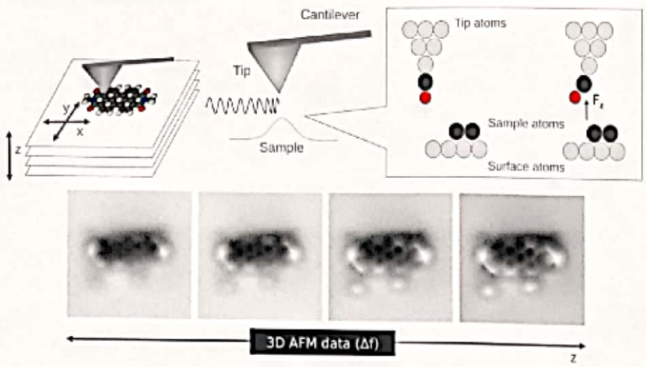
Molecule graph reconstruction from Atomic Force Microscopy images with machine learning

Niko Oinonen[1], Lauri Kurki[1], Alexander Ilin[2], Adam S. Foster[1,3]

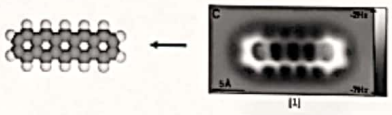
[1]Department of Applied Physics, Aalto University, Finland | [2]Department of Computer Science, Aalto University, Finland | [3]WPI Nano Life Science Institute (WPI-NanoLSI), Kanazawa University, Japan

Aalto University
School of Science

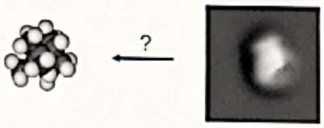
Atomic force microscopy (AFM)



In AFM, an extremely sharp tip "touches" a sample, resulting in a force interaction that causes a deviation in the motion of the tip. A scan over the sample produces an image of the sample, and scans at different distances from the sample result in different contrast. State-of-the-art AFM devices with CO-tip functionalization can image atomic features within individual molecules on surfaces [1]. However...



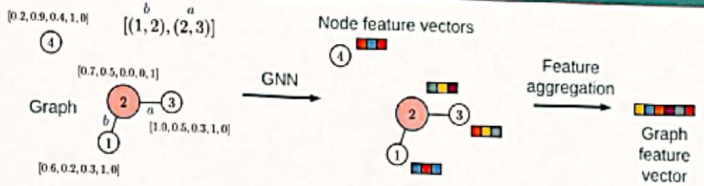
..., while AFM image of flat molecules are fairly easy to interpret...



..., images of more complicated 3D geometries are much more challenging to deal with, and elemental identification is difficult in general.

Challenge: Can we recover the (partial) geometry and elements of atoms from the AFM image?

Machine learning on graphs

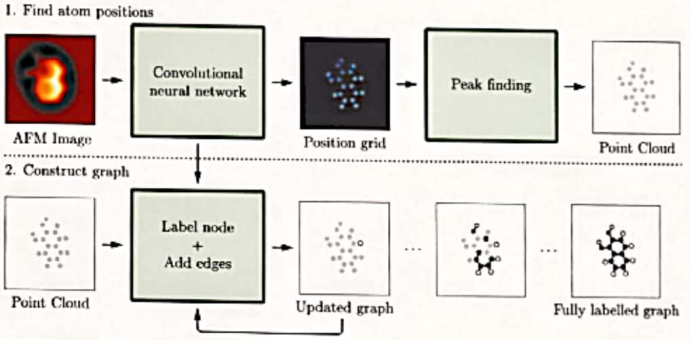


Graph neural networks (GNN) have recently emerged as a general framework for machine learning on graph-structured data [2]. GNNs can encode information in graphs without any handcrafted features in a similar way as convolutional neural networks have done for image data.

Molecules are quite naturally represented as graphs of atoms and bonds. A GNN can encode the positions and types of the atoms into a more abstract set of features for each atom, which can be further aggregated into a single fixed-size feature vector for the molecule. This type of feature vector is easily used for further inference using standard neural networks.

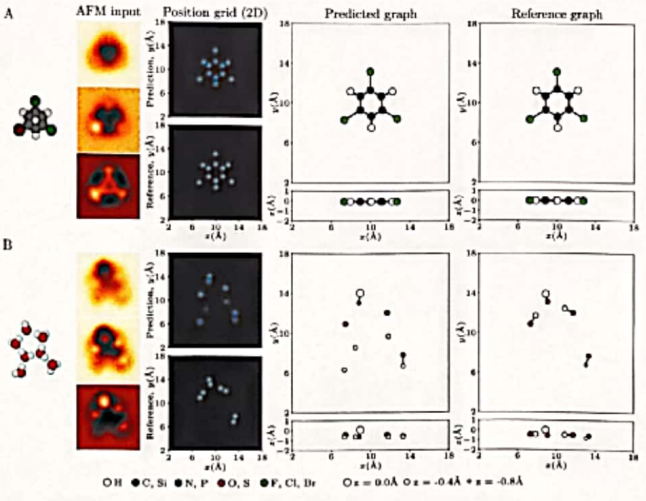
Graphs from AFM images

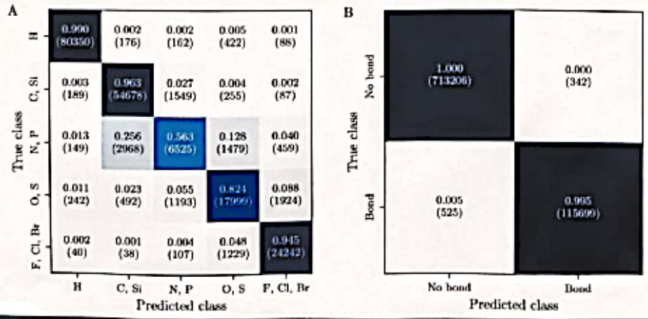
We propose a machine learning model for reconstructing the molecular geometry from a set of AFM images. The model first predicts the positions of the atoms in the AFM images via a CNN and a peak-finding algorithm and then the molecule graph is constructed iteratively by adding one atom and its bonds to the graph at a time using a GNN. The atoms are classified based on groups of the periodic table.



Results

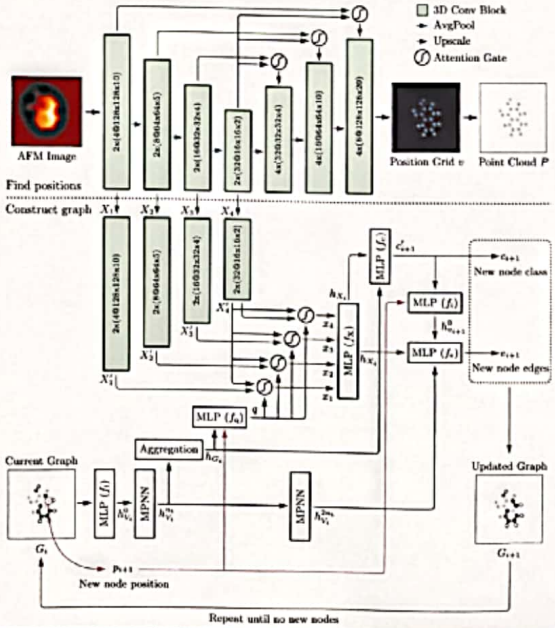
The model is trained and tested on AFM images simulated with the Probe particle model [3]. Atom positions are generally found very accurately, but there are some inconsistencies with deeper atoms. Atom element detection is very accurate for H-, C-, and F-group elements, but N-group elements are sometimes mistaken for C- or O-group. Atom bond detection is extremely accurate. Read more in an upcoming paper: [4].





Model architecture

The atom position detection is done by a U-net type CNN with attention gates in skip connections. The CNN predicts a 3D image grid where atom positions are marked by Gaussian peaks. A peak-finding algorithm finds the positions of these peaks and returns a point cloud of atom positions. The molecule graph is then constructed by picking one new node position at a time and adding it to the graph. At each step the current graph is encoded by a GNN into a feature vector that is combined with an encoding vector for the AFM image. The combined information is then used to predict the new node class and bond connections to the existing nodes in the graph.



References

[1] L. Gross, et al., "The Chemical Structure of a Molecule Resolved by Atomic Force Microscopy", Science, 325, pp. 1110-1114, 2009.
[2] P. Battaglia et al., "Relational inductive biases, deep learning, and graph networks", arxiv: abs/1806.01261.
[3] P. Hapala, et al., "Mechanism of high-resolution STM/AFM imaging with functionalized tips", Physical Review B 90, 085421 , 2014.
[4] N. Oinonen et al., "Molecule graph reconstruction from Atomic Force Microscopy images with machine learning" MRS Bulletin, 2022.

Acknowledgements

This research was supported by the Academy of Finland (Project no. 314877) and was a part of the Flagship Programmes under Projects No. 318930 and 318891 (Competence Center for Materials Bioeconomy, FinnCERES) and the Finnish Center for Artificial Intelligence FCAI. ASF has been supported by the World Premier International Research Center Initiative (WPI), MEXT, Japan. The neural networks were trained using computer resources within the Aalto University School of Science "Science-IT" project.