



OPEN

# Locating critical events in AFM force measurements by means of one-dimensional convolutional neural networks

Javier Sotres<sup>1,2✉</sup>, Hannah Boyd<sup>1,2</sup> & Juan F. Gonzalez-Martinez<sup>3</sup>

Atomic Force Microscopy (AFM) force measurements are a powerful tool for the nano-scale characterization of surface properties. However, the analysis of force measurements requires several processing steps. One is locating different type of events e.g., contact point, adhesions and indentations. At present, there is a lack of algorithms that can automate this process in a reliable way for different types of samples. Moreover, because of their stochastic nature, the acquisition and analysis of a high number of force measurements is typically required. This can result in these experiments becoming an overwhelming task if their analysis is not automated. Here, we propose a Machine Learning approach, the use of one-dimensional convolutional neural networks, to locate specific events within AFM force measurements. Specifically, we focus on locating the contact point, a critical step for the accurate quantification of mechanical properties as well as long-range interactions. We validate this approach on force measurements obtained both on hard and soft surfaces. This approach, which could be easily used to also locate other events e.g., indentations and adhesions, has the potential to significantly facilitate and automate the analysis of AFM force measurements and, therefore, the use of this technique by a wider community.

Since its invention<sup>1</sup>, the Atomic Force Microscope (AFM) has not only been revealed as a powerful tool to visualize surfaces at the nanoscale, but also to characterize different sample properties. The most common procedure for these characterizations is based on the acquisition of the so-called normal force curves, a method also known as force spectroscopy, where the force between the AFM probe and the investigated sample is registered at different positions along the sample normal direction<sup>2</sup>. AFM force curves have been widely used to perform a vast variety of e.g., chemical<sup>3,4</sup>, mechanical<sup>5-7</sup> and electrical<sup>8,9</sup> characterizations.

The analysis of force curves requires a number of processing steps, often with the goal to transform the experimental raw data i.e., position sensitive photodetector (PSD) values vs sample vertical positions, into a representation of forces vs probe-sample separations. This then allows, for example, the identification of the forces and separations at which certain events like adhesions and indentations occur, or to fit different sections of the curves to physical laws in order to quantify sample properties. While it is fairly easy to automate some of these required processing steps, for some others the input of AFM experienced users is often required. Moreover, because of the stochastic nature of force measurements and the lateral heterogeneity of most samples, the acquisition and analysis of a high number of force curves is typically required to obtain statistically significant results. This can result in the force spectroscopy experiments becoming an overwhelming task if their analysis are not automatized.

Here, we explored a Machine Learning (ML) approach for one of the most common processing steps required for AFM force curves analysis: the precise location of events of interest. As mentioned, this task can include locating indentation and adhesion events. In this work, we focused instead on locating the contact point. Accurately locating the contact point or point of zero distance is critical for representing a force curve in terms of force vs probe-sample separation. This then allows, for example, to correctly define indentation distances that in turn allow the characterization of mechanical properties of the samples<sup>10</sup>. Correctly determining probe-sample distances is also critical for the analysis of long-range interactions<sup>11</sup>. In the case of hard surfaces this is relatively simple. The contact point is defined as the point where two regions of the curve, which are characterized by

<sup>1</sup>Department of Biomedical Science, Faculty of Health and Society, Malmö University, 20506 Malmö, Sweden. <sup>2</sup>Biofilms-Research Center for Biointerfaces, Malmö University, 20506 Malmö, Sweden. <sup>3</sup>Department of Applied Physics, Technical University of Cartagena, 30202 Cartagena, Spain. ✉email: javier.sotres@mau.se

clearly differentiated slopes, intersect each other. This can be further facilitated in the case that a jump-to-contact event is observed. However, locating the contact point is significantly more difficult in the case of soft surfaces, and even more if they exhibit long-range interactions. In this case, the slope change between the contact and the non-contact regions might not be clearly differentiated and jump-to-contact events are not always observed. Typically, an accurate location of the contact point in these cases requires the attention of an experienced AFM user.

Different approaches have been proposed for the automated detection of the contact point in AFM force curves. In the literature there are two main trends. One focuses on determining the point for which the deflection of the cantilever increases above a threshold value set by the user, typically a multiple of the signal to noise ratio of the non-contact region of the force curve<sup>12,13</sup>. This approach can be complemented with fitting the contact region to different models that include the contact point as a fitting parameter, allowing the contact point to be re-defined in a second step<sup>14</sup>. This strategy has also been combined with additional algorithms to separate the non-linear and linear regions of the contact regime<sup>13</sup>. The second approach focuses on determining the point for which the first<sup>15</sup> or second<sup>16</sup> derivative of the force curve surpass a pre-defined threshold. Additionally, this approach is usually preceded by the prior application of smoothing filters. While powerful and well-tested, these approaches require the values of different parameters to be pre-defined e.g., window size for calculating the signal-to-noise ratio or the curve derivatives, threshold values and ratios of the total curve length corresponding to the non-contact region. However, the parameter values that allow to accurately locate the point of interest depend on the curve shape. This will in turn depend on specifics of the investigated sample e.g., mechanical properties, presence of indentation events and long-range interactions, as well as on several acquisition parameters e.g., ramp length of the non-contact region, maximum applied load, sampling frequency and number of averages per sample vertical position. Here, we explore a different approach to this problem: the use of ML algorithms, specifically one-dimensional (1D) convolutional neural networks (CNNs). Based on how neural networks excel in identifying patterns in heterogeneous data samples, we hypothesized that they would show better generalization i.e., an ability to accurately locate the contact point without the need to search for optimal values for parameters of the event location algorithm for each set of curves obtained under different environmental or acquisition conditions.

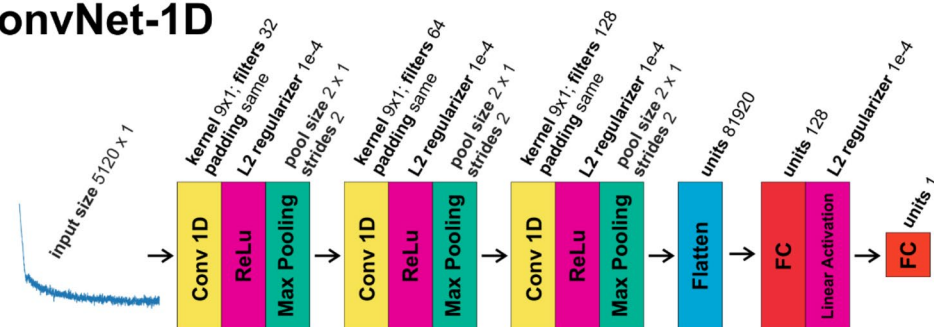
ML techniques have already started to be applied in the field of AFM, even though to a lower extent than in other microscopies. This includes works on the automation of AFM operation<sup>17–21</sup> but mostly on AFM data analysis, both in the context of image analysis<sup>22–29</sup> and force curves analysis<sup>30–32</sup>. However, in the context of force curves analysis the goal of ML techniques has mostly been to classify the curves e.g., according to the type of adhesion exhibited. Here, we were interested in inference instead, specifically in using these techniques to locate the position of critical events, with a specific focus on the contact point. As mentioned, for this goal we propose the use of 1D CNNs. Convolutional neural networks were developed<sup>33</sup> with the main goal of classifying two-dimensional (2D) data i.e., images. Since then, they have shown an extraordinary performance in this field<sup>34</sup>. Moreover, 2D CNNs have also been used for inference purposes. An example of relevance for this work being the location of specific objects within images<sup>35–38</sup>. Prompted by the success of 2D CNNs, 1D CNNs were developed for the context of analysis of 1D signals and rapidly achieved a high performance in a diversity of applications<sup>39</sup>. Here, we explored the performance of two different 1D CNN architectures to locate the contact point in force measurements performed in two different systems: (i) a hard surface in the presence of long-range forces (mica in aqueous solutions) and (ii) a soft surface in the presence of long-range forces (salivary pellicles in aqueous solutions).

## Results and discussion

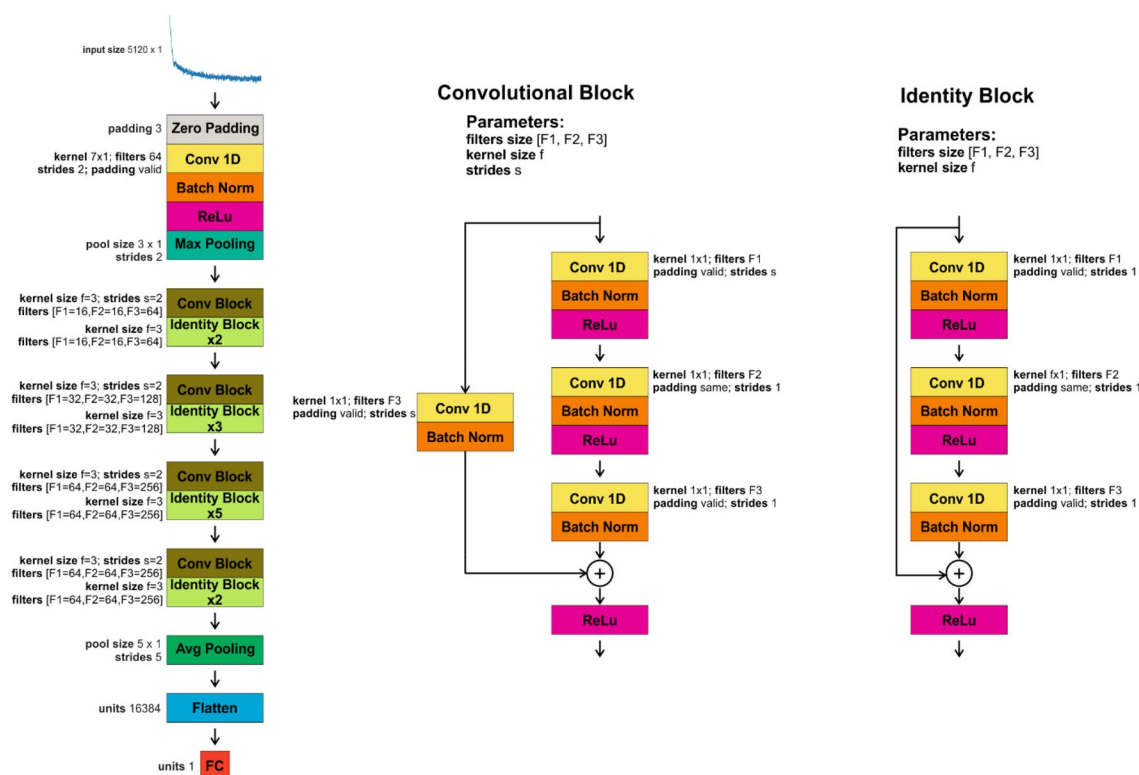
Since the development of convolutional neural networks, one of the most common basic architectures consisted of combining blocks containing convolutional layers followed by sub-sampling layers in the beginning of the network, these blocks then are followed by fully connected layers<sup>33,40</sup>. This type of architecture has also been proved successful in the analysis of 1D data just by using one-dimension convolutional layers rather than two-dimensional ones<sup>41–44</sup>. Subsequently, we also explored this architecture. Specifically, we report results obtained using the architecture depicted in Fig. 1a, from now on named ConvNet-1D. It consists in three blocks of 1D convolutional layers and sub-sampling (max-pooling) layers, followed by two fully connected layers (specific layers parameters are provided in Fig. 1a), the last one consisting of only one unit accounting for the predicted position of the critical event to be located, in this case the contact point within the experimental force curves.

We first investigated the performance of ConvNet-1D to locate the contact point in force curves performed with Si<sub>3</sub>N<sub>4</sub> tips on mica in aqueous solutions. Specifically, we experimentally collected a set of 1806 force curves obtained in 100 mM, 10 mM and 1 mM NaCl solutions using three different probe/sample pairs for each type of solution. Representative examples are shown in Fig. 2a,b. These force curves exhibit long range electrostatic forces when interacting at neutral pH values, as a result of both sample and tip being negatively charged<sup>8</sup>. At very short separations, the attractive van der Waals forces overcome the repulsive electrostatic force, often resulting in a mechanical instability known as jump-to-contact<sup>2</sup>. From this point, mechanical contact is established and the cantilever deflection increases approximately with the same rate as the sample displacement considering that mica surfaces (effective stiffness as measured by AFM  $\sim 82 \text{ N m}^{-1}$ )<sup>45</sup> are significantly harder than the employed cantilevers (nominal force constant  $0.1 \text{ N m}^{-1}$ ). First, we manually labelled the contact point in all collected force curves. Then, we randomly divided them into a training (70% of the curves) and a validation (30% of the curves) data sets. The curves shown in Fig. 2a,b correspond to the validation data set. Neural networks require a fixed input data size, and force curves are usually collected at a variety of sizes/points. This was also the case for those used in this work. To account for this, all curves were interpolated up to a total size of 5120 points, a value higher than the size of any of the collected experimental curves. Additionally, before feeding the curves to the networks, they were normalized both in the vertical (PSD values) and horizontal (sample displacement) axis.

## a) ConvNet-1D



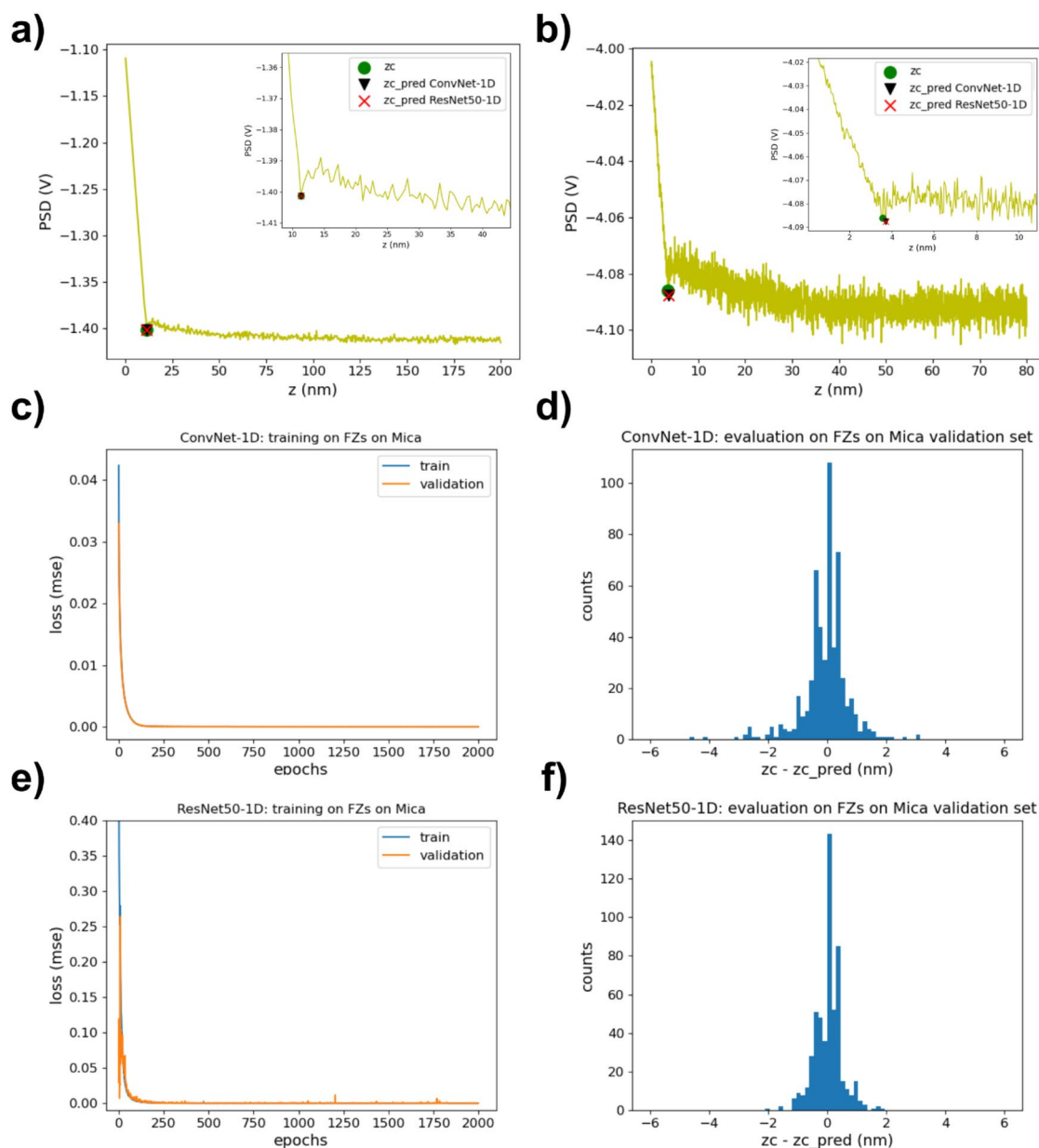
## b) ResNet50-1D



**Figure 1.** Convolutional neural networks used in this study: (a) ConvNet-1D and (b) ResNet-1D.

The output of the network will then be the normalized sample position of the contact point. Then, we trained the ConvNet-1D network on the training set using the mean square error (mse) loss function. The evolution of the mse loss while training the network over 2000 epochs is shown in Fig. 2c for both the training and validation data sets. No signs of exploding gradients, vanishing gradients or overfitting were observed. However, for avoiding overfitting we needed to incorporate L2 regularization in the different layers of the network. We then chose the weights obtained for the epoch with the minimum loss. As a quantity to validate the performance of the network, we used the difference between the sample positions of the hand-labelled,  $z_c$ , and the predicted,  $z_{c\_pred}$ , contact points (recalculated for the non-normalized force curves). The distribution of this quantity for the validation data set is shown in Fig. 2d. The mean and standard deviation values for this distribution were  $0.06 \pm 0.78$  nm (Table 1). Overall, we can see how a fairly simple and shallow architecture as that of ConvNet-1D provided excellent accuracy.

We then investigated the performance of ConvNet-1D for locating the contact point on force curves performed on salivary pellicles with  $\text{Si}_3\text{N}_4$  tips in aqueous NaCl solutions. Salivary pellicles are nm-thick proteinaceous films that rapidly form on almost any surface when exposed to saliva. They do not only confer oral surfaces with extraordinary protection against chemical and mechanical insults<sup>46,47</sup>. Salivary pellicles also constitute excellent aqueous boundary lubricants with the ability to provide low friction coefficients and high yield strengths to a wide variety of tribological contacts<sup>48–51</sup>. Several experimental results point towards a two-layer model for



**Figure 2.** (a) and (b) Representative examples of force curves obtained with  $\text{Si}_3\text{N}_4$  AFM tips on mica surfaces in NaCl aqueous solutions. The insets correspond to zooms of the contact point region. The curves also contain markers for the labelled contact point (green circle) as well as the contact points predicted by ConvNet-1D (black triangle) and by ResNet50-1D (red cross). (c) Evolution of the mse loss function i.e., the mean squared differences between labelled and ConvNet-1D predicted normalized sample positions corresponding to the contact point, obtained while training ConvNet-1D on force curves on mica. Loss values for both the training and validation sets are shown. Total training time = 1096 s ( $\sim 0.55$  s/epoch). (d) Distribution of the difference between the sample positions of the hand-labelled,  $z_c$ , and the ConvNet-1D predicted,  $z_{c\_pred}$ , contact points (recalculated for the non-normalized force curves). (e) Loss (mse) values for the same training and validation sets obtained in this case while training the ResNet50-1D network. Total training time = 4366 s ( $\sim 2.18$  s/epoch). (f)  $z_c - z_{c\_pred}$  distribution for the force curves on mica validation set as predicted by ResNet50-1D.

salivary pellicles reconstituted at solid–liquid interfaces, with an inner thin dense layer, mainly formed by low molecular weight proteins, and an outer thick diffuse layer mainly formed by the long glycoproteins known as mucins<sup>46</sup>. AFM force measurements are one of the main techniques that contributed to elucidate the structure of salivary pellicles<sup>11,52</sup>. AFM force measurements have also been used to determine how this structure is affected by oral health products like surfactants<sup>53</sup>. Force curves performed on salivary pellicles, for which representative examples are shown in Fig. 3a,b, resemble up to some extent those obtained on mica surfaces (Fig. 2a,b). When approaching probe and sample, a long range interaction with an exponential-like dependence with distance is initially observed. This is originated by the repulsive steric interaction that develops between the AFM probe and the diffuse pellicle outer layer. At shorter separations, a transition to a different regime is observed where the

	zc-zc_pred (nm)	
	ConvNet-1D	ResNet50-1D
Mica—validation set (NaCl 100 mM/10 mM/1 mM)	0.06 ± 0.78	0.03 ± 0.46
Pellicle on Mica—validation set (NaCl 100 mM/10 mM)	0.04 ± 1.19	0.01 ± 0.71
Pellicle on Mica—test set (NaCl 1 mM)	1.20 ± 1.80	0.20 ± 0.97

**Table 1.** Mean and standard deviation values for the difference between the sample positions of the labelled and predicted (by the two networks used in this work) contact points for all validation and test data sets.

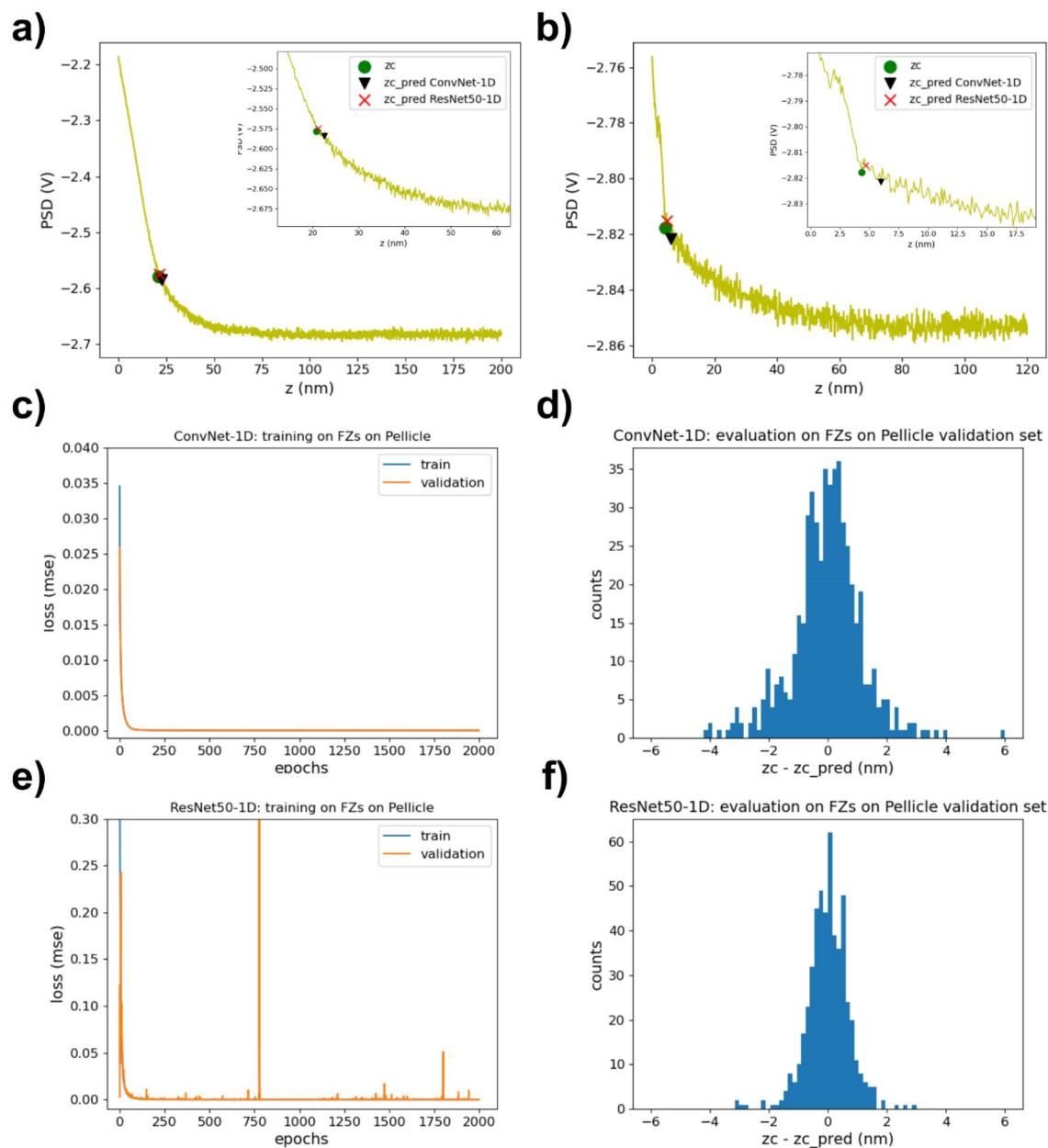
cantilever deflection exhibits a steeper dependence with the separation from the probe. This regime is typically well-fitted by mechanical contact models. In this system, the contact point delimits the transition between these two different regimes, and its accurate determination is critical for a correct interpretation of the mechanical properties of the pellicle and, therefore, of its structure. However, the position of the contact point in this system is not as clear as in the case of force curves performed on mica. Often, the slope difference between the steric and the mechanical contact regime is not clearly differentiated because of the sample softness. The jump-to-contact event is not as prominent either, or even sometimes absent. The location of the contact point can be facilitated though by the fact that the deflection noise typically decreases when entering the contact region<sup>54</sup>. Overall, the location of the contact point in this case is fairly difficult to determine by means of hard coded algorithms, and typically requires that an expert user analyzes each of the registered force curves.

We initially tested the performance of ConvNet-1D for locating the contact point on a set of force curves on salivary pellicles obtained both in NaCl 100 mM and NaCl 10 mM aqueous solutions, using three different probe/sample pairs for each type of solution. Specifically, we experimentally collected a total of 2584 force curves for this system. We then manually labelled the contact point in all curves and randomly divided them into a training (80% of the curves) and a validation (20% of the curves) sets (the ratio of the curves used for validation was lower than that used for the dataset of curves performed on mica surfaces as the total number of curves was higher in the current case). The curves in Fig. 3a,b correspond to the validation set. Before feeding the curves to the network, they were interpolated and normalized as previously described. The network was trained again over 2000 epochs. The mse loss for both the train and validation data sets during training is shown in Fig. 3c. As in the case of force curves obtained on mica, no signs of exploding gradients, vanishing gradients or overfitting were observed. Here, we also chose for prediction purposes the weights obtained for the epoch with the minimum validation loss. The distribution of the difference between the sample positions of the labelled and predicted contact points (recalculated for the non-normalized force curves) is presented in Fig. 3d. The mean and standard deviation values for this distribution were  $0.04 \pm 1.19$  nm (Table 1). Overall, our results show again a good performance of ConvNet-1D, even though the distribution of values was slightly wider than that obtained when applying the same network to locate contact points in force curves performed on mica. Of relevance is that ConvNet-1D was able to easily distinguish contact points from indentations, as observed in Fig. 3b.

Subsequently, we investigated the generalization of ConvNet-1D for locating contact points in a set of curves not used for training. For this, we tested the already trained network on a set of 153 force curves corresponding to a different salivary pellicle sample/Si<sub>3</sub>N<sub>4</sub> AFM tip pair than those used for training the network. This test set was also obtained in different conditions i.e., NaCl 1 mM. Representative curves are shown in Fig. 4a,b. The distribution of the difference between the sample positions of the labelled, zc, and ConvNet-1D predicted, zc\_pred, contact points (recalculated for the non-normalized force curves) is presented in Fig. 4c. The mean and standard deviation values for this distribution were  $1.20 \pm 1.80$  nm (Table 1). While still reasonable, it is clear that the performance of ConvNet-1D decreased when applied to data different than that used for training.

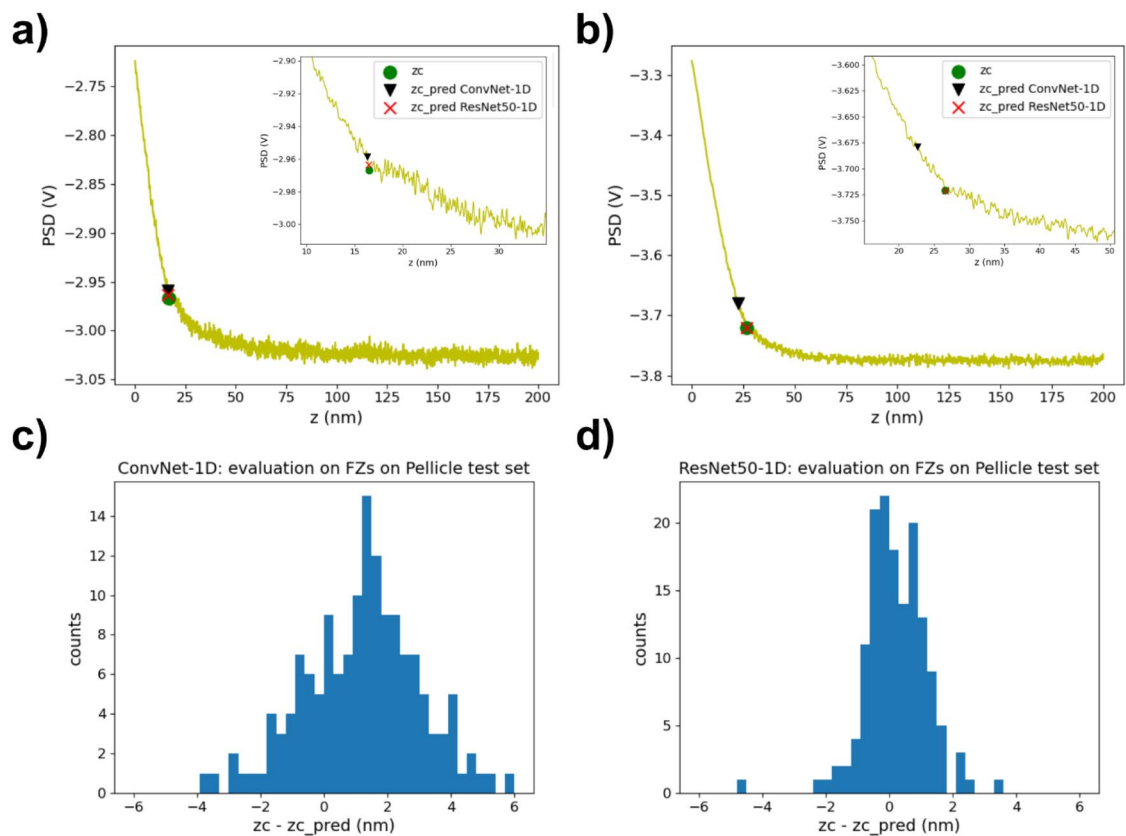
We then explored whether we could improve our results by means of a different network architecture. First, we explored different hyper-parameter values for ConvNet-1D, but we did not find a set of values that significantly improved its performance. Small increments in the network depth did not have a significant effect either. We then investigated the use of significantly deeper networks. Specifically, we focused on residual networks, which contain residual blocks that include short paths (Fig. 1b) with the ability to carry gradients throughout very deep networks mitigating in this way the problem of vanishing gradients<sup>55</sup>. We used as a starting point the ResNet-50 network<sup>56</sup> as implemented in Keras (<sup>57</sup>, <https://t.ly/uLupD>) and modified it in order to use 1D rather than 2D convolutional layers. The specific implemented architecture, from now on named ResNet50-1D, is shown in Fig. 1b. The train and validation mse losses obtained while training ResNet50-1D on force curves on mica and on salivary pellicles are shown in Figs. 2e and 3e respectively. The distribution of the difference between the sample positions of the hand-labelled and predicted contact points (recalculated for the non-normalized force curves) is presented in Figs. 2f and 3f for the validation sets of force curves performed on mica and salivary pellicles respectively. The mean and standard deviation values for these distributions were  $0.03 \pm 0.46$  nm for the mica validation set and  $0.01 \pm 0.71$  nm for the pellicle validation set (Table 1). Therefore, already on these sets ResNet50-1D exhibited a slightly better performance than ConvNet-1D, with narrower zc-zc\_pred distributions. We subsequently tested the performance of the already trained ResNet50-1D network on the test set consisting of force curves on salivary pellicles in NaCl 1 mM not used for training the network. The corresponding zc-zc\_pred distribution is shown in Fig. 4d. The mean and standard deviation values for this distribution was  $0.20 \pm 0.97$  nm (Table 1) i.e., the width of the distribution was approximately half than that obtained with ConvNet-1D.

Overall, our results prove that 1D CNNs can be used for accurately locating the contact point in AFM force curves. This was quantified by comparing the hand-labelled and the predicted contact point positions, zc-zc\_pred,



**Figure 3.** (a) and (b) Representative examples of force curves obtained with  $\text{Si}_3\text{N}_4$  AFM tips on salivary pellicles in NaCl (100 mM and 10 mM) aqueous solutions. The insets correspond to zooms of the contact point region. The curves also contain markers for the labelled contact point (green circle) as well as the contact points predicted by ConvNet-1D (black triangle) and by ResNet50-1D (red cross). (c) Evolution of the loss mse function while training ConvNet-1D on force curves on pellicles. Loss values for both the training and validation sets are shown. Total training time = 1615 s ( $\sim 0.81$  s/epoch). (d) Distribution of the difference between the sample positions of the hand-labelled,  $z_c$ , and the ConvNet-1D predicted,  $z_{c\_pred}$ , contact points (recalculated for the non-normalized force curves). (e) Loss (mse) values for the same training and validation sets obtained while training the ResNet50-1D network on the same data set. Total training time = 6760 s ( $\sim 3.38$  s/epoch). (f)  $z_c - z_{c\_pred}$  distribution for the force curves on the pellicles validation set using the values predicted by ResNet50-1D.

which was below 1 nm in most cases (Table 1). Common 1D CNNs networks with a relative simple and shallow architecture, like ConvNet-1D, can already accomplish this task with good accuracy. However, we also showed that deeper networks that make use of residual blocks, like ResNet50-1D, result in a more accurate location of this event as well as better generalization i.e., performance when used on data from a different set than that used for training. Obviously, this improved performance comes at a cost in terms of time. As reported, the time needed for training the networks was a factor  $\sim 4$  higher for ResNet50-1D ( $\sim 2.18$  s/epoch on the mica train & validation datasets;  $\sim 3.38$  s/epoch on the pellicle train & validation datasets) than for ConvNet-1D ( $\sim 0.55$  s/epoch on the



**Figure 4.** (a) and (b) Representative examples from the test set of force curves obtained with  $\text{Si}_3\text{N}_4$  AFM tips on salivary pellicles in NaCl 1 mM aqueous solutions. The insets correspond to zooms of the contact point region. The curves also contain markers for the labelled contact point (green circle) as well as the contact points predicted by ConvNet-1D (black triangle) and by ResNet50-1D (red cross). (c) and (d)  $z_c - z_{c\_pred}$  distributions for this test data set, obtained with the ConvNet-1D and ResNet50-1D networks respectively. The total inference time when using ConvNet-1D was 1.3 s ( $\sim 8.5$  ms/curve) whereas for ResNet50-1D it was 2.3 ms ( $\sim 15$  ms/curve).

mica training & validation datasets;  $\sim 0.81$  s/epoch on the pellicle training & validation datasets). Nevertheless, this increase in training time should not be a major limitation for choosing deeper networks (as seen in Figs. 2 and 3, for all datasets the losses for both models converged to an almost constant value after  $\sim 250$  epochs) unless one is dealing with very large datasets or only has access to low speed processors for training the networks. Inference time should also not be a major limitation for choosing ResNet50-1D like architectures ( $\sim 15$  ms/curve) over ConvNet-1D like ones ( $\sim 8.5$  ms/curve). It is worth highlighting the improvement this represents in comparison to a hand-labeling approach, where the user would spend in the best of cases few seconds on each curve.

The inference time is similar though to that achieved by other event locating algorithms. For instance, Offroy and co-workers<sup>13</sup> reported ( $\sim 23$  ms/curve) for their algorithm based on an increase of the cantilever deflection above the value of the signal-to-noise ratio of the non-contact region of the curve, or even probably slightly higher taking into account that they used a standard PC and that their algorithm also calculated the border between the linear and non-linear regimes of the contact region. However, it is also worth highlighting the generalization capability of the proposed ML-based approach. As stated in the Methods section, the datasets used in this work contained force curves obtained while randomly varying acquisition parameters like the ramp length, the maximum applied load, the sampling frequency and the ramp velocity. This heterogeneity is exemplified by the differences exhibited between the curves plotted in the figures of the manuscript for each dataset. Moreover, our datasets combined curves obtained in different environments (ionic strength), which result in different long-range interactions. Interestingly, our results show that 1D CNNs could accurately locate the contact point despite these differences, even in the presence of unexpected events like indentations (Fig. 3b). In contrast with previous proposed approaches, this could be achieved without pre-defining values for parameters of the event location algorithm.

The specific focus of this work was on locating the contact point. However, we see no limitations in using the same approach to locate other events in AFM force curves e.g., indentations and adhesions. It should also be possible to use architectures that would allow locating a variable number of different type of events within the same AFM curve by using similar approaches as the available object detector networks for 2D data. Overall, we foresee that these ML-based approaches will significantly facilitate AFM data analysis and, therefore, the use of this technique by a wider community.

## Methods

**Materials.** All water used to prepare salt solutions was of ultrahigh quality (UHQ), processed in an Elgastat UHQ II apparatus (Elga Ltd., High Wycombe, Bucks, England). Mica sheets were purchased from Electron Microscopy Sciences (Hatfield, PA).

For saliva experiments, unstimulated human whole saliva (HWS) was collected from two healthy donors, after they provided written informed consent, using the protocol described in<sup>58</sup>, pooled and then used immediately. Ethical approval was obtained from the Swedish Ethical Review Authority through the committee of research ethics at Lund University (2018/42), and saliva collection was performed following the protocol in the mentioned ethical permit, which was in accordance with the Declaration of Helsinki.

Unless otherwise specified, all other chemicals used were of at least analytical grade.

**Force spectroscopy.** A commercial Atomic Force Microscope (AFM) setup equipped with a liquid cell (MultiMode 8 SPM with a NanoScope V control unit, Bruker AXS, Santa Barbara CA) was employed for the acquisition of force curves. Rectangular silicon nitride levers with a nominal normal spring constant of  $0.1 \text{ N m}^{-1}$  were employed in all the experiments (OMLC-RC800PSA, Olympus, Japan). The choice of cantilevers with this spring constant was based on our previous experience, as they exhibit significant deflection values that allow investigating electrostatic forces on mica<sup>8</sup> as well as steric forces and deformation of salivary pellicles<sup>11,53</sup>. Choosing cantilevers with significantly higher spring constants would have resulted in a negligible influence of long-range forces or mechanical deformation on the measured cantilever deflection, in which case the location of the contact point becomes a straightforward problem to solve. Before every experiment, tips were rubbed against a clean freshly cleaved mica surface, a procedure that also leads to the removal of asperities and achieves a smooth tip surface<sup>59</sup>.

Force spectroscopy on mica experiments were performed on freshly cleaved mica sheets. For force spectroscopy on salivary pellicles, freshly cleaved mica sheets ( $\sim 1 \text{ cm}^2$ ) were drop-coated with 100  $\mu\text{L}$  of HWS, and subsequently rinsed with NaCl 100 mM after 1 h. Samples were then immediately placed on the AFM, ensuring they did not dry at any time. Force curves were then obtained while randomly varying acquisition parameters like the ramp length, the maximum applied load, the sampling frequency and the ramp velocity.

**ML analysis of force curves.** The contact point of all force curves was manually labelled with a custom made application available at <https://t.ly/Y0hC>.

Before feeding the force curves to the neural networks, they were interpolated to a final size of 5120 points, and subsequently normalized in the vertical and normal directions. The labelled contact point was also encoded by the normalized sample position at which it occurred,

The 1D CNNs architectures depicted in Fig. 1 were implemented using Keras<sup>57</sup>. The specific implementation is available in the repository of this work: <https://t.ly/MkVf>. For training, we used mse loss and Adam optimization with an initial learning rate of  $10^{-4}$ , a 1st momentum of 0.9 and a 2nd momentum of 0.999. All models were trained and validated on a system with the following characteristics: Intel Core i9 7980XE (18 cores, 36 threads) CPU, 64 GB of DDR4 RAM and a GPU NVIDIA 2080 RTX Ti 11 GB GDDR6 (4352 computation cores). Models were always trained and evaluated using the GPU.

## Data availability

All used force curves and code can be found in the repository of this work, available at: <https://t.ly/MkVf>.

Received: 6 April 2022; Accepted: 20 July 2022

Published online: 29 July 2022

## References

- Binnig, G., Quate, C. F. & Gerber, C. Atomic force microscope. *Phys. Rev. Lett.* **56**, 930–933. <https://doi.org/10.1103/PhysRevLett.56.930> (1986).
- Butt, H.-J., Cappella, B. & Kappl, M. Force measurements with the atomic force microscope: Technique, interpretation and applications. *Surf. Sci. Rep.* **59**, 1–152. <https://doi.org/10.1016/j.surfrep.2005.08.003> (2005).
- Hinterdorfer, P. & Dufrène, Y. F. Detection and localization of single molecular recognition events using atomic force microscopy. *Nat. Methods* **3**, 347–355. <https://doi.org/10.1038/nmeth871> (2006).
- Noy, A. Chemical force microscopy of chemical and biological interactions. *Surf. Interface Anal.* **38**, 1429–1441. <https://doi.org/10.1002/sia.2374> (2006).
- Garcia-Manyes, S. & Sanz, F. Nanomechanics of lipid bilayers by force spectroscopy with AFM: A perspective. *Biochim. Biophys. Acta Biomembr.* **1798**, 741–749. <https://doi.org/10.1016/j.bbmem.2009.12.019> (2010).
- Kuznetsova, T. G., Starodubtseva, M. N., Yegorenkov, N. I., Chizhik, S. A. & Zhdanov, R. I. Atomic force microscopy probing of cell elasticity. *Micron* **38**, 824–833. <https://doi.org/10.1016/j.micron.2007.06.011> (2007).
- Sotres, J., Jankovskaja, S., Wannerberger, K. & Arnebrant, T. Ex-vivo force spectroscopy of intestinal mucosa reveals the mechanical properties of mucus blankets. *Sci. Rep.* **7**, 7270. <https://doi.org/10.1038/s41598-017-07552-7> (2017).
- Sotres, J. & Baró, A. M. AFM imaging and analysis of electrostatic double layer forces on single DNA molecules. *Biophys. J.* **98**, 1995–2004. <https://doi.org/10.1016/j.bpj.2009.12.4330> (2010).
- Zhao, H. *et al.* Probing surface charge potentials of clay basal planes and edges by direct force measurements. *Langmuir* **24**, 12899–12910. <https://doi.org/10.1021/la802112h> (2008).
- Parra, A., Casero, E., Lorenzo, E., Pariente, F. & Vázquez, L. Nanomechanical properties of globular proteins: Lactate oxidase. *Langmuir* **23**, 2747–2754. <https://doi.org/10.1021/la062864p> (2007).
- Boyd, H. *et al.* A comparison between the structures of reconstituted salivary pellicles and oral mucin (MUC5B) films. *J. Colloid Interface Sci.* **584**, 660–668. <https://doi.org/10.1016/j.jcis.2020.10.124> (2021).
- Monclus, M. A., Young, T. J. & Di Maio, D. AFM indentation method used for elastic modulus characterization of interfaces and thin layers. *J. Mater. Sci.* **45**, 3190–3197. <https://doi.org/10.1007/s10853-010-4326-6> (2010).



13. Offroy, M., Razafitianamaharavo, A., Beaussart, A., Pagnout, C. & Duval, J. F. L. Fast automated processing of AFM PeakForce curves to evaluate spatially resolved Young modulus and stiffness of turgescent cells. *RSC Adv.* **10**, 19258–19275. <https://doi.org/10.1039/D0RA00669F> (2020).
14. Chang, Y.-R. *et al.* Automated AFM force curve analysis for determining elastic modulus of biomaterials and biological samples. *J. Mech. Behav. Biomed. Mater.* **37**, 209–218. <https://doi.org/10.1016/j.jmbbm.2014.05.027> (2014).
15. Benítez, R., Moreno-flores, S., Bolós, V. J. & Toca-Herrera, J. L. A new automatic contact point detection algorithm for AFM force curves. *Microsc. Res. Tech.* **76**, 870–876. <https://doi.org/10.1002/jemt.22241> (2013).
16. Nyland, L. R. & Maughan, D. W. Morphology and transverse stiffness of drosophila myofibrils measured by atomic force microscopy. *Biophys. J.* **78**, 1490–1497. [https://doi.org/10.1016/S0006-3495\(00\)76702-6](https://doi.org/10.1016/S0006-3495(00)76702-6) (2000).
17. Huang, B., Li, Z. & Li, J. An artificial intelligence atomic force microscope enabled by machine learning. *Nanoscale* **10**, 21320–21326. <https://doi.org/10.1039/C8NR06734A> (2018).
18. Liu, Y., Huang, C., Chen, H. & Fu, L. in *2019 12th Asian Control Conference (ASCC)*. 138–143.
19. Liu, Y. *et al.* Experimental discovery of structure-property relationships in ferroelectric materials via active learning. [arXiv:2108.06037v2](https://arxiv.org/abs/2108.06037v2) [cond-mat.mtrl-sci]. <https://doi.org/10.48550/arXiv.2108.06037> (2021).
20. Rashidi, M. & Wolkow, R. A. Autonomous scanning probe microscopy in situ tip conditioning through machine learning. *ACS Nano* **12**, 5185–5189. <https://doi.org/10.1021/acsnano.8b02208> (2018).
21. Sotres, J., Boyd, H. & Gonzalez-Martinez, J. F. Enabling autonomous scanning probe microscopy imaging of single molecules with deep learning. *Nanoscale* **13**, 9193–9203. <https://doi.org/10.1039/D1NR01109J> (2021).
22. Bai, H. & Wu, S. Nanowire detection in AFM images using deep learning. *Microsc. Microanal.* <https://doi.org/10.1017/s143192762002468x> (2020).
23. Checa, M., Millan-Solsona, R., Mares, A. G., Pujals, S. & Gomila, G. Fast label-free nanoscale composition mapping of eukaryotic cells via scanning dielectric force volume microscopy and machine learning. *Small Methods* **5**, 2100279. <https://doi.org/10.1002/smt.202100279> (2021).
24. Gordon, O. *et al.* Scanning tunneling state recognition with multi-class neural network ensembles. *Rev. Sci. Instr.* **90**, 103704. <https://doi.org/10.1063/1.5099590> (2019).
25. Gordon, O. M., Hodgkinson, J. E. A., Farley, S. M., Hunsicker, E. L. & Moriarty, P. J. Automated searching and identification of self-organized nanostructures. *Nano Lett.* **20**, 7688–7693. <https://doi.org/10.1021/acs.nanolett.0c03213> (2020).
26. Gordon, O. M., Junqueira, F. L. Q. & Moriarty, P. J. Embedding human heuristics in machine-learning-enabled probe microscopy. *Mach. Learn. Sci. Technol.* **1**, 015001. <https://doi.org/10.1088/2632-2153/ab42ec> (2020).
27. Kalinin, S. V. *et al.* Disentangling rotational dynamics and ordering transitions in a system of self-organizing protein nanorods via rotationally invariant latent representations. *ACS Nano* **15**, 6471–6480. <https://doi.org/10.1021/acsnano.0c08914> (2021).
28. Liu, Y. *et al.* General resolution enhancement method in atomic force microscopy using deep learning. *Adv. Theory Simul.* **2**, 1800137. <https://doi.org/10.1002/adts.201800137> (2019).
29. Sokolov, I. *et al.* Noninvasive diagnostic imaging using machine-learning analysis of nanoresolution images of cell surfaces: Detection of bladder cancer. *Proc. Natl. Acad. Sci. USA* **115**, 12920–12925. <https://doi.org/10.1073/pnas.1816459115> (2018).
30. Minelli, E. *et al.* A fully-automated neural network analysis of AFM force-distance curves for cancer tissue diagnosis. *Appl. Phys. Lett.* **111**, 143701. <https://doi.org/10.1063/1.4996300> (2017).
31. Müller, P. *et al.* nanite: Using machine learning to assess the quality of atomic force microscopy-enabled nano-indentation data. *BMC Bioinformatics* **20**, 465. <https://doi.org/10.1186/s12859-019-3010-3> (2019).
32. Zhou, F., Wang, W., Li, M. & Liu, L. in *2015 9th IEEE International Conference on Nano/Molecular Medicine & Engineering (NANOMED)*. 167–172.
33. Lecun, Y., Bottou, L., Bengio, Y. & Haffner, P. Gradient-based learning applied to document recognition. *Proc. IEEE* **86**, 2278–2324. <https://doi.org/10.1109/5.726791> (1998).
34. Esteva, A. *et al.* Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **542**, 115–118. <https://doi.org/10.1038/nature21056> (2017).
35. Girshick, R. Fast R-CNN. *IEEE International Conference on Computer Vision (ICCV)*, 1440–1448 (2015).
36. Girshick, R., Donahue, J., Darrell, T. & Malik, J. Region-based convolutional networks for accurate object detection and segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **38**, 142–158. <https://doi.org/10.1109/tpami.2015.2437384> (2016).
37. Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. You only look once: Unified, real-time object detection. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779–788 (2016).
38. Ren, S., He, K., Girshick, R. & Sun, J. Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**, 1137–1149. <https://doi.org/10.1109/tpami.2016.2577031> (2017).
39. Kiranyaz, S. *et al.* 1D convolutional neural networks and applications: A survey. *Mech. Syst. Signal Process.* **151**, 107398. <https://doi.org/10.1016/j.ymsp.2020.107398> (2021).
40. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems*, 1097–1105 (2012).
41. Chen, X., Chai, Q., Lin, N., Li, X. & Wang, W. 1D convolutional neural network for the discrimination of aristolochic acids and their analogues based on near-infrared spectroscopy. *Anal. Methods* **11**, 5118–5125. <https://doi.org/10.1039/C9AY01531K> (2019).
42. Kiranyaz, S., Ince, T. & Gabbouj, M. Real-time patient-specific ECG classification by 1-D convolutional neural networks. *IEEE Trans. Biomed. Eng.* **63**, 664–675. <https://doi.org/10.1109/TBME.2015.2468589> (2016).
43. Kiranyaz, S., Ince, T. & Gabbouj, M. Personalized monitoring and advance warning system for cardiac arrhythmias. *Sci. Rep.* **7**, 9270. <https://doi.org/10.1038/s41598-017-09544-z> (2017).
44. Li, X. *et al.* One-dimensional convolutional neural network (1D-CNN) image reconstruction for electrical impedance tomography. *Rev. Sci. Instrum.* **91**, 124704. <https://doi.org/10.1063/5.0025881> (2020).
45. Birdi, K. S. *Scanning Probe Microscopes: Applications in Science and Technology* (CRC Press, 2003).
46. Lindh, L., Aroonsang, W., Sotres, J. & Arnebrant, T. Salivary pellicles. *Monogr. Oral Sci.* **24**, 30–39. <https://doi.org/10.1159/000358782> (2014).
47. Hannig, M. & Joiner, A. *The teeth and their environment* Vol. 19, 29–64 (Karger Publishers, 2006).
48. Aguirre, A., Mendoza, B., Levine, M. J., Hatton, M. N. & Douglas, W. H. In vitro characterization of human salivary lubrication. *Arch. Oral Biol.* **34**, 675–677. [https://doi.org/10.1016/0003-9969\(89\)90024-1](https://doi.org/10.1016/0003-9969(89)90024-1) (1989).
49. Bongaerts, J. H. H., Rossetti, D. & Stokes, J. R. The lubricating properties of human whole saliva. *Tribol. Lett.* **27**, 277–287. <https://doi.org/10.1007/s11249-007-9232-y> (2007).
50. Hahn Berg, I. C., Rutland, M. W. & Arnebrant, T. Lubricating properties of the initial salivary pellicle—An AFM study. *Biofouling* **19**, 365–369. <https://doi.org/10.1080/08927010310001618571> (2003).
51. Sotres, J., Lindh, L. & Arnebrant, T. Friction force spectroscopy as a tool to study the strength and structure of salivary films. *Langmuir* **27**, 13692–13700. <https://doi.org/10.1021/la202870c> (2011).
52. Cárdenas, M., Valle-Delgado, J. J., Hamit, J., Rutland, M. W. & Arnebrant, T. Interactions of hydroxyapatite surfaces: Conditioning films of human whole saliva. *Langmuir* **24**, 7262–7268. <https://doi.org/10.1021/la800402s> (2008).
53. Boyd, H. *et al.* Effect of nonionic and amphoteric surfactants on salivary pellicles reconstituted in vitro. *Sci. Rep.* **11**, 12913. <https://doi.org/10.1038/s41598-021-92505-4> (2021).

54. Gonzalez-Martinez, J. F., Kakar, E., Erkselius, S., Rehnberg, N. & Sotres, J. Effect of relative humidity on the viscoelasticity of thin organic films studied by contact thermal noise AFM. *Langmuir* **35**, 6015–6023. <https://doi.org/10.1021/acs.langmuir.8b04222> (2019).
55. Veit, A., Wilber, M. & Belongie, S. Residual networks behave like ensembles of relatively shallow networks. *NIPS'16: Proceedings of the 30th International Conference on Neural Information Processing Systems*, 550–558 (2016).
56. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778 (2016).
57. Chollet, F. o. Keras. <https://github.com/keras-team/keras> (2015).
58. Dawes, C. Rhythms in salivary flow rate and composition. *Int. J. Chronobiol.* **2**, 253–279 (1974).
59. Qian, L., Xiao, X. & Wen, S. Tip in situ chemical modification and its effects on tribological measurements. *Langmuir* **16**, 662–670. <https://doi.org/10.1021/la9905618> (2000).

## Acknowledgements

The Swedish Research Council (Grant No. 2021-03879) is acknowledged for financial support.

## Author contributions

J.S. and J.F.G.M. designed the study. J.S. and H.B. performed the experimental AFM work. J.S. performed ML data analysis and wrote the first version of the manuscript. All authors reviewed and edited the final version of the manuscript.

## Funding

Open access funding provided by Malmö University.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to J.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022