Article

Predicting Mechanical Properties from Microstructure Images in Fiber-Reinforced Polymers Using Convolutional Neural Networks

Yixuan Sun [1,†] , Imad Hanhan [2,‡] , Michael D. Sangid [2] and Guang Lin [1,3,*]

1 School of Mechanical Engineering, Purdue University, West Lafayette, IN 47907, USA; sun650@purdue.edu
2 School of Aeronautics and Astronautics, Purdue University, West Lafayette, IN 47907, USA; ihanhan@purdue.edu (I.H.); msangid@purdue.edu (M.D.S.)
3 Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA
* Correspondence: guanglin@purdue.edu
† Current address: Argonne National Laboratory, Lemont, IL 60439, USA.
‡ Current address: Collins Aerospace, Charlotte, NC 28217, USA.

Abstract: Evaluating the mechanical response of fiber-reinforced composites can be extremely time-consuming and expensive. Machine learning (ML) techniques offer a means for faster predictions via models trained on existing input–output pairs and have exhibited success in composite research. This paper explores a fully convolutional neural network modified from StressNet, which was originally used for linear elastic materials, and extended here for a non-linear finite element (FE) simulation to predict the stress field in 2D slices of segmented tomography images of a fiber-reinforced polymer specimen. The network was trained and evaluated on data generated from the FE simulations of the exact microstructure. The testing results show that the trained network accurately captures the characteristics of the stress distribution, especially on fibers, solely from the segmented microstructure images. The trained model can make predictions within seconds in a single forward pass on an ordinary laptop, given the input microstructure, compared to 92.5 h to run the full FE simulation on a high-performance computing cluster. These results show promise in using ML techniques to conduct fast structural analysis for fiber-reinforced composites and suggest a corollary that the trained model can be used to identify the location of potential damage sites in fiber-reinforced polymers.

1. Introduction

Fiber-reinforced composites offer a transformative opportunity for the manufacturing of lightweight structures. One imposing impediment to the widespread adoption of these fiber-reinforced composites is the substantial time and cost necessary for certifying their structural integrity in engineering applications. Specifically, uncertainties regarding how the complex microstructures and defects within these materials influence the mechanical response have impeded their usage in structural applications. Agyei et al. [1] used a data-driven approach to provide a ranked order of the propensity of damage in fiber-reinforced composites caused by pores in the matrix, fiber orientation, fiber length, pores at fiber tips, resin-rich regions, and the resulting local stiffness variations, which provides a means to quantify the microstructure and detect features' role in damage. As discussed in the Materials Genome Initiative, a primary goal is to reduce the time and cost needed to bring new materials and structures to market by 50% [2]. To enable this goal, the NASA 2040 Vision discusses opportunities and needs for tool maturation, specifically articulating the importance of the development, verification, and validation of multiscale modeling approaches [3]. Towards this objective, this paper identifies rapid and efficient methods to

predict the mechanical response of fiber-reinforced composite materials with an emphasis on their microstructure, using machine learning techniques.

In composite materials, one approach that can offer reduced computational costs for microstructure-informed modeling is typically based on homogenization schemes or unit cell approaches. Broadly speaking, homogenization schemes rely on simulating the micromechanical behavior of a representative volume of the material and use the resulting material response to inform hierarchical modeling schemes at coarser length-scales [4–7]. In a similar fashion, the method of cells provides a periodically repeating unit of the material's microstructure and applies the continuity of the tractions and displacements over these cells to construct a continuum response of the solid [8–11]. In such modeling schemes, a reduction in the dimensionality of the modeling results in a loss of information, and as a consequence, these approaches are not able to capture local and rare events of damage. Furthermore, these schemes can be even more challenging for a thermoplastic matrix, since the matrix experiences significant non-linear deformation, which can be expensive to simulate. Therefore, these modeling events may not be appropriate to capture incipient damage.

Based on advancements in high-resolution characterization techniques coupled with microstructural sensitive modeling, a physics-based understanding and eventual prediction of the events that result in local damage is possible [12]; two examples are given here, which focus on damage initiating at an individual fiber, such as incipient damage or kink formation. For the damage of individual fibers in continuous fiber-reinforced composites, tomography has been used to study reorientation, breakage, and kink formation [13–16], which has led to the development of models at the fiber level [17,18]. Additionally, the formation and growth of incipient damage have been detected at the fiber tips of short discontinuous fibers, as the primary failure mechanism [19–22]. Through detailed modeling, Hanhan et al. individually meshed the microstructural features and determined that high hydrostatic stress was responsible as the primary driver of damage. In their analysis, the predictions of the site of damage through finite element (FE) simulations were in agreement with the same sites observed via in situ experimental observations [22]. While segmenting the fiber and void features identified through tomography and explicitly meshing these features for use within a FE model has provided promising results, it required 44.5 million elements to model a volume of 0.0096 mm^3 [22], which makes this method computationally impractical within an engineering workflow. For this reason, machine learning offers a promising direction for the exploration of computationally efficient tools and approaches.

Due to their extraordinary ability to learn non-linear mappings and their advances in computer vision tasks, deep neural networks, and a class of deep neural networks, convolutional neural networks (CNNs), have shown promising results in engineering research. CNNs operate on the input data by stacking multiple discrete convolutions with learnable kernels followed by nonlinear activations. This effectively captures the spatial dependencies of the data and extracts higher-level representations. Compared to other popular deep learning vision frameworks, such as Vision Transformers [23], Neural Operators [24], and diffusion models [25], CNNs have the advantage of lower memory and computation complexity while maintaining their efficacy in vision-related tasks and scientific applications [23,26], which makes them easier to train, more scalable, and faster during inference. CNNs have been widely used in solving engineering problems since AlexNet first used a deep convolutional network to perform ImageNet classification tasks [27]. Characteristics of CNNs, such as translation invariants and weight sharing, allow them to perform well for computer vision tasks. Fully connected neural networks were trained and evaluated on pairs between selected material treatment conditions or design variables and outputs to predict the compressive strength of heat-treated woods and concrete [28–30]. However, since the information of composite materials cannot be naturally stored in vectors that are mutually independent, careful feature selection is imperative to training a successful fully connected neural network. In recent studies, to relax this requirement, because of their parameter-sharing and translation invariant

characteristics, CNNs have been utilized for predicting material properties by looking directly at material topology images. A range of neural networks have been used to predict the homogenized properties from a composite's representative volume element [31–36]. Gu et al. built a CNN to predict the toughness of composites based on the topology of base materials [35]. Yang et al. adopted a 3D CNN to predict the stiffness of high-contrast composites from generated microscale volume elements [36]. Hanakata et al. used a traditional CNN structure to estimate the yield stress and strain based on graphene containing their kirigami designs [37]. While these studies showed high accuracy in the evaluation of the CNN models, they mainly focused on predicting a single value from the microstructure data. In order to investigate incipient damage and kink formation, it is of interest to utilize deep learning methods to quickly obtain the stress field of a composite. Deep neural networks have been used to predict the stress fields originating from a simulated microstructure [38,39], while the present paper uses an experimentally determined 3D microstructure.

Khadilkar et al. developed a two-stream CNN to predict the stress field for a stereolithography 3D printing process, in which the network outputs a long vector of $O(10^5)$ dimension that was reshaped to match the size of 2D slices of the specimen [40]. Nie et al. developed a fully convolutional network, StressNet, with an encoder–decoder structure to predict the stress field of linear elastic cantilevered structures in an end-to-end manner with additional injected load information and displacement boundary conditions [41]. This network can preserve the structural information of the input while extracting higher representations for accurate prediction.

In predicting the stress field of a materials given the microstructure, CNN-based deep learning models are suitable candidates because of their ability to capture hierarchical or structural information embedded in the microstructure. Based on its success with linear elastic cantilevered structures, in this work, we adopted and simplified the structure of StressNet, and extended it for a non-linear FE simulation to investigate its generalization ability on fiber-reinforced polymers. We only took the 2D segmented microstructure images as the network input without additional information such as load or boundary conditions. More specifically, the convolutional layers were regarded as feature extractors that slide through the input and perform the discrete convolution operations, generating feature maps containing high-level representations. During this process, the same convolutional filter was used for the entire input so that the trainable weights were shared throughout the input. This limited the number of trainable parameters that can make the training process more efficient. The downsampling operations in the pooling layers achieved translational invariance, making the absolute location of a certain pattern less important. A dataset containing 5321 2D microstructure slices sampled from segmented X-ray tomography images of a composite specimen and its corresponding FE simulation was used to train and validate the network. The trained model can make local stress predictions in a single forward pass from the given microstructure images within seconds on a laptop, compared to 92.5 h needed to run the full FE simulation on a high-performance computing cluster. According to the evaluation results, the trained network can output the stress distribution and capture important characteristics, especially on fibers, from the corresponding segmented microstructure 2D slices. The rest of the paper is organized as follow: Sections 2 and 3 describe the data generation process; Section 4 covers the workflow for training and evaluating the neural network; Section 5 shows the evaluation results and discussion; and Section 6 summarizes this work.

2. Experimental Methods

In order to provide a dataset for sampling that can be used for both neural network training and testing, the exact microstructure of a fiber-reinforced thermoplastic composite was analyzed. Specifically, the material studied through FE modeling was an injection-molded composite where the polymer matrix was polypropylene, and the fiber fillers were E-glass fibers. The E-glass fibers were approximately 10 μm in diameter, and had varying

lengths and orientations in the final composite specimen, due to the injection molding process [42]. The injection-molded part was a cylindrical rod with a diameter of 1.27 cm and a length of 45.72 cm, where the injection-molding direction was in the length direction of the cylinder (Z-axis in Figure 1). The cylindrical rod was then machined into a smaller dog-bone-shaped specimen with a gauge section diameter of 2.5 mm and a gauge section length of 5 mm.

In order to extract the 3D properties of the microstructure, as well as create a 3D FE model that can compute the local stresses within the microstructure, X-ray micro-computed tomography was conducted. The data acquisition was conducted at the Advanced Photon Source in Argonne National Laboratory at beam-line 2-BM using synchrotron X-rays. X-ray projections were acquired using X-ray energy of 25 keV and a detector exposure time of 100 ms. The detector used for capturing the X-ray projections was placed 75 mm downstream from the specimen. The specimen was rotated at 0.5°/s through a range of 180°, where an X-ray projection was acquired every 0.12°. The total 1500 X-ray projections were reconstructed using TomoPy [43] into a 3D image volume with dimensions of 3.33 by 3.33 by 1.61 mm (with a pixel size of 1.3 µm).

The set-up for the X-ray micro-computed tomography is shown in Figure 1, which is reprinted with permission from [44]. The segmentation and feature-identification steps are described by Agyei and Sangid [45], which also shows the fiber length distribution, fiber orientation distribution, and porosity distribution. The X-ray micro-computed tomography technique does have limitations (namely the characterization of a small region of interest that may divide fibers on the border of the characterized region, as well as restrictions on feature resolution resulting from a pixel size of 1.3 µm). The X-ray micro-computed tomography results have been directly compared to destructive optical characterization [46]. The discrepancies between the two methods have been quantified and the resulting uncertainties in the stiffness resulting from differences in the segmented microstructure and pore defects have been reported [46].
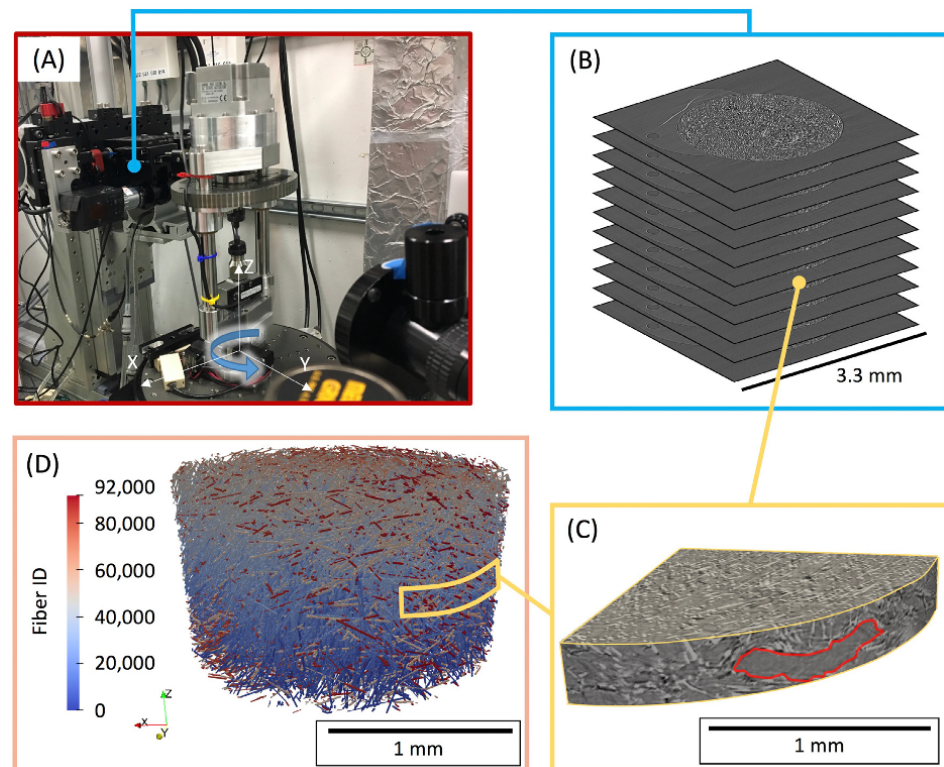

Figure 1. An overview of the X-ray micro-computed tomography: (**A**) set-up, (**B**) reconstructed tomography volume, (**C**) region of interest, and (**D**) segmented fibers in the volume. This figure is reprinted from Ref. [44], with permission from Elsevier.

3. Finite Element Simulation

The 3D image volume required several image-processing steps in order to extract the microstructural features of interest for the FE simulation. Specifically, a total of four feature types needed to be detected and extracted: the exterior edge of the specimen, the glass fibers, the porosity, and the polymer matrix. The exterior edge of the specimen was detected using an in-house MATLAB 2020b algorithm that uses an initial guess of the center and radius of the specimen. This is followed by automated image processing, which maps the image intensity values from a range of [0, 1] to a range of [0.4, 0.9], converts the image into a binary image using a threshold of 0.655 of the median intensity of the image, dilates the binarized image using a disk structural element with a radius of four pixels, and finally adjusts the binarized image by filling any holes [22]. The specimen edge detection was verified using ModLayer [47].

The glass fibers were detected using an iterative and supervised 2D and 3D combined algorithm [45]. Next, the porosity was detected using a combination of Weka machine learning segmentation [48] and manual correction using ModLayer [47]. Lastly, voxels within the interior of the specimen that were not classified as fibers or pores were labeled as the polymer matrix. The final detection of all the microstructural features was verified using ModLayer [47]. A region of interest was virtually down-selected and can be seen in Figure 2A. This region was chosen because it was determined to be a critical region that exhibited significant experimental damage within the microstructure [22].
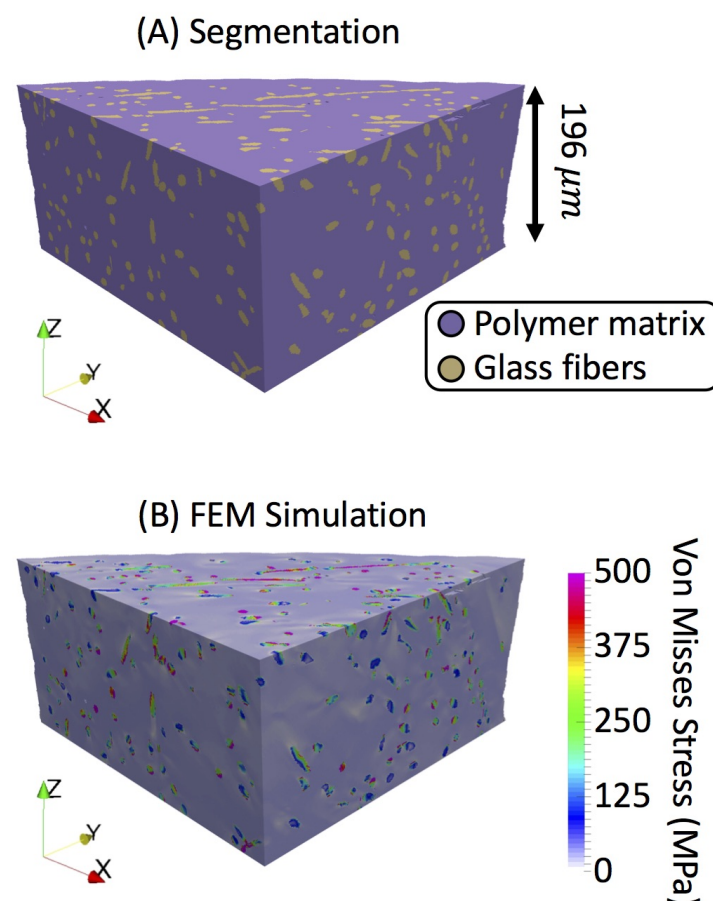
Figure 2. The segmented specimen microstructure is shown in (**A**), where the blue state represents the polymer and the yellow state represents fibers, and the FE simulation result is shown in (**B**), where the Von Mises stress is plotted.

The meshing of this down-selected region was conducted in ParaView. First, voxels that were characterized as porosity were removed from the 3D volume, effectively

creating gaps within the polymer matrix. The remaining voxels (which corresponded to the fibers and the matrix) were meshed using tetrahedral elements directly from the voxelated microstructure, which resulted in ideal and geometrically identical tetrahedral elements [22]. In total, the down-selected region of interest shown in Figure 2 contained 44.5 million elements.

The meshed fiber elements were assigned linear elastic mechanical properties with an elastic modulus of 72.4 GPa and a Poisson's ratio of 0.2 [49,50]. The meshed matrix elements were assigned non-linear mechanical behavior properties using a multi-linear isotropic hardening model [51]. The positive X, negative Y, and negative Z surfaces (Figure 2) were assigned roller boundary conditions. The positive Z surface was displaced positively by 7.8 μm [22]. The free surface of the specimen was allowed to remain a free surface with no boundary condition. The FE simulation was solved in Abaqus using 300 parallel processors with 1.92 TB of memory in 92.5 h. A sample of the result of the simulation is shown in Figure 2B, where the Von Misses stress is plotted. The result of the simulation has been validated by Hanhan et al., who showed that locations of high stress spatially matched with locations of local experimental damage initiation (in the form of micro-void nucleation) [22]. For the machine learning application in this work, the segmented microstructure and the FE stress in the loading direction, σ_{zz}, was used for data sampling, training, and testing.

4. Methodology: Deep Learning Framework

4.1. Data Sampling

Since the segmented microstructure shown in Figure 2A was represented by unit-less square voxels (with a voxel size of 1.3 by 1.3 by 1.3 μm), and the FE simulation was represented by tetrahedral mesh elements measured in μm, the stress field data were resized using the `scipy.ndimage.zoom` method via order 3 spline interpolation to match the segmented data. Data points were selected from the segmented microstructure and the corresponding stress field, with a sampling window size of 32×32 pixels. This particular sampling window size was chosen in order to (i) contain, on average, the entire fiber's cross-section, (ii) generate as many samples as possible, and (iii) serve the downsampling purpose in the network training process. A sampled data point was stored in a rank 3 array of the shape (32, 32, 2), with the first two axes indicating the spatial location of a voxel within the data point and the third axis representing the voxel microstructure type and stress value. During the sampling process, there was no overlapping between the sampling windows, which ensured that each data point was unique in the dataset. To investigate the importance of the sampling plane to the model performance, data points of equal size were sampled from three planes, namely the xy-plane, xz-plane, and yz-plane, and the data from different planes were used to train and evaluate the networks.

4.2. Preprocessing

In order to evaluate the model performance on unseen data, the sampled data points were randomly split into a training set and a testing set with no overlapping in between. The training set contained 80% of the entire dataset and the testing set contained 20%. In addition, since the standardization or normalization of the input data can accelerate the training process, the input of the model was standardized. In this work, all microstructure images from the sampled data points were standardized using the sample mean and standard deviation with the same shape as the input in the training set. Specifically, the input data, *X*, for both the training set and the testing set were standardized as shown in Equation (1),

$$X_{norm} = \frac{X - \mu_{train}}{\sigma_{train}} \tag{1}$$

where μ_{train} and σ_{train} are the sample mean and standard deviation from the data points in the training set, respectively. It is important to note that if the trained model is used for new predictions, the new input data would need to be standardized in the same manner.

4.3. Network Architecture

A CNN-based network, shown in Figure 3, was adopted from StressNet [41] for predicting the stress field given the microstructure. StressNet was originally developed to identify the stress field of linear elastic materials. In this study, StressNet was adopted and modified to extend to the non-linear mechanical behavior of fiber-reinforced polymers. The network is comprised of an input layer, an output layer, and 11 hidden layers in between. The input is the segmented microstructure array of size 32×32, denoted by X. A typical convolutional block consists of one or more convolutional layers and a pooling layer. In this section, we describe the convolution operations and pooling operations, which are accompanied by the details of the network architecture in Table 1.

Table 1. Network architecture description.

Operation Layers		Number of Filters	Kernel Size	Stride	Padding	Output Size
Input Segmented Microstructure		-	-	-	-	$32 \times 32 \times 1$
Convolution Layer	ReLU	32	3×3	1×1	SAME	$32 \times 32 \times 32$
Pooling	Max pooling	-	2×2	2×2	SAME	$16 \times 16 \times 16$
Convolution Layer	ReLU	64	3×3	1×1	SAME	$16 \times 16 \times 64$
Pooling	Max pooling	-	2×2	2×2	SAME	$8 \times 8 \times 64$
SE ResNet Layer	ReLU	64	3×3	1×1	SAME	$8 \times 8 \times 64$
SE ResNet Layer	ReLU	64	3×3	1×1	SAME	$8 \times 8 \times 64$
SE ResNet Layer	ReLU	64	3×3	1×1	SAME	$8 \times 8 \times 64$
SE ResNet Layer	ReLU	64	3×3	1×1	SAME	$8 \times 8 \times 64$
SE ResNet Layer	ReLU	64	3×3	1×1	SAME	$8 \times 8 \times 64$
Transposed Convolution	ReLU	64	3×3	2×2	SAME	$16 \times 16 \times 64$
Transposed Convolution	ReLU	32	3×3	2×2	SAME	$32 \times 32 \times 32$
Convolution Layer	-	1	3×3	1×1	SAME	$32 \times 32 \times 1$
Output Stress Field		-	-	-	-	$32 \times 32 \times 1$
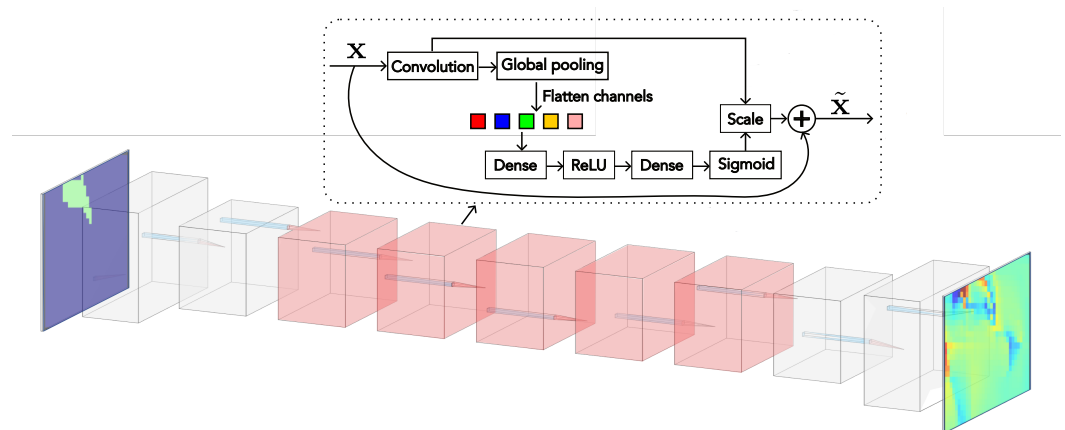
Figure 3. Convolutional neural network architecture with an encoder–decoder structure. This network takes microstructure images of the size 32×32 as input and outputs the corresponding stress field of the same size. The red highlighted blocks are squeeze-excitation residual blocks [52] and the rest are plain 2D convolution layers with MaxPooling.

Feature maps are obtained after convolving the input array with filters. This can be expressed as follows:

$$S_i = W_i \star X + b_i \qquad (2)$$

where i denotes the filter number in a layer, S_i is the ith feature map generated, W_i is the weight matrix associated with the ith filter, b_i is the bias matrix for the ith filter, and $\star$ is the convolution operator. At the end of each convolutional layer, an activation function was applied to the feature maps. Activation functions introduce non-linearity into the network, enabling the approximation of non-linear underlying functions. A rectified linear unit (ReLU) was placed as the activation function for each feature map due to its simplicity in calculating gradients and its ability to prevent gradient vanishing problems [53]. The formulation of the ReLU is

$$ReLU(s) = \begin{cases} s, & \text{if } s > 0 \\ 0, & \text{Otherwise} \end{cases} \tag{3}$$

where s is the entry of the feature map matrix S_i. The corresponding derivative with respect to s is

$$\frac{dReLU}{ds} = \begin{cases} 1, & \text{if } s > 0 \\ 0, & \text{Otherwise} \end{cases} \tag{4}$$

The final output of each convolutional layer is

$$O = ReLU(S_i) = ReLU(W_i \star X + b_i) \tag{5}$$

Pooling layers in convolutional blocks was performed as a means to down-sample the feature maps. Depending on the pooling kernel size, stride, and padding options, the original feature maps were transformed into an array with a smaller size according to a certain criterion, such as maximum, minimum, average, or global average. The pooling operations were all max-pooling in our network, in which the maximum value in each pooling window was preserved, except for the layers involving ResNet and SE blocks.

Among the hidden layers, some layers were made of residual blocks [54] and squeeze-excitation blocks [52]. These layers increased the network's representation power by capturing identical mapping and leveraging the importance of different channels.

4.4. Network Training

The implementation of the network was carried out in `Tensorflow v1.15`, a differential programming deep learning framework where the gradient-based optimization techniques, such as stochastic gradient descent (SGD) and Adam, can be easily achieved through auto-differentiation and back-propagation. The loss function was chosen as the mean squared loss (MSE) due to the nature of the regression problem. The optimizer for the network was Adam [55], which is a gradient-based, adaptive optimization method. Adam generally performs better than plain SGD due to its usage of momentum, leading to faster convergence and the ability to adaptively select a separate learning rate for each parameter as the training goes. A Tesla P100 GPU was utilized to train the network.

4.5. Evaluation Metrics

Due to the regression nature of the problem, we chose the coefficient of determination (R^2) as the metric for the model performance of the testing data. It is defined as

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{6}$$

where $SS_{res} = \sum (f - y)^2$ and $SS_{tot} = \sum (y - \bar{y})^2$. Here, f represents the predicted values from the CNN model, y represents the true values in the FE dataset, and $\bar{y}$ is the sample mean of y. Since R^2 measures the fraction of the variance in the data that can be explained by the model, a perfect model is expected to have an R^2 of 1.

5. Results and Discussion

In order to create a model that effectively uses the segmented microstructure as an input and generates σ_{zz} as the output, models with the same network architecture were trained on sampled data points from different planes. Data points sampled from the xy-plane, yz-plane, and xz-plane were used to train the models, and the models were evaluated on the designated testing set. The sampled datasets from the aforementioned three orthogonal planes were of the same size. They contained 5321 data points each, in which 20% were reserved for testing. All the networks were trained with 5000 epochs (the number of times that the network goes through the entire training set) with batch-normalization [56].

Figure 4A,B show the model training process on the 2D slices sampled from the xy-plane. As the number of epochs increased, the MSE values from both the training and testing decreased, and became stable after around 400 epochs, with a final testing error that was slightly higher than the training error. The opposite trend can be observed in the training process measured in R^2, where the scores on training and testing sets increased with the number of epochs. The curves also became flat after around 400 epochs, corresponding to a higher R^2 score for the training set compared to the testing set.
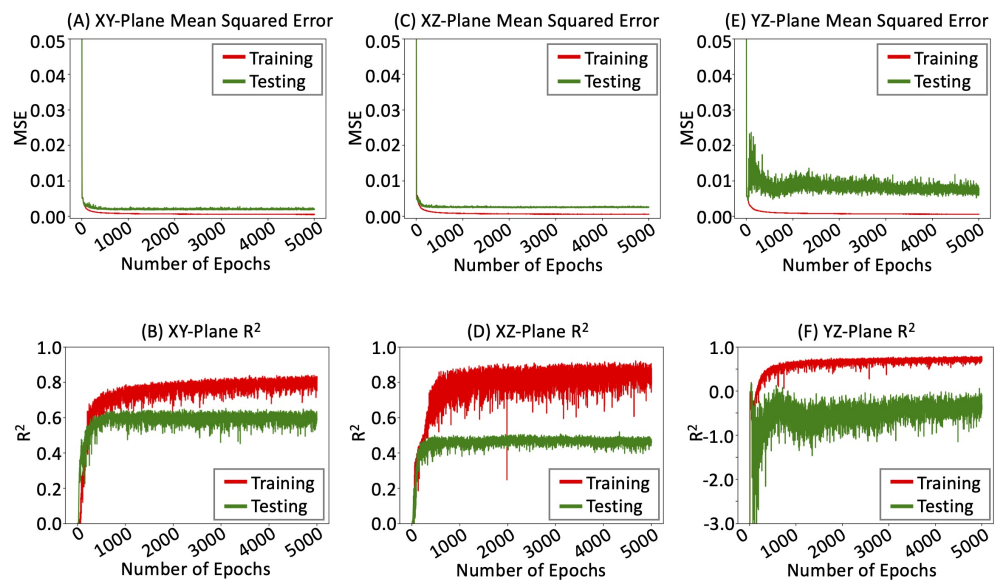
Figure 4. Mean squared error of training and testing losses (**top**) and curves of the coefficient of determination of model prediction on the training and testing sets (**bottom**) for each of the three orthogonal planes: (**A**,**B**) xy−plane, (**C**,**D**) xz−plane, and (**E**,**F**) yz−plane.

Figure 4C,D show the model training curves on the 2D slices sampled from the xz-plane. Similarly, the MSE values decreased as the number of epochs increased, while the gap between the training error and testing error was larger (after reaching stability) compared to the training process on the xy-plane data. In terms of the R^2 scores, although the score on the training set was comparable to that for the xy-plane shown in Figure 4B, the score on the testing set was much lower, making the data sampled from the xz-plane less reliable in predicting σ_{zz}.

The curves of the training process on slices sampled from the yz-plane are shown in Figure 4E,F. They generally follow the trend presented in the models trained on slices from the xy-plane and xz-plane. However, the model performance on the testing set was evidently worse than the previous models, although the training loss and R^2 score on the training set were close.

Table 2 shows the model's highest training and testing results using input data sampled from different planes throughout the training process to predict σ_{zz}. All models showed overfitting to some degree; however, the model that was trained on data sampled from the

xy-plane had the best performance, with an R^2 score on the testing set of 0.69, indicating that data sampled from the xy-plane contained the most relevant information for predicting the corresponding values of σ_{zz}.

Although the testing R^2 of the best performing model (using input microstructural data sampled from the xy-plane) was not 1, it was able to capture the overall σ_{zz} distribution. Figure 5 shows the visualization of the σ_{zz} fields (predicted and true) on select slices in the xy-plane. The overall predicted σ_{zz} on the testing data follows the corresponding stress fields, which were computed from the FE simulation, especially on areas with extreme values. The high-stress values consistently occurred within the fibers, and the trained model was able to accurately pinpoint the location where high stress was experienced within the fibers. Figure 6 shows the reconstructed composite block and 3D fibers from the stacked slices used in the network training and testing. Compared to the stress fields obtained from the FE simulation, the network-learned stress fields had less sharp edges and lost some details. Nonetheless, the locations and magnitudes of extreme stress values were consistent between the network prediction and FE simulation, especially relative to the fibers.
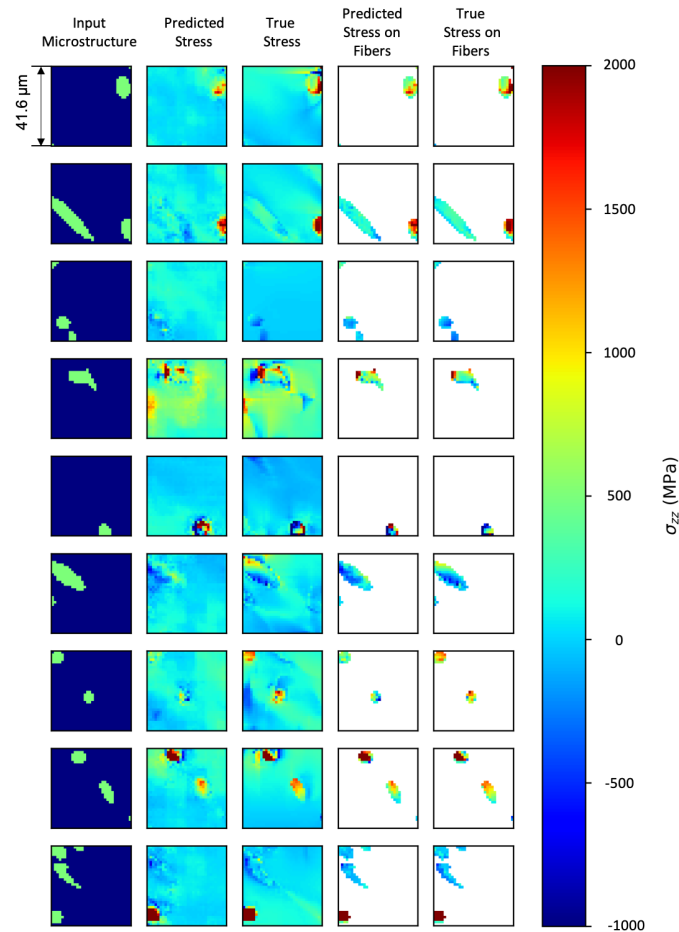
Figure 5. Visualization of the predicted stress fields from the CNN and the true stress fields from the FE simulation on 9 data points randomly selected from the testing set. The first column shows the input microstructure; the second column shows the corresponding predicted stress fields; the third column shows the true stress field obtained from FE simulation; and the fourth and fifth columns show the predicted and true stress within the fibers, respectively. All stress metrics correspond to the normal stress relative to the loading direction, σ_{zz}.

(a)

(b)
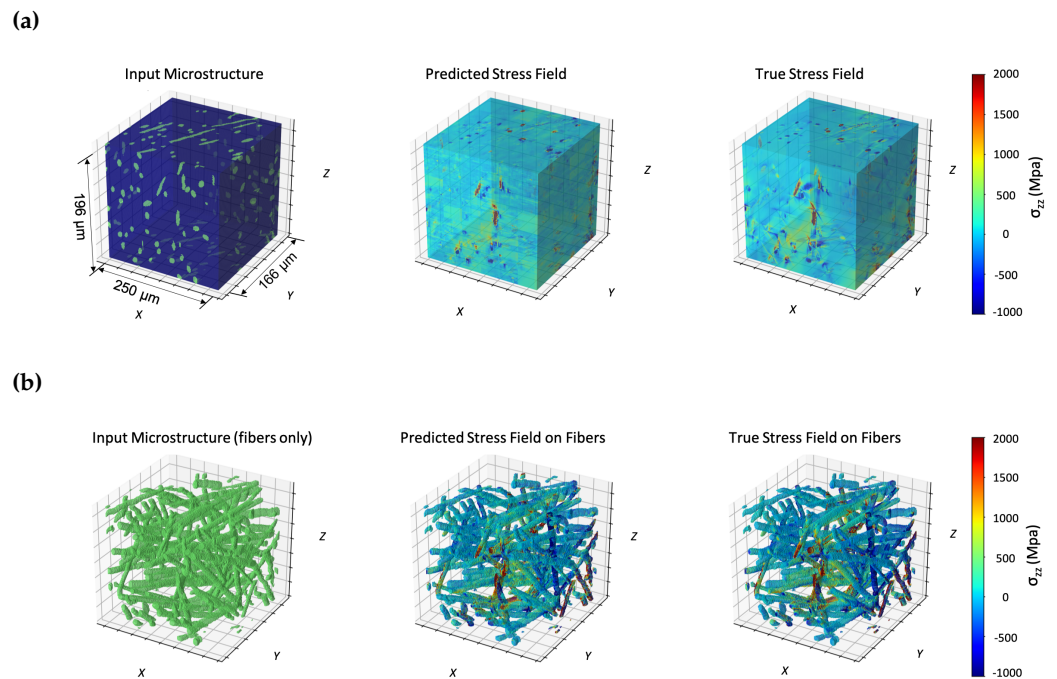
Figure 6. (**First column**) 3D phase reconstruction of the microstructure. Reconstructed stress fields (normal stress along the loading axis) from (**second column**) the predicted by the CNN model and (**third column**) the FE simulation dataset representing the training and testing sets. (**a**) The stress field over the entire sample volume of the composite. (**b**) The stress field within the discontinuous glass fibers.

The present study is based on the 3D microstructure that is directly determined from X-ray micro-computed tomography. The non-destructive X-ray micro-computed tomography imaging of the microstructure can take minutes on a synchrotron source and hours on a lab-source to acquire. Afterwards, image processing is necessary, which includes thresholding and segmentation, which can be time-consuming, while machine learning approaches for segmentation can greatly aid in this process. Afterwards, the FE simulation of the stress fields is reported to take 92.5 h to run on a supercomputer. The power of this method is that once trained, the CNN method takes seconds to run on an ordinary laptop, thus enabling the industrial use of process modeling for injection molding to be used [57,58] in combination with the stress fields produced by this CNN method to enable process parameter changes, design modifications, or aid in the model-based certification of the component.

Table 2. The coefficient of determination (R^2) of model performance of predicting the z-stress field on data sampled from different planes.

Stages	xy-Plane	xz-Plane	yz-Plane
Training	0.88	0.82	0.80
Testing	0.69	0.51	0.33

Some limitations exist in this study. First, due to the high computational cost of the FE simulation, there was only one fiber-reinforced composite specimen available for network training and testing. This not only poses difficulties in training a network that has a strong generalization ability, but it also creates challenges in validating the trained network on other composites. Second, no-load, displacement conditions or boundary conditions were taken into account as part of the network input (as the CNN input was restricted to the

stress field results from the FE simulation, which did not include the model set-up or evolution of the stress fields during loading), which might have prevented the network from learning a highly precise mapping to the stress field. It also limits the applicable scenarios of the trained network, since it is only expected to work when predicting the stress field of a composite that is under the same conditions as the one used in training. Lastly, the slices were only sampled in two dimensions and from orthogonal planes, resulting in a possible loss of information. These limitations are to be addressed in future studies.

6. Conclusions

In this work, a fully convolutional neural network with an encoder–decoder structure was used to predict the stress field from the 2D microstructure slices of a fiber-reinforced polymer specimen. The CNN model was trained on FE simulations of an exact 3D microstructure, which demonstrated that the areas of highest stress corresponded to damage initiation via in situ X-ray micro-computed tomography [22]. Hence, it is postulated as a corollary that the CNN model could identify regions of incipient damage relative to the microstructural features of a composite. Furthermore, the model was able to make fast predictions from the given 2D microstructure images via a single forward pass. The training and testing results show that the network performed best on segmented microstructural images sampled from the xy-plane to predict the normal stress field in the z-direction, σ_{zz}. Despite the suboptimal testing metric scores on the complete stress field, the network was able to capture the important characteristics of the stress distribution, especially on fibers, based on the visualized results. This network has proven to be useful in learning microstructure stress mapping for non-linear fiber-reinforced polymers. It can help accelerate the evaluation of the structural integrity and potentially assist in the identification of the locations where incipient damage formation could occur. Nevertheless, the trained network is only expected to work on fiber-reinforced composites under the same settings as the specimen used in this study. To expand the generalization ability of a trained network, some future work can involve (1) adding load and various boundary conditions into the training process using the same network; (2) adopting the Conditional Generative Adversarial Network (CGAN) [59] to learn stress fields conditioned on different boundary conditions so that new stress fields can be generated based on other boundary conditions; and (3) sampling 3D blocks and utilizing a 3D convolution operation to extract features, which is expected to preserve more spatial information of the data points.

Author Contributions: Y.S.: Methodology, Software, Validation, Formal analysis, Data Curation, Visualization, Writing—Original Draft. I.H.: Conceptualization, Methodology, Investigation, Data Curation, Visualization, Writing—Review and Editing. M.D.S.: Conceptualization, Investigation, Resources, Writing—Review and Editing, Supervision, Project administration, Funding acquisition. G.L.: Conceptualization, Investigation, Methodology, Resources, Writing—Review and Editing, Supervision, Project administration, Funding acquisition. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The processed data required to reproduce these findings are available to download from https://github.com/iamyixuan/stress_fiber_polymer, (accessed on 20 September 2024).

Conflicts of Interest: The authors declare no competing financial interests.

References

1. Agyei, R.; Hanhan, I.; Sangid, M.D. A data-driven microstructural rationale for micro-void nucleation in discontinuous fiber composites. *J. Thermoplast. Compos. Mater.* **2023**, *36*, 1694–1716. [CrossRef]
2. *Materials Genome Initiative for Global Competitiveness*; Executive Office of the President, National Science and Technology Council: Washington, DC, USA, 2011.
3. Liu, X.; Furrer, D.; Kosters, J.; Holmes, J. *Vision 2040: A Roadmap for Integrated, Multiscale Modeling and Simulation of Materials and Systems*; NASA Technical Report: Hampton, VA, USA, 2018.
4. Kammoun, S.; Doghri, I.; Adam, L.; Robert, G.; Delannay, L. First pseudo-grain failure model for inelastic composites with misaligned short fibers. *Compos. Part A Appl. Sci. Manuf.* **2011**, *42*, 1892–1902. [CrossRef]
5. Aboudi, J. Micromechanical analysis of thermo-inelastic multiphase short-fiber composites. *Compos. Eng.* **1995**, *5*, 839–850. [CrossRef]
6. Poulios, K.; Niordson, C.F. Homogenization of long fiber reinforced composites including fiber bending effects. *J. Mech. Phys. Solids* **2016**, *94*, 433–452. [CrossRef]
7. Fliegener, S.; Luke, M.; Gumbsch, P. 3D microstructure modeling of long fiber reinforced thermoplastics. *Compos. Sci. Technol.* **2014**, *104*, 136–145. [CrossRef]
8. Aboudi, J. Micromechanical analysis of composites by the method of cells. *Appl. Mech. Rev.* **1989**, *42*, 193–221. [CrossRef]
9. Aboudi, J.; Arnold, S.; Bednarcyk, B. *The Generalized Method of Cells Micromechanics. Micromechanics of Composite Materials—A Generalized Multiscale Analysis Approach*; Butterworth-Heinemann: Oxford, UK, 2013.
10. Böhm, H.J.; Eckschlager, A.; Han, W. Multi-inclusion unit cell models for metal matrix composites with randomly oriented discontinuous reinforcements. *Comput. Mater. Sci.* **2002**, *25*, 42–53. [CrossRef]
11. Modniks, J.; Andersons, J. Modeling the non-linear deformation of a short-flax-fiber-reinforced polymer composite by orientation averaging. *Compos. Part B Eng.* **2013**, *54*, 188–193. [CrossRef]
12. Sangid, M.D. Coupling in situ experiments and modeling–Opportunities for data fusion, machine learning, and discovery of emergent behavior. *Curr. Opin. Solid State Mater. Sci.* **2019**, *24*, 100797. [CrossRef]
13. Emerson, M.J.; Jespersen, K.M.; Dahl, A.B.; Conradsen, K.; Mikkelsen, L.P. Individual fibre segmentation from 3D X-ray computed tomography for characterising the fibre orientation in unidirectional composite materials. *Compos. Part A Appl. Sci. Manuf.* **2017**, *97*, 83–92. [CrossRef]
14. Garcea, S.; Sinclair, I.; Spearing, S.; Withers, P. Mapping fibre failure in situ in carbon fibre reinforced polymers by fast synchrotron X-ray computed tomography. *Compos. Sci. Technol.* **2017**, *149*, 81–89. [CrossRef]
15. Emerson, M.J.; Wang, Y.; Withers, P.J.; Conradsen, K.; Dahl, A.B.; Dahl, V.A. Quantifying fibre reorientation during axial compression of a composite through time-lapse X-ray imaging and individual fibre tracking. *Compos. Sci. Technol.* **2018**, *168*, 47–54. [CrossRef]
16. Wang, Y.; Burnett, T.L.; Chai, Y.; Soutis, C.; Hogg, P.J.; Withers, P.J. X-ray computed tomography study of kink bands in unidirectional composites. *Compos. Struct.* **2017**, *160*, 917–924. [CrossRef]
17. Wind, J.L.; Steffensen, S.; Jensen, H.M. Comparison of a composite model and an individually fiber and matrix discretized model for kink band formation. *Int. J. Non-Linear Mech.* **2014**, *67*, 319–325. [CrossRef]
18. Davidson, P.; Waas, A.M. A Micromechanics based model for the prediction of compression fatigue failure of fiber reinforced composites. In Proceedings of the AIAA Scitech 2020 Forum, Orlando, FL, USA, 6–10 January 2020; pp. 1420, 2103.
19. Sato, N.; Kurauchi, T.; Sato, S.; Kamigaito, O. Microfailure behaviour of randomly dispersed short fibre reinforced thermoplastic composites obtained by direct SEM observation. *J. Mater. Sci.* **1991**, *26*, 3891–3898. [CrossRef]
20. Rolland, H.; Saintier, N.; Robert, G. Damage mechanisms in short glass fibre reinforced thermoplastic during in situ microtomography tensile tests. *Compos. Part B Eng.* **2016**, *90*, 365–377. [CrossRef]
21. Hu, X.; Fang, J.; Xu, F.; Dong, B.; Xiao, Y.; Wang, L. Real internal microstructure based key mechanism analysis on the micro-damage process of short fibre-reinforced composites. *Sci. Rep.* **2016**, *6*, 34761. [CrossRef]
22. Hanhan, I.; Agyei, R.F.; Xiao, X.; Sangid, M.D. Predicting Microstructural Void nucleation in Discontinuous fiber composites through coupled in-situ X-ray tomography experiments and Simulations. *Sci. Rep.* **2020**, *10*, 3564. [CrossRef]
23. Han, K.; Wang, Y.; Chen, H.; Chen, X.; Guo, J.; Liu, Z.; Tang, Y.; Xiao, A.; Xu, C.; Xu, Y.; et al. A survey on vision transformer. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 87–110. [CrossRef]
24. Kovachki, N.; Li, Z.; Liu, B.; Azizzadenesheli, K.; Bhattacharya, K.; Stuart, A.; Anandkumar, A. Neural operator: Learning maps between function spaces with applications to pdes. *J. Mach. Learn. Res.* **2023**, *24*, 1–97.
25. Croitoru, F.A.; Hondru, V.; Ionescu, R.T.; Shah, M. Diffusion models in vision: A survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 10850–10869. [CrossRef] [PubMed]
26. Khan, A.; Rauf, Z.; Sohail, A.; Khan, A.R.; Asif, H.; Asif, A.; Farooq, U. A survey of the vision transformers and their CNN-transformer based variants. *Artif. Intell. Rev.* **2023**, *56*, 2917–2970. [CrossRef]
27. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems, Lake Tahoe, NV, USA, 3–6 December 2012; pp. 1097–1105.
28. Tiryaki, S.; Aydın, A. An artificial neural network model for predicting compression strength of heat treated woods and comparison with a multiple linear regression model. *Constr. Build. Mater.* **2014**, *62*, 102–108. [CrossRef]

29. Khademi, F.; Akbari, M.; Jamal, S.M.; Nikoo, M. Multiple linear regression, artificial neural network, and fuzzy logic prediction of 28 days compressive strength of concrete. *Front. Struct. Civ. Eng.* **2017**, *11*, 90–99. [CrossRef]

30. Young, B.A.; Hall, A.; Pilon, L.; Gupta, P.; Sant, G. Can the compressive strength of concrete be estimated from knowledge of the mixture proportions?: New insights from statistical analysis and machine learning methods. *Cem. Concr. Res.* **2019**, *115*, 379–388. [CrossRef]

31. Breuer, K.; Stommel, M. Prediction of short fiber composite properties by an artificial neural network trained on an RVE database. *Fibers* **2021**, *9*, 8. [CrossRef]

32. Gholami, K.; Ege, F.; Barzegar, R. Prediction of composite mechanical properties: Integration of deep neural network methods and finite element analysis. *J. Compos. Sci.* **2023**, *7*, 54. [CrossRef]

33. Stamopoulos, A.; Tserpes, K.; Dentsoras, A. Quality assessment of porous CFRP specimens using X-ray Computed Tomography data and Artificial Neural Networks. *Compos. Struct.* **2018**, *192*, 327–335. [CrossRef]

34. Jia, Y.; Yu, G.; Du, J.; Gao, X.; Song, Y.; Wang, F. Adopting traditional image algorithms and deep learning to build the finite model of a 2.5 D composite based on X-ray computed tomography. *Compos. Struct.* **2021**, *275*, 114440. [CrossRef]

35. Gu, G.X.; Chen, C.T.; Buehler, M.J. De novo composite design based on machine learning algorithm. *Extrem. Mech. Lett.* **2018**, *18*, 19–28. [CrossRef]

36. Yang, Z.; Yabansu, Y.C.; Al-Bahrani, R.; Liao, W.k.; Choudhary, A.N.; Kalidindi, S.R.; Agrawal, A. Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets. *Comput. Mater. Sci.* **2018**, *151*, 278–287. [CrossRef]

37. Hanakata, P.Z.; Cubuk, E.D.; Campbell, D.K.; Park, H.S. Accelerated search and design of stretchable graphene kirigami using machine learning. *Phys. Rev. Lett.* **2018**, *121*, 255304. [CrossRef] [PubMed]

38. Bhaduri, A.; Gupta, A.; Graham-Brady, L. Stress field prediction in fiber-reinforced composite materials using a deep learning approach. *Compos. Part B Eng.* **2022**, *238*, 109879. [CrossRef]

39. Shokrollahi, Y.; Nikahd, M.M.; Gholami, K.; Azamirad, G. Deep learning techniques for predicting stress fields in composite materials: A superior alternative to finite element analysis. *J. Compos. Sci.* **2023**, *7*, 311. [CrossRef]

40. Khadilkar, A.; Wang, J.; Rai, R. Deep learning–based stress prediction for bottom-up SLA 3D printing process. *Int. J. Adv. Manuf. Technol.* **2019**, *102*, 2555–2569. [CrossRef]

41. Nie, Z.; Jiang, H.; Kara, L.B. Stress Field Prediction in Cantilevered Structures Using Convolutional Neural Networks. *J. Comput. Inf. Sci. Eng.* **2020**, *20*, 011002. [CrossRef]

42. Bailey, R.; Kraft, H. A study of fibre attrition in the processing of long fibre reinforced thermoplastics. *Int. Polym. Process.* **1987**, *2*, 94–101. [CrossRef]

43. Gürsoy, D.; De Carlo, F.; Xiao, X.; Jacobsen, C. TomoPy: A framework for the analysis of synchrotron tomographic data. *J. Synchrotron Radiat.* **2014**, *21*, 1188–1193. [CrossRef]

44. Hanhan, I.; Sangid, M.D. Damage propagation in short fiber thermoplastic composites analyzed through coupled 3D experiments and simulations. *Compos. Part B Eng.* **2021**, *218*, 108931. [CrossRef]

45. Agyei, R.F.; Sangid, M.D. A supervised iterative approach to 3D microstructure reconstruction from acquired tomographic data of heterogeneous fibrous systems. *Compos. Struct.* **2018**, *206*, 234–246. [CrossRef]

46. Hanhan, I.; Agyei, R.; Xiao, X.; Sangid, M.D. Comparing non-destructive 3D X-ray computed tomography with destructive optical microscopy for microstructural characterization of fiber reinforced composites. *Compos. Sci. Technol.* **2019**, *184*, 107843. [CrossRef]

47. Hanhan, I.; Sangid, M.D. ModLayer: A MATLAB GUI Drawing Segmentation Tool for Visualizing and Classifying 3D Data. *Integr. Mater. Manuf. Innov.* **2019**, *8*, 468–475. [CrossRef]

48. Arganda-Carreras, I.; Kaynig, V.; Rueden, C.; Eliceiri, K.W.; Schindelin, J.; Cardona, A.; Seung, H.S. Trainable Weka Segmentation: A machine learning tool for microscopy pixel classification. *Bioinformatics* **2017**, *33*, 2424–2426. [CrossRef] [PubMed]

49. Agarwal, B.D.; Broutman, L.J.; Chandrashekhara, K. *Analysis and Performance of Fiber Composites*; John Wiley & Sons: Hoboken, NJ, USA, 1990.

50. Bauccio, M. *ASM Engineered Materials Reference Book*; CRC: Boca Raton, FL, USA, 1994.

51. Mohammadpour, E.; Awang, M.; Kakooei, S.; Akil, H.M. Modeling the tensile stress-strain response of carbon nanotube/polypropylene nanocomposites using nonlinear representative volume element. *Mater. Des.* **2014**, *58*, 36–42. [CrossRef]

52. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.

53. Xu, B.; Wang, N.; Chen, T.; Li, M. Empirical evaluation of rectified activations in convolutional network. *arXiv* **2015**, arXiv:1505.00853.

54. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

55. Kingma, D.P.; Ba, J. Adam: A method for stochastic optimization. *arXiv* **2014**, arXiv:1412.6980.

56. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* **2015**, arXiv:1502.03167.

57. Bay, R.S.; Tucker, C.L., III. Fiber orientation in simple injection moldings. Part I: Theory and numerical methods. *Polym. Compos.* **1992**, *13*, 317–331. [CrossRef]

58. VerWeyst, B.E.; Tucker, C.L., III; Foss, P.H.; O'Gara, J.F. Fiber orientation in 3-D injection molded features: Prediction and experiment. *Int. Polym. Process.* **1999**, *14*, 409–420. [CrossRef]
59. Mirza, M.; Osindero, S. Conditional generative adversarial nets. *arXiv* **2014**, arXiv:1411.1784.