# CNN-based surrogate for the phase field damage model: A study on its generalization across microstructure parameters for composite materials

Yuxiang Gao[1], Matthew Berger[2], and Ravindra Duddu[3]

[1]Ph.D. student, Department of Civil and Environmental Engineering, Vanderbilt University, TN 37235(Corresponding author). ORCID: https://orcid.org/0000-0003-0445-0685. E-mail: yuxiang.gao@vanderbilt.edu

[2]Assistant Professor, Department of Computer Science, Vanderbilt University, TN 37235. E-mail: matthew.berger@vanderbilt.edu

[3]Associate Professor, Department of Civil and Environmental Engineering, Vanderbilt University, TN 37235. ORCID: https://orcid.org/0000-0002-4205-0247. E-mail: ravindra.duddu@vanderbilt.edu

## ABSTRACT

We investigate the generalization of a CNN-based surrogate for the phase field model in predicting both damage and peak load under uniaxial tension, given the 2D microstructure image of a unidirectional fiber-reinforced composite. We first discuss the phase field model and the numerical procedure to generate training and test data from synthetic microstructures with different volume fractions and fiber radii. We next present a two-stage approach for predicting peak load, achieved by first transforming a given fiber-encoded microstructure image to a continuous damage field; and second, predicting peak load from the damage field. A key finding is that the direct approach for predicting peak load from the microstructure image using a standard regression model fails to generalize. Instead, the damage field, even if imperfectly predicted, provides valuable cues for the CNN in generalizing across new microstructures within the range of parameters used in training. We describe several case studies to demonstrate the capability of the surrogate model to predict damage and peak load, and to interpolate over fiber radii and volume fractions.

## INTRODUCTION

Characterizing the mechanical performance of heterogeneous composites materials requires robust and efficient computational methods for simulating damage evolution at the microstructure scale. The phase field damage model in conjunction with the finite element method (FEM) has been widely used to predict fracture in homogeneous quasi-brittle materials; however, it is computationally expensive for heterogeneous composite material microstructures, despite the use of parallel computing. While computational multi-scale mechanics models are popular, the need for optimizing design-space parameters has led researchers to pursue computationally inexpensive surrogate models, based on micromechanics and, more recently, machine learning (ML) approaches. However, studies focusing on the generalization of ML approaches are necessary to effectively augment/replace well-established numerical methods to address complex fracture problems in a robust and efficient manner. Herein, we assess the generalization of convolutional neural network (CNN) based ML models as an efficient surrogate for the phase field damage model to predict fracture patterns in unidirectional fiber-reinforced composite microstructures.

Predicting the initiation and propagation of cracks (damage) in fiber-reinforced composites is challenging owing to the intrinsic variability of microstructure configurations and the nonlinear nature of damage evolution (Shakiba et al. 2019; Tan and Martínez-Pañeda 2021). While the volume fraction of fibers is well correlated with the stiffness of the composite, the distribution of fibers, fiber size, shape and misalignment affect the damage evolution and failure response (Singh and Pal 2021; Ahmadian et al. 2019). The phase field fracture/damage model has been recently used to better understand the microscale failure mechanisms in quasi-brittle composite materials in relation to microstructure parameters (Kuhn and Müller 2016; Espadas-Escalante and Isaksson 2019; Zhang et al. 2019). The major advantage of the phase field damage model is that it provides a thermodynamically consistent variational framework (Francfort and Marigo 1998; Miehe et al. 2010; Lo et al. 2019) to simulate fracture initiation and propagation. However, its major drawback is the computational expense arising from mesh size and time step restrictions and iterative schemes to ensure accuracy and convergence. Despite recent work implementing quasi-Newton schemes,

monolithic solvers, adaptive mesh refinement, domain decomposition preconditioners, length-scale insensitive formulations, continuous/discontinuous Galerkin methods, (Kristensen and Martínez-Pañeda 2020; Bharali et al. 2022; Hirshikesh et al. 2021; Svolos et al. 2020; Svolos et al. 2022; Wu and Nguyen 2018; Gupta et al. 2022), the phase field damage model is computationally expensive for fast simulation of of fracture evolution in heterogeneous composite materials over size scales larger than a few millimeters, particularly in 3D.

In the past few years, new ideas infused into the computational fracture mechanics area inspired by the success of ML approaches in other areas of engineering, speech recognition and computer vision. Restricting our attention to ML approaches for the phase-field fracture and/or composite strength prediction, existing approaches can be classified as: physics informed (trained using the governing PDEs) and data driven (i.e. trained using only data from experiments and/or computer simulations). Physics-informed neural networks (PINNs), originally introduced by Raissi et al. (Raissi et al. 2019), are a class of coordinate-based networks that can be explicitly optimized to satisfy a governing PDE. Application of PINNs to phase field fracture simulation is quite limited, for example, Goswami et al. (Goswami et al. 2020) proposed transfer learning enhanced PINNs. Despite recent efforts to speed up, the computational expense to solve the PDEs using PINNs is comparable to that of the FEM. In contrast, the traditional data-driven surrogates utilize ML models trained on data generated by solving PDEs using other numerical methods, such as the FEM. For example, Pathan et al. (Pathan et al. 2019) presented a supervised ML model which can predict the elastic modulus and tensile strength of a fiber-reinforced composite given the image of its microstructure, using a gradient-boosted tree regression algorithm.

More recently, CNN-based deep learning models were developed to predict field quantities, such as stress, strain or damage, and material parameters, such as tensile strength. Croom et al. (Croom et al. 2022) built a UNet (Ronneberger et al. 2015) to predict the stress field in an additively manufactured metal directly from microstructure scans. Yang et al. (Yang et al. 2021a; Yang et al. 2021b) developed a conditional GAN to predict strain and stress fields from 3D printed microstructure geometries. Bhaduri et al. (Bhaduri et al. 2021) applied UNet architecture

and transfer learning technique to predict the stress field in fiber-reinforced composite with a weighted mean square loss function. Mohammadzadeh et al. (Mohammadzadeh and Lejeune 2021) proposed a MultiRes-WNet architecture based on convolutional blocks to predict binary damage field given an image of material distribution. Sepasdar et al. (Sepasdar et al. 2022) implemented UNet to predict stress field at the early stage of damage initiation and binary crack pattern for any given microstructure geometry, using a modified mean average error as the loss function. Zhou et al. (Zhou et al. 2021) used a CNN-based model for composite strength prediction given the microstructure RVE. However, to best of our knowledge, there are no studies focusing on surrogates for the phase field damage model and their generalization.

In this article, we present a fast CNN-based surrogate for the phase field damage model, which belongs to the class of data driven deep learning approaches. Based on our observation that CNNs poorly generalize for predicting peak load (i.e., material strength) if provided with only the 2D microstructure (e.g., binary image that encodes the matrix and fibers), we propose a two-stage approach for predicting peak load. In first stage, we translate the microstructure image to its corresponding damage field at peak load, and in the second stage, we predict the peak load from the damage field. We find that this two-stage approach generalizes better compared to the one-stage or direct approach that predicts peak load exclusively from the microstructure image. To provide guidance on the selection of training sets for future applications, we further investigate the generalization of our CNN-based surrogate model to predict damage fields and peak loads for microstructures with different volume fractions and fiber radii.

## METHODOLOGY

### Phase field damage model

The phase-field damage model employed here is an extension of Miehe's model (Miehe et al. 2010) proposed by Lo et al. (Lo et al. 2019). This model properly accounts for tension-compression asymmetry of damage evolution in brittle materials. To describe the model formulation, we consider an arbitrary linear elastic solid $\Omega \subset \mathbb{R}^2$ with external boundary $\partial\Omega$, as shown in Fig.1. The boundary $\partial\Omega$ is split into disjoint Dirichlet and Neumann boundaries denoted by $\partial\Omega_{\mathrm{D}}$ and $\partial\Omega_{\mathrm{N}}$, respectively.

The fiber domain $\Omega^{(\mathrm{f})}$ is assumed to be undamaged at all times. The damaged state at any position $\mathbf{x}$ and time $t$ in the matrix domain $\Omega/\Omega^{(\mathrm{f})}$ is described by a scalar phase-field variable $D(\mathbf{x}, t) \in [0, 1]$, where $D = 0$ represents undamaged state and $D = 1$ the fully damaged state.

In the hybrid phase field model formulation, there are two unknown field variables that need to be solved for: the displacement field $\mathbf{u}$ and the phase-field damage variable $D$. To determine the displacement field $\mathbf{u}$, we solve the governing equations of the elastostatic boundary value problem along with the corresponding boundary conditions over a space-time domain $Q = \Omega \times [0, T]$:

$$\begin{cases} \boldsymbol{\nabla} \cdot \boldsymbol{\sigma} = \mathbf{0} & \text{in } Q \\ \mathbf{u} = \mathbf{u}^* & \text{on } \partial\Omega_{\mathrm{D}} \times [0, T] \\ \boldsymbol{\sigma} \cdot \mathbf{n} = \mathbf{t}^* & \text{on } \partial\Omega_{\mathrm{N}} \times [0, T] \end{cases} \tag{1}$$

where $\mathbf{u}^*$ represents a prescribed displacement on the Dirichlet boundary $\partial\Omega_{\mathrm{D}}$, $\mathbf{t}^*$ represents a prescribed traction on the Neumann boundary $\partial\Omega_{\mathrm{N}}$, $\boldsymbol{\sigma}$ is the Cauchy stress tensor defined by the isotropic linear elastic constitutive law under the plane strain assumption, and $\mathbf{n}$ is the outward normal to the boundary $\partial\Omega_{\mathrm{N}}$. To determine the damage field $D$, we solve the governing equation for phase-field evolution in the rate form along with the associated boundary conditions:

$$\begin{cases} \eta\dot{D} = l_{\mathrm{c}}\Delta D - \dfrac{D}{l_{\mathrm{c}}} + 2(1 - D)\dfrac{H}{G_{\mathrm{c}}} & \text{in } Q \\ \boldsymbol{\nabla} D \cdot \mathbf{n} = 0 & \text{on } \partial\Omega \times [0, T] \end{cases} \tag{2}$$

where $\Delta$ is the Laplace operator, $\eta$ is the viscous regularization parameter, $H$ is the history field variable ensuring the monotonic increase of the phase-field variable defined as (Lo et al. 2019)

$$H(\mathbf{x}, t) = \max_{\tau \in [0,t]} [\psi^+(\boldsymbol{\varepsilon}(\mathbf{x}, \tau))], \tag{3}$$

$\boldsymbol{\varepsilon}$ is the small strain tensor, and $\tau$ is a pseudo-time variable related to the applied load/displacement rate. In the above equation, by defining the history field variable equal to the maximum value of

the tensile portion of the strain energy $\psi^+(\boldsymbol{\varepsilon})$ we ensure that damage is irreversible. We use the standard Galerkin finite element procedure to discretize and solve the equations using a staggered approach in the open source finite element software FEniCS (Alnæs et al. 2015). Full details of the procedure to solve these equations can be found in our recent article (Sun et al. 2021).

**Ground truth data generation by finite element method**

We consider an archetype problem of damage/fracture evolution within the 2-D representative volume element (RVE) of a unidirectional fiber-reinforced composite under applied uniaxial tension in plane strain. We generated several 2-D geometries of using the L-BFGS-B algorithm described in Section 2.3 of Pathan et al. (Pathan et al. 2017). Specifically, we chose square-shaped RVEs with side length of $L = 50$ $\mu$m and varied two microstructure parameters: the volume fraction $V_\mathrm{f}$ is varied from 0.2 to 0.5, and the fiber radii $R_\mathrm{f}$ is chosen as 2 $\mu$m, 4 $\mu$m or 6 $\mu$m. To prevent damage artifacts from initiating and propagating near fibers that are near or intersect with the external boundary, we limit the fiber positions to a square interior region with side length $L - (R_\mathrm{f} + 0.04L)$, where $R_\mathrm{f}$ is the chosen fiber radii and $0.4L$ is a small offset determined by trial-and-error. We assume that the fiber is a linear elastic material with a large stiffness (the Young's modulus is 74 GPa and the Poisson's ratio is 0.2); whereas the matrix is a linear elastic material with a small stiffness (the Young's modulus is 3.5 GPa and the Poisson's ratio is 0.35) and accrues damage (critical strain energy release rate is 10 J/m$^2$). For simplicity, the fiber and the matrix interface are assumed to be perfectly bonded, so failure by debonding does not occur. We take the phase-field viscosity (stabilization) parameter $\eta = 1.2$ and length scale parameter $l_\mathrm{c} = 0.8$ $\mu$m.

The RVEs were discretized with a triangular finite element mesh with edge length $\approx 0.1$ $\mu$m, using the open-source mesh generation software Gmsh (Geuzaine and Remacle 2009). As illustrated in Fig. 1, we enforce the traction free condition on the left and right boundary of the RVE, and free slip (or roller) condition on the bottom boundary. The bottom left corner is pinned to prevent rigid body motion. At each pseudo-time step, to simulate the uniaxial extension test we apply a uniform displacement increment of 0.01 $\mu$m on the top boundary. By summing the nodal reaction forces on the top boundary we obtain the applied load at each time step. We stop the simulation when the

applied load drops to 15 % of the maximum/peak load. By dividing the peak load by the length of the RVE, we can determine the tensile strength, although we prefer to compare peak load values for different cases. The peak load and the corresponding 2D damage field image obtained from the finite element simulations are stored and used as the ground truth data for training and testing the ML models, as discussed below.

As shown in Table 1, we generated five datasets with different volume fractions and fiber radii for training and evaluating the proposed ML model. For dataset A, as a baseline, we consider 800 samples in the training set and 200 samples in the test set. All samples in dataset A have the same fiber radius of 4 $\mu$m and volume fraction of 0.5. The volume fraction of the samples in dataset B and C are set to be 0.5. In the training set of dataset B, we have 2000 samples (1000 each) with fiber radii of 2 or 6 $\mu$m; whereas the training set of dataset C contains 2001 samples (667 each) with fiber radii of 2, 4, and 6 $\mu$m. The test sets of datasets B and C are the same as that of dataset A. The fiber radius of dataset D and E samples is set to be 4 $\mu$m. There are 2000 samples (1000 each) in the training set of dataset D with volume fractions of 0.2 or 0.5; whereas the training set of dataset E has 2080 samples (130 each) with a range of volume fractions from 0.2 to 0.5. The test set of the datasets D and E are the same, which has 200 samples with a volume fraction of 0.3 and fiber radius of 4 $\mu$m.

**Baseline model: Predicting peak load from microstructure**

We first consider a simple ML approach that can directly predict the peak load given the rasterized image of the RVE as input, as illustrated in Fig. 2a. The architecture of this CNN-based regression model is based on that proposed by Simonyan et al. (Simonyan and Zisserman 2015). The RVE image is converted to a $512 \times 512$ pixel binary image, where fibers are marked as 1 and matrix is marked as 0. The $1 \times 512 \times 512$ pixel input image is fed through a series of convolutional blocks, where each block performs: (1) a convolution with a set of learnable filters, in turn producing a multi-channel image; (2) batch normalization (Ioffe and Szegedy 2015); (3) a leaky ReLU activation (Khalid et al. 2020); and (4) spatial downsampling of the image by a factor of 2. We apply this series of operations until we reach a 512-channel image with spatial resolution

$2\times2$. This image is then flattened into a 2, 048-dimensional vector, and through a 2-layer multilayer perceptron (MLP), the model produces a single value as output. This model is trained to minimize the mean squared error (MSE) between the microstructure's ground-truth and predicted peak load.

To evaluate the accuracy of this baseline or one-stage ML model in predicting the peak load, we trained and tested it on a sample dataset with fixed microstructure design parameters, that is, with a fiber radius of 4 $\mu$m and a volume fraction of 0.5. The dataset includes 1000 paired RVE geometry images and the corresponding peak load values, and we split it into 80 % training set and 20 % test set. As evident from Fig. 2, this direct approach overfits to the training data (c.f. Fig. 2b), and fails to generalize to novel microstructures not observed during training (c.f. Fig. 2c). This suggests that microstructure geometry alone does not provide sufficient cues for this baseline CNN model to properly reason about the microstructure's peak load. These results motivate our two-stage approach, where we seek to provide a CNN-based peak load predictor with richer information as input, in the form of a damage field image.

**CNN-based image-to-image-to-value surrogate model**

To obtain an improved prediction of peak load, we now consider a two-stage ML approach that first predicts the damage state given the rasterized image of the RVE as input, and then uses the damage field image to predict the peak load. We hypothesize the damage field at peak load can provide better information for peak load prediction than just the RVE geometry. This is based on the rationale that the phase field damage includes information on the strain energy state, material fracture toughness, along with the RVE geometry. Because this two-stage approach can predict both the damage field and peak load, it can be used as a surrogate for the phase field damage model.

To predict damage field image from the RVE image, we employed a CNN-based image-to-image model with an encoder-decoder and skip connection architecture, similar to the UNet (Ronneberger et al. 2015), as shown in Fig. 3a. The encoder (left side) accepts the RVE image as input, a single-channel binary encoding of the fibers (c.f. Sec. Baseline model: Predicting peak load from microstructure). The purpose of the encoder is to produce a series of multi-channel images that increase in channel resolution, and decrease in spatial resolution, using a learned set of convolutional

blocks. The decoder (right side) accepts as input the last multi-channel image produced by the encoder, and performs a series of learnable convolutional blocks, each interleaved with spatial upsampling, until we reach a target image resolution of $512 \times 512$ pixels – it is this image that we treat as the predicted damage field. Skip connections are utilized, wherein multi-channel images produced by the encoder are concatenated, channel-wise, onto the images produced by the decoder, limited to layers that have equivalent spatial resolution.

The parameters of the model are found by minimizing a mean absolute error (MAE) objective between the predicted damage field and the ground-truth damage field. Because the ground-truth damage fields are obtained from FEM-based simulations, each field is defined over a triangular mesh, thus incompatible with the model's predicted image represented over a regular square grid. To resolve this issue, we resample the FEM damage field onto a $512 \times 512$ regular grid, which enables us to compute a pixel-wise MAE loss for optimization. Specifically, for each vertex on the triangular finite element mesh, we find the cell in the regular grid that contains this vertex and contribute this vertex's nodal value to each of the grid cell's four vertices. The contribution is obtained through a bilinear interpolation weight $w$ within each cell to calculate the grid point value $Q$, as described in Algorithm 1

**Algorithm 1** Resampling from FEM results to a rasterized image based on bilinear interpolation

**Require:** $M \times M$ image resolution, $(\mathbf{X}^{L\times 1}, \mathbf{Y}^{L\times 1}) \in [0,1] \times [0,1]$ (Normalized finite element nodal coordinates), $\mathbf{F}^{L\times 1}$ (Nodal values of the phase field damage).

Initialize $\mathbf{Q}^{M\times M} = \mathbf{0}, \mathbf{W}^{M\times M} = \mathbf{0}$
$\mathbf{X} \leftarrow \mathbf{X} \times (M-1), \mathbf{Y} \leftarrow \mathbf{Y} \times (M-1)$
$\hat{\mathbf{X}} \leftarrow \lfloor \mathbf{X} \rfloor, \hat{\mathbf{Y}} \leftarrow \lfloor \mathbf{Y} \rfloor$
**for** $p \leftarrow 1$ to $L$ **do**
    **for** $i \leftarrow \hat{X}_p$ to $\hat{X}_p + 1$ **do**
        **for** $j \leftarrow \hat{Y}_p$ to $\hat{Y}_p + 1$ **do**
            $w \leftarrow (1 - |i - X_p|)(1 - |j - Y_p|)$
            $Q_{ij} \leftarrow Q_{ij} + w f_p$
            $W_{ij} \leftarrow W_{ij} + w$
        **end for**
    **end for**
**end for**
Output the resampled image $\mathbf{Q} \leftarrow \mathbf{Q}/(\mathbf{W} + 10^{-12})$

We employ this CNN-based surrogate model with the architecture shown in Fig. 3(a) due to the following considerations. First, encoder-decoder architectures have modest memory requirements and are computationally efficient; this is especially advantageous for models tasked with transforming one high-resolution image to another high-resolution image. Second, skip connections allow us to exploit intricate relationships between the spatial distribution of fibers and the damage field. Damage evolution in the microstructure is a *localized* phenomenon that directly depends on the stress/strain distribution in the RVE. By designing the decoder (i.e. damage field) to condition on features produced at multiple resolutions by the encoder (i.e. RVE geometry), we allow for the model to reason damage localization based on fiber distribution, at different spatial resolutions.

Thus, the proposed two-stage approach consists of a pair of CNN-based deep learning models: (1) a CNN tasked with predicting the phase field damage, given a microstructure image; and (2) a CNN tasked with predicting the peak load, given a damage field as an image. We opt to use the same CNN model for predicting peak load given an image, as described in Section Baseline model: Predicting peak load from microstructure. We optimize for each of these CNN models separately, wherein the damage field model is optimized solely on instances of microstructure and ground-truth damage field images (c.f. Fig. 3b); whereas the peak load model is optimized on instances of the predicted damage field and corresponding peak load value (c.f. Fig. 3c). At inference time, given a microstructure image, we first predict the corresponding damage field, and then use this damage field to predict the peak load.

**Error metrics**

To determine the difference between the ground truth and the CNN-based ML model predictions, we utilize two commonly used error metrics, namely root mean square relative error (RMSRE) and normalized root mean square error (NRMSE).

Because the peak load for each RVE is a unique value (dependent on the spatial distribution of fibers), for comparing the error across different RVEs it is more appropriate to first evaluate the relative error for each RVE and then calculate its root mean square. Hence for a given number $N$

of RVEs, we define the RMSRE between ground truth $y$ and ML model predictions $\bar{y}$ as

$$\text{RMSRE} = \sqrt{\frac{1}{N} \cdot \sum_{i=1}^{N} \left( \frac{y_i - \bar{y}_i}{y_i} \right)^2}. \tag{4}$$

The damage field image is also uniquely determined by the spatial distribution of fibers, with zero damage within the fibers (i.e. fiber is linear elastic). However, the RMSRE is not suitable for evaluating the error in damage, because it can be zero at various locations within a given RVE, which means the relative error is not well defined. Therefore, for comparing the relative error of rasterized images of phase field damage between different RVEs, we define the NRMSE with respect to the RMS of the ground truth. Thus, for a ground truth image $Y$ and a predicted image $\bar{Y}$ with $N$ pixels, the NRMSE is defined as

$$\text{NRMSE} = \sqrt{\frac{\sum_{i=1}^{N} (Y_i - \bar{Y}_i)^2}{\sum_{i=1}^{N} Y_i^2}}. \tag{5}$$

Note that the average magnitude of the phase-field damage (i.e. the integral of damage over the RVE divided by its area) is different for each RVE. Therfore, to appropriately compare the error across different RVEs, we must use the NRMSE as defined above.

## RESULTS AND DISCUSSION

### Preliminary check: damage field and peak load prediction

We first test whether the CNN-based surrogate model is able to predict the damage field and peak load with reasonable accuracy, if only the spatial distribution of the fibers is varied. Therefore, we trained and tested the model on a sample dataset with fixed microstructure parameters of volume fraction 0.5 and fiber radius 4 $\mu$m. This dataset includes 1000 paired binary images of RVE geometry and damage field. We split the dataset into 80 % training set and 20 % test set.

Figs. 4a and 4b show the model and ground truth damage field predictions and the difference for the smallest and largest NRMSE in the entire test dataset, respectively. In Fig. 4c we find that the NRMSE of more than 80 % cases is smaller than 0.17, and the average NRMSE across all cases is

0.1289, suggesting good generalization behavior on novel microstructures. As evident in Fig. 4b, larger error tends to manifest as a difference in the magnitude of damage. Otherwise, the locations in the microstructure where damage occurs, and where damage is absent, are generally preserved well in the predicted field. This suggests that the damage field input (provided to the CNN model for predicting peak load) will contain useful image cues, even when the predicted input is imperfect.

Fig. 4e shows the comparison between the peak load results obtained from the FEM and the image-to-image-to-value (surrogate) model. The data points are distributed along the diagonal, which indicates that the regression model is able to learn/capture the relationship between the damage field and the peak load. Additionally, we notice that the surrogate model tends to underestimate the peak load value at the higher end and overestimate at the lower end. This is possibly due to the relatively smaller amount of training data corresponding to the peak load extremities. The baseline model (cf. Section Baseline model: Predicting peak load from microstructure) results are shown in Fig. 4f for comparison. Fig. 4d shows that the surrogate model has more cases within 2% relative error and smaller maximum relative error for peak load prediction than the baseline model. Quantitatively, the RMSRE between surrogate model prediction and ground truth of the peak load is 0.0212, and the maximum relative error is 0.0576, which are both smaller than the baseline model (RMSRE = 0.0336, Max relative error = 0.0941). Thus, the surrogate model demonstrates an improved capability for predicting peak load compared to the baseline model.

**Model generalization and training set selection**

Volume fraction and fiber radii are two microstructure design parameters that can be altered to improve the mechanical behavior of composite materials. In the above two sections, we only examined the accuracy of the surrogate model when testing and training on the dataset with the constant volume fraction and fiber radius. In this section, we address the generalization of the surrogate model across a range of volume fraction and fiber radii, which can be a formidable challenge. Particularly, it is useful to know how to select the appropriate training dataset so that the ML model performs well across different test sets spanning the range of microstructure parameters. We consider two relatively straightforward training strategies:

1. Extreme: train only using the dataset corresponding to the minimum and maximum values of microstructure parameters;
2. Range: train using the dataset covering the entire range of microstructure parameters.

For both strategies, we assess the accuracy of damaged field and peak load prediction using NRMSE and RMSRE metrics, respectively. We utilize a separate dataset within the range of parameters to test the ML model's ability to interpolate or generalize.

*Training with the extremes of fiber radius*

We trained the surrogate model with the datasets (consisting of RVE and damage field images and peak load values) corresponding to volume fraction $v_f = 0.5$ and the extremes of fiber radii $R_f = 2$ $\mu$m and 6 $\mu$m. Fig. 5(a) and Fig. 6(a) shows the model performance on the test dataset with $v_f = 0.5$ and intermediate fiber radius $R_f = 4$ $\mu$m. The surrogate model predicted the locations of damage initiation and the overall damage pattern reasonably well (results not shown here), despite the discrepancy in the magnitude of the damage field. The NRMSE of 96 % cases are smaller than 0.23 with an average NRMSE of 0.1589 for the entire distribution, as shown in Fig. 5(a). The relative error between the peak load predictions from the surrogate model and the finite element method (FEM) is quite small with RMSRE = 0.0262, which is comparable to that from our preliminary check (RMSRE = 0.0162) in Section Preliminary check: damage field and peak load prediction. Considering that the surrogate model was never trained on fiber radii of $R_f = 4$ $\mu$m, this study suggests that the two-stage approach can successfully generalize over the range of microstructure design parameters.

*Training with a range of fiber radii*

We next trained the surrogate model with the datasets of all three fiber radii $R_f = 2, 4$ and 6 $\mu$m and tested its performance for each fiber radii, separately. The purpose of this study is understand which of the two training strategies (i.e. extreme or range) improves accuracy of the damage field and peak load prediction. Figs. 5b–5d and 6b–6d show the results assessing the surrogate model performance on the test datasets of three different fiber radii. The surrogate model is able to predict

damage pattern reasonably well (results not shown here), even for the cases with the largest NRMSE in the respective test datasets. However, there are magnitude differences in the damage field images corresponding to surrogate model and ground truth.

Figures 5b–5d show that the NMRSE of the predicted damage field is less than 0.15 for more than 90% of cases across the three test sets. We notice that the surrogate model gives the smallest average NRMSE for damage field prediction on the test dataset with radius 2 $\mu$m (Fig. 5b). We suspect this is because the RVE contains a larger number of fibers distributed uniformly through the matrix, which reduces the variability or uncertainty arising from the spatial distribution of fibers.

Figures 6b–6d show that the surrogate model has more cases with < 2% relative error, demonstrating that it can generalize well when training with a range of fiber radii. Thus, the surrogate model is able to learn/capture the relationship between the damage field and the peak load.

*Model performance with different datasets*

We now examine the surrogate model's performance across the three different datasets A, B, and C, as shown in Table 1. To elaborate our aim here: in Section Preliminary check: damage field and peak load prediction, we trained and tested the surrogate model using the dataset A with volume fracture $v_{\mathrm{f}} = 0.5$ and fiber radii, $R_{\mathrm{f}} = 4$ $\mu$m, with an 80 % – 20 % split (see Table 1); in Section Model generalization and training set selection thus far, we trained and tested the surrogate model with datasets B and C, where we trained using the dataset with the extremes and range of the fiber radius, and tested using the dataset with an intermediate fiber radii. Thus, in all three datasets the test set is the same, which allows us to identify the best training strategy for the surrogate model.

From Table 2, it is evident that the lowest NRMSE error for damage field prediction and RMSRE for peak load prediction is obtained with dataset C, compared to datasets A and B. Thus, training the surrogate model with datasets of three different fiber radii improves its accuracy compared to training with the same fiber radius as the test dataset. Therefore, training the surrogate model with the range of fiber radii is a better choice for ensuring generalization.

*Training with the extremes of volume fraction*

We next trained the surrogate model with the datasets corresponding to fiber radius $R_f = 4$ $\mu$m and the extremes of fiber volume fraction $v_f = 0.2$ and 0.5. Fig. 5e and 6e show the model performance on the test dataset with $v_f = 0.3$ and $R_f = 4$ $\mu$m. This training and test dataset is denoted as dataset D in Table 1. The NRMSE of 85% of cases is smaller than 0.2, despite the maximum NMRSE being 0.35, as shown in Fig. 5e. Fig. 6e shows the relative errors between the peak load predictions from the surrogate model and the FEM are less than 4 % in most cases. This study illustrates that the surrogate model trained over the extremes of fiber volume fraction can be reasonably predictive at intermediate volume fractions.

*Training with the range of volume fraction*

We now investigate whether the model will exhibit improved generalization when training with the range of volume fraction. Therefore, we trained the surrogate model on a dataset consisting of 16 different volume fractions ranging from 20 % to 50 %. The fiber radius $R_f = 4$ $\mu$m is kept constant, so each additional fiber in the RVE increases the fiber volume fraction by 2%. To evaluate the accuracy of the two training strategies (i.e. extreme versus range), the surrogate model was tested on the same dataset with $V_f = 0.3$ and $R_f = 4$ $\mu$m, as in the previous section. The details of the two relevant datasets are given in Table. 1.

Fig. 5f and 6f shows the results assessing the surrogate model's performance when trained and tested with dataset E spanning the entire range of fiber volume fractions. The average NRMSE is 0.1481 when trained with dataset E, as opposed to 0.1477 when trained with dataset D (see Table 2). The RMSRE of peak load prediction for dataset E is 2.67 %, compared to 3.27 % smaller than with dataset D.

This study showed that training with a range of volume fractions slightly improves the surrogate model's capability to predict peak load, compared to training with the extremes of volume fractions. Interestingly, the improvement in peak load prediction is not accompanied by an improvement in predicting the damage field (c.f. Fig. 5e and Fig. 5f). This finding indicates that a diverse coverage of volume fraction is of more benefit for peak load generalization, and helps to mitigate the

overprediction bias phenomena previously.

*Discussion*

In all the above studies, we restricted our attention to "generalization" across microstructure parameters within the range used in training and to loading/boundary conditions corresponding to uniaxial tension under plane strain. To assess generalization outside of the parameter range, we performed a preliminary study by training the model on the dataset with $R_f = 2$ & $4$ $\mu$ m and testing on the dataset with $R_f = 6$ $\mu$ m. We find that the peak load error from the two-stage CNN model is larger for prediction outside the parameter range (NMRSE = 0.0542) compared to prediction within the parameter range (NMRSE = 0.0262). However, if we include a few RVEs $R_f = 6$ $\mu$ m in the training dataset, we find that two-stage ML model accuracy improves for prediction outside the parameter range (NMRSE = 0.0322). This is consistent with our observation that the error is the smallest when trained with the range of fiber radii or volume fraction. These results are included in Figure S6 in the supplementary material. Model generalization for the loading/BCs can be achieved by encoding these conditions as an additional channel of the input image; but this is not as simple for nonlinear damage prediction, unlike linear elastic stress prediction and will require more data for training and testing.

While non-circular carbon fiber reinforced plastics (CFRPs) may have superior performance, circular CFRPs are most commonly used. Therefore, we only considered circular fibers in 2D RVEs relying on the assumption that the fibers are well-aligned in the longitudinal (out-of-plane) direction. The proposed two-stage ML model illustrates that the "image-to-value" CNN model shows better performance when trained with rasterized damage images compared to the RVE image, for the same model architecture and hyperparameters. We also varied the model architecture and hyperparameters (e.g. by considering different initial channels and the number of convolutional blocks). The corresponding RMSREs given in Table S1 in the supplementary information illustrate that the two-stage model predictions have smaller RMSRE compared to that from the one-stage model. This supports our hypothesis that the damage field provides better cues to the CNN model in predicting the peak load.

## CONCLUSION

In this paper, our goal was to formulate and train a computationally inexpensive surrogate for the phase field damage model, which can quickly predict the damage field and peak load given the microstructure geometry of a unidirectional fiber-reinforced composite. To this end, we developed a two-stage CNN-based deep learning model – where the first stage predicts the damage field given a microstructure image, and the second stage takes in this predicted damage field and predicts the peak load value. Our preliminary check showed that the model was able to learn the relationship between fiber distribution, damage field and peak load, with fixed fiber radii and volume fraction. We next demonstrated that the direct prediction of peak load from the RVE image using the CNN-based image-to-value model is less accurate compared to the indirect prediction of peak load from the damage field image. We next assessed the generalization of the surrogate model to predict damage field and peak load for different fiber radii and volume fractions with different datasets. Although the model provided reasonably accurate predictions of the damage field and peak load when training with only the extremes of fiber radii and volume fraction, we find that training with the range of fiber radii and volume fraction improves model generalization.

## DATA AVAILABILITY STATEMENT

Dataset and codes for machine learning model training and testing used during the study are available in repositories online in accordance with funder data retention policies. (Codes on GitHub: https://github.com/YuxiangGao0321/Phase-field-damage-CNN, Datasets on Dryad: https://doi.org/10.5061/dryad.jh9w0vtfp.) Codes for solving the phase-field damage problem by FEniCS to generate machine learning datasets of this study are available from the corresponding author upon reasonable request.


## ACKNOWLEDGEMENT

This work was funded by a research grant from 3M Company. YG and RD acknowledge the financial support from NSF grant #PLR-1847173, which supported the development of the phase field model and the finite element implementation in FEniCS software.

## SUPPLEMENTAL MATERIALS

Figures S1 - S6, Table S1, and twelve datasets are available online https://doi.org/10.5061/dryad.jh9w0vtfp.

## REFERENCES


Ahmadian, H., Yang, M., Nagarajan, A., and Soghrati, S. (2019). "Effects of shape and misalignment of fibers on the failure response of carbon fiber reinforced polymers." *Computational Mechanics*, 63(5), 999–1017.

Alnæs, M., Blechta, J., Hake, J., Johansson, A., Kehlet, B., Logg, A., Richardson, C., Ring, J., Rognes, M. E., and Wells, G. N. (2015). "The FEniCS Project Version 1.5." *Archive of Numerical Software*, 3(100).

Bhaduri, A., Gupta, A., and Graham-Brady, L. (2021). "Stress field prediction in fiber-reinforced composite materials using a deep learning approach." *arXiv:2111.05271 [cond-mat]*.

Bharali, R., Goswami, S., Anitescu, C., and Rabczuk, T. (2022). "A robust monolithic solver for phase-field fracture integrated with fracture energy based arc-length method and under-relaxation." *Computer Methods in Applied Mechanics and Engineering*, 394, 114927.

Croom, B. P., Berkson, M., Mueller, R. K., Presley, M., and Storck, S. (2022). "Deep learning prediction of stress fields in additively manufactured metals with intricate defect networks." *Mechanics of Materials*, 165, 104191.

Espadas-Escalante, J. J. and Isaksson, P. (2019). "Mesoscale analysis of the transverse cracking kinetics in woven composite laminates using a phase-field fracture theory." *Engineering Fracture Mechanics*, 216, 106523.

Francfort, G. A. and Marigo, J. J. (1998). "Revisiting brittle fracture as an energy minimization problem." *Journal of the Mechanics and Physics of Solids*, 46(8), 1319–1342.

Geuzaine, C. and Remacle, J.-F. (2009). "Gmsh: A 3-D finite element mesh generator with built-in pre- and post-processing facilities: THE GMSH PAPER." *International Journal for Numerical Methods in Engineering*, 79(11), 1309–1331.

Goswami, S., Anitescu, C., Chakraborty, S., and Rabczuk, T. (2020). "Transfer learning enhanced physics informed neural network for phase-field modeling of fracture." *Theoretical and Applied Fracture Mechanics*, 106, 102447.

Gupta, A., Krishnan, U. M., Mandal, T. K., Chowdhury, R., and Nguyen, V. P. (2022). "An adaptive

mesh refinement algorithm for phase-field fracture models: Application to brittle, cohesive, and dynamic fracture." *Computer Methods in Applied Mechanics and Engineering*, 399, 115347.

Hirshikesh, H., Pramod, A. L. N., Waisman, H., and Natarajan, S. (2021). "Adaptive phase field method using novel physics based refinement criteria." *Computer Methods in Applied Mechanics and Engineering*, 383, 113874.

Ioffe, S. and Szegedy, C. (2015). "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift." *Proceedings of the 32nd International Conference on Machine Learning*, PMLR, 448–456, <https://proceedings.mlr.press/v37/ioffe15.html> (June).

Khalid, M., Baber, J., Kasi, M. K., Bakhtyar, M., Devi, V., and Sheikh, N. (2020). "Empirical Evaluation of Activation Functions in Deep Convolution Neural Network for Facial Expression Recognition." *2020 43rd International Conference on Telecommunications and Signal Processing (TSP)*, 204–207 (July).

Kristensen, P. K. and Martínez-Pañeda, E. (2020). "Phase field fracture modelling using quasi-Newton methods and a new adaptive step scheme." *Theoretical and Applied Fracture Mechanics*, 107, 102446.

Kuhn, C. and Müller, R. (2016). "A discussion of fracture mechanisms in heterogeneous materials by means of configurational forces in a phase field fracture model." *Computer Methods in Applied Mechanics and Engineering*, 312, 95–116.

Lo, Y.-S., Borden, M. J., Ravi-Chandar, K., and Landis, C. M. (2019). "A phase-field model for fatigue crack growth." *Journal of the Mechanics and Physics of Solids*, 132, 103684.

Miehe, C., Hofacker, M., and Welschinger, F. (2010). "A phase field model for rate-independent crack propagation: Robust algorithmic implementation based on operator splits." *Computer Methods in Applied Mechanics and Engineering*, 199(45), 2765–2778.

Mohammadzadeh, S. and Lejeune, E. (2021). "Predicting Mechanically Driven Full-Field Quantities of Interest with Deep Learning-Based Metamodels." *arXiv:2108.03995 [physics]*.

Pathan, M., Tagarielli, V., Patsias, S., and Baiz-Villafranca, P. (2017). "A new algorithm to generate representative volume elements of composites with cylindrical or spherical fillers." *Composites*

*Part B: Engineering*, 110, 267–278.

Pathan, M. V., Ponnusami, S. A., Pathan, J., Pitisongsawat, R., Erice, B., Petrinic, N., and Tagarielli, V. L. (2019). "Predictions of the mechanical properties of unidirectional fibre composites by supervised machine learning." *Scientific Reports*, 9(1), 13964.

Raissi, M., Perdikaris, P., and Karniadakis, G. (2019). "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations." *Journal of Computational Physics*, 378, 686–707.

Ronneberger, O., Fischer, P., and Brox, T. (2015). "U-Net: Convolutional Networks for Biomedical Image Segmentation." *Medical Image Computing and Computer-Assisted Intervention (MICCAI 2015)*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds., Cham, Springer International Publishing, 234–241.

Sepasdar, R., Karpatne, A., and Shakiba, M. (2022). "A data-driven approach to full-field nonlinear stress distribution and failure pattern prediction in composites using deep learning." *Computer Methods in Applied Mechanics and Engineering*, 397, 115126.

Shakiba, M., Brandyberry, D. R., Zacek, S., and Geubelle, P. H. (2019). "Transverse failure of carbon fiber composites: Analytical sensitivity to the distribution of fiber/matrix interface properties." *International Journal for Numerical Methods in Engineering*, 120(5), 650–665.

Simonyan, K. and Zisserman, A. (2015). "Very Deep Convolutional Networks for Large-Scale Image Recognition." *The 3rd International Conference on Learning Representations (ICLR2015)*, <http://arxiv.org/abs/1409.1556> (April).

Singh, A. and Pal, S. (2021). "Multi-phase field modeling for various fracture mechanisms in composites." *Engineering Fracture Mechanics*, 241, 107348.

Sun, X., Duddu, R., and Hirshikesh (2021). "A poro-damage phase field model for hydrofracturing of glacier crevasses." *Extreme Mechanics Letters*, 45, 101277.

Svolos, L., Berger-Vergiat, L., and Waisman, H. (2020). "Updating strategy of a domain decomposition preconditioner for parallel solution of dynamic fracture problems." *Journal of Computational Physics*, 422, 109746.

513 Svolos, L., Mourad, H. M., Manzini, G., and Garikipati, K. (2022). "A fourth-order phase-field
514 fracture model: Formulation and numerical solution using a continuous/discontinuous Galerkin
515 method." *Journal of the Mechanics and Physics of Solids*, 165, 104910.
516 Tan, W. and Martínez-Pañeda, E. (2021). "Phase field predictions of microscopic fracture and
517 R-curve behaviour of fibre-reinforced composites." *Composites Science and Technology*, 202,
518 108539.
519 Wu, J.-Y. and Nguyen, V. P. (2018). "A length scale insensitive phase-field damage model for brittle
520 fracture." *Journal of the Mechanics and Physics of Solids*, 119, 20–42.
521 Yang, Z., Yu, C.-H., and Buehler, M. J. (2021a). "Deep learning model to predict complex stress
522 and strain fields in hierarchical composites." *Science Advances*, 7(15), eabd7416.
523 Yang, Z., Yu, C.-H., Guo, K., and Buehler, M. J. (2021b). "End-to-end deep learning method to
524 predict complete strain and stress tensors for complex hierarchical composite microstructures."
525 *Journal of the Mechanics and Physics of Solids*, 154, 104506.
526 Zhang, P., Hu, X., Yang, S., and Yao, W. (2019). "Modelling progressive failure in multi-phase
527 materials using a phase field method." *Engineering Fracture Mechanics*, 209, 105–124.
528 Zhou, K., Sun, H., Enos, R., Zhang, D., and Tang, J. (2021). "Harnessing deep learning for physics-
529 informed prediction of composite strength with microstructural uncertainties." *Computational*
530 *Materials Science*, 197, 110663.

## List of Tables

**TABLE 1.** Details of the various training and test datasets

| No. | Training set | | | Test set | | |
|---|---|---|---|---|---|---|
| | $V_f$ | $R_f(\mu m)$ | Size | $V_f$ | $R_f(\mu m)$ | Size |
| A | 0.5 | 4 | 800 | 0.5 | 4 | 200 |
| B | 0.5 | 2, 6 | 1000+1000 | 0.5 | 4 | 200 |
| C | 0.5 | 2, 4, 6 | 667+667+667 | 0.5 | 4 | 200 |
| D | 0.2, 0.5 | 4 | 1000+1000 | 0.3 | 4 | 200 |
| E | 0.2, 0.22, ...,0.5 | 4 | 130×16 | 0.3 | 4 | 200 |

**TABLE 2.** Error metrics obtained for each test set

| Dataset No. | Avg. NRMSE for damage field prediction | RMSRE for peak load prediction |
|---|---|---|
| A | 0.1290 | 0.0212 |
| B | 0.1589 | 0.0262 |
| C | 0.1156 | 0.0172 |
| D | 0.1477 | 0.0327 |
| E | 0.1481 | 0.0267 |

## List of Figures

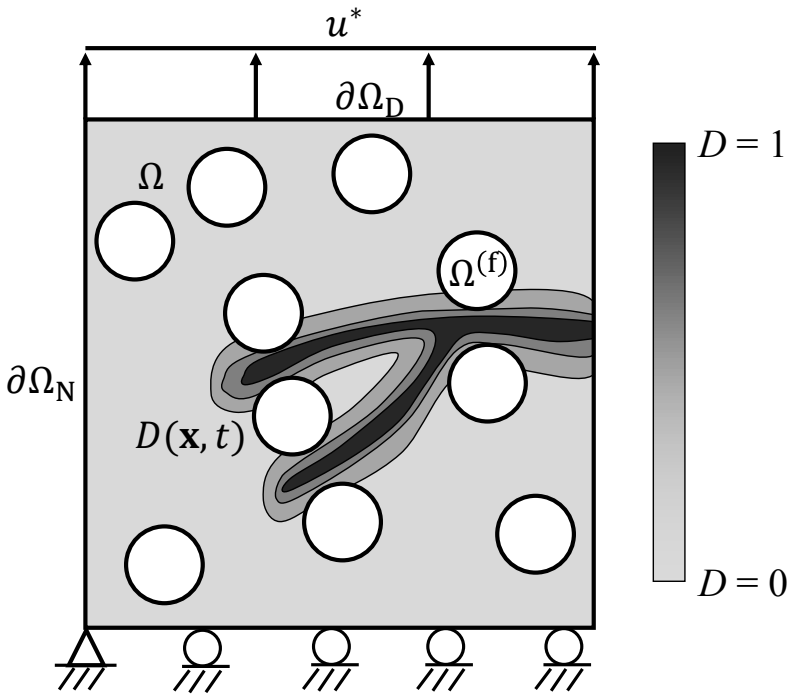


**Fig. 1.** Schematic of the solid domain $\Omega$ showing the representative volume element (RVE) of a fiber-reinforced composite and a diffused crack interface described by the phase field damage variable $D$.

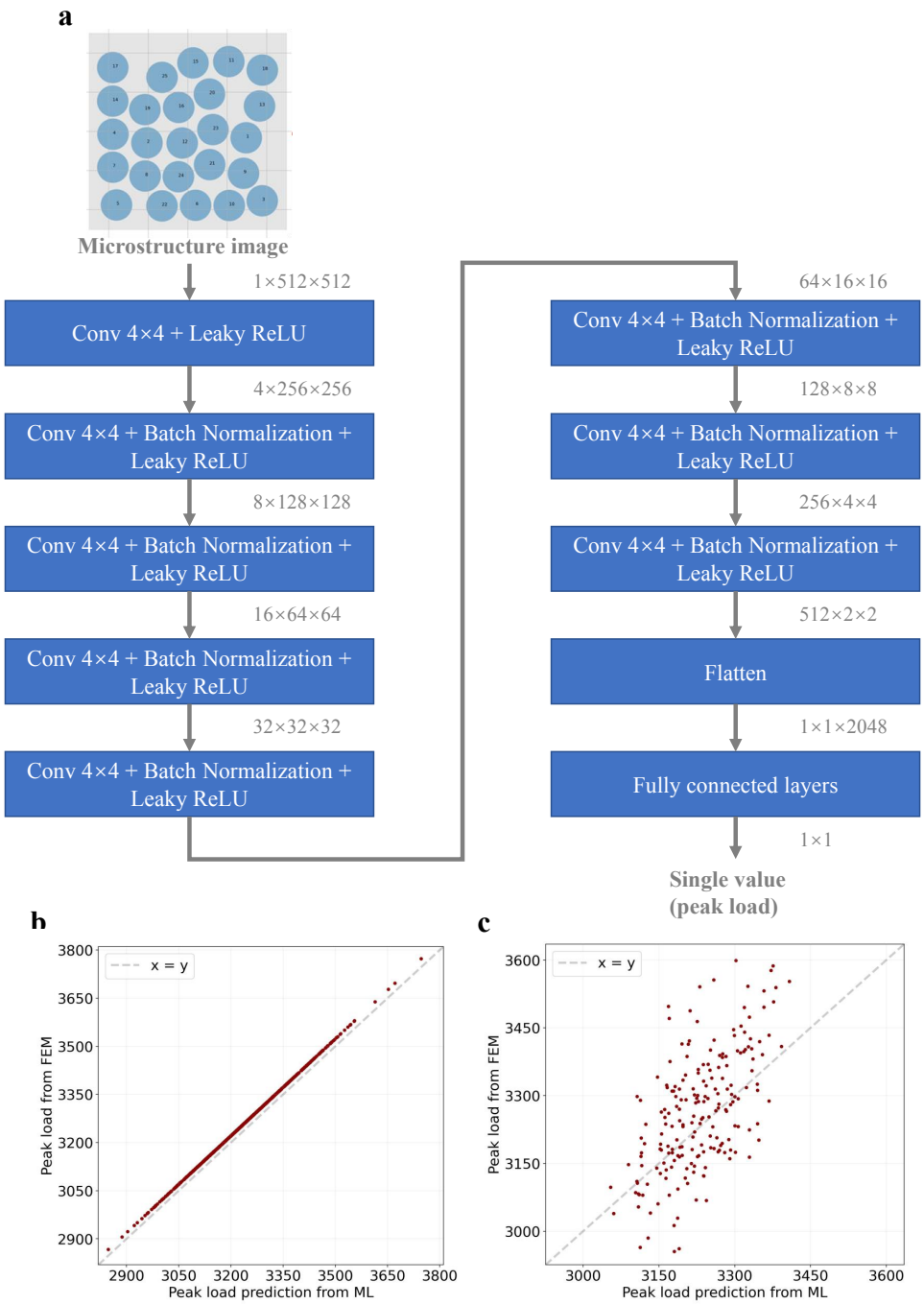


**Fig. 2.** (a) Architecture of the baseline or one-stage CNN-based model conditioned on images of the composite microstructure. Comparison of the peak load predicted by the model, and the peak load calculated by FEM evaluated on: (b) the training set and (c) the test set.

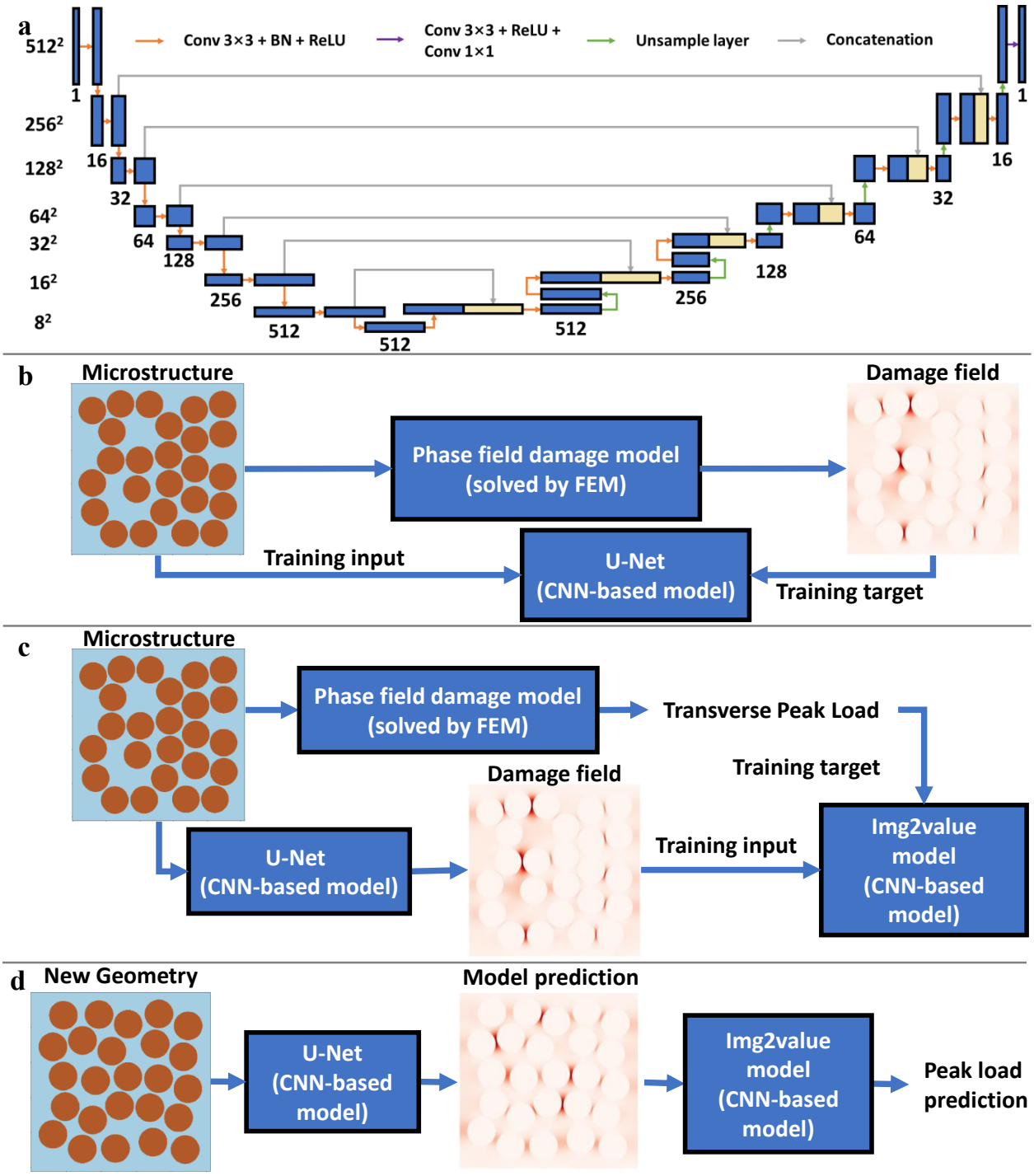


**Fig. 3.** (a) Architecture of the CNN based model that takes in microstructure geometry as input (left), and predicts its corresponding damage field at peak load (right); (b) training a CNN-based UNet for damage field prediction; (c) training a second CNN model for peak load prediction; (d) Model inference given a new RVE geometry.

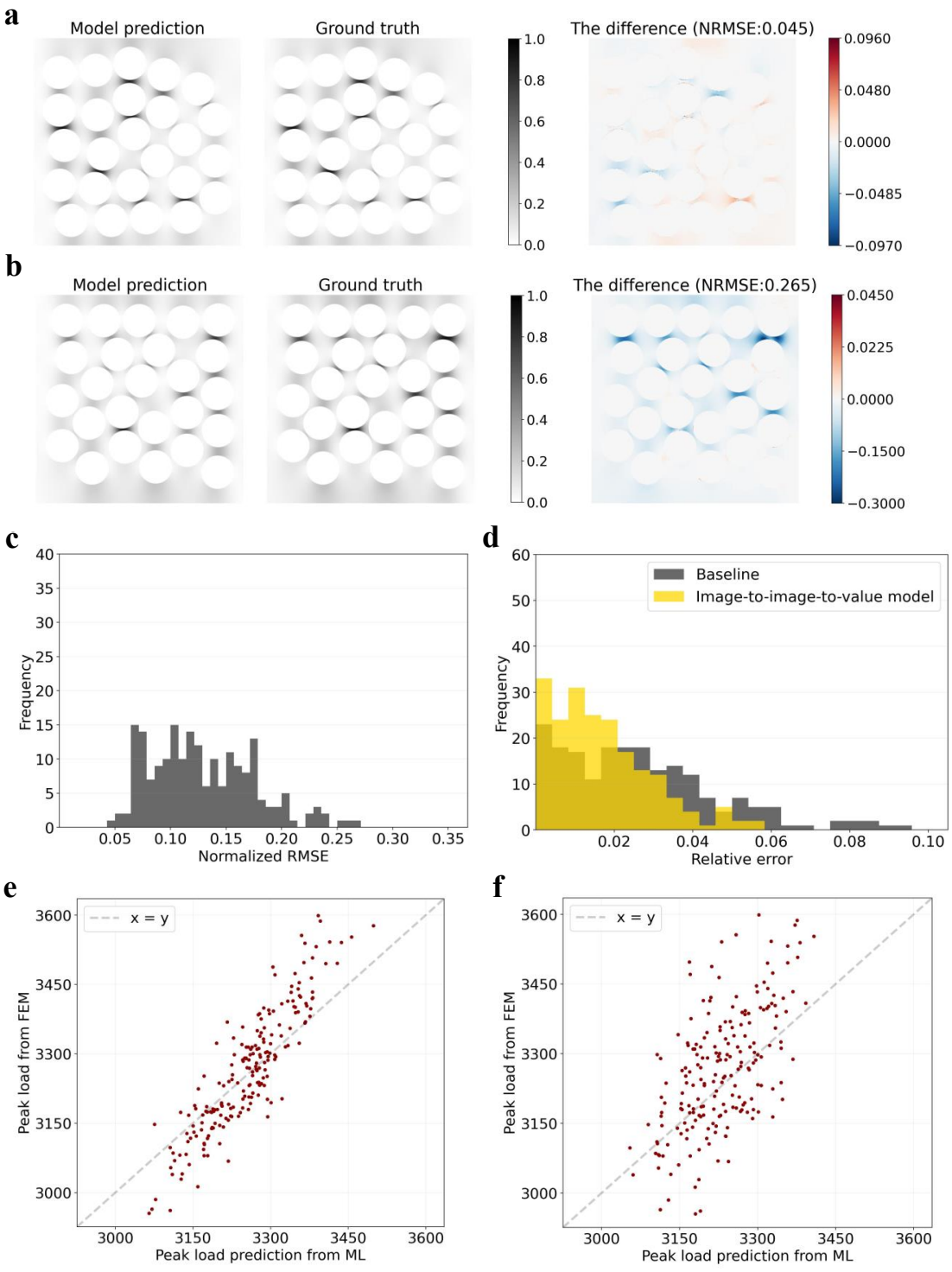


**Fig. 4.** Model performance evaluation on the test set with $R_f$ = 4 $\mu$m and $V_f$ = 0.5: Damage field predicted by the UNet model and the ground truth calculated by the FEM, and the difference between them for the case with (a) the smallest NRMSE and (b) the largest NRMSE; (c) The distribution of NRMSE; (d) The distribution of RMSRE for the peak load prediction. The peak load predicted by (e) the surrogate model and (f) the baseline model against that from the FEM.

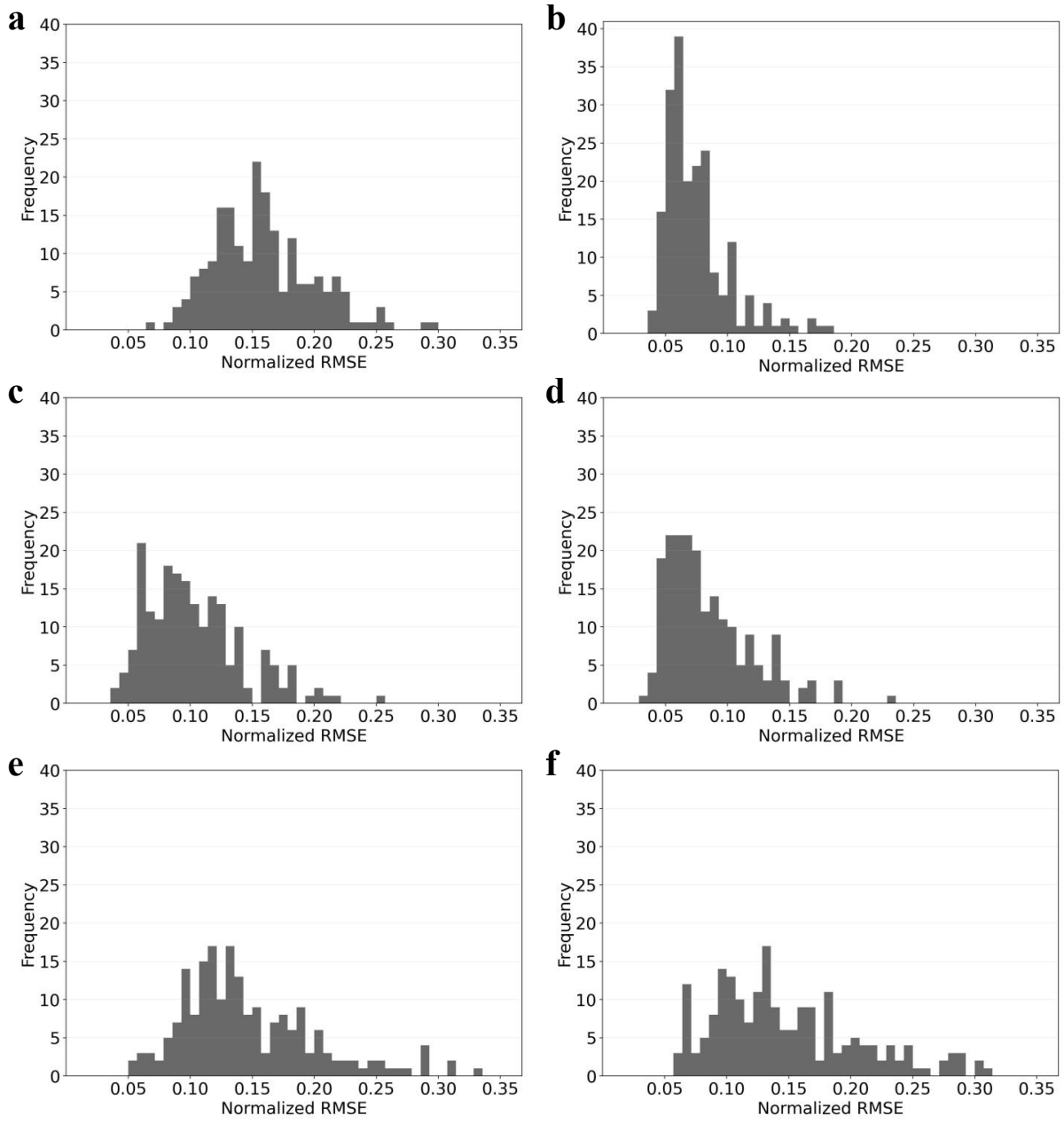


**Fig. 5.** The distribution of NRMSE for the damage field predicted by the model trained and evaluated with (a) dataset B; (b) dataset C (evaluated with $R_f = 2$ $\mu$m);(c) dataset C; (d) dataset C (evaluated with $R_f = 6$ $\mu$m);(e) dataset D; (f) dataset E.

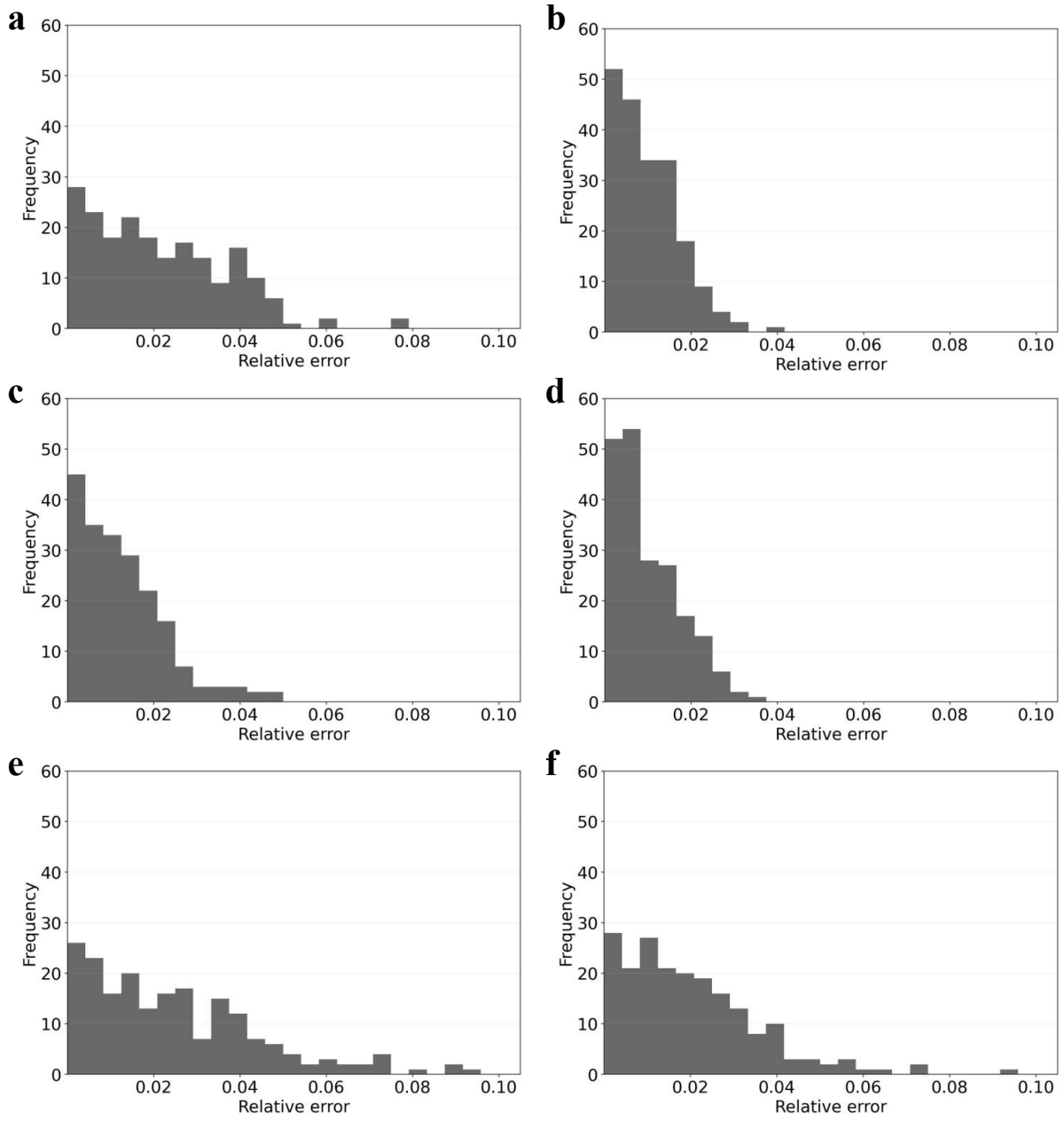


**Fig. 6.** The distribution of relative error for the peak load predicted by the model trained and evaluated with (a) dataset B; (b) dataset C (evaluated with $R_f = 2\ \mu m$);(c) dataset C; (d) dataset C (evaluated with $R_f = 6\ \mu m$);(e) dataset D; (f) dataset E.