MitoHub: Mitochondrial segmentation and Mobility estimation using YOLO and Optical flow Techniques

Shreyasvi Natraj[1,*], **Koen Wentinck**[1], **and Tatjana Kleele**[1]

[1]Institute of Biochemistry, Department of Biology, ETH Zürich, Zurich, 8093, Switzerland
[*]snatraj@ethz.ch

ABSTRACT

Mitochondria are central to cellular function and health, with their dynamics serving as crucial indicators of cellular status and overall vitality. In this study, we introduce MitoHub, a comprehensive pipeline designed to advance mitochondrial analysis through cutting-edge techniques. MitoHub integrates the YOLO neural network architecture for precise mitochondrial segmentation, enhancing the accuracy of identifying and delineating mitochondrial structures within complex microscopy images. To complement segmentation, MitoHub employs sparse optical flow algorithm, combining the Shi-Tomasi corner detection method with the Lucas-Kanade optical flow algorithm to accurately estimate mitochondrial mobility. This integration enables the detailed mobility estimation of mitochondrial movement, providing insights into their dynamic behavior. The effectiveness of MitoHub is demonstrated through rigorous experiments, which reveal its high performance in both segmentation and mobility estimation. Notably, the segmentation neural network models achieve a Mask Mean Average Precision ($mAP@50$) score of 93.2% using YOLOv8-X and a score of 93.5 using YOLOv9-E segmentation neural network architecture, underscoring its robustness and accuracy in mitochondrial segmentation. MitoHub's advanced capabilities offer a significant advancement in mitochondrial research, facilitating deeper insights into mitochondrial function and contributing to the broader understanding of cellular health and dynamics.

Introduction

Mitochondria are dynamic organelles essential for cellular energy production, apoptosis, and metabolic regulation. These multifaceted structures are the powerhouses of the cell, generating adenosine triphosphate (ATP) through oxidative phosphorylation, a process that is central to cellular metabolism and energy homeostasis[1]. Beyond their well-known role in bioenergetics, mitochondria are also key players in various cellular processes, including the regulation of metabolic pathways, signaling, and cell death mechanisms[2].

Mitochondrial dynamics refer to the continuous changes in mitochondrial shape, size, distribution, and movement within the cell. Mitochondrial mobility is a crucial aspect of these dynamics, involving the movement of mitochondria along the cytoskeleton to different regions of the cell where energy demand is high. This mobility is facilitated by motor proteins such as kinesin and dynein, which transport mitochondria along microtubules, and myosin, which moves them along actin filaments[3]. The ability of mitochondria to relocate within the cell is essential for maintaining cellular homeostasis and responding to metabolic needs, ensuring that regions of high energy demand, such as the synapses in neurons or the contractile apparatus in muscle cells, are adequately supplied with ATP[4].

The dynamic behavior of mitochondria is also vital for their quality control. Mitochondrial movement allows for the distribution of damaged or dysfunctional mitochondria to autophagic machinery for degradation, a process known as mitophagy. This selective removal of damaged mitochondria is critical for preventing the accumulation of mitochondrial defects, which can lead to cellular dysfunction and contribute to aging and disease[5]. Moreover, mitochondrial dynamics play a pivotal role in cellular responses to stress. In times of cellular stress, such as oxidative stress or nutrient deprivation, mitochondria can undergo morphological changes that help to optimize their function and protect the cell. For example, during periods of oxidative stress, mitochondria often elongate through fusion, which can help to mitigate damage by enhancing their ability to buffer calcium and produce ATP more efficiently[6].

Disruptions in mitochondrial dynamics are linked to a wide array of diseases. In neurodegenerative disorders such as Parkinson's and Alzheimer's diseases, impaired mitochondrial dynamics can lead to deficits in energy production, increased oxidative stress, and the accumulation of damaged mitochondria, contributing to neuronal death and disease progression[7]. In cancer, altered mitochondrial dynamics can influence metabolic reprogramming and resistance to apoptosis, enabling cancer cells to proliferate and survive in unfavorable conditions[8]. Given the centrality of mitochondrial dynamics to cellular health and disease, accurate detection and analysis of these processes are essential. Advances in imaging techniques, such as super-resolution microscopy and live-cell imaging, have allowed for detailed visualization of mitochondrial behavior in real-

time[9]. These techniques, combined with computational tools for image analysis, enable researchers to quantify mitochondrial dynamics with high precision, shedding light on the underlying mechanisms and potential therapeutic targets.

Mitochondrial mobility and dynamics are fundamental to cellular function and adaptation. They facilitate energy distribution, quality control, stress responses, and cellular signaling. Understanding these dynamic processes is crucial for unraveling the complex interplay between mitochondria and cellular health, providing insights into the pathogenesis of various diseases and informing the development of targeted therapies.

Studies such as Mitometer[10] and MitoGraph[11] have laid significant groundwork in the automated analysis of mitochondrial morphology and dynamics. Mitometer, for instance, employs traditional image processing techniques and has proven effective in quantifying mitochondrial shape and network properties. However, it lacks the real-time processing capabilities and advanced segmentation precision offered by modern deep learning approaches. Similarly, MitoGraph uses a combination of segmentation algorithms to extract mitochondrial networks, but it still relies heavily on user input for accurate analysis. Other tools like MiNA (Mitochondrial Network Analysis)[12] provide automated methods for assessing mitochondrial morphology and network characteristics, yet the quality and consistency of manual annotations often constrain them.

In the realm of mobility estimation, tools like ImageJ's TrackMate plugin[13] have been employed to track mitochondrial movements. TrackMate uses a feature-point detection approach combined with a frame-by-frame linking algorithm to estimate mitochondrial trajectories. However, its reliance on user-defined parameters can introduce variability and bias in the results. On the other hand, particle image velocimetry (PIV)-based methods, such as those implemented in the software package PIVlab[14], offer quantitative analysis of flow patterns in live cell imaging. Despite their utility, these methods often require extensive preprocessing and are not specifically tailored for mitochondrial analysis.

Furthermore, while other methods like Mitometer and MitoGraph have primarily focused on static images to analyze mitochondrial morphology, MitoHub's capability to process time-lapse microscopy data enables the continuous monitoring of mitochondrial dynamics. This is crucial for studying the transient and complex behaviors of mitochondria within living cells. The ability to capture these dynamic changes in real-time provides a more accurate and detailed understanding of mitochondrial functions and their alterations in disease states.

Other notable tools for mitochondrial segmentation include MitoSegNet[15], a deep learning-based tool that uses U-Net architecture for segmenting mitochondrial images. MitoSegNet has shown high accuracy in delineating mitochondria in various cell types but lacks real-time processing capabilities.

MitoHub distinguishes itself by leveraging the state-of-the-art YOLO-segmentation model, which significantly enhances the accuracy and speed of detecting mitochondrial fission and fusion events compared to traditional methods. The integration of YOLO[16] models into the pipeline allows for the simultaneous detection and segmentation of mitochondria in real-time, providing a substantial improvement over previous approaches that required sequential processing steps. Additionally, by incorporating sparse optical flow techniques, MitoHub not only detects but also quantifies mitochondrial mobility, offering a comprehensive analysis of mitochondrial dynamics that previous methods lack.

The novelty of MitoHub lies in its holistic approach, combining advanced deep learning techniques with sophisticated motion analysis tools to provide a robust platform for mitochondrial research. By automating the detection, segmentation, and quantification processes, MitoHub reduces the potential for human error and increases the efficiency and reproducibility of mitochondrial studies. This comprehensive pipeline not only streamlines the analysis workflow but also opens new avenues for research by enabling high-throughput and detailed investigations of mitochondrial dynamics in various pathological conditions.

Traditional methods for analyzing mitochondrial dynamics often rely on manual annotation, which is time-consuming and prone to human error. Advances in deep learning, particularly convolutional neural networks (CNNs), offer promising solutions for automating these tasks. YOLO (You Only Look Once) models[17], known for their fast and accurate segmentation capabilities, provide an efficient framework for segmenting mitochondrial structures and identifying their mobility dynamics.

In this paper, we introduce MitoHub, a comprehensive pipeline integrating segmentation and mobility estimation. The system incorporates state-of-the-art YOLO-based segmentation neural networks and sparse optical flow methods to quantify mitochondrial mobility, providing a holistic tool for studying mitochondrial dynamics.

In addition to the YOLOv8-X[18] segmentation model, this paper introduces the incorporation of the YOLOv9-E[19] segmentation model and compares the results between v8 and v9. The YOLOv9-E model, with its enhanced architecture and improved accuracy, further refines the detection and segmentation capabilities of MitoHub, making it a cutting-edge tool in the study of mitochondrial dynamics. This comparison not only highlights the advancements in deep learning technologies but also underscores the continuous evolution and improvement of methodologies for mitochondrial research.

Furthermore, MitoHub's user-friendly web interface and GPU-acceleration capability represent a significant advancement in the field of fast and precise mitochondrial analysis. By integrating the latest deep learning models and motion analysis techniques, it provides a comprehensive, real-time solution for studying mitochondrial dynamics. This approach not only surpasses traditional methods in terms of accuracy and efficiency but also facilitates new research possibilities, enhancing our understanding of mitochondrial function and its implications in health and disease.

Results

In our study, we aimed to analyze mitochondrial dynamics by leveraging a combination of optical flow-based mobility estimation and state-of-the-art transformer-based neural network segmentation models, specifically YOLOv8-X & YOLOv9-E. The optical flow method allowed us to capture and quantify the movement of mitochondria across sequential frames in live-cell imaging, providing crucial insights into mitochondrial mobility. Simultaneously, the YOLO models facilitated accurate segmentation of mitochondrial structures, which is essential for isolating individual mitochondria and analyzing their morphological changes over time.

YOLO Segmentation Neural Network

In this study, we employed two advanced versions of the YOLO (You Only Look Once) segmentation models, YOLOv8-X, YOLOv9-E and YOLOv11-X to accurately segment mitochondrial structures within high-resolution microscopy images. These models, known for their efficiency in real-time object detection and segmentation, were specifically adapted to handle the complex and varied morphology of mitochondria, which are often densely packed and overlapping in cellular environments.

Pre-processing	**Patched**						**Un-patched**					
Model	**YOLOv8-X**		**YOLOv9-E**		**YOLOv11-X**		**YOLOv8-X**		**YOLOv9-E**		**YOLOv11-X**	
Metrics	**Box**	**Mask**	**Box**	**Mask**	**Box**	**Mask**	**Box**	**Mask**	**Box**	**Mask**	**Box**	**Mask**
Precision	0.889	0.897	0.882	0.89	**0.884**	**0.893**	0.618	0.602	0.628	0.6	0.637	0.622
Recall	0.871	0.877	0.862	0.869	**0.872**	**0.88**	0.594	0.562	0.567	0.538	0.573	0.529
mAP@50	0.931	0.932	0.935	0.935	**0.937**	**0.94**	0.619	0.586	0.612	0.574	0.621	0.581
mAP@50-95	0.814	0.683	0.815	0.678	**0.817**	**0.68**4	0.453	0.287	0.438	0.269	0.455	0.281

Table 1. In the given table we present the performance of YOLOv9-E segmentation neural network model in comparison with YOLOv8-X segmentation neural network after training for 100 epochs at an upscaled image size of 640X640. We observe that YOLOv9-E segmentation neural network model (mAP@50 of 0.935 for bounding box and segmentation mask), performs slightly better than YOLOv8-X segmentation neural network model (mAP@50 of 0.931 for bounding box and 0.932 for segmentation mask) over the validation dataset containing 7651 patched images (where the metric is intersection over union (IoU). However, the difference between the two neural network's performance for the segmentation task is very less. When compared with the results without unpatching of the images, we can observe that there is a significant decrease in overall scores across the two neural network architectures (mAP@50 of for YOLOv8-X segmentation neural network model and a mAP@50 of 0.612 for YOLOv9-E segmentation neural network model). The score decreased more in the case of YOLOv9-E than YOLOv8-X. Here, the validation was carried out using 50 full images.

Dataset Preparation and Training Process Our dataset consisted of 218 annotated microscopy images of resolution 3400X2290. The training dataset used for both models consisted of converting these 218 images into patches of 340X299. This resulted in 17,855 image patches, each of size 340x299 pixels, extracted from large-scale microscopy images with a resolution of 3400X2290 pixels. These patches were carefully curated to include a wide range of mitochondrial shapes and sizes, ensuring that the models could learn to identify mitochondria in diverse scenarios. For validation, a separate set of 7,651 patches was used, allowing us to rigorously evaluate the performance of the models[20,21].

Both YOLOv8-X, YOLOv9-E, YOLOv11-X models were initially trained for 100 epochs. During training, the input image patches were upsampled to a resolution of 640x640 pixels to enhance the detection accuracy, particularly for smaller and more intricate mitochondrial structures. This upsampling allowed the models to capture finer details that are critical for precise segmentation. The training process involved stochastic gradient descent (SGD) optimization, coupled with a cosine annealing learning rate scheduler, which helped in efficiently converging the models while avoiding overfitting[22,23].

YOLOv8-X, YOLOv9-E and YOLOv11-X Model Architectures YOLOv8 represents a significant advancement in the YOLO family, introducing a balance between inference speed and detection accuracy that makes it highly suitable for a range of object detection applications. The architecture incorporates an improved backbone network, leveraging a CSPDarknet-based[24] structure that enhances feature extraction across different scales, ensuring more reliable detection of both small and large objects. Furthermore, YOLOv8 integrates advanced decoupled detection heads optimized for computational efficiency, allowing for faster processing without sacrificing accuracy. The model's anchor-free design simplifies the detection pipeline and reduces computational complexity, making it particularly well-suited for the segmentation of mitochondria. This capability is crucial given the complex and variable morphologies of mitochondria, which can include densely packed structures and irregular shapes. YOLOv8-X harnesses the power of convolutional neural networks (CNNs)[25] to learn and identify intricate patterns within the input images, facilitating precise delineation of mitochondrial boundaries, even in challenging scenarios involving clustered mitochondria.

Building on the foundation laid by YOLOv8-X, the YOLOv9-E model incorporates significant advancements that elevate its performance in various object detection tasks, including mitochondrial segmentation. One of the key enhancements in

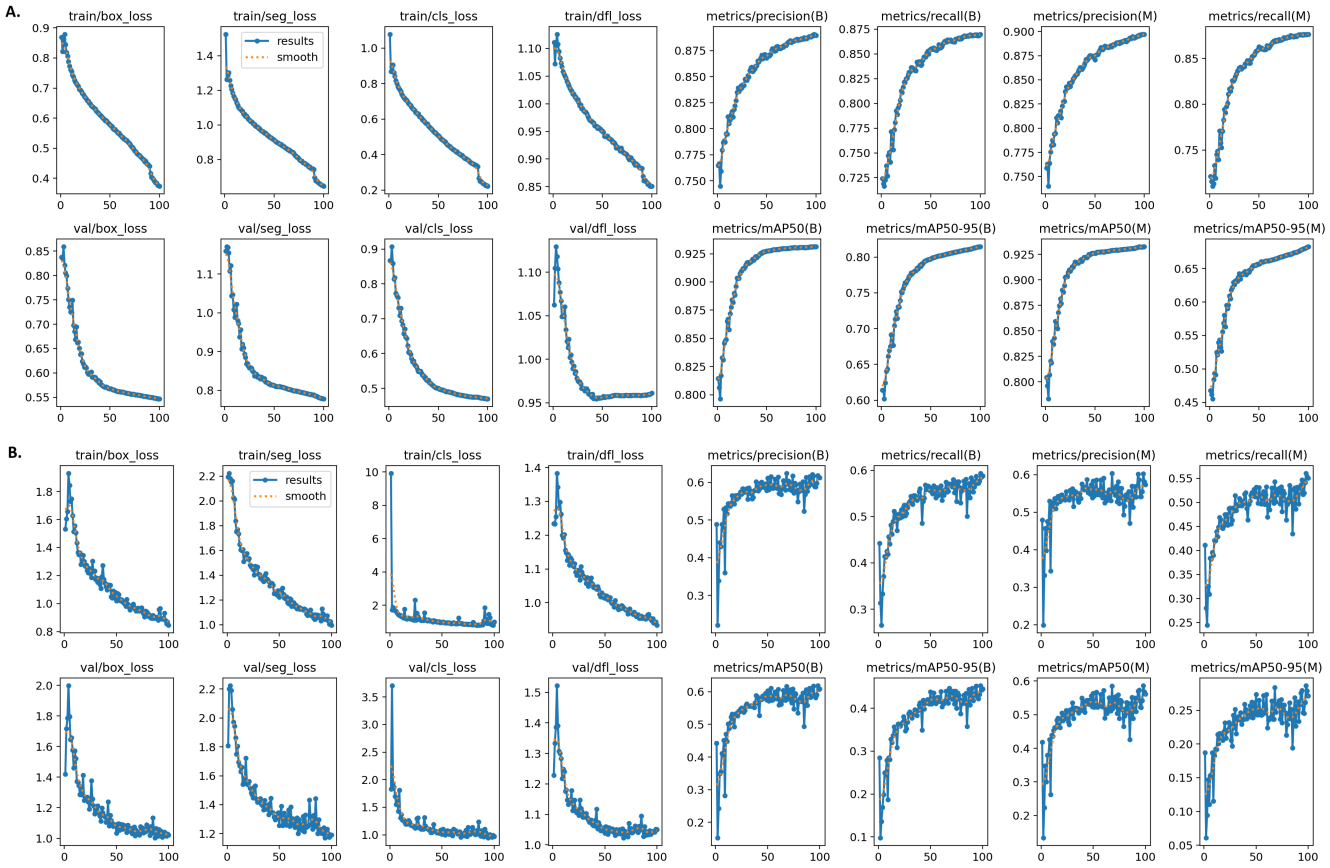

Figure 1. The given figure represents the performance metrics showcasing **A.** When patching of input image was performed and original images of size 3400X2290 were converted into patches of 340X299 resolution and later on upsampled into 640X640 resolution. Here, a batch size of 16 was used for 100 epochs using YOLOv8-X architecture **B.** represents YOLOv8-X segmentation neural network architecture where no patching was carried out and an input image of 1280 size was provided at 1 batch size and for 100 epochs. The reason behind training at a lower image and batch size was due to constraints on the RTX 4090 GPU memory. Upon training for more epochs, we did not observe any significant increase in neural network's performance (see Supplementary Table S1)

YOLOv9-E is the integration of a sophisticated Generalized Efficient Layer Aggregation Network (GELAN), which serves as an evolution of the traditional feature pyramid networks (FPN)[26]. This advanced architecture is designed to address the critical issue of information loss during deep network transmissions by retaining more complete information at multiple semantic levels. This capability is particularly beneficial for tasks like mitochondrial segmentation, where the detection of objects across a wide range of scales is crucial due to the varying sizes of mitochondria within the same image. Additionally, YOLOv9-E introduces a novel programmable gradient information (PGI) mechanism that improves the model's ability to focus on small and overlapping objects. The PGI mechanism provides reliable gradient information through auxiliary reversible branches, allowing YOLOv9-E to maintain key characteristics necessary for accurate segmentation. This design significantly enhances the model's capacity to capture intricate details in mitochondrial structures, resulting in more precise segmentation outcomes.

As a further enhancement to YOLOv9 and YOLOv10 architectures, YOLOv11-X employs a novel Enhanced Hierarchical Feature Fusion (EHFF) mechanism, which goes beyond traditional multi-scale feature integration techniques by incorporating hierarchical feature aggregation at different network depths. This design significantly reduces information loss while improving the model's ability to detect small and occluded objects. Additionally, YOLOv11-X integrates a Dynamic Spatial Attention (DSA) module that adjusts focus across different spatial regions based on the detected object sizes, improving the precision of detections in crowded scenes. These enhancements make YOLOv11-X particularly effective for tasks like detecting fruitlets in dense orchards, where diverse object sizes and occlusions present significant challenges. YOLOv11-X model has also been adapted for segmentation tasks, including mitochondrial segmentation in biological images. By integrating a refined version of the EHFF mechanism and incorporating boundary-aware loss functions, YOLOv11-X can accurately delineate the intricate shapes of mitochondria across varying scales. This makes it well-suited for biomedical use cases where precise segmentation is crucial for understanding mitochondrial morphology and detecting abnormalities in cellular structures. The

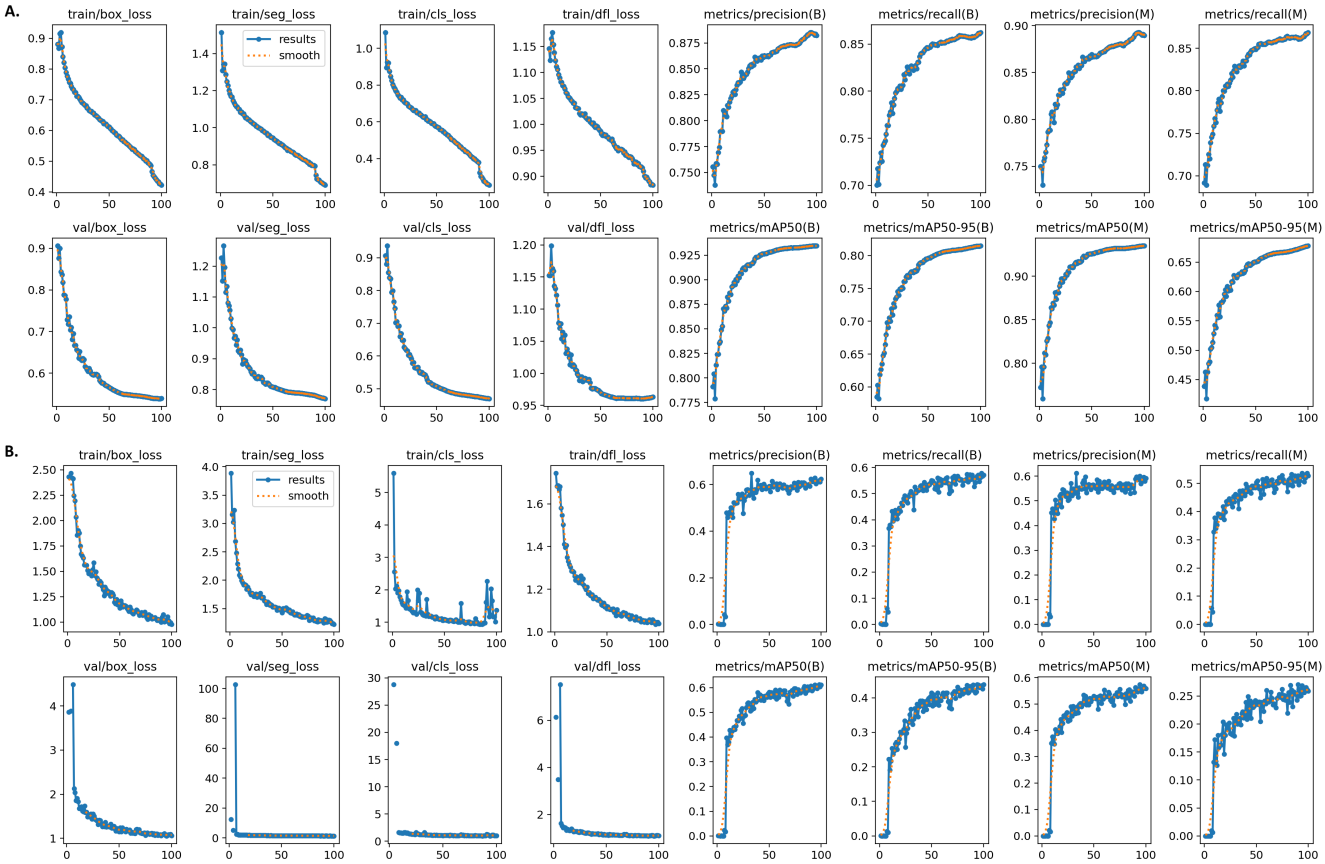

Figure 2. The given figure represents the performance metrics showcasing **A.** When patching of input image was performed and original images of size 3400X2290 were converted into patches of 340X299 resolution and later on upsampled into 640X640 resolution. Here, a batch size of 16 was used for 100 epochs using YOLOv9-E architecture **B.** represents YOLOv9-E segmentation neural network architecture where no patching was carried out and an input image of 1280 size was provided at 1 batch size and for 100 epochs. The reason behind training at a lower image and batch size was due to constraints on the RTX 4090 GPU memory. Upon training for more epochs, we did not observe any significant increase in neural network's performance (see Supplementary Table S1)

model's capability to handle complex backgrounds and overlapping regions enhances its effectiveness in identifying even the smallest mitochondrial features, leading to more accurate segmentation outcomes.

Performance Evaluation and Results The performance of YOLOv8-X, YOLOv9-E and YOLOv11-X was evaluated using the validation dataset of 7,651 image patches. The evaluation metrics included precision, recall, and mean Average Precision (mAP) at different Intersection over Union (IoU) thresholds, specifically mAP@50 and mAP@50-95.

As detailed in (see Table 1), YOLOv9-E achieved a slight edge over YOLOv8-X in terms of mAP@50, with both models performing exceptionally well. YOLOv9-E recorded a mAP@50 of 0.935 for both bounding box and segmentation mask tasks, compared to 0.931 and 0.932 for YOLOv8-X, respectively. The recall and precision scores were also comparable, with YOLOv9-E slightly outperforming YOLOv8-X. However, the difference in performance between the two models, while present, was minimal, particularly in the segmentation tasks where the mAP@50-95 scores were closely matched. However, YOLOv11-X was able to outperform both YOLOv9-E and YOLOv8-X models achieving a mAP@50 of 0.937 for bounding box and 0.94 for segmentation mask.

To provide a comprehensive overview of the training progress, we combined the training logs of both YOLOv8-X (see Figure 1), YOLOv9-E (see Figure 2) and YOLOv11-X (see Figure 3) into graphs. The graphs track several metrics across each epoch and highlight the convergence behavior of both models. YOLOv11-X and YOLOv9-E demonstrated a slightly faster convergence and reached a higher plateau of accuracy more quickly compared to YOLOv8-X, indicating its improved capability in learning complex mitochondrial patterns. YOLOv11-X and YOLOv9-E also exhibited more stable training dynamics, with fewer fluctuations in loss and accuracy metrics, which suggests better generalization to the validation data.

Impact of Patching on Model Performance The use of image patching during training had a noticeable impact on the models' performance. By dividing the large microscopy images into smaller patches, we effectively increased the resolution

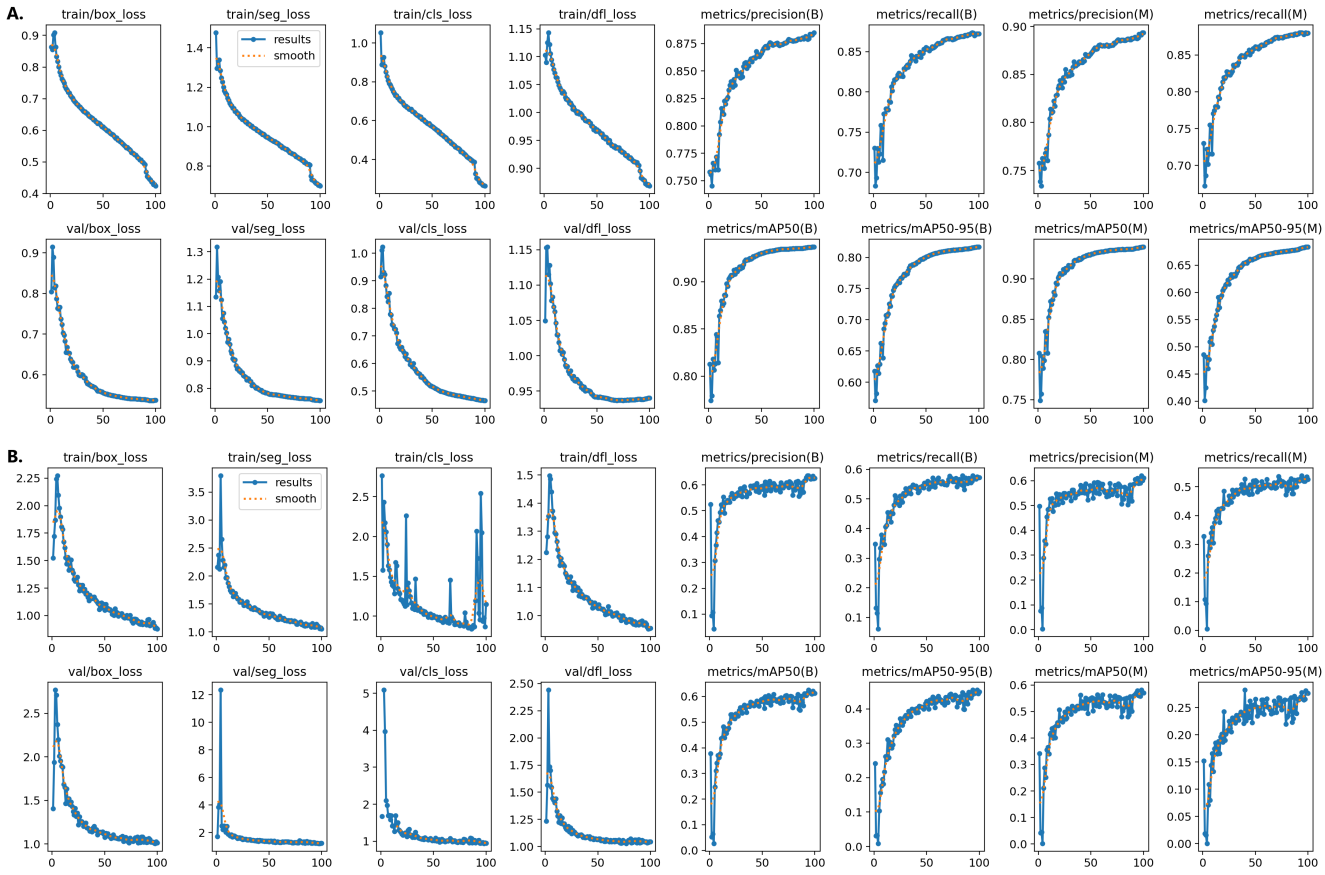

Figure 3. The given figure represents the performance metrics showcasing **A.** When patching of input image was performed and original images of size 3400X2290 were converted into patches of 340X299 resolution and later on upsampled into 640X640 resolution. Here, a batch size of 16 was used for 100 epochs using YOLOv11-X architecture **B.** represents YOLOv11-X segmentation neural network architecture where no patching was carried out and an input image of 1280 size was provided at 1 batch size and for 100 epochs. The reason behind training at a lower image and batch size was due to constraints on the RTX 4090 GPU memory. Upon training for more epochs, we did not observe any significant increase in neural network's performance (see Supplementary Table S1)

of the input data, allowing the models to detect and segment smaller and more intricate mitochondrial features with greater accuracy. This patching technique was particularly beneficial for YOLOv9-E, which, due to its enhanced architecture, was able to leverage the detailed information in the patches more effectively, leading to higher precision and recall in the segmentation tasks.

Both YOLOv8-X, YOLOv9-E and YOLOv11-X models demonstrated high efficacy in mitochondrial segmentation, with YOLOv11-X offering marginally superior performance due to its advanced features, such as the improved mechanism. The slight but consistent improvements observed with YOLOv11-X suggest that it is slightly better suited for applications in mitochondrial segmentation, especially when combined with image patching techniques. The results obtained in this study underscore the potential of YOLOv9-E as a powerful tool for real-time analysis of mitochondrial dynamics in high-resolution microscopy images.

Discussion

MitoHub represents a significant advancement in the automated analysis of mitochondrial dynamics. The integration of YOLO models provides efficient and accurate event detection and segmentation, while the use of sparse optical flow methods offers a quantitative assessment of mitochondrial mobility. These tools collectively enable comprehensive and high-throughput analysis of mitochondrial function, facilitating research into mitochondrial-related diseases.

In the field of mitochondrial analysis, current methods such as Mitometer have made significant strides in enhancing our understanding of mitochondrial dynamics. Mitometer, as detailed in recent studies, offers a robust platform for analyzing mitochondrial morphology and dynamics using advanced image processing techniques. It excels in providing high-resolution,

quantitative measures of mitochondrial shape, size, and distribution, which are crucial for studying various biological processes and diseases.

However, our approach, which integrates advanced YOLO segmentation with sparse optical flow techniques, offers several advantages over Mitometer. While Mitometer focuses on morphology and dynamics through a series of specific image processing steps, our system combines precise mitochondrial segmentation with detailed mobility tracking in a unified framework. This integration allows for a more comprehensive analysis, providing not only structural information but also dynamic insights into mitochondrial movement over time.

The YOLOv8 model employed in our system offers superior segmentation accuracy, particularly in handling complex, high-resolution microscopy images. Its advanced architecture enables real-time processing and precise delineation of mitochondrial boundaries, even in densely populated cellular environments. This contrasts with Mitometer's more traditional approach, which may not handle high-density or overlapping mitochondrial structures as effectively.

Additionally, our method's use of sparse optical flow, specifically through the Shi-Tomasi corner detection and Lucas-Kanade tracking, enhances the precision of movement analysis. This combination allows for fine-grained tracking of mitochondrial mobility, offering a more detailed view of their dynamic behavior. Mitometer's dynamic analysis capabilities, while effective, do not provide the same level of detailed motion tracking and quantification as our integrated approach.

The integration of mitochondrial segmentation and mobility estimation into a cohesive system, supported by a user-friendly web-based interface, represents a significant advancement in the analysis of mitochondrial dynamics. By combining precise segmentation with detailed tracking of mitochondrial movement, our system offers a powerful tool for exploring the complex behavior of these critical organelles in cellular environments. This integrated approach enhances our ability to gain insights into mitochondrial function and pathology, while the web-based interface ensures accessibility and ease of use for researchers.

The integration of segmentation and mobility estimation enables a more nuanced understanding of mitochondrial dynamics. Segmentation provides detailed information about the structure and boundaries of individual mitochondria, allowing for the precise identification of their size, shape, and spatial distribution. Once mitochondria are accurately segmented, mobility estimation can be applied to analyze their movement patterns. This combined approach allows researchers to study the dynamics of specific mitochondrial regions and gain insights into their functional roles within the cell.

For example, by examining the movement of mitochondria within distinct cellular compartments or in proximity to other organelles, researchers can infer how mitochondrial dynamics influence or are influenced by their surroundings. This can reveal patterns of mitochondrial behavior that are associated with cellular processes such as energy production, stress response, or apoptotic signaling. Additionally, the ability to track the mobility of mitochondria in relation to their segmented boundaries provides a deeper understanding of how mitochondrial morphology and movement are interrelated.

Using mobility estimation on segmented mitochondria also facilitates the investigation of mitochondrial health and dysfunction. In healthy cells, mitochondria exhibit specific movement patterns that are crucial for their proper function. Deviations from these patterns can indicate pathological conditions, such as neurodegenerative diseases or cancer. By analyzing how mitochondria move and interact with their environment, researchers can identify subtle changes that might be indicative of disease progression or therapeutic response.

For instance, in neurodegenerative diseases like Alzheimer's or Parkinson's, mitochondrial mobility may be disrupted, affecting neuronal health. Accurate segmentation combined with mobility analysis allows researchers to pinpoint specific mitochondrial abnormalities and assess their impact on cellular function. Similarly, in cancer research, examining how mitochondria alter their movement and morphology in response to metabolic demands provides insights into the mechanisms of tumor progression and potential therapeutic targets.

The web-based user interface (UI) is a crucial component of our system, designed to make the advanced capabilities of our integration accessible and manageable for researchers. The UI offers an intuitive platform where users can easily upload microscopy images, perform segmentation, and apply mobility estimation algorithms. Features include interactive visualization tools that allow users to view segmented mitochondria, track their movement, and analyze dynamic behavior in real-time.

The interface supports customization, enabling users to adjust parameters for segmentation and mobility estimation according to their specific research needs. This flexibility ensures that the system can be tailored to different experimental conditions and imaging modalities. Additionally, the UI provides comprehensive data export options, facilitating further analysis and integration with other research tools.

Future work could explore the integration of additional deep learning models and techniques to further enhance the accuracy and robustness of the pipeline. Additionally, expanding the dataset to include a wider variety of cell types and conditions could improve the generalizability of the models.

Methods

Our goal while developing their project was to create a system that excels in both quantification of the dynamic movement of mitochondria and segmenting them accurately within the complex environment of microscopy images. By leveraging the

strengths of optical flow techniques in tracking motion and the robustness of the YOLOv8-E and YOLOv9-E models in image segmentation, we aim to provide a comprehensive solution that meets the high demands of biological imaging and research.

Mobility Estimation using Sparse Optical Flow

Mitochondrial mobility is crucial for a variety of applications, including understanding cellular dynamics, assessing mitochondrial health, and investigating neurodegenerative diseases. Using this technique, we can accurately track the movement of individual mitochondria over time, even within complex and cluttered microscopy images. Sparse optical flow algorithms enable the identification of key feature points on mitochondria, allowing us to monitor their trajectories and analyze their motion patterns with high precision. This detailed tracking provides valuable insights into mitochondrial behavior, contributing to a deeper understanding of cellular processes and aiding in the development of targeted therapeutic interventions.

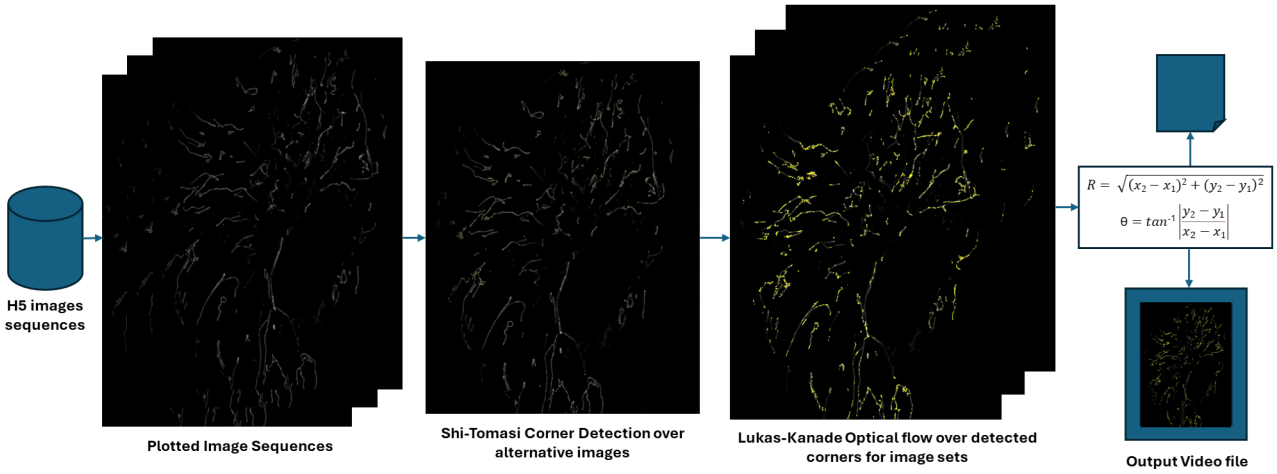

Figure 4. The given workflow shows the step-by-step workflow for converting the live-cell imaging h5 files into image sequences over which the Shi-Tomasi corner detection is implemented to identify significant features which are then tracked using Lukas-Kanade Optical flow estimation method. Furthermore, the magnitude and direction of motion is calculated from the initial (x_1, y_1) and final position (x_2, y_2) of the detected corner point. This can be used to effectively identify the magnitude of motion as well as direction in which the motion is taking place. In order to increase the sensitivity, number of corner points to detect can be increased. In the current example, we used 500 corner points which were detected and tracked for every alternate image sequence pairs.

In our implementation, we utilize the Shi-Tomasi corner detection method[27], a well-established technique for identifying significant feature points in an image. This method is particularly effective because it detects corners, which are regions in the image where the intensity changes significantly in multiple directions. These corner points are crucial for tracking because they represent distinct and stable features that can be reliably followed across successive frames.

To optimize our analysis, we have the flexibility to adjust the number of corner points to be tracked. This adjustment allows us to fine-tune the system to prioritize either sensitivity or specificity based on the specific needs of our study. For instance, increasing the number of corner points can enhance sensitivity, enabling the detection of more subtle movements, while reducing the number of points can improve specificity, focusing on the most prominent movements.

To improve the signal-to-noise ratio, we detect these corner points in every alternate frame rather than in every single frame. This approach helps in reducing noise and ensures that the detected points are more stable and reliable for tracking over time.

Once the corner points are identified, we employ the Lucas-Kanade optical flow method to track their movement across the sequence of images. The Lucas-Kanade method is particularly suited for this task because it assumes that the flow is essentially constant in a local neighborhood of the pixel under consideration, making it highly effective for tracking small, localized movements. By analyzing the displacement of these corner points between frames, we can accurately quantify the amount of motion and direction of movement of the mitochondria.

This robust combination of the Shi-Tomasi corner detection and Lucas-Kanade optical flow methods ensures precise and reliable tracking of mitochondrial movement. It allows us to capture detailed motion patterns and dynamics, providing valuable insights into mitochondrial behavior in various biological contexts. This capability is essential for advancing our understanding of cellular processes and for the development of targeted therapeutic interventions in diseases where mitochondrial dysfunction plays a critical role.

Mitochondrial Segmentation using YOLO models

Mitochondrial segmentation also plays a crucial role in identifying and analyzing the intricate structure and diverse functions of mitochondria within cells. This process is fundamental to several key areas of biological and medical research, as it allows scientists to visualize and quantify the detailed morphology of mitochondria. Understanding the precise shape, size, and distribution of mitochondria within the cellular environment is essential for studying how these organelles contribute to various cellular processes, including energy production, apoptosis, and calcium signaling.

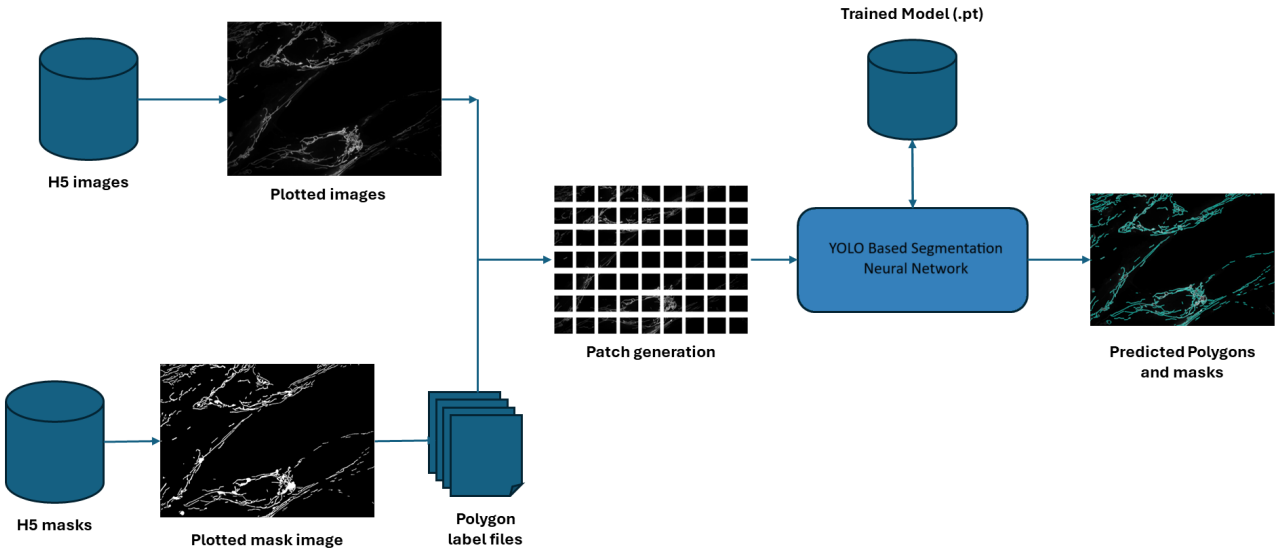

Figure 5. The given figure shows the workflow that was followed in order to convert de-convoluted imaging data into image files and segmentation mask data into first the segmentation image masks and later on polygon label files which were then converted into smaller patches. For our test case, we used an input image of 3400X2290 resolution and divided them into patches of 340X229 which no overlap for training. We then performed training using YOLOv8x and YOLOv9e-segmentation neural network models. Furthermore, for inference, we adjusted the images to have 1% overlap and carried out inference to obtain the segmented images.

Accurate mitochondrial segmentation enables researchers to dissect the complex interactions between mitochondria and other cellular components. For instance, by segmenting mitochondria, scientists can investigate their dynamic behavior and spatial relationships with the cytoskeleton, endoplasmic reticulum, and other organelles. This level of detail is vital for elucidating the mechanisms of mitochondrial fusion and fission, which are processes that maintain mitochondrial function and integrity.

Furthermore, mitochondrial segmentation is critical for assessing mitochondrial health and function in various physiological and pathological conditions. In healthy cells, mitochondria exhibit specific structural characteristics and distribution patterns, which can be altered in response to stress, disease, or genetic mutations. By segmenting and analyzing mitochondria, researchers can identify these alterations and gain insights into the underlying causes of mitochondrial dysfunction.

In the context of neurodegenerative diseases, such as Alzheimer's and Parkinson's, mitochondrial segmentation allows for the examination of mitochondrial abnormalities that are often associated with these conditions. Researchers can track changes in mitochondrial morphology and distribution, providing valuable information on how these organelles contribute to neuronal degeneration and death. Similarly, in cancer research, mitochondrial segmentation helps in understanding how cancer cells rewire their metabolism to support rapid growth and proliferation, often involving significant alterations in mitochondrial structure and function.

For this purpose, we utilized the YOLOv8 and YOLOv9 models, the current state-of-the-art deep learning frameworks renowned for their exceptional accuracy and speed in object detection and segmentation tasks. The YOLO (You Only Look Once) model is designed to efficiently process high-resolution images, making it particularly suitable for detailed and complex microscopy images.

The YOLOv8 model employs a single-stage detection architecture that enables it to perform both localization and classification of objects in real-time. This architecture is optimized to handle the intricacies of mitochondrial shapes and sizes, ensuring precise segmentation even in densely populated cellular environments. The model uses convolutional neural networks (CNNs) to learn and recognize patterns within the images, allowing it to accurately delineate the boundaries of each

mitochondrion. Moreover, the YOLOv8 model incorporates advanced techniques such as spatial pyramid pooling and path aggregation network, which enhance its ability to detect and segment objects at multiple scales. This multi-scale capability is particularly important for mitochondrial segmentation, as mitochondria can vary significantly in size and shape.

Building on the capabilities of YOLOv8, the YOLOv9 model introduces further enhancements in accuracy and speed. YOLOv9 incorporates an advanced feature pyramid network (FPN) for better multi-scale feature representation, which is crucial for detecting mitochondria of varying sizes. Additionally, YOLOv9 utilizes a more sophisticated attention mechanism to improve the detection of small and overlapping objects, thus offering superior performance in complex cellular imaging scenarios. This makes YOLOv9 particularly effective for mitochondrial segmentation in high-resolution microscopy images, providing even more precise and reliable analysis of mitochondrial dynamics.

Our approach to training and prediction involves a strategy that addresses the challenges associated with high-resolution microscopy images. We divide the large microscopy images into smaller, more manageable chunks or patches. This fine-grained approach allows us to perform detailed training and prediction processes, enhancing the model's ability to detect and segment mitochondria at a higher level of precision. Each patch is individually analyzed to capture intricate details of the mitochondrial structure, which contributes to a more accurate and comprehensive segmentation.

Moreover, to address the issue of segmenting mitochondria that are close to each other, we employ overlap between these image patches. This overlapping strategy ensures that the boundaries of adjacent patches are not lost, allowing for a seamless and continuous segmentation of mitochondria that may span across patch boundaries. By overlapping the patches, we can effectively bridge the gaps between adjacent segments and achieve a more coherent and accurate segmentation of closely situated mitochondria.

The combination of YOLO's advanced segmentation capabilities with our method of fine-grained training and overlapping image patches significantly enhances our ability to analyze mitochondrial structures. This approach provides a high-resolution, detailed segmentation that facilitates in-depth studies of mitochondrial morphology, dynamics, and their roles in various biological processes and diseases. It also improves the efficiency of data analysis, allowing researchers to derive meaningful insights and develop targeted therapeutic strategies based on a comprehensive understanding of mitochondrial function and pathology.

Integration of Segmentation and Mobility with user-friendly Web-UI

Our system's user-friendly web-based interface represents a significant advancement in making complex mitochondrial analysis accessible and efficient for researchers. Designed with an emphasis on ease of use, the interface allows biologists to seamlessly perform advanced mitochondrial segmentation, and analyze dynamic behavior without requiring extensive computational expertise.

The web-based platform is intuitively designed to support a broad range of functions. Researchers can easily upload their microscopy images, select parameters for segmentation and mobility estimation, and visualize results in real-time. This interactive environment streamlines the workflow, enabling users to quickly adjust settings, view segmented mitochondria, and track their movement with minimal effort. The interface provides interactive visualization tools that highlight key features, such as mitochondrial boundaries and movement trajectories, facilitating a more straightforward analysis process.

In addition to its user-friendly design, the web-based interface harnesses the power of GPU acceleration to significantly speed up inference and processing times. By leveraging GPU resources, the system can handle large-scale, high-resolution images efficiently, delivering rapid results without compromising accuracy. This capability is particularly beneficial for processing complex and voluminous datasets commonly encountered in microscopy studies, allowing for faster analysis and quicker turnaround times in research.

For biologists, the integration of a web-based interface with GPU acceleration offers several practical advantages. Firstly, it reduces the computational burden typically associated with advanced image analysis, enabling biologists to focus more on interpreting results rather than managing technical aspects of the analysis. The intuitive design of the interface also lowers the barrier to entry, making sophisticated analysis tools accessible to researchers without extensive training in computational techniques.

Moreover, the real-time feedback provided by the web-based platform allows for iterative analysis and immediate adjustments, which is crucial for experimental optimization. Researchers can visualize the impact of parameter changes on segmentation and mobility tracking in real time, facilitating a more responsive and iterative approach to data analysis.

References

1. Chen, H. & Chan, D. C. Physiological functions of mitochondrial fusion. *Annals New York Acad. Sci.* **1201**, 21–25 (2009).
2. Nunnari, J. & Suomalainen, A. Mitochondria: in sickness and in health. *Cell* **148**, 1145–1159 (2012).
3. Frederick, R. L. & Shaw, J. M. Mitochondria as a dynamic signaling platform. *Curr. opinion cell biology* **19**, 345–352 (2007).

4. Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A. & Hudspeth, A. J. *Principles of neural science* (McGraw-Hill New York, 2012).

5. Youle, R. J. & van der Bliek, A. M. Mitochondrial fission, fusion, and stress. *Science* **337**, 1062–1065 (2012).

6. Gomes, L. C., Di Benedetto, G. & Scorrano, L. A nexus for cellular homeostasis–the interplay between metabolic and stress signaling pathways. *The EMBO journal* **30**, 465–479 (2011).

7. Chan, D. C. Mitochondria and neurodegeneration: new insights and challenges. *Nat. Rev. Neurosci.* **21**, 358–368 (2020).

8. Vyas, S., Zaganjor, E. & Haigis, M. Mitochondria and cancer. *Cell* **166**, 555–566 (2016).

9. Jakobs, S. & Wurm, C. A. Super-resolution microscopy of mitochondria. *Curr. opinion chemical biology* **20**, 9–15 (2014).

10. Lefebvre, A. E. Y. T., Ma, D., Kessenbrock, K., Lawson, D. A. & Digman, M. A. Author correction: Automated segmentation and tracking of mitochondria in live-cell time-lapse images. *Nat. Methods* **19**, 770–770, DOI: 10.1038/s41592-022-01506-2 (2022).

11. Viana, M. P., Lim, S. & Rafelski, S. M. Chapter 6 - quantifying mitochondrial content in living cells. In Paluch, E. K. (ed.) *Biophysical Methods in Cell Biology*, vol. 125 of *Methods in Cell Biology*, 77–93, DOI: https://doi.org/10.1016/bs.mcb.2014.10.003 (Academic Press, 2015).

12. Valente, A. J., Maddalena, L. A., Robb, E. L., Moradi, F. & Stuart, J. A. A simple ImageJ macro tool for analyzing mitochondrial network morphology in mammalian cell culture. *Acta Histochem.* **119**, 315–326 (2017).

13. Ershov, D. *et al.* Trackmate 7: integrating state-of-the-art segmentation algorithms into tracking pipelines. *Nat. Methods* **19**, 829–832, DOI: 10.1038/s41592-022-01507-1 (2022).

14. Thielicke, W. & Stamhuis, E. J. PIVlab – towards user-friendly, affordable and accurate digital particle image velocimetry in MATLAB. *J. Open Res. Softw.* **2** (2014).

15. Fischer, C. A. *et al.* Mitosegnet: Easy-to-use deep learning segmentation for analyzing mitochondrial morphology. *iScience* **23**, 101601, DOI: https://doi.org/10.1016/j.isci.2020.101601 (2020).

16. Redmon, J., Divvala, S., Girshick, R. & Farhadi, A. You only look once: Unified, real-time object detection (2016). 1506.02640.

17. Vaswani, A. *et al.* Attention is all you need (2023). 1706.03762.

18. Jocher, G., Chaurasia, A. & Qiu, J. Ultralytics YOLO (2023).

19. Wang, C.-Y., Yeh, I.-H. & Liao, H.-Y. M. Yolov9: Learning what you want to learn using programmable gradient information (2024). 2402.13616.

20. Akyon, F. C., Onur Altinuc, S. & Temizel, A. Slicing aided hyper inference and fine-tuning for small object detection. In *2022 IEEE International Conference on Image Processing (ICIP)*, DOI: 10.1109/icip46576.2022.9897990 (IEEE, 2022).

21. Zhang, E. *et al.* Adaptive patching for high-resolution image segmentation with transformers (2024). 2404.09707.

22. Lu, J. Gradient descent, stochastic optimization, and other tales (2024). 2205.00832.

23. Loshchilov, I. & Hutter, F. Sgdr: Stochastic gradient descent with warm restarts (2017). 1608.03983.

24. Wang, C.-Y. *et al.* Cspnet: A new backbone that can enhance learning capability of cnn (2019). 1911.11929.

25. O'Shea, K. & Nash, R. An introduction to convolutional neural networks (2015). 1511.08458.

26. Lin, T.-Y. *et al.* Feature pyramid networks for object detection (2017). 1612.03144.

27. Mu, Z. & Li, Z. A novel shi-tomasi corner detection algorithm based on progressive probabilistic hough transform. In *2018 Chinese Automation Congress (CAC)*, 2918–2922, DOI: 10.1109/CAC.2018.8623648 (2018).

28. Bangare, S., Dubal, A. A., Bangare, P. & Patil, S. Reviewing otsu's method for image thresholding. *Int. J. Appl. Eng. Res.* **10**, 21777–21783, DOI: 10.37622/IJAER/10.9.2015.21777-21783 (2015).

Supporting Information

Supplementary Information

Understanding mitochondrial mobility across various cellular conditions provides crucial insights into mitochondrial dynamics and their role in cellular function. To assess mitochondrial mobility, we utilized an optical flow-based method, which involved tracking the movement of 200 Shi-Tomasi corner points across time-lapse sequences of high-resolution live-cell images. This approach allowed us to quantify mitochondrial movement with high accuracy under different experimental conditions, specifically focusing on cycling cells, cells under glucose starvation, and cells treated with cyclin-dependent kinase (CDK) inhibitor Palbociclib (Palbo) and the dual treatment of Palbociclib with mTOR inhibitor Torin.

Methodology for Optical Flow-Based Mobility Estimation

The Shi-Tomasi corner detection method was employed to identify robust and distinctive feature points within the mitochondria, which are ideal candidates for tracking over time. These corner points, which represent areas of high intensity variation, are crucial for accurately capturing the dynamics of mitochondrial movement. The Lucas-Kanade optical flow algorithm was then applied to track these points across sequential frames, enabling us to measure the displacement of each point and thus quantify the mobility of mitochondria within the cellular environment.

This optical flow approach is particularly well-suited for analyzing mitochondrial dynamics as it captures not only the directionality of movement but also the speed and consistency of mitochondrial motion. By applying this technique to different cellular conditions, we were able to discern significant variations in mitochondrial mobility that correspond to the cells' metabolic and signaling states.

Cycling vs. Glucose Starvation:

We first compared mitochondrial mobility between cycling cells and cells subjected to glucose starvation. Cycling cells, which are actively proliferating, exhibited significantly higher mitochondrial mobility compared to glucose-starved cells. In cycling cells, mitochondria were observed to move rapidly across the cytoplasm, reflecting the high energy demands and metabolic activity associated with cell division. This enhanced mobility is consistent with the role of mitochondria in distributing ATP to regions of the cell where it is most needed during proliferation.

In contrast, glucose-starved cells, which are in a metabolically deprived state, showed markedly reduced mitochondrial mobility. The mitochondria in these cells were more stationary, likely due to the diminished energy availability and the shift towards energy conservation mechanisms under nutrient scarcity. Even with the high contrast provided by live-cell imaging or when analyzed through binary mask processing, the mobility difference between cycling and glucose-starved cells was pronounced, underscoring the impact of metabolic state on mitochondrial dynamics.

To enhance the contrast and differentiate more clearly between the two conditions, we applied gamma correction to the images. This adjustment amplified the visibility of mitochondria, particularly in glucose-starved cells, allowing for better detection and tracking of mitochondrial movement. The application of gamma correction made the differences in mobility between cycling and glucose-starved cells even more pronounced (see Supplementary Figure S1).

Furthermore, we converted the images into binary format using Otsu thresholding, which provided a clear demarcation of mitochondrial boundaries. This binary conversion further validated the observed pattern: cycling cells consistently exhibited higher mitochondrial mobility than glucose-starved cells, regardless of the image processing technique employed. The consistency of these results across different image processing methods highlights the robustness of the mobility estimation technique across these two metabolic states under variability in microscopy images. We further find the mean mobility and standard deviation across all the samples belonging to a specific condition and obtain the overall mobility for comparative analysis (see Supplementary Table S2).

Macrophage Mobility Analysis in M1 vs. M2 Phases: We further extended our mobility analysis to macrophages, comparing mitochondrial dynamics across the M1 and M2 phases of macrophage activation. The analysis revealed that mitochondria in M1 macrophages exhibited significantly higher mobility compared to those in M2 macrophages. This increased mobility in M1 macrophages is likely associated with their heightened inflammatory and metabolic activity, which requires efficient mitochondrial distribution to meet energy demands. In contrast, M2 macrophages, which are involved in tissue repair and anti-inflammatory responses, showed reduced mitochondrial mobility, reflecting their lower metabolic activity and possibly more localized energy requirements (see Supplementary Figure S3).

Cycling vs. Palbociclib vs. Palbociclib-Torin Treatment:

Further comparisons were made between cycling cells, cells treated with Palbociclib (a CDK4/6 inhibitor), and cells treated with the combination of Palbociclib and Torin (an mTOR inhibitor). The analysis revealed a distinct hierarchy in mitochondrial mobility across these conditions.

Palbociclib Treatment: Cells treated with Palbociclib displayed the highest mitochondrial mobility among the three conditions. Palbociclib is known to induce cell cycle arrest in the G1 phase by inhibiting CDK4/6, which disrupts the normal progression of the cell cycle. Despite this arrest, mitochondria in Palbociclib-treated cells continued to exhibit high mobility, possibly due to the altered metabolic requirements and compensatory mechanisms activated during G1 arrest. This increased mobility might also reflect a cellular attempt to redistribute mitochondrial resources in response to stress induced by the drug.

Cycling Cells: As expected, cycling cells, which are actively progressing through the cell cycle, demonstrated substantial mitochondrial mobility, though it was slightly lower than that observed in Palbociclib-treated cells. This high mobility aligns with the energy demands of dividing cells, where mitochondria need to efficiently move to areas requiring ATP for processes such as mitosis and cytokinesis.

Palbociclib-Torin Treatment: The combination of Palbociclib with Torin, an mTOR inhibitor, resulted in the lowest mitochondrial mobility. mTOR is a key regulator of cellular growth and metabolism, and its inhibition by Torin likely led to a significant reduction in overall cellular activity and energy demand. The decreased mobility in this condition suggests that the dual inhibition of CDK4/6 and mTOR not only arrests the cell cycle but also dampens mitochondrial dynamics, possibly as a result of a broader suppression of metabolic processes. This reduction in mitochondrial movement could reflect a state of energy conservation or a shift in mitochondrial function towards maintenance rather than distribution.

Binning and Logarithmic Transformation of Mobility Data

To further analyze the differences in mitochondrial mobility across these conditions, we performed a binning process where the mobility data of the tracked corner points were divided into 100 bins. This binning allowed us to categorize the mobility of mitochondria into discrete ranges, making it easier to quantify and compare the number of high-mobility corner points across the three conditions. By plotting the binned data, we generated a frequency distribution graph of points versus the magnitude of mobility, which provided a clear visual representation of mitochondrial dynamics.

Given the inherent variability in mobility measurements—where some points exhibit significantly higher mobility than others—we applied a logarithmic transformation to the mobility magnitudes. This transformation normalized the data, reducing the impact of extreme values and allowing for a more accurate comparison of mitochondrial mobility across different treatments (see Supplementary Figure S4).

The resulting frequency distribution graph revealed distinct patterns in mitochondrial mobility. Cycling cells exhibited a broad range of mobility, with a significant number of high-mobility points. Upon treatment with Palbociclib, there was an observable increase in the number of high-mobility points, indicating that the drug induced a state of heightened mitochondrial dynamics, possibly as part of a stress response or a compensatory mechanism during cell cycle arrest. Conversely, the combination of Palbociclib and Torin treatment led to a marked decrease in mitochondrial mobility, with the distribution shifting towards lower mobility magnitudes. This reduction aligns with the dual inhibitory effects on both CDK4/6 and mTOR pathways, which together dampen the overall metabolic activity and reduce the need for mitochondrial redistribution.

To further assess the stability of mitochondrial mobility under Palbociclib and Palbociclib-Torin treatments over time, we estimated mobility using the same binning method after 2 days, 4 days, and 6 days of treatment. The results indicated that there was not much change in mobility across these time points. The frequency distribution graphs remained consistent, showing that the initial reduction in mobility caused by Palbociclib-Torin treatment persisted throughout the treatment duration. This suggests that the impact of these treatments on mitochondrial dynamics is established early and remains stable, reflecting a sustained suppression of mitochondrial mobility. Similarly, for Palbociclib alone, the heightened mobility observed initially did not show significant variation over time, indicating that the drug's effects on mitochondrial dynamics are also maintained over prolonged periods (see Supplementary Figure S5).

Implications of Mitochondrial Mobility Findings

These findings provide critical insights into how different cellular states and treatments impact mitochondrial dynamics. The high mobility observed in Palbociclib-treated cells suggests a unique mitochondrial response to CDK4/6 inhibition, possibly linked to altered metabolic pathways or stress responses. In contrast, the marked reduction in mobility under glucose starvation and dual Palbociclib-Torin treatment highlights the sensitivity of mitochondrial dynamics to both nutrient availability and signaling pathway inhibition.

By testing of variable contrast, binary mask, binning and logarithmic transformation, we were able to distinguish the subtle yet significant differences in mitochondrial mobility across different conditions. This approach not only provided a clearer understanding of the impact of different treatments on mitochondrial dynamics but also demonstrated the utility of advanced data processing techniques in enhancing the analysis of complex biological phenomena.

These variations in mitochondrial mobility across different conditions underscore the complexity of mitochondrial behavior and its tight regulation by cellular metabolic and signaling states. The ability to accurately quantify and compare mitochondrial mobility using optical flow methods offers a powerful tool for understanding the role of mitochondria in health and disease, particularly in contexts such as cancer therapy, where mitochondrial dynamics may influence treatment outcomes.

By tracking and analyzing 200 Shi-Tomasi corner points across these diverse conditions, and employing binning and logarithmic transformations, we have established a robust framework for assessing mitochondrial mobility, revealing key differences that correlate with cellular energy states and responses to therapeutic interventions. This approach not only provides valuable data on mitochondrial dynamics but also opens avenues for exploring how mitochondrial behavior can be modulated to optimize therapeutic strategies in various disease contexts.

Further tests for segmentation performance benchmarking using YOLO

In addition to the initial training of the YOLOv8-X and YOLOv9-E models, we conducted further experiments to evaluate whether extending the training duration or modifying the input image preprocessing would lead to any significant improvement in the performance of the segmentation neural networks. Specifically, we increased the number of training epochs beyond the initial 100 epochs, applied image patching, and also tested the models using full-resolution images rescaled to 1280x1280 pixels.

Despite these efforts, the extended training did not result in a significant increase in the performance of either YOLOv8-X or YOLOv9-E. Both models appeared to reach a performance plateau after the initial training, with minimal improvements in key metrics such as precision, recall, and mean Average Precision (mAP). This indicates that the models had already effectively learned the relevant features for mitochondrial segmentation during the initial training phase.

When we applied image patching, the segmentation performance remained consistent with the results obtained during the initial training. Patching the images into smaller segments was intended to enhance the models' ability to detect finer details and smaller objects within the high-resolution microscopy images. However, the results showed no significant improvement in the models' precision, recall, or mAP scores, suggesting that the models were already well-tuned to handle the intricacies of the mitochondrial structures in the original patches.

Similarly, training the models using full images rescaled to 1280x1280 resolution did not yield any performance gains. The increase in input resolution was expected to provide the models with more detailed information, potentially improving segmentation accuracy, especially for small or overlapping mitochondria. However, the mAP and other performance metrics remained largely unchanged, indicating that the models were already extracting sufficient information from the original image patches and did not benefit from the additional resolution.

These findings suggest that both YOLOv8-X and YOLOv9-E models were already optimized for the task of mitochondrial segmentation after the initial training period and that further modifications to the training regimen or input preprocessing did not provide additional benefits (see Supplementary Table S1). This stability in performance across different training conditions underscores the robustness of the models and their ability to generalize well to the validation dataset, even when the training parameters are altered.

Pre-processing	**Un-patched**		**Patched**					
Model Architecture	**YOLOv8-X**		**YOLOv8-X**		**YOLOv9-E**		**YOLOv11-X**	
Epochs	300		678		155		803	
Metrics	**Box**	**Mask**	**Box**	**Mask**	**Box**	**Mask**	**Box**	**Mask**
Precision	0.628	0.606	0.916	0.915	0.854	0.861	0.918	0.927
Recall	0.585	0.553	0.857	0.872	0.825	0.826	0.867	0.875
mAP@50	0.617	0.582	0.922	0.925	0.916	0.914	0.933	0.937
mAP@50-95	0.45	0.293	0.835	0.703	0.779	0.634	0.843	0.709

Table S1. The given table shows further training that was performed over unpatched image dataset which was downscaled to 1280X1280 image size and ran over 300 epochs using YOLOv8x architecture. There was no significant improvement in the model's performance upon increasing the number of epochs. The table further also shows the performance of YOLOv8-X, YOLOv9-E and YOLOv11-X segmentation neural network models which were trained over patched (340X299) image dataset that were upscaled to an image size of 640X640 for 2000 epochs. Due to no improvement in the performance of the model, the training stopped at 678, 155 and 803 epochs respectively. As can be observed from the table, training for higher number of epochs did no lead to a significant higher performance of the neural networks with the performance being closely similar to what was observed in the 100 epoch trained models.

Figure S1. In the given figure, **A.** Represents the mean mobility across different samples of mitochondria cultures during cycling and **B.** represents mean mobility across different samples of mitochondrial cultures upon glucose starvation when original live cell image sequences were used, a gamma correction was performed to increase the contrast of the image sets and when a binarization of the image sets was done using Otsu thresholding][28]. As can be observed, we see a higher mobility in **A.** compared to **B.** which is more clearly visible upon gamma correction and binarization of the image. However, it also leads to more variability in mobility estimation due to noise introduction.

Figure S2. The given figure represents the mean mobility along with standard deviation across all samples of mitochondrial cultures when cycling and when under glucose starvation. The bar plots also compare the change in mean mobility and standard deviation when a gamma correction was performed. As can be seen in both cases, during cycling there is more mobility in comparison to during glucose starvation.

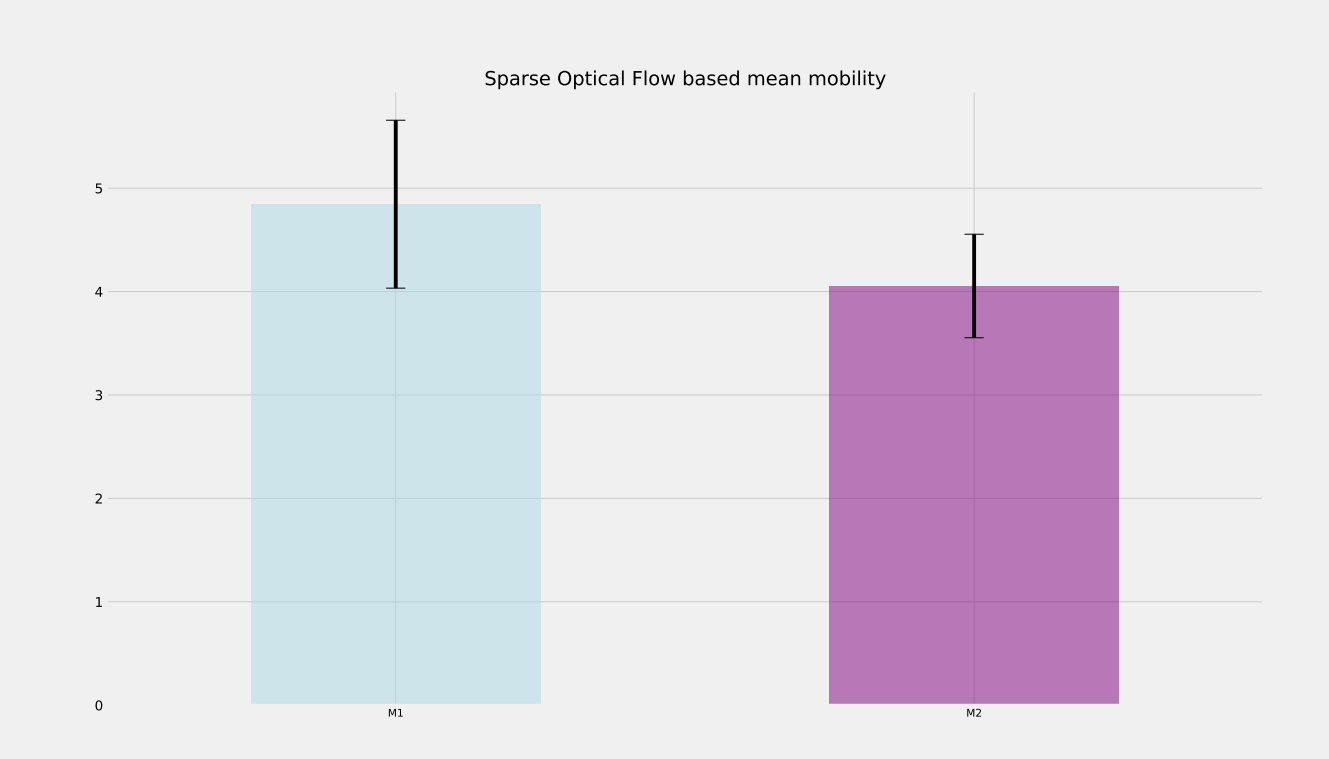

Figure S3. The figure represents the mean mobility and standard deviation of mitochondrial across different samples of macrophages during M1 and M2 phase. As it can be observed in the figures. We have a higher mobility in case of M1 macrophages compared to M2.

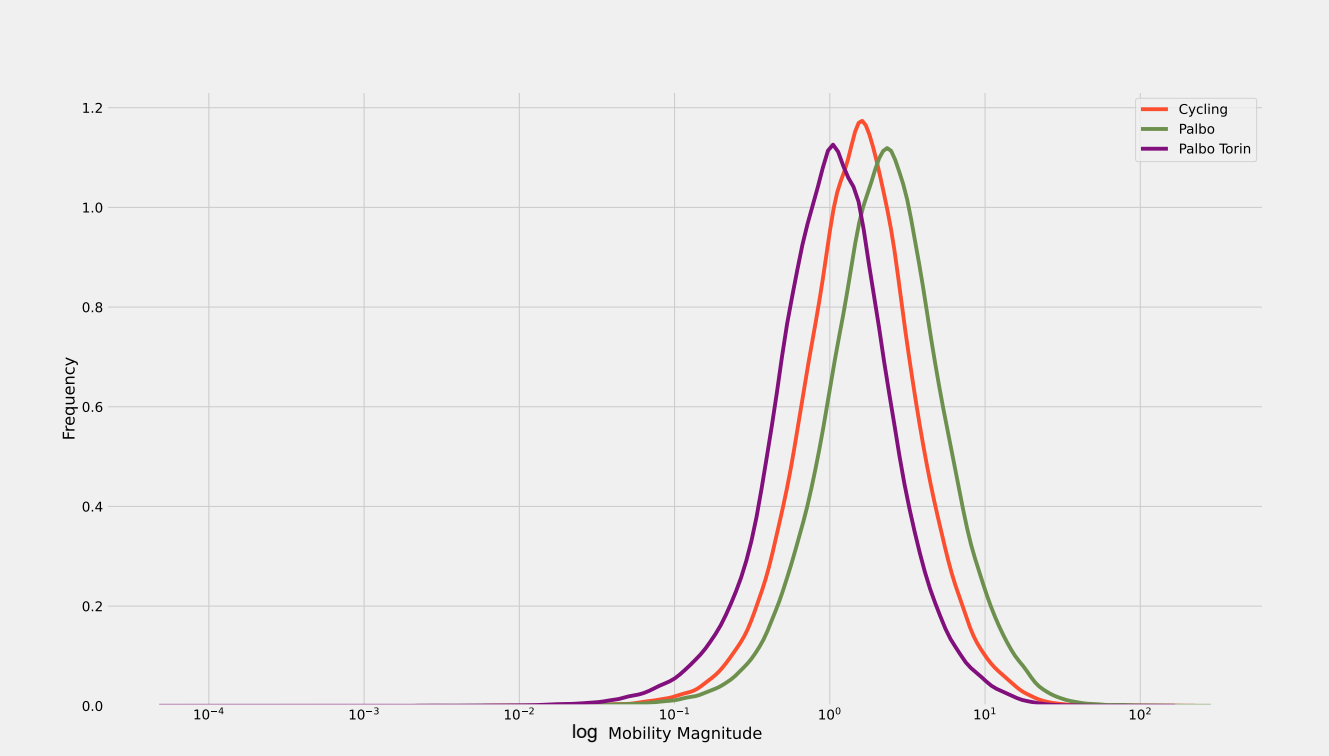

Figure S4. The figure here represents the mobility frequency of 100 binned corner points that were extracted through the Shi-Tomasi corner point detection. These corner points were then tracked using Lukas-Kanade optical flow method. In order to achieve better resolution and remove high variance in the mobility magnitude of different corner points, we took a common logarithm of the mobility magnitude to identify the difference between the three cultures (cycling, palbo and torin) and observed that upon palbo treatment, there is an increase in mobility compared to cycling and there is a reduction in mobility upon palbo+torin treatment.

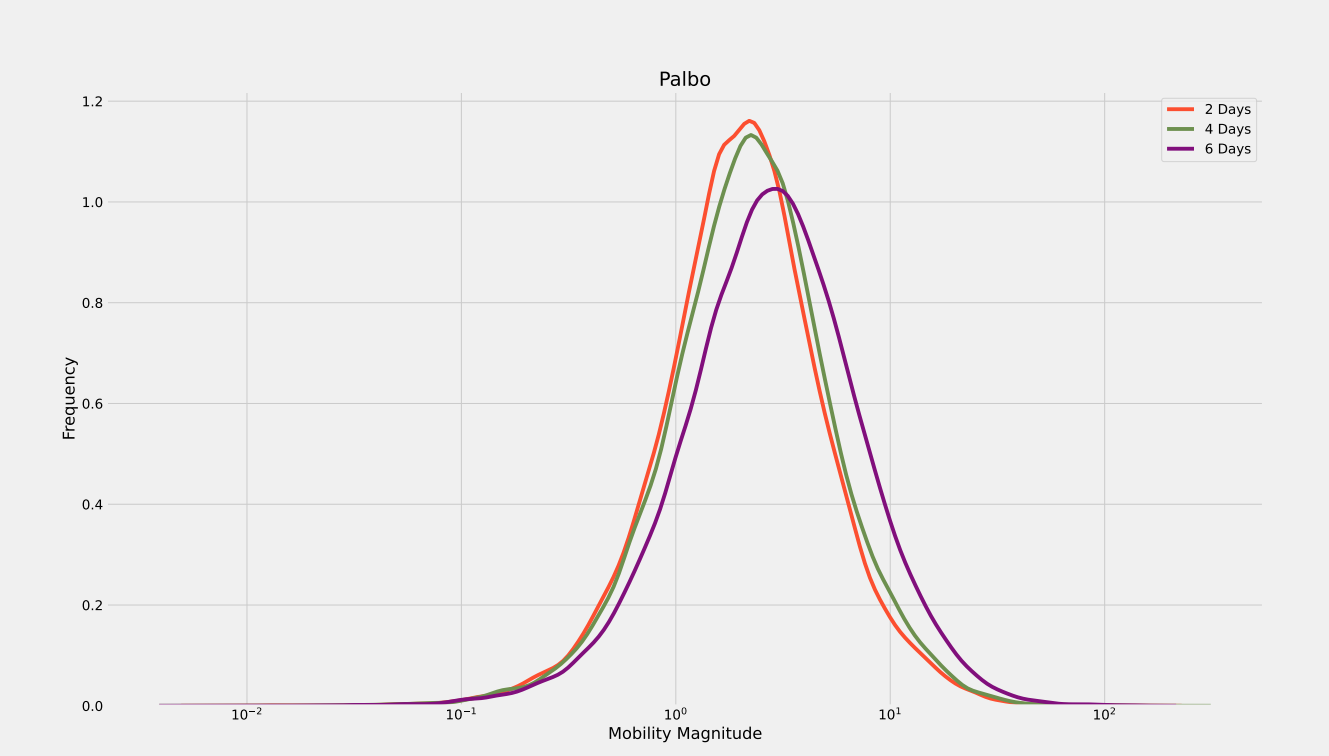

Figure S5. The figure here represents the mobility frequency of 100 binned corner points that were extracted through the Shi-Tomasi corner point detection. These corner points were then tracked using Lukas-Kanade optical flow method. In order to achieve better resolution and remove high variance in the mobility magnitude of different corner points, we took a common logarithm of the mobility magnitude to identify the difference between the Palbo at different time points (2 days, 4 days and 6 days) and observed that there is not a significant change in mitochondrial mobility upon palbo treatment when live-cell imaging was carried out at different time points.