BRUKER

Application Note

#155

Deep Learning to Classify and Establish Structure Property Predictions with PeakForce QNM Atomic Force Microscopy

As microscopists, we often hope – and intuit – that the microstructure we measure correlates somehow to bulk properties, but this correlation is difficult to quantify. Machine learning and specifically, deep learning, is a powerful tool to establish the presence (or absence) of such correlations with its ability to flesh out relationships and trends that are difficult to establish otherwise. This application note discusses the use of deep learning tools, such as convolutional neural nets (CNNs), to explore AFM phase and PeakForce QNM® images of impact copolymers, a polymer blend of polypropylene with micro-sized domains of rubber. First, CNNs are used to successfully classify AFM phase images of a variety of ICPs, where the rubber morphology and distribution varies. The model shows that the PeakForce QNM deformation channel provides the best accuracy in classification. Next, we use a regression-based CNN to correlate the AFM images with bulk mechanical properties, where both phase images and the PeakForce QNM deformation channel were used separately in the model. The results show that the PeakForce QNM data exhibits superior correlation with the bulk mechanical properties, with high accuracy for predicting flexural and Young's modulus, ultimate elongation, and impact toughness.

Machine Learning and AFM

Machine learning (ML), a subset of artificial intelligence, is permeating many sectors, from healthcare to transportation and drug and materials design. The story is the same for the microscopy world, as these new tools are being applied for the improvement of microscope operation and image analysis and interpretation. An example of application of ML tools to microscopy is the use of deep learning (DL) for atom segmentation and localization, noise reduction, and deblurring of STEM images.[1] Another example used DL for high-resolution fluorescence super-resolution microscopy to create a model that can improve low-resolution, low-SNR data.[2] Establishing structure-property relationships has been less common, although a group of researchers led by T. Young-Jin Han did use computer vision and ML techniques to predict compressive strength based on SEM images.[3]

Application of ML toward scanning probe microscopy has been more limited, partly because of the slow acquisition throughput that hampers the large datasets typically required for effective ML models. Some noteworthy examples include the use of ML to enable fast scanning probe microscopy with specific application to piezoresponse force microscopy[4] and the analysis of surface parameters of AFM images of bladder cells.[5]

The focus of this application note is the analysis of AFM images and their subsequent ability to predict bulk mechanical properties. Two kinds of AFM data were used in this study: phase data, which provides primarily the morphology of the material, and PeakForce QNM data, which provides both morphology and quantitative contrast related to mechanical properties, such as modulus, adhesion, and dissipation.

The material under investigation is an impact copolymer (ICP), a common polymer blend used in many consumer applications that require dual properties of stiffness and impact toughness. First, ML models were developed to classify the ICP based solely on its AFM image. Next, more sophisticated models were built to predict a variety of bulk mechanical properties of an ICP based on the AFM image. The PeakForce QNM data provided more accurate classification and prediction of the ICPs, in part due to the additional quantitative mechanical property information included with this AFM mode.

Experimental

All ICP samples were cryomicrotomed; one specimen was provided for each sample.

Phase Imaging

Phase images were collected on a series of five ICP samples (A2, B3, B4, C3, and C5). A total of eighty images were collected on each sample using a Bruker MultiMode® AFM with a variety of stiff probes (k~40N/m) on different days. Only images that were entirely net-repulsive and avoided significant cutting marks were included in the study. All the phase images were collected on a 10 µm x 10 µm length scale with a scan rate of 2 Hz.

PeakForce QNM Imaging

PeakForce QNM images were collected on a series of nine ICP samples (A2, B1, B2, B3, B4, B5, C3, C5, and E4) using a Bruker Icon® AFM. Since PeakForce QNM images often require more time to collect than phase images, it was decided to increase the duty cycle of the PeakForce QNM data acquisition through automation. AutoMET® software was used to program a routine to automatically image three polymer samples, followed by a tip-characterization sample. Five scans were collected on each of the polymer samples with a 4 µm scan size and scan rate of 1 to 1.5 Hz. The tip characterization sample was imaged once for every 15 polymer sample images using ScanAsyst® and a 2.5 µm scan size with a 2:1 aspect ratio. Approximately forty-five images were collected on each of the nine samples.

An RTESPA 150-30 probe with a nominal 5 N/m spring constant and 30 nm probe radius was used for all measurements. Note that this choice of probe is optimized toward the measurement of the rubber properties in the ICP and is not ideal for quantification of the polypropylene properties. A total of three probes were used to collect all the images in this study.

Analysis of the tip characterization samples was done using Bruker software and the tip-qualification feature. This analysis showed that for probes with an initial radius of 30 nm, changes to the tip diameter were minimal over the course of all the measurements, with changes of up to 2 nm (in either direction) recorded.

Some preprocessing of all the AFM images was required. For the AFM images, the phase scale was normalized so that absolute phase degrees information is discarded and only the relative phase difference between the two components is captured. For the

PeakForce QNM data, the deformation and dissipation channels were used "as is" with no additional processing. The DMT modulus channel contained some noise, which was removed prior to its use as an input to the DL model.

Machine Learning

CNNs are popularly used for image classification. CNNs with residual network architectures were used to build both the classification and the regression problems. Appropriate activation functions were used in the final layers of the classification and regression models so that they would predict a class and a value respectively. The set of images collected from AFM were randomly divided into 80% training data, and 20% was kept aside as test set to evaluate the DL model. The data size was relatively low for training purposes. Data augmentation techniques were used to increase the training set by five-fold. Each image was rotated 90 degrees, 180 degrees, and 270 degrees, and it was also flipped upside down, which generated a set of five images from one. A 50-layer-deep residual network (ResNet) was trained to build a classification model. A Tensorflow[6] module with a Keras backend was used to train all the models. The DL model structures were the same for the phase and the PeakForce QNM data. The models were either trained on phase data or PeakForce QNM data sets. While training the models to predict the material properties, the models were trained separately for each of the properties, which resulted in separate models for predicting different properties.

Confusion matrices are an effective tool to evaluate the accuracy of a machine learning to classify test data. In a confusion matrix, the true label is on the horizontal, while the ML model's predicted label is on the vertical axis. The ideal analysis would result in non-zero elements along the diagonal, and all zero off-diagonal elements indicating a 100% accuracy of the model's ability to classify test data correctly.

The models were then evaluated on the hold-out test data set. The classification models were evaluated on the basis of overall accuracy of prediction, and further prediction details were shown on a confusion matrix. The regression models were evaluated with an R2 score. An R2 score measures how much better or worse the model performance is compared to a baseline prediction of the mean of possible outcomes as the prediction every time. An R2 score of 1.0 implies perfect predictions, and the closer the R2 score is to 1.0 the better the model performance is. While building predictive models where outcomes (numeric values) are close to each other (which was the case for the correlation model for predicting material properties), the R2 score is an appropriate measure to evaluate model performance.

Results

A representative phase image of an impact copolymer (ICP) is shown in Figure 1a. There are three components to the ICP material: the yellow background that is comprised of the PP matrix, the dark brown domains of ethylene-propylene rubber, and the lighter ethylene inclusions within the rubber domains. There are many differences between the various ICPs explored in this study including: rubber domain/inclusion size, density, and distribution, and the quantitative mechanical measurements on the rubber domains. Figure 1b shows an analysis of the PeakForce QNM modulus data from three samples (A2, C3, and C5), revealing a significant difference in the DMT moduli between the rubbers in these three ICP samples, where A2 exhibits the lowest value (33 MPa) and C5 is almost 4x higher (130 MPa). This quantitative mechanical information is uniquely available from an AFM-based technique such as PeakForce QNM and can significantly enhance the accuracy of ML models for classification and prediction of bulk properties.

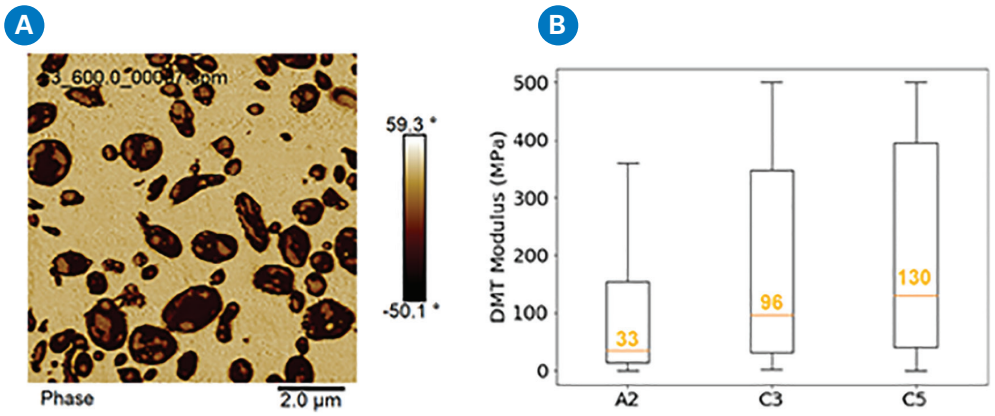

FIGURE 1

A) representative image of an impact copolymer (ICP) and B) distribution of DMT modulus of rubber domains in 3 ICPs (A2, C3, C5).

Classification

The ability of the ML models to classify the various ICPs was evaluated with results shown in confusion matrices (see Figures 2 and 3). First, the ability of the various AFM channels to accurately classify the various materials is shown in Figure 2, where phase data, the PeakForce QNM deformation channel, the PeakForce QNM dissipation channel, and the PeakForce QNM DMT modulus channel are all compared for a set of three samples (A2, C3, and C5). The ML model based on the PeakForce QNM deformation channel provided the best accuracy at 95.8%, followed by similar accuracy by the dissipation channel (87.5%) and DMT modulus channel (86.4%), with the phase data showing the lowest accuracy (78.4%). The model was then trained on a larger sample set comprised of five samples, including B3 and B4 that were challenging to differentiate by eye. Models built on phase data and PeakForce QNM deformation data are compared in Figure 3 for this larger sample set. The ML model based on deformation data accurately classified the five ICPs 65.8% of the time while the phase data exhibited accuracy for 58% of the test data.

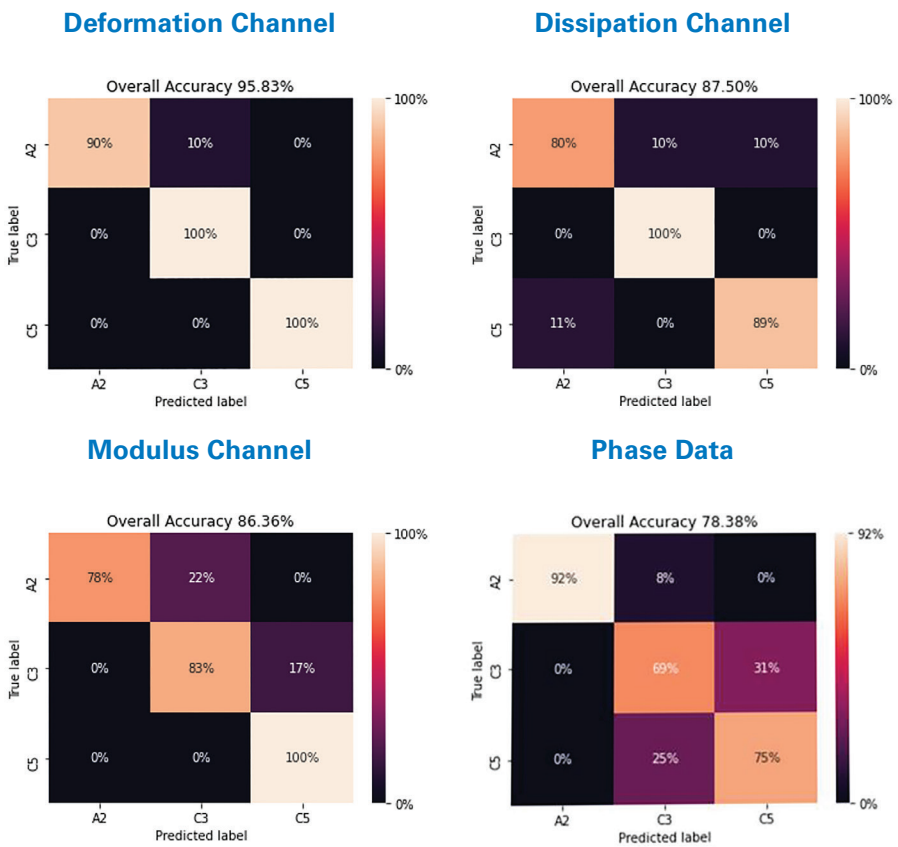

FIGURE 2

Confusion matrices showing results of a CNN classifying three ICPs (A2, C3, C5) using the various AFM data collected in this study: PeakForce QNM deformation, PeakForce QNM dissipation, PeakForce QNM, DMT modulus, and phase.

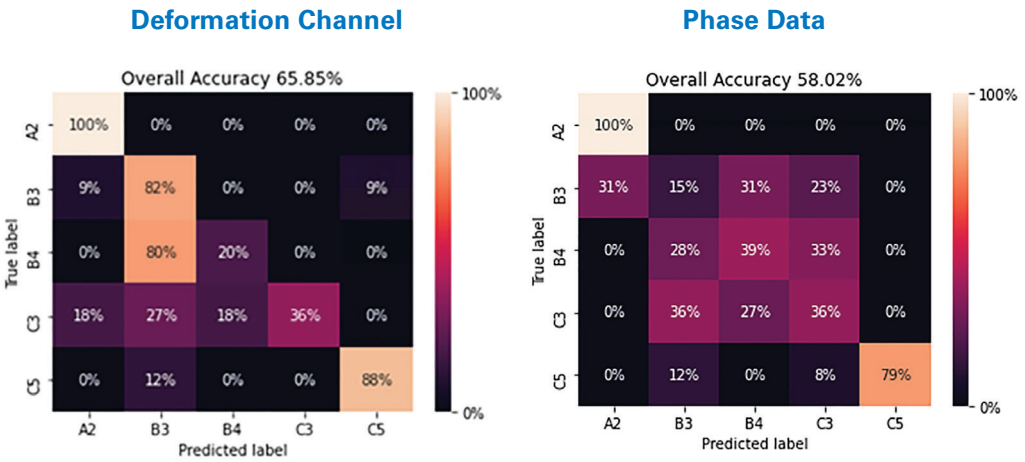

FIGURE 3

Confusion matrices showing results of a CNN on a five-way classification of five ICPs (A2, B3, B4, C3, C5) using the PeakForce QNM deformation channel and the phase data.

Bulk Property Prediction

Models based on phase data and PeakForce QNM deformation data were built and trained in order to predict the bulk mechanical properties of notched izod, flexural modulus, Young's modulus, and ultimate elongation. The results of the DL's prediction on test data for five samples (A2, B3, B4, C3, and C5) is shown in Figure 4, where the purple bars are the bulk property, the green bars are prediction from a model based on PeakForce QNM data, and the orange bars are the prediction from a model based on phase data. As described above, the R2 score is used to quantify the accuracy of the model's prediction for an individual property. As seen in Figure 4e, for all the bulk properties, the R2 score for models built on the deformation channel were higher than that of the phase data. The ultimate elongation (figure 4c) and Young's modulus (figure 4d) had the highest accuracy (both kinds of models), with notched izod (figure 4b) showing the lowest accuracy for both the PeakForce QNM and phase data. The general trend in prediction is similar between phase and PeakForce QNM data with the latter generally showing a better performance.

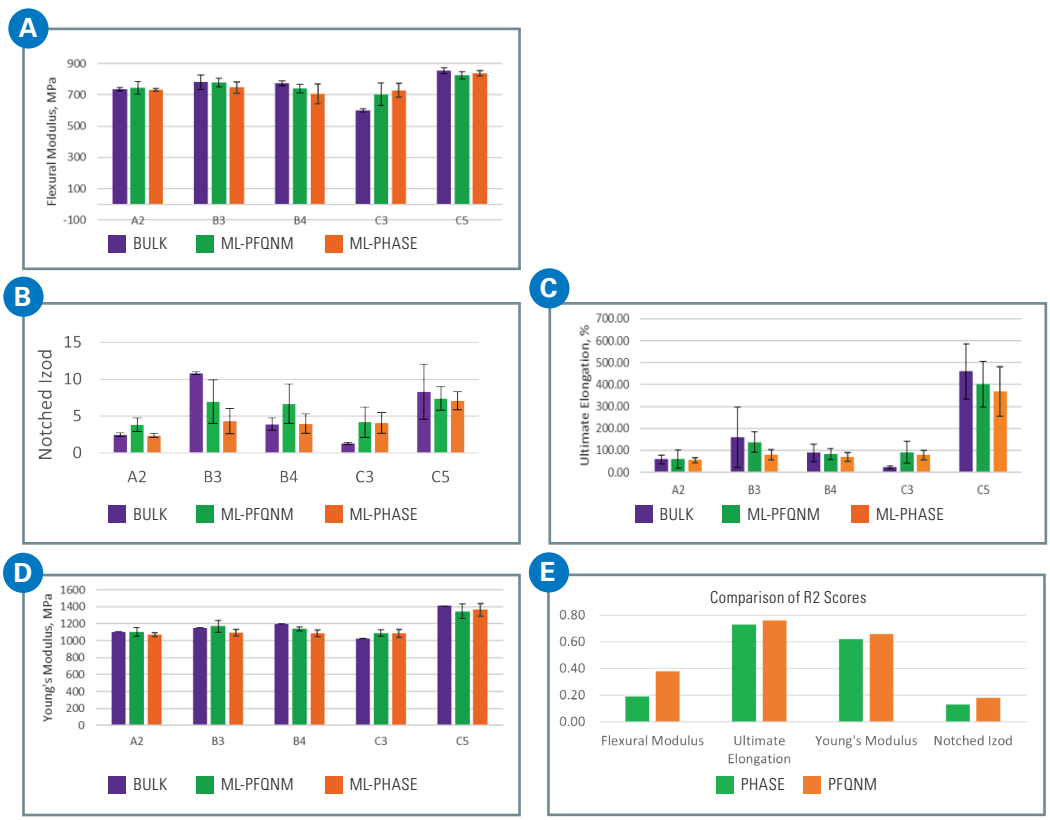

FIGURE 4

Results of a regression-based CNN for a five-sample ICP set to predict various bulk mechanical properties: A) flexural modulus, B) notched izod, C) ultimate elongation, D) Young's modulus, E) comparison of R2 scores for the ML-phase and ML-PFQNM models for the various bulk properties.

Given that the PeakForce QNM deformation channel provided the best accuracy, a ML model based on deformation was trained and tested on an expanded set of nine ICP samples (A2, B1, B2, B3, B4, B5, C3, C5, and E4). The results of the ML model's ability to predict a series of five mechanical properties – notched izod, flexural modulus, Young's modulus, yield strength, and ultimate elongation – are shown in figure 5. Note the addition of a model to predict yield strength is included in this larger set because it provided an appropriately broad spectrum in the yield strength value. The ML model is appropriate only when there is a wide range in the property it is trying to predict. In this larger sample set, the flex modulus and yield strength showed good ability of the ML model to predict the bulk property. Similar to the results in Figure 4, the notched izod was predicted with the lowest accuracy among all the material properties.

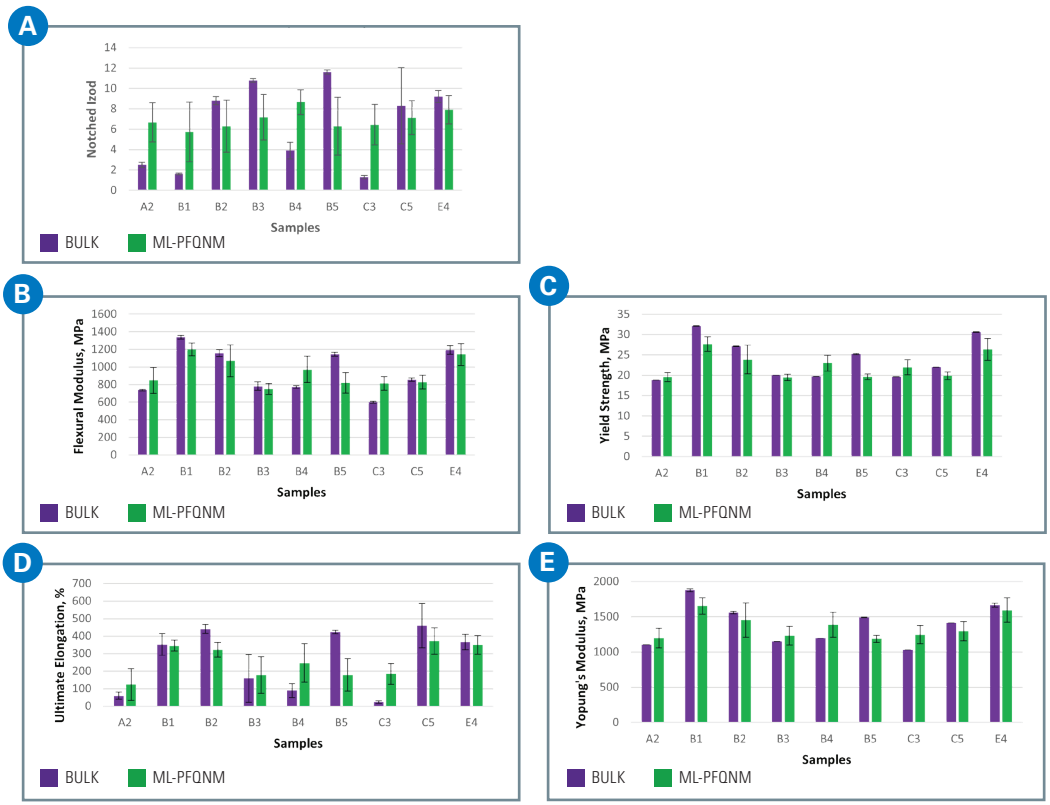

FIGURE 5

Results of a regression-based CNN for a nine-sample ICP set to predict various bulk mechanical properties: A) notched izod, B) flexural modulus, C) yield strength, D) ultimate elongation, E) Young's modulus.

Discussion

There are a few reasons why the deformation channel performed best among all the available AFM data. All the AFM data (phase and PeakForce QNM) are providing morphological information on the rubbers and their inclusions: their size, distribution, and density. The PeakForce QNM data, however, provides additional quantitative mechanical data as well, whether it is the DMT modulus of the rubber, the deformation of the cantilever into the rubber (that is related to its modulus), or the dissipation channel that is related to the energy dissipated by the cantilever into the material.

Among the PeakForce QNM channels, the deformation channel provides some unique advantages. First, it is the cleanest channel in terms of spurious noise relative to the dissipation and DMT modulus. The deformation channel is also less influenced by adhesion than other channels, such as the dissipation channel where small changes in tip chemistry could alter the measurement. Finally, the deformation channel is measured directly from the

force curve on the approach. This contrasts with the dissipation channel (measured between the approach and retract) or the DMT modulus (measured on the retract). Additionally, the DMT modulus error and noise increase if the change in deflection approaches the change in Z position or if there is a viscoelastic component to the deformation. All these reasons contribute to the deformation channel providing the best results for the ML model's ability to classify and predict the bulk properties of the various ICPs.

The bulk mechanical properties predicted in this study measure various elastic and plastic properties. Ultimate elongation is a plastic property that measures the percentage increase in length that occurs in a material before it breaks under tension. Yield strength is on the border of elastic-plastic behavior and is the measured stress at the onset of plastic deformation. The notched Izod test measures impact toughness and is a plastic property. Finally, Young's modulus and flexural modulus are purely elastic properties.

The flexural modulus and Young's modulus are predicted well by the regression-based ML model (see Figure 5), suggesting that the microstructure and quantitative deformation information are correlated to these properties. This is not surprising since the deformation is directly related to the modulus of the material, albeit being measured on the nanoscale in the PeakForce QNM channel. Alternatively, the impact toughness of the material as measured by the notched izod was not well predicted by the models, suggesting that the same information – at least on this 4 µm length scale – was not sufficient to predict impact toughness. This leaves open many other possibilities that can be explored to improve the prediction of impact toughness, such as combining different PeakForce QNM channels, collecting data on different length scales, and collecting quantitative mechanical measurements on the PP matrix, which would require repeating the measurements with a cantilever with a different spring constant that is better matched to the stiffer PP.

Conclusions

Deep learning models have been applied to classify a set of ICP materials and predict their bulk mechanical properties based on various types of AFM data. Both AFM phase data and PeakForce QNM data (DMT modulus, dissipation, and deformation) were used to build CNNs to perform this analysis. The deformation channel provided the best accuracy for the classification measurements on test ICP data for a series of five samples. Based on these results, CNNs built on AFM deformation data predicted the Young's modulus, flexural modulus, yield strength, notched izod, and ultimate elongation of an expanded set of nine ICP samples. The most accurate predictions were for yield strength, flexural modulus, and Young's modulus. Future predictions for other properties can be improved by exploring different length scales, adding additional PeakForce QNM data channels beyond deformation, and obtaining quantitative measurements on the polypropylene matrix material.

References

1. R.Z. Ruoqian Lin, Chunyang Wang, Xiao-Qing Yang, and Huolin L. Xin. "TEMImageNet training library and AtomSegNet deep-learning models for high-precision atom segmentation, localization, denoising, and deblurring of atomic-resolution images." *Scientific Reports* 2021, 11, 5386.

2. F.M. Linjing Fang, Sammy Weiser Novak, Lyndsey Kirk, Cara R. Schiavon, Seungyoon B. Yu, Tong Zhang, Melissa Wu, Kyle Kastner, Alaa Abdel Latif, Zijun Lin, Andrew Shaw, Yoshiyuki Kubota, John Mendenhall, Zhao Zhang, Gulcin Pekkurnaz, Kristen Harris, Jeremy Howard, and Uri Manor. "Deep learning-based point-scanning super-resolution imaging." *Nature Methods* 2021, 18, 406.

3. M.R. Brian Gallagher, Donald Loveland, T. Nathan Mundhenk, Brock Beauchamp, Emily Robertson, Golam G. Jaman, Anna M. Hiszpanski, and T. Yong-Jin Han. "Predicting Compressive Strength of Consolidated Molecular Solids Using Computer Vision and Deep Learning." *Materials & Design* 2020, 190, 108541.

4. K.P. Kelley; M. Ziatdinov, L. Collins, M.A. Susner, R.K. Vasudevan, N. Balke, S.V. Kalinin, and S. Jesse. "Fast Scanning Probe Microscopy via Machine Learning: Non- Rectangular Scans with Compressed Sensing and Gaussian Process Optimization." *Small* 2020.

5. I. Sokolov, M.E. Dokukin, V. Kalaparthi, M. Miljkovic, A. Wang, J.D. Seigne, P. Grivas, and E. Demidenko. "Noninvasive diagnostic imaging using machine-learning analysis of nanoresolution images of cell surfaces: Detection of bladder cancer." *PNAS* 2018, 115, 12920.

6. M. Abadi, et.al. "Tensorflow: A system for large-scale machine learning." *Symposium on Operating Systems Design and Implementation* 2016, 265-283.

Authors

Dalia Yablon (SurfaceChar LLC, Sharon MA);
Ishita Chakraborty (Stress Engineering Services, Houston, TX);
Hillary Passino, Krishnan Iyer, **Antonios Doufas**, and **Maksim Shivokhin**
(ExxonMobil Chemical Company, Baytown TX);
Bede Pittenger and **John Thornton**
(Bruker Nano Surfaces and Metrology, Santa Barbara, CA)

Bruker Nano Surfaces and Metrology Division

Santa Barbara, CA • USA
Phone +1.866.262.4040
productinfo@bruker.com

www.bruker.com/AFM-Imaging