**Machine Learning for Materials Design**

MS43040 - Project C

Group 12

**Atom Detection in Microscopy Images**  
TU Delft  
18th April 2025

Preyansh Srivastava (6057519)

Uku Luhari (6071740)

# Table of Contents

1. Introduction

1.1 Single-Atom Catalysts and Motivation

1.2 Importance of Atom Arrangement in SACs

1.3 Challenges in Studying SACs

2. Project Goals

3. Literature Review

3.1 Machine Learning in Materials Microscopy

3.2 Challenges in Atom Detection

4. Dataset Description

4.1 Class Labels for Small Image Crops

4.2 Expert-Labeled Atomic Coordinates for Full-Size Images

5. Model Architecture and Approach

5.1 CNN Model Description

5.2 Layer-by-Layer Overview

5.3 Loss Function and Activation Functions

5.4 Dimension Flow Summary

6. Theoretical Framework

6.1 Image Classification Using CNNs

6.2 Sliding Window Detection

6.3 Handling Imbalanced Data

6.4 Evaluation Metrics

6.5 Ground Truth and Expert Annotation

7. Results

7.1 Platinum

7.2 Nickel

8. Conclusion

# 1. Introduction

## 1.1 Single-Atom Catalysts and Motivation

Single-atom catalysts (SACs) are becoming an important class of materials, especially for applications related to green energy and sustainable chemical processes. Unlike conventional catalysts, which usually involve clusters of metal atoms or larger particles, SACs consist of individual metal atoms spread out across a support material like functionalized carbon or metal oxides.

There are a few major reasons why SACs are getting so much attention lately:

1. **Better Catalytic Efficiency**

* In SACs, every metal atom is exposed and active, leading to very high atomic efficiency.
* This typically makes them more catalytically active than catalysts based on nanoparticles.

1. **Adjustable Selectivity**

* The local environment around each atom can be fine-tuned, allowing researchers to favor particular reactions.
* Changing the way atoms are supported changes how they interact with molecules.

1. **Cost and Sustainability Advantages**

* Many important catalysts use expensive metals like platinum (Pt), palladium (Pd), or nickel (Ni).
* SACs use much less metal, making them both cheaper and more sustainable over the long term.

1. **Learning About Structure-Performance Links**

* Even small changes in how atoms are arranged can have big effects on performance.
* Understanding these relationships could help design much better catalysts in the future.

## 1.2 Importance of Atom Arrangement in SACs

Although SACs are based on isolated atoms, how those atoms are placed on the support still has a big effect. A few key points:

* **Nearby Atoms Can Interact**: Even though atoms are supposed to be isolated, if they’re close enough, they can still affect each other’s behavior.
* **Small Differences, Big Changes**: Tiny shifts in the distances between atoms can change catalytic activity and selectivity.
* **Long-Term Stability**: If atoms are positioned poorly, they might cluster together over time, which would reduce the catalyst’s effectiveness.

Because of all this, it’s critical to accurately figure out exactly where the atoms are located — and that’s where advanced microscopy techniques come in.

## 1.3 Challenges in Studying SACs

Aberration-Corrected Scanning Transmission Electron Microscopy (AC-STEM) is one of the main techniques used to visualize individual atoms on SACs. But analyzing these images manually isn't easy:

1. **Extremely Time-Consuming**

* Researchers often need to look through thousands of images by hand.
* Finding and labeling atoms in each image takes a huge amount of time.

1. **Errors and Inconsistencies**

* Different people might interpret the same image differently.
* Human mistakes and biases can make the results less reliable.

1. **Scaling Problems**

* As datasets get bigger, manual analysis just isn't practical anymore.
* New, automated methods are urgently needed.

**How Machine Learning Can Help**

To solve these issues, this project aims to develop a machine learning model that can automatically detect metal atoms in microscopy images. Using machine learning would bring several advantages:

* **Speed**: A well-trained model could process thousands of images in just a few seconds.
* **Consistency**: Automated detection avoids the subjectivity and errors that come with manual analysis.
* **Scalability**: It would make it possible to analyze large datasets much more easily.

# 2. Project Goals

1. Build a machine learning classifier that can decide whether a small image patch (21px × 21px) contains a metal atom or not.
2. Apply the classifier to full-size microscopy images (1024px × 1024px) to find the exact locations (x, y coordinates) of the atoms.
3. Test the model’s accuracy by comparing its predictions to expert-annotated data.
4. Extend the model so it can recognize different metals, particularly platinum (Pt) and nickel (Ni).

# 3. Literature Review

## 3.1 Machine Learning in Materials Microscopy

Recently, machine learning (ML) has become much more than just a tool for exploring materials microscopy—it’s now a key part of how images are analyzed. In electron microscopy especially, ML is being used for a wide range of tasks, including cleaning up noisy images, separating different regions, sorting images into categories, and spotting objects of interest.

Mitchell et al. (2021) were among the first to show that algorithms, when guided by domain-specific knowledge, could successfully pick out individual atoms in microscopy images. However, their methods didn’t adapt very well when the imaging conditions changed.

Later, Ni et al. (2023) made a big leap by applying deep learning, particularly convolutional neural networks (CNNs), to automate the detection of isolated catalyst atoms. By training on datasets with known labels, they managed to cut down the time needed for analysis and made the system more resistant to common imaging problems. They also stressed how important it is to use high-quality data and to preprocess it carefully to get the best results.

Rossi et al. (2024) took things even further by combining machine learning with spatial statistical analysis. Instead of just finding where the atoms were, they studied how atoms are clustered and what that means for their catalytic activity. Their work showed that ML could help not only with identifying atoms but also with understanding the bigger picture of how atomic arrangements affect performance.

## 3.2 Challenges in Atom Detection

Even with these improvements, detecting atoms in scanning transmission electron microscopy (STEM) images is still not straightforward. Some of the main challenges include:

1. **Noise and Contrast Differences**: Imaging conditions aren’t always the same—background noise, brightness, and contrast can all change from session to session. This makes it hard to create a detection model that works well across different images.
2. **Uneven Labels**: In most datasets, there are way more background pixels than atom pixels. This imbalance can mess with the model training and make it harder to detect atoms accurately.
3. **Visual Uncertainty**: Sometimes, even trained experts have trouble seeing atoms clearly, especially if they are partially buried or located near complex structures.
4. **Dataset-Specific Models**: A model trained to find platinum atoms might completely fail when looking for nickel atoms because of differences in their appearance and the textures around them.

Overcoming these obstacles will require better model designs, but also smart strategies for preparing and augmenting data, as well as post-processing the results.

# 4. Dataset Description

The dataset includes multiple CSV files that serve different purposes:

1. **Class Labels for Small Image Crops (for training/testing)**
2. **Expert-Labeled Atomic Coordinates for Full-Size Images (for model validation)**

Each file type is structured to provide essential information for training and evaluating the machine learning model.

## 4.1 Class Labels for Small Image Crops

These CSV files contain **labels** indicating whether a small (21px x 21px) image crop contains a metal atom (**positive: 1**) or does not (**negative: 0**).

**Files:**

* Pt-atom-crop.csv → Labels for Platinum (Pt) dataset.
* Ni-atom-crop.csv → Labels for Nickel (Ni) dataset.

**Structure:**

| **Image Name** | **Label (0 or 1)** |
| --- | --- |
| Pt\_crop\_001.png | 1 (metal atom present) |
| Pt\_crop\_002.png | 0 (no metal atom) |
| Ni\_crop\_010.png | 1 (metal atom present) |

**Usage:**

* These CSV files are **used during model training** to classify whether a given 21px x 21px image contains a metal atom.
* The machine learning model will be trained to **predict the label (0 or 1)** based on the image input.

## 4.2 Expert-Labeled Atomic Coordinates for Full-Size Images

These CSV files contain **manually annotated** (x, y) coordinates of metal atoms in full-size AC-STEM microscopy images. The coordinates represent the **ground truth** against which the ML model’s predictions will be compared.

**Files:**

* **For Platinum (Pt) Images:**
  + Pt-SAC-HX\_0007.csv → Coordinates for Pt-SAC-HX\_0007.png
  + Pt-SAC-HX\_0010.csv → Coordinates for Pt-SAC-HX\_0010.png
* **For Nickel (Ni) Images:**
  + Ni-HX\_0006.csv → Coordinates for Ni-HX\_0006.png

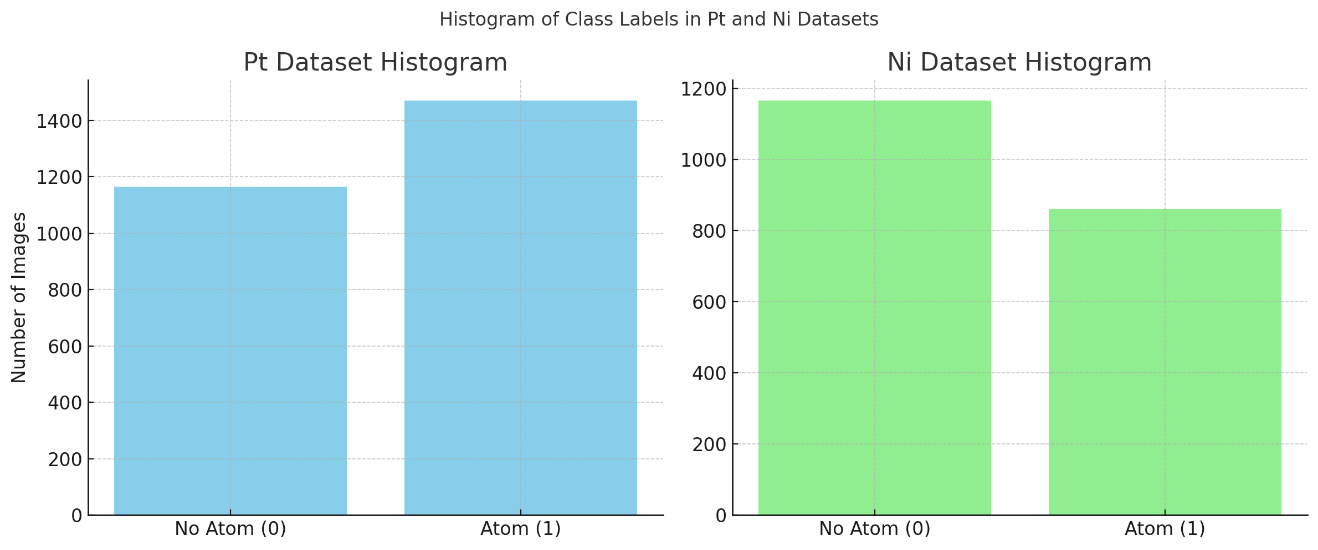
**Structure:**

| **x-coordinate** | **y-coordinate** |
| --- | --- |
| 125 | 300 |
| 678 | 432 |
| 950 | 784 |

Each row represents the **(x, y) position of a detected metal atom** in the full-size microscopy image.

**Usage:**

* The ML model will generate a **predicted list of (x, y) coordinates** for metal atoms in full-size images.
* These predicted coordinates will be compared to the **expert-labeled ground truth** from the CSV files.
* Performance metrics like **precision, recall, and localization error** will be calculated based on this comparison.

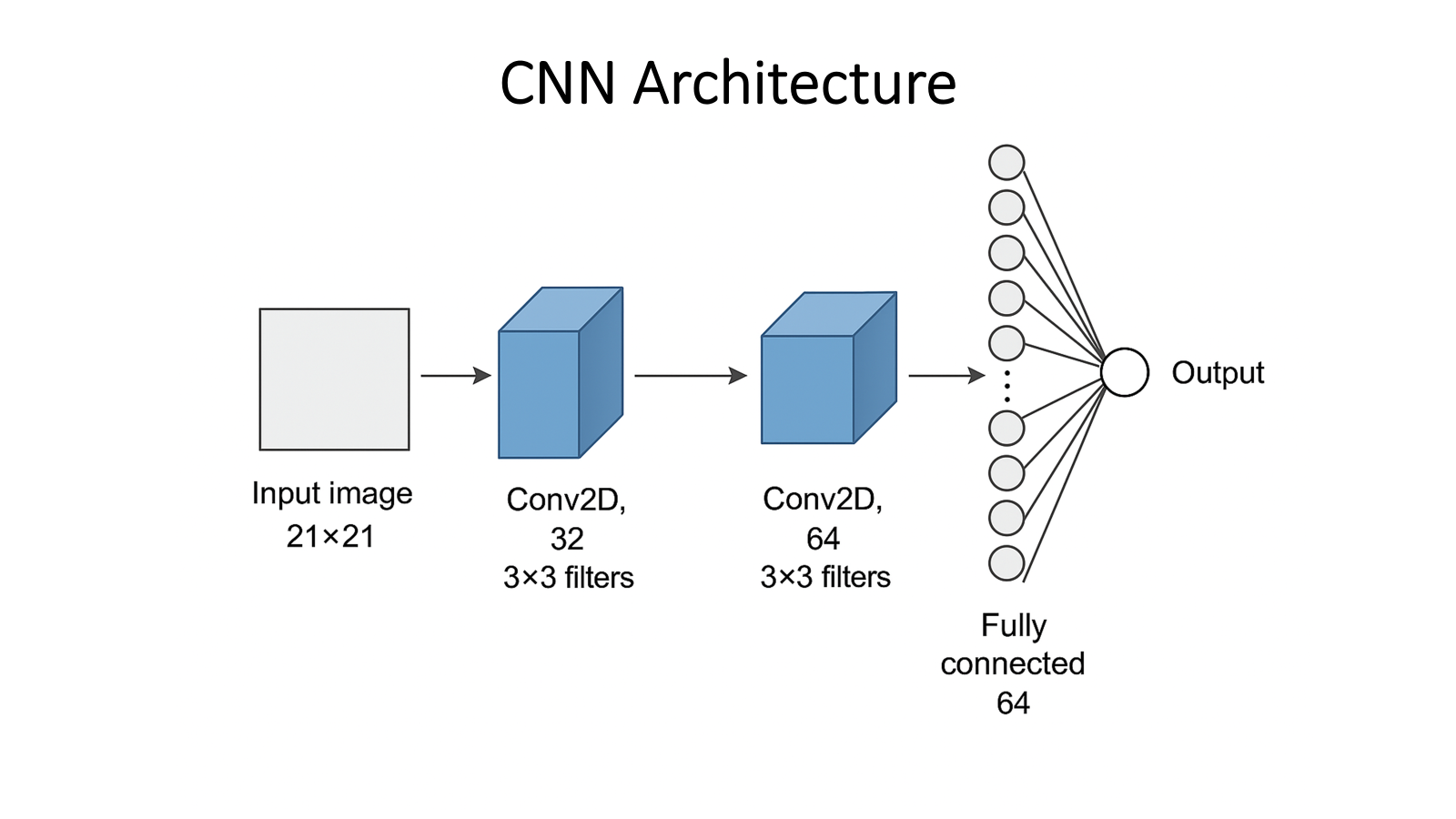


# 5. Model Architecture and Approach

This CNN model is designed for detecting the presence of a single atom in 21×21 pixel grayscale AC-STEM image patches.

## 5.1 CNN Model Description

Layer-by-layer description:  
- Input: Grayscale image (1, 21, 21)  
- Conv2D (1→32 filters, 3×3 kernel, stride=1, padding=1) → Output: (32, 21, 21)  
- ReLU activation  
- Conv2D (32→64 filters, 3×3 kernel, stride=1, padding=1) → Output: (64, 21, 21)  
- ReLU activation  
- AdaptiveAvgPool2d (Output: (64, 5, 5))  
- Flatten to (1600,)  
- Fully Connected (1600 → 64)  
- ReLU activation  
- Fully Connected (64 → 1)  
- Sigmoid activation → Output probability (0 to 1)



## 5.2 Layer-by-Layer Overview

Each convolutional layer preserves spatial dimensions due to padding. Adaptive average pooling forces the feature maps to a 5×5 size regardless of the input size, preparing features for the fully connected layers.

## 5.3 Loss Function and Activation Functions

Since the task is binary classification, binary cross-entropy loss (BCELoss) is applied.  
  
Loss Formula:  
Loss = -(1/N) \* Σ [ yᵢ log(ŷᵢ) + (1 - yᵢ) log(1 - ŷᵢ) ]

- yᵢ is the ground truth label (0 or 1)  
- ŷᵢ is the model's predicted probability  
  
This loss strongly penalizes incorrect confident predictions.  
  
- ReLU activation: Introduced after convolutional and first fully connected layers to introduce non-linearity.  
- Sigmoid activation: Used in the output layer to interpret the model's prediction as a probability score between 0 and 1.  
  
Adaptive pooling ensures consistent output dimensions before entering fully connected layers, improving robustness against slight variations in input image sizes and reducing parameter counts before the classifier.

## 5.4 Dimension Flow Summary

Input (1, 21, 21) → Conv2D (32, 21, 21) → Conv2D (64, 21, 21) → AdaptiveAvgPool2d (64, 5, 5) → Flatten (1600) → FC1 (64) → FC2 (1) → Sigmoid output (probability).

# 6. Theoretical Framework

## 6.1 Image Classification Using CNNs

CNNs have become the backbone of image recognition tasks in computer vision. Their architecture is inherently suited to spatially correlated data like images. A typical CNN model for binary image classification comprises the following layers:  
  
- \*\*Convolutional Layers\*\*: These layers apply a set of learnable filters to the input image to extract local features. The filters are able to capture structures like edges, corners, and blobs which are essential in distinguishing atoms from background noise.  
- \*\*Activation Functions\*\*: Functions like ReLU (Rectified Linear Unit) introduce non-linearity, enabling the network to learn complex patterns.  
- \*\*Pooling Layers\*\*: These reduce the dimensionality of the feature maps, enhancing computational efficiency while retaining crucial spatial information.  
- \*\*Fully Connected Layers\*\*: These act as the decision-making part of the network, aggregating the spatial features and mapping them to classification outputs.  
- \*\*Output Layer\*\*: For binary classification, a sigmoid activation function is used, producing a probability score indicating the presence of an atom.  
  
The model is trained by minimizing a loss function, such as binary cross-entropy, over a large number of labeled examples. During training, weights are updated using backpropagation and gradient descent to improve classification accuracy.

## 6.2 Sliding Window Detection

While CNNs perform well on cropped image patches, real-world application requires scaling up the model to detect atoms across full-sized (1024×1024) STEM images. This is achieved using a sliding window approach. The trained CNN is applied systematically across the image using a fixed stride, generating a heat map of prediction probabilities.  
  
This heat map is further processed using non-maximum suppression (NMS) or local maxima detection to isolate the most probable atom locations. Thresholding is applied to filter out low-confidence predictions. This approach enables comprehensive and automated atom detection over large images without the need for pixel-wise labeling.

## 6.3 Handling Imbalanced Data

Class imbalance is a prevalent issue in scientific imaging datasets. When a majority of training samples belong to the negative class (no atom), a naive classifier may become biased toward predicting the majority class. To address this, several strategies can be used:  
  
- \*\*Weighted Loss Functions\*\*: Assign higher weights to positive samples in the loss calculation to counteract their numerical underrepresentation.  
- \*\*Resampling Techniques\*\*: Either oversample the minority class (e.g., through duplication or synthetic generation using SMOTE) or undersample the majority class to achieve a more balanced distribution.  
- \*\*Augmentation\*\*: Introduce variability in the minority class through data augmentation techniques such as rotation, translation, scaling, or noise injection to expand the training dataset effectively.

## 6.4 Evaluation Metrics

To evaluate model performance in an imbalanced setting, reliance on accuracy alone is misleading. Instead, more informative metrics include:  
  
- \*\*Precision\*\*: The fraction of predicted positive instances that are actually correct.  
- \*\*Recall\*\*: The fraction of true positive instances that the model successfully identified.  
- \*\*F1 Score\*\*: The harmonic mean of precision and recall, representing a balance between them.  
- \*\*Receiver Operating Characteristic (ROC) Curve\*\* and \*\*Area Under Curve (AUC)\*\*: Visual and quantitative tools for evaluating classifier performance across varying thresholds.  
- \*\*Intersection over Union (IoU)\*\*: In full-image detection tasks, this metric compares predicted atom coordinates with ground-truth annotations, providing a measure of spatial accuracy.

## 6.5 Ground Truth and Expert Annotation

Ground truth data for training and evaluation is derived from expert annotations of full-size microscopy images. These annotations include precise (x, y) coordinates of observed metal centers, serving as a reference standard for evaluating model predictions.  
  
During validation, the proximity of predicted coordinates to ground-truth locations is assessed using Euclidean distance thresholds. Predictions falling within a predefined radius are counted as true positives. This method ensures that detection accuracy reflects real-world relevance and aligns with the domain expert's judgement.

# 7. Results

The main goals were to develop models capable of identifying metal atoms in 21x21 pixel images and extend this capability to full-size images. The model was specifically designed to detect nickel and platinum metal centers in single-atom catalysts.

During training, the model was evaluated using various performance metrics, including accuracy, precision, recall, and F1-score. These metrics provided insights into the model's ability to correctly identify metal atoms and minimize false positives and false negatives.

The training results were better in the case of platinum due to the better contrasting atoms on the AC-STEM images and some (barely noticeable) overfitting was seen in the case of nickel. The training losses dropped to ~0.20 for nickel and to ~0.17 for platinum. Platinum's validation loss curve followed the training curve tightly while nickel's validation curve was flattened at loss of around 0.3.

## 7.1 Full micrograph

All the patches delivered by the sliding window method will be analyzed and the suspected atoms locations will be picked out. Most of these predictions turn out to be false and they are removed to prevent incorrect guesses. Potential prediction “clouds” can be seen which correspond to the atom locations on the image. Then, a selection based on brightnesses, cloud shape and the distances between same-cloud atoms is made and the most likely coordinate pair is chosen. This process is repeated for the whole picture and some results are returned. Sometimes the predictions fall exactly onto the ground truth coordinates, but sometimes they fall close, too close for it to be a coincidence. The reason for the very close guess is likely a non-ideal algorithm that can not correctly predict the expert markings (which are guesses on their own). However it can often find “predictions” very close to expert markings and in the code provided, prediction was counted as correct if the distance between the prediction and ground truth was up to 10 pixels (which is also very close to the radius of the metal atoms).   
  
As mentioned earlier, platinum micrographs deliver much better results than those of nickel - mainly due to lower contrast between individual atoms on the nickel image. The platinum model was clearly above the competition with 87% of all final predictions being correct (assuming the 10px proximity threshold) and 82 % of all the correct positions found. Corresponding numbers for nickel were 29 % and 48 %. It was always possible to increase the atom detection threshold which would deliver more correct positions found, but the amount of correct predictions would go down in process - a good balance between these two should be found.

# 8. Conclusion

The machine learning model showed promising results in locating metal atoms in AC-STEM images, particularly for platinum. The model achieved high accuracy and precision for Pt detection, with a significant portion of predictions being correct and a high number of correct positions found. Nickel detection was less ideal, with lower accuracy and precision, and a smaller percentage of correct predictions.

These results suggest that while the model is effective for Pt detection, further improvements are needed for Ni detection. Potential enhancements could include refining the model, improving the full-picture analysis parameters, and exploring different initial coordinate finding methods such as blob or edge detection. Additionally, data augmentation techniques could be employed to improve the model's robustness and accuracy.

Overall, the project demonstrates the feasibility of using machine learning for atom detection in microscopy images, with significant potential for further development and application in the field of materials science.

References

* Mitchell, Sharon, et al. "Automated image analysis for single-atom detection in catalytic materials by transmission electron microscopy." *Journal of the American Chemical Society* 144.18 (2022): 8018-8029.
* Ni, Haoyang, et al. "Quantifying atomically dispersed catalysts using deep learning assisted microscopy." *Nano letters* 23.16 (2023): 7442-7448.
* Rossi, Kevin, et al. "Quantitative Description of Metal Center Organization and Interactions in Single‐Atom Catalysts (Adv. Mater. 5/2024)." *Advanced Materials* 36.5 (2024): 2470038.
* Tieu, Peter, et al. "Directly probing the local coordination, charge state, and stability of single atom catalysts by advanced electron microscopy: a review." *Small* 17.16 (2021): 2006482.
* Spurgeon, Steven R., et al. "Towards data-driven next-generation transmission electron microscopy." *Nature materials* 20.3 (2021): 274-279.