MSDE

PAPER

View Article Online
View Journal | View Issue

Cite this: *Mol. Syst. Des. Eng.*, 2020, 5, 962

Mining structure–property relationships in polymer nanocomposites using data driven finite element analysis and multi-task convolutional neural networks

Yixing Wang,[a] Min Zhang,[b] Anqi Lin,[b] Akshay Iyer,[a] Aditya Shanker Prasad,[c] Xiaolin Li,[d] Yichi Zhang,[a] Linda S. Schadler,[e] Wei Chen[a] and L. Catherine Brinson*[b]

Data-driven methods have attracted increasingly more attention in materials research since the advent of the material genome initiative. The combination of materials science with computer science, statistics, and data-driven methods aims to expediate materials research and applications and can utilize both new and archived research data. In this paper, we present a data driven and deep learning approach that builds a portion of the structure–property relationship for polymer nanocomposites. Analysis of archived experimental data motivates development of a computational model which allows demonstration of the approach and gives flexibility to sufficiently explore a wide range of structures. Taking advantage of microstructure reconstruction methods and finite element simulations, we first explore qualitative relationships between microstructure descriptors and mechanical properties, resulting in new findings regarding the interplay of interphase, volume fraction and dispersion. Then we present a novel deep learning approach that combines convolutional neural networks with multi-task learning for building quantitative correlations between microstructures and property values. The performance of the model is compared with other state-of-the-art strategies including two-point statistics and structure descriptor-based approaches. Lastly, the interpretation of the deep learning model is investigated to show that the model is able to capture physical understandings while learning.

Received 13th February 2020,
Accepted 1st April 2020

DOI: 10.1039/d0me00020e

rsc.li/molecular-engineering

Design, System, Application

The objective of this paper is to demonstrate a novel, data-centric framework to accelerate the development of next-generation nanostructured polymers with unprecedented and predictable combinations of properties. This study is an application of NanoMine (a nanocomposite data resource) on understanding structure–property relationship for polymer nanocomposites. Materials science is founded on a processing–structure–properties (**p–s–p**) paradigm. Our approach, focusing on understanding structure–property relationship, represents a significant advance to state-of-the-art materials design and is targeted and demonstrated across a broad class of materials: nano-reinforced polymeric composites. This work demonstrates an applicable path on developing data driven and machine learning methods for material mechanism understanding and design: starting from analyzing archived data, then moving to application of computational methods for data augmentation, followed by exploring qualitative relationship by restricting confounding parameters, and finally building quantitative relationship using machine learning models. This strategy can be applied to understand other material mechanisms such as the processing–structure relationship and guide the design of material with targeted performance.

1. Introduction

Polymer nanocomposites are defined as organic matrix materials containing nanoparticles with at least one dimension below 100 nm.[1] By including a small amount of fillers to the matrix, experiments have shown that polymer nanocomposites can achieve a significant improvement in dielectrical, mechanical and optical properties compared with their parent matrix system.[2–6] The enhancement in properties comes from the large interphase region resulting from the strong chemical and geometrical interactions between the particle surface and polymer area nearby.[7,8]

[a] *Department of Mechanical Engineering, Northwestern University, Evanston, IL, USA*
[b] *Department of Mechanical Engineering and Materials Science, Duke University, Durham, NC, USA. E-mail: brinson@duke.edu*
[c] *Department of Material Science and Engineering, Rensselaer Polytechnic Institute, Troy, NY, USA*
[d] *Theoretical and Applied Mechanics, Northwestern University, Evanston, IL, USA*
[e] *College of Engineering and Mathematical Sciences, The University of Vermont, Burlington, VT, USA*

To predict the properties of a nanocomposite given its microstructure, people have developed different types of numerical methods including continuum mechanics methods and multi-scale simulations.[9–11] A variety of micromechanical models such as Mori–Tanaka, Halpin–Tsai and the self-consistent scheme have been developed to predict the thermomechanical behavior of nanocomposites.[12] However, those analytical models are not sufficient to fully capture the dispersion state or the morphology information of the fillers although some of the models include structural parameters. Therefore, multi-scale simulations are often necessary. Finite element (FE) simulations are able to fully capture the structural information and accommodate non-homogenous material systems with explicit configurations of all material phases, which makes FE simulation a good candidate to analyze the behavior of nanocomposites. FE models have been developed to simulate the thermal and mechanical properties of polymer nanocomposites and investigate the impact of interphase, the altered polymer region closest to the nanoparticles.[13,14] It has also been shown that the interphase properties in the FE models can be represented by shifting factors based on the pure matrix properties for some of the material systems.[14,15] The necessary shift for a given experimental sample can be determined automatically through Bayesian optimization.[16]

In order to better understand the behavior of these materials and further design materials with target properties, researchers have proposed the paradigm of process–structure–property (PSP) linkage.[17] Data-driven approaches founded on the PSP linkage have attracted significant attention recently and become an important emerging area in material research.[15,18,19] Data-driven approaches combine material science, computer science and statistics with the goal to expediate material discovery by utilizing research data from experiments or computational simulations. To facilitate the development of data archiving, sharing and development of data-driven approaches, efforts have been made to create online databases for fast queries and reference.

In the nanocomposite domain, an open source, data-driven web-based platform, NanoMine, has been developed which archives experimental and computational data on nanocomposites, including composition, processing methods, microstructural images and measured properties.[20,21] NanoMine allows fast data queries, visualization, and sharing as well as a number of tools for analysis including microstructure descriptor identification and reconstruction tools.[20,21] Utilizing data in NanoMine, a data-driven approach was developed to model the processing–structure relationship for a class of nanocomposites, connecting the nanocomposite processing parameters and the structural descriptors.[15] In that work, quantitative processing–structure relationship is built applying linear regression models on a relatively small set of data (about 20 samples) coming from three well-controlled material systems with different matrixes and surface treatments. Although that work provides insights on building quantitative relationship on a small set of data to understand material mechanisms in polymer nanocomposites, the continuous development of NanoMine enables more comprehensive quantitative studies using a larger datasets with more sophisticated machine learning algorithms. In addition, in another work a simulation based data-driven framework was proposed for designing and modeling of new material systems and structures, which includes design of experiments, computational models and machine learning methods.[19] A common characteristic involved in almost all the data-driven approaches is that a sufficient pool of data is required so to apply machine learning and statistic methods can be applied to extract features from the data and build correlations.

Deep learning, one of the sub-field of machine learning, has dramatically improved the state-of-art in computer vision, natural language processing and many other fields including material science.[22] Deep learning approaches provide an end-to-end framework addressing automated feature extraction for a large set of potential features. Using these approaches, it is not necessary to design explicit features, which is usually required in traditional machine learning approaches. Deep learning has been applied in material science specifically for the case where structural images are involved. Yang *et al.* applied deep convolutional network to model structure–property linkages for high-contrast elastic 3-D composite microstructures.[18] In,[23–25] deep neural networks are applied to reconstruct microstructure images and implemented to design material microstructure with desired properties.

In this paper, we investigate the structure–property linkage for polymer nanocomposites using a database of simulated data. Simulations provide the flexibility to explore a wide range of structures and the ability to create and analyze a large amount of data. Although previous studies have demonstrated data-driven approaches to build structure–property linkages, a systematic study showing the impact of different factors (composition, dispersion, reconstruction method *etc.*) to different properties is still lacking. Additionally, most of the previous studies relied on traditional machine learning method using hand-crafted features, in which the quality of the prediction model depends on the quality of designed features based on material expertise. Finally, it is nontrivial to conduct feature selection from a large pool of microstructural features. In this study, by leveraging artificial generated microstructures and FE simulations, a sufficiently large database with different microstructures and associated properties can be obtained. The property of interest for this research is three mechanical properties of the bulk composites: $\tan\delta$ peak, glassy modulus and rubbery modulus. We aim to investigate the effect of microstructure to these target properties in a systematic and comprehensive way. The first part of the paper focuses on investigating the effect of microstructure reconstruction method, filler composition and dispersion, and interphase on the bulk properties. After that, we design and explore the application of a deep multi-task

convolutional network to quantitatively predict multiple property values given a microstructure image. This is the first time a deep multi-task learning model is applied to build structure–property linkage for prediction of mechanical properties of polymer nanocomposite. The results from proposed method is further compared with other methods including traditional approaches using physical descriptors and two-point statistics.

2. FE simulation and analysis

FE simulations are built using a library of microstructures created from several different reconstruction algorithms, which result in different degrees of dispersion. The properties of interests chosen for this study are the viscoelastic properties of tan δ peak magnitude, glassy modulus and rubbery modulus. In this section the results of direct simulation are presented for these different microstructures and the impact of an interphase of altered properties is also demonstrated. Machine learning methods to discover structure property relationships in the data sets are presented in the following section.

2.1 Experimental data limitations

We first begin with qualitative understanding of structure–property relationship by plotting structure parameters with the property values. Based on simple physical understanding, the glassy and rubbery moduli should increase as more stiff fillers are added to the polymer system. On the other hand, the tan δ peak should decrease with higher loading due to replacing a damping material (polymer matrix) with perfectly elastic particles with no damping. At first blush, one would expect to see these simple trends in any data set for nanocomposite properties with sufficient amount of data. Fig. 1 illustrates the experimentally measured tan δ peak as a function of volume fraction for all the samples currently stored in NanoMine with volume fraction less than 10%. The plot shows no statistically significant trend, which contradicts our initial hypothesis. While NanoMine contains thousands of samples, the data encompasses a wide range of matrix and filler materials, different processing conditions, environmental conditions and even measurement methods. Therefore, even with a relatively large amount of data, no consistent trend can be observed as there are too many confounding variable parameters. Additionally, the overall data volume in NanoMine is still too small to be able to narrow the search sufficiently to conduct a systematic study. While the curation process of data into resources such as NanoMine will provide sufficient data in the long term, to be able to explore structure–property relationships with machine learning methods now, a computational approach is needed to give the flexibility to explore a small, consistent set of structural parameter combinations to build a comprehensive linkage to their effect on the target property response.

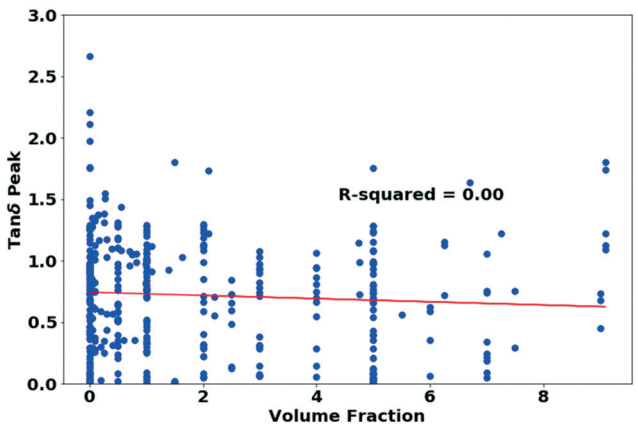

Fig. 1 Plot shows tan δ peak as a function of volume fraction using the data from NanoMine.

2.2 Computation method

In order to demonstrate the application of data driven approach in correlating the structure–property relationship in polymer nanocomposites, we generated a set of simulated data using a combination of microstructure reconstruction methods and finite element (FE) analysis. The FE work-flow is shown in Fig. 2. Polycarbonate (PC) is considered as the material property for the FEA and the frequency domain response of that material is measured by dynamic mechanical analysis (DMA).[26] Previous study has shown that neither the magnitude nor the anisotropy of nanoparticle modulus has a measurable impact on the viscoelastic properties of nanocomposites,[14] therefore, in this paper, for simplification, the silica nanoparticle in the model is assigned to be linear, elastic with young's modulus of 73 GPa and Poisson's ratio of 0.3. In the FE model, we can incorporate an interphase of altered properties near the filler[14] and will present results both with and without an interphase. It has been shown that the mobility of polymer chains can be restricted due to attractive interactions between nanoparticles and matrix.[27–30] In addition, both simulation and experiment results have shown that the glass transition temperature of nanocomposites can be tens of degrees higher than the pure matrix.[29,31,32] Based on the above observations and one of the previous study,[14] we hypothesize that interphase properties are obtained by shifting the matrix polymer master curve two decades lower in the frequency domain, which correlates to a strong interaction between the nanoparticle and the polymer matrix and locally stiffer response of the polymer near the particles. For this study, we choose tan δ peak, glassy modulus and rubbery modulus as the properties of interest and correlate those with the structural inputs. The output from FEA is the frequency domain response of the nanocomposite, based on which tan δ peak, glassy modulus and rubbery modulus can be extracted.

In this study, the structural inputs (referred to as a representative volume element or RVE) for the FEA are generated using different microstructure characterization

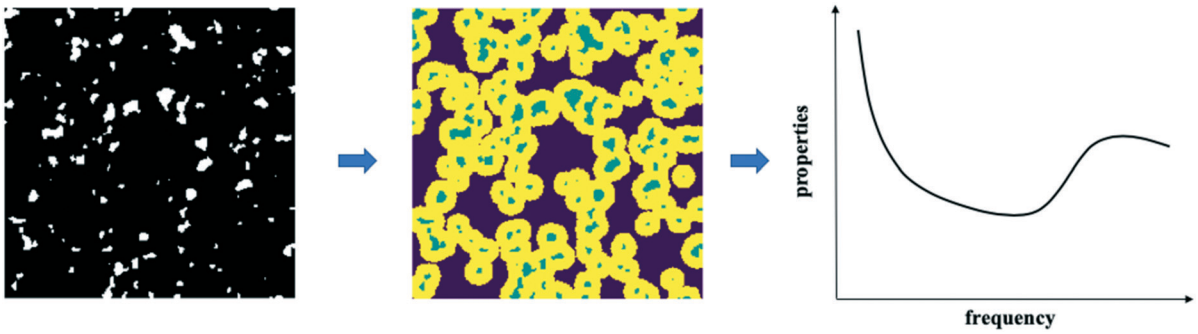

Fig. 2 FEA work-flow. Given a microstructure, an interphase layer is assigned assuming a fixed thickness and uniform stronger interaction between the nanoparticle and the polymer matrix, leading to enhanced interphase properties. Then the FE simulations for viscoelastic properties as function of frequency are run for samples with and without interphase.

and reconstruction (MCR) methods, which are commonly applied in predictive material modeling.[19,25,33] Three different types of MCR methods are applied to generate RVEs for the FE model: uniform dispersion, descriptor-based method, and spectral density function (SDF). The impact of these different types of microstructures on the structure–property relationship is investigated in the later section. Other studies also show more organized supramolecular structures, such as strings or sheets, which can be obtained by controlling grafting density and molecular weight of grafting chains.[34,35] In this paper, we do not focus on these super-structures cases, but rather on the simpler dispersion differences in overall isotropic systems, which forms the vast majority of the nanocomposite literature to date. Microstructures with uniform dispersion are generated using the random sequential adsorption (RSA) algorithm such that the centroid of all existing particles have to exceed a minimum value and therefore particles do not overlap/merge with one another. Similar methods[14] have been applied to generate the centers of spherical particles with same radius in the unit cell. Descriptor-based reconstruction generates statistically equivalent random digital microstructures by matching the prespecified structural descriptors (*i.e.* nearest neighbor distance, volume fraction *etc.*) through optimization.[15] Using a predefined set of four descriptors, volume fraction (VF), number of clusters (n), average nearest cluster distance ($\bar{r}_d$), aspect ratio (e_l), microstructures with different volume fraction and dispersion states can be obtained. The obtained microstructures have elliptical clusters in which particles can overlap with each other and therefore merge to form more complex filler shapes, in contrast to the uniform dispersion microstructures. Spectral density function (SDF) has been demonstrated to be able to characterize complex heterogeneous microstructures and reconstruction can be done through phase recovery techniques.[36] For the nanocomposite system, the SDF of all microstructures approximately follows an exponential distribution with two parameters – shape variable α and scale variable θ. For these systems, previous work has also shown that the shape parameter α varies in a small range and has very little influence on SDF profile and the consequent microstructure.[37,38] However, the scale parameter θ has a wide range and can greatly impact the decay of SDF and the dispersion state of microstructure.[36] Therefore, for this study, a set of microstructures are generated by controlling the volume fraction (VF) and scale parameter θ while the shape parameter α is fixed.

2.3 Impact of microstructure reconstruction method

The three representative microstructure reconstruction/generation methods are evaluated in terms of their impact on the properties. For each reconstruction method, 100 microstructures with different volume fraction (1–20%) and dispersion states are generated using optimal Latin hypercube sampling (OLHC).[39] Sample microstructures obtained from different methods are shown in Fig. 3. Here, we merely compare the impact of the reconstruction methods and there is no interphase layer added for the computation. Microstructures with uniform dispersion have spherical particles, which are uniformly distributed in the matrix and with no overlap. Physical descriptor reconstructed microstructures have elliptical clusters with different aspect ratio and the particles and clusters can overlap and form even larger clusters. Microstructures reconstructed using SDF have irregular shape clusters. All three microstructure types are realistic for different types of nanocomposite materials as shown in Fig. 4. The simulated properties *versus* volume fraction are shown in Fig. 5. The results show that for all the microstructures, the $\tan\delta$ peak decreases monotonically as a function of volume fraction while the glassy and rubbery modulus increases monotonically. This observation matches the physical understanding and result in previous work.[40] $\tan\delta$ peak is a measure of material damping and the reduction of the magnitude arises from the replacement of polymeric (damping) material with perfectly elastic participle (no damping). It is also noted that as the microstructure becomes more complex (from uniform dispersion to SDF), the r-squared values decreases, indicating a less monotonical trend for all three property values with increasing geometric complexity.

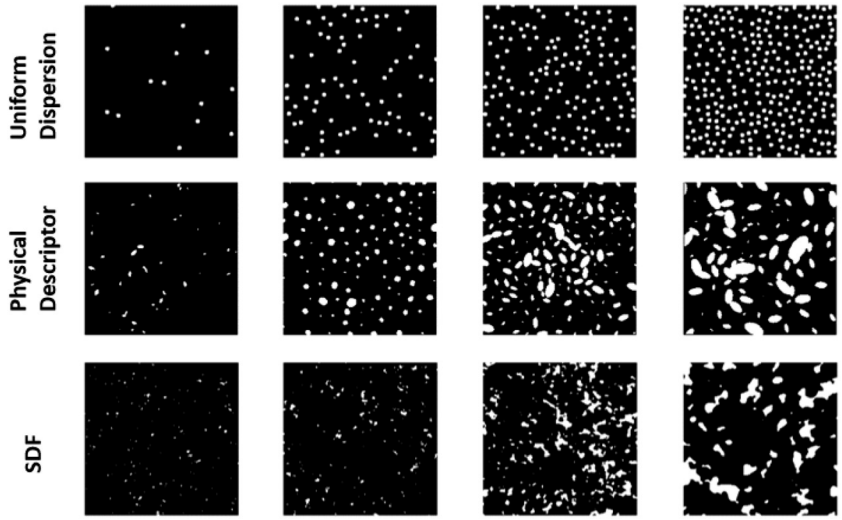

Fig. 3 Sample microstructures generated using different methods: first row, uniform dispersion; second row, physical descriptor; third row, spectral density function.

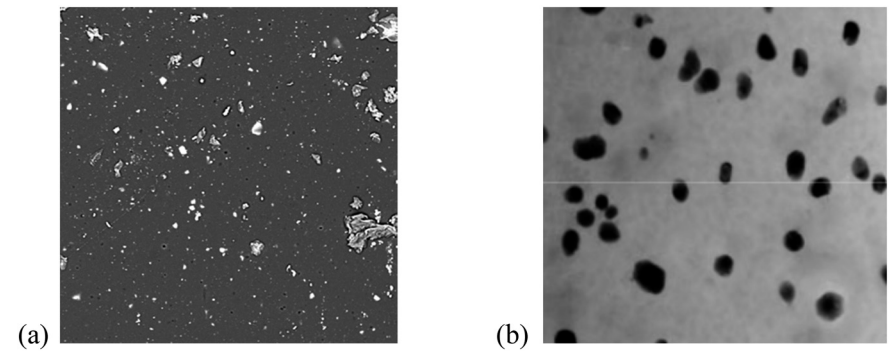

Fig. 4 Experimental images showing different types of microstructures. (a) SEM images of octyl-modified silica in PMMA, which could be well represented by SDF approach. (b) TEM images of Ag/C core/shell hybrid particles in the epoxy,[41] which could be well represented by descriptor approach.

2.4 Impact of interphase

Previous studies have shown that the interphase region plays an import role in the bulk composite properties.[10,13,14,16] The interphase properties can be represented by shift factors related to that of the matrix. In this study, we assume a fixed thickness of interphase whose properties are determined by shifting the master curve of the matrix two decades lower in the frequency domain as shown in Fig. 6. Although in this study, for simplicity the interphase properties are assumed to be uniform, gradient interphase properties can be considered in the future as well based on a new gradient interphase representation method in the FE model coming from observations from the local measurement of interphase.[13]

FE simulations are run using the same set of microstructures using SDF approach as in 2.2 with interphase region. The results are shown in Fig. 7. The property enhancement from the interphase layer exhibits an unbalanced reinforcement for the rubbery and glassy modulus, where the reinforcement of the rubbery modulus is more significant than that of the glassy modulus. While the variance remains similar, the glassy modulus exhibits almost the same range of increase with volume fraction as the no-interphase results while the rubbery modulus exhibits a much larger increase in magnitude compared to the no-interphase case. More significantly, the results for the $\tan\delta$ peak magnitude are dramatically different: $\tan\delta$ is nearly random with no discernable decrease with volume fraction, in contrast to the monotonic decrease for the simulations with no interphase. The results for the simulations with interphase for $\tan\delta$ are quite similar to the dataset from NanoMine for a wide range of experimental data (Fig. 1). This result can help explain the findings in Fig. 1 from NanoMine: since the data from experimental samples are wide ranging and in many cases the samples will contain an interphase of altered properties near the filler, the lack of trend in $\tan\delta$ can be expected over that broad data set.

To better understand the results for the $\tan\delta$ trends, the volume fraction and dispersion parameter are restricted

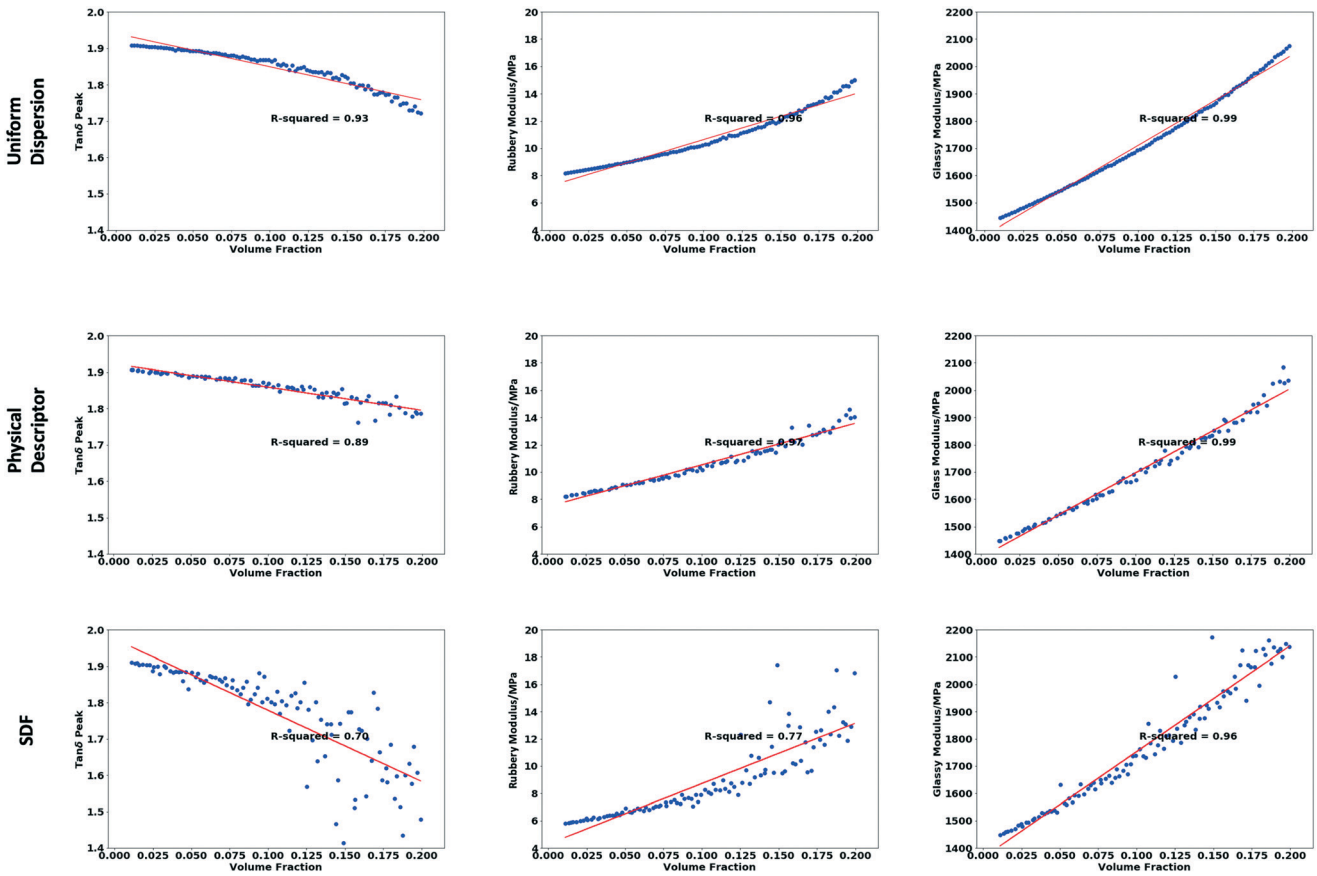

Fig. 5 Comparison of simulations using different type of microstructures for composites with polymer matrix, silica filler and no interphase. For each type of microstructure, three individual property value (tan δ peak, glassy modulus and rubbery modulus) is plot against volume fraction.

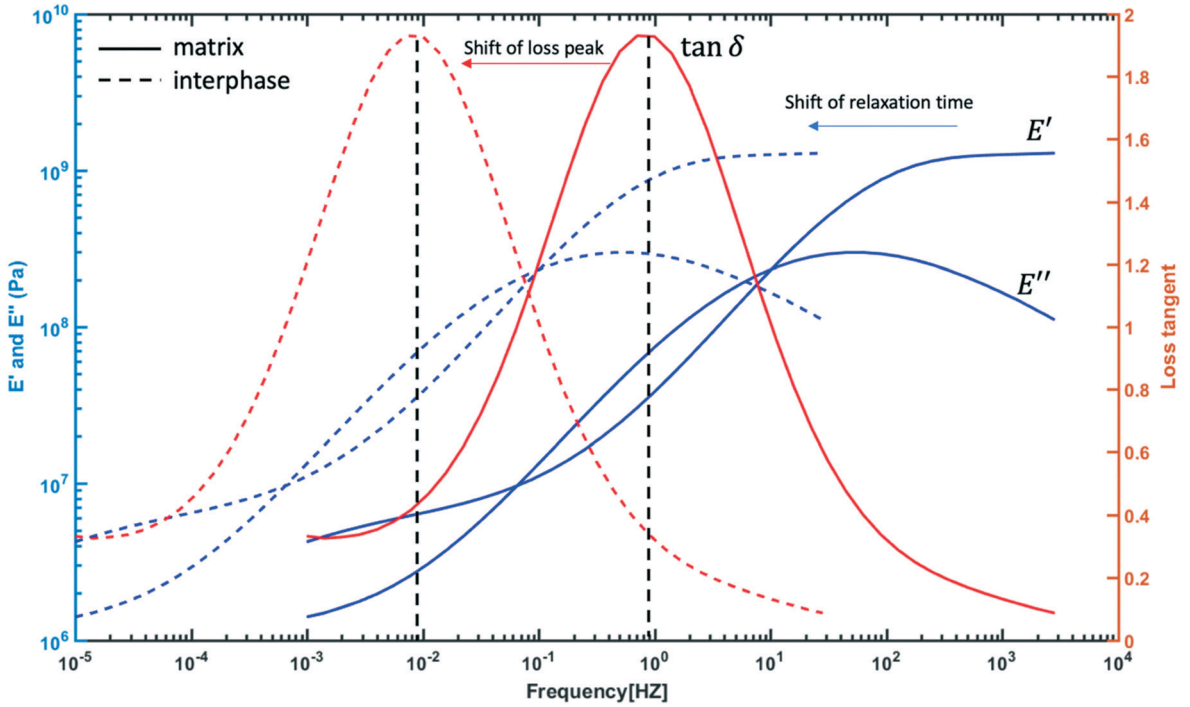

Fig. 6 Relationship between matrix and interphase properties. The interphase properties are determined by shifting the master curve of the matrix two decades lower in the frequency domain.

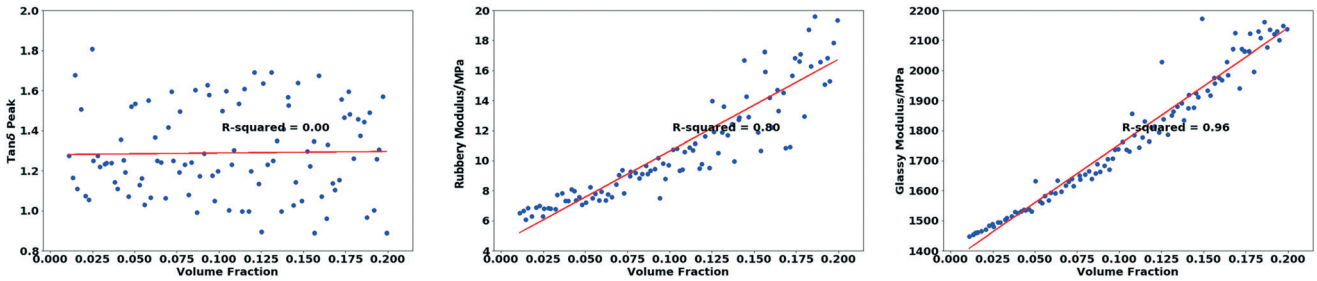

Fig. 7 Results for simulation with interphase region. Three individual property value (tan δ peak, glassy modulus and rubbery modulus) is plot against volume fraction using the microstructures from SDF approach as in 2.2.

individually in different trials. The dispersion state of microstructures is controlled by the factor theta, which varies from [0.1,1]. Fig. 8 shows the tan δ peak as a function of volume fraction for data under similar dispersion, where the SDF parameter theta is restricted to different ranges ([0.1,0.3], [0.2,0.4], [0.6,0.8]) respectively. With constraint on the dispersion, the tan δ peak exhibits a decreasing trend as a function of volume fraction at lower θ ranges ([0.1,0.3] and [0.2,0.4]), similar to the trend in the no-interphase calculations in 2.2. On the other hand, at a higher θ ranges, tan δ peak shows an opposite trend: increasing as a function of volume fraction. This can be explained by the fact that the behavior of the nanocomposite is gradually changing from being dominated by the matrix to dominated by the interphase. Interphase volume fraction can increase both with dispersion of the fillers in the matrix and a larger filler loading. However, in the case of poor dispersions, interphase layers will overlap, leading to minimal change in interphase fraction with increasing filler loading. Therefore, at lower dispersion levels, the tan δ peak value decreases with filler loading because the bulk property of the composite is controlled by matrix and filler and the replacement of polymeric (damping) material with perfectly elastic participle (no damping) decreases overall dissipation. At higher dispersion levels, the tan δ peak value increases because the property of the composite is dominated by the interphase as the interphase volume fraction exceeds a percolation threshold.[14] This result implies that for composites with interphase, both the loading and dispersion play a critical role in the determination of tan δ peak values.

We are also able to investigate the converse situation where the volume fraction is restricted to different small ranges ([1%,5%], [5%,10%], [10%,15%]) respectively and the effect of dispersion is varied. Fig. 9 shows the tan δ peak as a function of dispersion for data under similar volume fraction. The blue data points in Fig. 9 shows that at a lower volume fraction range ([0.01,0.05]), the tan δ peak decreases as the dispersion increases. This suggests that at this lower loading, despite microstructures with different dispersions, the volume fraction of interphase does not exceed the interphase percolation threshold and the relaxation behavior of the composite is still being dominated by the matrix, with the peak located at the PC matrix frequency location (see Fig. 6). As a result, as the dispersion improves, larger interphase volume reduces the tan δ magnitude at the matrix peak frequency location (broadening the peak by shifting some magnitude toward the interphase peak frequency location) and results in lower tan δ peak values. On the other

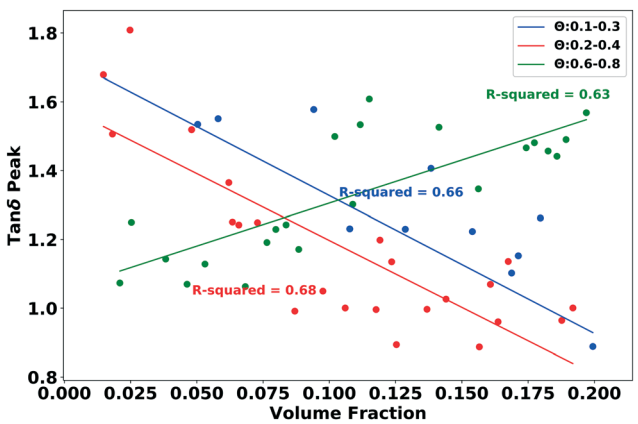

Fig. 8 tan δ peak as a function of volume fraction under different dispersion ranges. For each dispersion range: [0.1,0.3] (blue), [0.2,0.4] (red), and [0.6,0.8] (green), tan δ peak is plotted against volume fraction based on simulation results with interphase region.

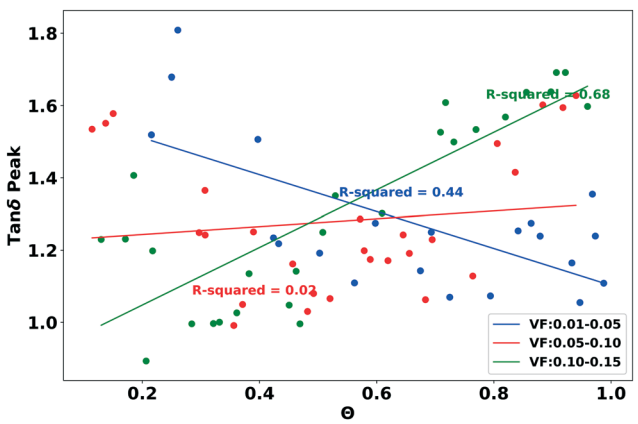

Fig. 9 tan δ peak as a function of dispersion under different volume fraction ranges. For each volume fraction range: [0.01,0.05] (blue), [0.05,0.01] (red), and [0.10,0.15] (green), tan δ peak is plotted against dispersion based on simulation results with interphase region.

hand, at higher volume fraction range (in green, [0.10,0.15]), the tan δ peak increases as the dispersion improves. This result can be explained by that fact that the property of the composite is dominated by the interphase as the interphase volume fraction exceeds the percolation thresholds at this higher loading and therefore more interphase leads to increasing the composite tan δ peak which for the percolated cases is located near the interphase tan δ peak frequency location. The red data points in Fig. 9 shows the tan δ peak first decreases and then increases as a function of dispersion, which provides evidence of the transition of composite property from being dominated by the matrix to being dominate by the interphase.

Based on two controlled experiments, overall, adding interphase to the system greatly impacts the tan δ peak value and the value is dependent on both the loading condition (volume fraction) and the dispersion state (theta). These results also further clarify the no-trend result in the experimental data from NanoMine: these data sets contain microstructures with very different dispersions conditions, which must be accounted for to understand the damping response of the composite.

3. Deep learning model and result

In the prior section, we have qualitatively shown that the mechanical properties of the composites vary according to microstructures with different filler loading and dispersion. The next task is to quantitatively predict the properties of interests (tan δ peak, glassy modulus and rubbery modulus) given a microstructure. The remainder of this paper focuses on building quantitative structure–property relationship using the interphase FE model with RVEs generated using SDF *via* a machine learning approach.

3.1 Deep learning model development

The objective is to predict three continuous property values given a 2D image, which can be considered as a regression problem using features from a 2D matrix. There are two common strategies that can be applied to build this regression model: one is to use hand-crafted features (usually geometrical descriptors) that could sufficiently capture the characteristics of a given microstructure; the other is to use feature-engineering free method such as deep learning. The quality of the prediction using hand-crafted features is heavily dependent on the quality of designed feature from material experts and these features are usually not transferable from one material system to the other. On the other hand, compared with traditional machine learning approach using hand-crafted features, deep learning method is considered as a feature engineering free method that is able to automatically learn critical features during the training process, and have shown significant improvement in learning ability, generalization and transferability.

Convolutional neural network (CNN) is a deep learning approach and has been widely applied for computer vision related task, such as image classification and object recognition. In those tasks, CNN outperforms conventional methods to a large extent due to its capability of extracting high-level abstractions of inputs through a series combination of non-linear transformations. A common CNN may include three different type of layers: convolution layer, pooling layer and fully connected layer. The objective of the convolution layer is to extract critical feature maps from the input images through applying filters with different number and size. The values in those filters are learned from available data. The pooling layer is often applied after convolution layers with the aim to reduce the dimension of the feature maps. Different down sampling strategies can be taken including max-pooling, average pooling and L2-norm pooling. By combining a series of convolution and pooling layers, a series of feature maps can be obtained and then fed to the fully connected layers for prediction of a class or a single value depending on the purpose of the task. The convolution and pooling layers are regarded as the feature extractor while the fully connected layers act as a non-linear regression model using the features from the first part. The nonlinearity in the network is introduced through applying activation functions, such as ReLU, sigmoid, tanh.

Our target problem involves multiple outputs: tan δ peak, glassy modulus and rubbery modulus. In order to predict multiple values given a microstructure image, two strategies can be taken: one is to train separate machine learning models for specific tasks (*i.e.* for our purpose, three separate machine learning models would be required to predict three different property values); the other strategy is to develop a single machine learning model that solves multiple learning tasks at the same time, which is also called multi-task learning (MTL). Fig. 10 shows two common ways to perform multi-task learning in deep neural networks: hard or soft parameter sharing.[42] The most common strategy in MTL is hard parameter sharing, where several hidden layers are shared across all tasks while task-specific layers are applied after to predict the value for different objectives. It has been shown that hard parameter sharing can reduce overfitting by including shared parameters. This advantage arises because the model must find hidden representations or features that captures all the tasks, which improves model generalization and reduces the chances of overfitting. For soft parameter sharing, each task has its own parameters. But the parameters for the different tasks are constrained to encourage the parameters to be similar. Different constrains are usually applied such as L2 distance for regularization and the trace form. MTL has also been demonstrated to be extremely helpful if tasks share some similarities and the data volume is relatively small. MTL appears to be a good candidate for our purpose as the three learning tasks share significant commonalities, which are to predict mechanical responses given a set of microstructure images. Additionally, because of the expensive cost of FE simulations, it is time-consuming to generate a very large volume of data. By utilizing MTL, it is aimed to implicitly augment the data,

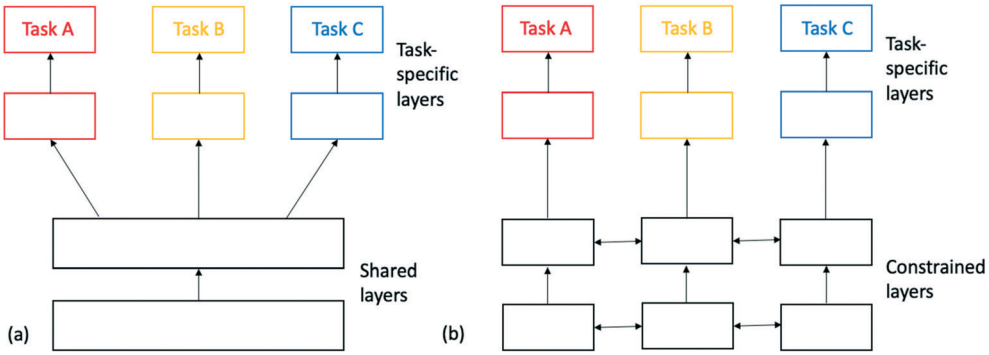

Fig. 10 Architectures of two types of MTL models. (a) Hard parameter sharing by applying hidden layers that shared across different tasks; (b) soft parameter sharing by having constrained individual parameters for different tasks to encourage similar parameters.

reduce the chances of overfitting and improve the model generalization.

Taking advantage of the MTL and CNN, in order to properly capture and extract high-level microstructure features and build the linkage of structure to multiple property values simultaneously, a multitask CNN model is proposed. The overall architecture of the model is shown in Fig. 11 and the detailed configurations and parameters of each layer is shown in Table 1. The input for this model is a 256 × 256 two phase binary image with zeros and ones, where zero represents the matrix and one represents the filler. Even though the FE model includes interphase layers in the calculation, the input for the deep learning model only consists two phases, filler and matrix. The image is pre-processed by replacing zero with −0.5 and 1 with 0.5, which is a useful method to improve the performance of convolution layers.[18,43] The model consists a number of shared convolution and pooling layers and three sets of tasks specific layers including convolution and fully connected layers. The shared layers, including convolution and pooling layers with different sizes, aim to extract the critical features from the images for determination of mechanical properties. Taking the features from shared layers, the task specific layers are designed to further extract task specific features and train different regressors for different outputs. The underlying reason for designing this architecture is that the extracted high-level features from convolutional layers should be shared across three highly related tasks while the weights for the later task specific layers are updated separately such that three different regressors are trained to predict different property values. The loss function for the model is formulated as the sum of mean absolute percentage error (MAPE) across three tasks:

$$\text{loss} = \sum_{j\in\{T,G,R\}} \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\overline{y_j^i} - y_j^i}{y_j^i}\right| \times 100\%$$

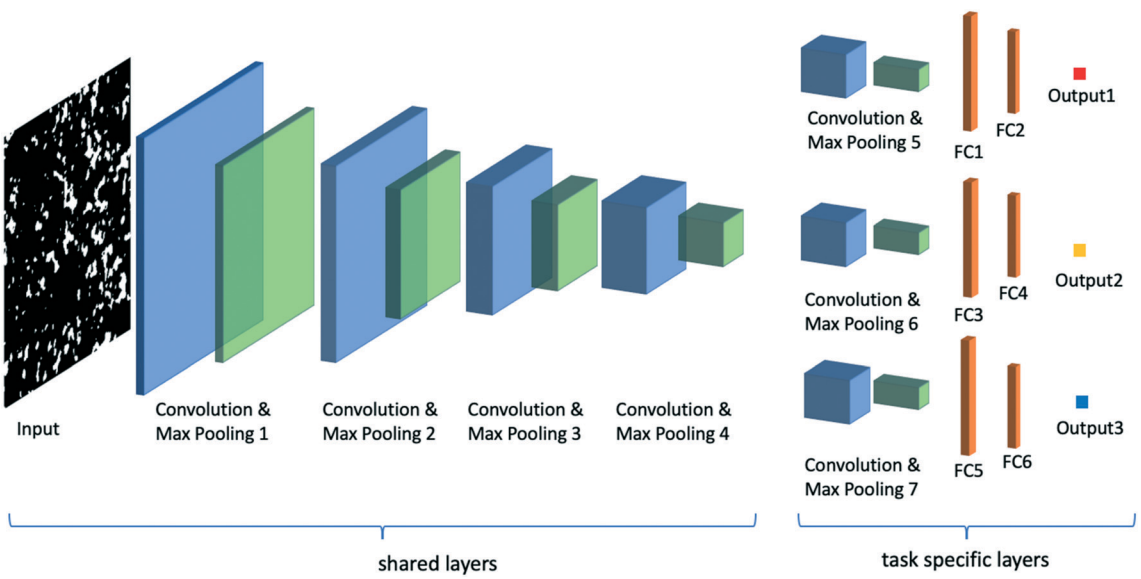

Fig. 11 Architecture of the proposed multi-task deep CNN model. The MTL is achieved through hard parameter sharing. The input image is first feed to a series of shared convolution and pooling layers to extract the high level shared structural features for different tasks. Following that, three sets of task specific layers, including one more convolution and pooling layers and two fully connected (FC) layers are applied to predict different output.

Table 1 Model configuration of the proposed multi-task deep CNN model. For each conv block, 3 × 3 convolution and ReLU activation with back normalization (l2 norm rate = 0.0005) is applied. The pooling size is 2 × 2. After all the convolution and pooling process, the weights are flattened and feed to two fully connected (FC) layers

Layer	Type	Dimension
Input	Shared	256 × 256 × 1
Conv and pooling 1	Shared	128 × 128 × 16
Conv and pooling 2	Shared	64 × 64 × 32
Conv and pooling 3	Shared	32 × 32 × 64
Conv and pooling 4	Shared	16 × 16 × 128
Conv and pooling 5/6/7	Task specific	8 × 8 × 128
FC 1/3/5	Task specific	1 × 1 × 512
FC 2/4/6	Task specific	1 × 1 × 256
Output 1/2/3	Task specific	1 × 1 × 1

where j represents the type of task, for our case it only has three values: $\tan\delta$ peak (T), glassy modulus (G) and rubbery modulus (R). $\overline{y_j^i}$ represents the i-th predicted value for task j and y_j^i represents the true value and N is the total size of training data. The training of this model is achieved by minimizing the loss function through back propagation and optimization. Additionally, in order to prevent over-fitting, common methods in deep learning including batch normalization and dropout is applied.

3.2 Datasets

In order to demonstrate the performance of proposed method, a data set with total size of 11 000 is created using generated microstructures and FEA model. Use of data from FEA simulations gives flexibility to explore a wide range of structures and the data can be accumulated in a relatively short period.

Each microstructure, with dimension 256 × 256 pixels, is generated from SDF and considered as an RVE for the FEA. As illustrated in previous section, the volume fraction and the dispersion are controlled by the volume fraction (VF) and scale variable θ respectively. By sampling 11 000 combinations of VF and θ, 11 000 RVEs with different volume fraction and dispersion are obtained. FEA is run for each RVE using Abaqus. Polycarbonate is chosen as the matrix material, whose master curve is measured through DMA.[26] The filler is silica nanoparticle with a linear elastic young's modulus of 73 GPa and Poisson's ratio of 0.3. The interphase layer has a thickness of 10 pixels, whose properties are determined by shifting the master curve of the matrix two decades lower in the frequency domain. The FEA outputs the frequency response of mechanical properties of the nanocomposite, from which three properties of interest $\tan\delta$ peak, glassy modulus and rubbery modulus can be extracted. To create this data set, 4 simulation jobs are run in parallel with each one having 16 CPUs on 2 work stations. The total computation time is about two weeks with each simulation taking about 20 minutes.

The 11 000 simulated data is further divided into a training, validation and testing set with a ratio of 7 : 1.5 : 1.5 : 7700 data are used to fit the deep learning model and 1650 is used to tune the hyperparameters in the model and the rest of data is reserved as testing data to evaluate the accuracy of the model. The accuracy of the model is evaluated using mean absolute percentage error (MAPE) which is given by:

$$e = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\overline{y^i} - y^i}{y^i}\right| \times 100\%$$

where $\overline{y^i}$ is the predicted value from the deep learning model and y^i is the true value from the FE model.

The FE model is run using Abaqus, a widely applied commercial software for FE simulations. The simulations are run on a work station with 192GB RAM and 16 core Intel(R) Xeon(R) CPU E5-2630 v3@2.40 GHz. To build the deep learning model, Python 2.7 and Keras with Tensorflow backend is adopted. The model is trained on a work station with a NVIDIA Quadro P5000 GPU with 16GB GPU RAM and 20 core Intel(R) Xeon(R) CPU E5-2630 v4@2.20 GHz with 192GB RAM.

3.3 Model performance

The performance of the model is benchmarked against two other common strategies for predicting structure–property linkage. Fig. 12 shows different strategies to predict the properties given a microstructure image. One of the traditional methods is to build correlations using hand crafted features such as volume fraction, nearest neighbor distance, aspect ratio *etc.*[19,44] Those features are usually designed by domain experts with the aim to capture the composition, dispersion as well as the geometric information of the microstructure. The structural features are then ready to be applied to fit a machine learning model to predict the property values. Another strategy to build the structure–property linkage is to compute two-point correlation functions from a microstructure and applied as features for the regression-based models.[45–47] Additionally, dimension reduction technique such as PCA is often employed on the two-point statistics because of the extremely high dimensionality. In order to evaluate the robustness and stability of the model, for every method, the model is trained and evaluated for ten times on different training testing splits. Additionally, to compare the performance of different methods, different models are given the same set of training, validation and testing set. The result is shown in Table 2.

The result shows that across all the tasks, the MTL-CNN model outperforms other methods. Additionally, the value of the standard deviation on all trials is also small indicating the robustness of the model. The method using geometrical descriptors even outperforms the two-point statistics especially for the predicting of rubbery modulus and $\tan\delta$ peak, which suggests that those hand-crafted descriptors designed by material express are really good at capture the

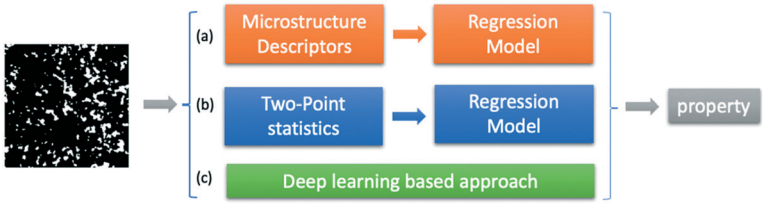

Fig. 12 Different types of methods to predict structure–property relationship. (a) Geometric descriptor-based approach (*i.e.* application of hand-crafted geometric features such as volume fraction, aspect ratio *etc.* to a regression model); (b) two-point statistic methods (using two-point statistics as features for the regression model); (c) deep learning-based approach (feature engineering free).

Table 2 Result comparison for different methods. The value shows the mean of MAPE together with the standard deviation across ten trials

Method	Glass modulus	Rubbery modulus	tan δ peak
Two-point statistics + PCA	1.23% ± 0.03%	5.40% ± 0.14%	8.99% ± 0.16%
Geometrical descriptors	1.16% ± 0.02%	4.54% ± 0.08%	4.55% ± 0.10%
MTL-CNN	0.68% ± 0.06%	3.12% ± 0.06%	3.58% ± 0.14%

characteristics of the microstructure and informative to the bulk properties of the composites. Compared with method using geometrical descriptors, deep learning model improves the accuracy for prediction of glass modulus by as much as 41.3%, rubbery modulus by 31.2%, and tan δ peak by 21.3%.

3.4 Model interpretation

As given in Table 2, deep learning model outperforms other traditional machine learning model for all the three target tasks. However, deep learning model is relatively hard to interpret compared with traditional method using hand crafted features. In another word, for traditional machine learning method, it is easy to extract the impact of every feature to the target while deep learning model works as a black box with complex architectures and millions of hidden parameters. Although it is not possible to interpret every parameter in a deep learning model, it is worth the effort to understand what the model learns and whether it captures some physical knowledge. Therefore, in this section, by modifying the inputs to the deep learning model, we try to interpret what the model learns from physical point of view.

In section 2, it has been shown that the glassy modulus increases monotonically as a function of volume fraction. This result suggests that the filler phase has larger impact on the glassy modulus of the composite than the matrix. To validate whether the proposed model captures this physical knowledge, the original microstructure input is modified by randomly removing a number of pixels from filler or matrix respectively and the prediction from the DL model on this partially incomplete microstructure is obtained. The removal process is achieved by setting the value of the chosen pixel to 0, which eliminates the activation of the model from that pixel and therefore its contribution to the target value. Fig. 13 shows an original image and the modified images after removal different number of pixels from filler or matrix. The modified responses from those two sets are compared with the known FEA response on the original, unmodified microstructure. The change of modified response is evaluated according to the residual equation below:

$$r = \frac{|\bar{y} - y|}{y} \times 100\%$$

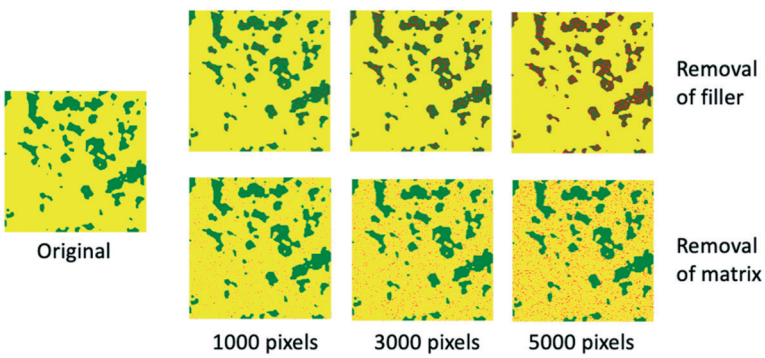

Fig. 13 Visualization of the original microstructure and the modified microstructure with different number of pixels removed from the original image. Yellow represents the matrix region while green represents the filler region. The pixels with red color are removed by setting value to zero.

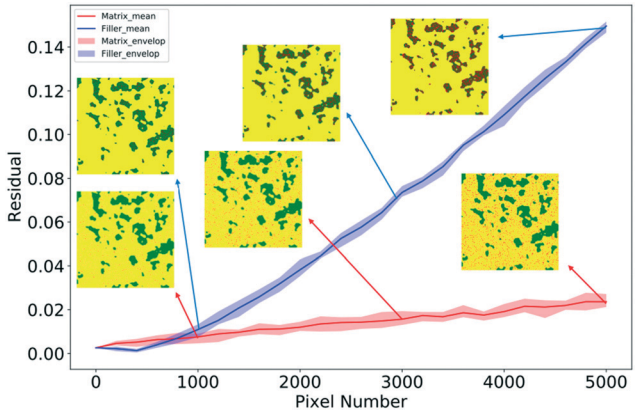

Fig. 14 The plot shows the residual as a function of removed pixels for different material phases on a single testing sample. The experiments are conducted ten times for robustness and accuracy. The colored area shows the distribution of values for ten trials and the solid line shows the average.

where $\bar{y}$ is the predicted value from the deep learning model and y is the true value from FE model.

Fig. 14 shows the contribution of matrix and filler in the trained deep learning model by plotting the residual against the number of removed pixels. The plot is based on a single testing sample and the experiments are conducted ten times for accuracy and stability. The plot shows that as the number of removed pixels from either from matrix or filler phases increases, the residual value increases monotonically, indicating that the deep learning model has lower accuracy as more pixels are eliminated. More importantly, the residual value from removal of filler pixels is much higher than the residual value from removal of matrix pixels, which suggests that the contribution from the filler is higher than matrix in determination of glassy modulus. This finding, corresponding to the knowledge illustrated earlier, were not explicitly introduced to the deep learning model during the training process. The proposed deep learning model is able to learn and capture this physical knowledge during the training process. Additionally, in order to show that this finding holds for other samples, the same experiments are done for ten random chosen testing samples and the average residual is calculated across those samples. Fig. 15 shows the residual for ten different samples. The plot shows that on average, the effect from the filler is higher than the matrix in determination of glassy modulus. Therefore, the trained deep learning model is able to implicitly learn and capture knowledge about the physical distinction between filler and matrix and use that information for more accurate predictions.

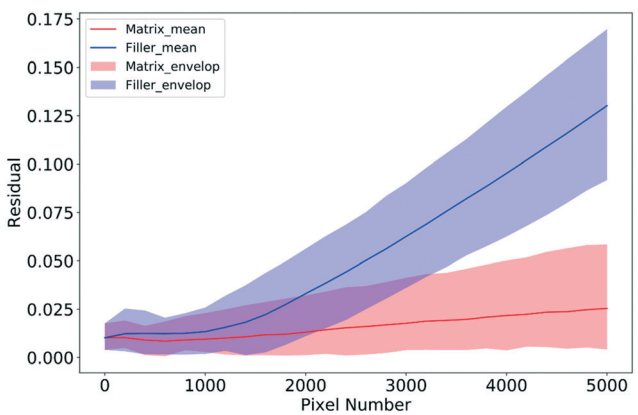

Fig. 15 Average residual plots for ten different testing samples. The colored area shows the distribution of average residual for different samples while the solid line shows the mean value.

In addition, the deep learning model is able to capture the implicit role of interphase in determination of different $\tan\delta$ trends as described earlier in section 2.4. This finding is demonstrated by generating a new data set of 100 microstructures where the $\tan\delta$ values are predicted based on the trained deep learning model without running FEA. Analysis is performed to investigate the trend of $\tan\delta$ as a function of volume fraction by controlling dispersion at different levels (similar to Fig. 8, data not shown here). Without running any direct simulations, the predicted values exhibit a similar trend as in section 2.4, capturing the interphase dominance at higher volume fractions. This test suggests that the deep learning model is able to implicitly learn the critical role of interphase and its percolation in determination of different $\tan\delta$ trends.

Conclusion

In this paper, we demonstrated a data driven and deep learning approach for modeling structure–property relationship for polymer nanocomposites. The analysis starts from analyzing archived experimental data from NanoMine, which motivates further analysis using computation models including microstructure reconstruction methods and finite element simulations. The first part of this paper focuses on qualitative relationship between microstructure descriptors and mechanical properties, resulting into new findings about the interplay of interphase, volume fraction and dispersion. By conducting an array of controlled simulations, it was demonstrated that the bulk property of nanocomposites is determined by confluent impacts of microstructure type, interphase, loading condition and dispersion. Specifically, adding interphase to the system was found to greatly impact the $\tan\delta$ peak value and cause the behavior of the bulk property to gradually shift from matrix dominant to interphase dominant as the interphase exceeds the percolation threshold. In this analysis, for simplicity, the interphase was assumed to be uniform with fixed thickness, while future work could include more comprehensive representations using gradient interphase and compound effect.[13] In addition, in this paper, $\tan\delta$ peak is chosen as the target property because of its capability of representation damping effect of composite. It is also acknowledged that the area under the $\tan\delta$ curve is another good representation of

the damping and energy absorption and it might be usefully to conduct similar analysis using this or other measures.

The second part of this paper focuses on building quantitative structure–property relationship using deep learning approach. A multi-task convolutional network is proposed to predict mechanical properties of polymer nanocomposites using microstructure images. A computational data set with size 11 000 is generated using finite element simulations. The performance of the model is benchmarked against two other state-of-the-art approaches using two-point statistics and structural geometrical descriptors. The result shows that the proposed deep learning model improves the accuracy for prediction of glass modulus by as much as 45.2%, rubbery modulus by 34.2%, and $\tan\delta$ peak by 19.7%. Additionally, by modifying the inputs to the deep learning model, we have shown that the deep learning model is able to capture physical understandings through learning, which are not explicitly introduced to the model before-hand. The proposed deep learning approach is a feature engineering free, high accuracy and generalization and interpretable model to study the structure–property linkage in polymer nanocomposites. Future study can also work on training the deep learning model using multi-phase images including interphase layers with gradient properties. Such a trained model could be applied to extract other material insights including the impact of gradient interphase layers to help better understand the role of interphase in polymer nanocomposites.

In general, this paper demonstrated an application of NanoMine on understanding structure–property relationship for polymer nanocomposites. Motivated by analyzing the archived Nanomine data and further application of data driven FEA and deep convolutional neural network, both qualitative and quantitative structure–property relationships are explored. This work also demonstrates an applicable path on developing data driven and machine learning methods for material mechanism understanding: starting from analyzing archived data, then moving to application of computational methods for data augmentation, followed by exploring qualitative relationship by restricting confounding parameters, and finally building quantitative relationship using machine learning models. In this paper, three representative viscoelastic properties are chosen as a starting point based on our research interest and availability of computation models for this case study. And given polymer nanocomposites exhibit a wide range of outstanding properties worthy investigation, this strategy can be generalized to understand other material mechanisms and properties as well as guide the design of material with targeted performance.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors gratefully acknowledge support of NSF CSSI (1835677), NSF DMREF 1818574, NSF DIBBS A12761, 1640840, NIST (70NANB14H012 Amd 5) and the CHiMaD center based at Northwestern University.

References

1 L. Schadler, L. Brinson and W. Sawyer, Polymer nanocomposites: a small part of the story, *JOM*, 2007, **59**(3), 53–60.
2 M. Moradi, J. A. Mohandesi and D. F. Haghshenas, Mechanical properties of the poly (vinyl alcohol) based nanocomposites at low content of surfactant wrapped graphene sheets, *Polymer*, 2015, **60**, 207–214.
3 S. M. Hosseini and M. Razzaghi-Kashani, Vulcanization kinetics of nano-silica filled styrene butadiene rubber, *Polymer*, 2014, **55**(24), 6426–6434.
4 S. Mirzaee, S. F. Shayesteh and S. Mahdavifar, Anisotropy investigation of cobalt ferrite nanoparticles embedded in polyvinyl alcohol matrix: a Monte Carlo study, *Polymer*, 2014, **55**(16), 3713–3719.
5 Y. Zare and H. Garmabi, Attempts to simulate the modulus of polymer/carbon nanotube nanocomposites and future trends, *Polym. Rev.*, 2014, **54**(3), 377–400.
6 M. Norouzi, Y. Zare and P. Kiany, Nanoparticles as effective flame retardants for natural and synthetic textile polymers: application, mechanism, and optimization, *Polym. Rev.*, 2015, **55**(3), 531–560.
7 T. Tanaka, *et al.*, Proposal of a multi-core model for polymer nanocomposite dielectrics, *IEEE Trans. Dielectr. Electr. Insul.*, 2005, **12**(4), 669–681.
8 D. Ciprari, K. Jacob and R. Tannenbaum, Characterization of polymer nanocomposite interphase and its impact on mechanical properties, *Macromolecules*, 2006, **39**(19), 6565–6573.
9 Y. Huang, *et al.*, Predicting the breakdown strength and lifetime of nanocomposites using a multi-scale modeling approach, *J. Appl. Phys.*, 2017, **122**(6), 065101.
10 Y. Huang, *et al.*, Modeling of charge transport in nanodielectrics using a coupled finite element and Monte Carlo approach, in *2016 IEEE Conference on Electrical Insulation and Dielectric Phenomena (CEIDP)*, IEEE, 2016.
11 X. Bai, *et al.*, High-fidelity micro-scale modeling of the thermo-visco-plastic behavior of carbon fiber polymer matrix composites, *Composite Structures*, 2015, **134**, 132–141.
12 I. M. Daniel, *et al.*, *Engineering mechanics of composite materials*, Oxford university press, New York, 1994, vol. 3.
13 X. Li, *et al.*, Rethinking Interphase Representations for Modeling Viscoelastic Properties for Polymer Nanocomposites, *Materialia*, 2019, 100277.
14 R. Qiao and L. C. Brinson, Simulation of interphase percolation and gradients in polymer nanocomposites, *Compos. Sci. Technol.*, 2009, **69**(3–4), 491–499.

15 I. Hassinger, *et al.*, Toward the development of a quantitative tool for predicting dispersion of nanocomposites under non-equilibrium processing conditions, *J. Mater. Sci.*, 2016, **51**(9), 4238–4249.
16 Y. Wang, *et al.*, Identifying interphase properties in polymer nanocomposites using adaptive optimization, *Compos. Sci. Technol.*, 2018, **162**, 146–155.
17 G. B. Olson, Computational design of hierarchically structured materials, *Science*, 1997, **277**(5330), 1237–1242.
18 Z. Yang, *et al.*, Deep learning approaches for mining structure-property linkages in high contrast composites from simulation datasets, *Comput. Mater. Sci.*, 2018, **151**, 278–287.
19 M. Bessa, *et al.*, A framework for data-driven analysis of materials under uncertainty: Countering the curse of dimensionality, *Comput Methods Appl Mech Eng*, 2017, **320**, 633–667.
20 H. Zhao, *et al.*, Perspective: NanoMine: A material genome approach for polymer nanocomposites analysis and design, *APL Mater.*, 2016, **4**(5), 053204.
21 H. Zhao, *et al.*, NanoMine schema: An extensible data representation for polymer nanocomposites, *APL Mater.*, 2018, **6**(11), 111108.
22 Y. LeCun, Y. Bengio and G. Hinton, Deep learning, *Nature*, 2015, **521**(7553), 436.
23 Z. Yang, *et al.*, Microstructural materials design via deep adversarial learning methodology, *J. Mech. Des.*, 2018, **140**(11), 111416.
24 X. Li, *et al.*, A transfer learning approach for microstructure reconstruction and structure-property predictions, *Sci. Rep.*, 2018, **8**(1), 1–13.
25 X. Li, *et al.*, A deep adversarial learning methodology for designing microstructural material systems, in *ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, American Society of Mechanical Engineers, 2018.
26 F. Fisher, *et al.*, Spectral response and effective viscoelastic properties of MWNT-reinforced polycarbonate, *Adv. Compos. Lett.*, 2004, **13**(2), 096369350401300201.
27 G. D. Smith, *et al.*, A molecular dynamics simulation study of the viscoelastic properties of polymer nanocomposites, *J. Chem. Phys.*, 2002, **117**(20), 9478–9489.
28 J. S. Smith, D. Bedrov and G. D. Smith, A molecular dynamics simulation study of nanoparticle interactions in a model polymer-nanoparticle composite, *Compos. Sci. Technol.*, 2003, **63**(11), 1599–1605.
29 C. Wei, D. Srivastava and K. Cho, Thermal expansion and diffusion coefficients of carbon nanotube-polymer composites, *Nano Lett.*, 2002, **2**(6), 647–650.
30 C. J. Ellison and J. M. Torkelson, The distribution of glass-transition temperatures in nanoscopically confined glass formers, *Nat. Mater.*, 2003, **2**(10), 695–700.
31 T. Ramanathan, H. Liu and L. C. Brinson, Functionalized SWNT/polymer nanocomposites for dramatic property improvement, *J. Polym. Sci., Part B: Polym. Phys.*, 2005, **43**(17), 2269–2279.
32 A. Eitan, *et al.*, Reinforcement mechanisms in MWCNT-filled polycarbonate, *Compos. Sci. Technol.*, 2006, **66**(9), 1162–1173.
33 A. Cecen, *et al.*, Material structure-property linkages using three-dimensional convolutional neural networks, *Acta Mater.*, 2018, **146**, 76–84.
34 P. Akcora, *et al.*, Anisotropic self-assembly of spherical polymer-grafted nanoparticles, *Nat. Mater.*, 2009, **8**(4), 354–359.
35 G. Munaò, *et al.*, Molecular structure and multi-body potential of mean force in silica-polystyrene nanocomposites, *Nanoscale*, 2018, **10**(46), 21656–21670.
36 S. Yu, *et al.*, Characterization and design of functional quasi-random nanostructured materials using spectral density function, *J. Mech. Des.*, 2017, **139**(7), 071401.
37 A. Iyer, *et al.*, Data-Centric Mixed-Variable Bayesian Optimization For Materials Design, arXiv preprint, 2019, arXiv:1907.02577.
38 W. Chen, *et. al.*, Materials Informatics and Data System for Polymer Nanocomposites Analysis and Design, in *Handbook on Big Data and Machine Learning in the Physical Sciences*, 2019, pp. 65–125.
39 J.-S. Park, Optimal Latin-hypercube designs for computer experiments, *J. Stat. Plan. Inference*, 1994, **39**(1), 95–111.
40 C. D. Wood, *et al.*, Understanding competing mechanisms for glass transition changes in filled elastomers, *Compos. Sci. Technol.*, 2016, **127**, 88–94.
41 Y. Shen, *et al.*, High dielectric performance of polymer composite films induced by a percolating interparticle barrier layer, *Adv. Mater.*, 2007, **19**(10), 1418–1422.
42 S. Ruder, An overview of multi-task learning in deep neural networks, arXiv preprint, 2017, arXiv:1706.05098.
43 Z. Yang, *et al.*, Establishing structure-property localization linkages for elastic deformation of three-dimensional high contrast composites using deep learning approaches, *Acta Mater.*, 2019, **166**, 335–345.
44 Y. Zhang, *et al.*, Microstructure reconstruction and structural equation modeling for computational design of nanodielectrics, *Integr. Mater. Manuf. Innov.*, 2015, **4**(1), 14.
45 N. H. Paulson, *et al.*, Reduced-order structure-property linkages for polycrystalline microstructures based on 2-point statistics, *Acta Mater.*, 2017, **129**, 428–438.
46 A. Gupta, *et al.*, Structure–property linkages using a data science approach: application to a non-metallic inclusion/steel composite system, *Acta Mater.*, 2015, **91**, 239–254.
47 R. Kondo, *et al.*, Microstructure recognition using convolutional neural networks for prediction of ionic conductivity in ceramics, *Acta Mater.*, 2017, **141**, 29–38.