Automated particle detection and quantitative analysis from electron microscopy images

Mohammad Javad Raei, Seyed Aref Golsorkhi, Frances Joan Alvarez, Marc Mamak

Traditional cryo-EM analysis is challenging because imaging physics and sample conditions make the rims inconsistent and hard to segment. Low SNR obscures boundaries, defocus/CTF effects from intentional underfocus create halos and edge delocalization, so the same particles can look sharp in one region but faint or broken in another, and beam-induced motion further blurs fine features despite frame alignment. In crowded samples also touch or overlap in 2D projection, so intensity-based masks and morphology tend to merge neighboring particles into a single object. For this reason, particle quantification is not merely counting, but a materials-characterization task where correctly distinguishing true particles from ice texture, contaminants/aggregates, or filament-like debris is essential for meaningful statistics, because misclassification and merging/split errors directly bias the measured diameter or axis-length distributions that are used to relate microstructure to macroscopic performance. Classical pipelines (e.g., thresholding with contrast enhancement, median filtering, dilation/fill-hole, Canny edges, and Hough-based circle finding) can improve visibility but often amplify background texture, fragment rims, and become highly parameter-sensitive in low contrast, overlap, or deformation-leading to false positives and missed or merged particles. Together, these limitations motivate an instance-aware workflow that separates overlapping particles like vesicles and then fits circles/ellipses to extract reliable size/shape statistics for formulation and processing decisions.

Figure 1 summarizes the NanORange end-to-end workflow for automated vesicle/nanoparticle quantification from Cryo-EM micrographs, combining Gemini-based AI agents with deterministic image processing. Raw low-dose images (optionally tiled for large fields of view) are first passed to an AI parameter optimizer, which selects a compact set of hyperparameters that govern enhancement and segmentation behavior. An AI contrast enhancer then denoises and boosts local contrast to make particle rims more consistent under variable ice thickness/defocus. The enhanced output is thresholded to produce an initial binary mask with basic cleanup to suppress speckle and avoid unintended merging. Next, an AI boundary colorizer generates color-coded instance boundaries to better separate adjacent objects and support downstream instance separation. Individual contours are then processed by shape fitting (circle/ellipse least-squares fits) to extract quantitative descriptors such as equivalent diameter, major/minor axes, aspect ratio, and fit residuals/quality flags. Finally, the pipeline exports annotated overlays and structured measurements (PNG/CSV/HTML) and provides an optional human-in-the-loop review step that feeds parameter adjustments back to the optimizer for iterative refinement.

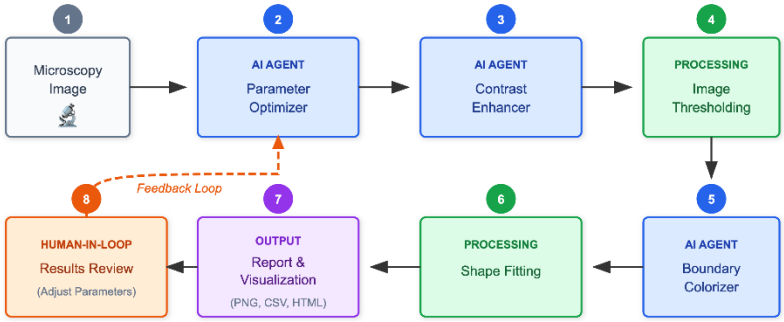

Figure 1 NanORange analysis workflow for Cryo-TEM particle/vesicle quantification

Figure 2 benchmarks the same multi-stage vesicle-quantification pipeline on two distinct cryo-EM datasets Liquid Fabric Enhancer (LFE1, LFE2), highlighting how image variability propagates through detection and why an instance-aware workflow is needed. In the raw micrographs, vesicle rims are barely visible because of low-dose contrast and spatially varying background; this is especially evident in the LFE1 example, where strong nonuniform background features and a prominent diagonal, filament-like/contaminant structure obscure vesicle boundaries and create regions where vesicles blend into the ice/background. After AI denoising + contrast enhancement + adaptive thresholding, vesicle rims become much more continuous and measurable, but the output also reveals a key challenge: complex background texture and overlapping vesicles produce dense edge networks (LFE1) and crowded rim fields (LFE2), where simple binary masks would merge neighboring vesicles and bias any subsequent size analysis. The colored-boundary step addresses this by explicitly separating individual vesicle instances, even in crowded regions where multiple vesicles touch or overlap in 2D projection. This is crucial in both datasets, but for different reasons: LFE1 contains many partially occluded vesicles near the diagonal feature and heterogeneous contrast regions that would otherwise create fragmented or merged contours, while LFE2 contains a broad size distribution with many small vesicles packed around large ones, making separation the dominant difficulty. Finally, the fitted-shape overlay shows the quantitative endpoint of the workflow: once instances are isolated, circle/ellipse fits can be applied consistently across dense fields to extract diameter/axis-length statistics and enable downstream size-distribution analysis.

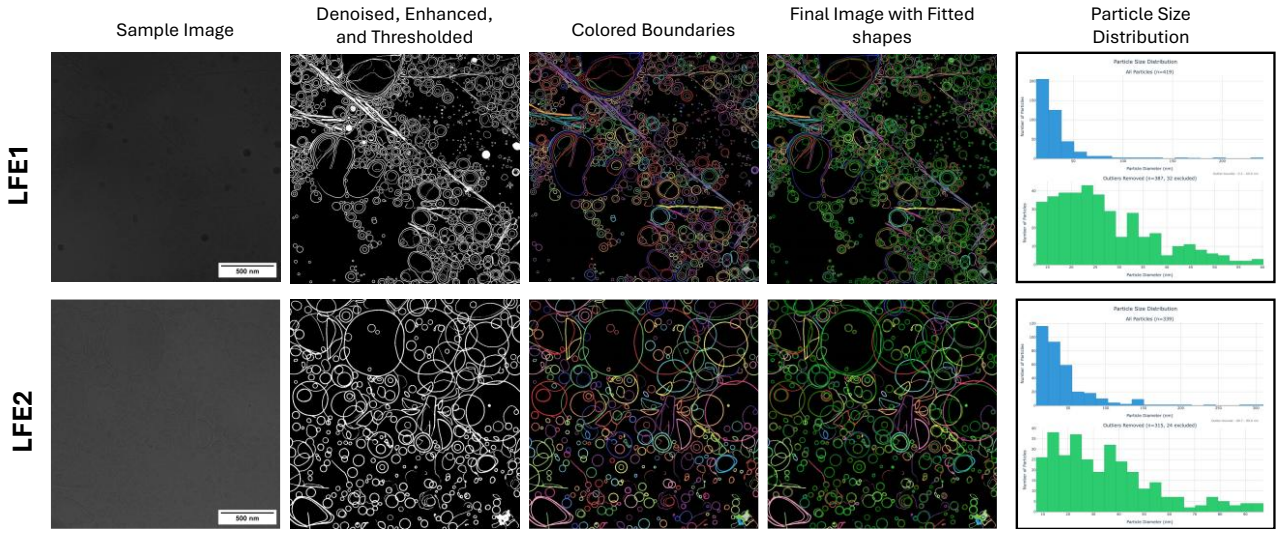

Figure 2 AI-assisted vesicle detection and geometric quantification on two cryo-EM datasets.

Future work will leverage the modular design of this pipeline, which can be easily adjusted to new imaging conditions and sample types with minimal changes, making it a strong candidate for large-scale, semi-automated labeling across diverse EM datasets. By using NanORange to rapidly generate consistent instance boundaries and size annotations, we can build a high-volume labeled dataset that would be impractical to produce manually and then use that dataset to train a dedicated deep learning segmentation/detection model for faster, fully automated vesicle picking and more robust generalization across instruments, contrast regimes, and particle morphologies.