998).
2. Brockenbrough, J. R., Hinkle, A. J., Magnusen, P. E. & Bucci, R. J. Microstructurally based model of fatigue initiation and growth. https://ntrs.nasa.gov/citations/19950008045 (1994).
3. Tuegel, E. The Airframe digital twin: some challenges to realization. https://arc.aiaa.org/doi/abs/10.2514/6.2012-1812 (2012).
4. Glaessgen, E. H. & Stargel, D. S. The digital twin paradigm for future NASA and U.S. Air Force vehicles. https://ntrs.nasa.gov/citations/20120008178 (2012).
5. Yeratapally, S. R., Leser, P. E., Hochhalter, J. D., Leser, W. P. & Ruggles, T. J. A digital twin feasibility study (Part I): non-deterministic predictions of fatigue life in aluminum alloy 7075-T651 using a microstructure-based multi-scale model. *Eng. Fract. Mech.* **228**, 106888 (2020).
6. Castelluccio, G. M. & McDowell, D. L. A mesoscale approach for growth of 3D microstructurally small fatigue cracks in polycrystals. *Int. J. Damage Mech.* **23**, 791–818 (2014).
7. Rovinelli, A. et al. Assessing reliability of fatigue indicator parameters for small crack growth via a probabilistic framework. *Model. Simul. Mater. Sci. Eng.* **25**, 045010 (2017).
8. Proudhon, H., Li, J., Ludwig, W., Roos, A. & Forest, S. Simulation of short fatigue crack propagation in a 3D experimental microstructure. *Adv. Eng. Mater.* **19**, 1–9 (2017).
9. Phung, B. et al. Predicting microstructurally sensitive fatigue-crack path in WE43 magnesium using high-fidelity numerical modeling and three-dimensional experimental characterization. *Fatigue Fract. Eng. Mater. Struct.* **47**, 862–883 (2024).
10. Van Paepegem, W. *The Cycle Jump Concept for Modelling High-cycle Fatigue in Composite Materials* (Elsevier Ltd, 2015).
11. Turon, A., Costa, J., Camanho, P. P. & Dávila, C. G. Simulation of delamination in composites under high-cycle fatigue. *Compos. Part A Appl. Sci. Manuf.* **38**, 2270–2282 (2007).
12. Hofman, P., van der Meer, F. P. & Sluys, L. J. Modeling of progressive high-cycle fatigue in composite laminates accounting

for local stress ratios. *Compos. Part A Appl. Sci. Manuf.* **183**, 108219 (2024).
13. Sally, O., Laurin, F., Julien, C., Desmorat, R. & Bouillon, F. An efficient computational strategy of cycle-jumps dedicated to fatigue of composite structures. *Int. J. Fatigue* **135**, 105500 (2020).
14. Babu Rao, V., Phung, B. R., Johnsson, B. T. & Spear, A. D. Statistical analysis of microstructurally small fatigue crack growth in three-dimensional polycrystals based on high-fidelity numerical simulations. *Eng. Fract. Mech.* **307**, 110282 (2024).
15. Vaswani, A. et al. Attention is all you need. *Adv. Neural Inf. Process. Syst*. **30**, https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf (2017).
16. Devlin, J., Chang, M. W., Lee, K. & Toutanova, K. BERT: pre-training of deep bidirectional transformers for language understanding. https://aclanthology.org/N19-1423 (2019).
17. Dosovitskiy, A. et al. An image is worth 16 x 16 words: transformers for image recognition at scale. https://arxiv.org/abs/2010.11929 (2021).
18. Jumper, J. et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
19. Pierson, K., Rahman, A. & Spear, A. D. Predicting microstructure-sensitive fatigue-crack path in 3D using a machine learning framework. *JOM* **71**, 2680–2694 (2019).
20. Pierson, K. D., Hochhalter, J. D. & Spear, A. D. Data-driven correlation analysis between observed 3D fatigue-crack path and computed fields from high-fidelity, crystal-plasticity, finite-element simulations. *JOM* **70**, 1159–1167 (2018).
21. Rovinelli, A., Sangid, M. D., Proudhon, H. & Ludwig, W. Using machine learning and a data-driven approach to identify the small fatigue crack driving force in polycrystalline materials. *npj Comput. Mater.* **4**, 35 (2018).
22. Hsu, Y.-C., Yu, C.-H. & Buehler, M. J. Using deep learning to predict fracture patterns in crystalline solids using deep learning to predict fracture patterns in crystalline solids. *Matter* **3**, 197–211 (2020).
23. Babu Rao, V. & Spear, A. D. A deep learning framework to predict microstructurally small fatigue crack growth in three-dimensional polycrystals. *Comput. Methods Appl. Mech. Eng.* **437**, 117689 (2025).
24. National Aeronautics and Space Administration. Nondestructive evaluation requirements for fracture-critical metallic components. *NASA Technical Standard (NASA-STD-5009B)*. https://standards.nasa.gov/sites/default/files/standards/NASA/C/0/Historical/nasa-std-5009b.pdf (2019).
25. Thuerey, N. et al. Physics-based deep learning. https://arxiv.org/abs/2109.05237 (2025).
26. Phung, B. R. & Spear, A. D. A voxel-based remeshing framework for the simulation of arbitrary three-dimensional crack growth in heterogeneous materials. *Eng. Fract. Mech.* **209**, 404–422 (2019).
27. Phung, B. R., He, J. & Spear, A. D. A surface-mesh gradation tool for generating gradated tetrahedral meshes of microstructures with defects. *Comput. Mater. Sci.* **197**, 110622 (2021).
28. Hochhalter, J. D. et al. A geometric approach to modeling microstructurally small fatigue crack formation: II. Physically based modeling of microstructure-dependent slip localization and actuation of the crack nucleation mechanism in AA 7075-T651. *Model. Simul. Mater. Sci. Eng.* **18**, 045004 (2010).
29. Kuleshov, V., Fenner, N. & Ermon, S. Accurate uncertainties for deep learning using calibrated regression. In *Proc. 35th International Conference on Machine Learning*, Vol. 80 of *Proceedings of Machine Learning Research* (eds Dy, J. & Krause, A.) 2796–2804 (PMLR, 2018).
30. Meinshausen, N. Quantile regression forests. *J. Mach. Learn. Res.* **7**, 983–999 (2006).
31. Wen, R., Torkkola, K., Narayanaswamy, B. M. & Madeka, D. A multi-horizon quantile recurrent forecaster. https://www.amazon.science/publications/a-multi-horizon-quantile-recurrent-forecaster (2017).
32. Rahmati, O. et al. Predicting uncertainty of machine learning models for modelling nitrate pollution of groundwater using quantile regression and uneec methods. *Sci. Total Environ.* **688**, 855–866 (2019).
33. Kasraei, B. et al. Quantile regression as a generic approach for estimating uncertainty of digital soil maps produced from machine-learning. *Environ. Model. Softw.* **144**, 105139 (2021).
34. Angelopoulos, A. N. & Bates, S. A gentle introduction to conformal prediction and distribution-free uncertainty quantification. https://arxiv.org/abs/2107.07511 (2022).
35. Romano, Y., Patterson, E. & Candès, E. J. *Conformalized Quantile Regression*. (Curran Associates Inc., 2019).

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. CMMI-1752400. The authors would like to acknowledge the computational resources provided by the University of Utah's Center for High Performance Computing.

Author contributions

A.D.S. conceptualized, acquired funding, and supervised the project. V.B.R. ran the physics-based simulations, trained the deep-learning models, analyzed the results, and wrote the manuscript. A.D.S. reviewed and edited the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41524-025-01741-z.

Correspondence and requests for materials should be addressed to Ashley D. Spear.