INVESTIGATION OF INTERFACIAL PROPERTY WITH IMPERFECTION: A MACHINE LEARNING APPROACH

Sanjida Ferdousi

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

July 2023

APPROVED:

Yijie Steven Jiang, Major Professor
Hector R. Siller, Committee Members
Yunwei Xu, Committee Members
Herman Shen, Chair of Department of Mechanical Engineering
Paul S. Kruger, Dean of College of Engineering
Victor Prybutok, Dean of the Toulouse Graduate School

Ferdousi, Sanjida. *Investigation of Interfacial Property with Imperfection: A Machine Learning Approach*. Master of Science (Mechanical and Energy Engineering), July 2023, 37 pp., 13 figures, 85 numbered references.

Interfacial mechanical properties of adhesive joints are very crucial in board applications, including composites, multilayer structures, and biomedical devices. Establishing traction-separation (T-S) relations for interfacial adhesion can evaluate mechanical and structural reliability, robustness, and failure criteria. Due to the short range of interfacial adhesion such as micro to nanoscale, accurate measurements of T-S relations remain challenging. The advent of machine learning (ML) became a promising tool to predict materials behaviors and establish data-driven mechanical models. In this study, we integrated a state-of-the-art ML method, finite element analysis (FEA), and standard experiments to develop data-driven models for characterizing the interfacial mechanical properties precisely. Macroscale force-displacement curves are derived from FEA with incorporation of double cantilever beam tests to generate the dataset for ML model. The eXtreme Gradient Boosting (XGBoost) multi-output regressions and classifier models are used to determine T-S relations with R2 score of 98.8% and locate imperfections at the interface with accuracy of around 80.8%. The outcome of the XGBoost models demonstrated accurate predictions and fast calculation speed, outperforming several other ML methods. Using 3D printed double cantilever beam specimens, the performance of the ML models is validated experimentally for different materials. Furthermore, a XGBoost model-based package is designed to obtain different adhesive materials T-S relations without creating a database or training a model.

ACKNOWLEDGEMENTS

This was a wonderful period in my life, and I am grateful to have chosen this path to complete my master's degree. First of all, I would like to express my most sincere gratitude to my advisor Dr. Yijie Jiang for his continuous support and guidance. I am truly grateful that I have the opportunity of working with such an outstanding advisor. When I started, I did not have much of research experience and he guided me through each step from the beginning. He always appreciated my efforts, exposed me to some fascinating research and taught me many new skills. I would also like to thank my dissertation committee members for their insightful suggestions, encouragement, and immense support. I am deeply grateful to Dr. Hector R. Siller and Dr. Yunwei Xu for being part of my dissertation committee and providing valuable feedback. I received great support from both committee members for arranging the defense on short notice. It is also my pleasure to work with incredibly talented co-authors for their contributions to the papers that partly result from this dissertation. I would like to thank: Qiyi Chen, Mehrzad Soltani, Jiadeng Zhu, Pengfei Cao, Wonbong Choi, Rigoberto Advincula for their significant contributions to this research. I have received enormous help from my lab mates and friends throughout my master's journey and want to thank all of them for their supports. My deepest gratitude belongs to my beloved father Md Akkas Ali, my mother Nazma Begum, and my other family members who cared for, encouraged, and supported me along the way. Finally, I wanted to thank the most significant person my husband Md Mosharaf Hossain who supported and helped me in every step from the beginning. He always comforted me and guided me in developing skills. Without his immense support, help, and encouragement, I would not be able to continue my master's degree. I would like to dedicate this achievement to my husband for always being by my side.

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

USE OF PREVIOUSLY PUBLISHED MATERIAL

Material in sections 4.1-4.3, 5.1-5.4, and 6.1 is reproduced from the following publication:

> Sanjida Ferdousi, Qiyi Chen, Mehrzad Soltani, Jiadeng Zhu, Pengfei Cao, Wonbong Choi, Rigoberto Advincula & Yijie Jiang. (2021). “Characterize traction–separation relation and interfacial imperfections by data-driven machine learning models.” *Sci Rep* **11**, 14330, https://doi.org/10.1038/s41598-021-93852-y

CHAPTER 1

INTRODUCTION

1.1 Overview

Composite materials contain two or more distinctive property materials combined to achieve a new improved properties material.[1] Fiber-matrix adhesion in composites, interphase of multilayered structures is required to enhance the mechanical properties in the application of vehicle structures, soft robotics, and aerospace.[2,3] To mitigate interlaminar failure, delamination in composites, and fiber-matrix debonding, interlaminar failure theory has become quite popular to investigate the adhesion property. To evaluate the structural reliability of composites, the interfacial property plays a crucial role as it directly impacts the mechanism of load transfer.[4] Furthermore, enhanced mechanical properties of the composite can be ensured by resilient interface adhesion.

Due to the advent of composite materials, multi-layer structures in commercial manufacturing, the reliability of fiber-matrix adhesion has expanded to carry materials with higher performance levels. The cohesive zone method has been used to determine the damage mechanism through the traction-separation (T-S) relationship.[5] Obtaining an exact T-S relationship is difficult due to the complex relationship and microscale interaction involvement. For a comprehensive understanding of the fundamental mechanism of adhesion, extensive experimental explorations of cohesive behavior are required alongside finite element modeling.[6–9] Recently, the implementation of machine learning (ML) is becoming an effective approach in accurately capturing and predicting the highly complex data set.[10–14] Modeling the interlayer delamination using finite element analysis with the combination of ML approach can facilitate a better understanding of the behavior of the interface.

1.2 Problem Statement

Interfacial mechanical properties are important in composite materials and their applications, including vehicle structures, soft robotics, and aerospace. Determination of traction-separation (T-S) relations at interfaces can lead to evaluations of structural reliability, mechanical robustness, and failures in composites. Acquiring precise measurements of T-S relations remains challenging, since many studies indicates that intrinsic interfacial interaction generally happens at several nm range.[15,16] Furthermore, determining T-S relations experimentally requires extra measurements with an extensive and time-consuming setup.

1.3 Objective

The purpose of this study is to characterize the interfacial properties of a scalable wide range of material systems to overcome intricate, time-consuming setup procedures. A data-driven model for predicting the interfacial behavior of composite materials and establishing their mechanical models has become more and more efficient with the advent of machine learning (ML). This research aims to achieve the following objectives:

1. Quantify T-S relationships by developing a data-driven ML approach based on finite element analysis. Validate the ML model performance using experimental data on different materials.
2. Identify the location of interfacial imperfections via mechanical property fluctuation.
3. Develop a package based on the XGBoost regression model to facilitate users with interface T-S relations using F-D data without additional database setup.

CHAPTER 2

LITERATURE REVIEW

Interfacial mechanical properties quantify adhesion interactions between two surfaces at the microscale and are critical to the performance of heterogeneous engineering materials, such as multilayer structural materials and fiber reinforced composites.[3,4,17,18] Multilayer structural materials rely on interfacial mechanical properties to resist stress concentrations and to maintain structural integrity.[3] In addition, fiber reinforced composites highly rely on the interface between fiber and matrix, as it determines strength and stiffness.[19,20] Key interfacial properties, including fracture toughness and interfacial strength, can be measured via combination of numerical simulations and multiscale physical experiments, such as standard double cantilever beam (DCB) tests to *in-situ* scanning electronic microscope (SEM) fiber pull-out tests.[5,6,21–23] The DCB tests measure the fracture toughness of the interface and interfacial adhesion by calculating the critical energy release rate of the specimen by inducing controlled loads during crack propagation.[24–26] Alternatively, the in-situ fiber pull-out test assess the direct measurement of the interfacial strength by pulling out fibers from the matrix simulating the controlled conditions and characterize the interfacial shear strength, and energy dissipation mechanisms.[27] In addition to parameters such as total toughness and averaged interfacial strength, establishing a traction-separation (T-S) relation can fully quantify interfacial adhesion, understand crack propagation and fracture,[7] and thus lead to the novel design and fabrication of high-performance composites.[28–31]

Cohesive zone approach is employed to analyze the localized fracture and crack propagation of the material to capture the crack growth and fracture characteristics of the materials. Bonded interface in cohesive zone is divided into two region damage initiation in the

crack tip and damage evolution in the separated material. Figure 2.1 represent the cohesive zone in the composite layer and the behavior of the traction-separation (T-S) law. The cohesive zone is characterized by the T-S law which represents the relation between the applied force and corresponding displacement of interfaces in materials. Under loading conditions, the T-S law describes the behavior of materials at interfaces when cracks propagate. Delamination causes degradation of material stiffness which can be obtained from the traction-separation law. In addition, T-S law specifies the cohesive behavior of material in localized areas of damage where debonding occurred.

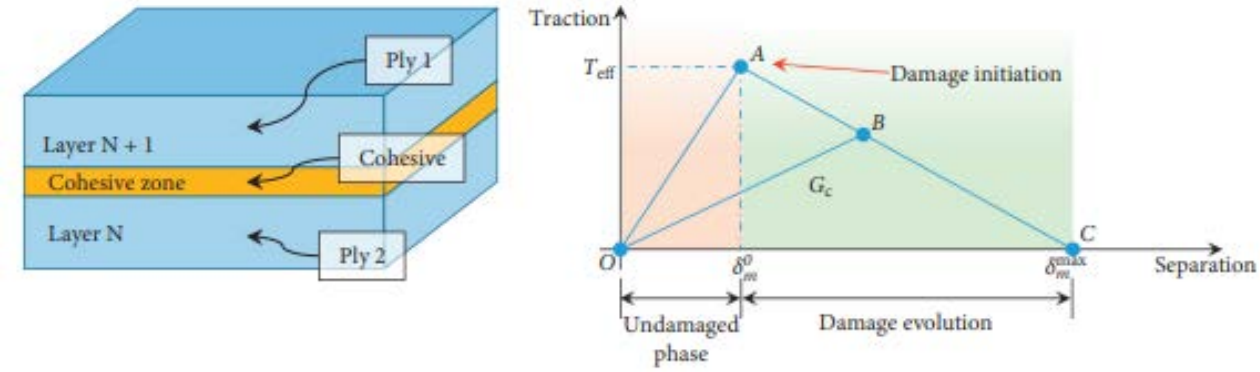

Figure 2.1: Representation of 3D cohesive element along with the triangular form of T-S law.[32]

The T-S law can be expressed by defining the critical tensile failure stress of the interface, the shape of the traction-separation law and the critical energy release rate. There are several shapes of T-S laws, including (a) Dugdale form, (b) Triangular form, (c) Lennard-Jones form, (d) polynomial form, (e) Linear form and so on, shown in Figure 2.2. T-S law is commonly defined in the triangular form, where normal traction and crack opening separation are used for fracture mode in order to define the relationship.[33] There are four modes of interlamination failure - (a) Mode-1 failure, (b) Mode-2 failure, (c) Mode-3 failure and (d) Mixed mode failure.

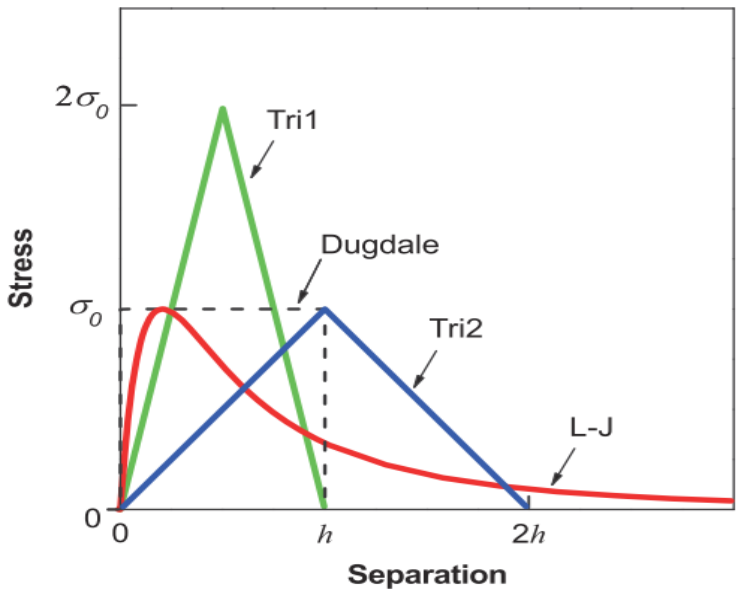

Figure 2.2: Stress-separation relationship for Dugdale form, Lennard-Jones form and two triangular form shape of the T-S law.[8]

Mode I failure is the most common type of failure where the crack propagates along the adhesive layer. Mode II failure is caused by tension in the adhesive layer and mode III failure is driven by shear failure in the adhesive layer. Mixed mode failure is caused by a combination of two or more of these modes.[34–36]

The cohesive T-S law of triangular form is expressed in equation (1).[33]

$$t_n = \frac{d}{1-d}\frac{\Delta_n}{\Delta_{nc}}\frac{\sigma_{max}}{d_{int}} \tag{1}$$

where, t_n is cohesive normal traction, Δ_n is the crack opening displacement, Δ_{nc} is the critical opening displacement, d is the damage parameter, d_{int} is the initial damage parameter, and σ_{max} is the tensile cohesive failure strength. The damage parameter can be define as-[37]

$$\mathrm{d} = \frac{\Delta_f(r^t - d_{int})}{r^t(\Delta_f - d_{int})} \tag{2}$$

where, Δ_f represent the displacement for fully damaged interface, r^t represent the time dependent damage threshold.

The total energy release rate of mode I failure is expressed as equation (3).[37]

$$G_{Ic} = \frac{1}{2} k_I \Delta_n \Delta_{nc} \tag{3}$$

where k_I represent the penalty stiffness for mode I.

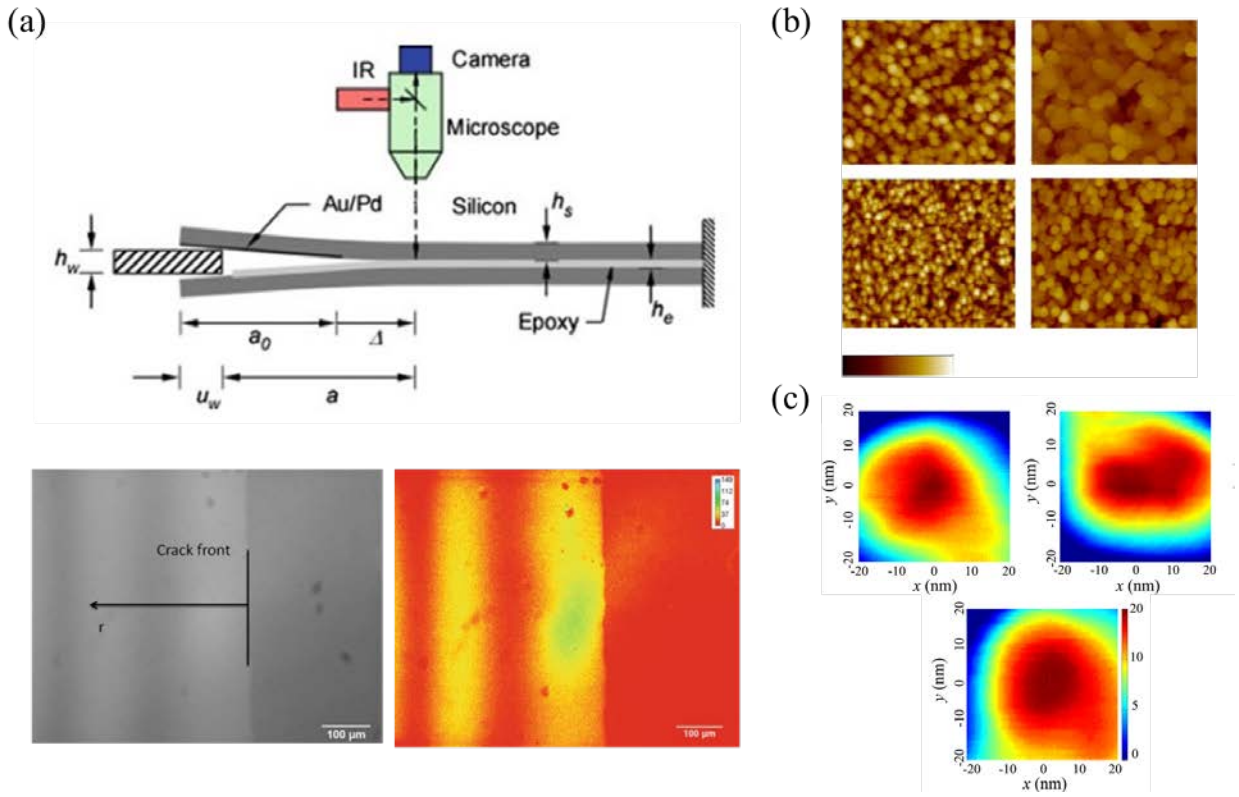

Figure 2.3: Experimental setup to determine the interface T-S relation. (a) Schematic of the setup of DCB test along with grey scale and color interferometric images. AFM tip images of (b) flat surface, (c) height maps.[6,8,40]

Multiple types of T-S relations have been proposed[8,38,39] based on mathematical approximations (such as Dugdale and triangular forms)[8,39,40] or physical observations (such as Lennard-Jones potential) to analyze the interfacial fracture characteristics.[41,42] The delamination behavior of interfaces has been analyzed using the cohesive zone method, constitutive models by simulating crack initiation and growth to emphasize the parameters.[43] Due to the nature of short adhesion range, it requires high resolution experiments and imaging, such as atomic force microscopy (AFM)[8,40,44,45] and optical interferometric measurements,[6,7] together with numerical and theoretical analysis to fully establish T-S relations, shown in Figure 2.3. For example, to

determine an intrinsic T-S relation of a single asperity contact, both AFM experimental pull-off forces and results from FEA and Maugis-Dugdale-*n* analytical model[9] are obtained. Then, iterative fittings using varying work of adhesion and range of adhesion are performed to match numerical and theoretical results with experimental data and finally establish the T-S relation.[8,40] Another example that to establish T-S relations in relating with the characterized crack tip, DCB tests with *in-situ* interferometric measurements (~20 nm resolution) of normal crack opening displacements along with J-integral analytical solution have been performed and compared with FEA approach.[6,7] In another instance, strength degradation and energy dissipation were evaluated based on the proposed new energy-based criteria to address the interdependence between damage and fracture criteria for mixed-mode loading.[43]

Recently, predictive machine learning (ML) is an emerging research area and a promising tool in prediction of the mechanical properties and design of materials.[46–48] The ML methods effectively learn on experimental or simulated data, and efficiently predict complicated data patterns or trends.[49] Supervised learning algorithms are a subset of ML methods, where models are trained with both input and output in training dataset. The underlying ML algorithms establish a pattern and predict targets based on input in testing dataset.[50] In interfacial science, recent works have advanced the prediction of interface fracture patterns, crack propagations, and interfacial thermodynamic constraints through supervised ML methods, [51–55] such as neural and deep material networks. While previous studies focus on specific types of materials, failure around known defects, or algorithm developments, still lacking are data-driven models that can establish intrinsic T-S relations and capture imperfections and can be generalized for different material interfaces.

In this study, a data-driven models are developed based on FEA, ML, and mechanical

experiments that quantify the T-S relations as well as identify locations of interfacial imperfections. The details of FEA simulation, ML model creation and procedure of the experimental validation are discussed in Chapter 4. The established eXtreme Gradient Boosting (XGBoost) ML models provided an efficient and convenient approach for characterizing interfacial properties in a wide range of material systems and discussed in Chapter 5. In Chapter 5, multi-output XGBoost models are demonstrated where XGBoost outperformed other ML models and predicted the microscale T-S relations with high precision through training on macroscale force-displacement (F-D) curves. Furthermore, experimental evaluations on several 3D printed material systems are performed and obtain agreeable F-D curves from experiments and FEA results using ML learned T-S relations. Additionally, to locate interfacial imperfections, a classification model is developed with a higher accuracy. To make impact out of this research work, a package has been developed based on the XGBoost regression model, which can facilitate the users to obtain T-S relations for their own F-D data without additional database establishment or model training.

CHAPTER 3

MACHINE LEARNING APPROACHES

ML model is a method which learns patterns and relationships in the data by utilizing the algorithm to generalize and predict the unforeseen data with accuracy. There are several types of ML approaches including supervised, unsupervised and reinforcement learning. Among them supervised learning approach are quite popular for labeled training data which consists of input features and corresponding target variables. The datasets are trained for learning the example and determining the underlying patterns and relationships between the features and variables. In addition, regression models are a subset of supervised learning algorithms designed for predicting continuous numerical values. In regression model, a mathematical relationship captures and learn the trend between the input features and the target variable.

3.1 XGBoost

In the field of mechanical engineering, the application of XGBoost has been quite efficient in predicting mechanical properties with the incorporation of generalized micromechanics. As an example, the XGBoost ML model conducted the parametric study to reduce the extensive experimentation in finding the optimal design variables and effect of mechanical properties.[56] XGBoost stands for Extreme Gradient Boosting, where it uses gradient boosting algorithm with regularization techniques to obtain robust predictive model. Gradient Boosting is an ensemble learning approach where multiple weak learners are trained sequentially to improve the prediction accuracy of the previous models. The main goal of the XGBoost models is to build an iterative ensemble of decision trees by minimizing objective functions, shown in Figure 3.1. The minimization of the objective function is a combination of the loss function and regularization.

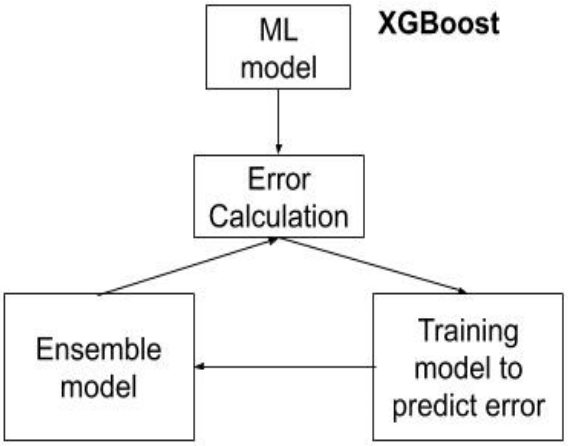

Figure 3.1: Workflow of XGBoost Model

The objective functions are shown in equation (4).[57]

$$Z = \sum_{i=1}^{n} l(o_i, \widetilde{o_i}) + \sum_{p=1}^{t} \Omega\left(f_p\right) \quad (4)$$

where, *n* is the total numbers of data in the *p*-th tree, *i* is the sample of dataset, and *t* is the total number of trees. $l(o_i, \widetilde{o_i})$ represent the individual loss for each sample in the training set for dependent variable o_i and its predicted value $\widetilde{o_i}$ shown in equation (5). $\Omega\left(f_p\right)$ represents the regularization term, which penalizes the complexity of the model and mitigate the tendency of the overfitting, where f_p represents the *t*-th tree in the ensemble, shown in equation (6).[57]

$$l(o_i, \widetilde{o_i}) = \sum_{i-1}^{n} (o_i - \widetilde{o_i})^2 \quad (5)$$

$$\Omega\left(f_p\right) = \gamma T + \frac{1}{2}\lambda \sum_{b-1}^{T} w_b^2 \quad (6)$$

where, T is the total number of leaf nodes in the decision tree, b represents each leaf in a node, w is the weight of the leaf node, γ maintain the complexity of the trees and λ are the regularization parameters.

The objective function iterates at *j*-th times with gradient function to create a boosted models for independent variable/features x_i, shown in equation (7).[57]

$$Z^j = \sum_{i=1}^{n} \left[l\left(o_i, \widetilde{o_i}^{j-1}\right) + \partial_{\widetilde{o_i}^{j-1}} l\left(o_i, \widetilde{o_i}^{j-1}\right) f_j(x_i) + \frac{1}{2} \partial^2_{\widetilde{o_i}^{j-1}} l\left(o_i, \widetilde{o_i}^{j-1}\right) f_j^2(x_i) \right] + \Omega\left(f_j\right) \quad (7)$$

Here, $\partial_{\widetilde{o_i}^{j-1}} l\left(o_i, \widetilde{o_i}^{j-1}\right) f_j(x_i)$ and $\partial^2_{\widetilde{o_i}^{j-1}} l\left(o_i, \widetilde{o_i}^{j-1}\right) f_j^2(x_i)$ represent the first and second order derivative of the loss function for corresponding features in each iteration.

3.2 Random Forest

Random forest has been implemented to predict composite mechanical properties for optimizing design space considering the framework of advanced material design.[58] For instances, random forest was employed to predict delamination locations in composite plates accurately and facilitate the design of advanced materials structures.[59] Random forest is an ensemble model where it integrates prediction of multiple decision tree to capture the relation between features and targeted variables. A random forest is a series of decision trees trained using bagging/ bootstrap aggregating in order to increase the overall accuracy of prediction. Furthermore, Random Forest incorporates randomness to feature selections while expanding the length of decision trees, since it searches for the best features in a random subset during the splitting of nodes. In each node, the algorithm finds the most optimal split to minimize the predicted error based on the loss function such as mean squared error or mean absolute error. Each decision tree expands using different bootstrap sample and random feature subset and all the decision tree prediction combine in the final ensemble.

The forest average prediction of the random forest model can be expressed as-[60]

$$f_n^M(x) = \frac{1}{M} \sum_{i=1}^{M} \frac{1}{NA_n(x)} \sum Y_i \quad (8)$$

Here, M represents the number of trees, $NA_n(x)$ the number of the estimation points of each tree, Y_i denotes the prediction of the tree.

3.3 Artificial Neural Network

The application of ML techniques, particularly artificial neural networks (ANN), in interface-related problems demonstrate the potential to accurately capture the complex nature of interfacial phenomena. To illustrate, ANN has been deployed for non-intrusive materials informatics methods in synthetic microstructure space to study interface impact.[61] Furthermore, the application of ANN has been used for predicting composite mechanical properties based on process parameters, for molecular dynamics simulation at atomic scale with traction-separation relation of grain boundary interfaces.[62,63] ANN is a ML algorithm that consists of interconnected neurons organized through layers, and the model is trained to adjust the weights and biases of these neurons to minimize prediction error. Typically, neural networks have three layers: an input layer, a hidden layer, and an output layer. Each layer is interconnected with adjacent layers through neurons moving in the forward direction. In forward propagation, inputs are fed into the network and pass through each layer where an activation function is applied to calculate the biases and weighted sum of the inputs. A feed forward process iterates for a defined period to reach the best prediction with a reduced amount of computation power and processing time. The neuron function model can be defined as –[64]

$$y = \emptyset(A(x)) \tag{9}$$

where $\emptyset$ is nonlinear function with interval [-1,1]. $A(x)$ represents the activation function of the neurons and simplification term shown in equation (10).[64]

$$A(x) = \theta + \sum_{i=1}^{N} w_i x_i \tag{10}$$

Here, x_i is the *i*-th input among the *N* neuron possessed, w_i is the *i*-th weight, and θ is the neuron threshold.

The effect of the weight in the loss function for feed-forward propagation can be

expressed using the partial derivative, shown in (11) equation.[64]

$$\frac{\partial L}{\partial w_1} = \frac{\partial L}{\partial \hat{y}} \frac{\partial \hat{y}}{\partial A_1} \frac{\partial A_1}{\partial w_1} \tag{11}$$

Here, L represent the loss function which can be calculated using the mean squared error. $\hat{y}$ represent the prediction.

3.4 Support Vector Regression

In the domain of interfaces, the interfacial properties such as surface energy and adhesion strength, can enable the design and optimization of tailored interfaces with desired characteristics for specific applications deploying ML techniques. Support vector regression (SVR) can predict composite mechanical properties, demonstrating the effectiveness of ML techniques in capturing complex interactions at material interfaces and accurately estimating material behavior. For example, SVR has been utilized to generalize the characterization of nanoscale adhesion using topological roughness parameters for different materials and interfaces.[55] SVR is a regression based ML model to predict accurately continuous numeric value and capture the complex pattern in the data. A key objective of SVR is to find an optimal hyperplane in a high-dimensional feature space by minimizing the loss function and specifying the margin tolerance. Furthermore, SVR can be applied to nonlinear relationships using kernels that transform the data into a higher dimensional space and capture the relationship. The objective function of the SVR expressed in equation (12).[65]

$$Z = C \sum_{i=1}^{N} L(y_i, \hat{y_i}) + \frac{1}{2} \|w\|^2 \tag{12}$$

Here, $\hat{y_i} = f(x_i) = w^T x_i + b$ and $C = {}^{1}/_{\lambda}$, is the regularization term. The margin constraint can be represented as equation (13).[65]

$$|y_i - w_i x_i| \leq \varepsilon + |\xi_i| \tag{13}$$

After adding the margin/ slack constraint the objective function can be rewritten in equation (14).[65]

$$Z = C \sum_{i=1}^{N} |\xi_i| + \frac{1}{2} \|w\|^2 \quad (14)$$

CHAPTER 4

MATERIALS AND METHOD

4.1 Database Collection via FEA

Double cantilever beam (DCB) specimens were established as 2D plane strain models in Abaqus® 2020. A layer of cohesive elements (70 mm × 0.1 mm, element type: COH2D4) was sandwiched between two solid beams (100 mm × 5 mm, element type: CPE4R). Mesh convergence check was done by carrying out simulations with 4 meshing sizes of about 1.5 times differences in sequence, namely 0.23, 0.15, 0.1, and 0.07 mm, shown in Figure 4.1. To mimic the DCB tests, a vertical displacement of 10 mm was applied on one beam at a distance of 25 mm away from the initial crack tip and a fixed boundary condition was applied at the corresponding location on another beam. Force was collected as a function of displacement from the simulations. To establish the database, triangular traction-separation relation was used with normal stiffness ranging from 20 kPa/mm to 2 GPa/mm, maximum strength from 0.2 kPa to 200 MPa, and energy dissipation from 2 mJ/m^2 to 60 kJ/m^2. The interfacial normal stiffness is selected uniformly by 30 intervals. For each value of interfacial stiffness, for example 100 MPa/mm, 10 uniformly distributed cases of interface strength from 1 to 10 MPa are selected. Then, for each interfacial stiffness and strength, 5 cases of energy dissipation are selected uniformly, specifically $G = {c\sigma^2}/{2E_n}$ where c is a prefactor varying from 2 to 6 uniformly. The upper and lower bounds of these parameters were extensively selected to allow the database to include large enough ranges for ML. In total, 1500 simulations were performed, and their corresponding T-S relations and F-D curves were discretized into 500 and 1000 features, respectively, for the purpose of ML.

For non-triangular traction-separation relations, including Lennard-Jones potential and

ML results from experiments, tabular function was utilized when defining the damage evolution for cohesive elements. For interfaces with imperfections, a MATLAB® code was generated to read Abaqus input files, randomly select and remove interfacial cohesive zone elements, and rewrite new input files by batch. Additional 780 cases with interfacial imperfection ratio 5% to 30% were produced. For each case, a F-D curve and the locations of imperfections were generated for ML.

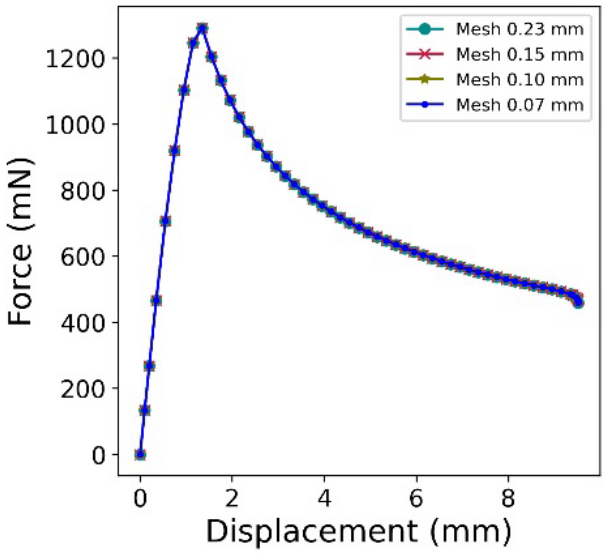

Figure 4.1: Convergence tests results of the four different sizes of FEA meshes (0.23 mm, 0.15 mm, 0.1 mm, and 0.07 mm).

4.2 Predictive ML Models

Common supervised ML regression models include SVR, RF, ANN, and boosted tree based models, such as XGBoost.[66] In this work, we particularly concentrated on the XGBoost algorithm as it showed competitive performance compared to other advanced ML algorithms.[67] Additionally, XGBoost algorithm delivers a good generalization and speed through building tree-based ensemble technique.[68] We established our multi-output regression models according to XGBoost[68] and scikit-learn[69] ML packages. The XGBoost model was an association of gradient decision trees with enhanced speed and performance.[68] It created a model where residuals were

calculated from the prior model and combined to forecast the final prediction. We split the datasets by 70/30 rule, where we trained the ML model using 70% data. The remaining 30% was the testing data to evaluate each model. Before training the models, we performed an extensive grid search using training data to determine optimal parameters. In grid search, we set maximum depth in a range between 6 - 18, learning rate 0.06-0.4, and subsample 0.2-0.6 with 4 fold-cross validation in order to obtain reliable results. We determined the optimal parameters for XGBoost as linear regression for the objective function, 16 for the maximum depth of the decision trees, 0 for γ, 0.5 for subsample to control overfitting, and 0.08 for the learning rate. We calculated median R^2 score for training data along with the test data. The median of R^2 score for the training and test data were 0.999 and 0.988, respectively. The difference between training and test data was about 1% and R^2 score on test was close to 1, which indicates the model was neither overfitting nor underfitting.[70,71] ML models were applied to 1500 samples, where each sample contained 1000 input features of F-D curves and 500 output features of T-S relation. To deal with multiple features, we tested both direct and chained multi-output regression methods. The detailed algorithm flows can be found in literatures.[72,73] In direct multi-output regression, each feature in the output was treated as an independent target variable for prediction. Each sample was divided into 500 (number of output features) separate problems. In chained multi-output regression, it creates a linear sequence of models for each sample, where predicted output features in previous steps were used together with input features to predict next feature.[69,74–76] Therefore, our multi-output models with 1500 samples and 500 output features each are equivalent to a data size of 1500×500 = 750,000 in single output models. We have applied both approaches and the direct multioutput regressor shows slightly better performance. Convergence tests were carried out on XGBoost model, using 10 sample sizes from 105 to 1050 samples and

resulting in a converged R^2 score at 0.988, shown in Figure 4.2(a).

In addition to XGBoost, we employed other three algorithms for comparison, namely SVR, ANN, and RF.[77–79] Optimal parameters for all these models were selected via grid search. In SVR, we used radial basis function (rbf) kernel with C=1000 as regularization parameter and ε=0.001 that associated with a training loss function. In RF algorithm, multiple decision trees were constructed to calculate the mean from all the trees. We found that 100 decision trees and maximum depth of 10 for those trees were optimal from a grid search. Finally, we used ANN that had capabilities of learning from sequential data and consisted of several hidden layers.[80] In this work, we applied 2 hidden layers with 200 neurons in each.

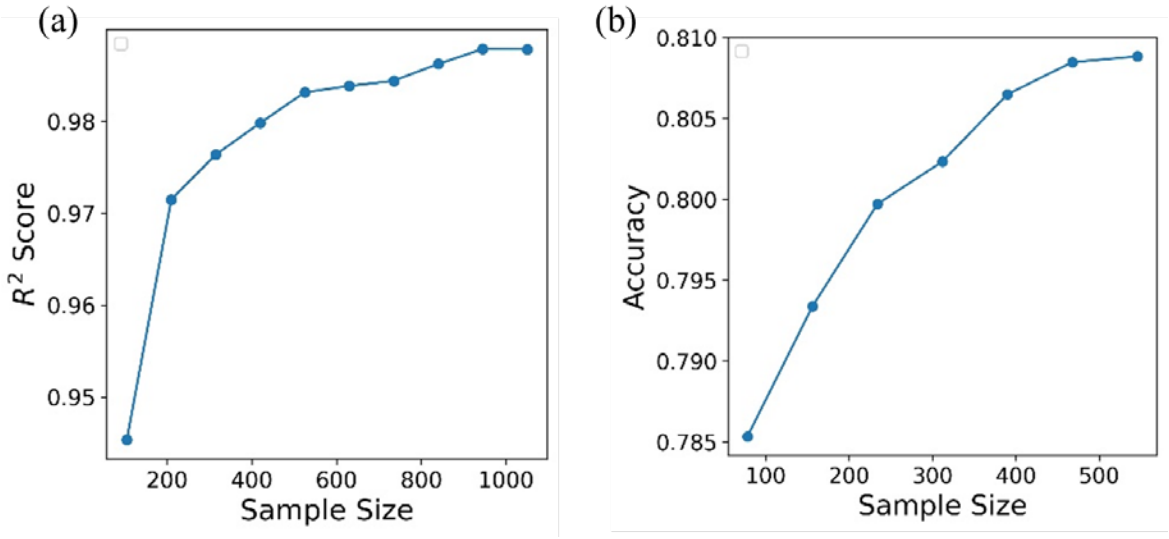

Figure 4.2: Convergence tests result for XGBoost ML model (a) R^2 scores in the T-S predictions, and (b) accuracy results in the imperfections predictions.

Regarding the interface imperfection task, we used XGBoost classifier for predicting locations of imperfection, where the label 1 indicated as perfectly bonded interface and 0 indicated as an imperfection (void). Similar to the regression models, we used 780 samples and 350 output features each to establish our multi-output classification models. The convergence tests of 7 sample sizes from 78 to 546 samples were performed and accuracy converged to 0.808,

shown in Figure 4.2(b). To work with the classification task, we set logistic regression as the objective function and applied another grid search. The grid search set a range of 200-350 for the number of trees, 6-10 for the maximum depth, 0.005-0.1 for the learning rate, and 0.5-7 for L2 regularization, respectively. As a result, we used linear regression with 300 for optimal number of trees, 6 for maximum depth, 0.008 for the learning rate, and 5 for L2 regularization. In addition, three other ML models, specifically SVC, ANN classification, and RF classification, were tested. We utilized a grid search mechanism as before to obtain the optimal parameters. For RF classifier, we used 100 as number of trees in the forest, 6 for maximum depth of trees, 5 for minimum number of samples required to split an internal node, 2 for minimum number of samples to be required at a leaf node. In ANN classification, we applied one hidden layer with 100 neurons, 'tanh' activation function for the hidden layer, α= 0.0001 for L2 regularization, and a stochastic gradient descent (sgd) solver for weight optimization. For SVC model, we used the rbf kernel, γ= 1 for kernel coefficient, and C= 0.1 for regularization parameter.[69,77–79]

4.3 Double Cantilever Beam (DCB) Experiments

DCB specimens of different material systems were prepared by 3D printing. The material systems included epoxy-epoxy, epoxy-carbon fiber (CF) composites, and polymer-silicone interfaces. We 3D printed epoxy and its composites into DCB specimens using direct ink writing (DIW) method. The inks were prepared based on our previously reported literature.[81] Briefly, epoxide resin (Epon 826®) was mixed with cross-linkers (Epikure W®, 25 wt%, and Epikure 3140®, 5 wt%), and silica nanoparticles (EH-5®, 13 wt%) using a planetary mixer. Carbon fibers (10 wt%) were added to the mixture to prepare the epoxy-CF composite ink. The as-prepared inks were 3D printed by DIW technique using a Hyrel 3D SR Engine® printer with a nozzle size of 0.3 mm, layer height of 0.15 mm, and a printing speed of 10 mm/s. After printing, the

materials were thermally cured at 150 ºC for 10 hrs. For polymer-silicone samples, we used stereolithography (SLA) to 3D print polymer beams (Formlab® Tough resin), which were then washed by isopropyl alcohol (IPA) and cured at 60 ºC for 60 min. Then, a layer of silicone glue was applied at interface and cured at room temperature for at least 2 hrs. The polymer-silicone samples had a beam thickness of 5 mm, a beam width of 20 mm, a length of 100 mm and an initial crack length of 25 mm (ASTM D5528 standard). A scaling factor of 0.8 was used for epoxy and epoxy-CF samples to reduce the 3D printing time. Different material properties were considered in FEA validations, where E_{pol}=1.5 GPa for SLA printed polymer beams, E_{epoxy}=1.6 GPa and $E_{epoxy-CF}$=2.2 GPa for DIW printed epoxy and epoxy-CF samples, respectively. DCB tests were performed via a universal tensile machine under displacement control (0.1 mm/s) up to 10 mm. F-D curves are collected at 10 pts/s.

CHAPTER 5

RESULTS AND DISCUSSIONS

An overview of our workflow is shown in Figure 5.1, including three key components: (i) establishing F-D and T-S relations database by FEA; (ii) training ML models; and (iii) validating ML models by experimental data.

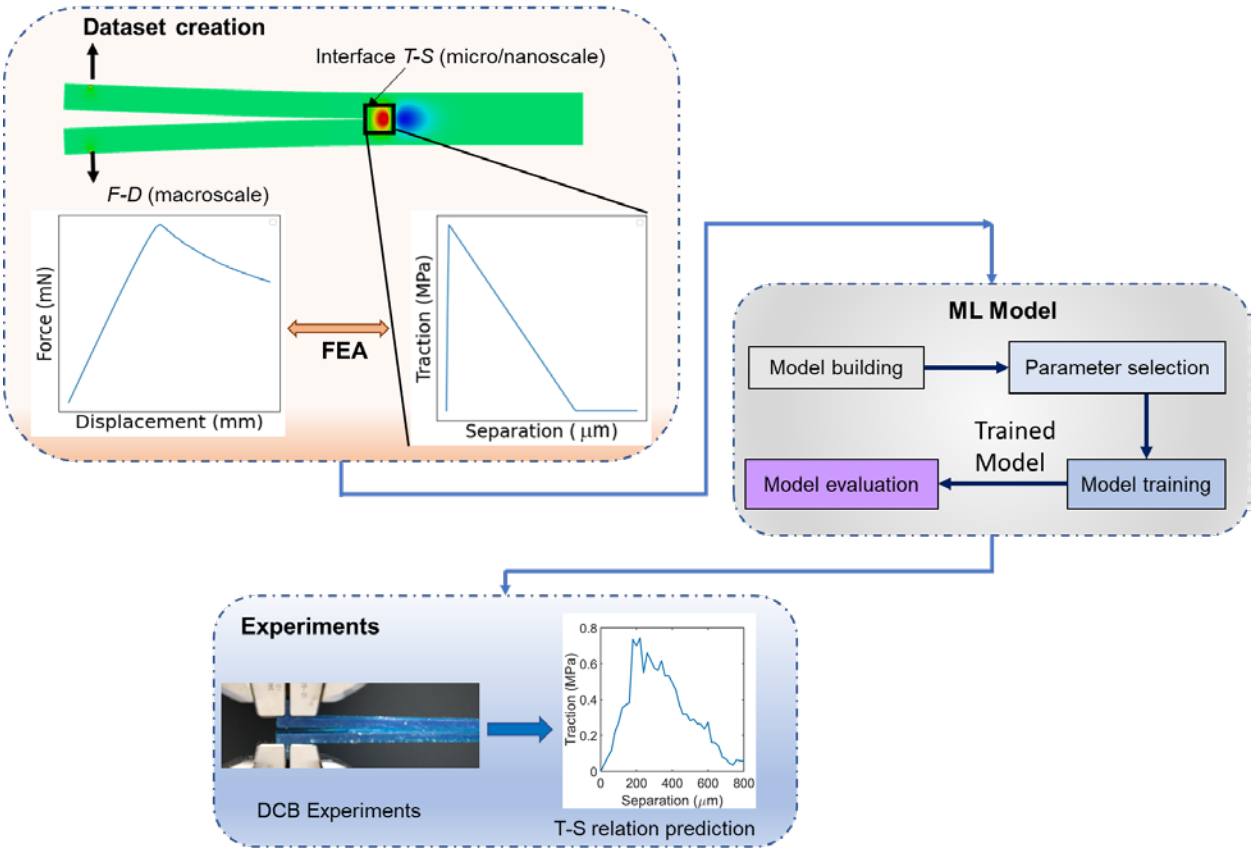

Figure 5.1: An overview schematic of the development of a data-driven method for establishing interfacial properties. The method includes FEA database generation, multi-output ML models training, and experimental validations.

Firstly, we perform FEA simulating the DCB tests to collect a series of F-D curves correlating to T-S relations with interfacial stiffness and strength spanning 5 and 6 orders of magnitude, respectively, shown in top section of Figure 5.1. Secondly, we use several ML methods including SVR, RF, ANN, and XGBoost to train the models based on FEA database. Each ML model performance evaluate to find the optimal model where XGBoost ML models

outperform the other models. Scalable relations from F-D curves to T-S relations are established in trained ML models, shown in middle section of Figure 5.1. Finally, we perform experimental validations that harness the trained ML models on experimental F-D curves to establish T-S relations at the interfaces of multiple material systems, shown in bottom section in Figure 5.1.

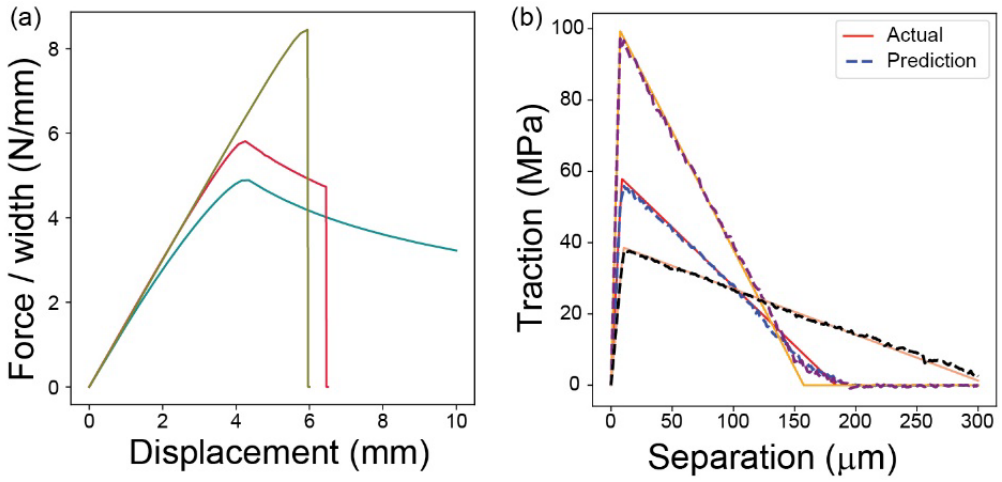

Figure 5.2: Machine learning on FEA data. (a) Different types of force-displacement curves, (b) comparison between triangular traction-separation relations predicted by XGBoost ML and the actual cases preset in FEA.

5.1 ML Results on T-S Relations

Different interfacial failure behaviors are observed from F-D curves. Several representative F-D curves are shown in Figure 5.2(a), including gradual crack propagation, propagation followed by catastrophic failure, and immediate failure after ultimate strength. It is challenging to correlate F-D curves and T-S relations directly as the measurements of F-D curves (macroscale) and T-S relations (microscale) are at different length scales[51] and there is a curve to curve relation between them. However, through the training of the multi-output XGBoost model, we are able to accurately predict the interfacial T-S relations. XGBoost is a scalable version of gradient boosting framework[82] that builds a sequential ensemble method to reduce the error of predecessor trees by updating the residual errors.[67,68] With the integration of decision tree

method, XGBoost can enhance the accuracy, speed, and performance of a model. For testing dataset, the median of R^2 score is 0.988, manifesting the near perfect matching case where $R^2=1$. Comparisons of example T-S relations between FEA preset values and the predicted results by the XGBoost model are exhibited in Figure 5.2(b). The solid lines indicate the FEA preset values, whereas dashed lines represent predicted T-S relations using XGBoost model. The ML model results in accurate prediction for the triangular form and key parameters, including interfacial stiffness, strength, and range of adhesion. It is worth noting that the predicted T-S relations are output as discrete traction data as a function of separation (Methods section) and a triangular form is not prescribed in our ML model.

More interestingly, we investigate our model performance on a different form of T-S relation, namely 3-9 Lennard-Jones (L-J) potential,[42] which is not included in the training process and the L-J potential form is not given during ML prediction. For conventional fittings or inverse properties extractions, predefined forms for the target are generally required. This is a constraint that one needs to make assumptions or use trial and error to determine a proper function form for fitting. In contrast, our ML model can be used directly for prediction instead of predetermining the form for the targeted T-S relation. The ML model is trained based on purely triangular T-S relations, in which the L-J potential form is not given during the training process. Three examples of F-D curves generated via different L-J potentials are shown in Figure 5.3(a) and their corresponding predicted results are in Figure 5.3(b). Although variations are observed at the softening process around separation of 100-250 μm, the predictions capture key features of L-J potential, including the gradual transition near maximum traction and long tails of the T-S relation (Figure 5.3(b)). This indicates the advantage of data-driven ML that the data features can be captured without making assumptions of targeted fitting forms, as well as a capability to

predict different forms of T-S relations out of the training data domain.

To generalize our ML model to different materials and dimensions of DCBs, we introduce the normalized force and displacement that $\bar{F} = Fa^2/(Eh^3t)$ and $\bar{d} = d/a$, where E is the Young's modulus of the beams, a is the initial crack length, h is beam thickness, and t is beam width. Since the FEA database for ML model training is established based on the linear elastic deformation assumption and the maximum normalized force of $0.001<\bar{F}_{max}<0.04$, our ML model would be applicable to predict the T-S relations within these limitations.

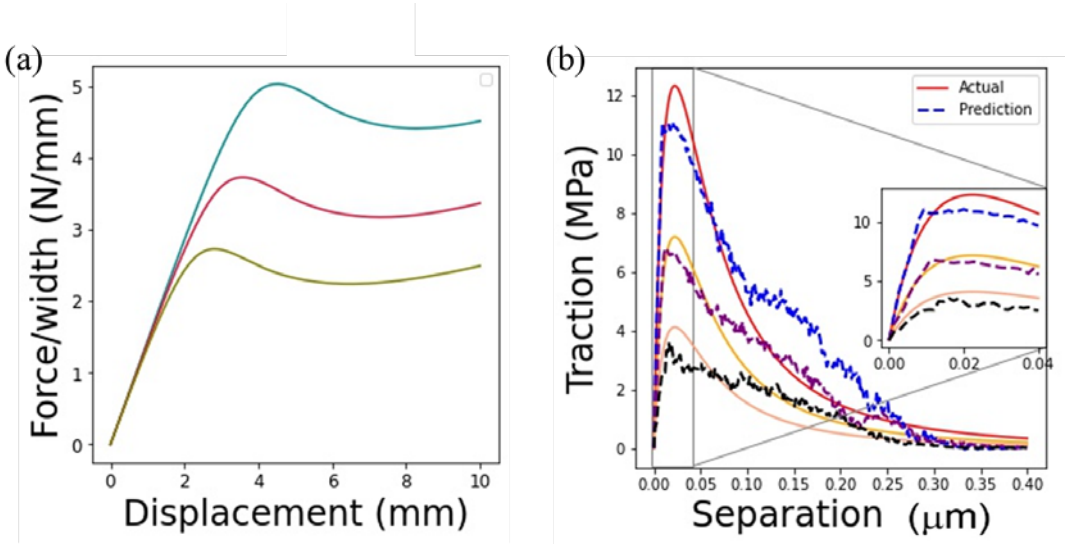

Figure 5.3: (a) F–D curves calculated based on different Lennard-Jones T–S relations and (b) ML prediction results for Lennard-Jones traction-separation relations.

5.2 Experimental Validation

To validate the T-S relation established from the XGBoost ML approach, we perform DCB experiments for three material systems, namely epoxy, epoxy-CF composite, and polymer-silicone interfaces. These experiments yield the F-D curves plotted in solid lines in Figure 5.4(a). The F-D curve of polymer-silicon interface has a compliant stiffness and a gradual softening process after maximum force at around 1 mm, indicating a gradual debonding at interface along the crack. Both pure epoxy and epoxy-CF composite have stiff and brittle interface. The drops in

F-D curves indicate catastrophic failures of interfaces in both materials. These three material systems are selected to validate the application of our ML model on both hard and soft interfaces, as well as single component and composite materials.

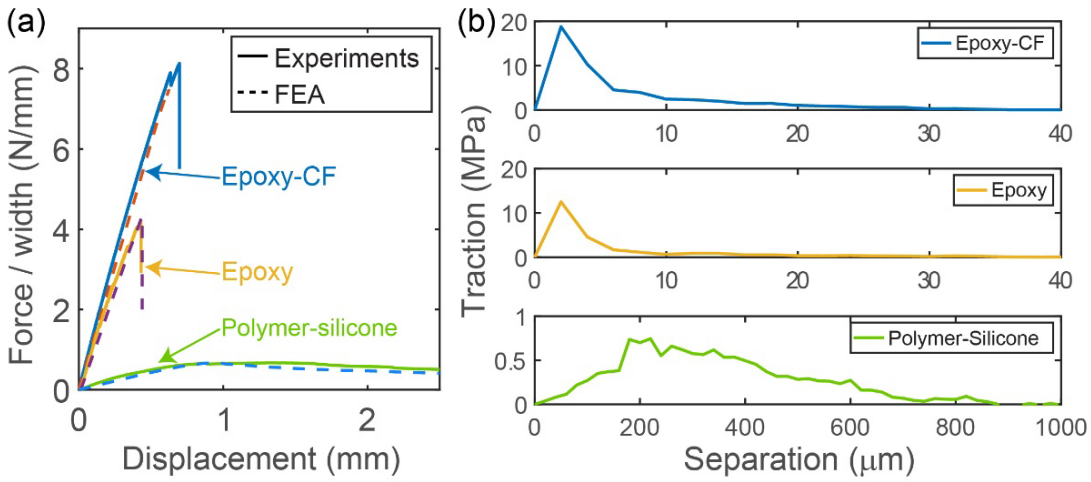

Figure 5.4: DCB experimental data for ML model validation. (a) Experimental F-D curves and FEA results using (b) T-S relations trained from experimental F-D curves for epoxy, epoxy-CF, and polymer-silicone interfaces.

Harnessing our trained ML model, T-S relation for each material system is generated individually based on the experimental F-D curves (Figure 5.4(b)). Stiff and short range interface interaction is seen for epoxy and epoxy-CF, with strength of the former (12.5 MPa) is 66% of the later (18.8 MPa). Polymer-silicone interface shows a compliant and long range T-S relation, which is expected because the silicone at interface can be largely deformed before breaking. Subsequently, the T-S relations are used in FEA simulations to obtain F-D curves (dashed lines in Figure 5.4(a)). We observe that the F-D curves from simulations are in agreement with experimental F-D curves in all three cases. The comparisons demonstrate that the accurate T-S relations are predicted from experimental data using our ML model.

5.3 Model Performance

We investigate other ML models that are SVR, ANN, and RF to compare with the

XGBoost model. The performance of models is evaluated on the basis of coefficient of determination (R^2 score) and root mean square error (RMSE), which is normalized by maximum interfacial traction in each test case. As illustrated in Figure 5.5(a) and (b), XGBoost outperforms other models with the highest R^2 score (R^2=0.988) and lowest normalized RMSE. Though comparable performance is seen in RF, XGBoost is the most efficient and takes only half of training time than others.

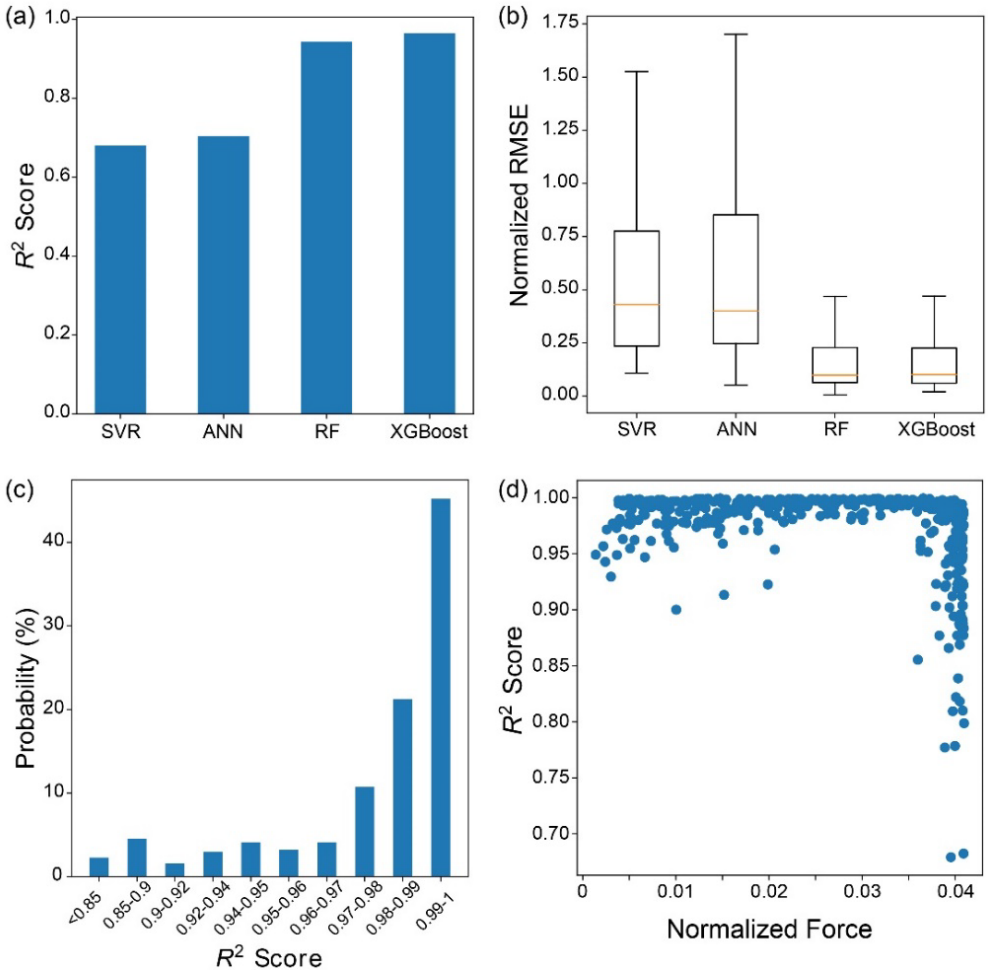

Figure 5.5: Performance of different machine learning methods. (a) R^2 score and (b) normalized RMSE for four different methods. XGBoost shows the highest R^2 score and lowest RMSE. (c) Distribution of R^2 score for the XGBoost model and (d) relation between R^2 and maximum normalized forces.

To further evaluate the performance in XGBoost model, we show the distribution of R^2

scores for all testing samples in Figure 5.5(c). R^2 scores are concentrated at high values (near to 1) and more than 40% predicted samples lie in the 0.99-1 range. More than 84% samples have a R^2 score more than 0.95. To identify the low performance cases, we exhibit individual R^2 score for each test case as a function of maximum normalized force in Figure 5.5(d). The normalized force is $\bar{F} = Fa^2/(Eh^3t)$, where a=25 mm is the distance of loading location to initial crack tip, E=2 GPa is the Young's modulus, h= 5 mm is the thickness, and t=1 mm is the width in our DCB models for simulations. When the test data has a maximum force approaching the limits of force range in our training database ($0.001<\bar{F}<0.04$), the R^2 score decreases significantly. For lower bound, negative R^2 scores are observed (not shown in Figure), indicating the prediction is even worse than a simple average of data. For upper bound, R^2 scores distribute from 1 to less than 0.7. Between these two bounds, the ML predictions have R^2 between 0.95 and 1 with only a few exceptional cases.

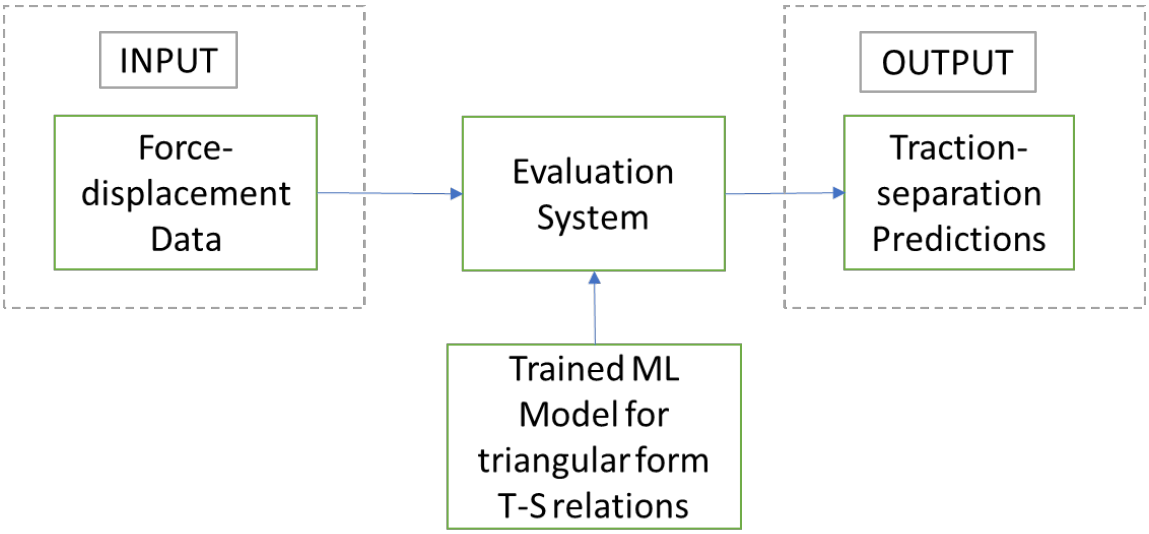

Figure 5.6: A schematic view of evaluation system architecture.

Based on our trained XGBoost model, we develop a code package that allows users to predict T-S relations by simply inputting F-D curves from DCB experiments or simulations. The schematic of the evaluation process is shown in Figure 5.6. This package, including trained

model file, python source codes, example data, and user guideline, is available to be downloaded.[83] This program can establish T-S relations without setting up a new database or training ML models again.

5.4 Predictions of Interface Imperfections

The imperfections can affect the stress distribution at the interface remarkably and thus leads to different crack propagation and failure behaviors.[84,85] In addition to characterizing T-S relation at perfect interfaces, we harness the FEA-ML method on identifying interface imperfections. As one example shown in Figure 5.7(a), interfacial voids are generated in our FEA models as imperfections. Local stress concentration depends on imperfection distribution, so does the macroscale F-D curves. Both the ratio and the locations of these imperfections result in different maximum forces, loading slopes, and fluctuations in F-D curves (Figure 5.5(b), (c)). Such force fluctuation may simply be smoothed or considered as noise, especially in experiments, without relating such behavior to interface imperfections, not mention to precisely identify imperfection locations.

To identify the locations of interfacial imperfections, the interface (70 mm in length) is discretized into 0.2 mm intervals and analyzed as target features. Then, we use four different classification methods to establish ML models trained on F-D curves and imperfection locations correlated data. The four methods, including support vector classification (SVC), ANN classification, RF classification, and XGBoost classification, result in close prediction accuracy between 81-82% (see Methods section for details). The XGBoost classification model results in an average accuracy of 80.9±8.6% predicting bonded and imperfect locations along the interface for 312 random cases. Five cases with prediction accuracy from 67-92% are shown as examples in Figure 5.5(d), indicating correct and wrong predictions along interface and their locations.

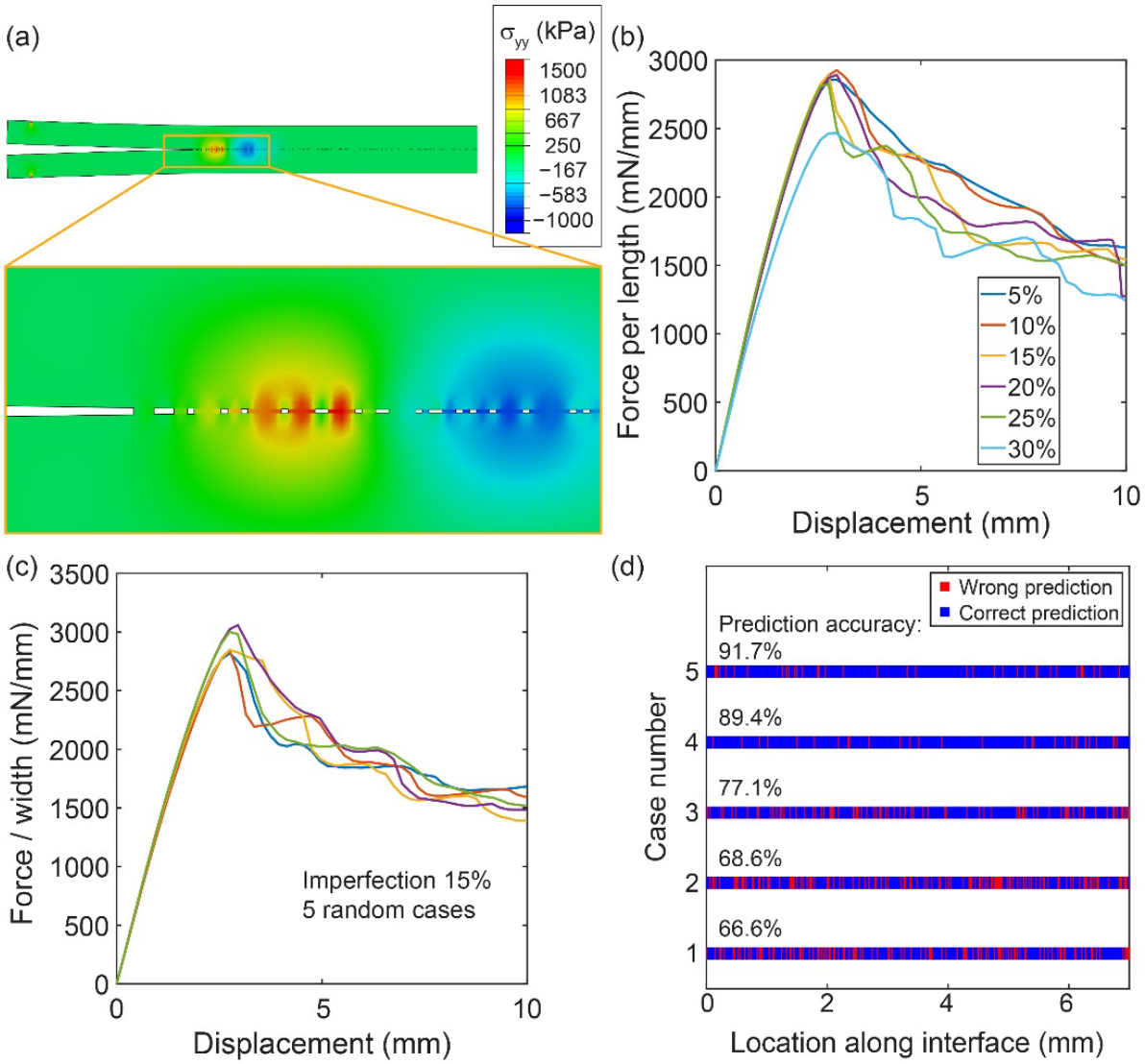

Figure 5.7: Prediction of interfacial imperfection by machine learning. (a) An example of stress distribution at interface with imperfections in a FEA DCB test. Force-displacement curves obtained from cases with (b) different imperfection ratio and (c) different locations of imperfections. (d) Examples of imperfection locations predictions using XGBoost classifier. The average prediction accuracy is 80.9% over 312 testing cases.

CHAPTER 6

CONCLUSION AND FUTURE WORK

6.1 Conclusion

In conclusion, we have demonstrated the process of characterizing interface adhesive properties and imperfections using a combination of FEA and ML. We use advanced ML models framing with multi-output regressors to train on a FEA database, simulating DCB tests with interfacial stiffness and strength over 5 and 6 orders of magnitude, respectively. Our optimized XGBoost model predicts T-S relations with a coefficient of determination of 0.988. 3D printed DCB specimens of epoxy, epoxy-CF, and silicone-polymer interfaces are used to experimentally measure F-D curves, extract T-S relations by our ML model, and then validated by the F-D curves that are consistent in both experiments and FEA. Furthermore, we leverage this FEA-ML approach to illustrate the capability of detecting imperfection locations along interface with an accuracy of 80.8±8.6%. Finally, we develop a Python code package including our trained ML model for user customizable applications for establishing T-S relations for other material systems.

6.2 Future Work

Further research into the real life application of interface properties can be explored for different failure criteria. Incorporation of different materials properties, microstructural arrangements, hybrid composites interfacial characteristics have the potential for further studies. Integrating constitutive continuum mechanics into ML models can also contribute to a more comprehensive understanding of interfacial delamination.

REFERENCES

1. Swain, R. E., Reifsnider, K. L., Jayaraman, K. & El-Zein, M. Interface/Interphase Concepts in Composite Material Systems. *J. Thermoplast. Compos. Mater.* **3**, 13–23 (1990).

2. Gan, Y. X. Effect of Interface Structure on Mechanical Properties of Advanced Composite Materials. *Int. J. Mol. Sci.* **10**, 5115–5134 (2009).

3. Mukhopadhyay, T., Mahata, A., Adhikari, S. & Zaeem, M. A. Effective mechanical properties of multilayer nano-heterostructures. *Sci. Rep.* **7**, 1–13 (2017).

4. Mamedov, A. A. *et al.* Molecular design of strong single-wall carbon nanotube/polyelectrolyte multilayer composites. *Nat. Mater.* **1**, 190–194 (2002).

5. Harper, P. W. & Hallett, S. R. Cohesive zone length in numerical simulations of composite delamination. *Eng. Fract. Mech.* **75**, 4774–4792 (2008).

6. Gowrishankar, S., Mei, H., Liechti, K. M. & Huang, R. A comparison of direct and iterative methods for determining traction-separation relations. *Int. J. Fract.* **177**, 109–128 (2012).

7. Wu, C., Gowrishankar, S., Huang, R. & Liechti, K. M. On determining mixed-mode traction–separation relations for interfaces. *Int. J. Fract.* **202**, 1–19 (2016).

8. Grierson, D. S., Liu, J., Carpick, R. W. & Turner, K. T. Adhesion of nanoscale asperities with power-law profiles. *J. Mech. Phys. Solids* **61**, 597–610 (2013).

9. Zheng, Z. & Yu, J. Using the Dugdale approximation to match a specific interaction in the adhesive contact of elastic objects. *J. Colloid Interface Sci.* **310**, 27–34 (2007).

10. Xiong, J., Shi, S.-Q. & Zhang, T.-Y. Machine learning of phases and mechanical properties in complex concentrated alloys. *J. Mater. Sci. Technol.* **87**, 133–142 (2021).

11. Liu, X., Athanasiou, C. E., Padture, N. P., Sheldon, B. W. & Gao, H. A machine learning approach to fracture mechanics problems. *Acta Mater.* **190**, 105–112 (2020).

12. Daghigh, V. *et al.* Machine learning predictions on fracture toughness of multiscale bio-nano-composites. *J. Reinf. Plast. Compos.* **39**, 587–598 (2020).

13. Mahmoodzadeh, A. *et al.* Prediction of Mode-I rock fracture toughness using support vector regression with metaheuristic optimization algorithms. *Eng. Fract. Mech.* **264**, 108334 (2022).

14. Yin, B. B. & Liew, K. M. Machine learning and materials informatics approaches for evaluating the interfacial properties of fiber-reinforced composites. *Compos. Struct.* **273**, 114328 (2021).

15. Mortezaei, M., Famili, M. H. N. & Kokabi, M. The role of interfacial interactions on the glass-transition and viscoelastic properties of silica/polystyrene nanocomposite. *Compos. Sci. Technol.* **71**, 1039–1045 (2011).

16. Xie, L., Cui, X., Gong, L., Chen, J. & Zeng, H. Recent Advances in the Quantification and Modulation of Hydrophobic Interactions for Interfacial Applications. *Langmuir* **36**, 2985–3003 (2020).

17. Munch, E. *et al.* Tough, bio-inspired hybrid materials. *Science.* **322**, 1516–1520 (2008).

18. Drzal, L. T. The role of the fiber-matrix interphase on composite properties. *Vacuum* **41**, 1615–1618 (1990).

19. Totry, E., Molina-Aldareguía, J. M., González, C. & LLorca, J. Effect of fiber, matrix and interface properties on the in-plane shear deformation of carbon-fiber reinforced composites. *Compos. Sci. Technol.* **70**, 970–980 (2010).

20. Whitney, J. M., Browning, C. E. & Hoogsteden, W. A Double Cantilever Beam Test for Characterizing Mode I Delamination of Composite Materials. *J. Reinf. Plast. Compos.* **1**, 297–313 (1982).

21. Manoharan, M. P. *et al.* The interfacial strength of carbon nanofiber epoxy composite using single fiber pullout experiments. *Nanotechnology* **20**, 295701 (2009).

22. Zhou, X. F., Wagner, H. D. & Nutt, S. R. Interfacial properties of polymer composites measured by push-out and fragmentation tests. *Compos. - Part A Appl. Sci. Manuf.* **32**, 1543–1551 (2001).

23. Kumar, M., Kumar, P. & Bhadauria, S. S. Numerical simulation of delamination growth in fiber reinforced polymer laminates using cohesive zone modeling. *Mech. Adv. Mater. Struct.* 1–17 (2020).

24. Whitney, J. M. Stress analysis of the double cantilever beam specimen. *Compos. Sci. Technol.* **23**, 201–219 (1985).

25. Blackman, B. R. K., Hadavinia, H., Kinloch, A. J., Paraschi, M. & Williams, J. G. The calculation of adhesive fracture energies in mode I: revisiting the tapered double cantilever beam (TDCB) test. *Eng. Fract. Mech.* **70**, 233–248 (2003).

26. Bennati, S., Colleluori, M., Corigliano, D. & Valvo, P. S. An enhanced beam-theory model of the asymmetric double cantilever beam (ADCB) test for composite laminates. *Compos. Sci. Technol.* **69**, 1735–1745 (2009).

27. Tsuda, T., Ogasawara, T., Deng, F. & Takeda, N. Direct measurements of interfacial shear strength of multi-walled carbon nanotube/PEEK composite using a nano-pullout method. *Compos. Sci. Technol.* **71**, 1295–1300 (2011).

28. Narducci, F. & Pinho, S. T. Exploiting nacre-inspired crack deflection mechanisms in CFRP via micro-structural design. *Compos. Sci. Technol.* **153**, 178–189 (2017).

29. Kim, S. *et al.* Soft nanocomposite electroadhesives for digital micro- And nanotransfer printing. *Sci. Adv.* **5**, eaax4790 (2019).

30. Minsky, H. K. & Turner, K. T. Composite microposts with high dry adhesion strength. *ACS Appl. Mater. Interfaces* **9**, 18322–18327 (2017).

31. Selezneva, M. *et al.* The brittle-to-ductile transition in tensile and impact behavior of hybrid carbon fibre/self-reinforced polypropylene composites. *Compos. Part A Appl. Sci. Manuf.* **109**, 20–30 (2018).

32. Di Caprio, F., Saputo, S. & Sellitto, A. Numerical-Experimental Correlation of Interlaminar Damage Growth in Composite Structures: Setting Cohesive Zone Model Parameters. *Adv. Mater. Sci. Eng.* (2019) doi:10.1155/2019/2150921.

33. Abdel Wahab, M. M. 11 - Simulating mode I fatigue crack propagation in adhesively-bonded composite joints. in *Fatigue and Fracture of Adhesively-Bonded Composite Joints* (ed. Vassilopoulos, A. P.) 323–344 (Woodhead Publishing, 2015). doi:https://doi.org/10.1016/B978-0-85709-806-1.00011-2.

34. Zhu, Y., Liechti, K. M. & Ravi-Chandar, K. Direct extraction of rate-dependent traction–separation laws for polyurea/steel interfaces. *Int. J. Solids Struct.* **46**, 31–51 (2009).

35. Tamrakar, S., Ganesh, R., Sockalingam, S. & Gillespie, J. W. Rate dependent mode II traction separation law for S-2 glass/epoxy interface using a microdroplet test method. *Compos. Part A Appl. Sci. Manuf.* **124**, 105487 (2019).

36. Sockalingam, S., Dey, M., Gillespie, J. W. & Keefe, M. Finite element analysis of the microdroplet test method using cohesive zone model of the fiber/matrix interface. *Compos. Part A Appl. Sci. Manuf.* **56**, 239–247 (2014).

37. Álvarez, D., Blackman, B. R. K., Guild, F. J. & Kinloch, A. J. Mode I fracture in adhesively-bonded joints: A mesh-size independent modelling approach using cohesive elements. *Eng. Fract. Mech.* **115**, 73–95 (2014).

38. Heidari-Rarani, M. & Ghasemi, A. R. Appropriate shape of cohesive zone model for delamination propagation in ENF specimens with R-curve effects. *Theor. Appl. Fract. Mech.* **90**, 174–181 (2017).

39. Maugis, D. Adhesion of spheres: The JKR-DMT transition using a dugdale model. *J. Colloid Interface Sci.* **150**, 243–269 (1992).

40. Jiang, Y. & Turner, K. T. Measurement of the strength and range of adhesion using atomic force microscopy. *Extrem. Mech. Lett.* **9**, 119–126 (2016).

41. Jacobs, T. D. B., Wabiszewski, G. E., Goodman, A. J. & Carpick, R. W. Characterizing nanoscale scanning probes using electron microscopy: A novel fixture and a practical guide. *Rev. Sci. Instrum.* **87**, 013703 (2016).

42. Yu, N. & Polycarpou, A. A. Adhesive contact based on the Lennard-Jones potential: A correction to the value of the equilibrium distance as used in the potential. *J. Colloid Interface Sci.* **278**, 428–435 (2004).

43. de Oliveira, L. A. & Donadon, M. V. Delamination analysis using cohesive zone model: A discussion on traction-separation law and mixed-mode criteria. *Eng. Fract. Mech.* **228**, 106922 (2020).

44. Mello, A. W. & Liechti, K. M. The effect of self-assembled monolayers on interfacial fracture. *J. Appl. Mech. ASME* **73**, 860–870 (2006).

45. Safari, A. *et al.* Interfacial separation of a mature biofilm from a glass surface - A combined experimental and cohesive zone modelling approach. *J. Mech. Behav. Biomed. Mater.* **54**, 205–218 (2016).

46. Schmidt, J., Marques, M. R. G., Botti, S. & Marques, M. A. L. Recent advances and applications of machine learning in solid-state materials science. *npj Comput. Mater.* **5**, 1–36 (2019).

47. Butler, K. T., Davies, D. W., Cartwright, H., Isayev, O. & Walsh, A. Machine learning for molecular and materials science. *Nature* **559**, 547–555 (2018).

48. Bartók, A. P. *et al.* Machine learning unifies the modeling of materials and molecules. *Sci. Adv.* **3**, e1701816 (2017).

49. Rovinelli, A., Sangid, M. D., Proudhon, H. & Ludwig, W. Using machine learning and a data-driven approach to identify the small fatigue crack driving force in polycrystalline materials. *npj Comput. Mater.* **4**, 1–10 (2018).

50. Nasteski, V. An overview of the supervised machine learning methods. *Horizons. b.* **4**, 51–62 (2017).

51. Fernández, M., Rezaei, S., Rezaei Mianroodi, J., Fritzen, F. & Reese, S. Application of artificial neural networks for the prediction of interface mechanics: a study on grain boundary constitutive behavior. *Adv. Model. Simul. Eng. Sci.* **7**, 1–27 (2020).

52. Mudunuru, M. K., Panda, N., Karra, S. & Srinivasan, G. Surrogate models for estimating failure in brittle and quasi-brittle materials. *Appl. Sci.* **9**, 2706 (2019).

53. Liu, Z. Deep material network with cohesive layers: Multi-stage training and interfacial failure analysis. *Comput. Methods Appl. Mech. Eng.* **363**, 112913 (2020).

54. Zhang, J., Wei, C. & Wu, C. Thermodynamic consistent neural networks for learning material interfacial mechanics. in *NeurIPS 2020 Workshop on Interpretable Inductive Biases and Physically Structured Learning* (2020).

55. Khatri, N., Ji, X., Minsky, H. & Jiang, Y. Understanding Nanoscale Topology–Adhesion Relationships Via Support Vector Regression. *Adv. Mater. Interfaces* **8**, 2100175 (2021).

56. Guo, P., Meng, W., Xu, M., Li, V. C. & Bao, Y. Predicting Mechanical Properties of High-Performance Fiber-Reinforced Cementitious Composites by Integrating Micromechanics and Machine Learning. *Materials (Basel).* **14**, 3143 (2021).

57. Kiangala, S. K. & Wang, Z. An effective adaptive customization framework for small manufacturing plants using extreme gradient boosting-XGBoost and random forest ensemble learning algorithms in an Industry 4.0 environment. *Mach. Learn. with Appl.* **4**, 100024 (2021).

58. Shah, V., Zadourian, S., Yang, C., Zhang, Z. & Gu, G. X. Data-driven approach for the prediction of mechanical properties of carbon fiber reinforced composites. *Mater. Adv.* **3**, 7319–7327 (2022).

59. Lim, D. K., Mustapha, K. B. & Pagwiwoko, C. P. Delamination detection in composite plates using random forests. *Compos. Struct.* **278**, 114676 (2021).

60. Denil, M., Matheson, D. & De Freitas, N. Narrowing the Gap: Random Forests In TheDenil, M., Matheson, D., & De Freitas, N. (2014). Narrowing the Gap: Random Forests In Theory and In Practice. Proceedings of The 31st International Conference on Machine Learning, (1998), 665–673. Retrieved from ht. *Proc. 31st Int. Conf. Mach. Learn.* 665–673 (2014).

61. Attari, V. & Arroyave, R. Machine learning-assisted high-throughput exploration of interface energy space in multi-phase-field model with CALPHAD potential. *Mater. Theory* **6**, 5 (2022).

62. Fazilat, H. *et al.* Predicting the mechanical properties of glass fiber reinforced polymers via artificial neural network and adaptive neuro-fuzzy inference system. *Comput. Mater. Sci.* **58**, 31–37 (2012).

63. Fernández, M., Rezaei, S., Rezaei Mianroodi, J., Fritzen, F. & Reese, S. Application of artificial neural networks for the prediction of interface mechanics: a study on grain boundary constitutive behavior. *Adv. Model. Simul. Eng. Sci.* **7**, 1 (2020).

64. Geretti, L. & Abramo, A. Chapter 1 - The Synthesis of a Stochastic Artificial Neural Network Application Using a Genetic Algorithm Approach. in *Advances in Imaging and Electron Physics* (ed. Hawkes, P. W.) vol. 168 1–63 (Elsevier, 2011).

65. Smola, A. J. & Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **14**, 199–222 (2004).

66. Caruana, R. & Niculescu-Mizil, A. An empirical comparison of supervised learning algorithms. *ACM Int. Conf. Proceeding Ser.* **148**, 161–168 (2006).

67. Zhang, W., Zhang, R., Wu, C., Goh, A. T. C. & Wang, L. Assessment of basal heave stability for braced excavations in anisotropic clay using extreme gradient boosting and random forest regression. *Underground Space* (2020).

68. Chen, T. & Guestrin, C. XGBoost: A scalable tree boosting system. in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 785–794 (2016).

69. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011).

70. Ghojogh, B. & Crowley, M. The Theory Behind Overfitting, Cross Validation, Regularization, Bagging, and Boosting: Tutorial. (2019).

71. Ying, X. An Overview of Overfitting and its Solutions. *J. Phys. Conf. Ser.* **1168**, 022022 (2019).

72. Melki, G., Cano, A., Kecman, V. & Ventura, S. Multi-target support vector regression via correlation regressor chains. *Inf. Sci. (Ny).* **415**–**416**, 53–69 (2017).

73. Montiel, J., Read, J., Bifet, A. & Abdessalem, T. Scikit-multiflow: A Multi-output Streaming Framework. *J. Mach. Learn. Res.* **19**, 1–5 (2018).

74. Lu, P. *et al.* A novel spatio-temporal wind power forecasting framework based on multi-output support vector machine and optimization strategy. *J. Clean. Prod.* **254**, 119993 (2020).

75. Joly, A., Wehenkel, L. & Geurts, P. Gradient tree boosting with random output projections for multi-label classification and multi-output regression. *arXiv Prepr.* (2019).

76. Watt, J., Borhani, R. & Katsaggelos, A. Regression. in *Machine Learning Refined: Foundations, Algorithms, and Applications* 45–72 (Cambridge University Press, 2016). doi:10.1017/CBO9781316402276.005.

77. Hecht-nielsen, R. Theory of the backpropagation neural network. *Neural networks Percept.* 65–93 (1992).

78. Awad, M. & Khanna, R. Support Vector Regression. *Effic. Learn. Mach.* 67–80 (2015).

79. Breiman, L. Random Forests. *Mach. Learn.* **45**, 5–32 (2001).

80. Nguyen, D. L. H., Thanh Do, D. T., Lee, J., Rabczuk, T. & Nguyen-Xuan, H. Forecasting damage mechanics by deep learning. *Comput. Mater. Contin.* **61**, 951–977 (2019).

81. Chen, Q. *et al.* A dual approach in direct ink writing of thermally cured shape memory rubber toughened epoxy. *ACS Appl. Polym. Mater.* **2**, 5492–5500 (2020).

82. Friedman, J. H. Greedy Function Approximation: A Gradient Boosting Machine. *Ann. Stat.* **29**, 1189–1232 (2001).

83. Software package. https://drive.google.com/file/d/1Mv3duk75Fezz6u8FQgx9BdV3yAXW7mfy/view?usp=sharing.

84. Bui, V. Q., Marechal, E. & Nguyen-Dang, H. Imperfect interlaminar interfaces in laminated composites: delamination with the R-curve effect. *Compos. Sci. Technol.* **60**, 2619–2630 (2000).

85. Needleman, A. An analysis of decohesion along an imperfect interface. *Int. J. Fract.* **42**, 21–40 (1990).