Machine Learning to Reveal Nanoparticle Dynamics from Liquid-Phase TEM Videos

Lehan Yao, Zihao Ou, Binbin Luo, Cong Xu, and Qian Chen*

ABSTRACT: Liquid-phase transmission electron microscopy (TEM) has been recently applied to materials chemistry to gain fundamental understanding of a wide range of phenomena that are otherwise elusive. For example, by taking TEM videos of nanoparticle growth in real-time and real-space by liquid-phase TEM. A diversity of properties for differently shaped anisotropic nanoparticles are mapped, including the anisotropic interaction landscape of nanoprisms, curvature-dependent and staged etching profiles of nanorods, and an unexpected kinetic law of first-order chaining assembly of concave nanocubes. These systems representing properties at the nanoscale are otherwise experimentally inaccessible. Compared to the prevalent image segmentation methods, U-Net shows a superior capability to predict the position and shape boundary of nanoparticles from highly noisy and fluctuating background—a challenge common and sometimes inevitable in liquid-phase TEM videos. We expect our framework to push the potential of liquid-phase TEM to its full quantitative level and to shed insights, in high-throughput and statistically significant fashion, on the nanoscale dynamics of synthetic and biological nanomaterials.

INTRODUCTION

The recent proliferation of liquid-phase transmission electron microscopy (TEM) studies has offered great opportunities to fill knowledge gaps on nanoscale structural and functional dynamics by generating massive amounts of multidimensional data, which necessitate methods to extract quantitative information from them. On one hand, compared with ensemble measurements (e.g., small-angle X-ray scattering, UV–vis), liquid-phase TEM captures real-space videos at the nanometer, or sometimes atomic, resolution, which provides insights on a wide range of phenomena that are otherwise elusive. For example, by taking TEM videos of nanoparticle growth in solution, researchers have elucidated various growth mechanisms, such as nucleation, oriented attachment, and coalescence. Similarly, nanoparticle assembly and superlattice crystallization pathways have been studied by liquid-phase TEM, where kinetically stable intermediates were captured to understand phase transition. On the other hand, liquid-phase TEM videos are huge data files consisting of temporal stacks of spatial images. With the advent of fast detectors capable of capturing hundreds to thousands of frames per second, the size of liquid-phase TEM videos grows ever rapidly. Methods for automated image processing and analysis, with high tolerance to noise and high precision, thus become instrumental to converting raw TEM images of spatial intensity profiles to chemical and physical properties, and thus to materials design rules.

However, achieving automated and robust processing of liquid-phase TEM videos is still in its infancy, due to technical challenges associated with low quality videos of spatial heterogeneity, which as we argue below are often unavoidable and generic to liquid-phase TEM. First, the liquid layer and window (e.g., graphene, silicon nitride, graphene oxide) of a vacuum-sealed liquid chamber interact with incident imaging electrons, making the TEM contrast between the objects of interest and background lower than that in the TEM images taken at dry and windowless conditions. Second, numerous studies have showed that the imaging electrons can cause artifacts on the nanoscale dynamics by modifying the shape, interaction, and chemical stability of samples due to
effects such as ionization and radiolysis.\textsuperscript{13−16} As a result, using low dose rates emerges as a standard protocol in liquid-phase TEM studies.\textsuperscript{17} Shot noise of a Poisson nature increases with lowered dose rates,\textsuperscript{18} which together with the low image contrast makes liquid-phase TEM videos often with a low signal-to-noise ratio (SNR). For the low SNR TEM videos, conventional image analysis algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background). Recently developed advanced segmentation algorithms based on single intensity threshold easily gives false “segmentation” (identifying features from the background).

Here we draw upon the great success of machine learning in applications, such as facial recognition,\textsuperscript{22} self-driving automobile,\textsuperscript{23} biomedical image diagnosis,\textsuperscript{24} and the prominent example of atomic level structure analysis on scanning transmission electron microscopy (STEM) images,\textsuperscript{25,26} and apply machine learning for the first time to analyzing liquid-phase TEM videos. Compared with current liquid-phase TEM analysis methods,\textsuperscript{27,28} the key advantage of machine-learning methods is that no intensity threshold or input parameters other than the raw TEM videos themselves are needed. Specifically, we use the U-Net convolutional neural network (NN) among the many machine learning algorithms, which was purposed in 2015 by Olaf Ronneberger and co-workers.\textsuperscript{29} This method does not require hand-crafted features or predetermined input parameters and works by automatically learning a hierarchy of increasingly complex features directly from the training data. U-Net has made major breakthroughs in identifying irregular cellular features from gray scale optical microscopy images.\textsuperscript{30} It has not yet been applied to TEM images or liquid-phase TEM videos, which require a customized workflow we present here, particularly on the generation and choice of training data with well-defined ground truth. We show as proof-of-concept three types of nanoscale dynamics, including motion, chemical reaction, and self-assembly of nanoparticles. We extract a series of parameters from the analyzed TEM videos to reveal: (i) a highly anisotropic interparticle interaction landscape based on

Figure 1. U-Net-based image segmentation for liquid-phase TEM videos. (a) The workflow of segmentation for time-lapse TEM images through U-Net. (b) Representative images showing the data flow in U-Net models. From left to right: Simulated TEM images as inputs for training (nanorods, triangular nanoprisms, concave nanocubes, bipyramids), images obtained at different convolution layers (the middle 6 columns), and the prediction after training. (c) Convergence of the prediction results to ground truth over epochs. The red color in the images denotes deviation from ground truth. Here one epoch corresponds to one round of training cycles for all the training images (800 in this example). Training sets used are the simulated liquid-phase TEM images at $1 \text{ e}^{-1}\text{Å}^2\text{s}^{-1}$ (top) and $10 \text{ e}^{-1}\text{Å}^2\text{s}^{-1}$ (bottom), respectively. (d) Plot of intersection-over-union (IoU) during training over number of epochs and dose rates. A total of 10 models are trained to account for 10 different dose rates (training data simulated at the corresponding dose rates). Scale bars: 20 nm in (a); 20 nm, 100 nm, 200 nm, 100 nm from top to bottom in (b); and 100 nm in (c).
Figure 2. Statistical analysis of the diffusion, structure, and interaction landscape of gold triangular nanoprisms imaged by liquid-phase TEM, enabled by accurate segmentation based on the trained U-Net model. (a) Raw liquid-phase TEM snapshot of gold triangular nanoprisms taken at a dose rate of 3.7 $\text{e}^{-}\text{Å}^{-2}\text{s}^{-1}$ in a SiN$_x$ liquid chamber (see Materials and Methods in Supporting Information for experimental details). (b) Histogram of the pixel intensity of the liquid-phase TEM image (a) which shows a single-peaked distribution (black line) without a clear valley differentiating the nanoprisms and the background. Insets: Binarized images based on different single intensity thresholds as denoted by the arrows. (c) Binarized images of (a) by different combinations of Gaussian blur and thresholding by built-in functions in MATLAB. (d) Tracked nanoprisms (the same color denoting the same nanoprism) in time-lapse liquid-phase TEM images based on U-Net predictions. (e) Temporal trajectories of the tracked centroid positions of nanoprisms colored by their mean-squared displacement (MSD) at a time interval of 10 s. (f) Natural logarithm of the relative probability $\ln(P)$ of displacements of the tracked nanoprisms (d) at a time interval of 0.77 s with a parabola fitting at small displacements and a linear fitting at large displacements, where $P$ denotes the probability of a displacement of $d$ and $P_d$ denotes the probability of $d = 0$. Dash lines denote the extended fitting results. (g) The 2D distribution map of the occurrence of other nanoprisms when one nanoprism is positioned at the coordinate center. (h) The radial distribution function and “effective” interaction landscape based on the Boltzmann inversion rule as derived from (g). Dash lines in the plot correspond to data from the dashed box in (g), and solid lines correspond to data from the solid box in (g). (i) The 2D histogram (right) of the relative orientations (defined by $\theta_1$ and $\theta_2$ combined, left) of two nanoprisms at a particle center-to-center distance $r < 200$ nm. Scale bars: 200 nm in (a) and (e); 100 nm in (g).

The trajectory sampling of around 300,000 pairs of interacting nanoprisms, (ii) curvature-dependent and staged etching profiles of gold nanorods from the complete tracking of nanoparticle boundaries, and (iii) thermodynamically favored structural motifs and kinetic laws of chain formation from the self-assembly of concave nanocubes. These three systems, all new experimental results themselves, cover representative scenarios and challenges in the analysis of liquid-phase TEM videos. Notably, for a TEM video of hundreds of frames, it only takes < 1 min for a trained NN model to finish segmentation, which fully embraces the emerging paradigm shift to automated imaging and big data analysis in the field of liquid-phase TEM as well as the related disciplines where new knowledge is being learned, which is only possible based on statistical mechanics-based analysis of a significant data set.

## RESULTS AND DISCUSSIONS

### The Segmentation Workflow for Liquid-Phase TEM Videos Using U-Net.

In this work, we focus on using U-Net for the first and foremost step of image analysis—segmentation—to delineate objects of interest in the liquid-phase TEM videos (Figure 1a), though U-Net can also be extended to other pixel classification tasks. We recognize that one critical step in using U-Net for liquid-phase TEM videos is the generation and choice of training data. In previous applications, annotated (mostly manually) experimental images were used as training data. However, liquid-phase TEM videos taken at low dose rates have low SNR, making annotations difficult even for human experts. Inspired by machine-learning-based analysis of STEM images, we present a workflow to create a huge, in principle unlimited, number of training data sets consisting of simulated liquid-phase TEM images. As shown in Figure S1 and Figure 1b (the leftmost column), the simulated liquid-phase TEM images are generated by randomly placing nanoparticles on a substrate. The nanoparticles are modeled as polyhedrons in a three-dimensional (3D) liquid chamber (Machine Learning Workflow, Supporting Information). The pixel intensity in the simulated image is calculated as follows to resemble the contrast and background in an experimental liquid-phase TEM image. First, Beer’s law is applied to obtain the generic TEM contrast of the sample due to sample–electron interaction, following

$$I_b = \exp\left(-\sum \frac{I_i}{I_s}\right),$$

where $I_b$ is incident beam intensity, $I_s$ is the beam intensity after interacting with the sample, and $I_i$ and $I_s$ are...
the thickness and inelastic mean free path of species $i$ (e.g., nanoparticles, solvent, window) along the beam path. Next, Poisson noise is superimposed to consider the noise coming from the electron counting process (i.e., shot noise) that scales with dose rates, and Gaussian noise is applied to account for all other noises from the electronic devices and processes after the shot noise in a detector (Figure S1a). Last, a modulation transfer function measured from an experimental liquid-phase TEM image of a blank liquid sample is further superimposed to the simulated TEM image to reproduce the pixel—pixel intensity correlation (the visible “graininess” in a TEM image, Figure S2).

Using this workflow, we find that fewer than 1000 simulated TEM images are sufficient to validate training for the variety of nanoparticle shapes presented in Figure 1b. These nanoparticles cover the different geometric shapes one might encounter in TEM analysis, from two-dimensional (2D) (nanoprisms), three-dimensional (3D) convex (nanorods, bipyramids) to 3D concave (concave nanocubes) polygons. The workflow works for all these nanoparticles because the generation of simulated images poises in principle no limits on the object size or shape. This workflow of generating training data is now posted on GitHub as open-access codes (see Supporting Information).

The U-Net NN has a symmetric encoder-decoder architecture as shown in Figure 1a and Figure S3. The down-sampling process allows the NN to identify the features at different scales, while the up-sampling process reconstructs the extracted features at the original image scale to complete the forward propagation. The result is compared with ground truth and then proceeds reversely to the beginning to complete the backpropagation and update every parameter inside the U-Net model; this concludes one training cycle. We find that the number of epochs needed for sufficient training depends on the SNR (as determined mostly by the electron dose rate) of the input images. For images of high SNR (Figure 1c bottom, $10 \text{e}^{-}\cdot \text{Å}^{-2}\cdot \text{s}^{-1}$), the IoU, namely the fraction of the particle prediction overlapping with ground truth, increases rapidly from 0.22 to 0.92 after the first epoch, while images of a low SNR (Figure 1c top, $1 \text{e}^{-}\cdot \text{Å}^{-2}\cdot \text{s}^{-1}$) take 15 epochs to reach a similar IoU (0.90). We plot the IoU evolution during training as a function of dose rates and number of epochs (Figure 1d). Within 15 epochs of training, the U-Net model is capable of high-precision segmentation since the SNR of the input images is too low (at a dose rate lower than $1 \text{e}^{-}\cdot \text{Å}^{-2}\cdot \text{s}^{-1}$) (Figures S4 and S5). Next, experimental liquid-phase TEM videos are fed to the trained NN without preprocessing, which yield time series of pixel-level probability maps representing the likelihood of each pixel belonging to the objects of interest, as the prediction images for further analysis.

Based on this unified U-Net segmentation workflow, we present three different systems as below. In each system, we discuss the challenges in segmentation using conventional threshold-based algorithms and how our U-Net method overcomes these challenges. Meanwhile, each system represents one type of liquid-phase TEM studies, including nanoparticle diffusion and interaction, chemical reaction, and self-assembly, demonstrating the potency and versatility of our method.

**System 1: Mapping the Interaction Landscape of Gold Triangular Nanoprisms.** In the first system, we focus on mapping an effective interaction landscape at the nanoscale, which governs fundamentally the phase behaviors of nanosized objects in the solution but has remained challenging to model or predict due to nonadditive multiscale coupling effects. Specifically, we study a model system of gold triangular nanoparticles with anisotropic shapes to show how U-Net-based segmentation enables interaction mapping. The small thickness of the nanoprisms (7.5 nm) and the low dose rate ($3.7 \text{e}^{-}\cdot \text{Å}^{-2}\cdot \text{s}^{-1}$) used to minimize beam artifacts lead to low SNR in the raw liquid-phase TEM images (e.g., Figure 2a). As a result, the typical intensity histogram of the images exhibits only a single peak (Figure 2b). In conventional intensity threshold-based algorithms, a threshold at the valley between two intensity peaks of the histogram is chosen to divide the image pixels into background and foreground (objects of interest). In this liquid-phase TEM image, the intensities of background and foreground cannot be distinguished, thereby resulting in blurry or incorrect segmentation regardless of the choice of thresholds (insets in Figure 2b). Alternative tricks such as Gaussian blur can reduce the background noise and make more structured prediction (Figure 2c, right), but the segmented boundaries are also blurred during this process, resulting in rounded corners and merged shapes. Other smart thresholding algorithms provided by MATLAB, such as automatic global thresholding (Figure 2c, top) and adaptive thresholding (Figure 2c, bottom), also show insufficient segmentation. In contrast, when we use the same liquid-phase TEM image as input for the trained U-Net, the predicted image is significantly improved in determining the probability of the pixels belonging to nanoprisms (Figure 2d). The predicted image is further binarized given the better-defined intensity histogram (Figure 5h). The centroid positions and boundaries of the nanoprisms (∼50 particles per frame) are extracted to track the continuous motions of nanoprisms (Figure 2e, Movie S1).

Statistical analysis of the trajectories shows the presence of two components in the motions of nanoprisms. The logarithm of displacement probability of the nanoprisms (∼28,000 displacements with a time interval $\Delta t$ of 0.77 s) is fitted with a parabola trend at a small displacement ($d < 4 \text{ nm}$) and a linear relationship at large displacement (Figure 2f). The parabola region corresponds to a Gaussian-distributed Brownian motion with a diffusion coefficient of 0.5 nm$^2$·s$^{-1}$, while the linear region suggests an exponential tail which could be attributed to hopping of the nanoprisms, consistent with the adsorption and desorption of nanoparticles to the SiN$_2$ substrate as discussed in previous studies. Such a two-component fitting has been observed in the motions of micron-sized objects in a crowded environment imaged by optical microscopy as they are temporarily confined by the surrounding particles. Here, we retrieve similar behaviors at the nanoscale as the nanoprisms are partially bounded by the substrate.

Meanwhile, the smooth boundaries of the nanoprisms tracked by our U-Net enable us to analyze the stable motifs of two interacting nanoprisms and measure an effective interaction landscape at nanometer resolution. In Figure 2g, we enumerate all the nanoprism pairs (around 300,000) in the TEM video, put one nanoprism in the pair at the coordinate center, and plot the counts of a pixel $(x, y)$ in the 2D plane occupied by the other nanoprism in the pair (Figure S6). The as-obtained 2D distribution map shows a high-occurrence, short-range structure of two nanoprisms in the side-by-side motif, which agrees with earlier simulations on the densest packing of triangular nanoprisms. From the 2D distribution, we map the radial distribution function $g(r)$ along two different directions as marked in Figure 2g. Clearly $g(r)$ exhibits an anisotropy (Figure 2h), differing from the radially symmetric $g(r)$ in the gold nanorod case studied earlier. Further analysis of a different 2D histogram on the relative orientation of two close nanoprisms ($r$
< 200 nm) (Figure 2i) confirms that the nanoprisms are preferentially positioned in the side-by-side configuration (Figure 2i, upper right). The other slightly high probability motif corresponds to a sawtooth configuration where the two tips of two nanoprisms connect with a 120° bond angle (Figure 2i, center). This sawtooth configuration can be attributed to balanced van der Waals attraction and electrostatic repulsion we previously discussed. 17 Finally, based on the Boltzmann inversion rule, namely more populated states corresponding to lower free energies in equilibrium, we map an effective interaction landscape. As shown in Figure 2h, the energy valley occurs at different locations along different directions of the nanoprisms, with a net attraction strength of $-0.55 k_B T$ ($r = 100$ nm) perpendicular to the side of the center nanoprism (dashed box in Figure 2g) and of $-0.39 k_B T$ ($r = 125$ nm) along the tip of the center nanoprism (solid box in Figure 2g). Real-space imaging of trajectories allows for mapping anisotropic interaction.

Note that our measured interaction landscape is phenomenological and can be a combined result contributed by factors such as entropic effects that favor closest packing of nanoprisms, and enthalpic effects (e.g., van der Waals attraction, electrostatic repulsion, steric repulsion of the surface ligands). Nevertheless, the direct measurement of the net interparticle interaction by delineating the boundaries of anisotropic nanoparticles via U-Net-based segmentation serves as a general route for arbitrarily shaped or solvated nanosized objects. The identification of motifs (i.e., different configurations in Figure 2i) provides essential guidance for designing more sophisticated assembly structures from anisotropic nanoparticles. Our demonstrated tracking capability of a massive number of anisotropic nanoparticles not only enables statistical mechanics-based analysis of the energetics of the system but also reveals the dynamics of individual nanoparticles, which can be potentially useful for understanding questions ranging from diffusions in geometrically or chemically confined environment, non-equilibrium self-propulsion and collective motion as driven by external fields, to aggregations of heterogeneous and polydisperse systems such as micro- and nanoparticles.

**System 2: Curvature-Dependent, Staged Etching of Gold Nanorods.** In the second system, we focus on the chemical reaction kinetics of nanoparticles captured by liquid-phase TEM, where mechanisms such as existence of non-equilibrium intermediates, nucleation, growth, and corrosion have been suggested, and our U-Net method pushes for complete quantitativeness. We choose an established system of gold nanorods dispersed in FeCl$_3$ solution as a representation of studies on beam-triggered nanoparticle transformation. Although this experiment does not require low-dose imaging because electrons are used as the reactant, leading to liquid-phase TEM video of high SNR, conventional thresholding-based segmenta-

---

Figure 3. Multiobject tracking of particle boundaries using U-Net to reveal the curvature-dependent, staged etching dynamics of gold nanorods captured by graphene liquid cell TEM. (a) Time-lapse liquid-phase TEM images overlaid with nanorod boundaries predicted by the trained U-Net model (each color denoting a different nanorod). (b) The intensity histograms of all the four gold nanorods (coloring matching with a) and the background (gray) in the TEM images as tracked by U-Net. (c) Nanorod boundaries colored to their local surface curvatures during etching. (d) Nanorod boundaries colored to their local etching rates. The height of the boundary profiles corresponds to the etching time. (e) The three stages of etching based on the distinctive relationships between the etching rate and curvature at a local site. Coloring denotes the nanorods in (a). 0–40 s is regarded as etching stage 1 for the green, yellow, blue, and red nanorods. The etching stage 2 is 40–49 s, 40–62 s, and 40–68 s for the green, blue, and red nanorods, respectively, and stage 3 covers the rest of the etching. There is no stage 2 or 3 for the yellow nanorods. Scale bars: 50 nm.
tion still remains insufficient. The thickness and orientation of each nanorod are different, generating different TEM contrasts in the same image, which makes a single threshold unsuitable for all the nanorods (Figure S7): The nanorods of low TEM contrast are tracked with the boundaries enveloping a smaller area or completely missed (Figure S7i). Even the intensities of the same nanorods undergo temporal changes due to motions, shifting the intensity histograms over time (Figure 3b). This shift makes the segmentation inconsistent, which causes an unphysical “dip” in the nanorod area versus time curve (Figure S7j). Note that such intensity fluctuation in the objects can happen inevitably in some liquid-phase TEM videos and causes difficulties in analyzing those videos. Though it is feasible to manually adjust the threshold from time to time or set different thresholds for different nanorods, these modifications require additional parameter optimization, which can be low throughput and may easily bias the segmentation based on the operator’s own subjective judgment.

Our trained U-Net model shows robust segmentation even in the presence of the contrast fluctuation among individual nanorods and with time because it requires no determination of thresholds (Figure 3a). The greatly improved accuracy on tracking particle boundaries is confirmed by the nanorod area versus time curve (Figure S7j). Note that such intensity fluctuation in the objects can happen inevitably in some liquid-phase TEM videos and causes difficulties in analyzing those videos. Though it is feasible to manually adjust the threshold from time to time or set different thresholds for different nanorods, these modifications require additional parameter optimization, which can be low throughput and may easily bias the segmentation based on the operator’s own subjective judgment.

In our effort to elucidate a quantitative interdependence of the local curvature and local etching rate of nanorods, we find unexpectedly a three-stage process, which is a characteristic otherwise hidden but emerges from the significant data set of the nanorod boundaries we now have. As shown in Figure 3e, the first 40 s is stage 1, where isotropic etching dominates. We see no statistical difference of local etching rates over the whole curvature range, possibly due to a large excess of oxidative species at the beginning that overwhelms the curvature dependence. In the intermediate stage 2, the etching becomes directional, with the higher-curvature sites etched faster, following a positive relationship between the local curvature and local etching rate as noted in Figure 3e. This regime can be attributed to the high reactivity of less coordinated atoms at the high-curvature sites, consistent with previous discussions where high-curvature sites signature lower ligand density. At the end of the etching process (stage 3), the etching rates no longer depend on local curvatures but undergo a sudden drop, possibly due to the local accumulation of gold ions (products of the etching) or local depletion of oxidative species in a sealed liquid environment, both of which slow the etching. By monitoring more than one gold nanorod, we also found that the etching profiles are not the same for every particle. As shown in Figure 4d, one gold nanorod (bottom right, shown in yellow in Figure 4a,b,e) is different from the other three nanorods, exhibiting a

Figure 4. Substrate-mediated self-assembly of gold concave nanocubes. (a) Time-lapse liquid-phase TEM snapshots of chaining concave nanocubes with the particle centroid labeled as red, tracked by a trained U-Net model. (b) The conventional thresholding directly “cuts” the intensity (for easier comparison, the intensity is inverted) in the original image, giving irregular shaped or connected particle boundaries. (c) Applying a threshold to the U-Net prediction produces more accurate particle boundaries than that in (b). (d) The 2D histogram (right) of the relative orientations (defined by θ₁ and θ₂ combined as shown on the left) of two close concave nanocubes at a surface-to-surface distance < 35.7 nm. (e) Plot of ln(Xₙ(t)) and t. Straight line is the linear fitting. Inset: Plot of Xₙ(t)−t showing a nonlinear trend. (f–h) Liquid-phase TEM snapshots during the chaining process. Green denotes desorbed particles, and red denotes adsorbed particles on the SiNx window. Scale bars: 300 nm in (a); 100 nm in (f–h).
curvature independent, single-staged etching profile. We attribute this heterogeneity on the single particle level to a heterogeneous local reaction environment surrounding the nanorods. For example, nanoparticles at different positions in the imaging area might be exposed to different concentrations of oxidative species as the electron beam modulates the network of redox reactions. Or the local depletion or excess of FeCl₃ concentration could be different if the initial nanorod’s size and shape are varied. We see this subtle difference showcasing the need for our demonstrated capability to track the etching trajectories of many nanorods.

Our complete mapping of the staged and curvature dependent etching of gold nanorods is only possible by the huge amount of data we can derive with accurate boundary tracking. Our method can be applied to identify broadly multistep processes involved in chemical reactions or structural transformation of nanosized entities. In comparison to the single-step etching presented in previous work, the detailed understandings on the staged etching mechanism suggest additional routes to control the etching kinetics by intervening different stages, which can be combined with regrowth methods to synthesize nanoparticles with novel shapes or hierarchical structures.

**System 3: “Individualizing” Connected Assemblies into Basic Units to Elucidate Assembly Kinetics.** In the third system, we focus on the kinetic pathways of nanoparticle assembly and show the potency of U-Net in solving one key challenge in the analysis of such liquid-phase TEM videos: Distinguishing individual building blocks in a connected/bonded assembled structure. The system concerns an aqueous suspension of dispersed, charged gold concave nanocubes sandwiched in a SiNₓ liquid chamber. Upon the flow of salted buffer, the concave nanocubes have their electrostatic repulsion screened and assemble into chains as shown in the liquid-phase TEM images (Materials and Methods, Supporting Information, Figure S4) and Movie S3. The chaining process resembles qualitatively bifunctional monomers polymerizing into one-dimensional polymers. In order to quantify the kinetic laws and chain topology (e.g., connection scheme, bond angle) during assembly, we extract the position and orientation of the concave nanocubes consisting the assembled chains. Conventional thresholding method fails to disconnect the “bridge” region between bonded concave nanocubes while keeping the particle shape integrity. As shown in Figure 4b, a low threshold keeps the bridge region between two concave nanocubes and thus does not separate them, while a high threshold gives rise to unphysical and corrugated particle shape. In contrast, using the trained U-Net model, we retrieve mostly smooth particle boundaries of individual concave nanocubes. Remarkably, in the image predicted after U-Net processing (Figure 4c), the intensities of the pixels in the “bridge” region are much lower than those in the “particle” region. We attribute this advantage of our U-Net model over single thresholding to the fact that U-Net weights collectively many factors such as intensity, shape, and local surroundings to achieve the final segmentation, while the thresholding method concerns only pixel intensity. The centroid position and orientation of each concave nanocube in growing chains are shown in Movie S3, so that we can digitalize the liquid-phase TEM images into nanoparticle “polymers” with well-defined topology and bond angle. The bond angle statistics on the bonded pairs show preference on the tip-to-tip configuration of the concave nanocubes (Figure 4d), likely due to the smaller electrostatic repulsion at the position where tips are connected.

The immediate surprise that emerges is a first-order reaction mechanism in the polymerization kinetics as we analyze the length evolution of growing chains. We use the number-averaged degree of polymerization \( \langle X_n(t) \rangle \) at a given time \( t \) as the average number of concave nanocubes in polymer chains. The \( X_n(t) - t \) curve shows a nonlinear dependence (Figure 4e), different from the linear relationship observed in previous studies associated with step-growth polymerization.\(^{17,64}\) Instead, \( X_n(t) \) relates phenomenologically to \( t \) following a mathematical form of \( \ln(X_n(t)) = k t + 0.17 \), in which \( k \) is the assembly reaction constant is fitted to be 0.029 s\(^{-1}\). Such a linear relationship between \( \ln(X_n(t)) \) and \( t \) can be rationalized by a first-order reaction accounting for the linking of monomers into polymers: \(-d[M]/dt = k[M]\), where \([M]\) is the instantaneous concentration of monomers and chains (see derivation in Image Processing and Data Analysis in Supporting Information).

The first-order reaction mechanism seems counterintuitive first because the growth of linear polymer chains requires two monomers/chains to approach and connect. We find that the adsorption of concave nanocubes onto the SiNₓ window likely due to a friction aging effect reported before\(^{65}\) explains this unexpected behavior: The concave nanocubes spend \(~98%\) of the time adsorbed, and the chain growth only occurs for the occasionally desorbed particles (Figure 4f–h). By neglecting the rare cases where two desorbed particles bond with each other, the chain growth can be simplified into two steps:\( M_m \xrightarrow{k_1} M_m^* \) (step 1) and \( M_m^* + M_n \xrightarrow{k_1} M_{m+n} \) (step 2), where \( M_m \) stands for adsorbed species consisting \( m \) monomeric particles, and \( M_m^* \) stands for the corresponding desorbed species. Given our observation that the time between desorption events (~16 s) is much longer than the average time required for a desorbed particle to diffuse and attach to growing chains (~0.5 s), i.e., \( k_1 \ll k_d \), step 1 is the rate-limiting step and defines a linear dependence of reaction rate on the concentration \([M]\). In other words, our fitted \( k (0.029 \text{ s}^{-1}) \) essentially measures particle desorption rate \( k_d \). This kinetic model proposed here works for the scenario where particles adsorption on the substrate is involved in the solution-phase assembly, a typical case for nanoparticle assembly in liquid-phase TEM studies and solvent-evaporation driven self-assembly.\(^{66,67}\) This capability to quantitatively track the kinetic pathways during nanoparticle assembly offers mechanistic insights that can unravel previously unrecognized routes to manipulate the kinetics. The switching between the desorption controlled first-order growth mechanism demonstrated here and a diffusion controlled second-order step-growth mechanism discussed extensively in previous studies\(^{17}\) can potentially lead to new self-assembly structures for advanced applications.

**CONCLUSION**

In this work we map the interaction landscape, anisotropic etching, and the self-assembly kinetic laws using a U-Net-based workflow to achieve nanoparticle segmentation in liquid-phase TEM videos. U-Net can efficiently and precisely identify the boundary of nanoparticles, with a high tolerance over low SNR, spatial heterogeneity in particle intensity, and particle bridging, all of which are long-standing obstacles in the analysis of TEM data using conventional tracking methods, and we now solve all at once. Going beyond proof-of-concept systems of inorganic nanoparticles, we expect machine learning as a potent toolset
can also enable the analysis of liquid-phase TEM videos of soft materials such as polymers, micelles, or proteins or even detect prenucleation clusters in molecular crystals that are otherwise too difficult to distinguish using traditional methods.

**ASSOCIATED CONTENT**

**Supporting Information**
The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acscentsci.0c00430.

Details of synthesis, sample preparation and liquid-phase TEM imaging methods; whole workflow of the machine-learning-based video processing and data analysis with figures; additional figures of performance comparison and neural network architecture (PDF)

Movie S1: Dynamic diffusion of concentrated gold nanoparticles and the particle segmentation result (AVI)

Movie S2: Etching of gold nanorod sand the particle segmentation result (AVI)

Movie S3: Real-time self-assembly of gold concave nanocubes with tracked particle boundaries and bonding structure (AVI)

**AUTHOR INFORMATION**

**Corresponding Author**
Qian Chen — Department of Materials Science and Engineering, Research Laboratory, Beckman Institute for Advanced Science and Technology, and Department of Chemistry, University of Illinois at Urbana—Champaign, Urbana, Illinois 61801, United States
Email: qchen20@illinois.edu

**Authors**
Lehan Yao — Department of Materials Science and Engineering, University of Illinois at Urbana—Champaign, Urbana, Illinois 61801, United States
Zihao Ou — Department of Materials Science and Engineering, University of Illinois at Urbana—Champaign, Urbana, Illinois 61801, United States
Binbin Luo — Department of Materials Science and Engineering, University of Illinois at Urbana—Champaign, Urbana, Illinois 61801, United States
Cong Xu — Department of Materials Science and Engineering, University of Illinois at Urbana—Champaign, Urbana, Illinois 61801, United States

Complete contact information is available at: https://pubs.acs.org/doi/10.1021/acscentsci.0c00430

**Author Contributions**
Q.C., B.L., and C.X. designed and performed the experiments. L.Y., Z.O., and Q.C. performed the machine learning and the analysis of experimental data. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

**Notes**
The authors declare no competing financial interest.

**ACKNOWLEDGMENTS**
This research was supported by the National Science Foundation under grant no. 1752517. The workstation used for U-Net-based segmentation was supported by the Air Force Office of Scientific Research Defense Research Instrumentation Program (under AFOSR Project FA9550-18-0393). Experiments were carried out in part in the Materials Research Laboratory (MRL) Central Research Facilities, University of Illinois. We thank Prof. Jian-Min Zuo at University of Illinois for the useful discussions on the simulation of liquid-phase TEM images. MATLAB code (and the associated user graphic interface) used to generate simulated liquid-phase TEM images for training data sets as well as an example script for model training are accessible on the Internet: https://github.com/chenlabUIUC/simulatedLPTEM and https://github.com/chenlabUIUC/U-NetExample.

**REFERENCES**

1. Lob, N. D.; Sen, S.; Bosman, M.; Tan, S. F.; Zhong, J.; Nijhuis, C. A.; Král, P.; Matsuda, P.; Mirmasoud, U. Multistep nucleation of nanocrystals in aqueous solution. Nat. Chem. 2017, 9, 77–82.
2. Zhu, C.; Liang, S.; Song, E.; Zhou, Y.; Wang, W.; Shan, F.; Shi, Y.; Hao, C.; Yin, K.; Zhang, T.; et al. In-situ liquid cell transmission electron microscopy investigation on oriented attachment of gold nanoparticles. Nat. Commun. 2018, 9, 421.
3. Song, M.; Zhou, G.; Lu, N.; Lee, J.; Nakouzi, E.; Wang, H.; Li, D. Oriented attachment induces fivefold twins by forming and decomposing high-energy grain boundaries. Science 2020, 367, 40–45.
4. Niu, K.-Y.; Liao, H.-G.; Zheng, H. Visualization of the coalescence of bismuth nanoparticles. Microsc. Microanal. 2014, 20, 416–424.
5. Luo, B.; Smith, J. W.; Ou, Z.; Chen, Q. Quantifying the self-assembly behavior of anisotropic nanoparticles using liquid-phase transmission electron microscopy. Acc. Chem. Res. 2017, 50, 1125–1133.
6. Ou, Z.; Wang, Z.; Luo, B.; Luijten, E.; Chen, Q. Kinetic pathways of crystallization at the nanoscale. Nat. Mater. 2020, 19, 450–455.
7. Yuk, J. M.; Park, J.; Ercius, P.; Kim, K.; Hellebusch, D. J.; Crommie, M. F.; Lee, J. Y.; Zettl, A.; Alivisatos, A. P. High-resolution EM of colloidal nanocrystal growth using graphene liquid cells. Science 2012, 336, 61–64.
8. Ye, X.; Jones, M. R.; Frechette, L. B.; Chen, Q.; Powers, A. S.; Ercius, P.; Dunn, G.; Rotkoff, G. M.; Nguyen, S. C.; Adiga, V. P.; et al. Single-particle mapping of nonequilibrium nanocrystal transformations. Science 2016, 354, 874–877.
9. Park, J.; Zheng, H.; Lee, W. C.; Geissler, P. L.; Rabani, E.; Alivisatos, A. P. Direct observation of nanoparticle superlattice formation by using liquid cell transmission electron microscopy. ACS Nano 2012, 6, 2078–2085.
10. Lee, D.; Park, H.; Ko, Y.; Park, H.; Hyeon, T.; Kang, K.; Park, J. Direct observation of redox mediator-assisted solution-phase discharging of Li-O2 battery by liquid-phase transmission electron microscopy. J. Am. Chem. Soc. 2019, 141, 8047–8052.
11. De Clercq, A.; Dachraoui, W.; Margeat, O.; Pelzer, K.; Henry, C. R.; Giorgio, S. Growth of Pt-Pd nanoparticles studied in situ by HRTEM in a liquid cell. J. Phys. Chem. Lett. 2014, 5, 2126–2130.
12. Jungnhohn, K. L.; Evans, J. E.; Aguiar, J. A.; Arslan, I.; Browning, N. D. Atomic-scale imaging and spectroscopy for in situ liquid scanning transmission electron microscopy. Microsc. Microanal. 2012, 18, 621–627.
13. Ramachandramoorthy, R.; Bernal, R.; Espinosa, H. D. Pushing the envelope of in situ transmission electron microscopy. ACS Nano 2015, 9, 4675–4685.
14. Kim, J.; Jones, M. R.; Ou, Z.; Chen, Q. In situ electron microscopy imaging and quantitative structural modulation of nanoparticle superlatice. ACS Nano 2016, 10, 9801–9808.
15. Weiner, R. G.; Chen, D. P.; Usovic, R. R.; Srkalic, S. E. Impact of membrane-induced particle immobilization on seeded growth monitored by in situ liquid scanning transmission electron microscopy. Small 2016, 12, 2701–2706.
16. Woehl, T. J.; Evans, J. E.; Arslan, I.; Ristenpart, W. D.; Browning, N. D. Direct in situ determination of the mechanisms controlling nanoparticle nucleation and growth. ACS Nano 2012, 6, 8599–8610.
