Bridging experiments with theory: ELIT to HPC simulations

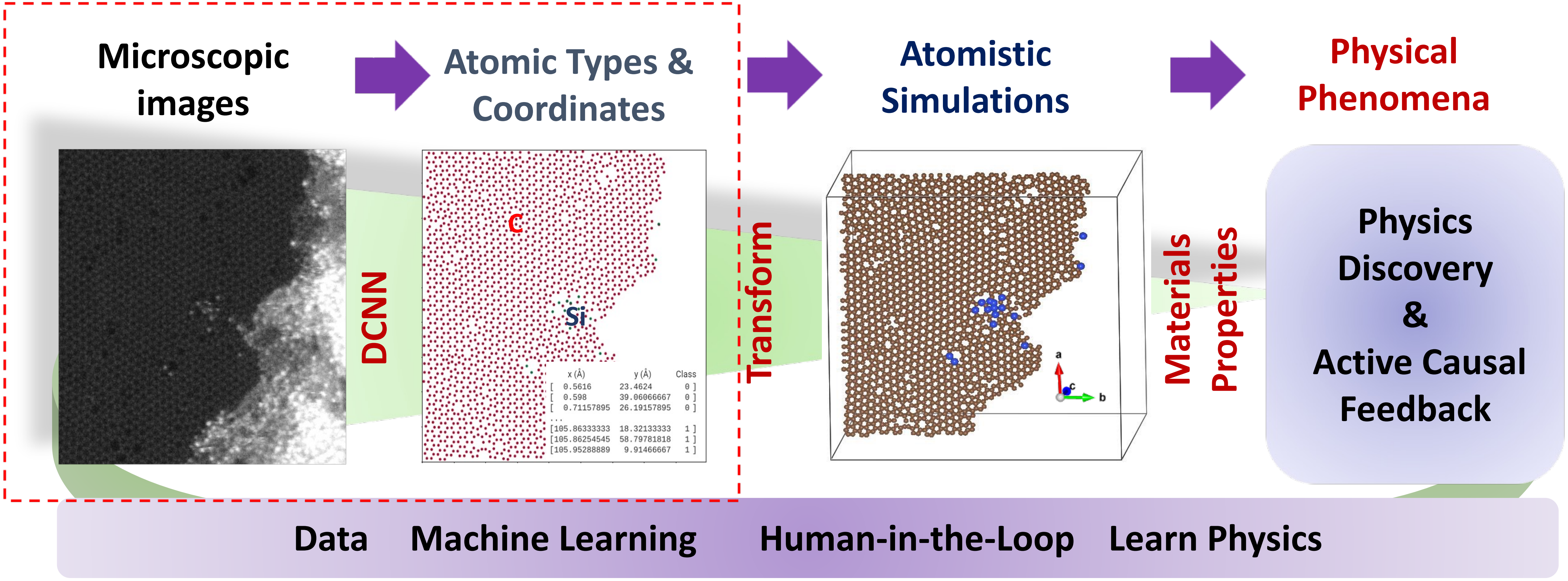
Ayana Ghosh

Research Scientist

Computational Sciences & Engineering Division

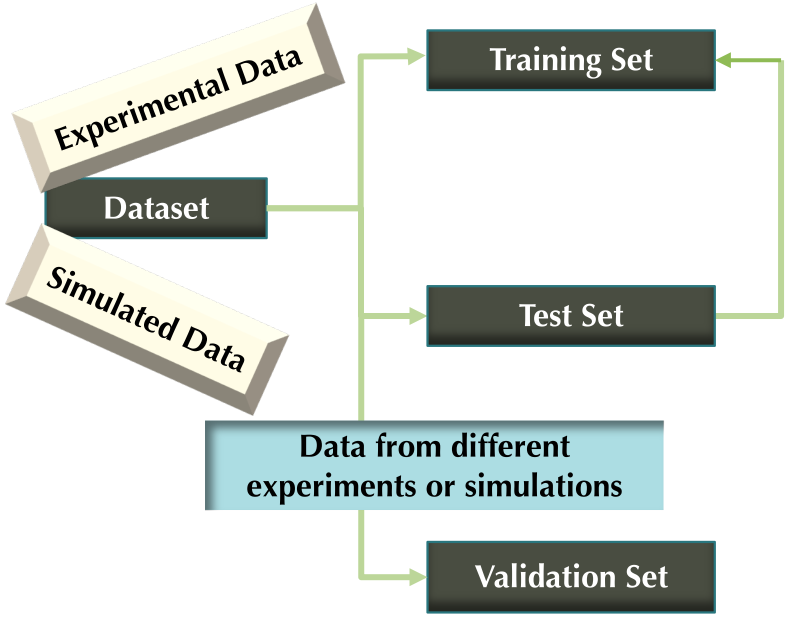
Oak Ridge National Laboratory

Workflows



Microscopic images
DCNN
Atomic Types & Coordinates
C
Si
x (Å) y (Å) Class
[0.5616 23.4624 0]
[0.598 39.06066667 0]
[0.71157895 26.19157895 0]
...
[105.86333333 18.32133333 1]
[105.86254545 58.79781818 1]
[105.95288889 9.91466667 1]
Transform
Atomistic Simulations
a
c
b
Materials Properties
Physical Phenomena
Physics Discovery & Active Causal Feedback
Data Machine Learning Human-in-the-Loop Learn Physics

Quick reminders: working with image data



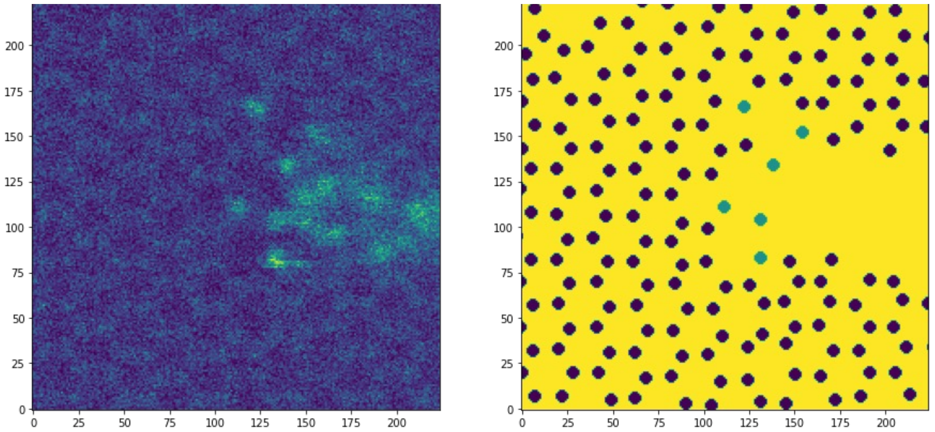
Fully-labeled

Weakly-labeled

Unlabeled

Image

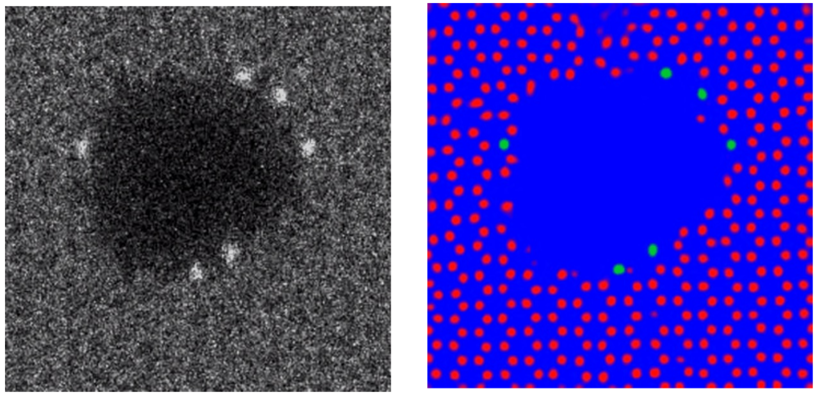
Ground truth



Theory in the loop to connect with experiments combined with ML, beyond traditional way of theory-experimental matching

Experiment

Neural network output



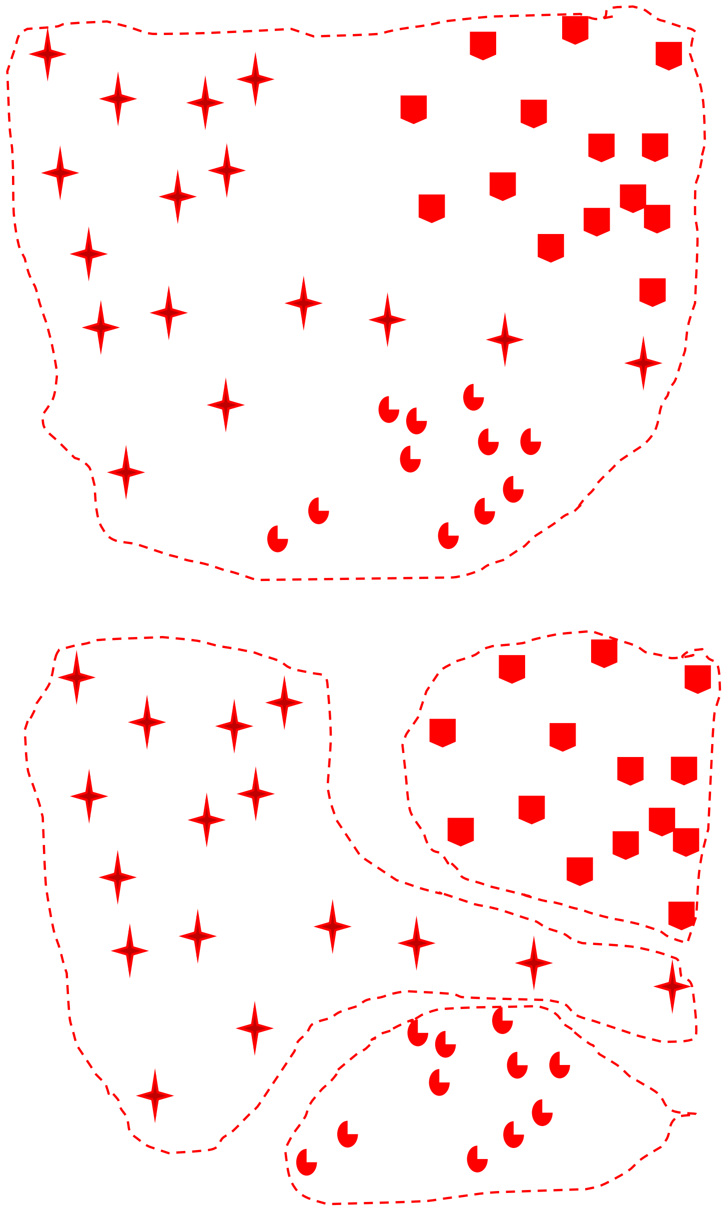
M. Ziatdinov, A. Ghosh, T. Wong, & S. V. Kalinin. "*AtomAI: a deep learning framework for analysis of image and spectroscopy data in (Scanning) transmission electron microscopy and beyond*." Nat. Machn. Intell. (2023).

Ensemble Learning Iterative Training (ELIT) framework

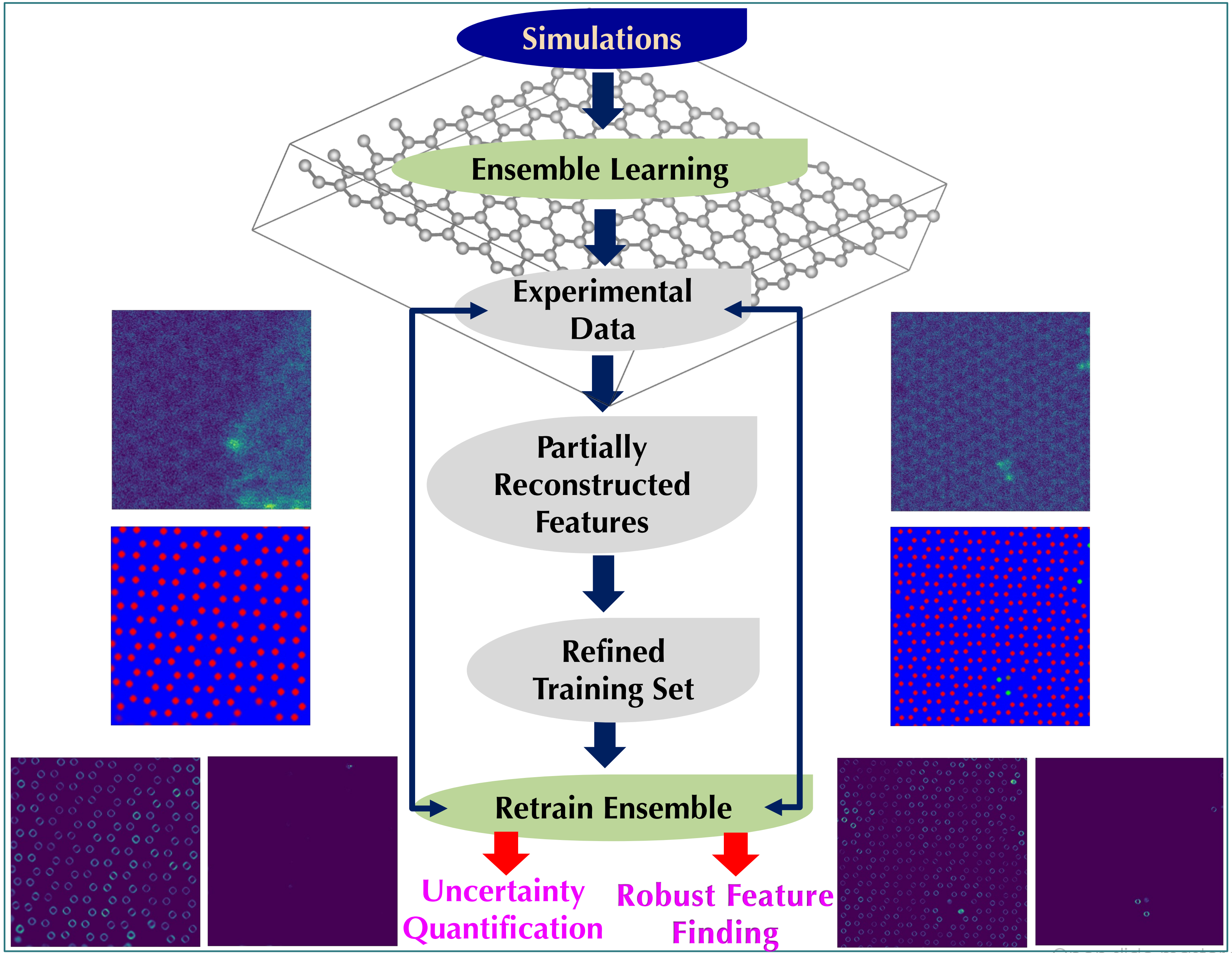
- Deep learning strategies applied to categorize natural images into different classes, e.g., CIFAR, MNIST datasets

- Each class generally contains thousands of samples with large variability

- Atom/particle/defect identification in STEM
  - Find nearly identical objects
  - Dataset shifts, out-of-distribution data
  - Feature finding becomes much more challenging

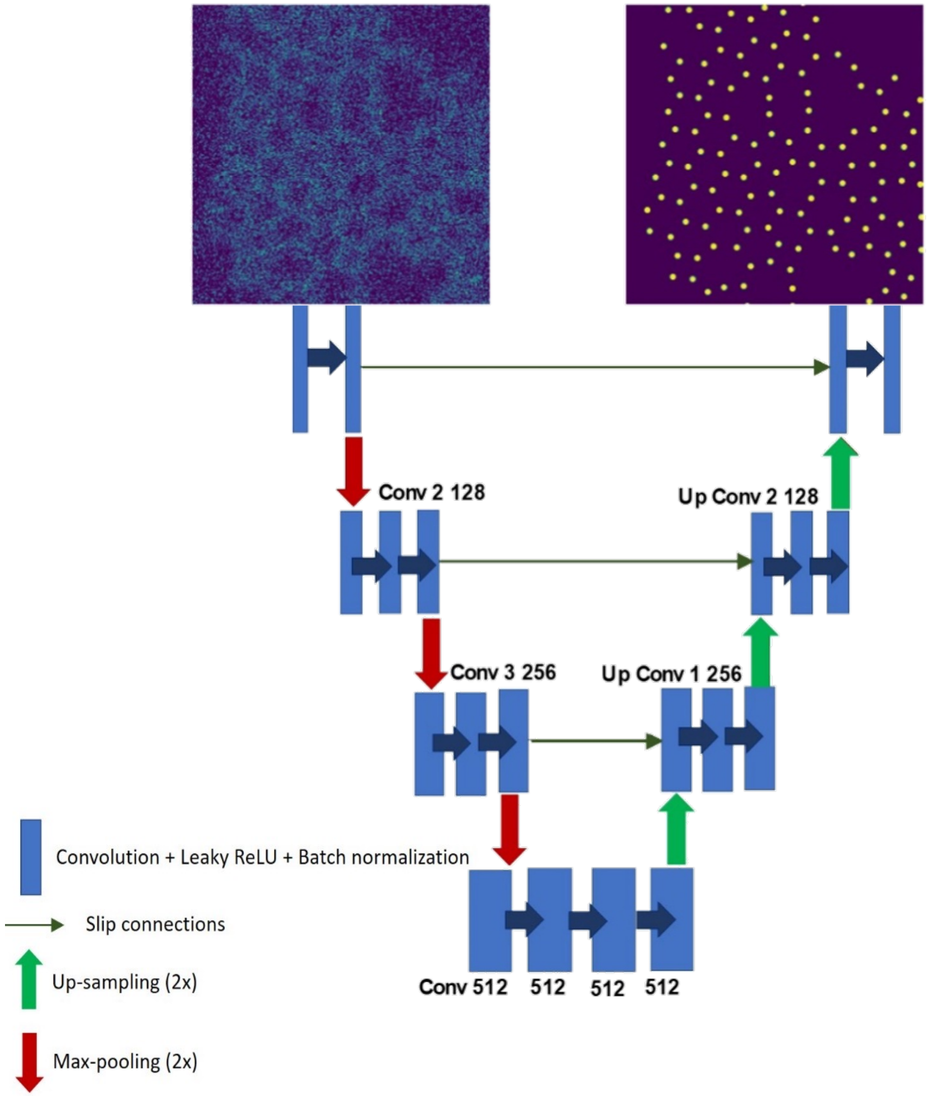


ELIT framework



ELIT framework

Deep Network & Algorithm



Algorithm 1 Steps to train in ELIT framework

$E_{initial} \leftarrow EnsTrain(D_{sim})$ ▷ Train ensemble of models
$P_{Ens} \leftarrow EnsPredict(E_{initial}, D_{exp})$ ▷ Apply ensemble to expt data
$E_0 \leftarrow ArtifactFree(E_{initial}, P_{Ens})$ ▷ Choose artifact-free models
for i in $\{1, \ldots, N\}$ **do**
    $D_i \leftarrow TrainData(E_{i-1})$ ▷ Generate training data
    $E_i \leftarrow EnsTrain(D_i)$ ▷ Train a new ensemble of models
    $P_{Ens} \leftarrow EnsPredict(E_{recent}, D_{exp})$ ▷ Apply ensemble to expt data
done

Optional Steps:

$D_{MC} \leftarrow MultiClassData(E_N)$ ▷ Generate multi-class training set
$E_{final} \leftarrow EnsTrain(D_{MC})$ ▷ Train a new ensemble of models

Link to open-access notebooks, tutorials!

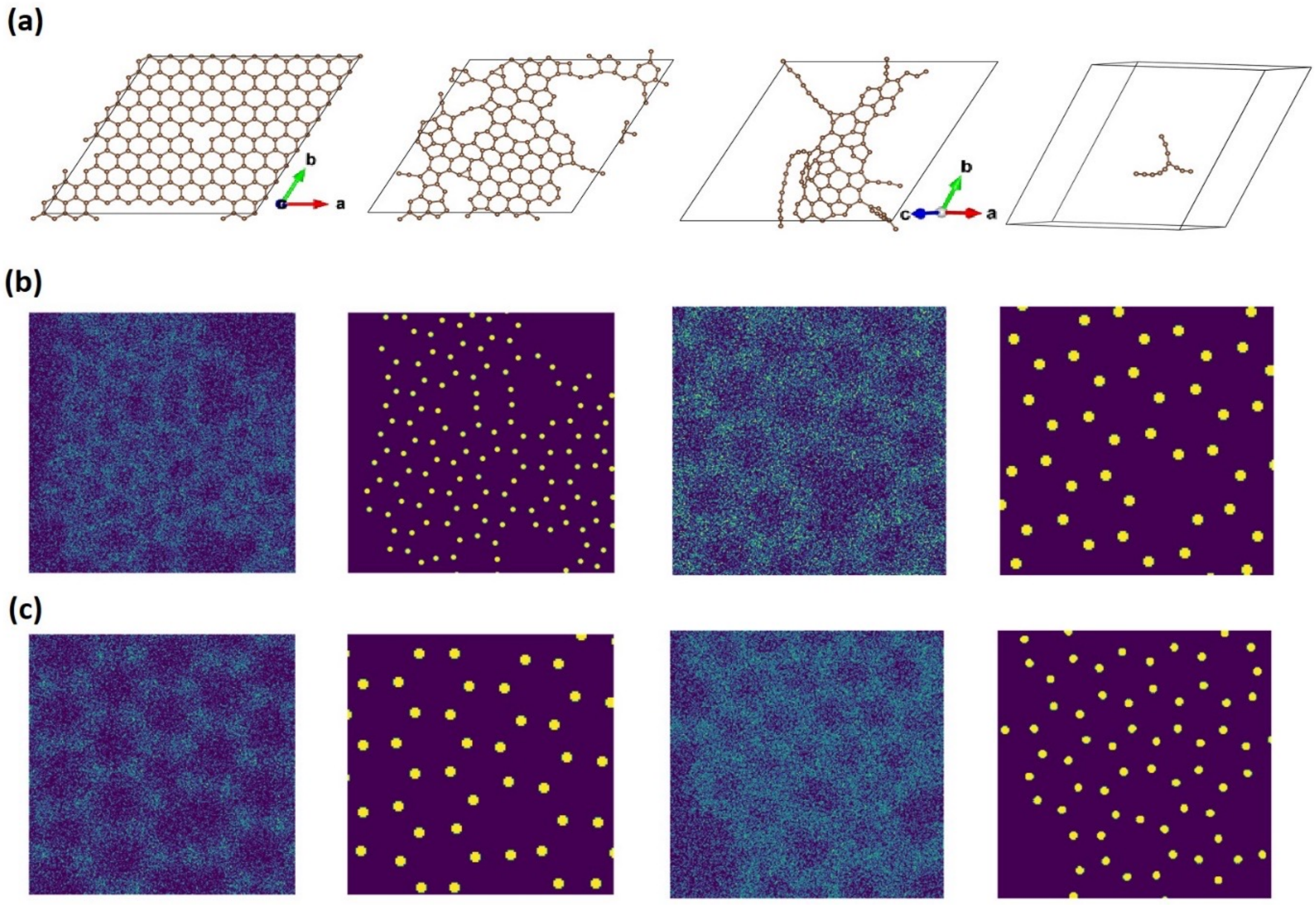
https://github.com/aghosh92/ELIT

https://github.com/aghosh92/ELIT_tutorial

https://github.com/SergeiVKalinin/ML-ElectronMicroscopy-2023

ELIT framework

Image Augmentation

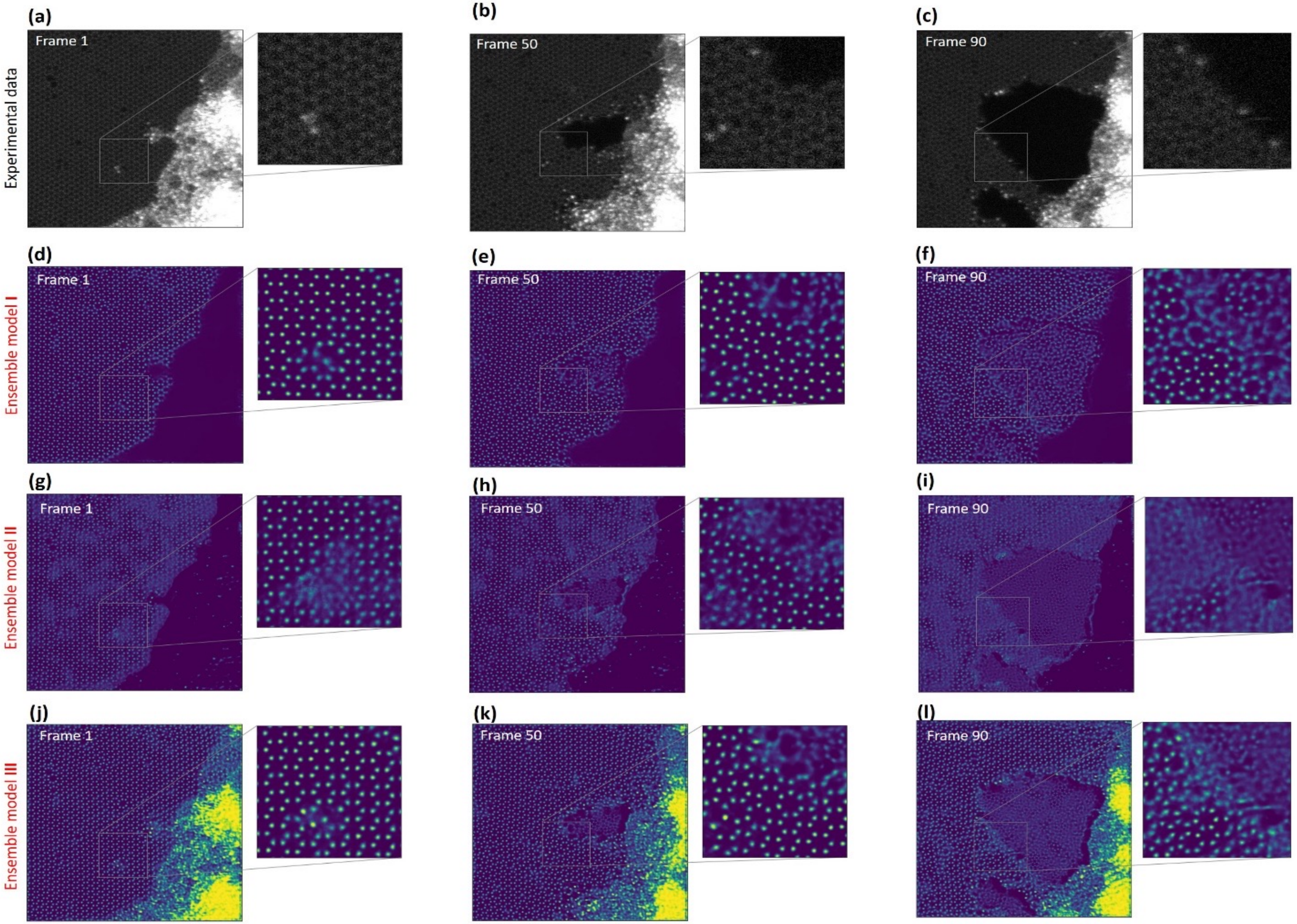


A. Ghosh, B.G. Sumpter, O Dyck, S.V. Kalinin and M. Ziatdinov, *Ensemble learning and iterative training (ELIT) machine learning: applications towards uncertainty quantification and automated experiment in atom-resolved microscopy*, **npj Comput. Mater.** 7, 1-8 (2021).

ELIT framework

Results & Discussions

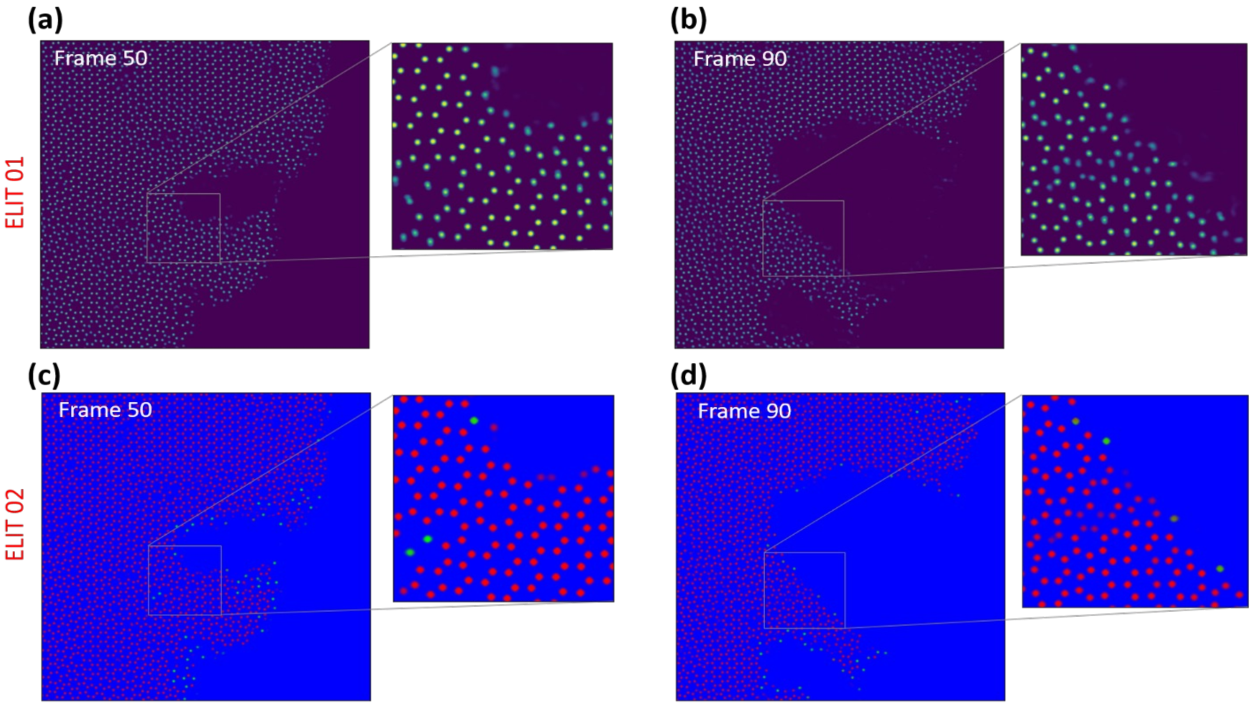
Application of ensemble models trained on simulated data to real data



Oak Ridge National Laboratory

ELIT framework

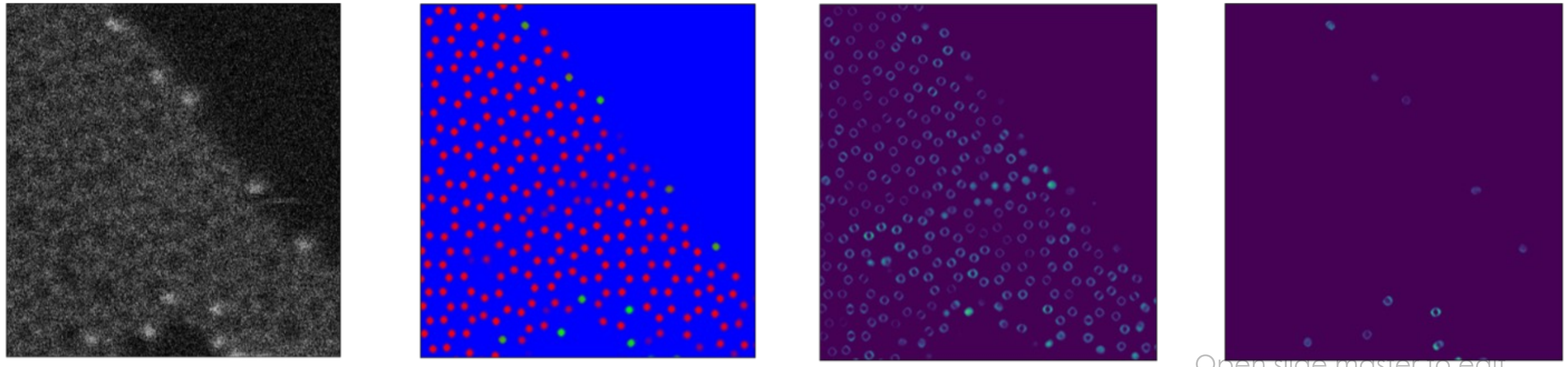
Results & Discussions



Mean DL prediction (semantically segmented map) from the ensemble after the first iteration

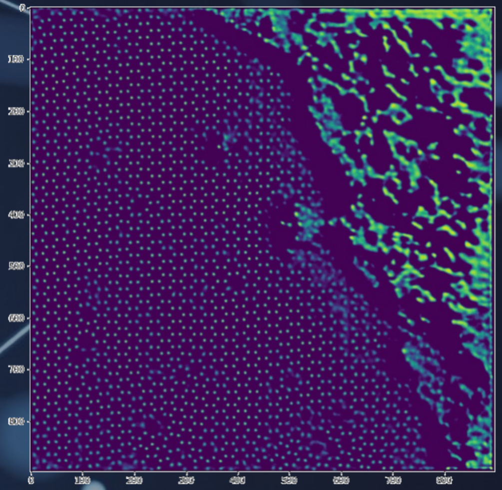
Prediction (semantically segmented map) after next iteration using multi-class segmentation masks with training data randomly sampled from the entire movie

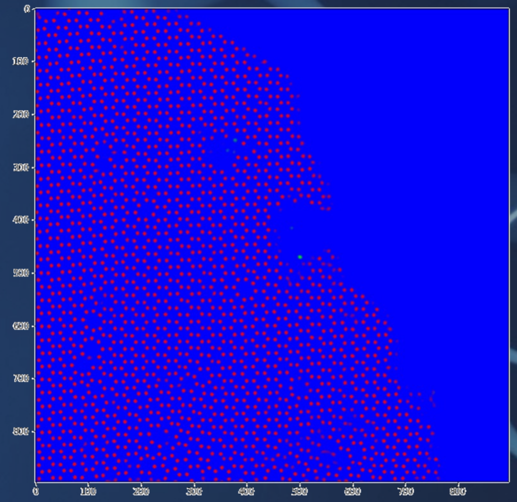
Frame 90 | **Prediction** | **Uncertainty in 'C' channel** | **Uncertainty in 'impurity' channel**



Estimation of uncertainty in ensemble predictions on the level of individual pixels

Takeaways…





A. Ghosh, B.G. Sumpter, O Dyck, S.V. Kalinin and M. Ziatdinov, *Ensemble learning and iterative training (ELIT) machine learning: applications towards uncertainty quantification and automated experiment in atom-resolved microscopy*, **npj Comput. Mater.** 7, 1-8 (2021).

K. M Roccapriore, M. G. Boebinger, O. Dyck, A. Ghosh, R. R. Unocic, S. V Kalinin and M. Ziatdinov, *Probing Electron Beam Induced Transformations on a Single-Defect Level via Automated Scanning Transmission Electron Microscopy*, **ACS Nano** 2022, 16, 10, 17116–17127 (2022).

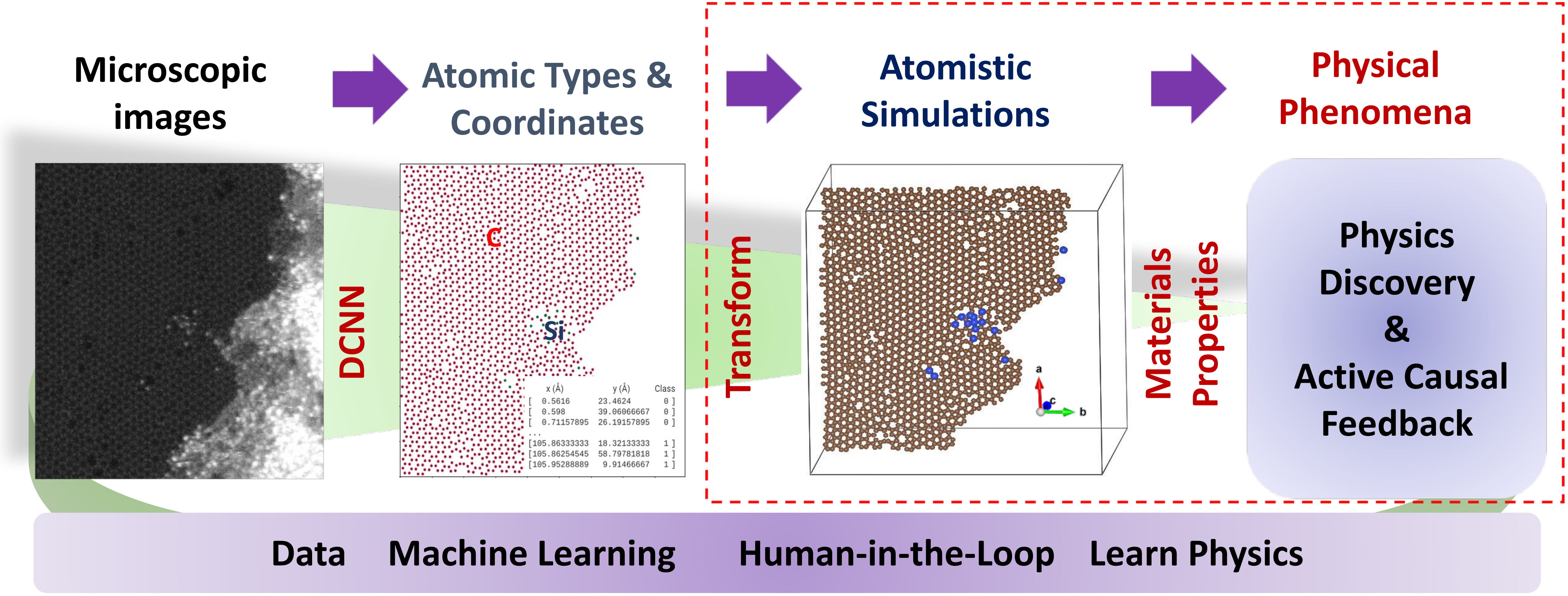
S.V Kalinin, R. Vasudevan, Y. Liu, A. Ghosh, K. Roccapriore and M. Ziatdinov, *Probe microscopy is all you need,* **Mach. Learn.: Sci. Technol.** 4 023001 (2023).

Let's start going through some notebooks

- Data preparation (Notebook_I)
- Training ELIT models (Notebook_II)

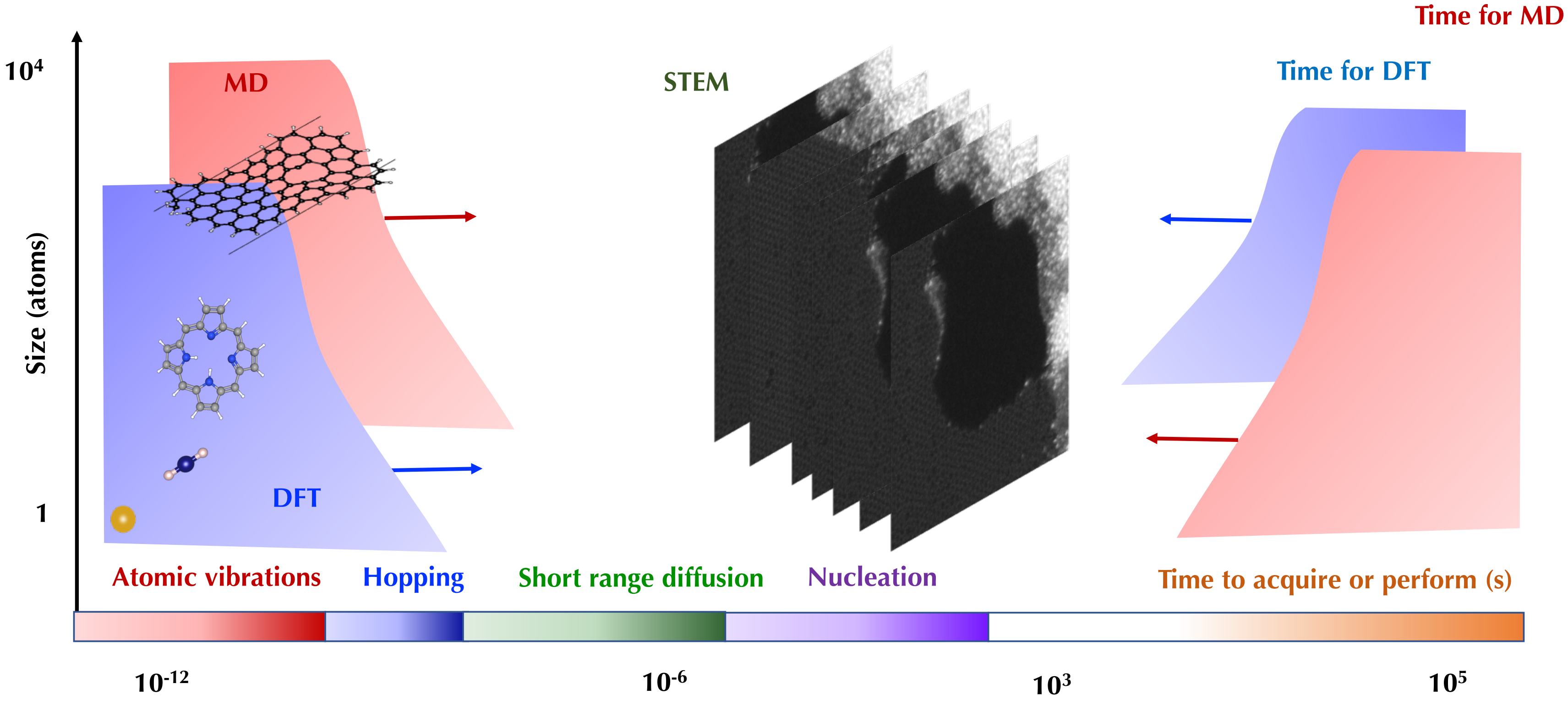
https://github.com/SergeiVKalinin/ML-ElectronMicroscopy-2023

Workflows



Microscopic images
DCNN
Atomic Types & Coordinates
C
Si
x (Å) y (Å) Class
[0.5616 23.4624 0]
[0.598 39.06066667 0]
[0.71157895 26.19157895 0]
...
[105.86333333 18.32133333 1]
[105.86254545 58.79781818 1]
[105.95288889 9.91466667 1]
Transform
Atomistic Simulations
a
c
b
Materials Properties
Physical Phenomena
Physics Discovery & Active Causal Feedback
Data
Machine Learning
Human-in-the-Loop
Learn Physics

Bridging microscopic images with simulations via deep learning

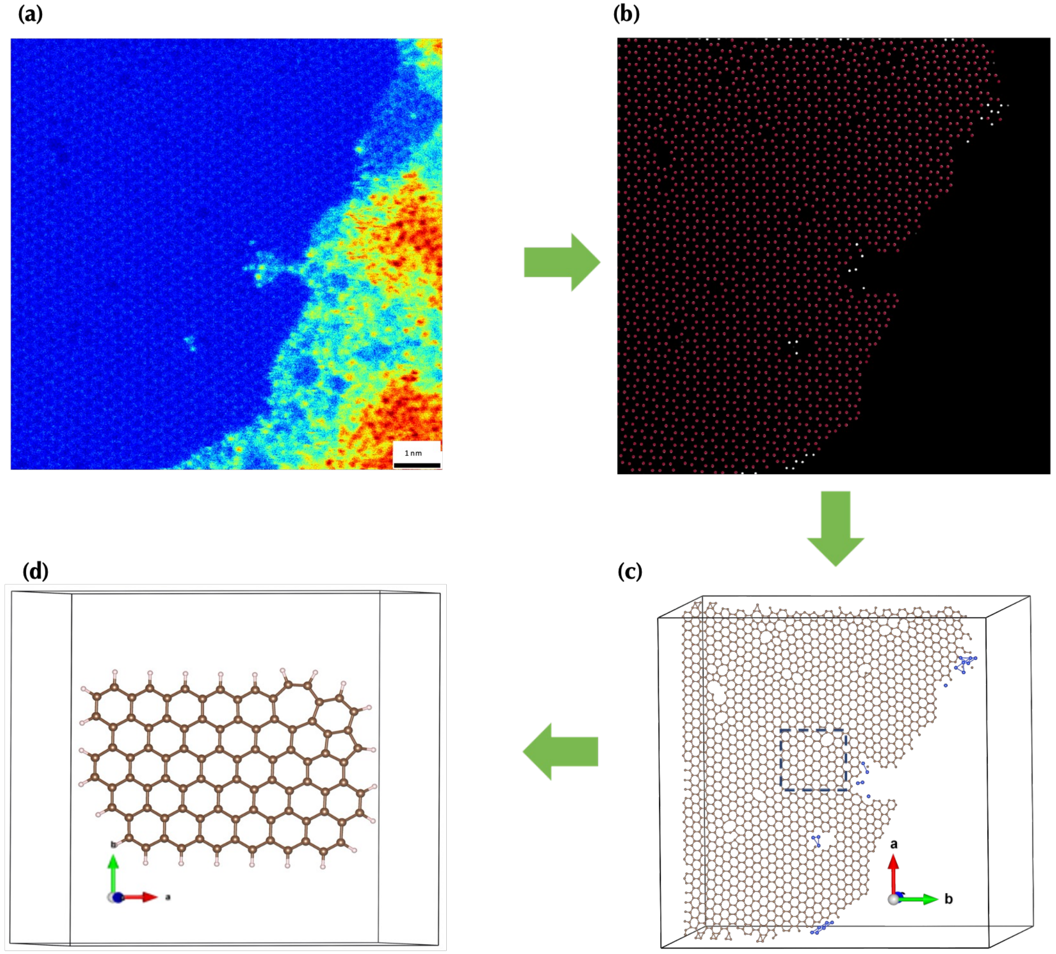


Roadblocks in direct piping of the STEM data into the simulation environment

- Disparity in time and length scales
- Identification of regions of interest
- On the fly analyses

A. Ghosh,* M. Ziatdinov, O. Dyck, B.G. Sumpter, and S. V. Kalinin, *Bridging microscopy with molecular dynamics and quantum simulations: An AtomAI based pipeline*, **npj Comput. Mater**. 8, 1-11 (2022).

What did we do ?

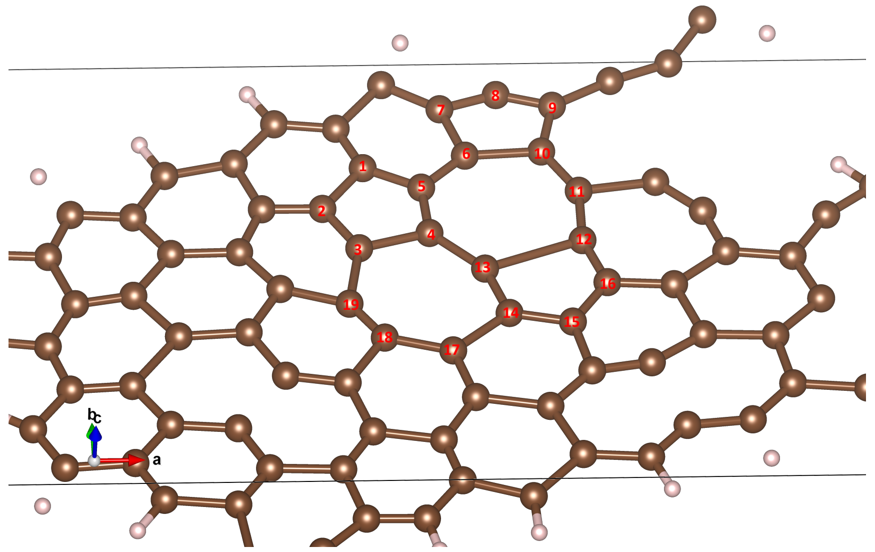


- Use ELIT or other methods to find coordinates of atoms or features using AtomAI from a STEM image
- Build an object (file with coordinates) that is readable by commonly used Molecular Dynamics (MD) or *ab initio* MD packages such as ASE, LAMMPS, VASP
- Perform MD, *ab initio* MD, DFTB simulations as required for geometry reconstruction

https://github.com/SergeiVKalinin/ML-ElectronMicroscopy-2023

Bond length and angle analyses

- Defects begin to propagate and rearrange themselves to form 5-7-7-5 defects at higher temperatures
- The average bond length of the 7-members and 5-members rings to be 1.430 Å and 1.606 Å, respectively
- The average percentage error in the coordinates along x and y directions are <5% between DL outputs, simulations



Temperature (K)	C_{1-2} (Å)	C_{2-3} (Å)	C_{3-4} (Å)	C_{4-5} (Å)	C_{5-6} (Å)	C_{6-7} (Å)	C_{7-8} (Å)
4000	1.529	1.403	1.628	1.358	1.378	1.474	1.324
Temperature (K)	C_{8-9} (Å)	C_{9-10} (Å)	C_{10-11} (Å)	C_{11-12} (Å)	C_{12-13} (Å)	C_{13-14} (Å)	C_{6-10} (Å)
4000	1.270	1.390	1.411	1.470	2.336	1.413	1.693
Temperature (K)	C_{14-15} (Å)	C_{15-16} (Å)	C_{14-17} (Å)	C_{17-18} (Å)	C_{18-19} (Å)	C_{19-3} (Å)	C_{13-4} (Å)
4000	1.412	1.423	1.698	1.544	1.256	1.674	1.596
Temperature (K)	C_{16-12} (Å)						
4000	1.351						

Oak Ridge National Laboratory

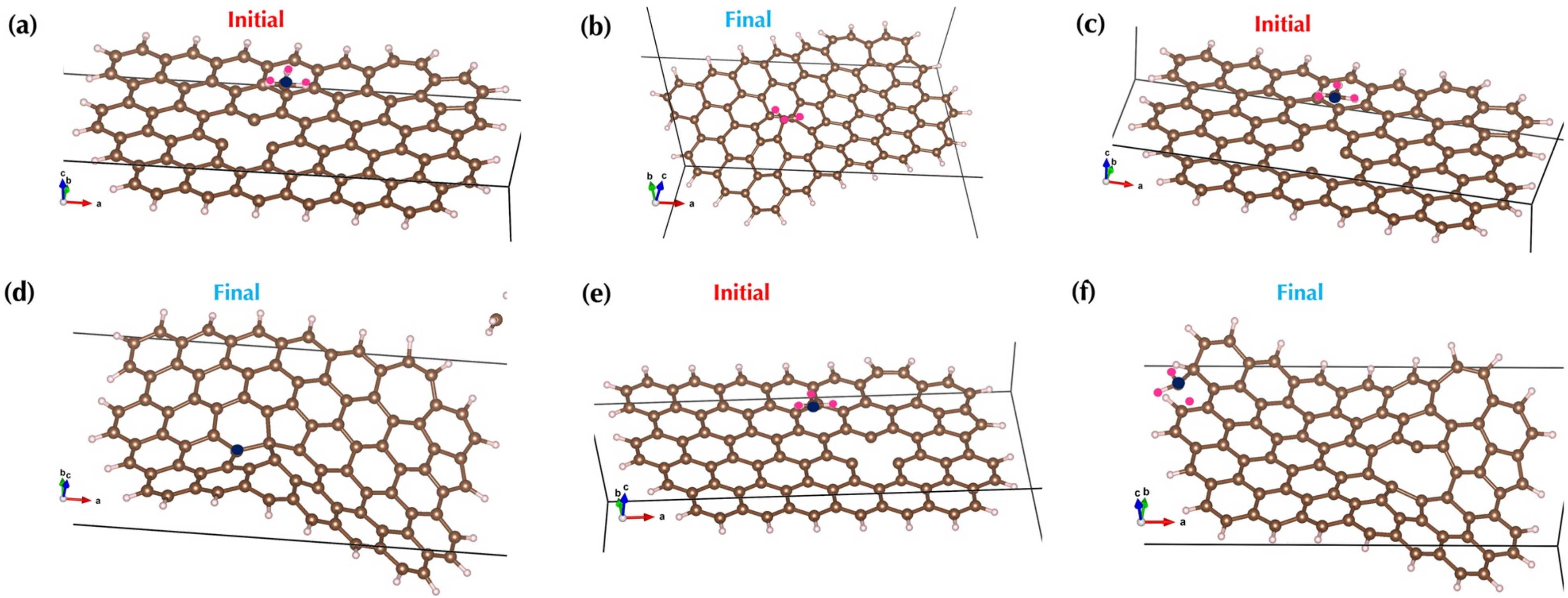
Graphene healing effects

Temperature Dependent Dynamics

Ab initio MD

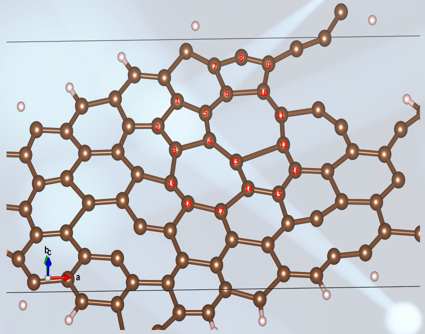
CH_3 molecule

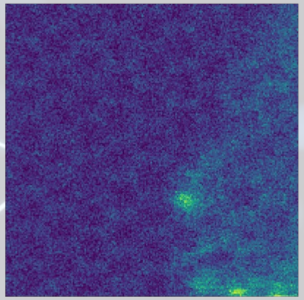
- Graphene holes have healed completely or partially (not all hexagonal rings can be achieved) in all cases depending on the variations in the energy landscapes
- Interactions of these molecules with the edge atoms can still be observed
- Strong dependence of choosing the initial configuration of the atoms and distances between the molecule and the surface along with temperature

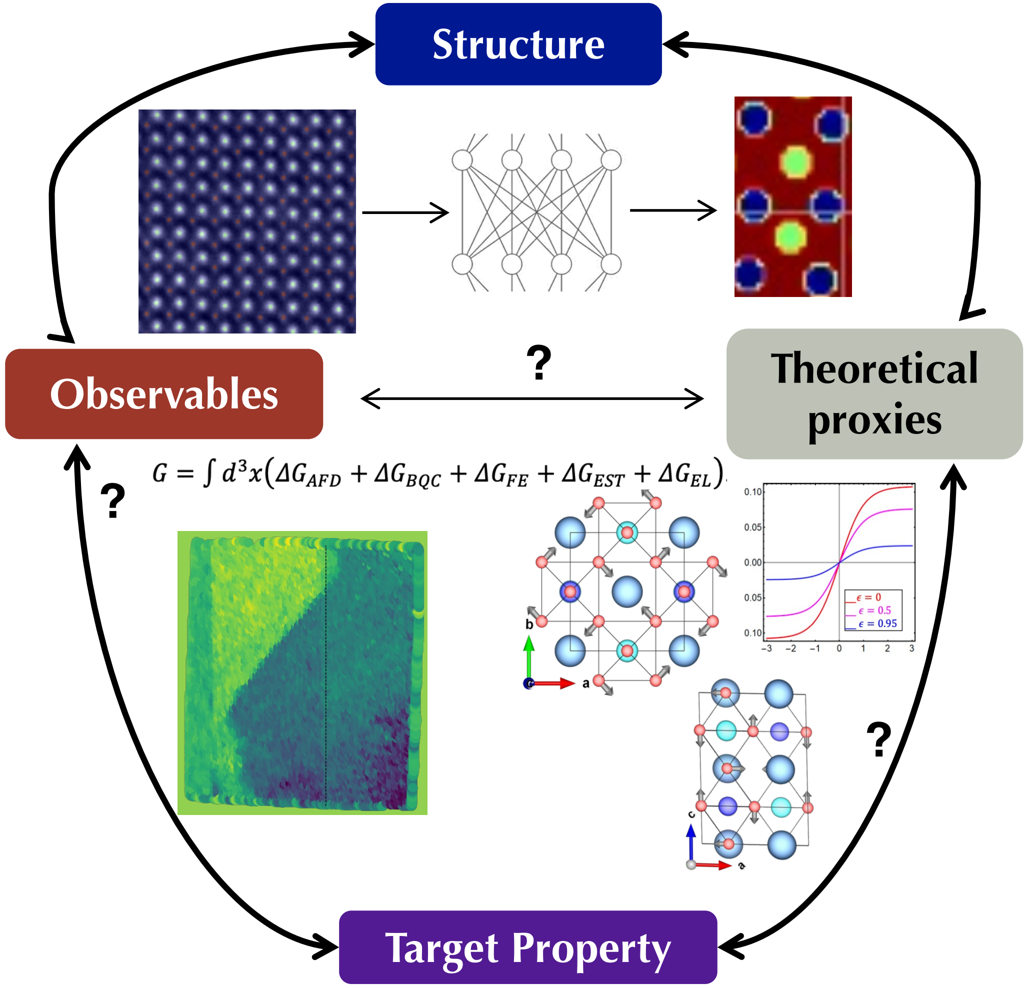


OAK RIDGE National Laboratory

Takeaways and into the future…







A. Ghosh,* M. Ziatdinov, O. Dyck, B.G. Sumpter, and S. V. Kalinin, *Bridging microscopy with molecular dynamics and quantum simulations: An AtomAI based pipeline*, **npj Comput. Mater**. 8, 1-11 (2022).

C. Nelson, A. N. Morozovska, M. A. Ziatdinov, E. A. Eliseev, X. Zhang, I. Takeuchi, & S. V. Kalinin, Mapping causal patterns in crystalline solids , arXiv preprint arXiv:2103.01951 (2021).

Y. Liu, M. Ziatdinov, S.V. Kalinin, *Exploring Causal Physical Mechanisms via Non-Gaussian Linear Models and Deep Kernel Learning: Applications for Ferroelectric Domain Structures*. **ACS Nano**, 16, 1250– 1259 (2022).

A. Ghosh,* G. Palanichamy, D. P. Trujillo, M. Shaikh and S. Ghosh, *Insights into cation ordering of double perovskite oxides from machine learning and causal relations*, **Chem. Mater.** 34, 7563–7578 (2022).

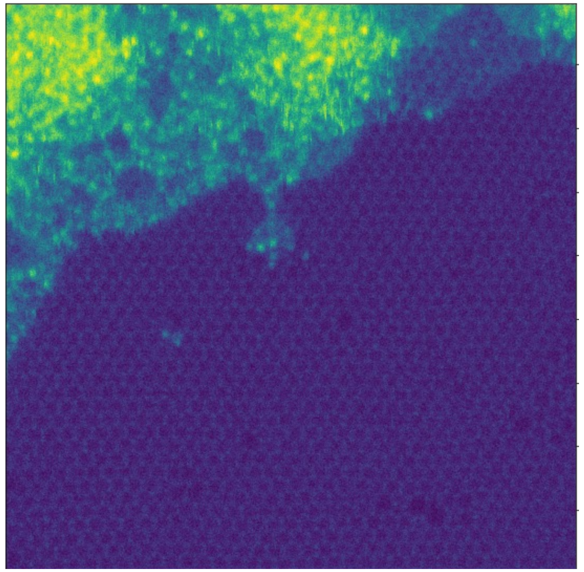
A. Ghosh,* *Towards physics-informed explainable machine learning for materials research and development, submitted* (2023).

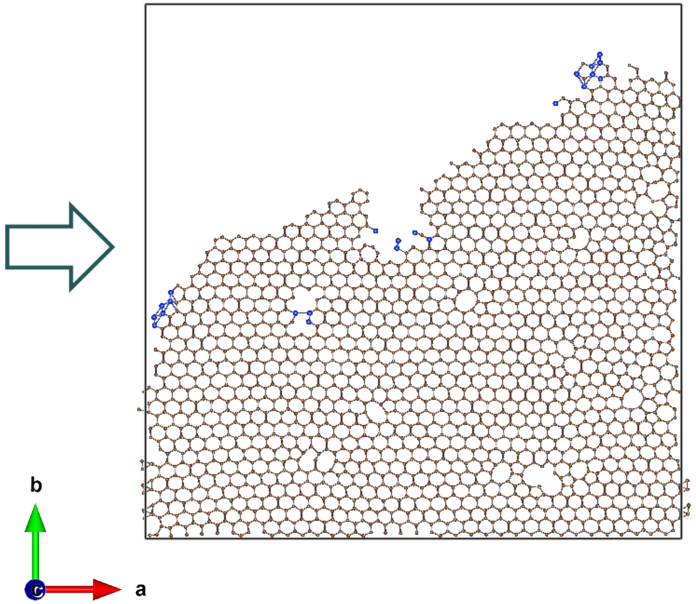
Let's start going through some notebooks

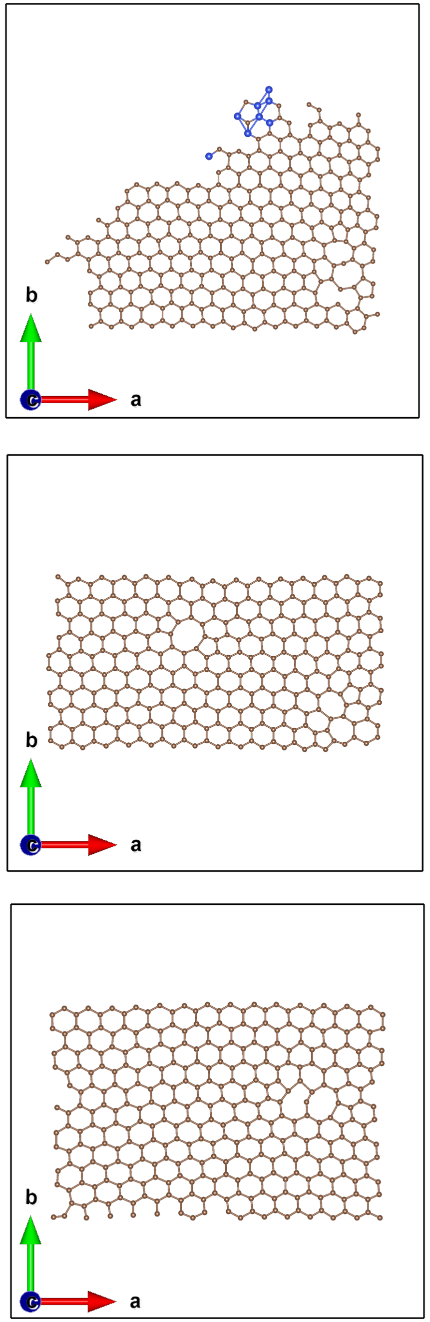
- Piping NN outputs to perform atomistic simulations (Notebook_III)
- Performing simulations (Optional: Notebook_IV, Notebook_V)

https://github.com/SergeiVKalinin/ML-ElectronMicroscopy-2023

Workflow: STEM image → feature finding → simulation object







Demo workflow notebook reference:

https://github.com/SergeiVKalinin/ML-ElectronMicroscopy-2023

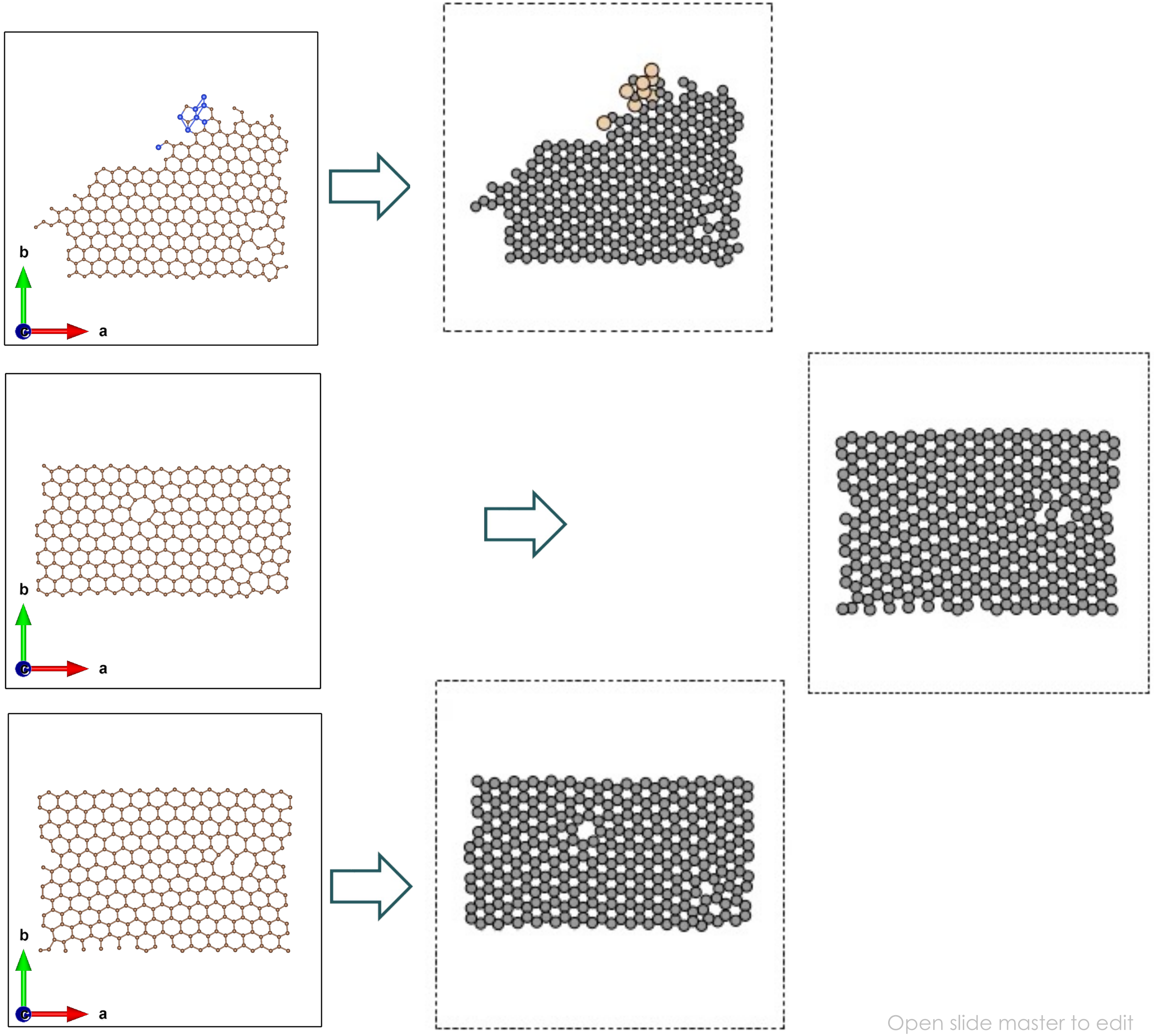
Workflow: example simulations

Example simulations:

- **Standard MD simulations with Tersoff potential for 200 fs with 1 fs each time step at T=300 K**
- **No geometry optimization here**

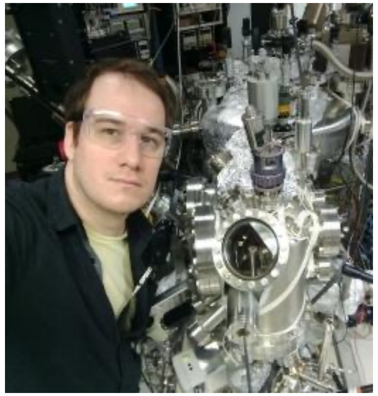
Demo workflow notebook reference:

https://github.com/SergeiVKalinin/ML-ElectronMicroscopy-2023



Acknowledgements

Dr. Sergei V. Kalinin (UTK)



Dr. Maxim Ziatdinov (ORNL)

Dr. Bobby G. Sumpter (ORNL)

Dr. Stephan Irle (ORNL)

Dr. Rama Vasudevan (ORNL)

Virtual summer school on Machine Learning in Electron Microscopy 2023