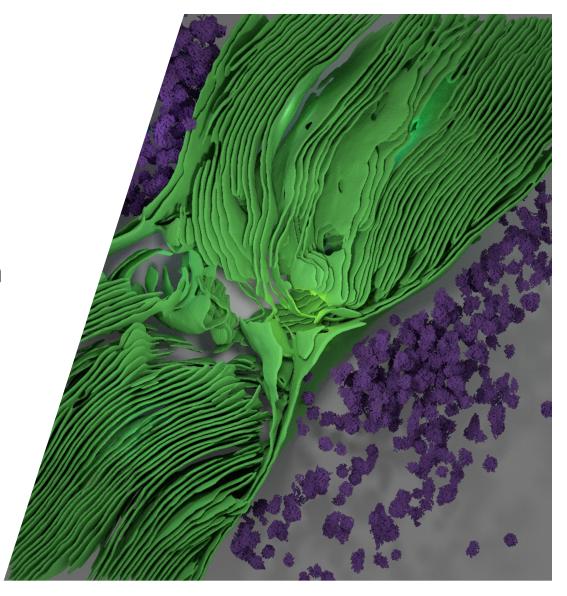


# **Segmentation of Cryo-Electron Tomograms**

Jessica Heebner

June 6, 2025

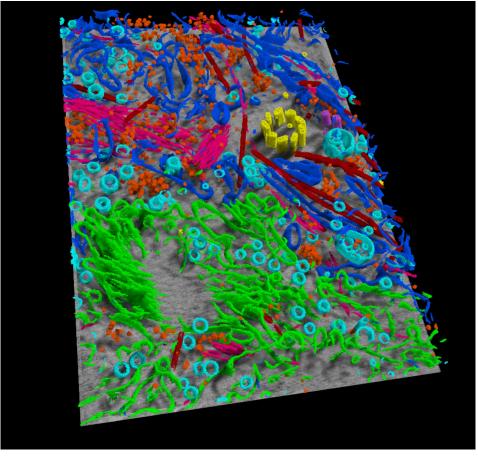




#### The Plan



- 1 Introduction
- The Project That Started It All
- 3 Segmentation Fundamentals
- 4 Why Segment?
- 5 Useful Resources



Data courtesy of Petr Chlanda, University of Heidelberg

#### Who am I?



Veterinary Oncology Nurse

Graduated July 2023



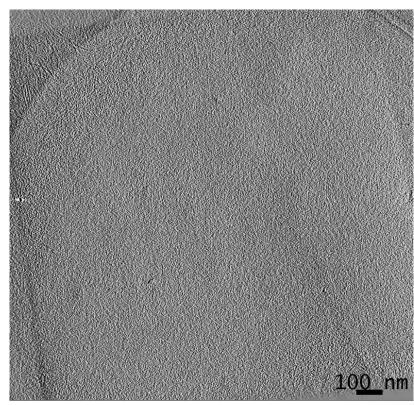
Joined the Apps Team Sept 2023



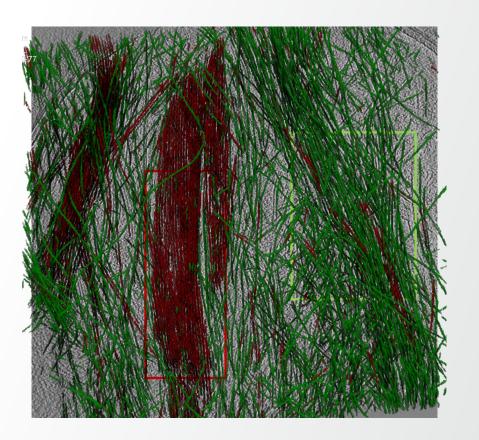


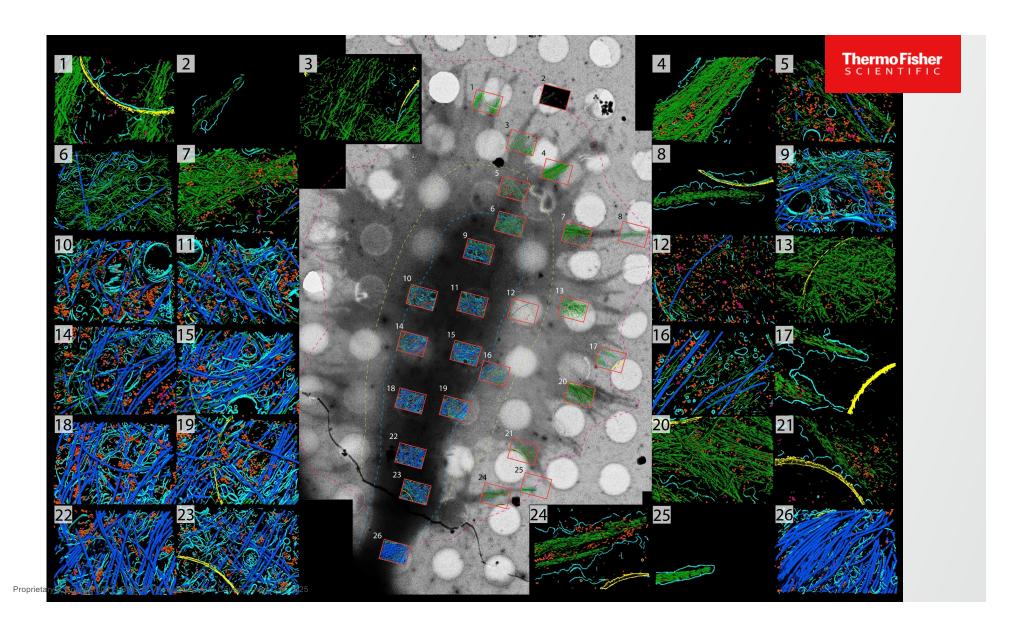
# **How did I get started?**

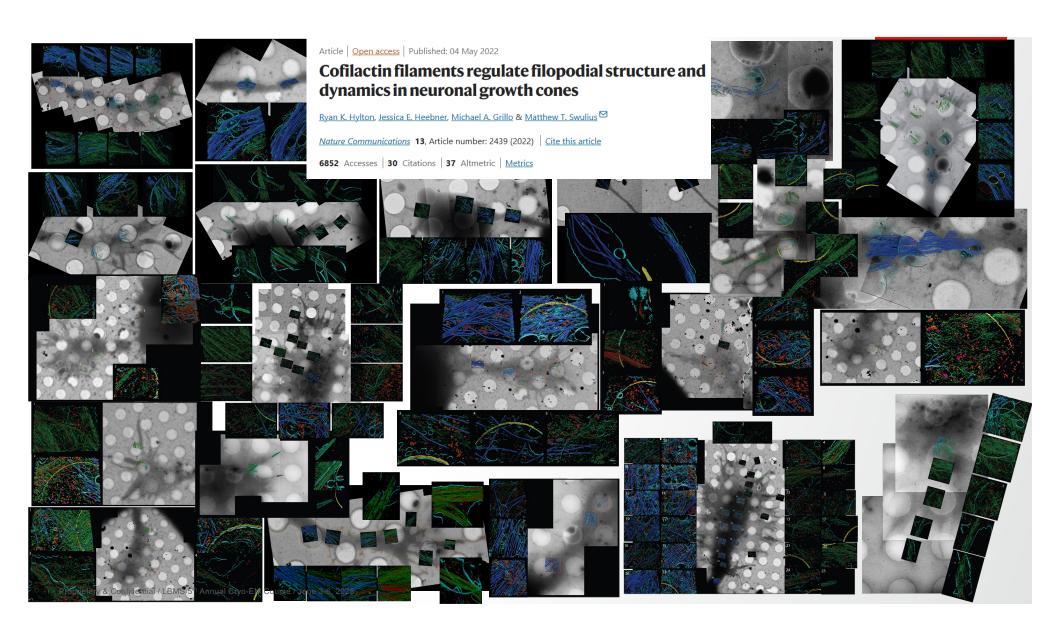




Data courtesy of Matt Swulius, Penn State College of Medicine





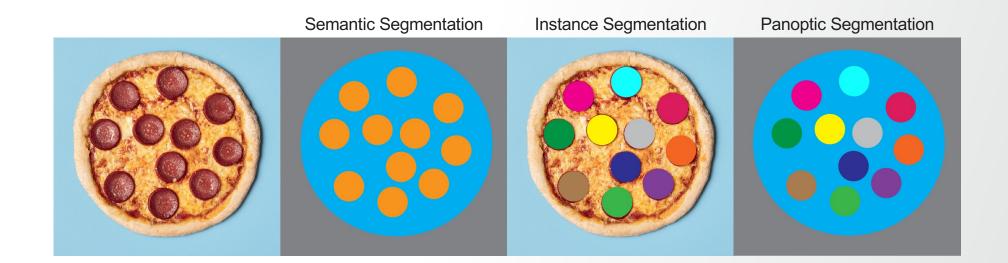




# **Terminology**



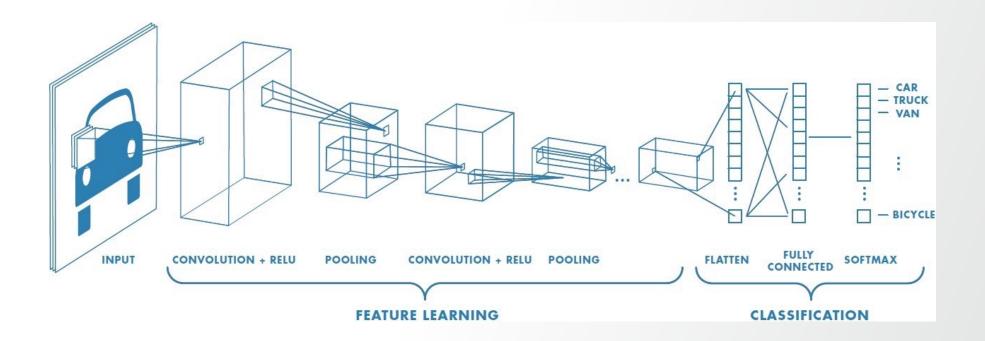
What exactly is segmentation?



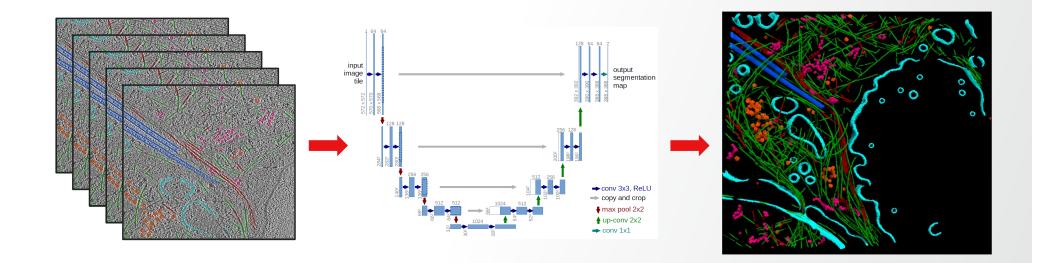


# **Deep Learning**

Convolutional Neural Networks excel at identifying features and labeling images



# **Basic Training Workflow**



#### **Open Source Options**







- Napari Installation guide
  - Python based image viewer
  - Manual and deep learning segmentation plugins





- IMOD <u>Installation guide</u>
  - A suite of software for tomogram reconstruction and processing
  - Manual segmentation tools and a custom CNN that can be trained with your data





- Fiji/ImageJ Download page
  - Software for processing and analyzing scientific images
  - Manual and deep learning segmentation plugins

## Thermo Fisher

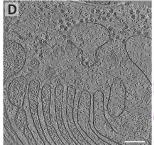
#### **Commercial Options**

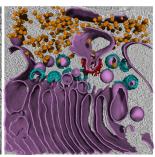
Thermo Fisher Scientific



Dedicated segmentation workroom with tools for manual and AI assisted annotation, as well as the option to train your own CNNs.

• Free Trial



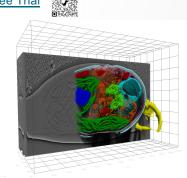


**Comet Technologies** 



Built in segmentation tools for both manual annotation and a framework for training your own CNNs.

• Free Trial



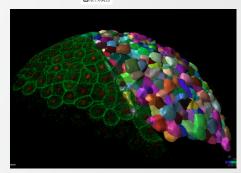
Oxford Instruments



Offers segmentation tools and the ability to train CNNs for your own data.

Free Trial





#### Out of the box solutions?

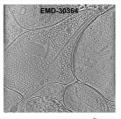
Thermo Fisher

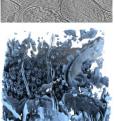
Very few solutions exist that don't require you to provide training data

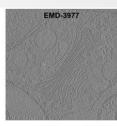
- Membrain-Seg V2 Github repo
- - 3D UNet for segmenting membranes in cryoET data.
  - Pretrained network is provided that works well on many datasets.
  - Can also be trained using your own data.
- DeePiCt Github repo

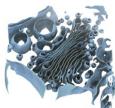


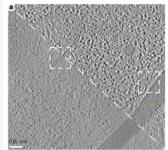
- 2D UNet for segmenting cellular compartments and 3D UNet for segmenting continuous structures (membrane and filaments) and particle localization in cryoET data.
- Pretrained networks are provided that are shown to work on datasets the networks are naïve to in the paper
- Can also be trained using your own data.

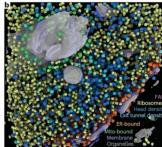










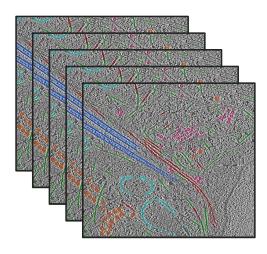


## Thermo Fisher SCIENTIFIC

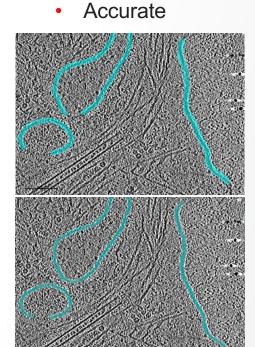
## I've chosen a software, now what?

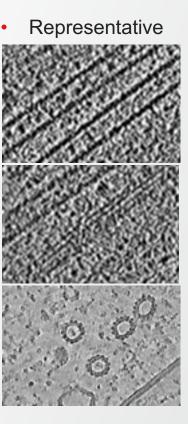
A CNN is only as good as its training data.

How much is necessary?



Start with less, but more is generally better. 3-5 slices usually sufficient for a 2D or 2.5D CNN.





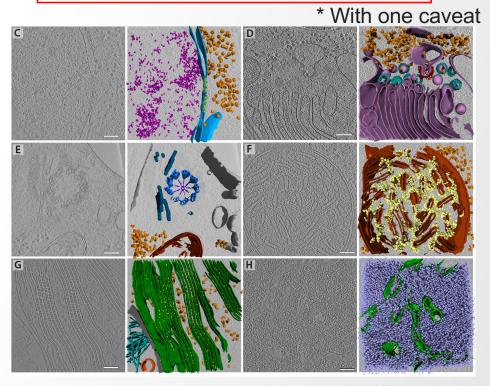
#### What network should I train?

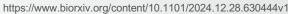
Thermo Fisher SCIENTIFIC

You name it, a UNet exists for it...

- UNet
- 3D UNet
- Attention UNet
- Dense UNet
- Trans UNet
- UNet++
- UNet3+
- Inception UNet
- TernausNet
- R2-UNet
- nnUNet ("No New UNet")
- ... you get the idea

For cryoET data, in most cases a basic UNet performs excellently.







#### Thermo Fisher SCIENTIFIC

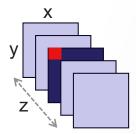
#### **UNet Dimension is Important**

2D



- The network sees and segments small portions of one slice of data at a time.
- Patch size of a network defines how big the "small portion" is in X and Y

2.5D



- The network segments one slice of data at a time but can see pixel intensity from one or two slices above and below to make a more informed segmentation decision.
- Patch size defines the same area as in 2D network.

3D

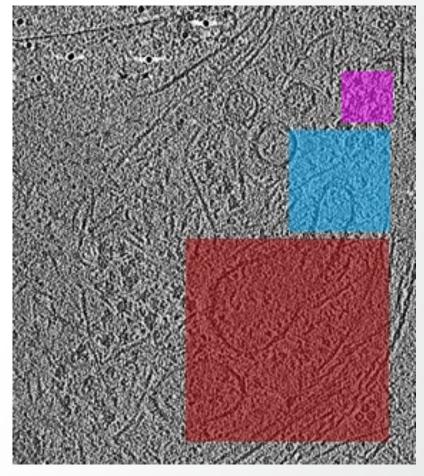


- The network sees and segments a full cube of data.
- Patch size defines the 3D cube.

## How do I determine the right patch size?



- More context is better
- Trade off computationally
  - · Fewer resources required to train a network with small patch size
  - Try to find the largest patch size that will reliably train on the computer you are using.



64 pixels

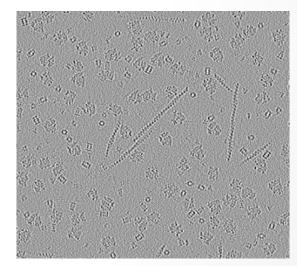
128 pixels

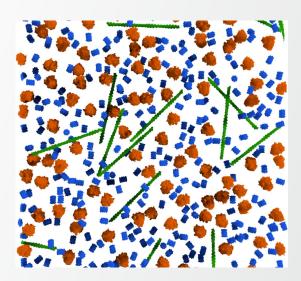
256 pixels

# What about synthetic training data?











**PolNet** 

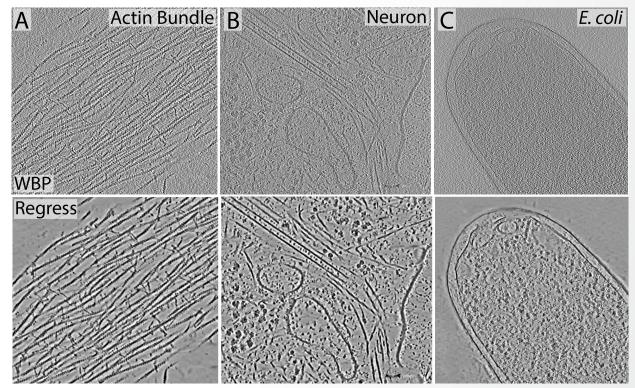


**CryoTomoSim** 

# Thermo Fisher SCIENTIFIC

# **Denoising?**

Should you denoise your data before segmentation?

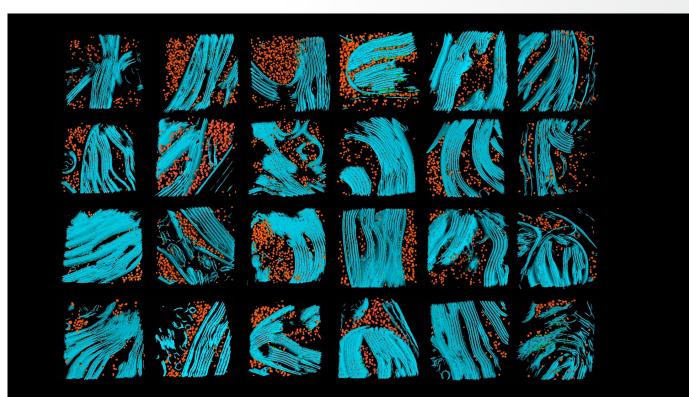


Useful for manual annotation, not necessary for accurate segmentation

## How to optimize for inference/generalizability?



• Representative



Inference relies on a representative training set, often requiring training data from multiple tomograms

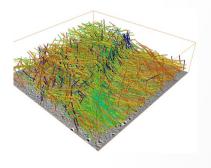
# **Instance Segmentation?**



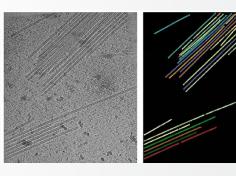
Instance Segmentation



Amira 3D

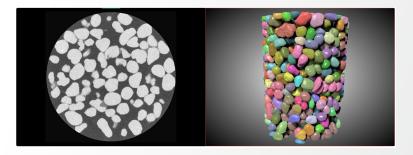


**TARDIS** 





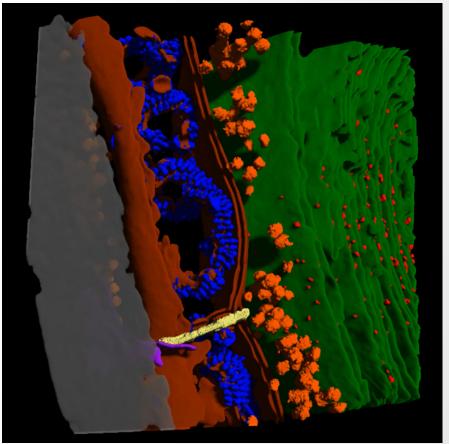
Watershed Transform



## What do you do once the data is segmented?



- When we see more, we learn more
- Analysis requires localization
- CryoET data is exponentially more engaging when it is viewed in 3D



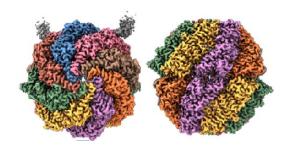
Data from EMD 11830

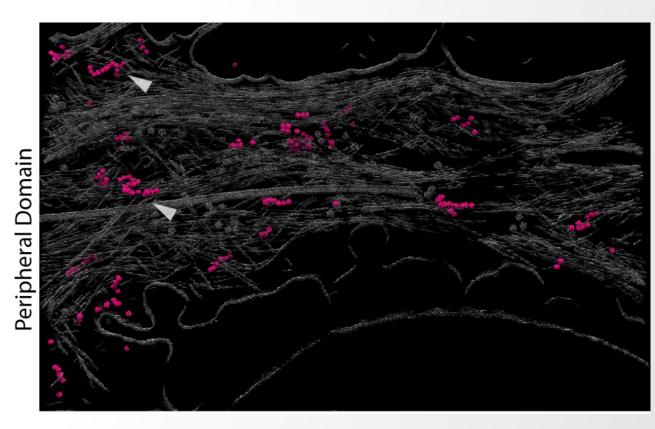
## When we see more, we learn more





#### TRIC/CCT



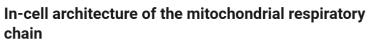


Data courtesy of Matt Swulius, Penn State College of Medicine

#### Thermo Fisher SCIENTIFIC

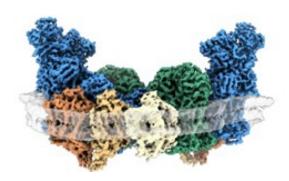
# Respirasomes

RESEARCH ARTICLE | CELL BIOLOGY

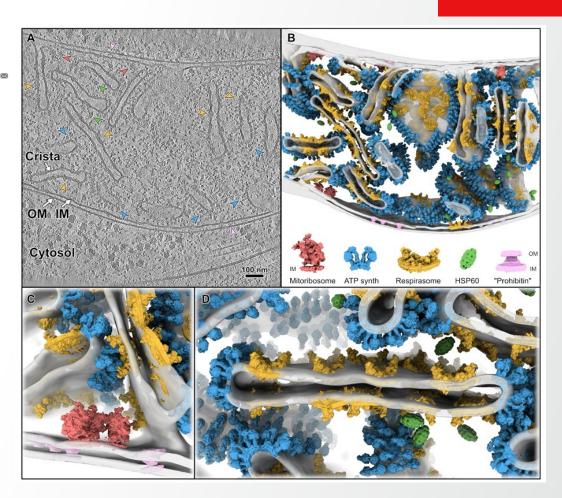


FLORENT WALTZ 📵 , RICARDO D. RIGHETTO 📵 , LORENZ LAMM 📵 , THALIA SALINAS-GIEGÉ 📵 , RON KELLEY 📵 , XIANJUN ZHANG 📵 , MARTIN OBR 📵 , SAGAR KHAVNEKAR (6), ABHAY KOTECHA (6), AND BENJAMIN D. ENGEL (6) Authors Info & Affiliations

SCIENCE · 20 Mar 2025 · Vol 387, Issue 6740 · pp. 1296-1301 · <u>DOI: 10.1126/science.ads8738</u>

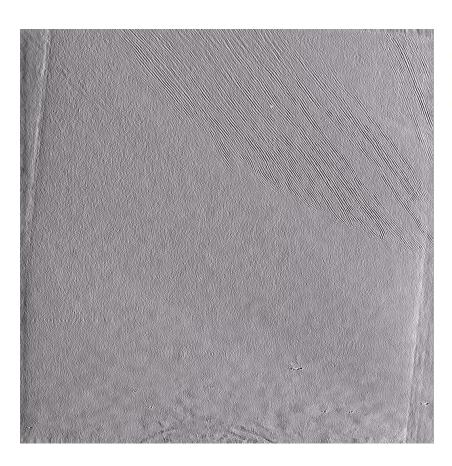






#### Segmentation for in situ structural biology





Towards community-driven visual proteomics with large-scale cryo-electron tomography of Chlamydomonas reinhardtii

🔟 Ron Kelley, 🔟 Sagar Khavnekar, 🔟 Ricardo D. Righetto, 🔟 Jessica Heebner, 🔟 Martin Obr, 🗓 Xianjun Zhang, Saikat Chakraborty, @ Grigory Tagiltsev, @ Alicia K. Michael, @ Sofie van Dorst, @ Florent Waltz,

O Caitlyn L. McCafferty, Lorenz Lamm, Simon Zufferey, Philippe Van der Stappen, Hugo van den Hoek,

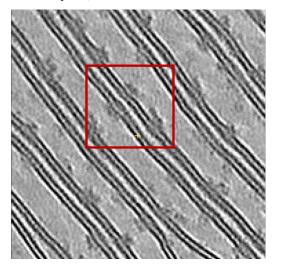
10 Wojciech Wietrzynski, 10 Pavol Harar, 10 William Wan, 10 John A.G. Briggs, 10 Jürgen M. Plitzko,

Benjamin D. Engel, Abhay Kotecha

doi: https://doi.org/10.1101/2024.12.28.630444

This article is a preprint and has not been certified by peer review [what does this mean?].

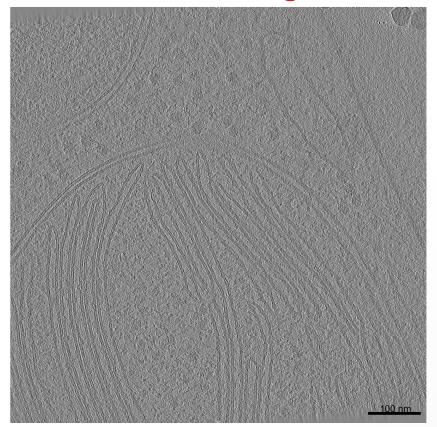
Membrane embedded complexes PSI, PSII, Cyt b<sub>6</sub>, etc.

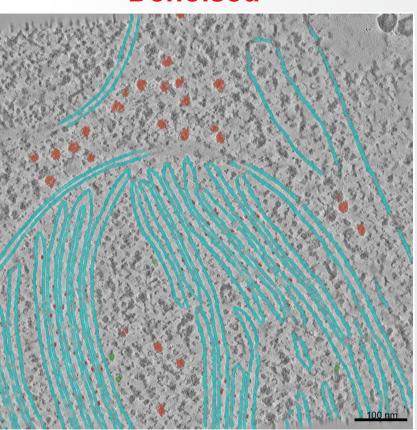


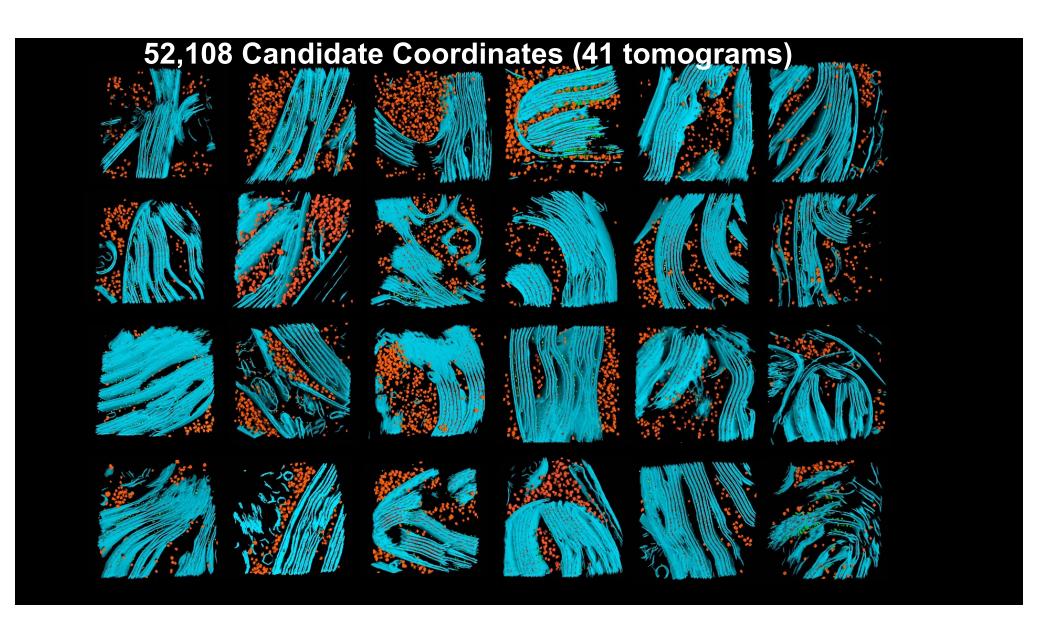


# **Unfiltered Tomogram**

#### **Denoised**

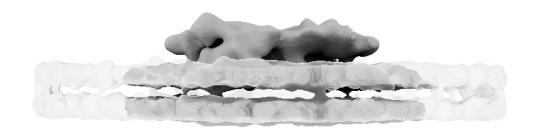






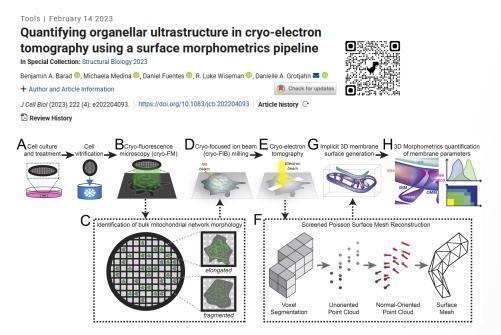
# Photosystem II, 19 Å

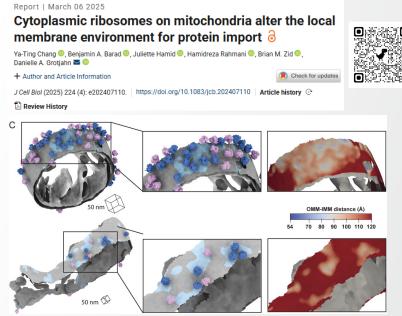






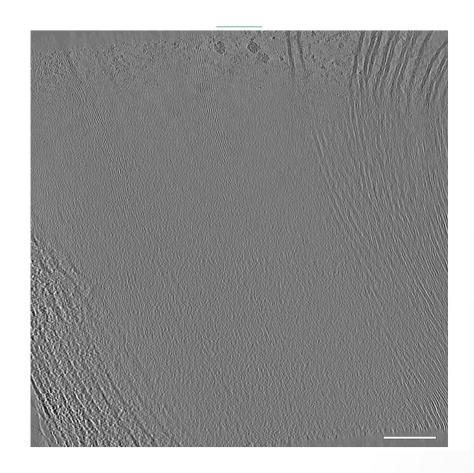
#### **Quantifying organellar ultrastructure**

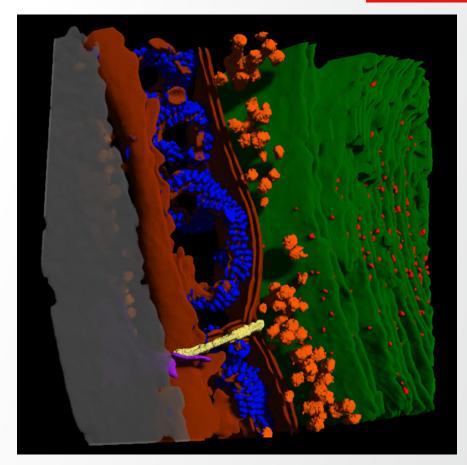




#### **Mitochondrial Fission**

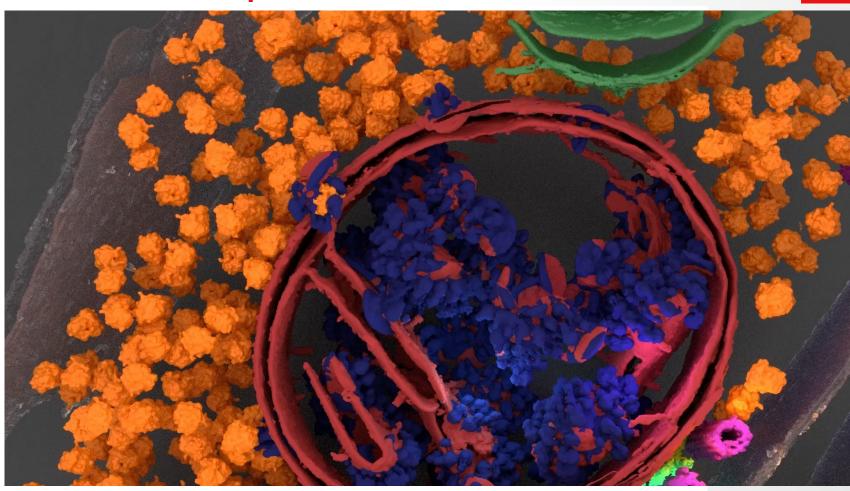






# **Mitochondria Transport**





#### **Useful Resources**



MemBrain-seg **Installation Guide** and Tutorials



Napari Installation Guide and **Tutorials** 



Image.sc Forum



**Amira Tutorials** 

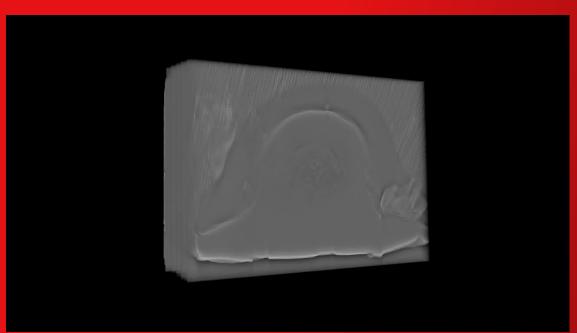


**Dragonfly Tutorials** 



**JoVE Segmentation Tutorial** 





# Questions

**EMPIAR 11275** 

Thermo Fisher