Surfaces and Interfaces at the Nanoscale (SIN)

Aalto University School of Science

Enhancing AFM Image Analysis Through Machine Learning with Style Translation and Data Augmentation

Jie Huang[a], Niko Oinonen[a,b], Lauri Kurki[a], Adam S. Foster[a,c]

[a] Department of Applied Physics, Aalto University, Finland

[b] Nanolayers Research Computing Ltd., London N12 0HL, United Kingdom

[c] WPI Nano Life Science Institute (WPI-NanoLSI), Kanazawa University, Japan

jie.huang@aalto.fi

Atomic Force Microscopy (AFM) plays a crucial role in characterizing nanostructures. The **Probe Particle Model (PPM)** [1] provides an efficient method for simulating AFM images.

Motivation

How can we enhance the performance of machine learning (ML) models on experimental AFM images?

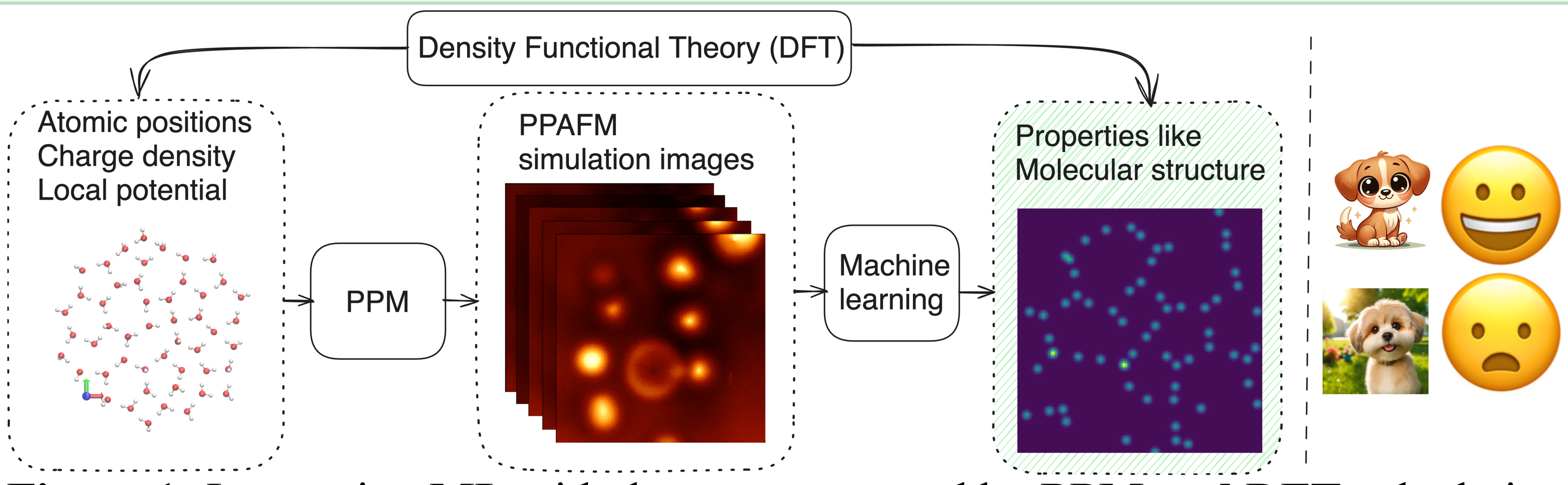

Figure 1: Integrating **ML** with datasets generated by **PPM and DFT** calculations enables the prediction of properties such as **molecular structures** [2, 3] and **electrostatic force potential** [4]. However, when applying the trained models to actual experimental AFM images, the performance often falls short compared to their performance on simulation images.

Hypothesis

The gap between the simulation and experimental AFM images hinders the performance of ML models.

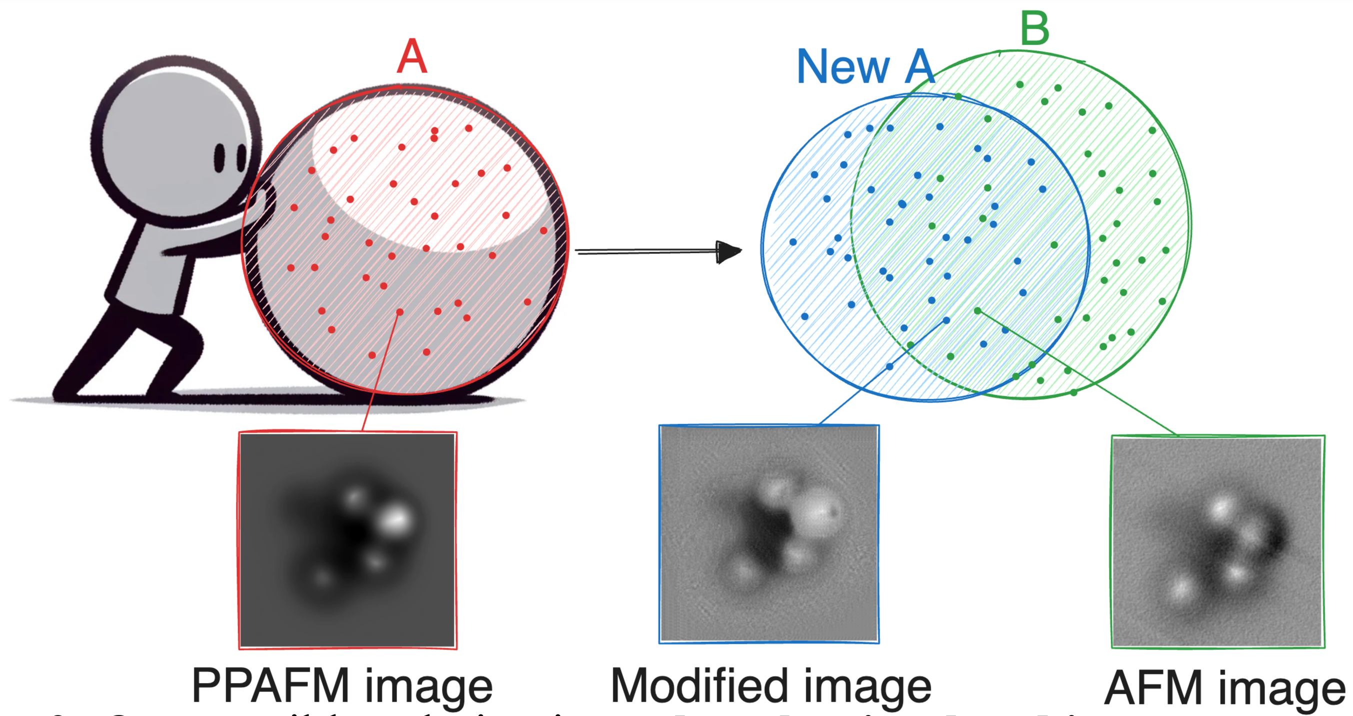

Figure 2: One possible solution is to **alter the simulated images** to more closely resemble experimental images, incorporating these modifications into the training dataset (**Data Augmentation**).

CycleGAN

Experimental-like AFM images can be obtained by feeding simulation images into a trained image generator.

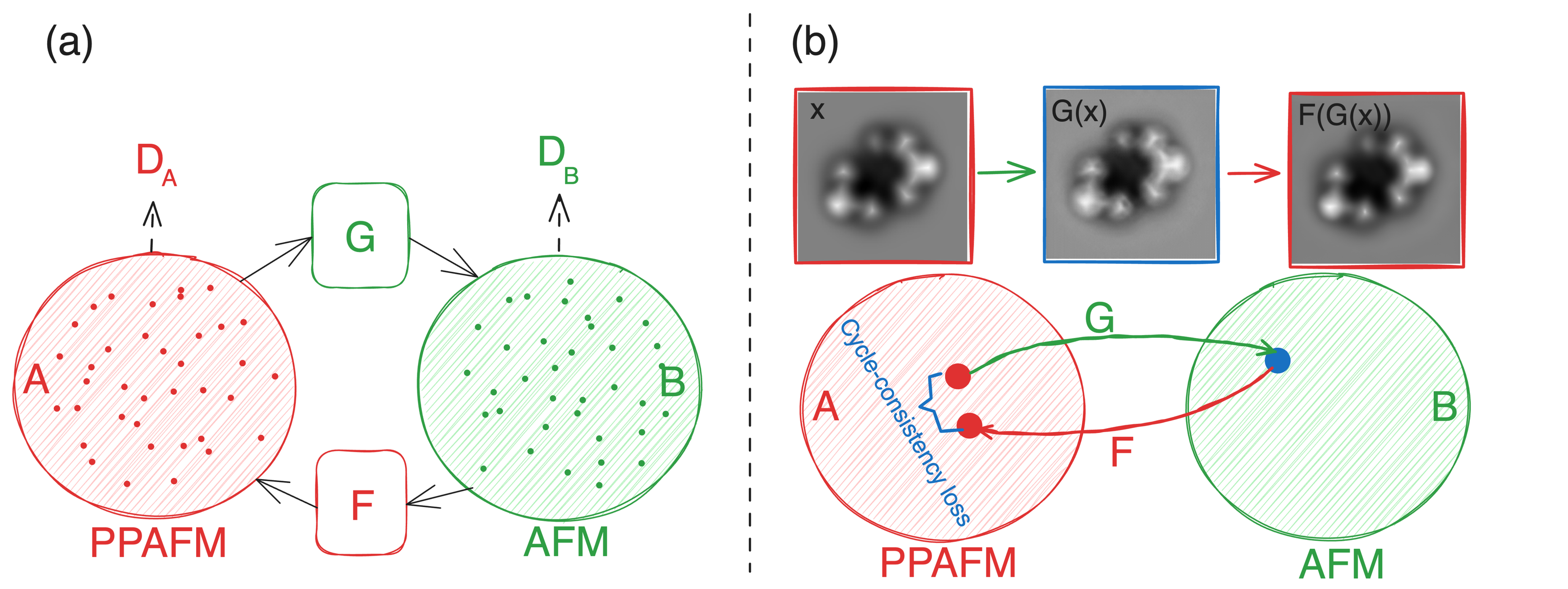

Figure 3: (a) **CycleGAN [5]** includes two mapping functions G: A → B and F: B → A, and associated adversarial discriminators D_A and D_B, which encourages **G to translate A into outputs indistinguishable from domain B**, and vice versa for D_A and F. (b) **Cycle onsistency** ensures that converting from one domain to another and back again returns to the original starting point.

Generated Dataset Evaluation

The generated experimental-like images receive high scores on a authenticity test compared to simulation images.

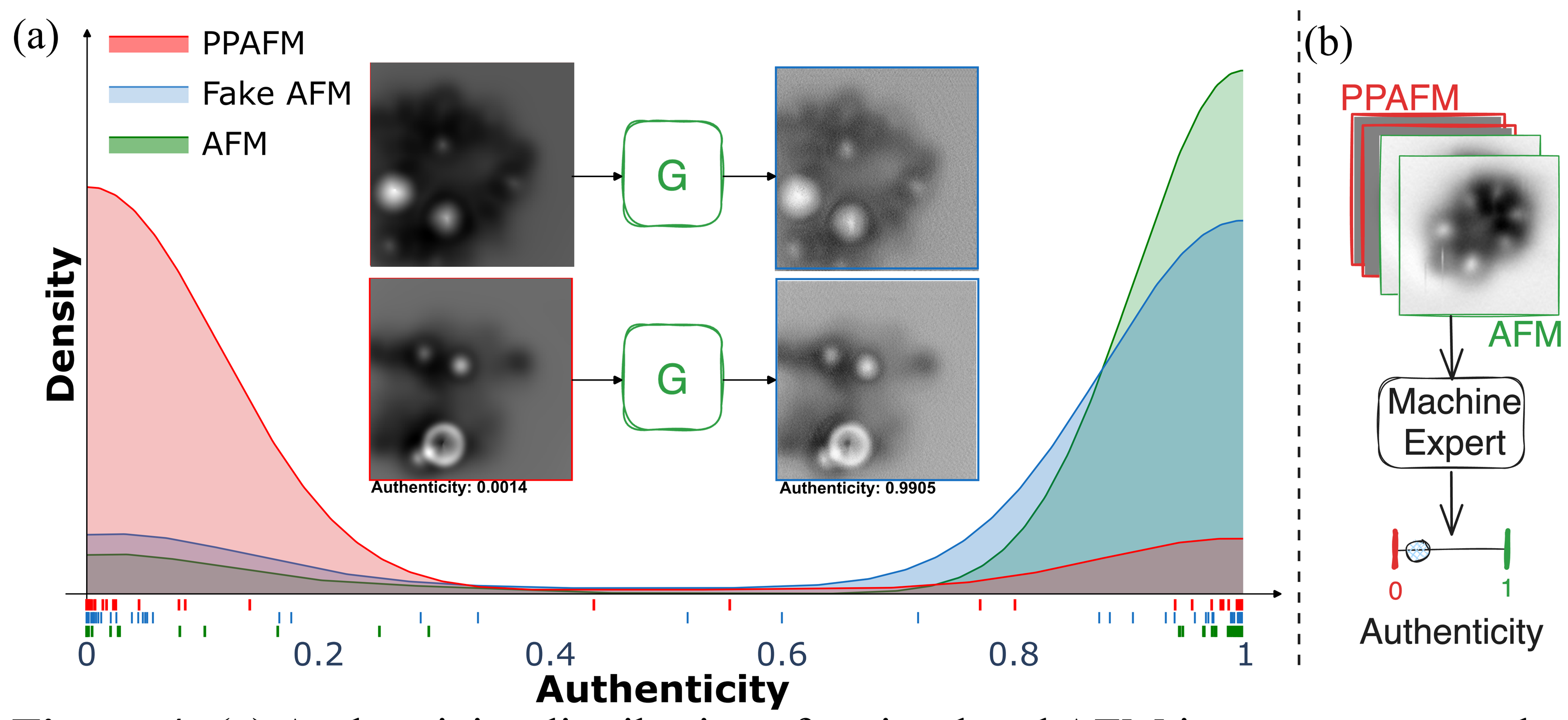

Figure 4: (a) Authenticity distributions for simulated AFM images, generated style-translated fake AFM, and real AFM images. (b) The score of authenticity is given by a machine expert trained by PPAFM and experimental AFM images.

ML Model Improvement

The structure discovery models trained on the datasets utilizing style-translated fake AFM images perform better on experimental AFM images.

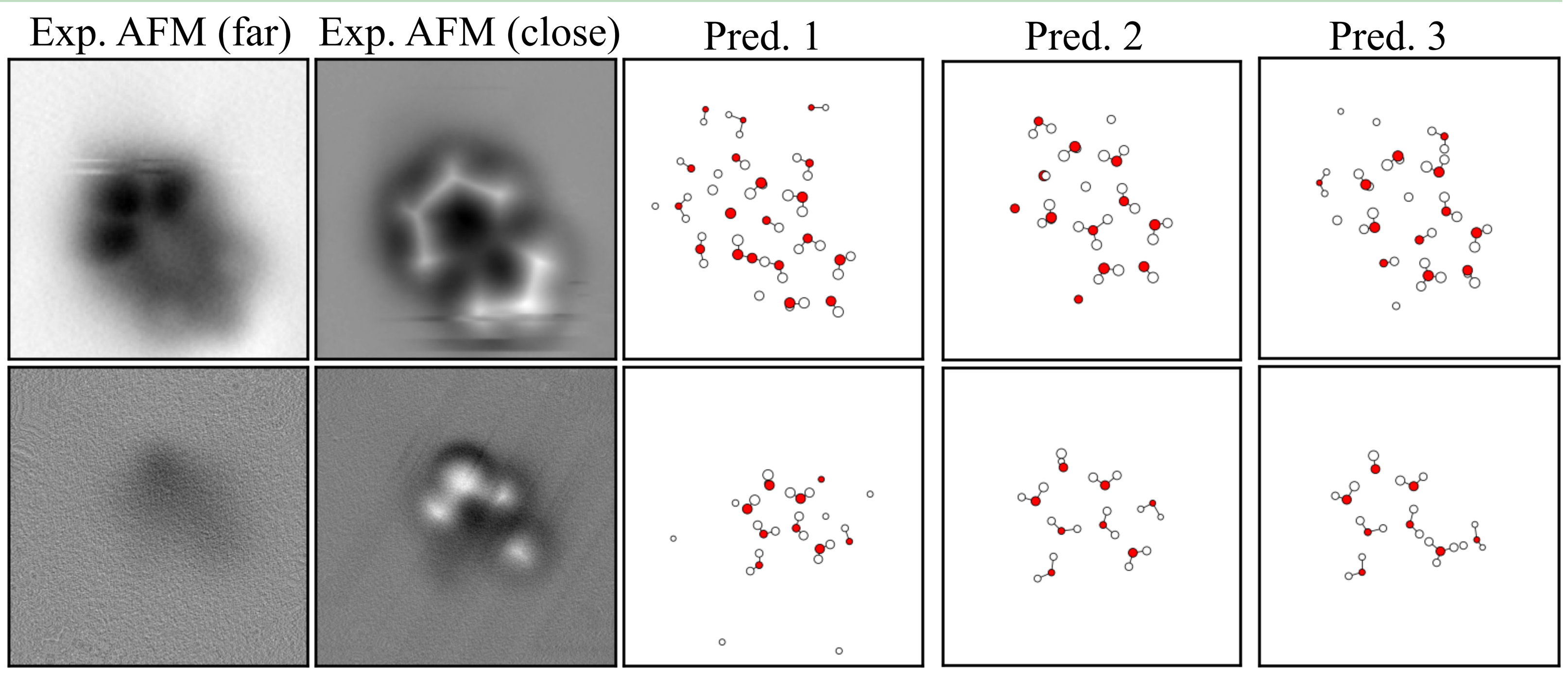

Figure 5: Three predictions (Pred. 1 to Pred. 3) for the same two systems are given by three different models trained by PPAFM, style-translated fake AFM, and PPAFM+style-translated fake AFM images.

Future Directions

1. Apply this workflow to additional systems.
2. Explore using the reverse generator as a denoising step.

Acknowledgements

This work was supported by World Premier International Research Center Initiative (WPI), MEXT, Japan, and by the Academy of Finland (Projects 347319 and 346824). The authors acknowledge the computational resources provided by the Aalto Science-IT Project and CSC, Helsinki.

References

[1] P. Hapala, R. Temirov, F. S. Tautz, P. Jelínek. Origin of high-resolution iets-stm images of organic molecules with functionalized tips. *Phys. Rev. Lett.*, 113:226101, 2014.

[2] B. Alldritt, P. Hapala, N. Oinonen, F. Urtev, O. Krejčí, F. F. Canova, J. Kannala, F. Schulz, P. Liljeroth, A. S. Foster. Automated structure discovery in atomic force microscopy. *Sci. Adv.*, 6(9), 2020.

[3] F. Priante, N. Oinonen, Y. Tian, D. Guan, C. Xu, S. Cai, P. Liljeroth, Y. Jiang, A. S. Foster. Structure Discovery in Atomic Force Microscopy Imaging of Ice. *ACS Nano*, 18(7):5546-5555, 2024.

[4] N. Oinonen, C. Xu, B. Alldritt, F. F. Canova, F. Urtev, S. Cai, O. Krejčí, J. Kannala, P. Liljeroth, A. S. Foster. Electrostatic discovery atomic force microscopy. *ACS Nano*, 16(1):89–97, 2022.

[5] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks, *arXiv:1703.10593*, 2017.