Made available for a limited time for personal research and study only <u>License</u>.

Machine Vision Automated Chiral Molecule Detection and Classification in Molecular Imaging

Jiali Li,[||] Mykola Telychko,[||] Jun Yin, Yixin Zhu, Guangwu Li, Shaotang Song, Haitao Yang, Jing Li, Jishan Wu, Jiong Lu,* and Xiaonan Wang*

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Scanning probe microscopy (SPM) is recognized as an essential characterization tool in a broad range of applications, allowing for real-space atomic imaging of solid surfaces, nanomaterials, and molecular systems. Recently, the imaging of chiral molecular nanostructures via SPM has become a matter of increased scientific and technological interest due to their imminent use as functional platforms in a wide scope of applications, including nonlinear chiroptics, enantioselective catalysis, and enantiospecific sensing. Due to the time-consuming and error-prone image analysis process, a highly efficient analytic framework capable of identifying complex chiral patterns in SPM images is needed. Here, we adopted a state-of-the-art machine vision algorithm to develop a one-image-one-system deep learning framework for the analysis of SPM images. To demonstrate its accuracy and versatility, we employed it to determine the chirality of the molecules comprising two supramolecular self-assemblies with two distinct chiral organization patterns. Our framework accurately detected the position and labeled the chirality of each molecule. This framework underpins the tremendous potential of machine learning algorithms for the automated recognition of complex SPM image patterns in a wide range of research disciplines.
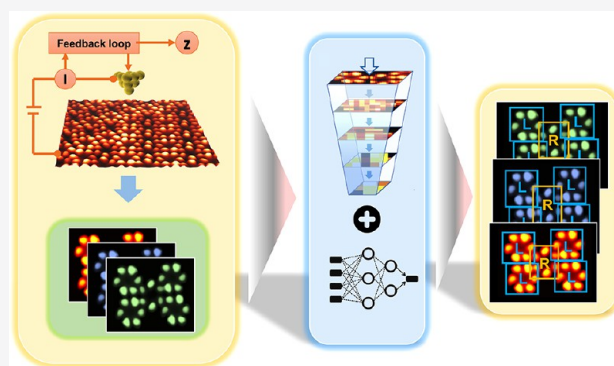
INTRODUCTION

Scanning probe microscopy (SPM), including scanning tunnelling microscopy (STM) and atomic force microscopy (AFM), has been widely used for characterization of the structural and electronic properties of nanomaterials and molecules with an unprecedented spatial resolution.[1−5] A wide range of systems suitable for the SPM imaging span from single atoms,[1] molecules,[2] and their assembled nanostructures[3] to the large biopolymers and complex biological proteins.[4] Recent breakthroughs in the field of SPM imaging are exemplified by visualization of the intramolecular chemical bonds in a variety of molecular functional nanostructures using tip-functionalized AFM and bond-resolved STM imaging techniques.[2] These research advances establish SPM as a highly prominent tool in a broad range of disciplines, including materials science, chemistry, physics, and biology.

Molecular chirality imaging and recognition of supramolecular assemblies are detrimental for their use as functional platforms in a wide range of applications, including enantioselective heterogeneous catalysis,[6] chiral separation,[7] chiroptical materials,[8] and fundamental biochemical processes.[9] Therefore, the study of the chiral organization of the self-assembled molecular structures using SPM is a matter of increased scientific and technological interests.[10] However, recognition of the chiral

patterns of a densely packed molecular structure, revealed by SPM images, is usually a time-consuming and error-prone process, which relies on the ability of a human to analyze, classify, and interpret ultimately fine variations of the STM contrast at the nanoscale. Therefore, it hinges on the need for a highly efficient analytic tool capable of accurate and autonomous recognition of the complex chiral patterns of SPM images, requiring minimal human supervision.

To this end, a number of analytic tools with traditional machine vision as well as machine learning algorithms have been developed to identify molecular patterns in SPM images with reduced human supervision.[11−14] These algorithms typically separate the molecular pattern recognition process into two steps, namely, molecule detection and molecule classification. The molecule detection step is typically achieved using traditional machine vision algorithms, including normalized cross-correlation[15] and scale-space theory.[16] The molecule
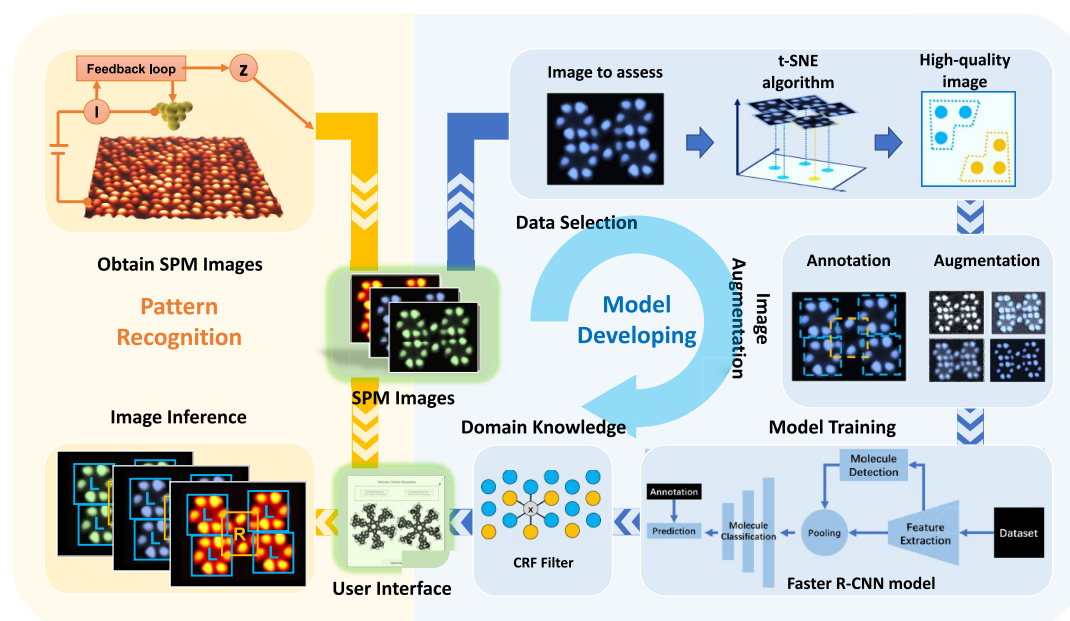
Figure 1. Overview of the automated chiral molecule detection and identification workflow. This flow diagram illustrates the overall process of the study, showing the completion of molecular chirality recognition through a series of modules from an original SPM image. First, the experimental STM images are obtained (where *I* denotes tunneling current and *Z* denotes topography channel). Then image quality is rated by an unsupervised learning method called t-distributed stochastic neighbor embedding (t-SNE). The data with the desired quality is augmented to generate a training data set for the core prediction model. Afterward, an optional conditional random field filter based domain knowledge incorporation module can be utilized to enhance performance for nonideal STM images. Lastly, the trained model is utilized for the image inference tasks for automated pattern analysis of the SPM images.

classification step is attained by other supervised machine learning (ML) methods.[17] However, it is still challenging for traditional machine vision algorithms to detect and classify molecules of complex SPM patterns. The substantial drawback of the traditional machine vision approaches is that different algorithms and model parameters are required for different molecular patterns to achieve the desired performance, and a long trial-and-error process is required to determine viable algorithms and parameters.[18,19] For example, the normalized cross-correlation algorithm cannot accurately detect the position of each molecule in a complex densely packed system due to the similarity between interspaces and molecules. Besides, the quality and resolution of the SPM images can vary significantly even for the same molecular system, depending on an intimate interplay between the electronic and structural properties of the sample, as well as the status of scanning probe and instrument parameters applied for image acquisition (e.g., bias voltage, tunneling current, or frequency shift set point). Furthermore, the acquisition of high-quality SPM images of complex supramolecular self-assemblies is normally time-consuming, so the data acquisition rate is limited, making the machine vision model fall short of data and, thus, present limited performance. As a result, it is of high value to address these challenges through a more accurate, robust, and data-efficient deep learning framework.

Here, we develop a one-image−one-system deep learning framework consisting of a state-of-the-art deep learning machine vision algorithm, a uniquely designed data selection method, and effective data augmentation techniques to provide a versatile solution capable of the automated recognition of a complex SPM pattern. Our framework uniquely combines a designed data selection method and data augmentation techniques,[20] which not only allows for a highly accurate molecule detection and classification by a regional-based convolutional neural network (Faster R-CNN) but also reduces demand for large data sets to one single SPM image for each molecular system. We then showcase our framework to analyze the STM images of the two densely packed molecular self-assemblies with different chiral recognition patterns. Despite ultimately fine variations of the STM contrast of the individual molecules with different chirality and the varying quality of the STM images, our framework automatedly labeled the chirality configuration of each molecule within self-assemblies with unprecedented accuracy and robustness. Moreover, the generality of our framework is illustrated by analyzing nonideal STM images, and an additional domain knowledge incorporation module is used to further improve the generalization capability. Our results underpin the tremendous potential of our one-image−one-system deep learning framework for highly accurate detection and classification of complex SPM patterns with a limited data set.

■ RESULTS

Molecular Systems. As a proof-of-concept demonstration of the generality of our framework, we selected two different supramolecular assemblies consisting of star-shaped hexadimethylphenylbenzene (HPB) and fluorine-substituted hexadimethylphenylbenzene (F-HPB) molecules (Figure S1), respectively. Both HPB and F-HPB molecules adopt a highly nonplanar configuration due to a steric-hindrance between dimethylphenyl rings (Figure S1 in this study and see also Figure 3 and Figure S7 in ref 21). The steric orientations of dimethylphenyl rings render a prochiral left-handed (L) or right-handed (R) enantiomeric configurations upon their adsorption on Ag(111) (Figure S1). Furthermore, the different chemical composition of HPB and F-HPB molecules leads to
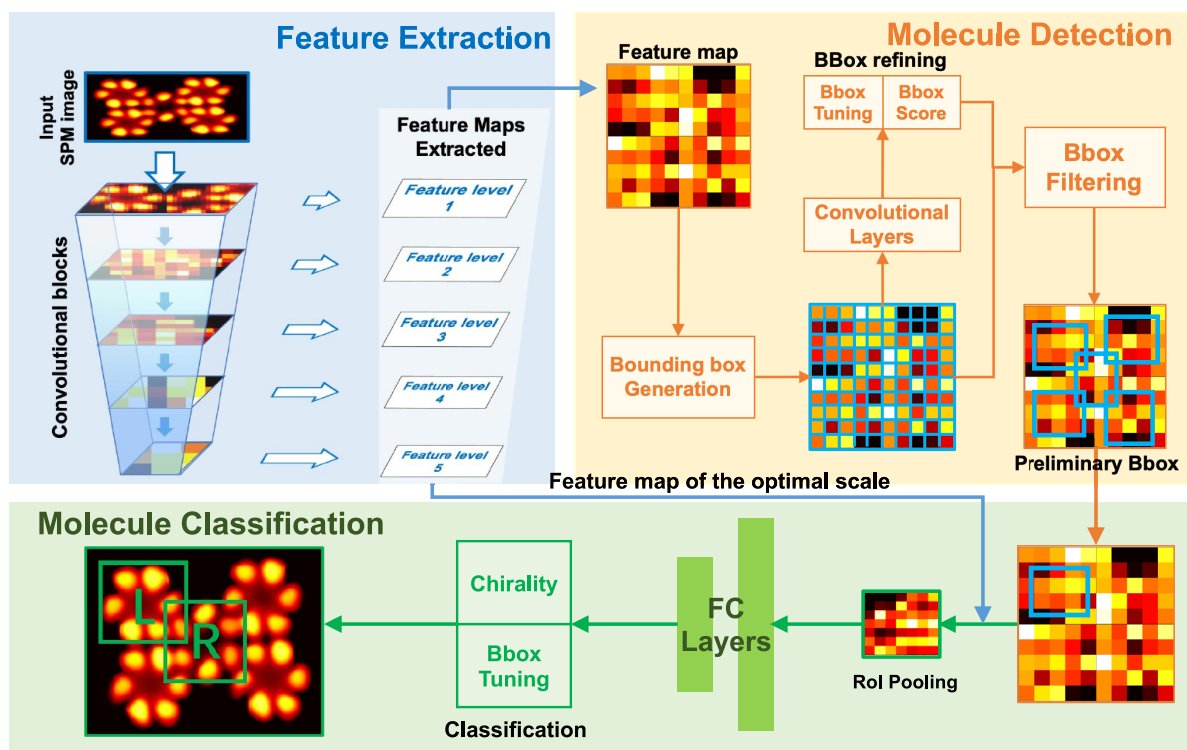
Figure 2. Schematics of the Model Training module, which consists of three main stages: Feature Extraction (blue block), Molecule Detection (orange block), and Molecule Classification (green block). Bounding box: boxes around objects being checked for a collision.

remarkably distinct chiral organization patterns of their self-assemblies.

The STM imaging reveals that both HPB and F-HPB molecules exhibit a 6-fold symmetry manifested by six protrusions arranged in a hexagonal-like manner, which are attributed to the six methyl groups attached to dimethylphenyl rings. It is noted that individual HPB and F-HPB molecules exhibit expected clockwise or anticlockwise appearances (Figure S1) attributed to the L- or R-handed enantiomeric configurations, respectively. The chiral configurations of molecules (i.e., L or R) are manifested by sub-Å variations of the STM contrast. Therefore, the detection and classification of individual molecules are formidably challenging due to barely discernible contrast variations and the ambiguous borderline of each molecule residing in the densely packed self-assembly. The highly relevant and complex problems of the chirality recognition in molecular self-assemblies and ML-driven SPM image analysis make HPB and F-HPB self-assemblies ideal playgrounds to showcase abilities and potential of the developed one-image—one-system deep learning framework in automating the molecule detection and chirality classification of the STM images, which we further introduce below.

Overall Framework and Model Architecture. We attain the ML-driven chirality recognition of the two molecular self-assemblies (HPB and F-HPB) through the combination of the four intelligent modules, namely, Data Selection, Data Augmentation, Model Training, and Domain Knowledge incorporation as shown schematically in Figure 1.

Module 1. Data Selection. The first module selects high-quality STM images for training. A t-distributed stochastic neighbor embedding (t-SNE) algorithm is adopted for unsupervised data quality rating and visualization. In our chirality recognition problem, the STM image is classified as a high-quality one if a clear difference associated with distinct

molecule chirality can be revealed by t-SNE. The details of data selection are introduced in the Training Data Selection section. At this stage, the low-quality STM images are discarded for better training purposes due to their adverse impact on the model's performance.

Module 2. Data Augmentation. The second module generates a large data set from the limited high-quality experimental STM images. A combination of effective augmentation techniques is chosen such that the most optimized model's performance is achieved, as explained in detail in the Data Augmentation section.

Module 3. Model Training. The core prediction model to realize molecule detection and chirality classification is the Faster Region-based Convolutional Neural Network[22] (Faster R-CNN), which is an object detection architecture widely used in the ML field for the target object detection and classification. A general block chat diagram of this core model (illustrated in Figure 2) consists of three stages, namely, feature extraction, molecule detection, and molecule classification. The details of the model's working principle are discussed in the Experimental Section. A brief explanation of each stage is also provided below:

- Feature extraction stage (the first stage) aims to extract useful abstract features of STM images. The abstract features contain information including edges, textures, and shapes of the molecules in STM images and, thus, can be used for molecule detection and classification. This process is highlighted by the blue block in Figure 2.

- Molecule detection stage (the second stage) can then detect the preliminary bounding boxes, which indicate the position and size of each molecule in STM images, based on the abstract features extracted from the first stage as shown in the orange block in Figure 2. These bounding
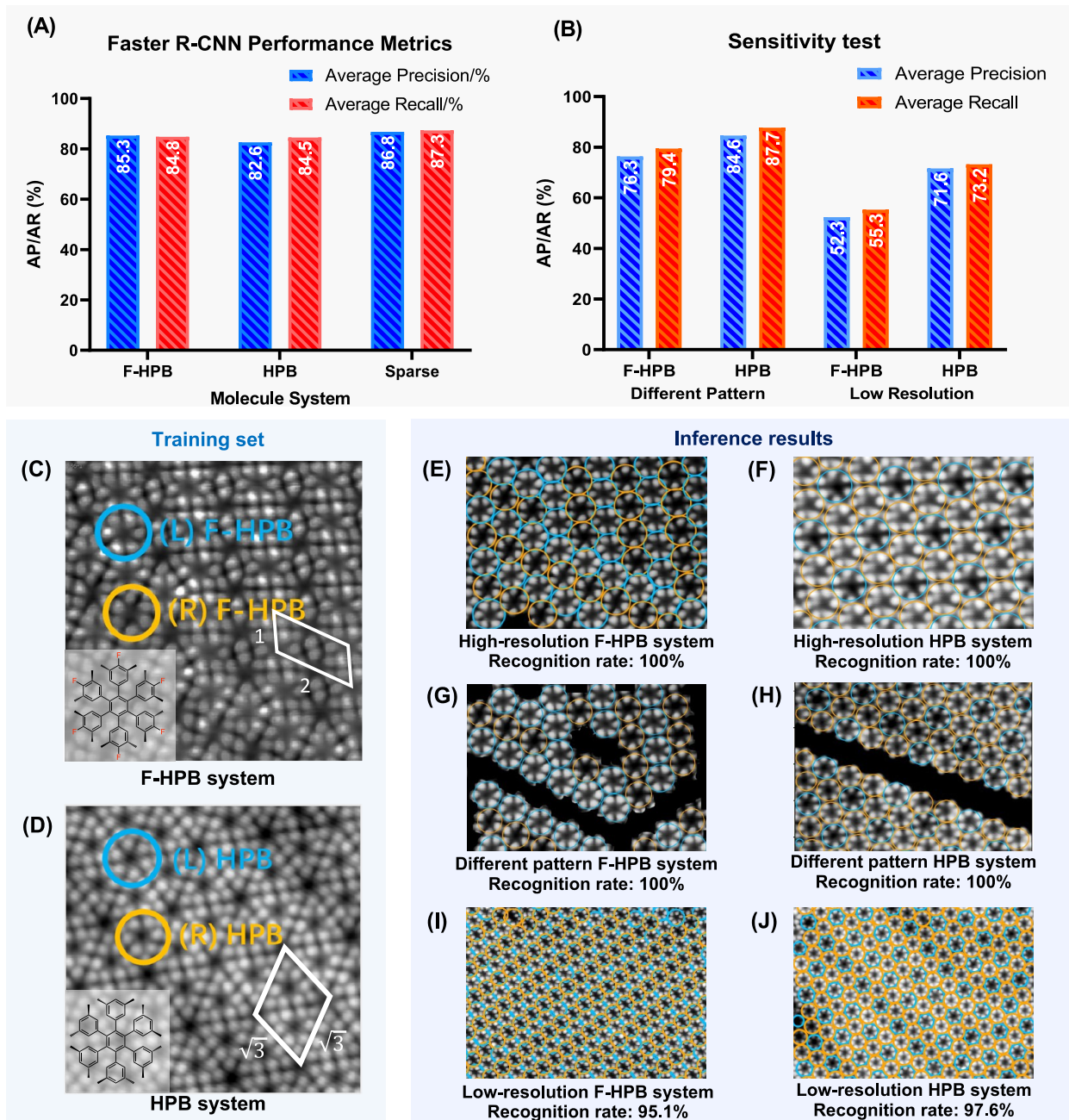
(A) Faster R-CNN Performance Metrics
- Average Precision/%
- Average Recall/%

(B) Sensitivity test
- Average Precision
- Average Recall

(C) Training set — F-HPB system

(D) HPB system

Inference results

(E) High-resolution F-HPB system
Recognition rate: 100%

(F) High-resolution HPB system
Recognition rate: 100%

(G) Different pattern F-HPB system
Recognition rate: 100%

(H) Different pattern HPB system
Recognition rate: 100%

(I) Low-resolution F-HPB system
Recognition rate: 95.1%

(J) Low-resolution HPB system
Recognition rate: 97.6%

Figure 3. Overall system performance and robustness analysis. (A) Performance metrics of the Faster R-CNN applied to three target systems. (B) Robustness test performed on F-HPB and HPB systems. Model's performance on STM images with low resolution and the different pattern was tested. (C, D) Original experimental STM images of each molecular system, which were used to train Faster R-CNN models. The molecular structures of F-HPB and HPB with their unit cells are indicated in the insets of figures. (E−J) The inference results were obtained upon analysis of high-resolution (E, F), different pattern (G, H), and low-resolution (I, J) experimental STM images from the two trained Faster R-CNN models. Blue and orange circular markers represent L-handed and R-handed molecules, respectively. The bounding boxes are converted from rectangles as per model outputs into their inscribed circles to fit into the shape of molecules.

boxes are fine-tuned and filtered to yield more accurate results.

- Molecule classification stage (the third stage) finally classifies each molecule into its corresponding chirality class (L- or R-chirality) according to the abstract features extracted from the first stage and the position of each molecule detected by the second stage. It also fine-tunes the position of each molecule, as shown in the green region of Figure 2.

Module 4. Domain Knowledge Incorporation. It is an additional stage to enhance model performance on "nonideal" STM images containing scanning artifacts, edges of self-assemblies, or surface steps, by utilizing the conditional random field method. It is not a necessary component when the framework is used only for post-analysis of high-quality STM images. However, it allows reduction of the false positive rate and improvement of the recognition rate during analysis of nonideal STM images. The key results of this additional module are discussed for nonideal STM images in Figure 6.
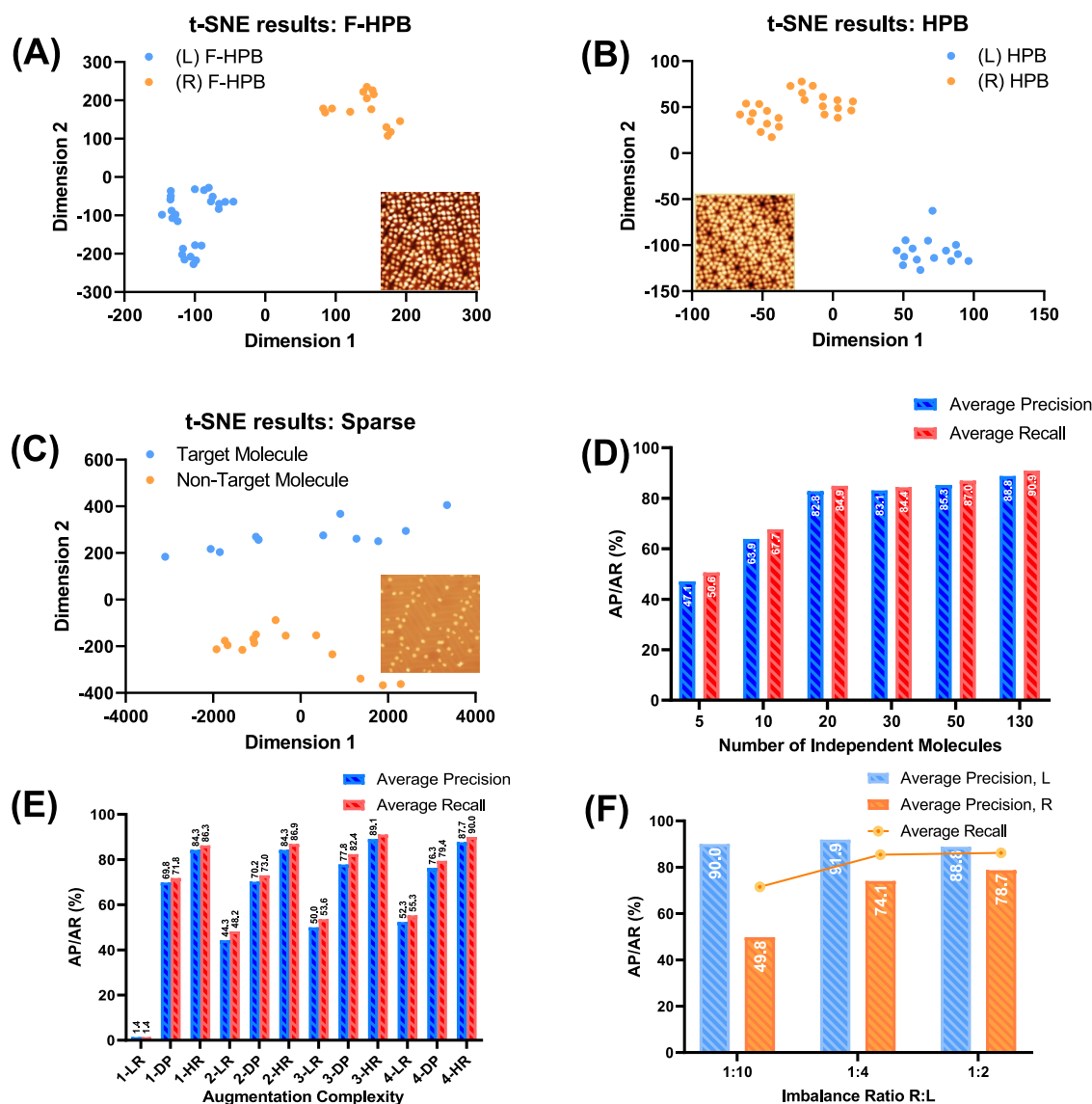
Figure 4. Training data selection and data augmentation performance. (A−C) t-SNE results of F-HPB (A), HPB (B), and sparsely packed (B) molecular system obtained from images of single molecule cropped from the STM images used for a model training. The clear boundary between two clusters in the t-SNE results of both the F-HPB and HPB systems indicates a significant distinction between prochiral molecules. The results of the sparsely packed system show a clear boundary between two clusters (target molecules and nontarget molecules) in t-SNE results, which indicates that the selected STM image is suitable for model training. (D) Correlation between the performance metrics of Faster R-CNN models for the F-HPB system and the number of independent molecules in an STM image used for a data set generation. (e) Correlation between the complexity of augmentation techniques used for data sets generation and the performance metrics of Faster R-CNN models for the F-HPB system tested on STM images in different conditions. LR, DP, and HR represent low resolution, different patterns, and high resolution, respectively. Prefix numbers 1−4 denote augmentation complexity level. (e.g., 1-LR represents the model trained on the data set generated by augmentation complexity 1 tested on low resolution STM image). Augmentation complexity is dependent on the number of image augmentation techniques on an incremental basis used for data set generation. More details can be found in Supplementary Note S5. (F) Performance of Faster R-CNN models against data set with different ratios of R- and L-handed molecules.

After completing the model training stage, arbitrary STM images with various resolutions containing hundreds of prochiral molecules can be processed automatically with high accuracy. Sample output is highlighted by the green block in Figure 2, wherein each molecule is labeled by its bounding box and corresponding chiral category (L or R). Further, as shown in Figure 1, well-trained models can be incorporated into the software with user interface (UI) designed for users' convenience. More detailed information is provided in Supplementary Note S14. The source code and developed software can be found in the Zenodo database.

System's Performance. In this study, a well-trained automated analytic tool can be obtained within a short time frame (a few hours) by using our one-image−one-system deep learning framework. The core prediction Faster R-CNN model was trained on a set of 1000 STM images generated by data augmentation from only one high-quality STM image containing approximately 50 molecules as shown in Figure 3C,D. The inference time for chiral pattern recognition is notably low, i.e., 0.1 s per STM image. Figure 3A shows the performance metrics of the trained Faster R-CNN models applied to the analysis of the high-resolution experimental STM

images of F-HPB and HPB densely packed self-assemblies. In this study, an STM image is classified as a high-resolution image if one molecule occupies a 50×50 pixels area in the image. For the F-HPB system, the model trained on the data set generated from one single experimental high-quality STM image containing 53 independent molecules (see Table S1) achieves an average precision (AP) of 85.3% and an average recall (AR) of 84.8%. The AP and AR are performance metrics that are commonly used in object detection[23] for a quantitative evaluation of the accuracy and precision of both bounding box and classification results. Here, AP is defined as the percentage of the correct chirality predictions among all predictions, while AR reflects a percentage of the correct bounding box and chirality predictions among all molecules in an STM image. The detailed definition of the AP and AR can be found in Supplementary Note S2.

In addition, we evaluated the model's performance based on recognition rate, which is defined as the percentage of correctly recognized molecules from visual inspection, as shown in Figure 3E−J where each recognized molecule is overlaid by a circle (blue or orange) depending on chirality type (L or R). Figure 3E,F shows that the recognition rate of high-resolution STM images is exceptionally high, namely, 100%, which means that all molecules in the STM images can be correctly detected and classified.

Robustness Performance. *Low Resolution.* After the success of our model in chirality recognition of high-resolution STM images, we evaluated its robustness upon analysis of the low-resolution STM images, wherein one molecule occupies about 30×30 pixels area. Surprisingly, we found that despite a low-resolution of the STM images, the recognition rates of F-HPB and HPB systems are still remarkably high, namely, 95.1% and 97.6%, respectively, Figure 3I,J, although the model achieves AP of 52.3% in the F-HPB system, presumably due to imperceptible bounding box error. As a result, the molecular patterns in low-resolution images are still accurately decoded by the Faster R-CNN models trained on the high-resolution STM images.

Molecular Pattern. Next, to demonstrate that the recognition accuracy of our Faster R-CNN approach can be applied for classification of any arbitrary molecular pattern, i.e., distinct from the patterns in the training data set, a robustness test was performed on STM images containing a domain-boundary, Figure 3G,H. The sensitivity test results in Figure 3B suggest that the AP and AR accuracy of the Faster R-CNN model is insensitive to the different molecular patterns. Estimated inferencing results obtained upon analysis of STM images of both F-HPB and HPB self-assemblies with different patterns are provided in Figure 3G,H, which shows a high recognition rate of 100%.

A similar test was also performed on STM images of a sparsely packed molecular system shown in Figure S2. The low quality of STM image (a large scan range and low pixelization) precludes classification of the target molecules by human eyes. The model's performance of this system, which was evaluated based on molecule detection accuracy without chirality classification step, reveals a remarkably high AP of 86.8% and an AR of 87.3. The extended inference results of the sparsely packed system are discussed in Supplementary Note S1.

Training Data Selection. As we demonstrated above, the quality of STM images is essential for the training of the Faster R-CNN models, as the model's performance is highly sensitive to the quality of training sets.[24] Specifically, faster R-CNN hardly learns the distinctions between L- and R-molecules when trained on the low-quality STM images. Therefore, high-quality STM images are preferred to achieve the optimal model's performance. A. Krull et al.[11] have developed a high-accuracy CNN-based method to assess the quality of STM images acquired on the bare metal surface. However, the method has not yet been generalized to assess the quality of images of complex systems (i.e., molecular self-assemblies) by a single CNN. Furthermore, such a quality rating CNN requires an extensive data set containing large variations of different quality STM images for training. Here we adapted t-distributed stochastic neighbor embedding (t-SNE)[25] to determine the quality of STM images by inspecting the extent of the distinction between molecules of different chirality in STM images.

t-SNE is an unsupervised machine learning algorithm used for embedding high-dimensional data (including images) into low-dimensional space for a visualization purpose. When images are used as inputs, similar images are clustered in the low-dimensional space while dissimilar images are repelled from each other from visual inspection. Therefore, if the images of prochiral molecules (L- and R-handed) in an STM image can be successfully separated into two clusters by the t-SNE algorithm, there is a significant distinction between molecules possessing different chirality in this STM image and the quality of the STM image is high. On the other hand, if the images of prochiral molecules are assigned to only one cluster after using the t-SNE algorithm, the molecules of different chirality are highly similar and the quality of the STM image is low.

To assess the quality of the training sets used in this study, images of target molecules under interest were cropped from their corresponding experimental STM images and used as t-SNE input. Figure 4A−C shows the t-SNE outputs of the three training sets for the three different molecular systems. As shown in Figure 4A,B, the target molecules' STM images of F-HPB and HPB systems can be separated into two clusters with a clear boundary, and by postanalysis, the two clusters are with different chirality. The two models, trained from the data sets generated from these two STM images, can detect and classify the chirality of F-HPB and HPB molecules with high accuracy. For the sparsely packed molecular system, it is beyond human capability to recognize the chirality of target molecules. However, it could still be possible to distinguish the target molecules from other molecules in this system. Although most of the molecules on the STM images in this system are highly similar to human eyes, the t-SNE results given by Figure 4C suggest that the target molecules show a significant distinction among other molecules. Indeed, the model trained on the STM image of the sparsely packed molecular system can detect the target molecules with high recognition rate as shown in Figure S2. However, one cannot tell the difference among the target molecules to identify the chirality of them from both human inspection and t-SNE results as shown in Figure S3. Thus, it is impossible for Faster R-CNN model, trained on this STM image, to achieve the molecular chirality classification task. Therefore, it is recommended to select higher quality STM images for data set generation if chirality recognition is required. More details of t-SNE and the hyperparameters used in this experiment can be found in Figure S7 and Table S2.

Data Augmentation. Experimental generation of a large data set to train a deep CNN based models is a time-consuming and impractical task. It has been reported that image augmentation can successfully increase the size of data sets and improve the performance of CNN based models in many recent studies.[26,27]
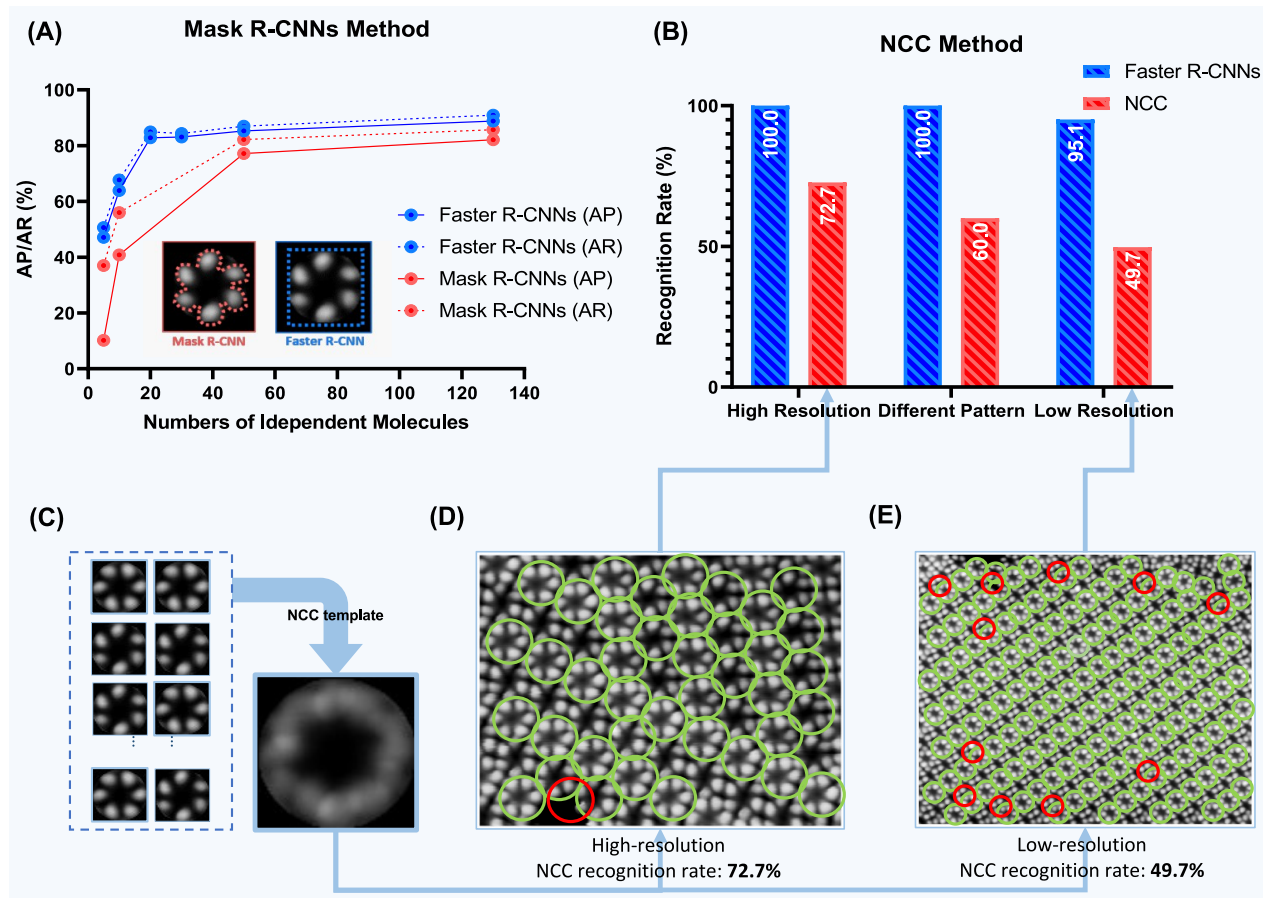
Figure 5. Method comparison. (A) Comparison of the performance metrics of Mask R-CNN and Faster R-CNN models trained by data set generated from the different numbers of independent molecules and tested on a high-resolution experimental STM image of the F-HPB system. (B) Comparison between the recognition rates of the traditional NCC method and Faster R-CNN tested on high-resolution, different pattern, and low-resolution experimental STM images of the F-HPB system. (C) NCC template, generated by an average of images of single molecule cropped from the STM image. (D, E) Inference results of the NCC method on high-resolution and low-resolution STM images of F-HPB system in panel B, where green and red circle markers represent correctly and falsely recognized molecules, respectively. The inference results of Faster R-CNN on the same images can be found in Figure 3.

In consideration of the limited number of experimental SPM images, a feasible method of data set generation is to perform image data augmentation. It is worthy to note that there is still the possibility to work with small data without doing augmentation;[28] however, it is preferred to perform the augmentation when training data is as small as one single image. Therefore, we provide a general guideline on the minimal quantity of experimental STM images and the extent of augmentation required to obtain a well-trained Faster R-CNN model. In our case, the image augmentation is particularly effective in reducing the number of required experimental STM images because the molecules of each class are highly identical (i.e., reveal minor variations), which could be simulated by proper augmentation techniques. These variations, shown in synthetic images, allow Faster R-CNN to extract key features of molecules more precisely and, thus, improve the detection and classification robustness of Faster R-CNN applied to STM images acquired in different conditions.

Since standard methods to quantify the complexity of image augmentation do not exist to date, we generate four data sets from the single high-quality STM image containing 53 independent molecules using different augmentation techniques. The number of augmentation techniques applied to the four data sets is on an incremental basis, which can ensure that

augmentation performed on the subsequent data set is more complex than that on the previous one. As shown in Figure 4E, the model's performance generally improves upon increasing complexity of augmentation used for a data set generation, especially when a model is tested on the low-resolution STM images. While the model, trained on the data set generated from high-quality STM images using augmentation complexity 1, barely recognizes 1.4% of the molecules on the tested low-resolution image, the recognition rate of the same STM image dramatically increases to 95.1% when augmentation complexity of 4 is used (inference results in Figure 3E). More details about combinations of augmentation techniques and augmentation complexity levels can be found in Table S3 and Supplementary Note S5.

It is also noted that the performance of Faster R-CNN is highly sensitive to the number of independent molecules in the original experimental STM image (see Figure 4D). Six Faster R-CNN models were trained on data sets generated by the same combination of data augmentation techniques (complexity 3) from high-quality STM images containing different numbers of molecules. The models were then tested on the high-resolution STM images. The results suggest that the model's performance increases with the number of independent molecules used in the training set as shown in Figure 4D. An experimental STM image
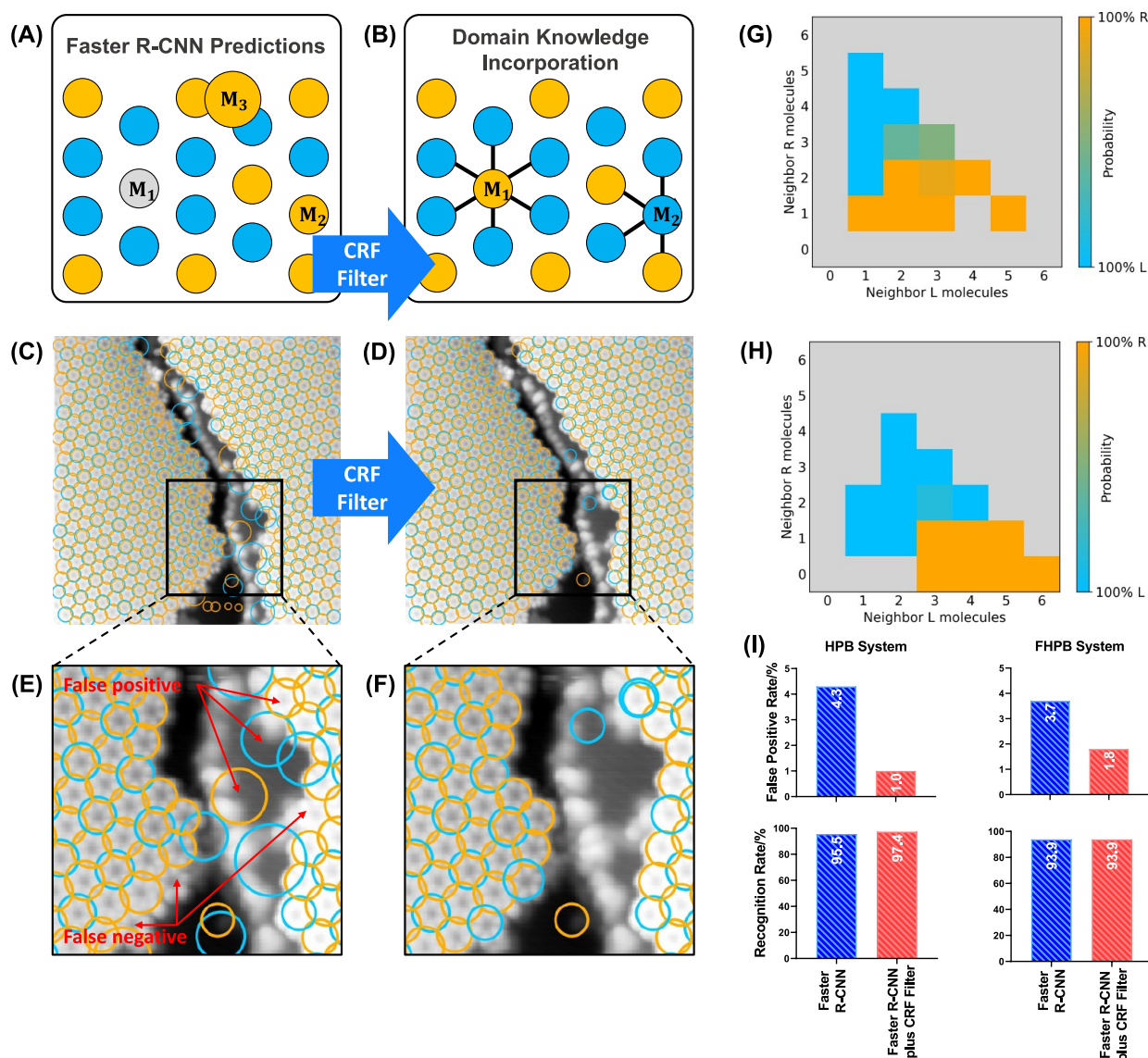
Figure 6. Conditional random field filter. (A) Chirality and position of each molecule predicted by Faster R-CNN. M means predicted molecules. (B) Chirality and position of each molecule updated by a conditional random field filter generated from domain knowledge. (C, D) Inference results of a sample low-resolution STM image of HPB system with steps before and after applying the CRF field. (E, F) Zoomed-in views of (C) and (D), which show the reduction of false positives and false negatives after CRF filter. (G, H) Sample distribution probabilities at different neighboring conditions for FHPB and HPB systems learned from the images. (I) Comparison of recognition rates and false positive rates before and after applying the CRF field on the two systems.

containing more than 20 molecules for each chirality is sufficient to train a desired Faster R-CNN model reaching AP = 82.8 and AP = 84.9.

It should be noted that each Faster R-CNN of different systems was trained on an STM image with the approximately equal number of L and R prochiral molecules. The class balance has a significant impact when developing the training data set samples to get satisfying results.[29] In our study, the number of molecules of different chirality is comparable in both molecular self-assemblies. However, for systems with a minor presence of specific prochiral molecules, it is important to maintain L and R molecule balance in the training set. Otherwise, if the distribution of each class differs substantially, the accuracy of the model would be low when classifying the specific target with a low portion of data due to the imbalance of L and R molecules. The performance metrics of models, trained on the imbalanced

data sets (i.e., L/R ratio of 1:10), are shown in Figure 4F, where the average precision is calculated and shown separately for each chirality. From the figure, the model cannot detect and classify (R) F-HPB molecules in the STM image with high accuracy if (R) F-HPB molecules are in minor presence in the training sets.

Different Machine Vision Methods Comparison. At last, we compare the performance of our framework to two other recently developed methodologies, namely, traditional normalized cross-correlation (NCC) and Mask R-CNN, as discussed below.

Traditional NCC Methodology. First, we employ our F-HPB system to test the performance of the NCC method. Previously, Ziatdinov et al.[17] employed NCC for an accurate analysis of experimental STM images of molecular self-assemblies. As shown in Figure 5C, we use a template constructed from an average of all F-HPB molecules cropped from the STM image

containing 53 molecules to perform normalized cross-correlation on experimental STM images. The results in Figure 5B show that the recognition rate of the NCC method applied to high-resolution STM images of the F-HPB system is unsatisfactory. Inference results of the NCC method on high-resolution and low-resolution STM images are also shown in Figure 5D,E. While our Faster R-CNN approach achieves a recognition rate of 100% and 95.1% on high-resolution and low-resolution STM images, respectively, the NCC method detects only 72.7% and 49.7% of molecules in these two STM images.

Compared to HPB assembly, the STM image of the F-HPB system shows a more complex pattern, which results in a greater difficulty of detection. From the results shown in Figure 5D,E, it is evident that the traditional machine vision-based molecule detection method is prone to confuse backgrounds or interspaces between molecules with a hollow-center shape. The substantial disadvantage of traditional approaches lies in the fact that different algorithms and model parameters are required for different molecular patterns to achieve the desired performance, and a long trial-and-error process is required to decide viable algorithms and parameters.[18] Therefore, the introduction of R-CNN significantly improves the detection rate of molecules with a complex pattern.

Mask R-CNN Methodology. Compared with Faster R-CNN, Mask R-CNN[30] extends Faster R-CNN by adding a parallel branch for predicting an object mask. The object mask indicates the pixels of objects in the bounding box. It is also reported that Mask R-CNN has improved detection pooling.[22] However, results shown in Figure 5A suggest that Faster R-CNN achieved better performance in this experiment with a higher recognition rate. While the Faster R-CNN model is dedicated to predicting the position and chirality of each molecule, Mask R-CNN model yields additional information about molecular masks, which is not useful for molecular pattern recognition and leads to more computational burdens. More details about Mask R-CNN can be found in Figure S9.

Domain Knowledge Incorporation and Generality Test. The domain knowledge of molecular assemblies could also be incorporated into molecule detection and classification. As molecular assemblies would usually adopt one or a few specific patterns, the abnormal predictions from the Faster R-CNN model, which disobey and break the pattern, could be re-evaluated by an additional conditional random field (CRF) filter based on the domain knowledge. While the Faster R-CNN model can already achieve exceptionally high performance on STM images of various conditions as shown in previous experiments, the domain knowledge incorporated in the CRF filter could complement the pure machine learning model and further enhance the overall model performance. The filter would be especially effective on "nonideal" STM images with defects, where the Faster R-CNN could not yield a 100% recognition rate. "Nonideal" STM images are defined as STM images with imperfection features (e.g., artifacts arising from imaging instability or STM images containing the step edges or grain boundaries). In this study, the capability of the CRF filter will be evaluated on a set of defected STM images, as shown in Figure 6.

Figure 6A,B shows the mechanism of the CRF filter. The filter receives the predictions from the Faster R-CNN model and generates a set of distribution probabilities for each molecule to be a certain chirality when its neighbors are observed to be in a certain condition. This step allows the CRF filter to learn the regional molecular patterns within the image. The CRF will be subsequently used to identify the less confident (i.e., both the classification categories are smaller than 0.7) predictions given by the Faster R-CNN model and to correct them based on their neighboring molecules to match the molecular patterns shown on the image. Figure 6C,D provides the inference results of a low-resolution STM image of the HPB system with steps before and after applying the CRF field. Figure 6E,F provides zoomed-in views, which show the reduction of false positives and false negatives after CRF filter.

In addition to the single case shown in Figure 6, many other nonideal STM images as well as STM images of a completely different molecular self-assembly were tested and shown in Supplementary Note S10 and Note S11 respectively. The good performances for both the nonideal scenarios and the completely different molecular system have proved further the generality of our framework. Moreover, the average recognition rates and false positive rates before and after applying the CRF filter on the two systems in this study are provided in Figure 6I, which shows an increase in recognition rates and a significant decrease in false positive rates after applying the CRF filter on defected images. This indicates the effectiveness of this additional domain knowledge incorporated module. More details about the CRF filter can be found in the Experimental Section.

■ DISCUSSION

In summary, we have developed a one-image–one-system deep learning framework, which encompasses a uniquely designed data selection method, effective data augmentation techniques, and a Faster R-CNN detection model for semiautomated molecule detection and chirality identification of densely packed molecular self-assemblies in STM experiments. Our deep learning framework enables an efficient and well-trained automated analytic tool with a recognition rate of over 90% using a single STM image. Due to the high performance, robustness, and low-data requirement, our framework serves as a general automation solution for the detection and classification of complex SPM patterns. Our current framework can potentially remarkably accelerate the routine tasks of SPM researchers and the database development at the same time. It is expected to accelerate advanced materials discovery, involving supramolecular assemblies, for a wide range of applications such as photovoltaics, photocatalysts, organic semiconductors, drug delivery, cell culture, and molecular tuning.[5]

While our framework realizes fast deployment of an automation solution to any molecular system with limited experimental data supply, there is still a need to train separate models for different tasks and molecular systems. A general model that could be applied to different tasks and molecular systems for SPM imaging would be possible if a huge amount of versatile training data are provided. Such general models are usually available in other fields of deep learning where abundant training data can be collected, such as GPT-3 in natural language processing.[31] This amount of data could be possible from SPM image simulation methods[17,32−34] or automated SPM imaging tools.[11,35,36] In fact, simulated STM images were recently used for the training of the CNN.[33,34] Nevertheless, STM contrast depends on multiple variables, including tip–sample distance, orbital texture of the tip-apex, bias voltage, etc. Frequently, these variables cannot be correctly captured by STM simulations, leading to attenuated agreement between simulated and real STM images. In addition, the automation SPM imaging tool is still in the early stage, which can only work for a specific system.[11] Though a large amount of SPM images as a training set

for a general model are currently not feasible, we foresee this kind of general-purpose SPM intelligent model will be a meaningful future direction, and a general model will be developed when versatile data set is available.

■ EXPERIMENTAL SECTION

Faster R-CNN. Faster R-CNN is presented by a combination of two neural networks, named regional proposal network[26] (RPN) and fast region-based convolutional neural networks (Fast R-CNN). A general block chat diagram of this key model is illustrated in Figure 2, consisting of (i) fast region-based convolutional neural networks network (FPN)[23] backbone (i.e., feature extraction stage), (ii) RPN (i.e., molecule detection stage), and (iii) region of interest (RoI) module (i.e., molecule classification stage). In the following section, we provide a detailed explanation of each element of the Faster R-CNN architecture.

FPN extracts feature maps of the different scales from input SPM images and, hence, enables Faster R-CNN to detect molecules of various sizes, revealed by SPM images with different resolutions. As highlighted in blue in Figure 2, FPN includes five convolutional blocks connected in series built with a total of 50 convolutional and ReLu layers (ResNet-50[37]). The detailed architecture is provided in Supplementary Note S7.

For a common three-channel SPM image of $[3, H, W]$, three channels are used to store color information (red, green, and blue); H is the height of the image, and W is the width of the image. After the SPM image is processed, five 256-channel feature maps at five different scales can be extracted from different layers of FPN. 256 channels are used to store extensive abstract information on the input SPM image that is further used for molecular recognition. Different scales are designed to detect molecules of various sizes, which is explained in depth in the RoI section below. Feature map level is a number ranging from 1 to 5, determined by the layer location where a map feature is extracted. As the level of feature maps increases, the scale and resolution of feature maps decrease as shown in Figure 2, blue part. In the meantime, the spatial coverage and abstract information on feature maps gradually increase. Small objects like molecules are usually detected in low-level feature maps, which mainly contain high-resolution information essential for detection and location. The top-down structure of FPN allows high-level feature maps, which include more abstract information essential for classification, to be incorporated into low-level feature maps.[23]

RPN generates regional proposals, or bounding boxes, on the extracted feature maps and gives a score for each bounding box to contain objects of interest. For each input image, massive square anchors of five various sizes centered by each pixel of feature maps are first generated, as shown in Figure 2. It aims to ensure that all objects present in the feature maps are fully or partially covered by one or multiple anchors. The shape and sizes of anchors are deliberately customized for our small square-shaped molecules. After a two-stage fine-tuning and filtering in RPN and RoI (introduced later), anchors are outputted as bounding boxes, which can accurately predict the position of molecules on the input SPM images. A visualization of anchor processing schematics in this section and subsequent RoI section is shown in Figure 2. After the generation of a large number of possible anchors, a CNN based anchor proposal network is used. This network will output two kinds of information: the first one is the score of the proposed anchors, and the second is the tuning parameters for adjusting the proposed anchors. The higher the score, there is a higher possibility for the anchor to contain objects (i.e., molecules). Afterward, bad quality anchors (i.e., low score, out of the region, etc.) are removed. Besides, duplicate anchors are also removed by the nonmaximum suppression (NMS) method which is introduced later. The proposal information is combined with raw feature information from FPN and passed into the RoI module.

The RoI module will first convert the combined information from RPN and FPN into fixed-size vectors. The fixed-size characteristic allows proposals of different sizes to be classified by the same fully connected (FC) layers. The feature maps of the fixed shape are flattened and then processed by FC layers similar to a typical CNN.[38] The FC layers would output two matrixes: (i) a three-channel matrix containing the scores for the proposals to be each of the three classes (left-handed molecule (L-chiral), right-handed molecule (R-chiral), or a background); (ii) a four-channel matrix containing four final bounding box fine-tuning parameters. The probability distribution for a detected molecule to be a certain chirality (L- or R-chirality) is obtained from the scores by using the SoftMax function.

The hyperparameters used in this study were deliberately chosen for SPM image and small objects detection. However, we noticed that the instancing results might not be sensitive to hyperparameters since the complexity of molecular chirality classification is intrinsically low as suggested by t-SNE results. This developed method shows great improvements with higher performance, better accuracy, lower computational cost, and faster detection speed compared to traditional approaches.

NMS. It works as the anchor with the highest score is retained when multiple anchors have a large intersection over union (IoU). The above circumstance indicates those anchors probably cover the same molecule. RPN then outputs the remaining anchors with high scores as proposals. Each proposal is a bounding box obtained by preliminarily fine-tuning an anchor which likely contains a molecule. However, most anchors would be discarded as duplicates or backgrounds at this stage. More details about NMS operation can be found in Supplementary Note S8.

Transfer Learning and Training. Due to the data limitation in our study, we employed the transfer learning strategy that was widely used when using faster-RCNN. The transfer learning method to accelerate convergence was adapted by starting from a ResNet-50-FPN model[23] pretrained on COCO train2017 data set[39] comprising images of daily life objects. This approach can be used since the bottom of the backbone network FPN would be similar regardless of the detection targets of Faster R-CNN.[40] During the training process, the weights of model parameters from convolutional and fully connected layers are adjusted by the backpropagation and gradient descent method to minimize a loss function. The loss function mainly includes two types of components which are SoftMax cross-entropy classification loss between the actual chirality of molecules and model predictions and L1 bounding box regression loss between the actual positions, sizes of molecules, and model predictions. The detailed mathematical expression is provided in Supplementary Note S9.

t-SNE. t-SNE is an unsupervised machine learning algorithm often used for clustering and visualization of high-dimensional data. The algorithm starts by calculating the conditional probability of similarity between high-dimensional data points and also between their low-dimensional counterparts by the Euclidean distances of data points. It then attempts to minimize a cost function, which is defined as a single Kullback−Leibler divergence between joint probability distributions in the high-dimensional space and the low-dimensional space, by using a Student's t-distribution with a single degree of freedom to compute the similarity between two points in the low-dimensional space. Kullback−Leibler divergence measures the distance between two random distributions. When two random distributions are the same, their Kullback−Leibler divergence is equal to zero. When the difference between two random distributions increases, their Kullback−Leibler divergence also increases.[41]

CRF Filter. Conditional random field filter is a statistical modeling method that is used to classify a sample (a molecule in this study) based on its neighboring samples. The filter will first identify the intermolecular distance of the assembly and the neighbors of each molecule from the predictions of the Faster R-CNN. The molecular patterns on the STM image can be automatically learned by the CRF filter by generating a set of probability distributions. These probability distributions record the probability for a molecule to be R or L chirality at different neighboring conditions. On the basis of both the Faster R-CNN predictions and the probability distributions, the CRF filter can then re-evaluate all less confident predictions by the following eq 1:

$$P(\text{R}) = \frac{1}{Z} P_0(\text{R}) P_{\text{CRF}}(\text{R}|\text{N}_i) \tag{1}$$

where $P_0(R)$ is the prior probability for a molecule to be R chirality predicted by Faster R-CNN, $P_{CRF}(R|N_i)$ is the probability for the molecule to be R chirality given its neighboring condition N_i, Z is a normalizing constant, and $P(R)$ is the posterior probability for the molecule to be R chirality. The chirality of each molecule will be predicted based on the posterior probability in the final inference results. The source code can be found in the Zenodo database.

STM Measurements and Fabrication of Molecular Self-Assemblies. The synthetic procedures of the HPB and F-HPB compounds have been reported in previous studies.[42,43] A Knudsen cell was used to deposit the HPB and F-HPB precursor molecules (at 160 and 180 °C, respectively) onto a Ag(111) surface held at room temperature. The STM experiments were performed in ultrahigh vacuum conditions (base pressure of $>5 \times 10^{-11}$ mbar) at 4.4 K using a commercial Omicron low temperature STM system.

■ ASSOCIATED CONTENT

Ⓢ Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacs.1c03091.

> STM and SPM images and results, molecular structures, interference results, t-SNE results, details and results of modeling, architecture information, and guidelines for users (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

Jiong Lu − *Department of Chemistry, National University of Singapore, Singapore 117543, Singapore; Centre for Advanced 2D Materials (CA2DM), National University of Singapore, Singapore 117546, Singapore;* Ⓞ orcid.org/0000-0002-3690-8235; Email: chmluj@nus.edu.sg

Xiaonan Wang − *Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117585, Singapore;* Ⓞ orcid.org/0000-0001-9775-2417; Email: chewxia@nus.edu.sg

Authors

Jiali Li − *Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117585, Singapore*

Mykola Telychko − *Department of Chemistry, National University of Singapore, Singapore 117543, Singapore;* Ⓞ orcid.org/0000-0002-8156-8938

Jun Yin − *Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117585, Singapore*

Yixin Zhu − *Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117585, Singapore*

Guangwu Li − *Department of Chemistry, National University of Singapore, Singapore 117543, Singapore*

Shaotang Song − *Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117585, Singapore;* Ⓞ orcid.org/0000-0003-3487-3566

Haitao Yang − *Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117585, Singapore*

Jing Li − *Department of Chemistry, National University of Singapore, Singapore 117543, Singapore;* Ⓞ orcid.org/0000-0002-5627-4153

Jishan Wu − *Department of Chemistry, National University of Singapore, Singapore 117543, Singapore;* Ⓞ orcid.org/0000-0002-8231-0437

Complete contact information is available at:
https://pubs.acs.org/10.1021/jacs.1c03091

Author Contributions

‖Jiali Li and Mykola Telychko contributed equally to this work.

Notes

The authors declare no competing financial interest.
Codes are provided at the Zenodo DOI link https://zenodo.org/badge/latestdoi/280320042.

■ ACKNOWLEDGMENTS

■ REFERENCES

(1) Khajetoorians, A. A.; Wegner, D.; Otte, A. F.; Swart, I. Creating Designer Quantum States of Matter Atom-by-Atom. *Nat. Rev. Phys.* **2019**, *1* (12), 703−715.

(2) Jelínek, P. High Resolution SPM Imaging of Organic Molecules with Functionalized Tips. *J. Phys.: Condens. Matter* **2017**, *29* (34), 343002.

(3) Mali, K. S.; Pearce, N.; De Feyter, S.; Champness, N. R. Frontiers of Supramolecular Chemistry at Solid Surfaces. *Chem. Soc. Rev.* **2017**, *46* (9), 2520−2542.

(4) Wu, X.; Delbianco, M.; Anggara, K.; Michnowicz, T.; Pardo-Vargas, A.; Bharate, P.; Sen, S.; Pristl, M.; Rauschenbach, S.; Schlickum, U.; et al. Imaging Single Glycans. *Nature* **2020**, *582* (7812), 375−378.

(5) Ariga, K.; Nishikawa, M.; Mori, T.; Takeya, J.; Shrestha, L. K.; Hill, J. P. Self-Assembly as a Key Player for Materials Nanoarchitectonics. *Sci. Technol. Adv. Mater.* **2019**, *20* (1), 51−95.

(6) Baleizao, C.; Garcia, H. Chiral Salen Complexes: An Overview to Recoverable and Reusable Homogeneous and Heterogeneous Catalysts. *Chem. Rev.* **2006**, *106* (9), 3987−4043.

(7) Raval, R. Chiral Expression from Molecular Assemblies at Metal Surfaces: Insights from Surface Science Techniques. *Chem. Soc. Rev.* **2009**, *38* (3), 707−721.

(8) Chen, Z.; Lu, X. Self-Assembly of Plasmonic Chiral Super-structures with Intense Chiroptical Activity. *Nano Express* **2020**, *1* (3), 032002.

(9) Brooks, W. H.; Guida, W. C.; Daniel, K. G. The Significance of Chirality in Drug Design and Development. *Curr. Top. Med. Chem.* **2011**, *11* (7), 760−770.

(10) Ernst, K.-H. Molecular Chirality in Surface Science. *Surf. Sci.* **2013**, *613*, 1−5.

(11) Krull, A.; Hirsch, P.; Rother, C.; Schiffrin, A.; Krull, C. Artificial-Intelligence-Driven Scanning Probe Microscopy. *Commun. Phys.* **2020**, *3* (1), 54.

(12) Rashidi, M.; Wolkow, R. A. Autonomous Scanning Probe Microscopy in Situ Tip Conditioning through Machine Learning. *ACS Nano* **2018**, *12* (6), 5185−5189.

(13) Kalinin, S. V.; Strelcov, E.; Belianinov, A.; Somnath, S.; Vasudevan, R. K.; Lingerfelt, E. J.; Archibald, R. K.; Chen, C.; Proksch, R.; Laanait, N.; Jesse, S. Big, Deep, and Smart Data in Scanning Probe Microscopy. *ACS Nano* **2016**, *10* (10), 9068−9086.

(14) Wang, S.; Zhu, J.; Blackwell, R.; Fischer, F. R. Automated Tip Conditioning for Scanning Tunneling Spectroscopy. *J. Phys. Chem. A* **2021**, *125* (6), 1384−1390.

(15) Yoo, J.-C.; Han, T. H. Fast Normalized Cross-Correlation. *Circuits, Syst. signal Process.* **2009**, *28* (6), 819.

(16) Lindeberg, T. *Scale-Space Theory in Computer Vision*; Springer Science & Business Media, 2013; Vol. 256.

(17) Ziatdinov, M.; Maksov, A.; Kalinin, S. V. Learning Surface Molecular Structures via Machine Vision. *npj Comput. Mater.* **2017**, *3* (1), 31.

(18) O'Mahony, N.; Campbell, S.; Carvalho, A.; Harapanahalli, S.; Hernandez, G. V.; Krpalkova, L.; Riordan, D.; Walsh, J. Deep Learning vs. Traditional Computer Vision. In *Science and Information Conference*; Springer, 2019; pp 128−144.

(19) Ghosh, A.; Sumpter, B. G.; Dyck, O.; Kalinin, S. V.; Ziatdinov, M. Ensemble learning and iterative training (ELIT) machine learning: applications towards uncertainty quantification and automated experiment in atom-resolved. *arXiv* **2021**, arXiv:2101.08449v2 (https://arxiv.org/abs/2101.08449v2, accessed Jun 6, 2021).

(20) Shorten, C.; Khoshgoftaar, T. M. A Survey on Image Data Augmentation for Deep Learning. *J. Big Data* **2019**, *6* (1), 60.

(21) Telychko, M.; Wang, L.; Hsu, C.-H.; Li, G.; Peng, X.; Song, S.; Su, J.; Chuang, F.-C.; Wu, J.; Wong, R. M. W.; Lu, J. Tailoring Long-Range Superlattice Chirality in the Molecular Selfassemblies via Weak Fluorine-Mediated Interactions. *ChemRxiv* **2021**, DOI: 10.26434/chemrxiv.14743116.v1, (accessed Jun 8, 2021).

(22) 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.* **2017**, *39* (6), 1137−1149.

(23) Lin, T.-Y.; Dollár, P.; Girshick, R.; He, K.; Hariharan, B.; Belongie, S. Feature Pyramid Networks for Object Detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE, 2017; pp 2117−2125.

(24) Dodge, S.; Karam, L. Understanding How Image Quality Affects Deep Neural Networks. *Proceedings*, 2016 Eighth International Conference on Quality of Multimedia Experience (QoMEX); IEEE, 2016; pp 1−6.

(25) van der Maaten, L.; Hinton, G. Visualizing Data Using T-SNE. *J. Mach. Learn. Res.* **2008**, *9* (Nov), 2579−2605.

(26) Mikołajczyk, A.; Grochowski, M. Data Augmentation for Improving Deep Learning in Image Classification Problem. *2018 international interdisciplinary PhD workshop (IIPhDW). IEEE* **2018**, 117−122.

(27) Ziatdinov, M.; Dyck, O.; Maksov, A.; Li, X.; Sang, X.; Xiao, K.; Unocic, R. R.; Vasudevan, R.; Jesse, S.; Kalinin, S. V. Deep Learning of Atomically Resolved Scanning Transmission Electron Microscopy Images: Chemical Identification and Tracking Local Transformations. *ACS Nano* **2017**, *11* (12), 12742−12752.

(28) Pelt, D. M.; Sethian, J. A. A Mixed-Scale Dense Convolutional Neural Network for Image Analysis. *Proc. Natl. Acad. Sci. U. S. A.* **2018**, *115* (2), 254−259.

(29) Oksuz, K.; Cam, B. C.; Kalkan, S.; Akbas, E. Imbalance Problems in Object Detection: A Review. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, DOI: 10.1109/TPAMI.2020.2981890.

(30) He, K.; Gkioxari, G.; Dollár, P.; Girshick, R. Mask R-CNN. In *Proceedings of the IEEE International Conference on Computer Vision*; IEEE, 2017; pp 2961−2969.

(31) Brown, T. B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A. Language models are few-shot learners. *arXiv* **2020**, arXiv:2005.14165 (https://arxiv.org/abs/2005.14165, accessed Jun 6, 2021).

(32) Choudhary, K.; Garrity, K. F.; Camp, C.; Kalinin, S. V.; Vasudevan, R.; Ziatdinov, M.; Tavazza, F. Density functional theory and deep-learning to accelerate data analytics in scanning tunneling microscopy. *arXiv* **2019**, arXiv:1912.09027 (https://arxiv.org/abs/1912.09027, accessed Jun 6, 2021).

(33) Alldritt, B.; Hapala, P.; Oinonen, N.; Urtev, F.; Krejci, O.; Canova, F. F.; Kannala, J.; Schulz, F.; Liljeroth, P.; Foster, A. S. Automated Structure Discovery in Atomic Force Microscopy. *Sci. Adv.* **2020**, *6* (9), eaay6913.

(34) 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.* **2020**, *20* (10), 7688−7693.

(35) Gordon, O. M.; Moriarty, P. J. Machine Learning at the (Sub) Atomic Scale: Next Generation Scanning Probe Microscopy. *Mach. Learn. Sci. Technol.* **2020**, *1* (2), 023001.

(36) Leinen, P.; Esders, M.; Schütt, K. T.; Wagner, C.; Müller, K.-R.; Tautz, F. S. Autonomous Robotic Nanofabrication with Reinforcement Learning. *Sci. Adv.* **2020**, *6* (36), eabb6987.

(37) He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE, 2016; pp 770−778.

(38) Albawi, S.; Mohammed, T. A.; Al-Zawi, S. Understanding of a Convolutional Neural Network. In *2017 International Conference on Engineering and Technology (ICET)*; IEEE, 2017; pp 1−6, DOI: 10.1109/ICEngTechnol.2017.8308186.

(39) Lin, T.-Y.; Maire, M.; Belongie, S.; Bourdev, L.; Girshick, R.; Hays, J.; Perona, P.; Ramanan, D.; Zitnick, C. L.; Dollár, P. Microsoft COCO: Common Objects in Context. In *European Conference on Computer Vision*; Springer, 2014; pp 740−755.

(40) Girshick, R.; Donahue, J.; Darrell, T.; Malik, J. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE, 2013; pp 580−587.

(41) Perez-Cruz, F. Kullback-Leibler Divergence Estimation of Continuous Distributions. In *2008 IEEE International Symposium on Information Theory*; IEEE, 2008; pp 1666−1670, DOI: 10.1109/ISIT.2008.4595271.

(42) Telychko, M.; Li, G.; Mutombo, P.; Soler-Polo, D.; Peng, X.; Su, J.; Song, S.; Koh, M. J.; Edmonds, M.; Jelínek, P.; et al. Ultrahigh-Yield on-Surface Synthesis and Assembly of Circumcoronene into a Chiral Electronic Kagome-Honeycomb Lattice. *Sci. Adv.* **2021**, *7* (3), eabf0269.

(43) Li, G.; Han, Y.; Zou, Y.; Lee, J. J. C.; Ni, Y.; Wu, J. Dearomatization Approach Toward a Superbenzoquinone-Based Diradicaloid, Tetraradicaloid, and Hexaradicaloid. *Angew. Chem.* **2019**, *131* (40), 14457−14464.