scientific reports

Check for updates

OPEN

Deep learning-driven automated mitochondrial segmentation for analysis of complex transmission electron microscopy images

Chan Jang[1,3], Hojun Lee[2,3], Jaejun Yoo[1]✉ & Haejin Yoon[2]✉

Mitochondria are central to cellular energy production and regulation, with their morphology tightly linked to functional performance. Precise analysis of mitochondrial ultrastructure is crucial for understanding cellular bioenergetics and pathology. While transmission electron microscopy (TEM) remains the gold standard for such analyses, traditional manual segmentation methods are time-consuming and prone to error. In this study, we introduce a novel deep learning framework that combines probabilistic interactive segmentation with automated quantification of mitochondrial morphology. Leveraging uncertainty analysis and real-time user feedback, the model achieves comparable segmentation accuracy while reducing analysis time by 90% compared to manual methods. Evaluated on both benchmark Lucchi++ datasets and real-world TEM images of mouse skeletal muscle, the pipeline not only improved efficiency but also identified key pathological differences in mitochondrial morphology between wild-type and mdx mouse models of Duchenne muscular dystrophy. This automated approach offers a powerful, scalable tool for mitochondrial analysis, enabling high-throughput and reproducible insights into cellular function and disease mechanisms.

Keywords Mitochondrial morphology, Deep learning segmentation, Transmission electron microscopy imaging, Interactive segmentation, Uncertainty analysis, Automated quantification

Mitochondria are central to cellular energy production, primarily through oxidative phosphorylation, where they regulate cellular metabolism and adapt to various environmental conditions[1]. Besides producing ATP, mitochondria are involved in many other cellular processes including calcium homeostasis, apoptosis regulation, and reactive oxygen species (ROS) status[2]. The morphology and structure dynamics of mitochondria are critical determinants of mitochondrial efficiency in energy production and overall mitochondrial functional performance. Notably, increased mitochondria size is associated with ATP synthesis, reflecting an upregulation in cellular energy storage[3]. Moreover, alterations in mitochondrial size and structure are indicative of cellular metabolic status, influencing key processes such as calcium signaling, apoptotic pathways, and ROS regulation[4]. Therefore, the assessment of mitochondrial parameters is important to understand cellular metabolism and its optimization.

Scanning electron microscopy (SEM) and transmission electron microscopy (TEM) have become indispensable tools for the qualitative characterization of morphological structures at the cellular level[5]. Among these, TEM, alongside SEM, is particularly effective in visualizing mitochondrial ultrastructure with exceptional resolution. This capability allows for a detailed analysis of morphological changes associated with different cellular metabolic states, offering critical insights into mitochondrial function and dysfunction[6]. Among the various technologies available for assessing mitochondrial structural parameters, TEM is particularly valued for its precision in measuring mitochondrial size and shape. Traditionally, the quantitative analysis of mitochondrial parameters has been conducted manually, which, although useful, is a labor-intensive process and is susceptible to variability between researchers[7]. These challenges in manual approaches can hinder both the efficiency and reproducibility of mitochondrial measurements, underscoring the need for more standardized and automated methods in mitochondrial research.

[1]Graduate School of Artificial Intelligence, Ulsan National Institute of Science & Technology, Ulsan 44919, Republic of Korea. [2]Department of Biological Sciences, Ulsan National Institute of Science & Technology, Ulsan 44919, Republic of Korea. [3]Chan Jang and Hojun Lee contributed equally to this work. ✉email: jaejun.yoo@unist.ac.kr; haejinyoon@unist.ac.kr

Recent advancements in deep learning (DL) have driven the development of several tools aimed at automating and expediting mitochondrial segmentation, significantly improving both analytical efficiency and accuracy. However, traditional supervised DL models[8–14], despite their success, often face significant limitations in real-world scenarios due to data scarcity, primarily stemming from the high costs and labor required labeled training data. To address this limitation, particularly critical in bioimaging—several strategies have emerged. One promising approach is self-supervised pre-training[15–19], which allows leveraging large-scale unlabeled datasets, such as CEM500k[20], to extract meaningful representations for downstream tasks. Another strategy is active learning, which reduces labeling demands by selectively identifying the most informative data points[21,22].

However, challenges persist, particularly in mitochondrial segmentation using electron microscopy (EM). The inherent complexity and variability of EM data often lead to suboptimal segmentation, requiring further manual verification and correction to ensure reliable analysis. Furthermore, since many segmentation methods exclude user intervention during processing, users lack control over the model's output. As a result, the only option to improve segmentation quality is labor-intensive post-processing after receiving the output, which undermines the intended goal of automation—achieving high-quality segmentation and maintaining high controllability over the system with minimal human effort.

In this study, we propose an interactive segmentation framework built on a probabilistic segmentation model[23]. A key feature of this framework is its ability to analyze uncertainty in the segmentation results, which plays a crucial role in refining the model's predictions. By leveraging uncertainty analysis, enabled by the capability of the model to generate multiple hypotheses and to highlight regions where variability is predicted, our framework identifies areas prone to errors, as uncertainty strongly correlates with these regions[24]. This enables the framework to effectively suggest where user attention or feedback should be focused during inference, allowing for a more efficient and accurate refinement process. In addition, we developed an automated tool for mitochondrial morphology quantification based on the segmentation mask, integrating both segmentation and quantification into a unified pipeline to improve overall efficiency.

To evaluate the efficacy and reliability of the proposed framework, we assessed the performance of the entire pipeline, including both segmentation and quantification processes, by comparing the results with manual methods. These evaluations were conducted using the publicly available Lucchi++ mouse hippocampus EM[25] and mouse skeletal muscle TEM dataset, which we measured and collected ourselves. This enabled validation of the approach in both benchmark and real-world settings. Importantly, using real-world data collected in-house, we demonstrated the practical utility of the framework in experimental conditions. The framework exhibited robustness in mitochondrial segmentation, along with substantial time savings. The quantification results closely mirrored those obtained via manual methods, but with a substantial reduction in time costs. These findings suggest that the proposed framework has the potential to accelerate mitochondrial research and enable comprehensive assessments of cellular energetics.

Methods and material

Mice

C57BL/6J and C57BL/10ScSn-Dmdmdx/J (mdx strain #001801) mice were obtained from The Jackson Laboratory and housed in the Animal Facility at Ulsan National Institute of Science and Technology (UNIST). All *in vivo* experiments were performed under the guidance of the Institutional Animal Care and Use Committee of the Ulsan National Institute of Science and Technology (UNIST IACUC #24-059) in this study. Mice were housed at 20–22 degrees on a 12 h light/dark cycle with ad libitum access to food (LabDiet 5053) and water. Only male mice were used for the experiment. Genotyping was performed by PCR using genomic DNA obtained from the tails. All animal studies were performed in accordance with protocols approved by the Institutional Animal Care and Use Committee, the Standing Committee on Animals at UNIST. This study is reported in accordance with the ARRIVE 2.0 guidelines to ensure transparent and reproducible reporting of animal research.

Electron microscope

Specimens were fixed with 2% glutaraldehyde—2% paraformaldehyde in 0.1 M phosphate buffer (pH 7.4) for 24 h at 4 °C and washed in 0.1 M phosphate buffer, post-fixed with 1% OsO4 in 0.1 M phosphate buffer for 1.5 h, then dehydrated in graded alcohols (50–100%), infiltrated with propylene oxide for 10 min, and embedded using a Poly/Bed 812 kit (Polysciences, USA) for 18 h. Thin sections were cut at 65 nm with a diamond knife and stained with 5% uranyl acetate for 10 min and 1% lead citrate for 5 min. The sections were analyzed using a transmission electron microscope (JEM-1011, JEOL, Japan) at an 80 kV accelerating voltage and were photographed with a digital CCD camera (RADIUS, EMSIS, Germany).

User procedure for mitochondrial segmentation and morphological quantification

Six participants with a background in biology were recruited for the experiment. They were randomly assigned to two groups, ensuring no significant differences in key factors such as experience level. Three participants performed semi-automatic segmentation and manual quantification using Adobe Photoshop and ImageJ, respectively, while the other three used a web-based graphical user interface (GUI) for a probabilistic interactive segmentation model. To prevent carryover effects and maintain consistency, each participant was assigned a unique subset of images, while all participants followed identical conditions and procedures under the guidance of the authors to minimize variability.

Segmentation evaluation

The segmentation accuracy was assessed by comparing results to gold standard masks for each dataset. In the Lucchi++ dataset, quantitative analysis was conducted by comparing generated masks with expert-provided manual labels. For mouse skeletal muscle images, which lacked a pre-existing gold standard, a qualitative

assessment was performed on both manual and deep learning (DL)-based segmentation outcomes. In manual segmentation, Adobe Photoshop was utilized, where mitochondria were identified using the "Object Selection Tool" and refined with the "Eraser Tool". For consistency in validation, the images were converted to black and white.

Quantification evaluation

Segmentation masks produced through the analysis were quantified using an automated module for mitochondrial morphological analysis. Morphological parameters manually measured from electron microscopy (EM) images using ImageJ were used as a gold standard to assess the accuracy of automated quantification. Specifically, width, height, area, and count were measured for evaluation. Given the mitochondria's asymmetric elliptical shape, two perpendicular axes were drawn: the longest axis defined as width, and the longest axis perpendicular to it as height. The outline of each mitochondrion was used to define its area. To quantify the density of mitochondria within muscle tissue samples, we calculated the number of mitochondria per unit area in transmission electron microscopy (TEM) images. Density was defined as the number of mitochondria per square micrometer (n°/μm^2), providing a measure of the concentration of mitochondria within a defined area and allowing for direct comparison between groups (Table S1). The "Straight Line" and "Freehand Selections" tools in ImageJ were applied to delineate these key morphological parameters for each EM image. To evaluate the efficiency of the analysis process, the time taken to complete each task was recorded.

Statistical analysis

Data were analyzed using R software (version 4.4.1) and GraphPad Prism 10.2.3 (GraphPad Software). The Shapiro–Wilk test was used to assess normality, and Levene's test was employed to evaluate the equality of variances. For comparisons between two samples with normally distributed data, an independent two-sample t-test (two-tailed, unpaired) was performed. For non-normally distributed data, the Mann–Whitney U test or Wilcoxon rank-sum test was applied. In cases involving comparisons among more than two groups, one-way ANOVA or Welch's ANOVA followed by a T3 multiple comparisons test was used for normally distributed data. For non-normally distributed data, the Kruskal–Wallis test followed by Dunn's post hoc test was applied. Data are presented as mean ± standard deviation (SD) in scatter plots and box plots. Statistical significance was indicated as follows: $P < 0.05$ (*), $P < 0.01$(**), and $P < 0.001$(***).

Implementation for probabilistic and interactive segmentation model and quantification method

The deep learning-driven mitochondria analysis pipeline comprises two main modules: probabilistic interactive segmentation and mask-based automated quantification. Both modules are implemented within a Python-based environment, utilizing PyTorch (version 1.11.0) and OpenCV (version 4.8.0). Additionally, the graphical user interface for the segmentation model is developed by Gradio[26] (version 4.38.1).

The segmentation model's encoder is pre-trained on the CEM 500k dataset and further fine-tuned on the training data from two datasets: Lucchi++ (165 images) and mouse skeletal muscle (21 images). During the fine-tuning phase, the model is trained using an automated feedback collection protocol to simulate user feedback. This process involves gathering feedback iteratively through the model predictions based on previous feedback, continuing until the number of feedback reaches a predefined number ($n \leq 20$ in this study). Each feedback point corresponds to the coordinate with the highest uncertainty in the model's predictions, with the feedback sign identified by the label.

For input images, a resizing operation (with a maximum resolution of 1024×1024) is applied, and patch-based training with a fixed size of 256×256 is utilized to ensure consistent learning across datasets with varying resolutions, such as Lucchi++ (764×1024) and skeletal muscle (2048×2048). Additionally, the quantification module outputs pixel-level quantitative values based on the input image resolution when measuring morphological parameters. More detailed information on the training process and model implementation is provided in Sect. 1 of the supplementary materials.

Code availability

The training and evaluation code, as well as the GUI code for our pipeline, are available at https://github.com/LAIT-CVLab/DeepPI-EM. This repository contains the necessary scripts and documentation for reproducing the results presented in this study. Any modifications or updates to the code will be maintained through this version-controlled repository.

ARRIVE compliance statement

This study is reported in accordance with the ARRIVE 2.0 guidelines for reporting animal research. The ARRIVE Essential 10 and Recommended Set have been followed to ensure transparency, reproducibility, and methodological rigor. Relevant details, including study design, sample size, randomisation, blinding, outcome measures, statistical methods, animal characteristics, housing and husbandry, and experimental procedures, are thoroughly described in the Methods and Supplementary sections. Ethical approval was obtained from the Institutional Animal Care and Use Committee (IACUC) of UNIST, and all procedures complied with institutional and national guidelines for animal welfare.

Results

Probabilistic interactive segmentation and mask-based auto quantification

Our pipeline for mitochondrial morphology analysis was developed using a probabilistic interactive segmentation model combined with a mask-based automated quantification module (Fig. 1, Figure S1). The segmentation

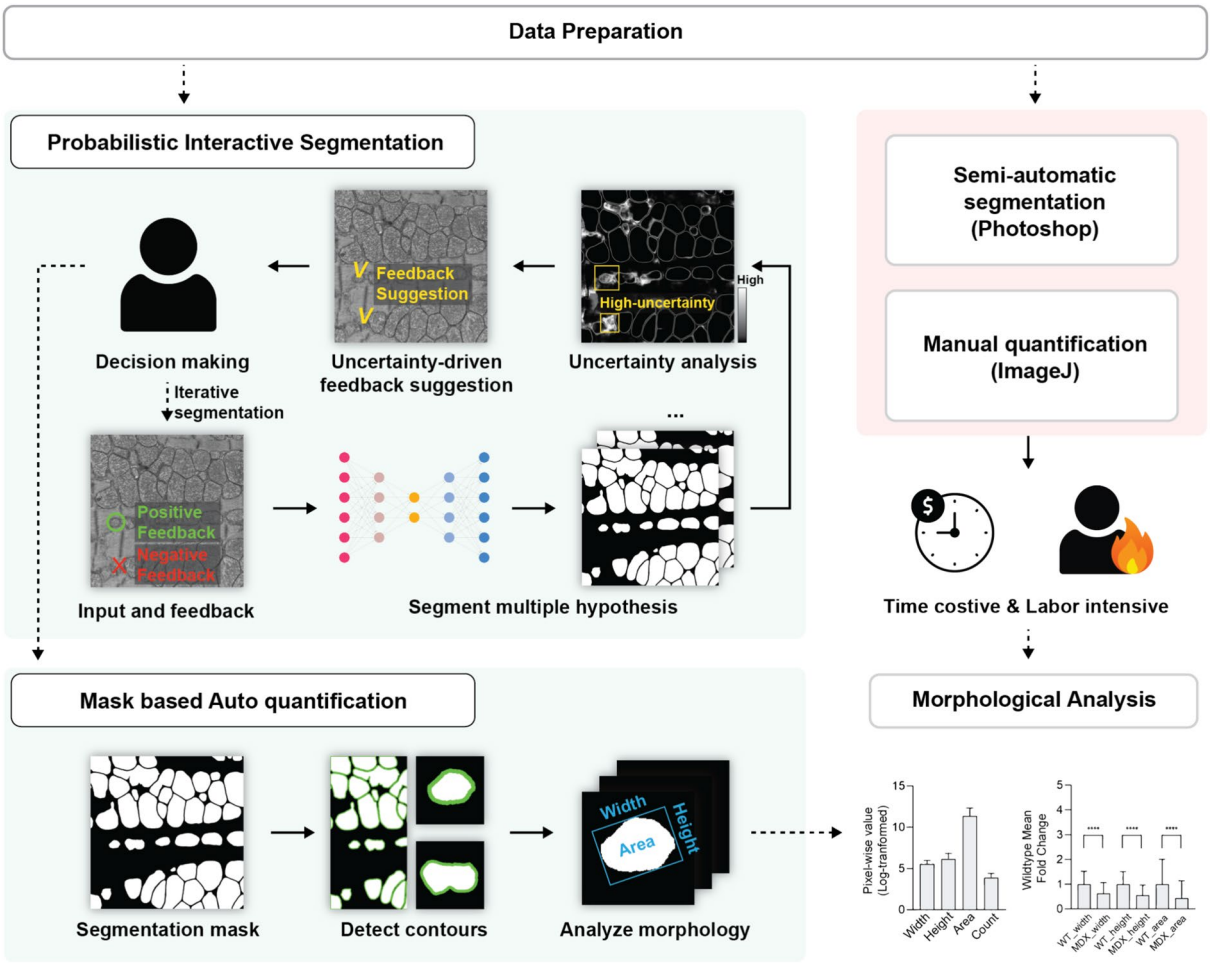


Fig. 1. Workflow comparison of probabilistic interactive segmentation and mask-based automatic quantification for mitochondrial morphology analysis with manual segmentation and quantification approach. The diagram presents two approaches for segmenting and quantifying mitochondria. The Probabilistic Interactive Segmentation method (left) integrates user feedback to generate multiple segmentation hypotheses, identifying regions of uncertainty that require further refinement. Based on this analysis, the system suggests areas for user input to refine the segmentation. The user can respond to these suggestions with positive or negative feedback, or alternatively, provide input independently. The Mask-based Automatic quantification method (bottom) performs contour detection on masks using a connected component-based approach, followed by morphological quantification of each contour through minimum bounding rectangles derived from the rotating calipers algorithm. In contrast, the Manual Segmentation and Quantification method (right) involves segmentation in Photoshop and quantification in ImageJ, demanding significant manual effort and time.

model integrates user-driven interactive segmentation with a probabilistic framework that leverages uncertainty analysis to enhance feedback incorporation. The model architecture is based on a probabilistic U-Net[23], which merges the structural advantages of a U-Net[14] with the generative capability of a conditional variational autoencoder[27] (cVAE). The U-Net encoder was pre-trained on the CEM 500k[20] dataset following a self-supervised learning protocol[17], ensuring robust feature extraction. This probabilistic segmentation framework generates multiple potential segmentation outcomes, with the final segmentation obtained by averaging the pixel-wise probabilities across all predictions. This approach improves segmentation accuracy by reducing the impact of outliers and enhancing overall consistency.

To further refine the segmentation results, a feedback suggestion module was implemented. This module identifies high-uncertainty regions within the segmentation by calculating pixel-wise standard deviations across multiple output predictions. In our study, regions representing the top 0.015% of uncertainty were targeted for further inspection. Pixel coordinates in these regions were extracted, and K-means clustering ($k=3$) was applied to identify the centroids of these uncertain areas, providing users with targeted feedback recommendations. User feedback, whether positive or negative, is converted into a distance map with the same resolution as the input image using a disk-based method, which assigns uniform binary value of 1 within the predefined radius ($r=20$) around the feedback point. This distance map was concatenated with the input image along the channel dimension and fed back into the segmentation model, improving iterative performance by addressing

uncertainty directly. Notably, the iterative interactive segmentation process based on user feedback is an optional module that users can selectively apply based on their judgment.

In the mask-based automated quantification module, mitochondrial contours were accurately identified from the binary mask results. This enabled precise quantification of key mitochondrial features, including width, height, area, and count. Contour identification was performed using the boundary-following algorithm developed by Suzuki and Abe[28], allowing for detailed object delineation. The width and height of each mitochondrion were measured using the rotating calipers algorithm[29], providing a robust framework for morphological assessment. This automated approach ensures high reproducibility and accuracy in the measurement of mitochondrial morphology, contributing to more reliable downstream analyses.

Improving mitochondrial segmentation reliability through uncertainty-driven interactive feedback

To improve the reliability of deep learning-based mitochondrial segmentation, we developed and implemented a novel interactive feedback mechanism driven by uncertainty analysis. This mechanism identifies segmentation errors and directs targeted user interventions, thereby enhancing the iterative performance of the model. Our probabilistic interactive segmentation model demonstrated significant improvements over the baseline U-Net model (Fig. 2A). By incorporating user corrections through a probabilistic interactive framework, the final segmentation is more closely aligned with the ground truth, particularly in cases of high morphological complexity or sparse training data, where the model performance is constrained, as seen in the skeletal muscle dataset. We quantitatively evaluated these improvements by comparing the Intersection over Union (IoU) scores between the previously proposed MitoSegNet[10], the baseline U-Net model, and the feedback-enhanced model using the Lucchi++ benchmark dataset (train: 165, test: 15) and the skeletal muscle dataset (train: 21, test: 6). For the Lucchi++ dataset, MitoSegNet and the baseline U-Net achieved average IoU scores of 0.91 and 0.89, respectively, while the model incorporating uncertainty-driven feedback improved this to 0.93 (Figure S2A). Similarly, for the skeletal muscle dataset, MitoSegNet and the baseline U-Net achieved average IoU scores of 0.85 and 0.87, respectively, whereas the feedback-enhanced model further improved this to 0.88 (Figure S2B). The superior segmentation performance of our pipeline was also reflected in the measurement of mitochondrial morphological parameters, demonstrating high consistency with manually annotated gold standard values across both the Lucchi++ and skeletal muscle datasets, outperforming MitoSegNet in terms of reliability and stability (Figure S4A, B).

To further assess the impact of user interactions, we first evaluated our interactive model in a user study where feedback was provided at the user's discretion without predefined constraints. While this approach demonstrated the adaptability of the model, its effectiveness is inherently dependent on user input. To provide a more standardized evaluation, we also conducted a feedback simulation experiment using ground truth masks, allowing for a controlled assessment of the pipeline's performance. Under these conditions, our model achieved superior IoU scores of 0.95 and 0.92 on the Lucchi++ and skeletal muscle datasets, respectively, using 20 feedback inputs (Figure S2C,D). Furthermore, to evaluate the generalization capability of our pipeline, we conducted additional performance assessments using the feedback simulation approach on the publicly available Kasthuri[30] (train: 85, test: 75) and MitoEM-H[31] (train: 400, test: 100) datasets. As a result, while the baseline U-Net achieved IoU scores of 0.80 and 0.90 on the Kasthuri and MitoEM-H datasets, respectively, our pipeline was able to improve the results through iterative feedback. With 20 feedback inputs, our model achieved IoU scores of 0.85 and 0.96 on the Kasthuri and MitoEM-H datasets, respectively, demonstrating its strong generalization capability (Figure S2E,F). These results highlight the effectiveness of the feedback mechanism in reducing segmentation errors and enhancing overall reliability.

A key factor in this improvement was the uncertainty-driven feedback mechanism, which systematically guided the user toward regions requiring correction. Uncertainty was quantified by calculating pixel-wise standard deviations across multiple segmentation outputs, identifying areas with the highest variation as high-uncertainty regions. By analyzing the relationship between uncertainty and error, we found a strong correlation, with high-uncertainty areas often corresponding to error-prone regions in the segmentation mask (Fig. 2B). This uncertainty-guided feedback was paired with a user interaction module, where users provided feedback through positive and negative clicks. Positive clicks reinforced correct segmentations, while negative clicks corrected errors. This feedback was incorporated as additional input into the segmentation model, alongside the original image, for use in subsequent predictions. The iterative process progressively refined the segmentation with minimal user intervention. The integration of uncertainty-driven interactive feedback significantly enhanced the deep learning-based mitochondrial segmentation process. This approach not only reduced segmentation errors but also improved the overall efficiency of the process by directing user interventions to high-uncertainty areas. This helps reduce the cognitive load on users and facilitated more meaningful feedback. Moreover, this framework significantly improved segmentation accuracy and efficiency, even under limited data conditions. As illustrated in Figure S3, models trained with only a few annotated samples (e.g., 10-shot or 15-shot) achieved substantial performance gains through interactive feedback, closely approaching the performance of a fully trained model. Taken together, the approach enabled scalable and accurate segmentation of mitochondrial morphology, making this pipeline a valuable tool for large-scale mitochondrial analysis.

Interactive mitochondria analysis pipeline achieves manual level reliability with improved efficiency

To validate the effectiveness of our proposed mitochondrial morphological analysis pipeline, we evaluated its segmentation and quantification performance using the Lucchi++ benchmarking dataset. For segmentation performance, we compared the traditional semi-automatic (SA) manual approach with the deep learning-based probabilistic interactive (PI) method. Importantly, both methods incorporate manual correction through

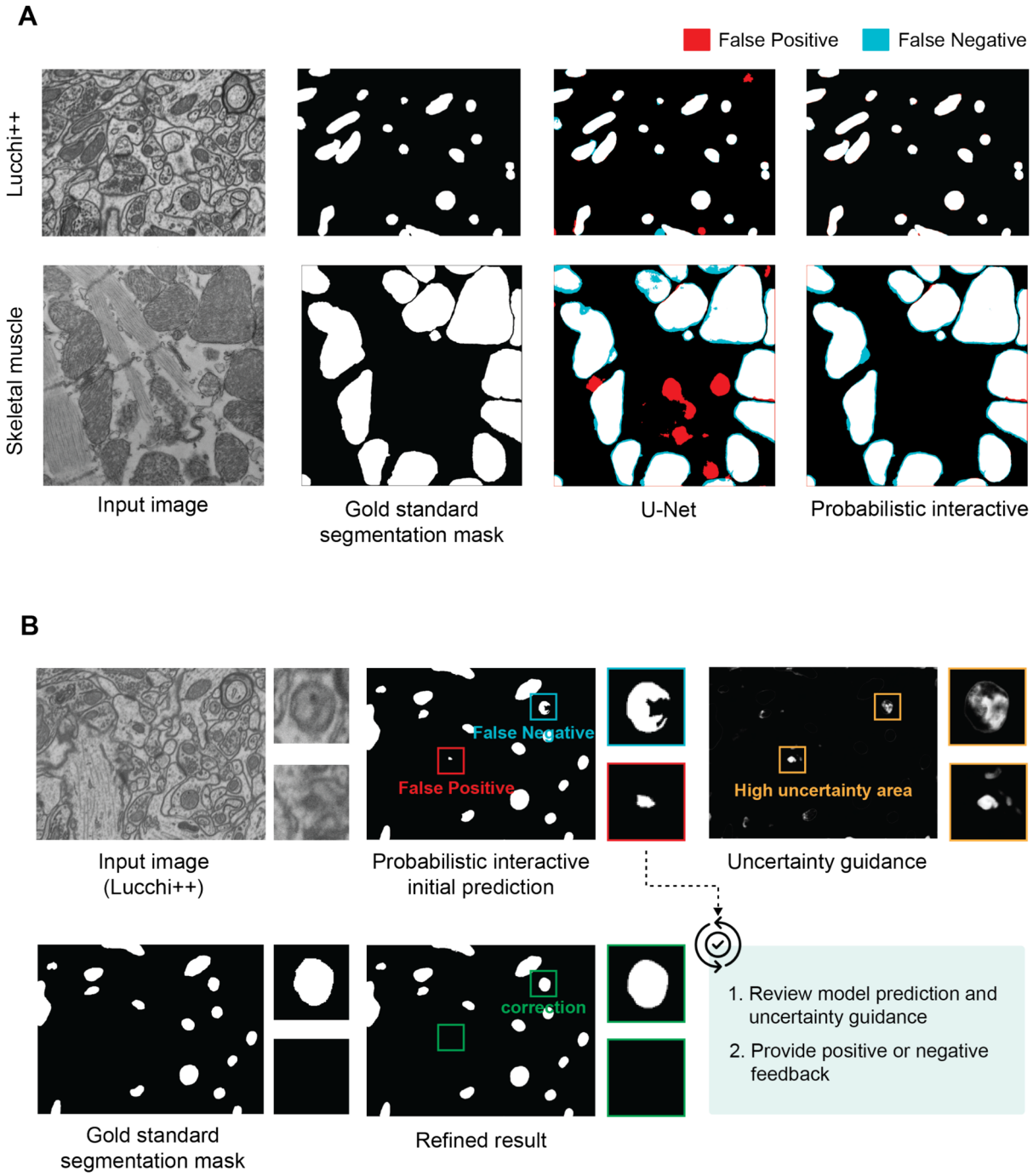


Fig. 2. Probabilistic interactive segmentation improves mitochondrial segmentation accuracy by refining uncertain regions through user feedback. (A) Segmentation results from the Lucchi++ and skeletal muscle datasets, comparing the performance of the standard U-Net model with the probabilistic interactive segmentation approach. Both models are pre-trained on the CEM500k dataset. For the Lucchi++ dataset, the gold standard segmentation mask is derived from the expert-labeled masks included in the dataset, while for the skeletal muscle dataset, it is based on manually segmented masks. (B) Visualization demonstrates how user feedback during interactive segmentation progressively enhances segmentation accuracy. Initial prediction errors are corrected with positive (blue) and negative (red) feedback clicks. The prediction errors, represented as differences between the ground truth and the model predictions, are highlighted in the uncertainty map to guide further refinements.

user intervention following the initial image segmentation, ensuring improved accuracy and reliability. For quantification performance, we compared the results from mask-based automatic quantification with those from manual quantification (Fig. 3A). This comparison provides a comprehensive understanding of how our pipeline can improve both the precision and efficiency of mitochondrial analysis, which is particularly crucial in studies where manual methods can be time-consuming and prone to errors.

Qualitatively, SA segmentation showed inconsistencies in areas such as image edges and ambiguous mitochondrial structures, where human error was more likely to occur (Fig. 3B). In contrast, PI segmentation demonstrated significantly lower error rates across all tested samples, particularly in reducing false negatives, which are common in mitochondrial segmentation due to their complex and variable morphology (Fig. 3C). This improvement can be attributed to the uncertainty-guided feedback mechanism, which directs user corrections toward regions with high uncertainty, enabling more focused and user-driven refinements. Additionally, segmentation accuracy for mitochondria detection was measured using Intersection over Union (IoU) between the predicted segmentations and the gold standard. The PI method consistently outperformed the SA segmentation approach, with a mean IoU difference of 0.1484 ± 0.02262 (SEM). This improvement was statistically significant, indicating the robustness of our method (Fig. 3D). To further evaluate instance-level performance, we measured mean average precision (mAP) across multiple IoU thresholds. On the Lucchi++ dataset, the PI model achieved a markedly higher mAP of 0.888 compared to 0.676 for the SA method, with particularly large margins at stricter thresholds (Table S2). Finally, to assess reproducibility of the segmentation model within our proposed pipeline, we conducted five repeated training runs on the Lucchi++ dataset, evaluating the baseline model's performance prior to user intervention. The results demonstrated a high level of consistency, with an accuracy of 0.9070 and a standard deviation of 0.0019, highlighting the stability of the model across multiple training iterations.

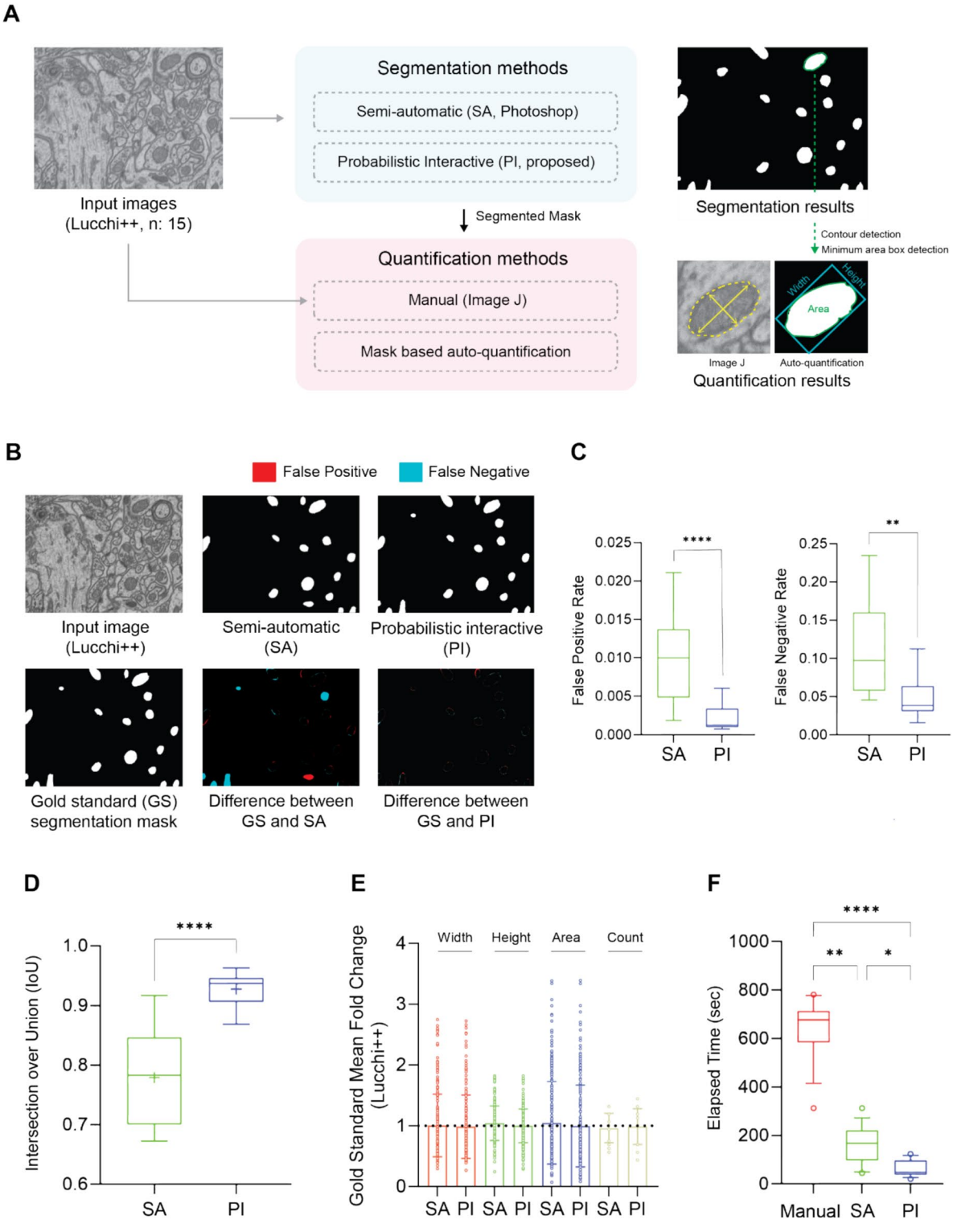
To evaluate the accuracy of our pipeline for mitochondrial morphological analysis, we assessed its performance in measuring key parameters such as mitochondrial width, height, area, and count. Manual quantifications performed in ImageJ were treated as the gold standard and compared to the results generated by the automated quantification module applied to both SA segmentation masks and PI segmentation masks. The automated quantification applied to segmentation masks yielded results that were not significantly different from those obtained via manual ImageJ measurements, even when relying solely on uncertainty-driven feedback (Fig. 3E and Figure S4C). This close alignment underscores the reliability of the mask-based automated quantification approach, demonstrating that automation can maintain high fidelity to gold standard-level quantification techniques. Moreover, the PI segmentation and automated quantification pipeline provided substantial time savings compared to both the manual quantification and segmentation methods (Fig. 3F). These results highlight the efficiency of the pipeline, making it a valuable tool for large-scale mitochondria morphological analysis.

Consistent performance of interactive mitochondria analysis in complex skeletal muscle EM image

To assess the scalability and robustness of the proposed mitochondrial morphology analysis pipeline, we applied it to six transmission electron microscopy (TEM) images of mouse skeletal muscle. The analysis followed the same methodological framework as with the Lucchi++ dataset (Fig. 3A). The results from the TEM images showed similar trends in reliability, and time efficiency to those observed with scanning electron microscopy (SEM) images (Fig. 3), although the increased complexity of the TEM dataset revealed several noteworthy insights.

In terms of segmentation performance, the probabilistic interactive (PI) segmentation model demonstrated superior precision, particularly in capturing intricate and irregular mitochondrial boundaries. When mitochondrial structures were clearly delineated, the PI models outperformed semi-automatic (SA) methods, especially in recognizing complex, irregularly shaped mitochondria. However, in cases where mitochondria appeared ruptured or exhibited poorly defined boundaries, the PI models struggled to segment these regions accurately. Despite these challenges, the PI model not only demonstrated a clear advantage in handling most mitochondrial structures but also provided uncertainty-based cues to guide users toward further review, facilitating more efficient and accurate identification of mitochondria (Fig. 4A and Figure S6). A quantitative segmentation performance analysis based on the Intersection over Union (IoU) metric showed that the PI model achieved an average IoU score of 0.88, which was nearly equivalent to the performance of the SA method (Fig. 4B). Furthermore, as shown in Table S2, the PI model achieved a higher mean average precision (mAP) of 0.750 compared to 0.720 for the SA method in the instance-level segmentation evaluation, with notably better AP scores at higher IoU thresholds. Regarding the reproducibility of the segmentation model within our proposed pipeline, we conducted 5 repeated training runs on the skeletal muscle TEM dataset, evaluating our proposed model's performance before user intervention. The results demonstrated a high level of consistency, with an average IoU score of 0.8745 and a standard deviation of 0.0035, indicating the robustness and stability of the model across different training iterations. This findings suggest that uncertainty-guided segmentation is a scalable and effective strategy for analyzing for complex tissue environments.

For mitochondrial morphological quantification, the results show high consistency across the approaches, with no significant differences in the measurements even when using only uncertainty-driven feedback (Fig. 4C and Figure S4D). These results confirm that the PI segmentation can reliably replicate the accuracy of gold standard quantification, even in more complex tissue environments where fewer training samples are available. In addition, the automated method dramatically reduced the time required for quantification compared to the manual method (Fig. 4D). To sum up, this increased efficiency is critical for large-scale studies, where manual quantification would be prohibitively time-consuming.



Pathological difference between mdx mouse and wild type is uncovered via interactive mitochondria analysis pipeline

Duchenne muscular dystrophy (DMD) is a fatal X-linked disorder caused by the absence of dystrophin, leading to progressive muscle degeneration and weakness[32]. Mitochondrial dysfunction plays a key role in DMD pathology. Notably, mitochondrial size and density were significantly reduced in mdx mice, indicating compromised bioenergetics in DMD[33]. To examine mitochondrial alterations, we utilized quadriceps muscle samples from mdx mice, a well-established DMD model, and wild-type (WT) controls. Our analysis pipeline designed for high-resolution transmission electron microscopy (TEM) images, described in Fig. 4, enabled accurate mitochondrial segmentation and quantification for mitochondrial morphological analysis. The dataset included 18 images from mdx (muscular dystrophy X-linked) mice and 27 images from WT controls (Fig. 5A), analyzed using a probabilistic segmentation framework integrated with an uncertainty-guided feedback

◀**Fig. 3**. Comparison of segmentation and quantification methods for Lucchi++ benchmark dataset validating the efficiency of the mask-based automated quantification module. (**A**) Overview of segmentation and quantification methods tested on the Lucchi++ dataset. The input image is segmented using a manual method using the semi-automated tool or the deep learning (DL)-based probabilistic interactive model. The resulting segmentation masks are used for mask-based automated quantification, while manual quantification with ImageJ is performed directly on the input image. (**B**) Segmentation results compared to the gold standard. Segmentation masks are displayed for the semi-automatic method and probabilistic interactive model. Differences from the gold standard are shown, with false positives in red, false negatives in blue, true positives in white, and true negatives in black. (**C**) The use of a probabilistic interactive model resulted in a reduction of false positives and false negatives were reduced through (mean ± SD, $p = 2.0 \times 10^{-8}$, $p = 2.0 \times 10^{-4}$, Mann–Whitney U test. N = 15). (**D**) Comparison of IoU values derived from semi-automated (SA) and probabilistic interactive (PI) segmentation (whiskers show min-to-max range, $p = 4.4 \times 10^{-6}$, independent two-sample t-test. N = 15). Arithmetic mean is indicated by a + symbol in each box plot. (**E**) Comparison of fold changes in quantified mitochondrial morphological parameters (height, width, area, and count) relative to the gold standard. Multiple parameters/variables were extracted from each image. The gold standard is based on manual quantification using ImageJ, whereas the semi-automatic (SA) segmentation and probabilistic interactive (PI) model results were derived from automatic quantification modules applied to segmentation masks obtained by each respective method. (mean ± SD, independent two-sample t-test. N = 230–235). (**F**) Box and whisker plot illustrating the elapsed time for each method (whiskers show 10–90 percentile, $p = 2.0 \times 10^{-8}$, Kruskal–Wallis test followed by Dunn's multiple comparisons test. N = 15). The average time for the manual quantification method was 641.9 s, while the probabilistic interactive and automatic quantification method required an average of 62.33 s, indicating a 90.3% reduction in analysis time when using the method. Statistical significance is indicated by asterisks.

mechanism. To quantify mitochondrial density within muscle tissue samples[34], we calculated the number of mitochondria per unit area in TEM images under each experimental condition (Table S1).

Our framework effectively segmented mitochondria in both WT and mdx mouse models (Figure S7), revealing statistically significant reductions in key mitochondrial metrics (width, height, area, and density) in mdx mice compared to WT controls (Fig. 5B). These findings confirm that mdx mice exhibit altered mitochondrial morphology consistent with previous findings. Through these measurements, we can conclude that our analytical framework enables precise measurement of changes directly associated with these morphological characteristics, providing a robust quantitative assessment of how pathological traits manifest within biological systems. By identifying specific alterations in physiological processes driven by pathological factors, this framework offers a powerful tool for evaluating the impact of various diseases on biological function.

Discussion

In this study, we introduced an interactive mitochondrial analysis model that integrates deep learning with user feedback to enhance the efficiency of mitochondrial segmentation and quantification. Our probabilistic interactive segmentation model, in combination with a mask-based automatic quantification module, demonstrated a substantial improvement in both benchmark and real-world datasets. Conventional deep learning-based pipelines for mitochondria analysis, which do not support user intervention during segmentation , our approach enables real-time user refinement, achieving expert-level segmentation accuracy within a deep learning framework. Furthermore, the uncertainty-driven feedback mechanism optimizes user effort by providing targeted suggestions, facilitating an efficient and reliable segmentation process, ultimately making it a valuable tool for scalable mitochondrial analysis, particularly in pathologically relevant models such as the mdx mouse model of Duchenne muscular dystrophy. One key advantage of this framework is its ability to direct user interventions toward high-uncertainty regions, allowing for more efficient corrections and refined segmentation outcomes. However, an inherent limitation is that regions requiring correction are not always accurately captured by the uncertainty map. Although the model's uncertainty analysis identifies many error-prone areas, certain regions may still require adjustments even when they are not flagged as high uncertainty. For instance, regions with subtle boundaries or low-contrast regions relative to the surrounding cellular structures may evade detection by the uncertainty-guided feedback process. During our experiments, such cases were indeed observed, especially in areas with complex mitochondrial morphology or where mitochondria bordered other densely packed organelles.

To address this limitation in future iterations, incorporating additional image features, such as texture or intensity gradients[35], could help improve the identification of challenging regions. Another potential solution could be to allow for user-defined adjustments based on visual assessment, even in low-uncertainty regions. This would provide flexibility in managing cases where automated uncertainty metrics are insufficient.

Acknowledging this limitation clarifies that, while our framework effectively reduces the need for extensive manual intervention, complete automation, as well as the analysis of fine structures within mitochondria that lack separate labels, remains challenging in highly complex cellular environments. Future improvements to the probabilistic segmentation model could involve incorporating domain-specific self-supervised techniques[36–38] or adaptation techniques[39–41] to better handle ambiguous structures with limited labeled data.

Overall, our interactive deep learning framework represents a significant advancement in mitochondrial analysis and holds promise extension to other organelles or application domains, offering a scalable approach for high-throughput applications. However, acknowledging its limitations and exploring strategies to address these

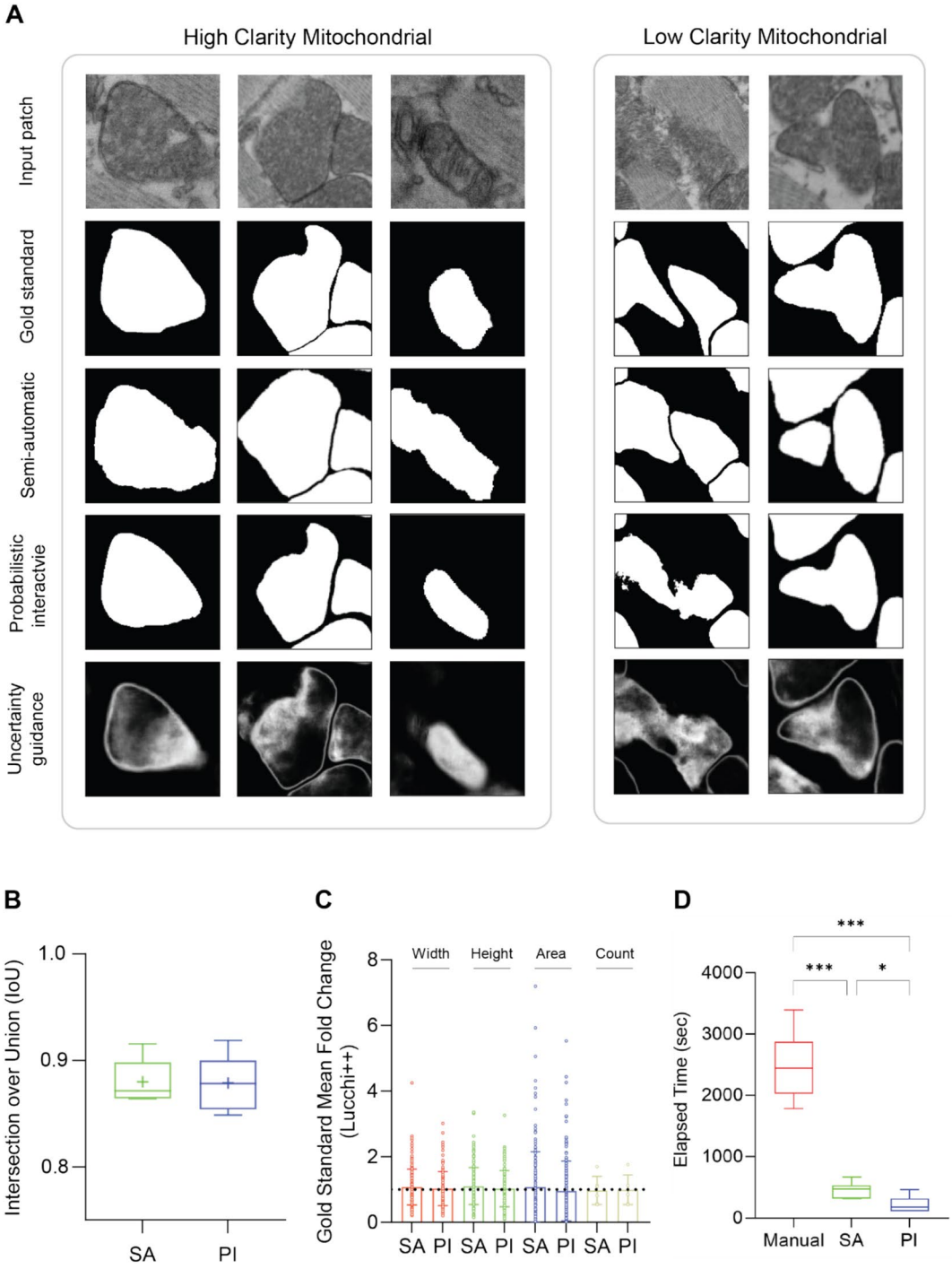


Fig. 4. Uncertainty-guided segmentation ensures efficient mitochondrial quantification in complex skeletal muscle TEM images. (**A**) Comparison between semi-automatic and probabilistic interactive segmentation results for mitochondria with varying levels of structural clarity. The bottom row shows the uncertainty maps generated by the probabilistic interactive model, where brighter regions indicate areas of higher uncertainty, guiding user attention. (**B**) Comparison of IoU values derived from semi-automated (SA) and probabilistic interactive (PI) segmentation (whiskers show 10–90 percentile range, N = 6). The arithmetic mean is indicated by a + symbol in each box plot. (**C**) Fold change comparison of mitochondrial morphological parameters (height, width, area, and count) relative to the gold standard. Multiple parameters were measured for each image. Manual quantification using ImageJ was considered the gold standard, while semi-automatic (SA) and probabilistic interactive (PI) methods produced results via automated quantification based on segmentation masks from each method (mean ± SD, independent two-sample t-test, N = 172–176). (**D**) Box-and-whisker plot displaying the elapsed time for each method (whiskers represent the 10–90 percentile, $p = 1.1 \times 10^{-8}$, Welch ANOVA followed by Dunnett's T3 multiple comparisons test, N = 6). The average time for manual segmentation was 2480 s, while the probabilistic interactive method averaged 214 s, representing a 91.4% reduction in analysis time. Statistical significance is indicated by asterisks.

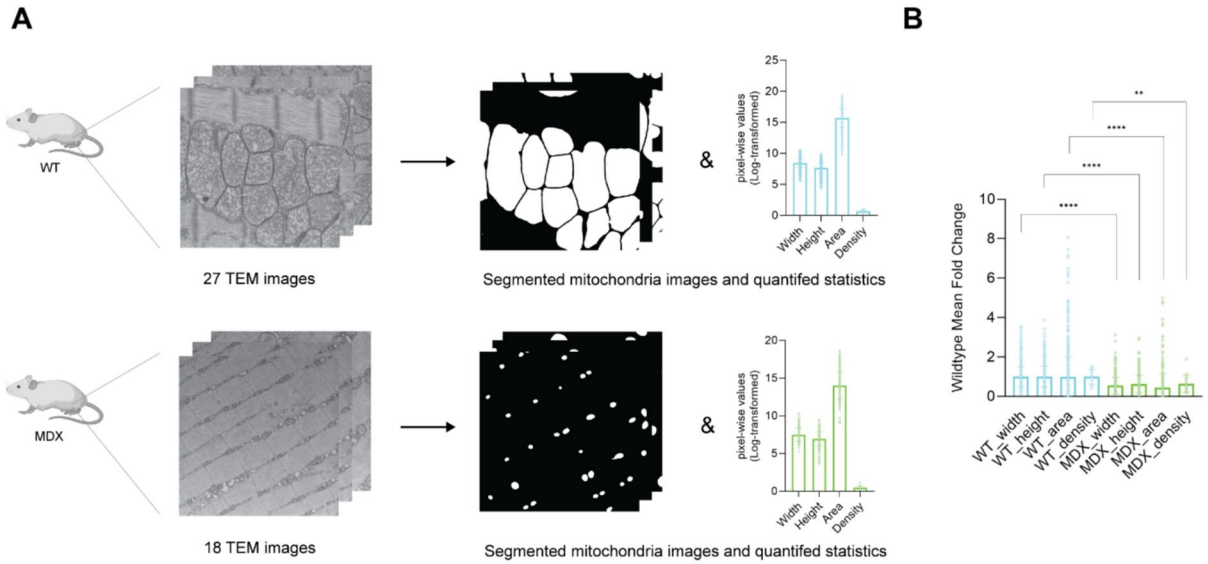


Fig. 5. Mitochondrial analysis pipeline reveals pathological differences between wild type and mdx mice. (**A**) Schematic figure of mitochondrial segmentation and morphological quantification from TEM images of the quadriceps muscle in wild type and mdx mice. Left are representative TEM images of mitochondria from wild-type and mdx mice. Right are corresponding results from mitochondrial segmentation and morphological quantification (Log-transformed). (**B**) Box and whiskers plot showing morphological differences in mitochondria between wild-type and mdx mice. Mdx mice exhibit significantly altered mitochondrial features compared to wild-type mice (mean ± SD, $p = 2.2 \times 10^{-16}$, independent two-sample t-test. N = 18–27). Statistical significance is indicated by asterisks. To accommodate values below 1 and to prevent negative outputs during log transformation, we added a constant offset of 1 to all measurements before applying the log transformation.

challenges will be essential for further optimization and robustness of this methodology in diverse biological contexts.

Data availability

The datasets generated during and/or analysed during the current study are available in the following link: https://yoonlab.unist.ac.kr/index.php/research/mitochondria-tem-dataset/

Received: 27 January 2025; Accepted: 20 May 2025
Published online: 30 May 2025

References

1. Spinelli, J. B. & Haigis, M. C. The multifaceted contributions of mitochondria to cellular metabolism. *Nat. Cell Biol.* **20**, 745–754 (2018).
2. Paupe, V. & Prudent, J. New insights into the role of mitochondrial calcium homeostasis in cell migration. *Biochem. Biophys. Res. Commun.* **500**, 75–86 (2018).
3. Neurohr, J. M., Paulson, E. T. & Kinsey, S. T. A higher mitochondrial content is associated with greater oxidative damage, oxidative defenses, protein synthesis and ATP turnover in resting skeletal muscle. *J. Exp. Biol.* **224**, jeb242462 (2021).
4. Hajnóczky, G. et al. Mitochondrial calcium signalling and cell death: Approaches for assessing the role of mitochondrial Ca2+ uptake in apoptosis. *Cell Calcium* **40**, 553–560 (2006).
5. Neikirk, K. et al. Call to action to properly utilize electron microscopy to measure organelles to monitor disease. *Eur. J. Cell Biol.* **102**, 151365 (2023).
6. Collins, H. E. et al. Mitochondrial morphology and mitophagy in heart diseases: qualitative and quantitative analyses using transmission electron microscopy. *Front. Aging* **2**, 670267 (2021).
7. Lam, J. et al. A universal approach to analyzing transmission electron microscopy with ImageJ. *Cells* **10**, 2177 (2021).
8. Zaghbani, S., Pranti, R. K., Faber, L. & Garcia-Saez, A. J. MitoSkel: AI tool for semantic segmentation and quantification of mitochondria from light microscopy images. *Biomed. Signal Process. Control* **106**, 107762 (2025).
9. Vojtová, J. et al. A fully automated morphological analysis of yeast mitochondria from wide-field fluorescence images. *Sci. Rep.* **14**, 30144 (2024).
10. Fischer, C. A. *et al.* MitoSegNet: easy-to-use deep learning segmentation for analyzing mitochondrial morphology. *iScience* **23**, (2020).
11. Franco-Barranco, D., Pastor-Tronch, J., González-Marfil, A., Muñoz-Barrutia, A. & Arganda-Carreras, I. Deep learning based domain adaptation for mitochondria segmentation on EM volumes. *Comput. Methods Programs Biomed.* **222**, 106949 (2022).
12. Hultgren, N. W., Zhou, T. & Williams, D. S. Machine learning-based 3D segmentation of mitochondria in polarized epithelial cells. *Mitochondrion* **76**, 101882 (2024).
13. Nawabi, A. K. et al. Segmentation of drug-treated cell image and mitochondrial-oxidative stress using deep convolutional neural network. *Oxid. Med. Cell. Longev.* **2022**, 5641727 (2022).
14. Ronneberger, O., Fischer, P. & Brox, T. U-Net: convolutional networks for biomedical image segmentation. *Med. Image Comput. Comput. Assist. Interv. MICCAI 2015*, 234–241 (2015). https://doi.org/10.1007/978-3-319-24574-4_28.

15. Sekh, A. A. et al. Physics-based machine learning for subcellular segmentation in living cells. *Nat. Mach. Intell.* **3**, 1071–1080 (2021).
16. Chen, T., Kornblith, S., Norouzi, M. & Hinton, G. A Simple Framework for Contrastive Learning of Visual Representations. In *Proceedings of the 37th International Conference on Machine Learning* **119**, 1597–1607 (2020).
17. Chen, X., Fan, H., Girshick, R. & He, K. Improved Baselines with Momentum Contrastive Learning. *arXiv* preprint arXiv:2003.04297 (2020). https://doi.org/10.48550/arXiv.2003.04297.
18. He, K. *et al.* Masked Autoencoders Are Scalable Vision Learners. *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.* 16000–16009 (2022).
19. Oquab, M. *et al.* DINOv2: Learning robust visual features without supervision. *arXiv* preprint arXiv:2304.07193 (2024). https://doi.org/10.48550/arXiv.2304.07193.
20. Conrad, R. & Narayan, K. CEM500K, a large-scale heterogeneous unlabeled cellular electron microscopy image dataset for deep learning. *Elife* **10**, e65894 (2021).
21. Rombaut, B., Roels, J. & Saeys, Y. BioSegment: Active Learning segmentation for 3D electron microscopy imaging. *arXiv* preprint arXiv:2203.14289 (2022). https://doi.org/10.48550/arXiv.2203.14289.
22. Suga, S. et al. An interactive deep learning-based approach reveals mitochondrial cristae topologies. *PLoS Biol.* **21**, e3002246 (2023).
23. Kohl, S. *et al.* A Probabilistic U-Net for segmentation of ambiguous images. *Adv. Neural Inf. Process. Syst.* **31** (2018).
24. Czolbe, S., Arnavaz, K., Krause, O. & Feragen, A. Is segmentation uncertainty useful?. *Inf. Process. Med. Imaging* **12829**, 715–726 (2021).
25. Casser, V., Kang, K., Pfister, H. & Haehn, D. Fast mitochondria detection for connectomics. *Proc. 3rd Conf. Med. Imaging Deep Learn.* 111–120 (2020).
26. Abid, A. *et al.* Gradio: Hassle-Free Sharing and Testing of ML Models in the Wild. *arXiv* preprint arXiv:1906.02569 (2019). https://doi.org/10.48550/arXiv.1906.02569.
27. Higgins, I. *et al.* beta-vae: Learning basic visual concepts with a constrained variational framework. *ICLR Poster* **3**, (2017).
28. Suzuki, S. & Abe, K. Topological structural analysis of digitized binary images by border following. *Comput. Vis. Graph. Image Process.* **30**, 32–46 (1985).
29. Toussaint, G. T. Solving geometric problems with the rotating calipers. *Proc. IEEE Melecon* **83**, A10 (1983).
30. Kasthuri, N. et al. Saturated Reconstruction of a Volume of Neocortex. *Cell* **162**, 648–661 (2015).
31. Wei, D. et al. MitoEM dataset: Large-scale 3D mitochondria instance segmentation from EM images. *Med. Image Comput. Comput. Assist. Interv.* **11767**, 66–78 (2019).
32. Verhaart, I. E. C. & Aartsma-Rus, A. Therapeutic developments for Duchenne muscular dystrophy. *Nat. Rev. Neurol.* **15**, 373–386 (2019).
33. Percival, J. M., Siegel, M. P., Knowels, G. & Marcinek, D. J. Defects in mitochondrial localization and ATP synthesis in the mdx mouse model of Duchenne muscular dystrophy are not alleviated by PDE5 inhibition. *Hum. Mol. Genet.* **22**, 153–167 (2013).
34. Federico, M. et al. CaMKII activation in early diabetic hearts induces altered sarcoplasmic reticulum-mitochondria signaling. *Sci. Rep.* **11**, 20025 (2021).
35. Wang, X. *et al.* MitoStructSeg: A Comprehensive Platform for Mitochondrial Structure Segmentation and Analysis. *bioRxiv* preprint 2024.06.28.601295 (2024). https://doi.org/10.1101/2024.06.28.601295.
36. Wagner, R., Lopez, C. F. & Stiller, C. Self-supervised pseudo-colorizing of masked cells. *PLoS ONE* **18**, e0290561 (2023).
37. Wolf, S., Lalit, M., McDole, K. & Funke, J. Unsupervised Learning of Object-Centric Embeddings for Cell Instance Segmentation in Microscopy Images. In *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 21263–21272 (2023).
38. Xie, R., Pang, K., Bader, G. D. & Wang, B. MAESTER: Masked Autoencoder Guided Segmentation at Pixel Resolution for Accurate, Self-Supervised Subcellular Structure Recognition. *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.,* 3292–3301 (2023).
39. Archit, A. & Pape, C. Probabilistic Domain Adaptation for Biomedical Image Segmentation. *arXiv* preprint arXiv:2303.11790 (2023). https://doi.org/10.48550/arXiv.2303.11790.
40. Liu, D. et al. PDAM: A panoptic-level feature alignment framework for unsupervised domain adaptive instance segmentation in microscopy images. *IEEE Trans. Med. Imaging* **40**, 154–165 (2021).
41. Stenger, A. *et al.* Fast and interpretable unsupervised domain adaptation for FIB-SEM cell segmentation. *Proc. IEEE Int. Symp. Biomed. Imaging (ISBI),* 1–5 (2023). https://doi.org/10.1109/ISBI53787.2023.10230327.
42. Percie du Sert, N. *et al.* The ARRIVE guidelines 2.0: Updated guidelines for reporting animal research. *PLoS Biol.* **18**, e3000410 (2020).

Acknowledgements

We thank the members of Haejin Yoon's lab and Jaejun Yoo's lab for their valuable feedback on project development. We also extend our gratitude to David Joon Ho (Department of Cancer AI & Digital Health, Graduate School of Cancer Science and Policy, National Cancer Center) for his insightful discussions on this project. We also thank Hongjae You, San Kwon, Minseo Hong, Jaehyun Kim, Yoonseo Joung for their contributions to mitochondria segmentation and quantification. We further acknowledge Sungho Hwang and Eunji Lee for their dedicated efforts in generating Ground Truth annotations for skeletal muscle dataset. H.Y. is funded by Settlement Research Fund (1.220105.01) of UNIST (Ulsan National Institute of Science & Technology), as well as by the National Research Foundation of Korea(NRF) grant funded by the Korea government (MSIT) (2022R1C1C1010618), and supported by Basic Science Research Program through National Research Foundation of Korea (NRF) funded by the Ministry of Education (2022R1A4A1021318). Also, J. Y is funded by National Research Foundation of Korea (NRF) grant (No.2022R1C1C100849612) and Institute of Information & communications Technology Planning & Evaluation (IITP) grant (No.2022-0-00959 / RS-2022-II220959, (Part 2) Few-Shot Learning of Causal Inference in Vision and Language for Decision Making), supported by the Korea government (MSIT). We thank Match Finder Co., Ltd. for their valuable support concerning electron microscopy.

Author contributions

H.Y. and J.Y. designed experiments, wrote the paper, and analyzed data. C.J. generated a pipeline for deep learning-driven automated mitochondrial segmentation for analysis of TEM images. H.L. performed and analyzed mitochondrial TEM images.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-03311-1.

Correspondence and requests for materials should be addressed to J.Y. or H.Y.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

© The Author(s) 2025