Image analysis of biomedical volume electron microscopy data using deep learning and crowdsourced training data

Martin Jones

Electron Microscopy
Science Technology Platform

The Francis Crick
Institute

@martinjones78
@EtchACell

The Francis Crick Institute

THE FRANCIS CRICK INSTITUTE

- We are a biomedical discovery institute researching the biology underlying human health
- Six founding partners
- Regular exhibitions
- Based in London, near St Pancras (12 minutes walk from IOP HQ)

CANCER RESEARCH UK

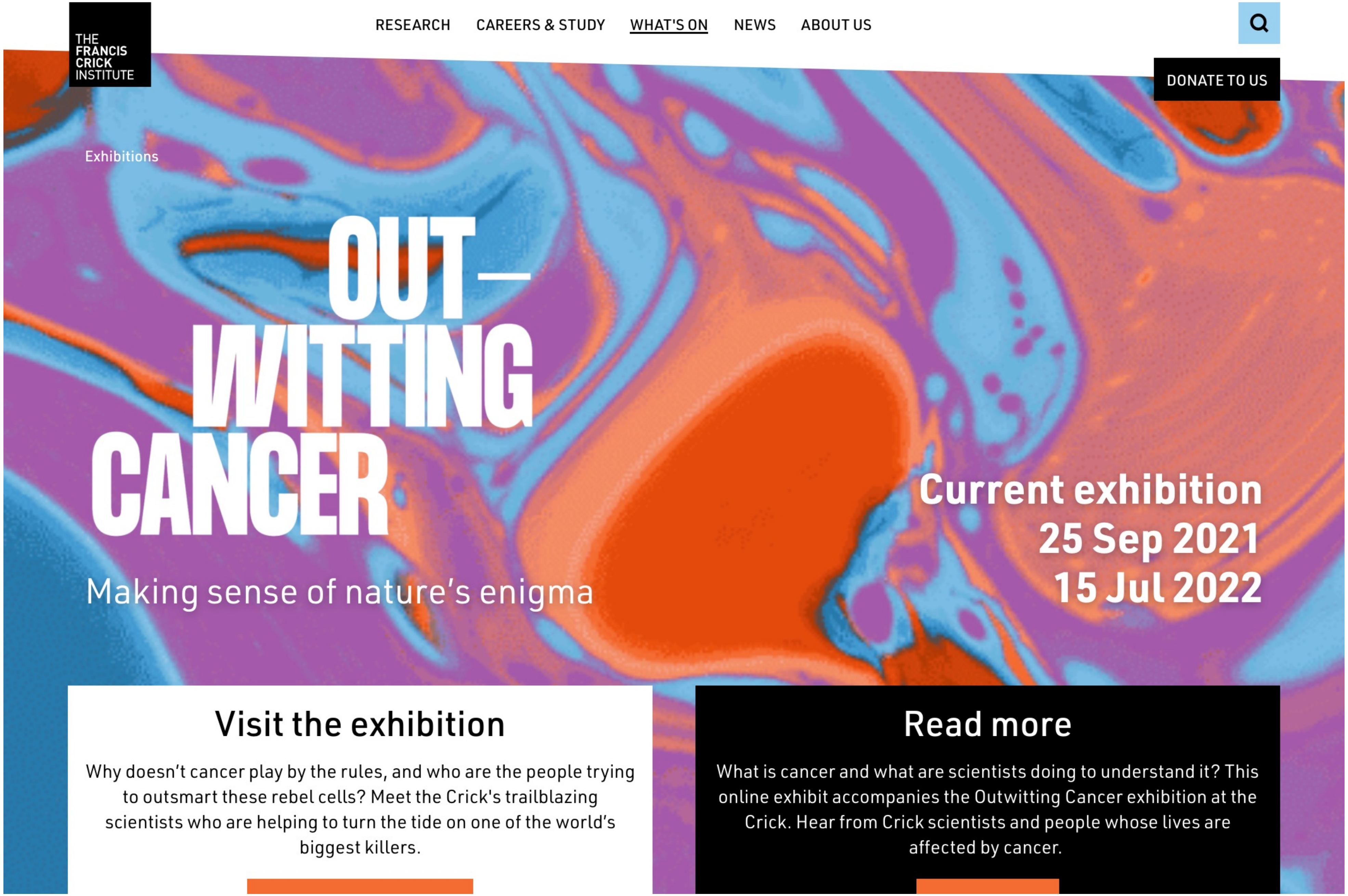
wellcome

MRC Medical Research Council

KING'S College LONDON

Imperial College London

UCL

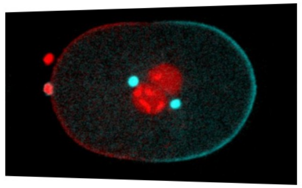


THE FRANCIS CRICK INSTITUTE

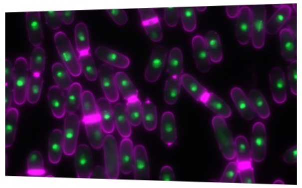
100+ labs, 1250+ researchers, 17 Science Technology Platforms

THE FRANCIS CRICK INSTITUTE

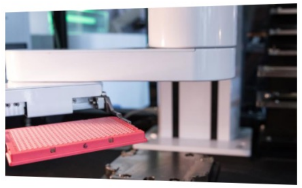
Biochemistry & Proteomics



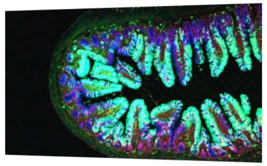
Cell Biology



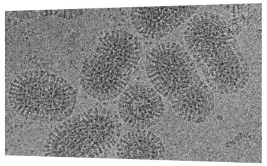
Cell Cycle & Chromosomes



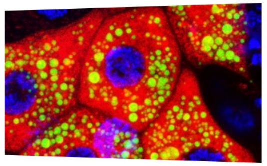
Chemical Biology & High Throughput



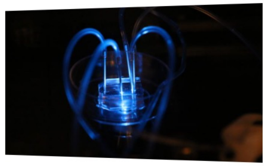
Immunology



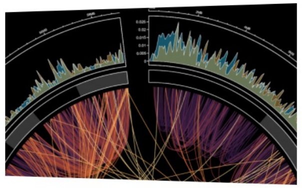
Infectious disease



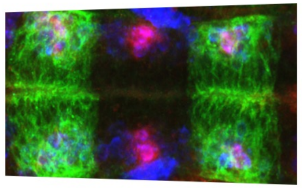
Metabolism



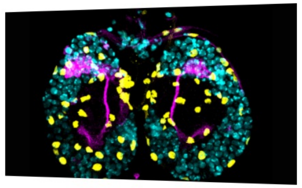
Microfabrication & Bioengineering



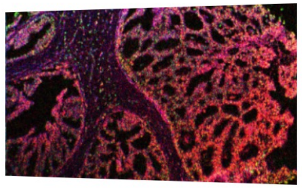
Computational & Systems Biology



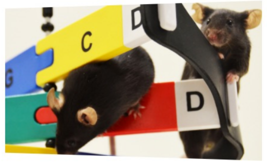
Developmental Biology



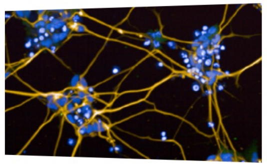
Ecology, Evolution & Ethology



Gene Expression

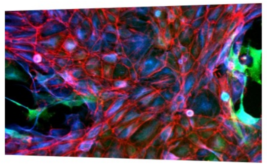


Model organisms

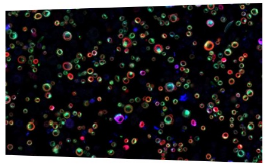


Neurosciences

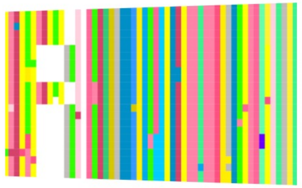
It remains a major research challenge of the 21st century to understand how neural circuits in the brain and nervous system process information; respond to the environment, injury or disease; and give rise to our actions and behaviour.



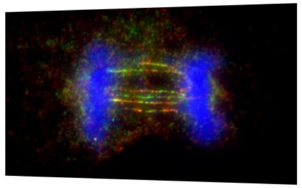
Signalling & Oncogenes



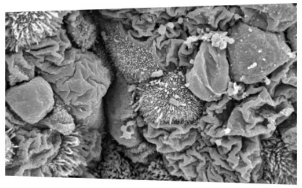
Stem Cells



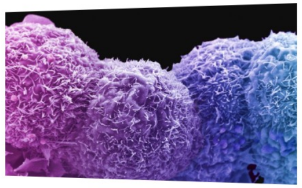
Genetics & Genomics



Genome Integrity & Repair

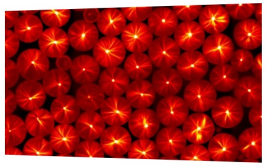


Human Biology & Physiology

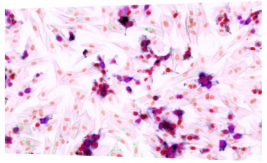


Imaging

Structural Biology & Biophysics



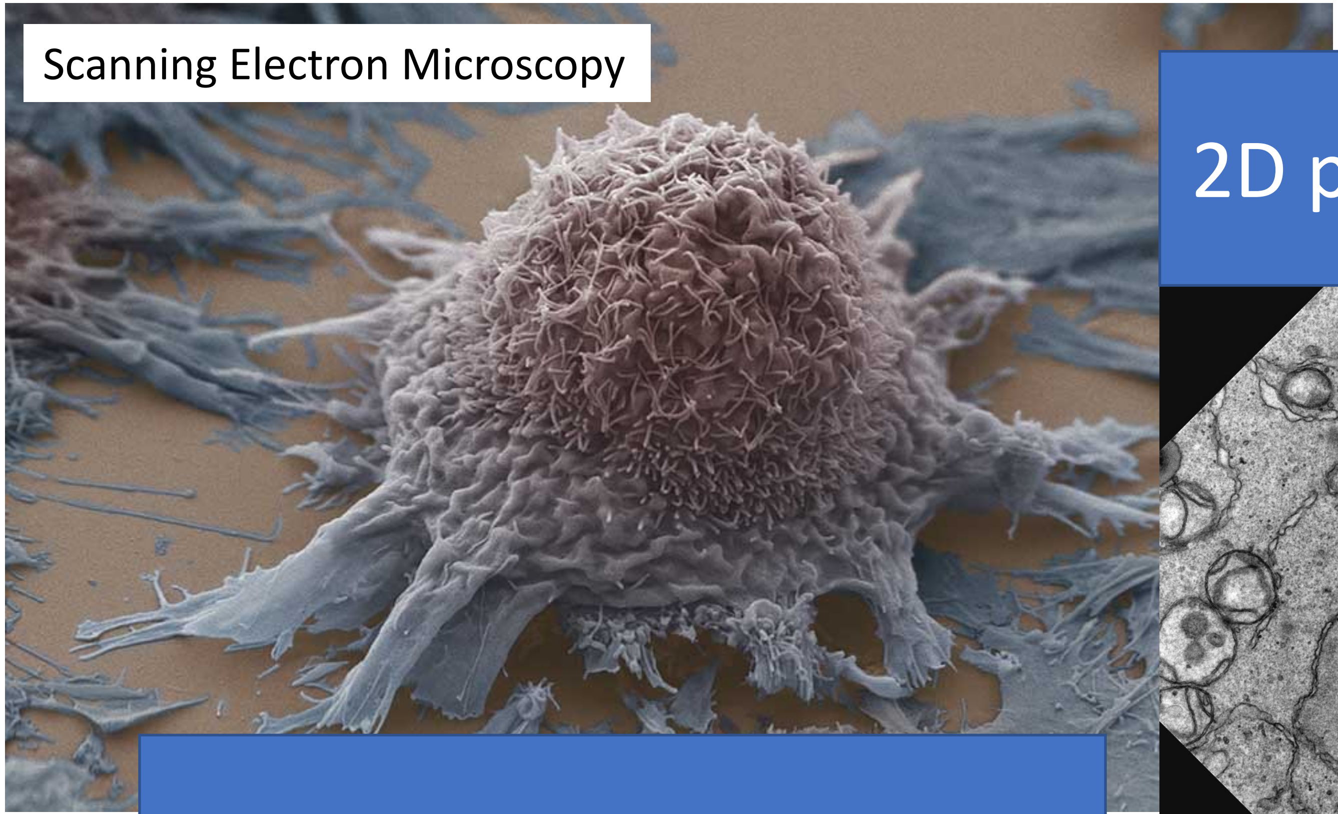
Synthetic Biology



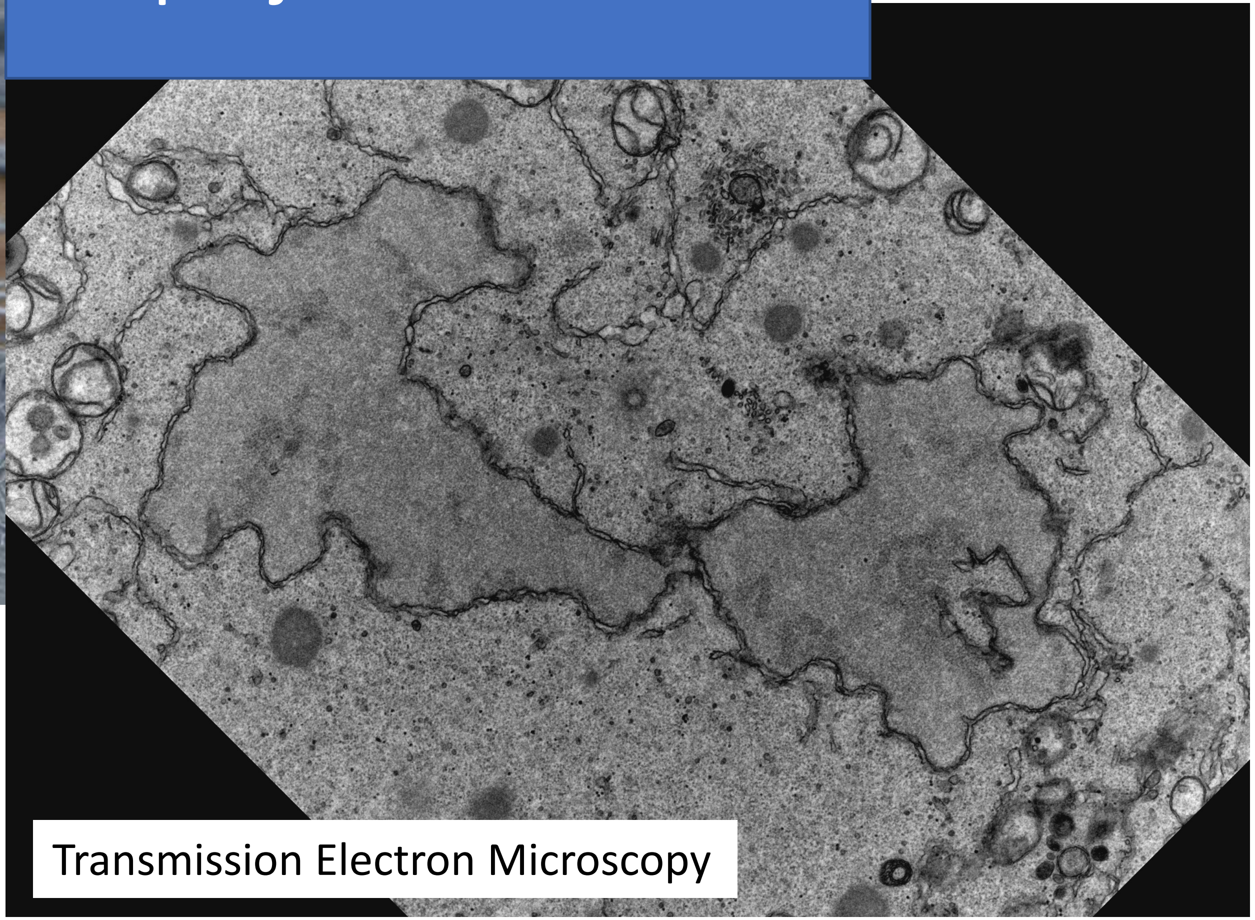
Tumour Biology

“Traditional” Electron Microscopy

THE FRANCIS CRICK INSTITUTE



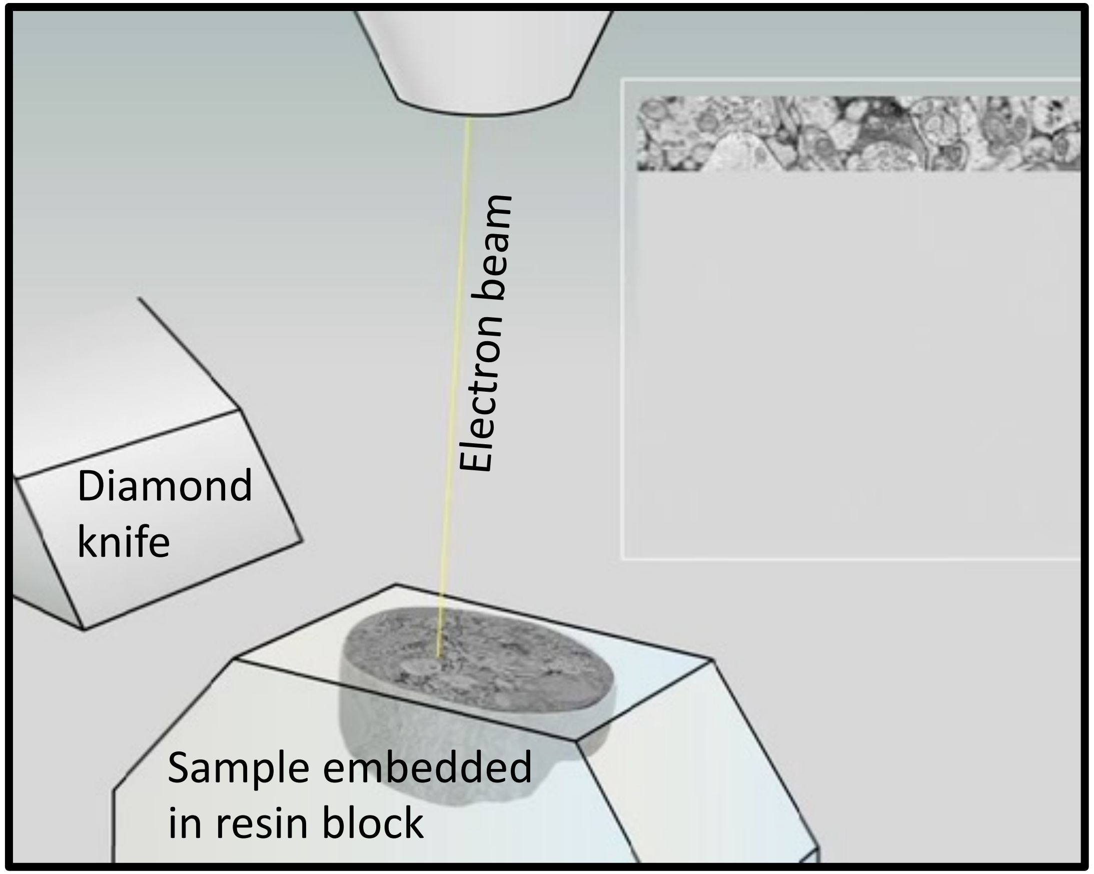
2D projections or slices



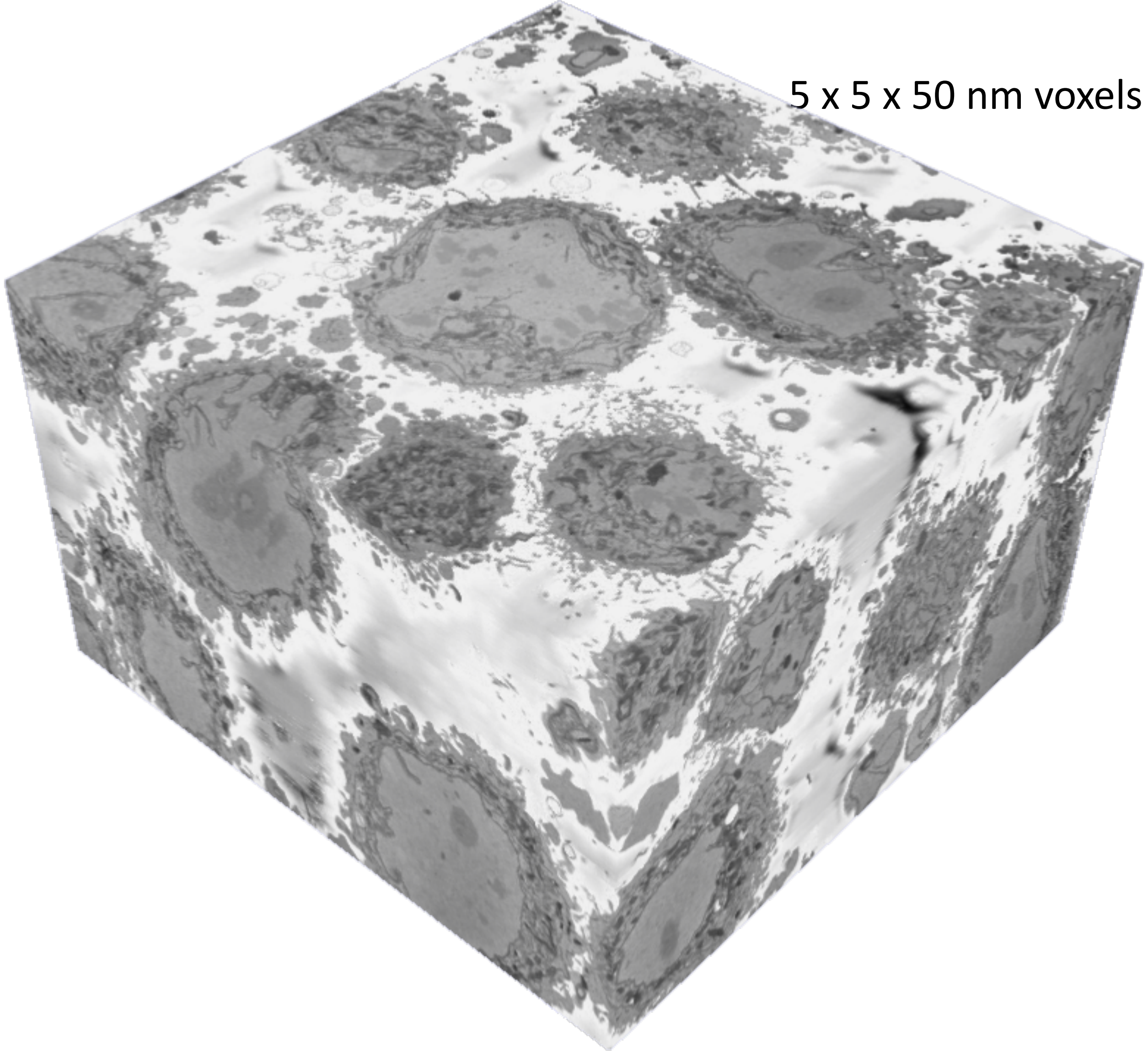
Manual analysis is tough, but images arrive relatively slowly

Automated volume electron microscopy

Serial Block Face Scanning Electron Microscopy

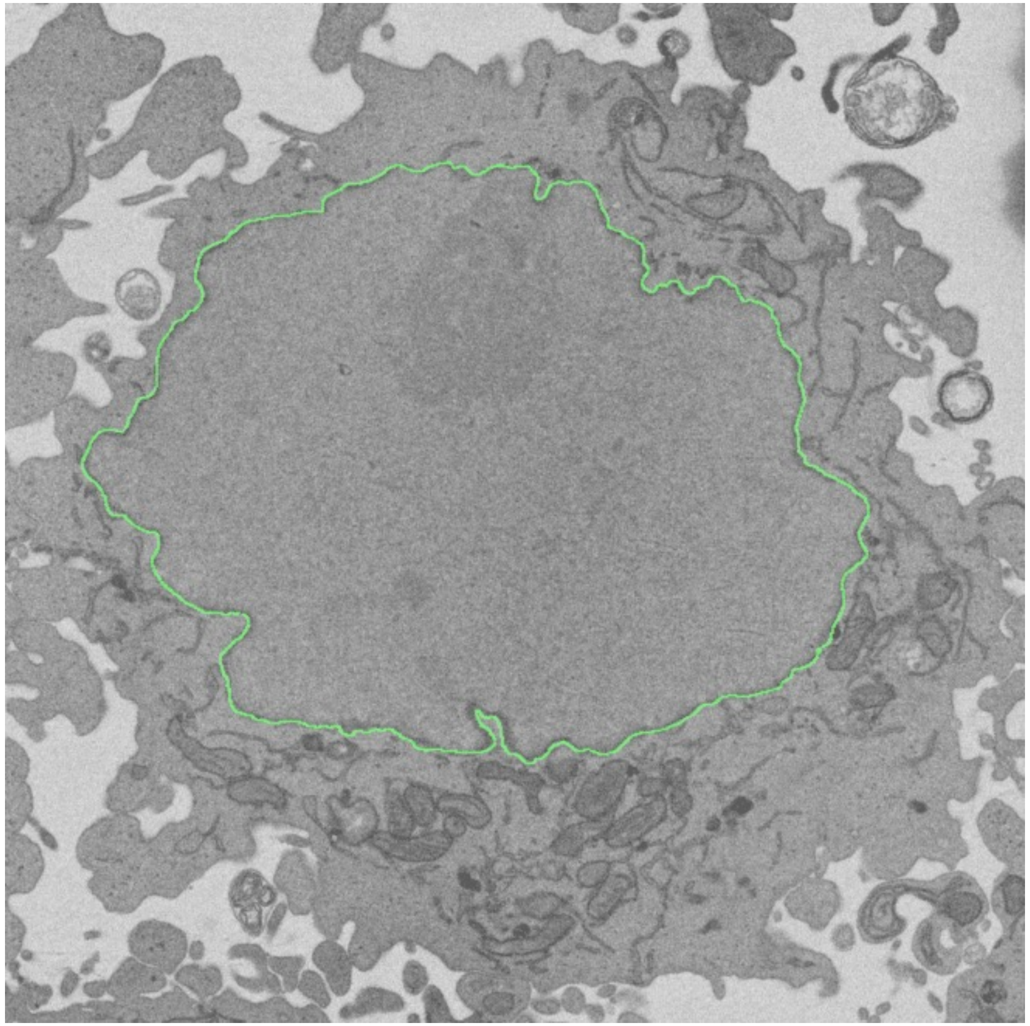


http://www.gatan.com/products/sem-imaging-spectroscopy/3view-system



5 x 5 x 50 nm voxels

Extracting meaning – Segmentation



- Study and quantification of shape requires segmentation – delineating objects of interest from background

- Allows quantitative comparison of healthy and diseased cells and tissues

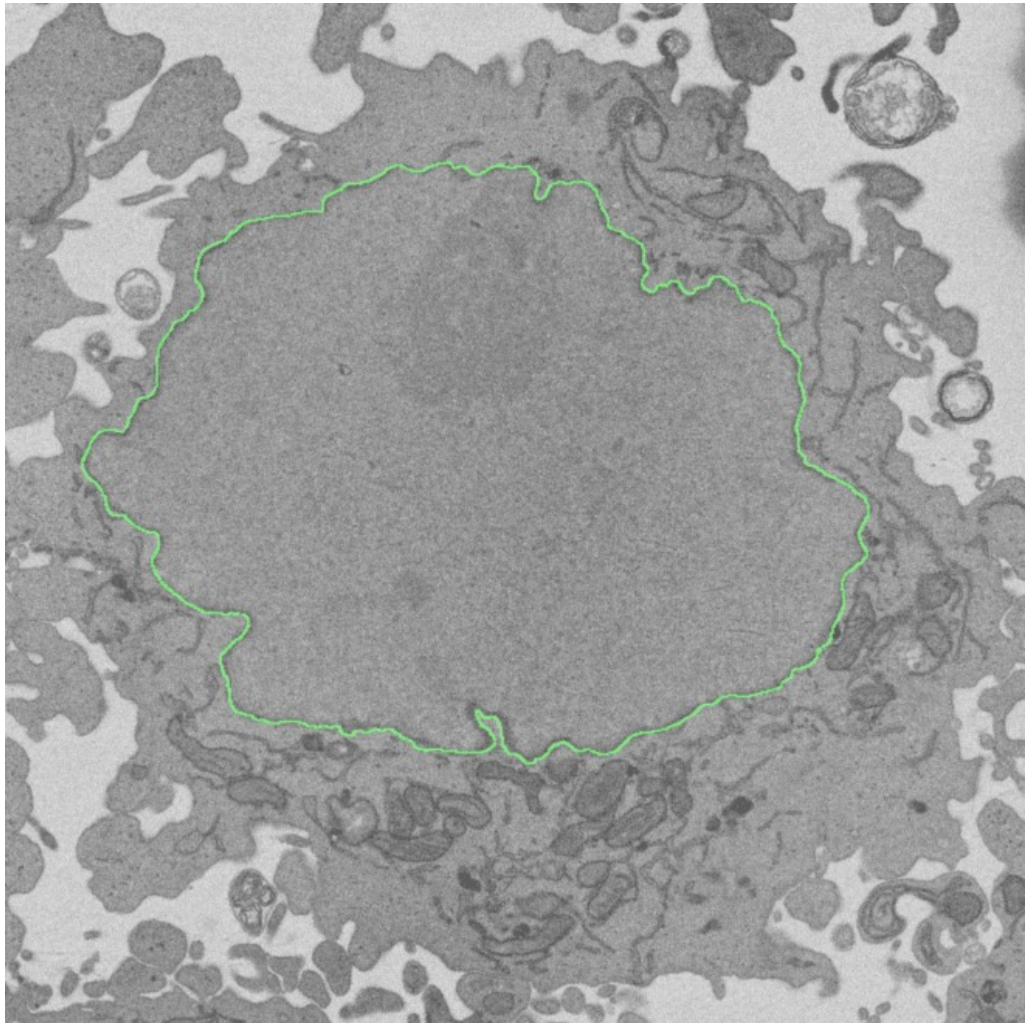
Image segmentation

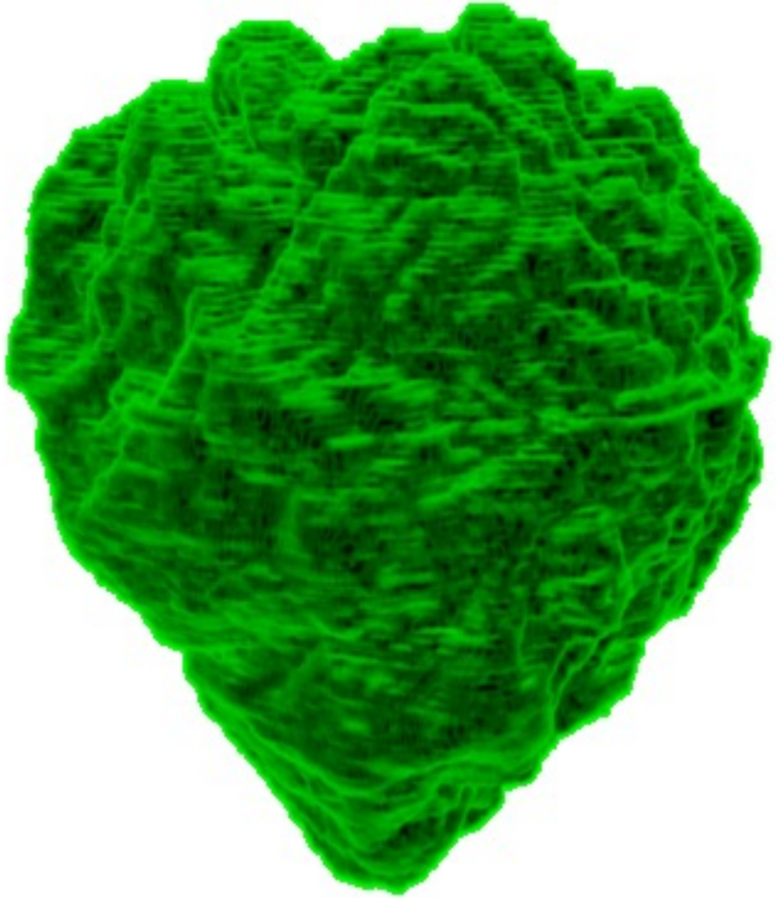
2D views of 3D scenes can be misinterpreted

THE FRANCIS CRICK INSTITUTE

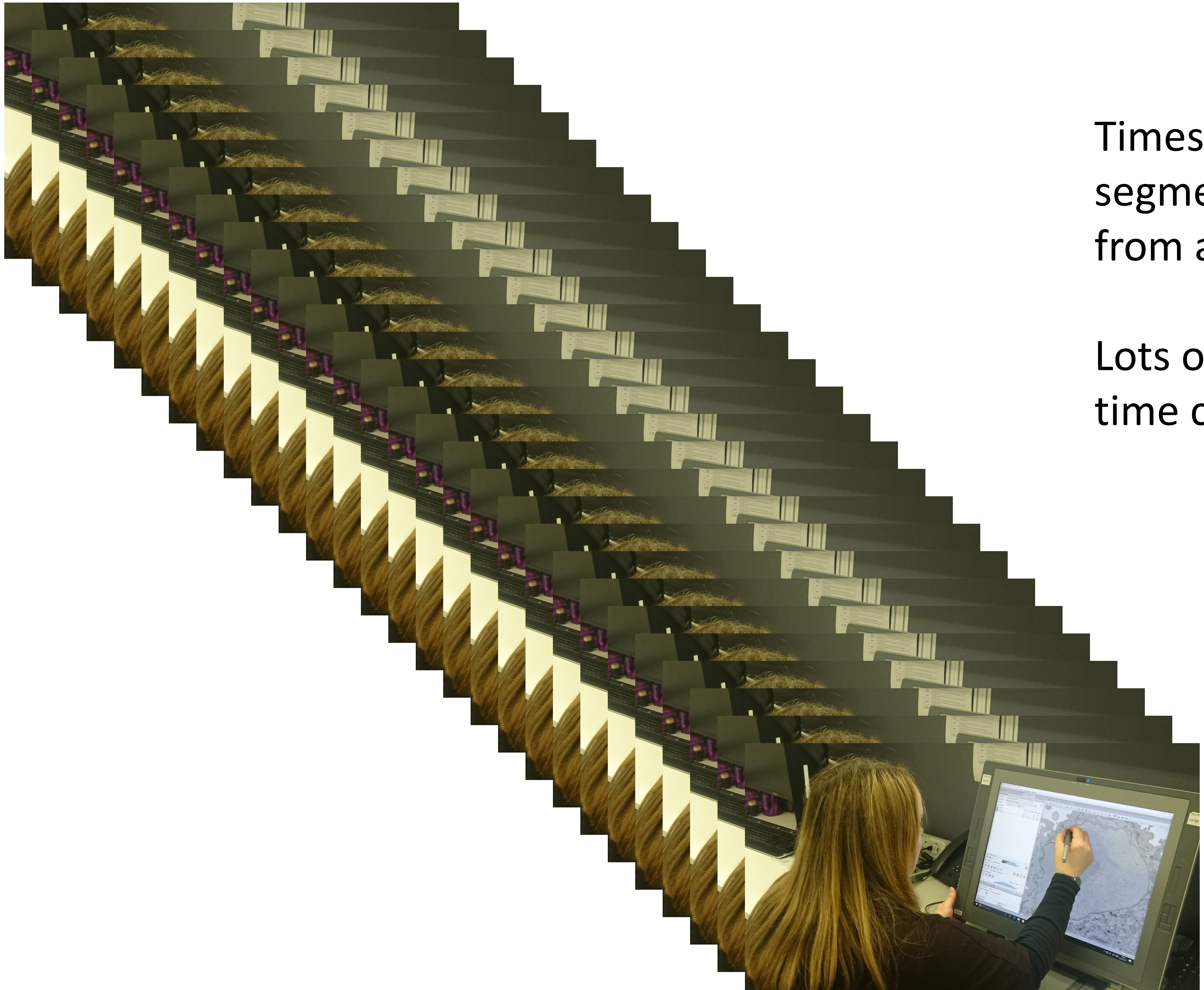
Extracting meaning – Segmentation in 3D

THE FRANCIS CRICK INSTITUTE





Expert segmentation – low throughput



Timescales can be days/weeks/months for segmentation of a single type of structure from a single day's acquisition

Lots of information is left un-mined due to time constraints

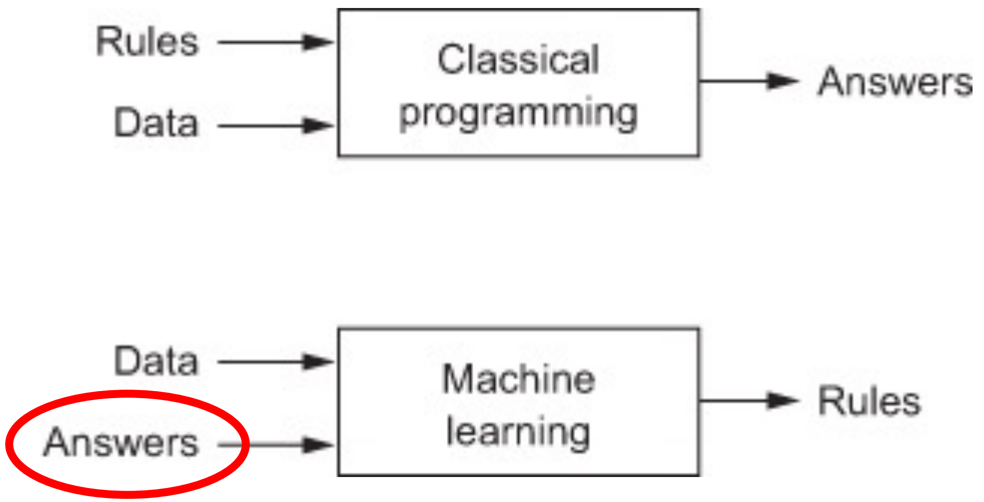
THE FRANCIS CRICK INSTITUTE

Sounds like a job for Deep Learning!

Revolutionising many image analysis tasks

Can be trained to generalise

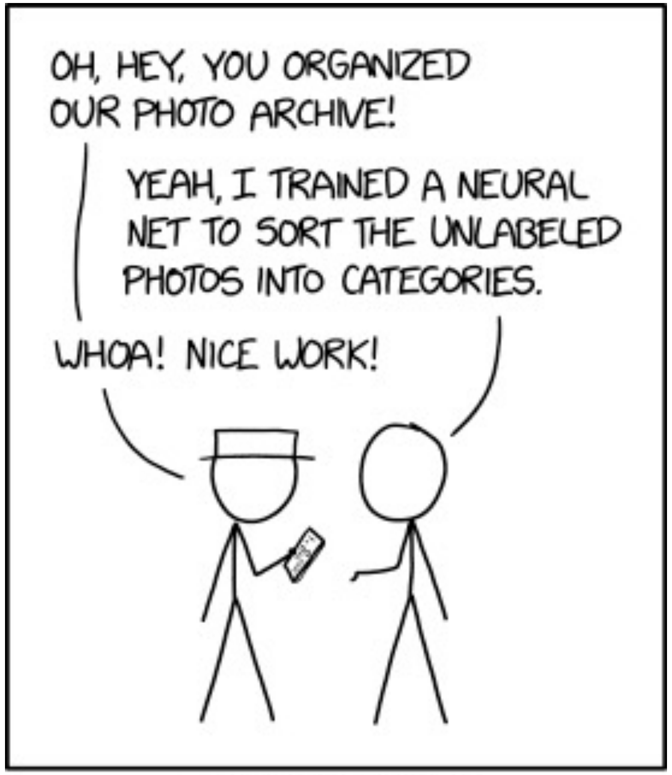
BUT Requires lots of training data



From Deep Learning with Python by Francois Chollet

- For many tasks, human analysis is still the gold standard
  - So, we should use humans!

- But we have ***lots*** of images to analyse
  - So we need ***lots*** of humans!



ENGINEERING TIP:
WHEN YOU DO A TASK BY HAND, YOU CAN TECHNICALLY SAY YOU TRAINED A NEURAL NET TO DO IT.

https://www.xkcd.com/2173/

THE FRANCIS CRICK INSTITUTE

Do you *always* need a PhD to do this?

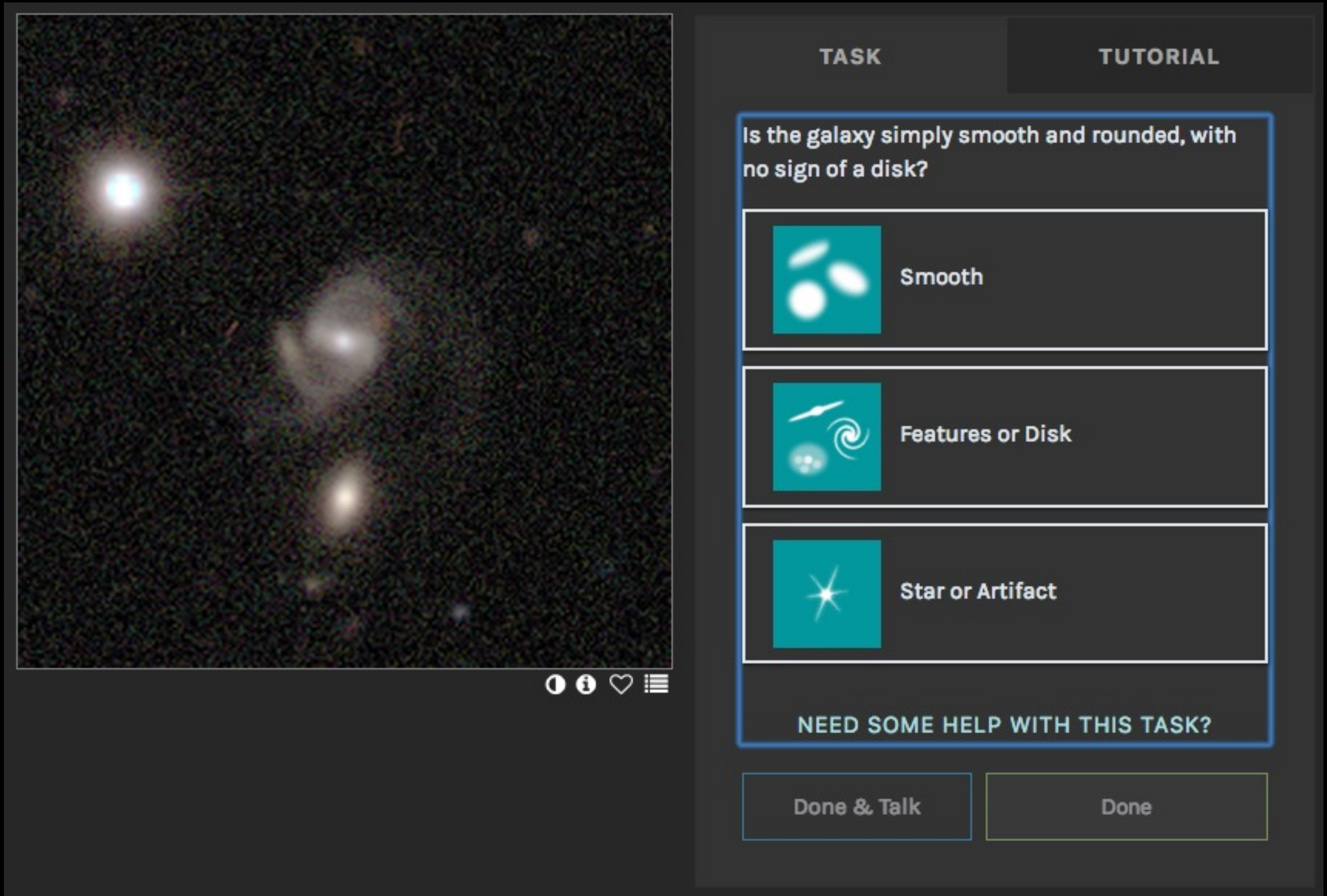
THE FRANCIS CRICK INSTITUTE

- Can question be phrased simply?
  - *"if you see a line like this, trace it"*
  - *"how many objects that look like this can you see?"*
  - *"has this image been corrupted in some way?"*

- For "Human-easy : Computer-hard" problems, *Crowdsourcing* is a way to get many non-experts to contribute to research

Galaxy Zoo (galaxyzoo.org)

THE FRANCIS CRICK INSTITUTE



Etch a Cell (bit.ly/etchiverse)

@EtchACell

THE FRANCIS CRICK INSTITUTE

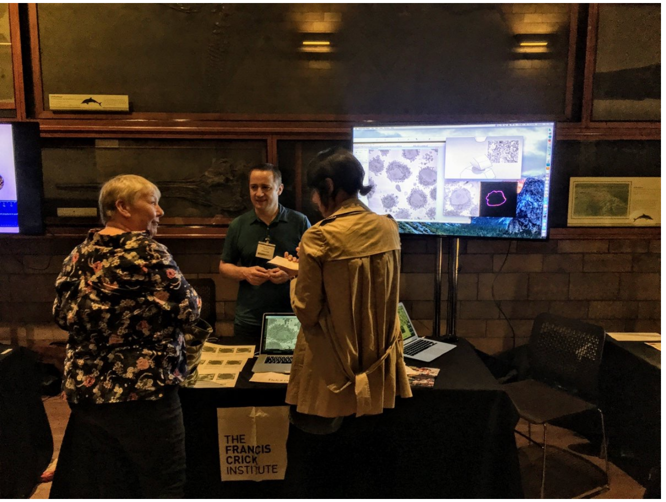
Nuclear envelope

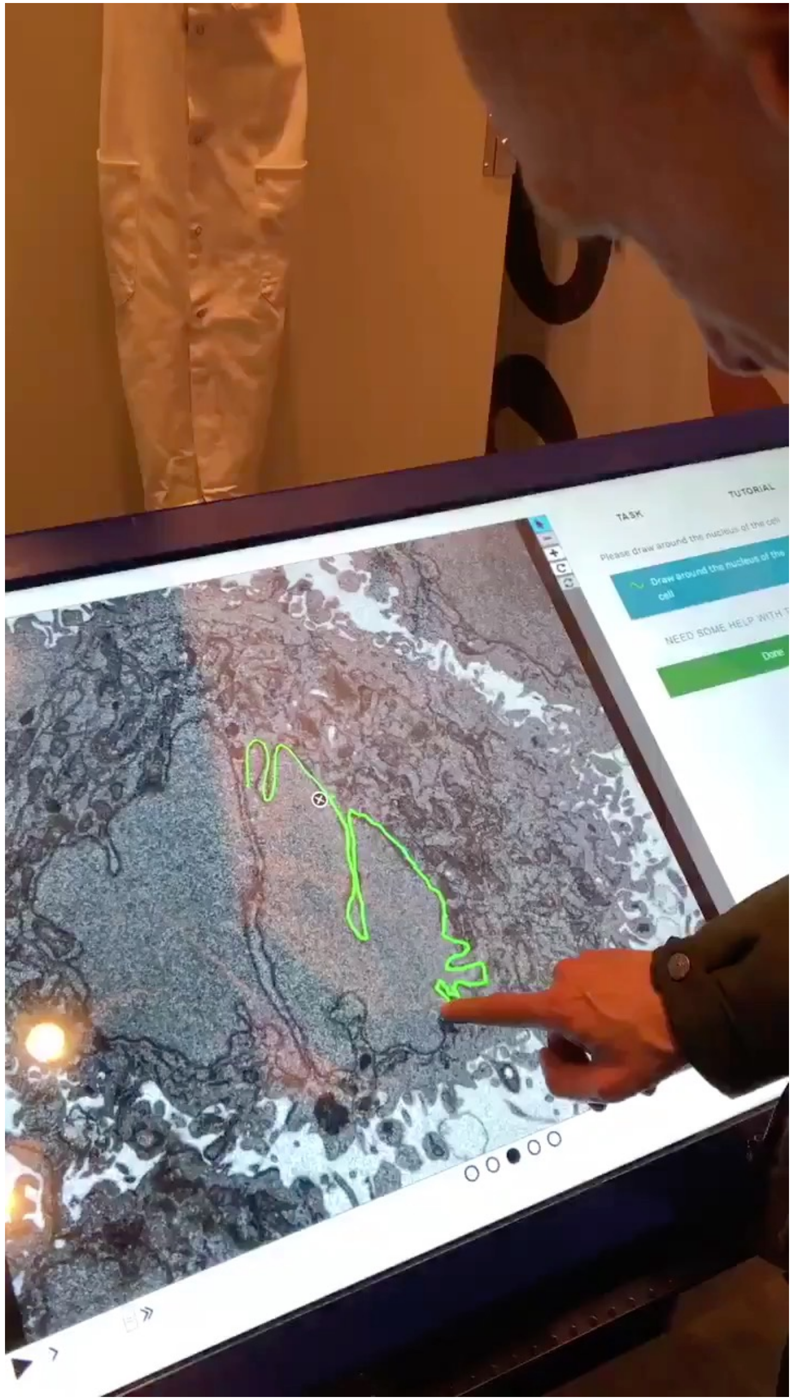
Mitochondria

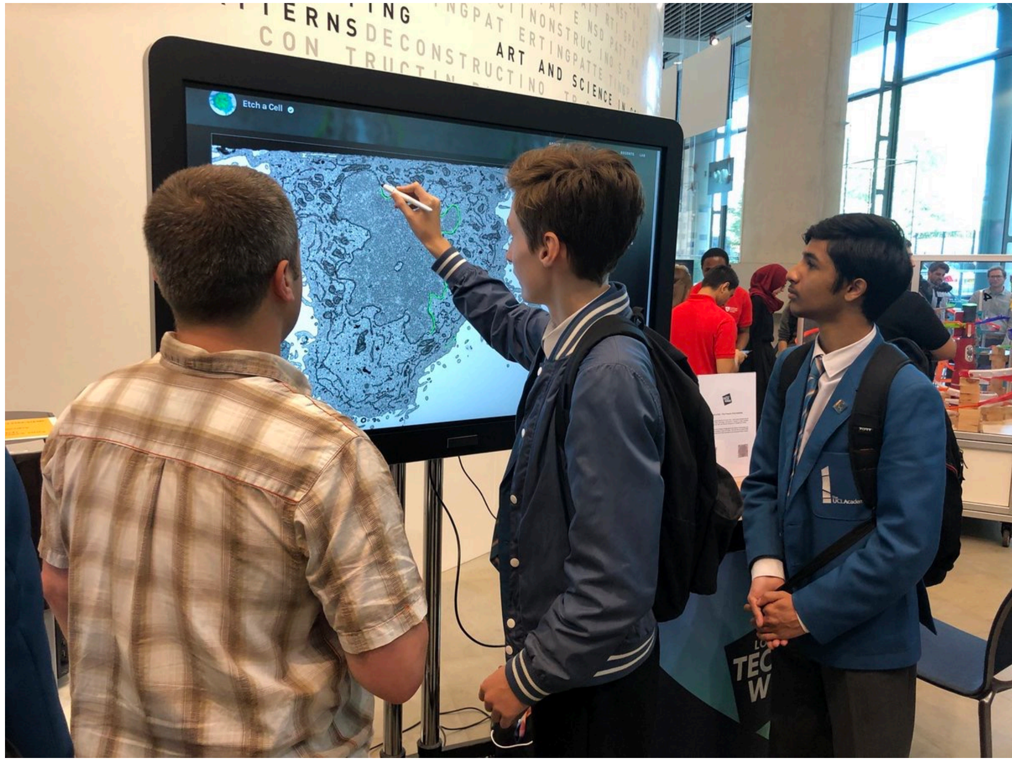
Mito + more

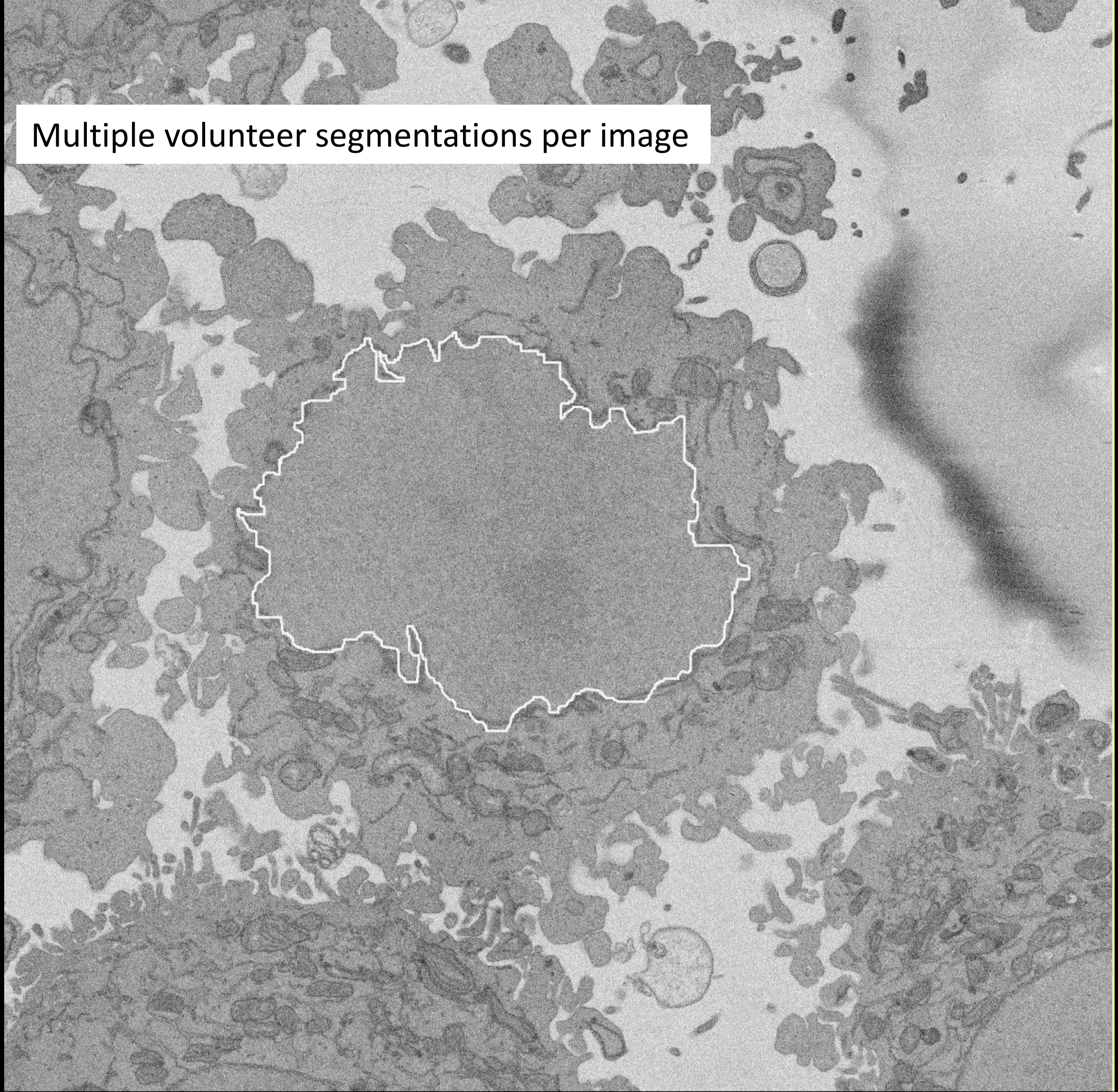
ER


LET'S BEAT CANCER SOONER
CANCER RESEARCH UK
OUR RESEARCHERS IN MANCHESTER ARE BEATING CANCER


THE FRANCIS CRICK INSTITUTE

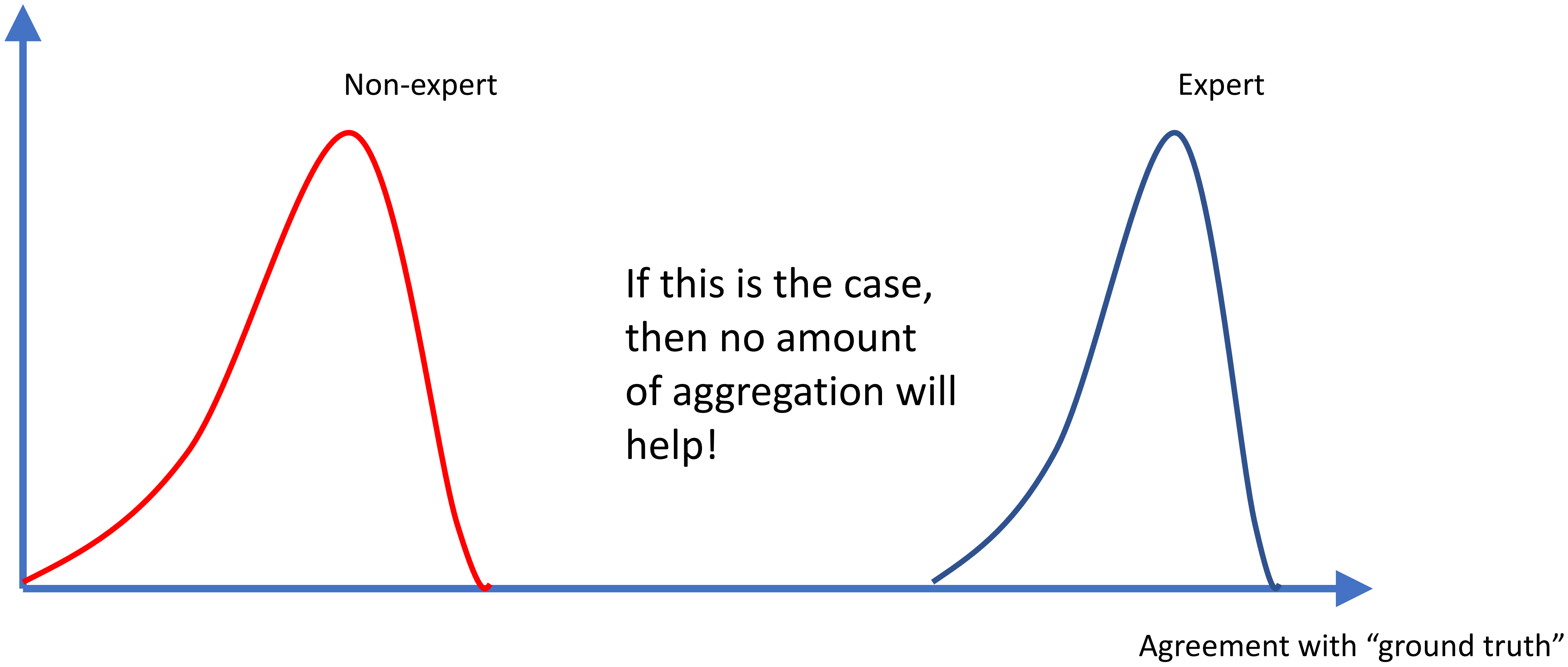

TASK
TUTORIAL
Done


Etch a Cell
ART AND SCIENCE



THE FRANCIS CRICK INSTITUTE

Million dollar question...

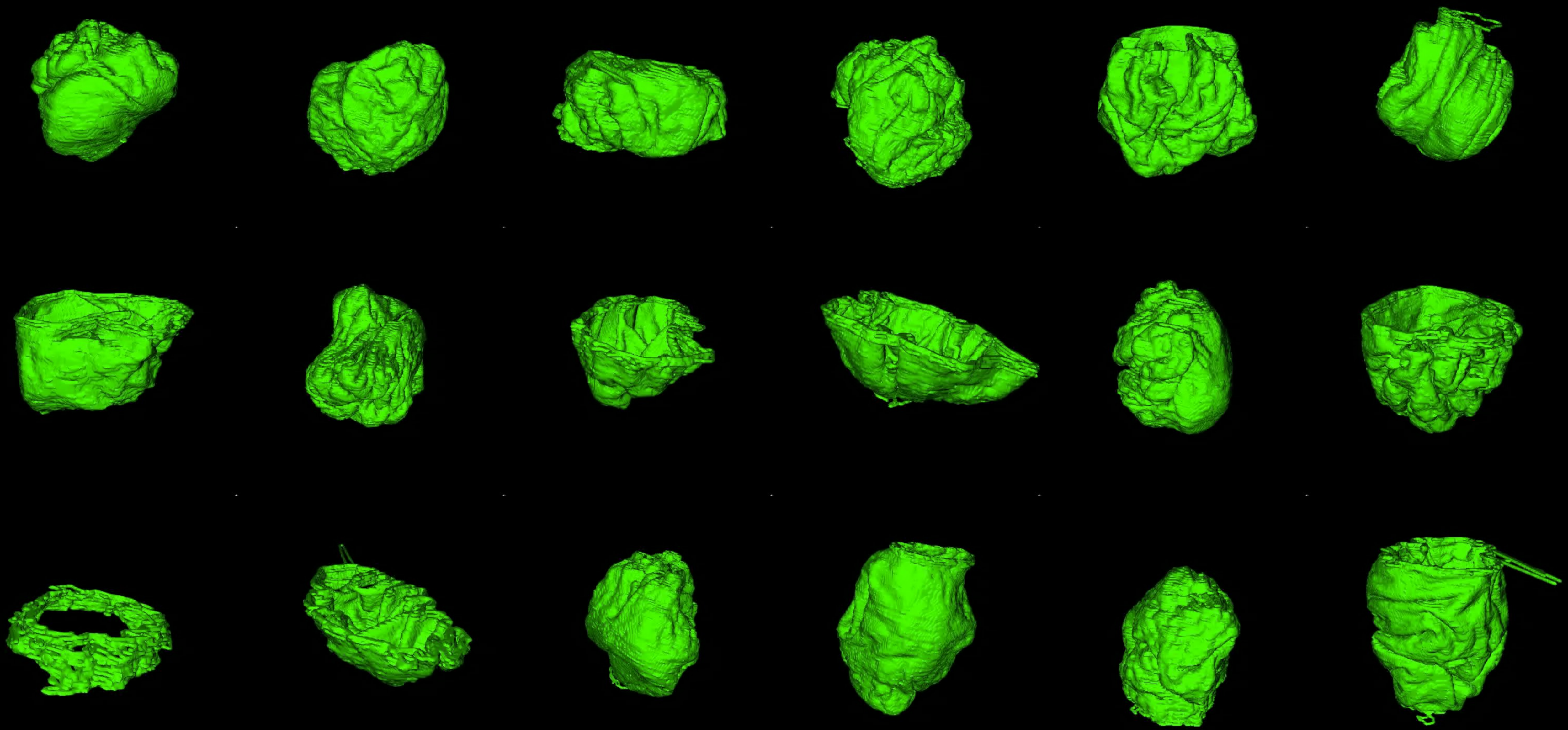


If this is the case, then no amount of aggregation will help!

Million dollar question…

Aggregation

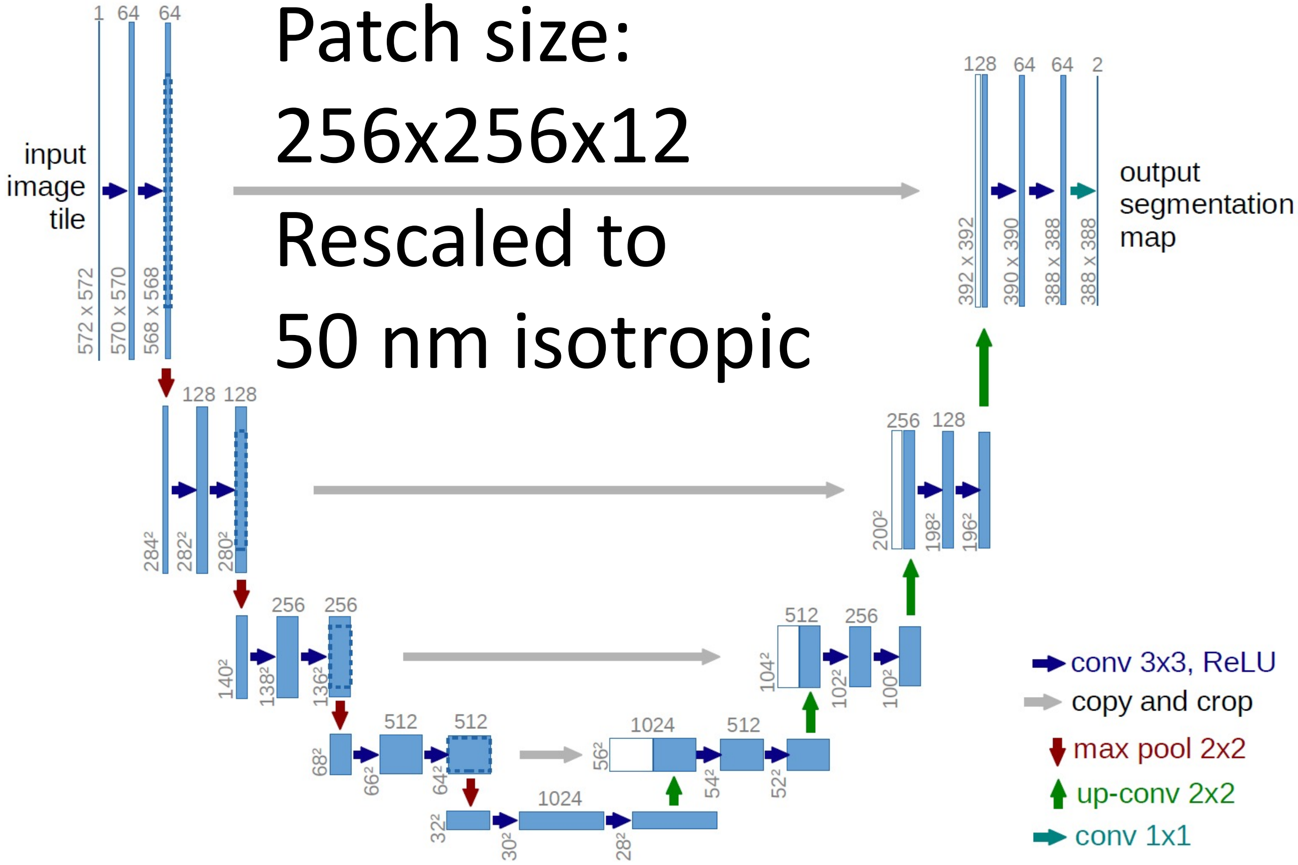
THE FRANCIS CRICK INSTITUTE



THE FRANCIS CRICK INSTITUTE

Convolutional Neural Network (U-Net)

Patch size:
256x256x12
Rescaled to
50 nm isotropic



Falk et al. Nature Methods 2019

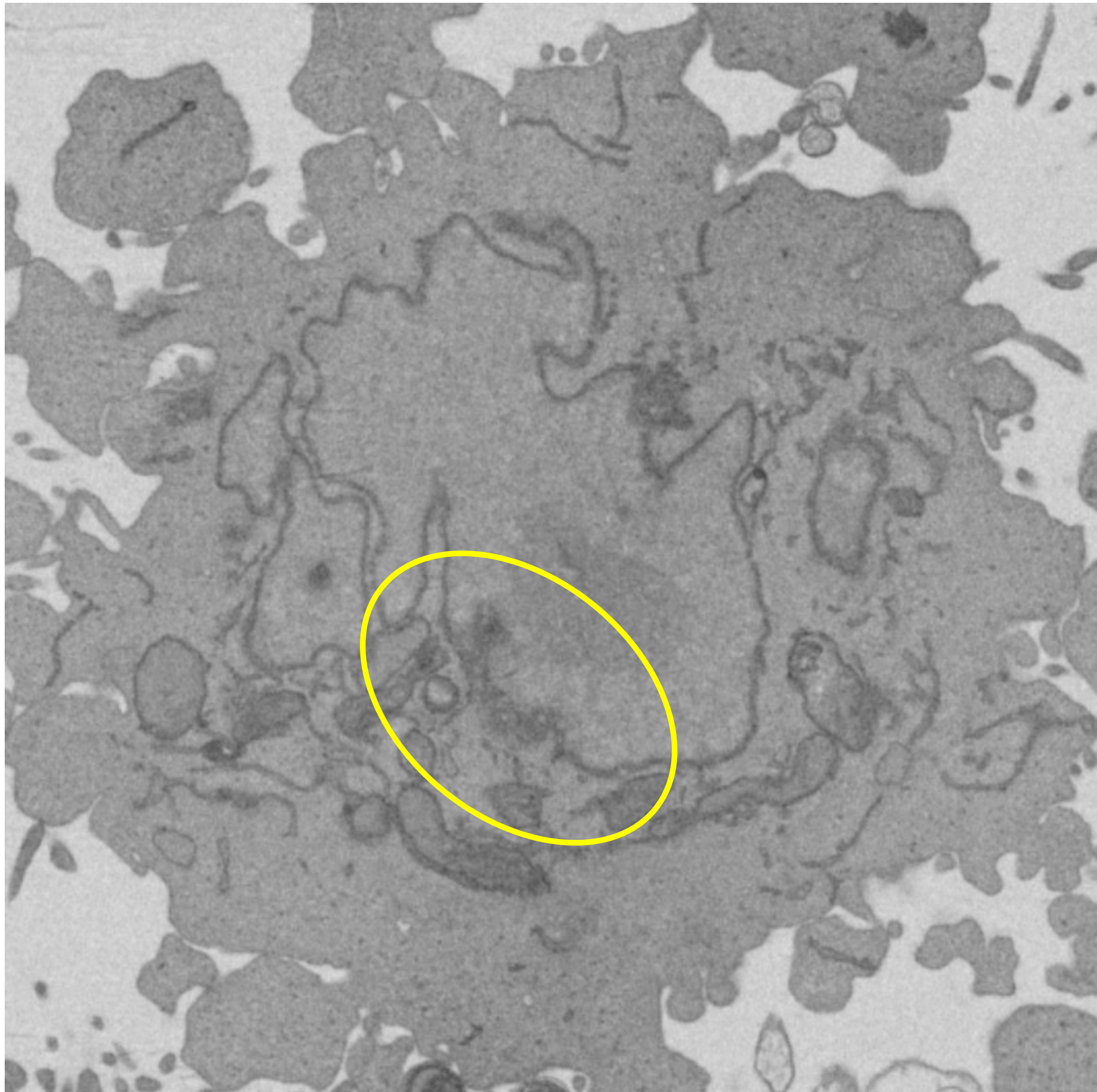
Harry Songhurst
Luke Nightingale
Joost de Folter

Membrane prediction

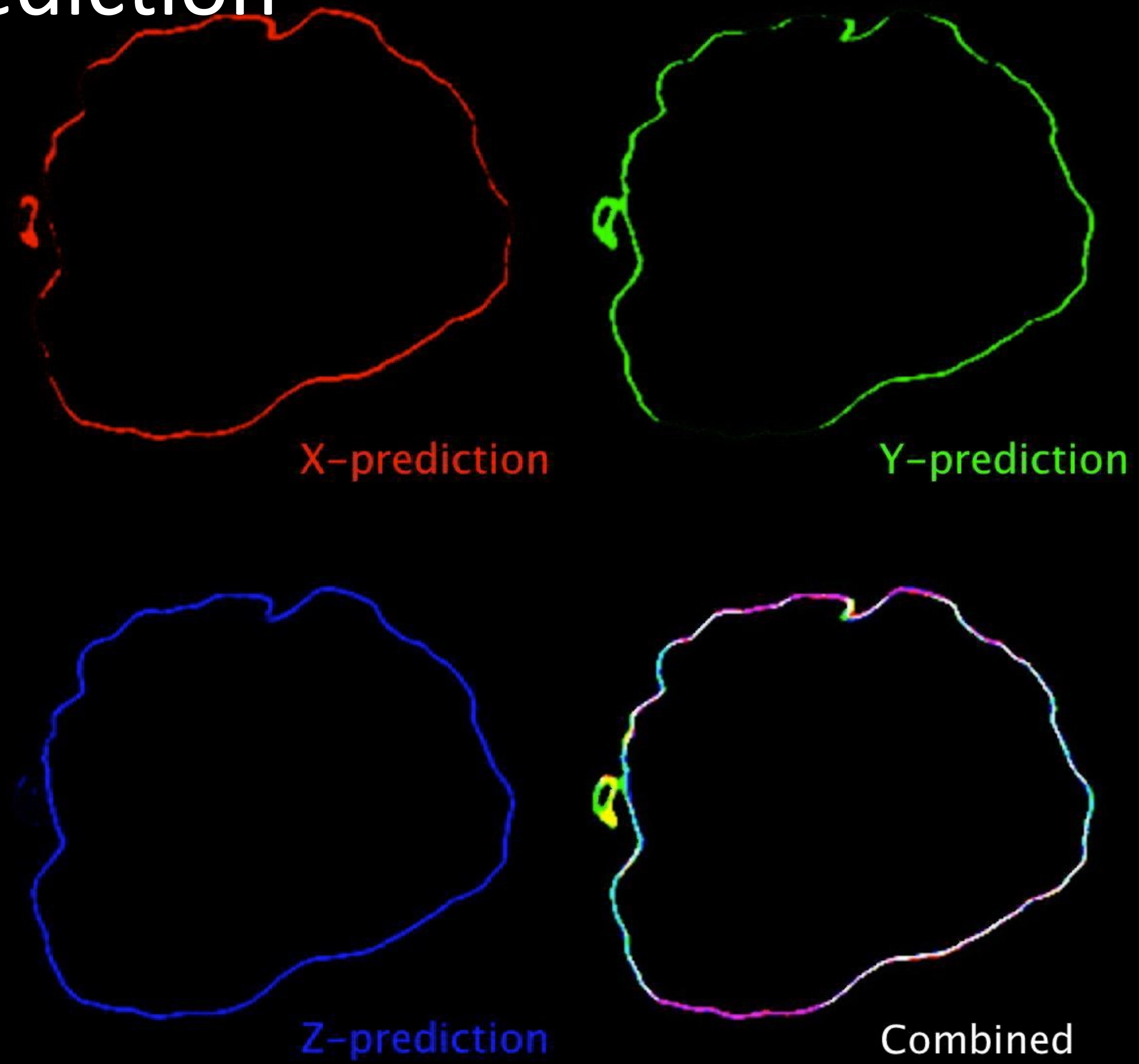
- Membranes oriented parallel to cutting/imaging plane appear as diffuse blurred lines

- Easily missed with typical workflows

- By predicting over 3 orthogonal axes, errors can be minimised



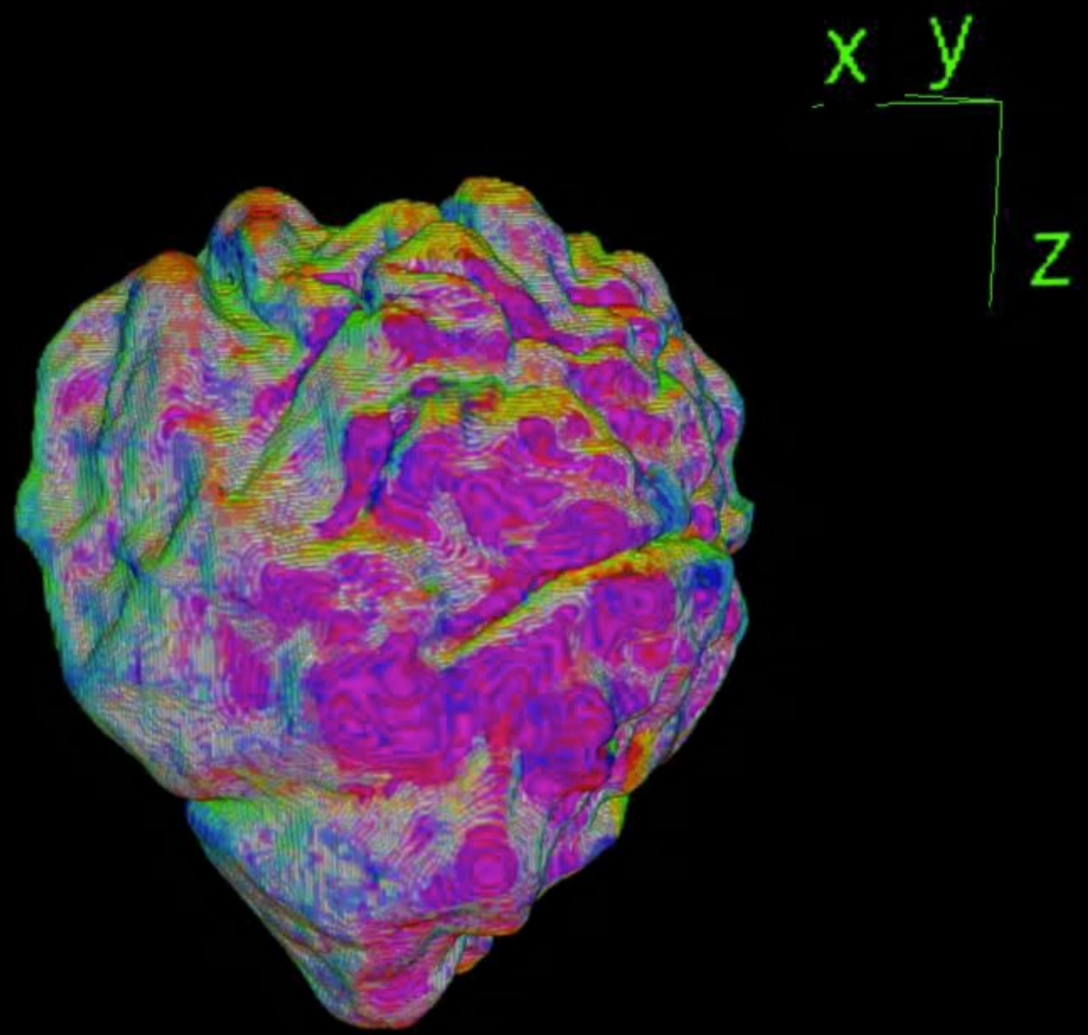
Tri Axis Prediction



THE FRANCIS CRICK INSTITUTE

Tri Axis Prediction

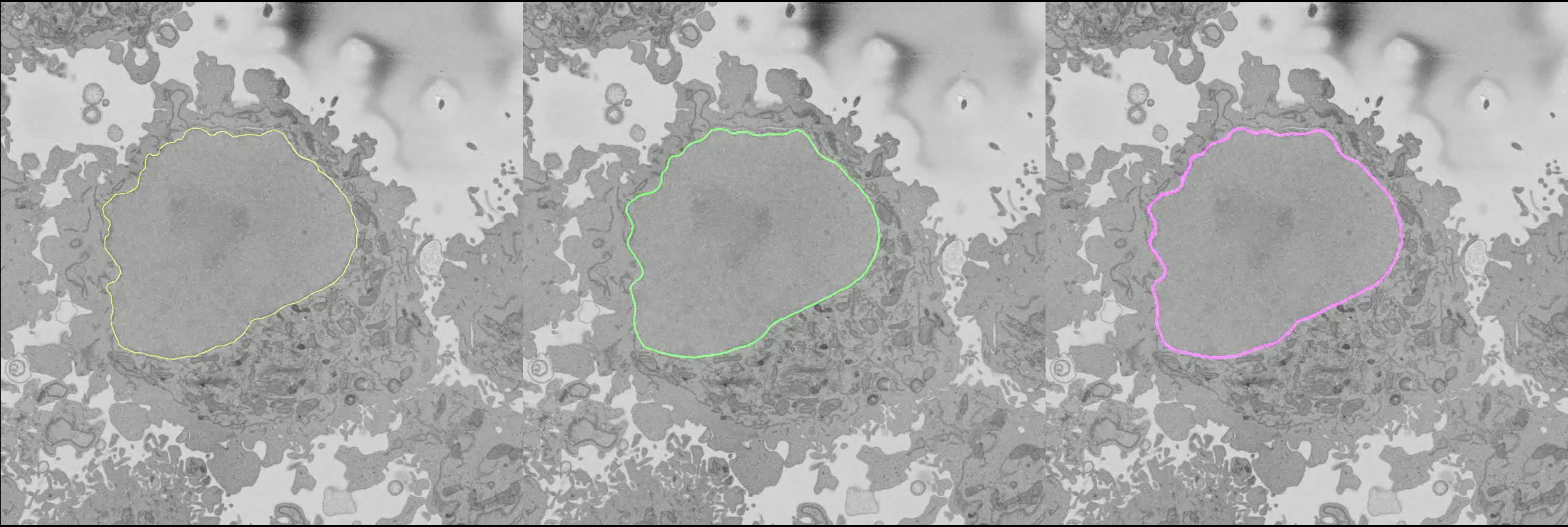
THE FRANCIS CRICK INSTITUTE



Expert

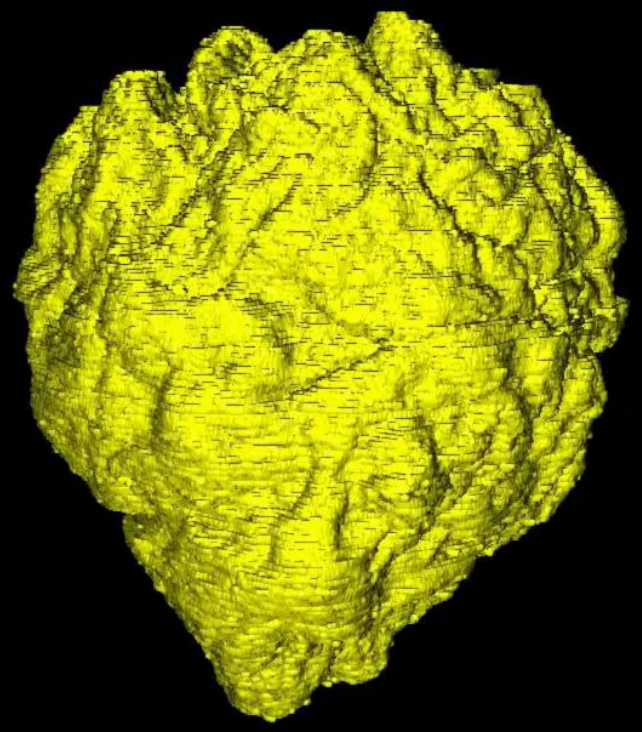
Aggregated
Cit Sci

U-Net

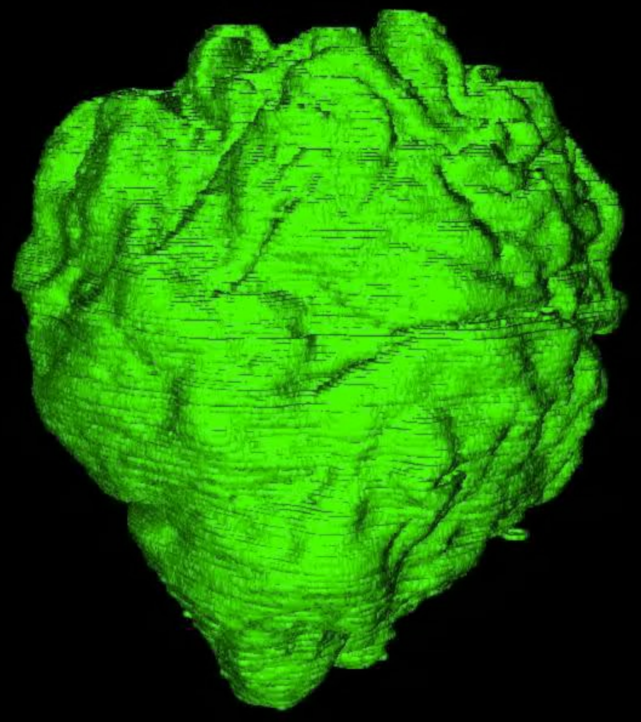


THE FRANCIS CRICK INSTITUTE

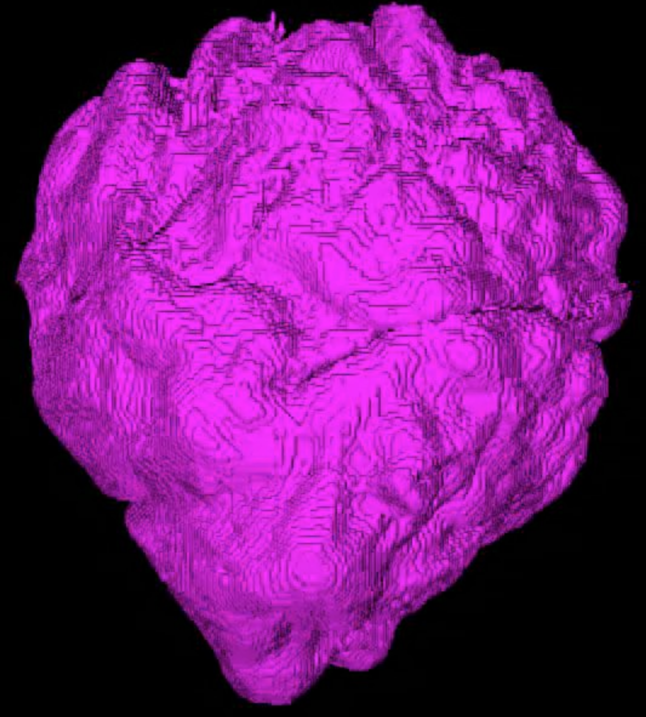
Expert

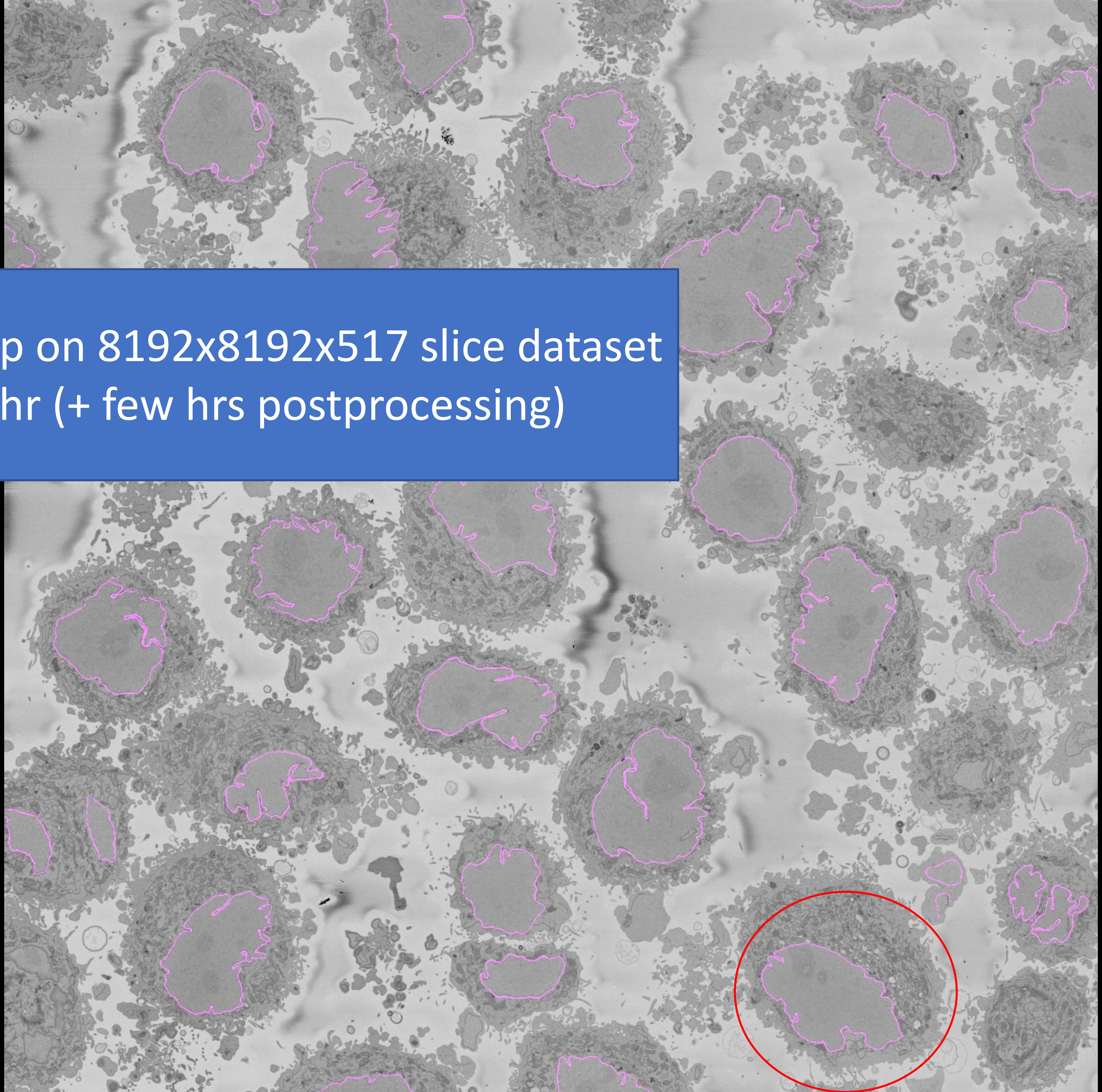


Aggregated Cit Sci



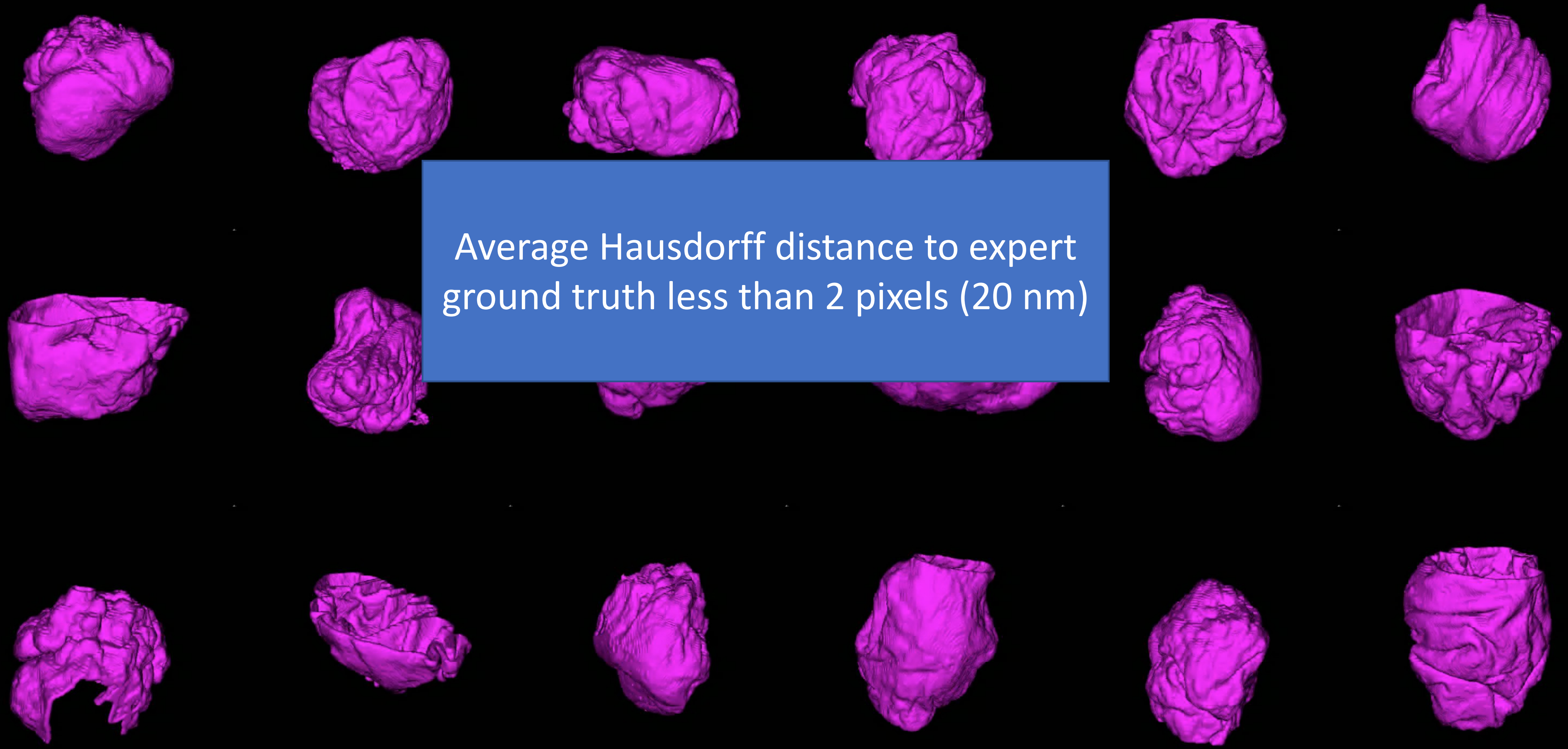
U-Net





Prediction step on 8192x8192x517 slice dataset takes < 1hr (+ few hrs postprocessing)

THE FRANCIS CRICK INSTITUTE


Average Hausdorff distance to expert ground truth less than 2 pixels (20 nm)

THE FRANCIS CRICK INSTITUTE

All data and code available openly

Traffic
The moving front of cell biology

TOOLBOX | Open Access

Deep learning for automatic segmentation of the nuclear envelope in electron microscopy data, trained with volunteer segmentations

Helen Spiers, Harry Songhurst, Luke Nightingale, Joost de Folter, The Zooniverse Volunteer Community, Roger Hutchings, Christopher J. Peddie, Anne Weston, Amy Strange, Steve Hindmarsh, Chris Lintott, Lucy M. Collinson, Martin L. Jones ... See fewer authors

First published: 29 April 2021 | https://doi.org/10.1111/tra.12789

The Zooniverse Volunteer Community: This publication has been made possible by the participation of volunteers in the Etch A Cell project. Their contributions are acknowledged at https://www.zooniverse.org/projects/h-spiers/etch-a-cell/about/results.
Funding information: Cancer Research UK, Grant/Award Number: FC001999; Medical Research Council: FC001999; Wellcome Trust: FC001999

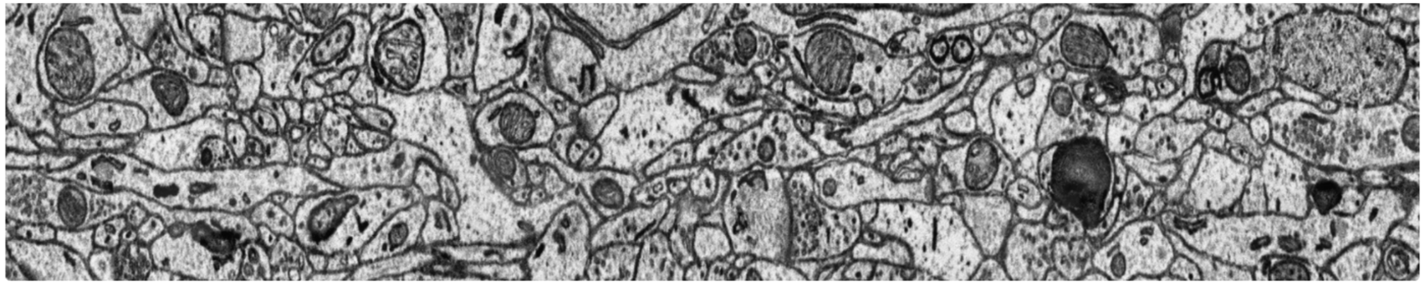
Reusing the model

- Network trained on nuclear envelope encodes a lot of features of EM data

- Can be repurposed to new problems by using as *initialization* of a training run on different data

- E.g. training nuclear envelope model to segment mitochondria

THE FRANCIS CRICK INSTITUTE



Info | Forum | Leaderboard | Join

Challenge

Evaluation

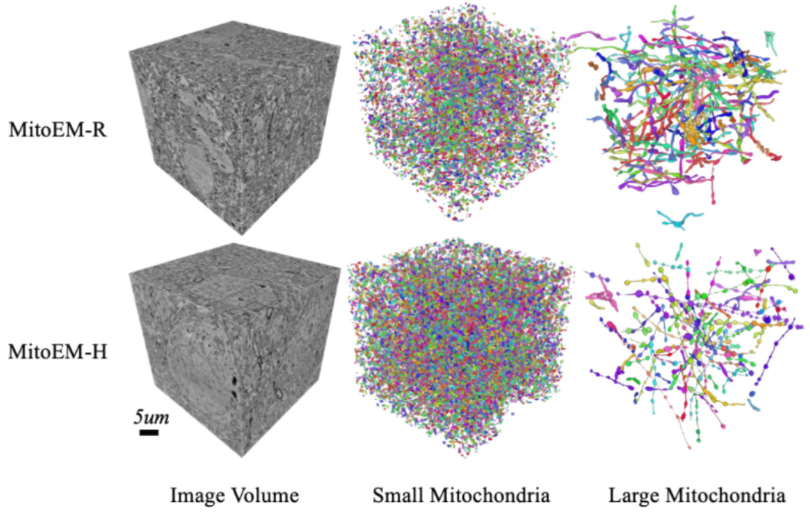
MitoEM Challenge: Large-scale 3D Mitochondria Instance Segmentation

This challenge is part of ISBI 2021 taking place on April 13-16th 2021. The top teams will be invited for writing an overview paper after the challenge. More information about the MitoEM dataset can be found in our MICCAI 2020 paper.

Note: we keep the online evaluation open for new submissions, but only results submitted before March 1, 2021, are considered for the ISBI challenge workshop.

Explore it in your browser! [MitoEM-H], [MitoEM-R]

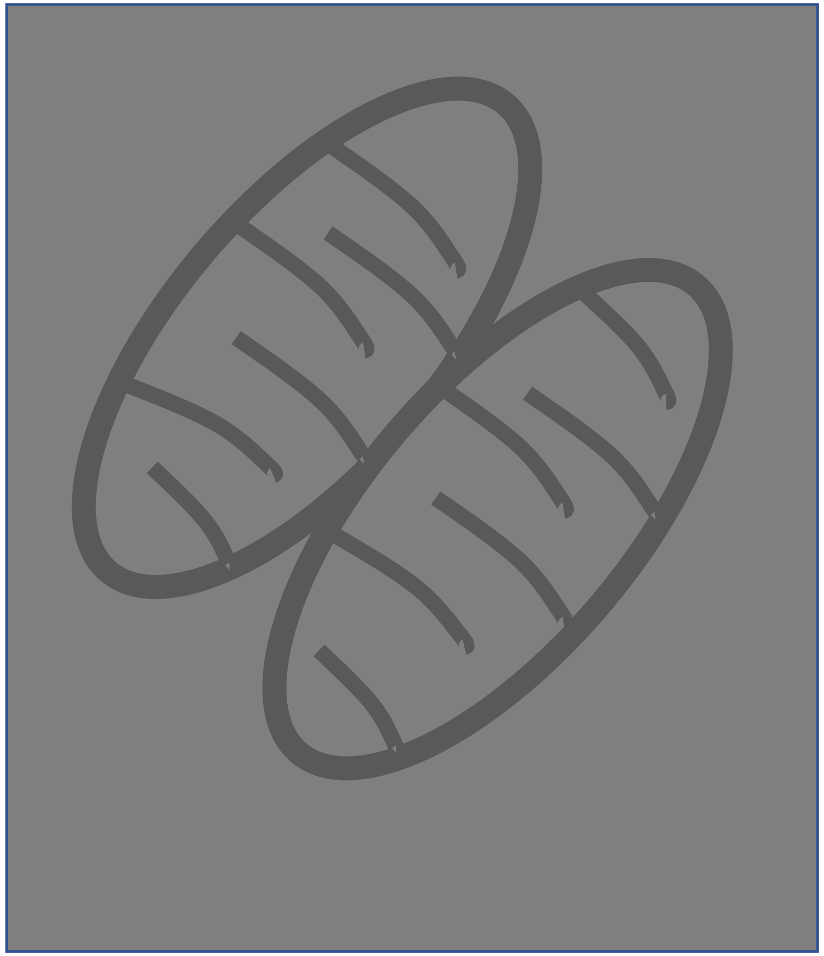
(navigation tips: [manual] [youtube video])



Instance segmentation

THE FRANCIS CRICK INSTITUTE

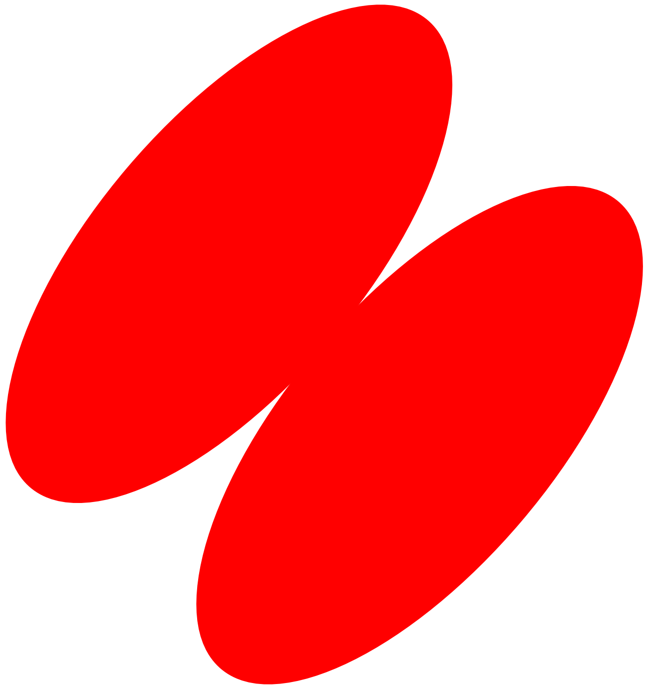
- Each mitochondrion given a unique label
- Area + boundary predictions combined with watershed algorithm to create instance segmentation



Instance segmentation

- Each mitochondrion given a unique label
- Area + boundary predictions combined with watershed algorithm to create instance segmentation

Area (semantic)



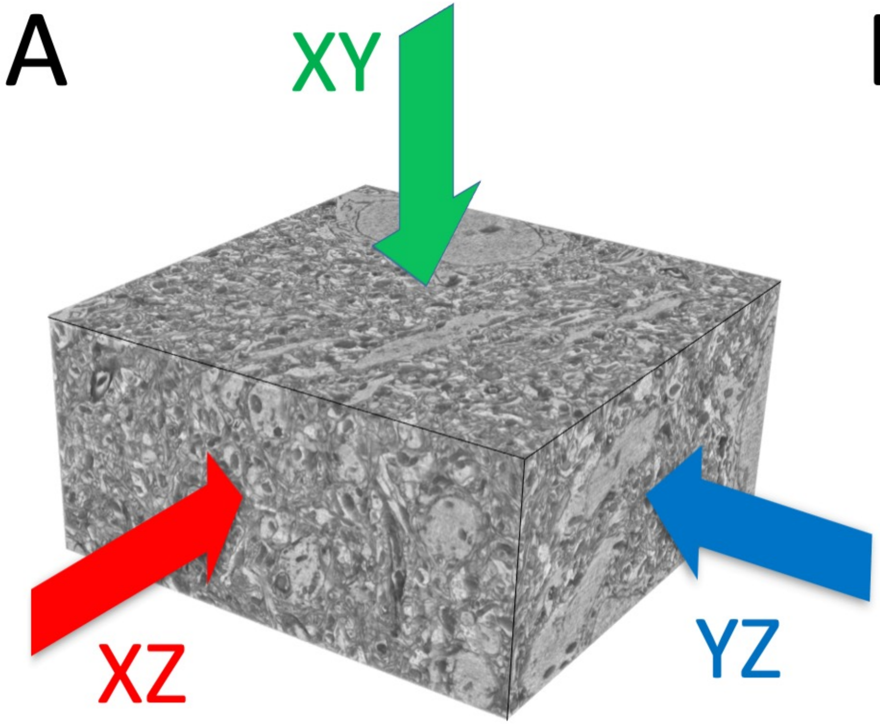
Instance segmentation

- Each mitochondrion given a unique label

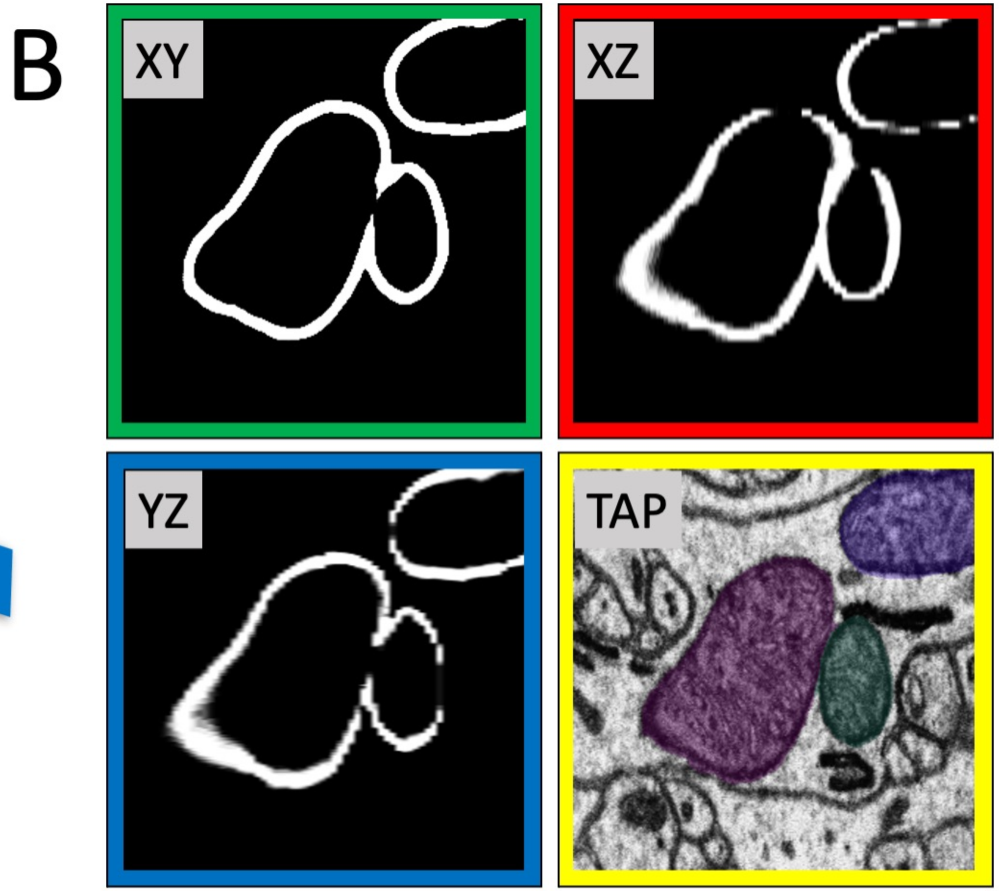
- Area + boundary predictions combined with watershed algorithm to create instance segmentation

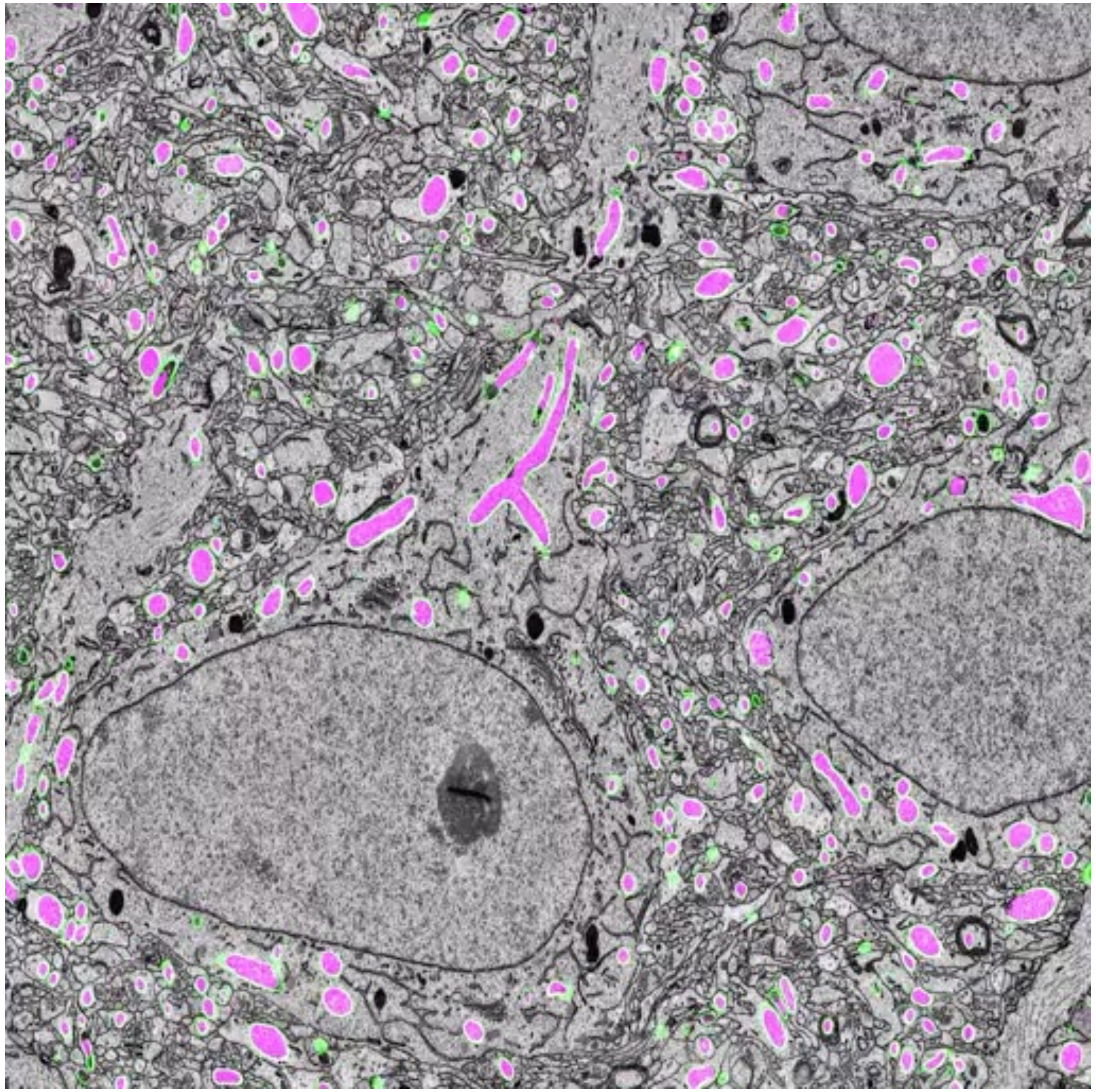
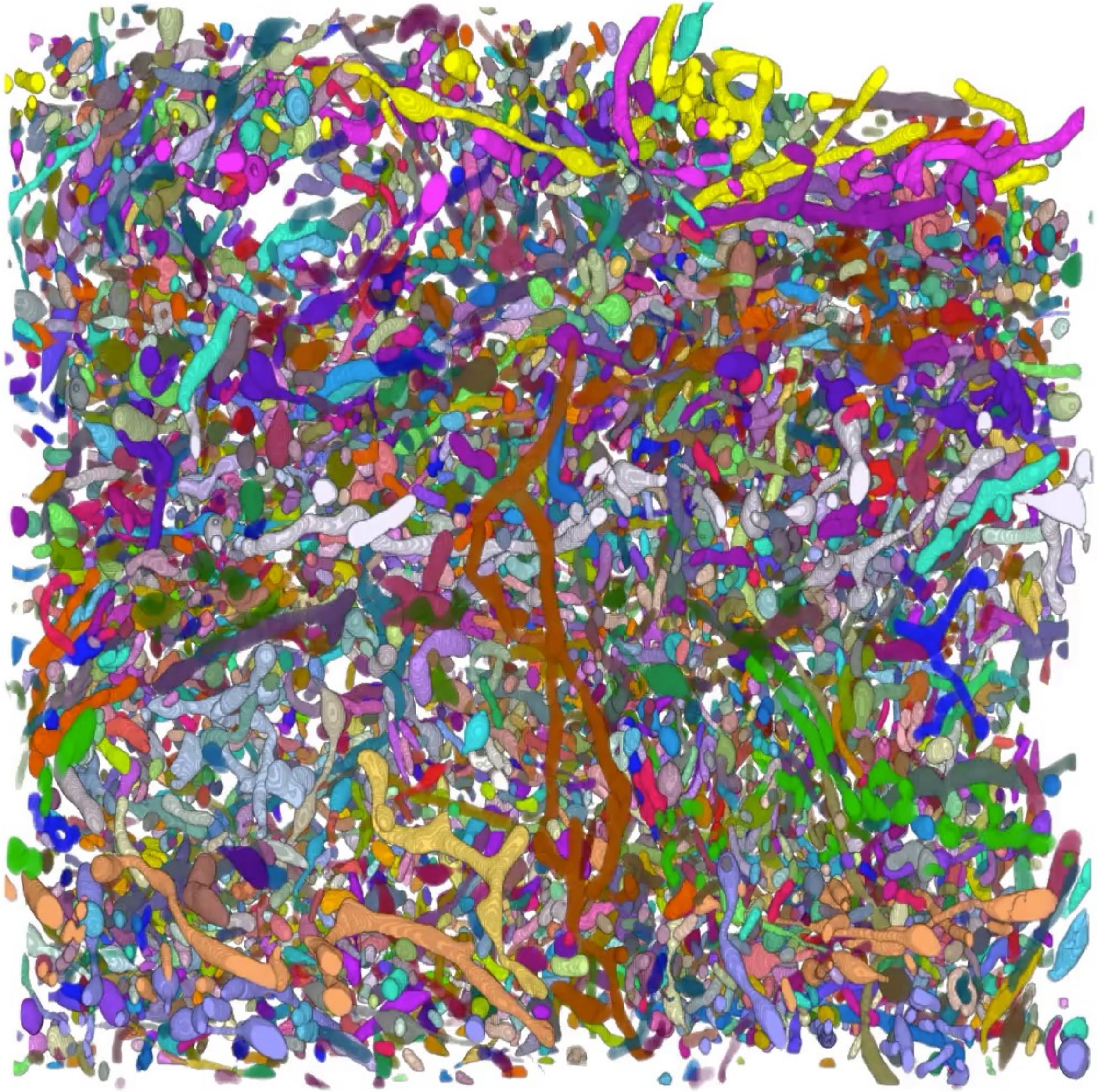
Tri-axis prediction for non-isotropic data

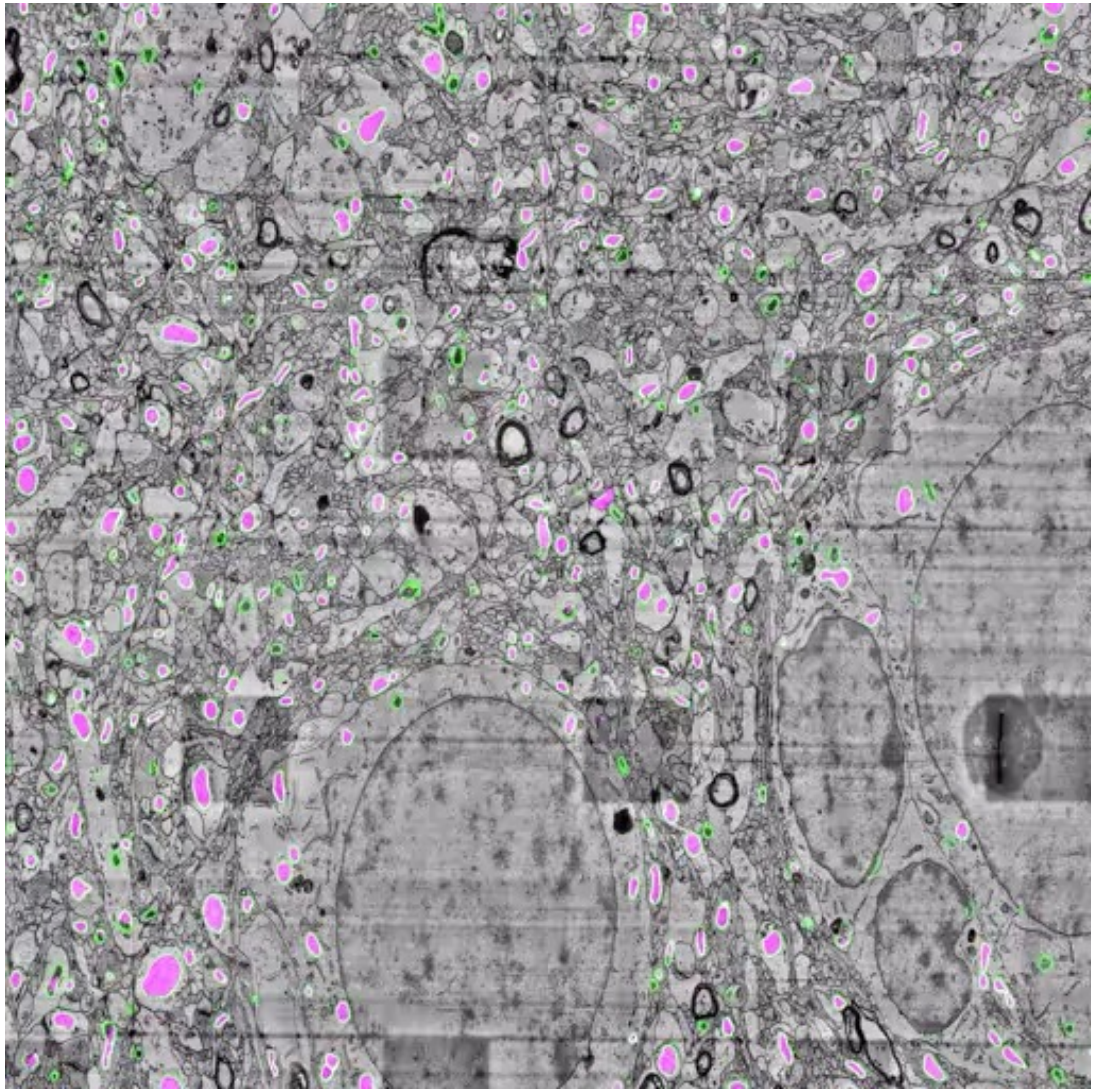
THE FRANCIS CRICK INSTITUTE

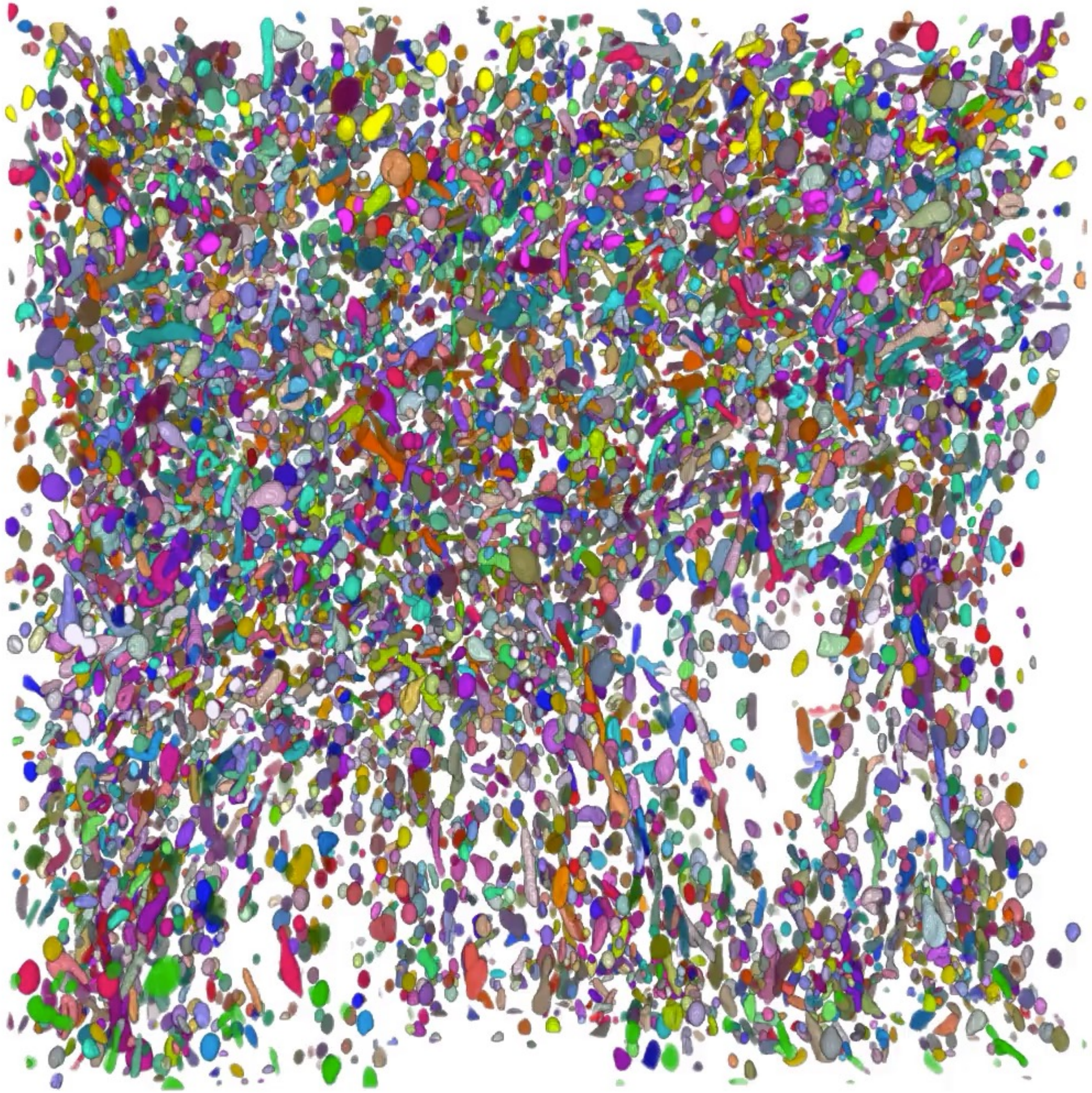


Raw data 8x8x30 nm voxels







bioRxiv preprint doi: https://doi.org/10.1101/2021.05.24.444785; this version posted May 24, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

All data and code available openly

AUTOMATIC INSTANCE SEGMENTATION OF MITOCHONDRIA IN ELECTRON MICROSCOPY DATA

Luke Nightingale[1], *Joost de Folter*[1], *Helen Spiers*[1,2], *Amy Strange*[1], *Lucy M Collinson*[1], *Martin L Jones*[1]

[1] The Francis Crick Institute, London, NW1 1AT, United Kingdom
[2] University of Oxford, Oxford, OX1 3RH, United Kingdom

ABSTRACT

We present a new method for rapid, automated, large-scale 3D mitochondria instance segmentation, developed in response to the ISBI 2021 MitoEM Challenge. In brief, we trained separate machine learning algorithms to predict (1) mitochondria areas and (2) mitochondria boundaries in image volumes acquired from both rat and human cortex with multi-beam scanning electron microscopy. The predictions from these algorithms were combined in a multi-step post-processing procedure, that resulted in high semantic and instance segmentation performance. All code is provided via a public repository.

Index Terms— Electron microscopy, Mitochondria, Instance Segmentation, Deep Learning, 3D

1. INTRODUCTION

Electron microscopy has undergone a radical advancement in data generation capacity through the development of a series of automated systems that are able to rapidly collect thousands of serial images through whole cell and tissues. This has displaced the research bottleneck in this domain to data analysis – one of the greatest challenges faced by modern imaging is how to efficiently and quickly extract meaningful insight from the incredibly rich data produced.

However, the development of novel analytical pipelines is often hampered by a scarcity of relevant, publicly available, well labelled datasets. In response, the MitoEM dataset

U-Nets (Sections 2.4, 2.6, 2.7) to predict mitochondria areas and boundaries (Section 2.8) in both the rat and human data. The predictions from these algorithms were combined in a multi-step post-processing procedure (Section 2.9), that resulted in high semantic and instance segmentation performance (Table 1). An overview of the approach used is presented in Figure 1.

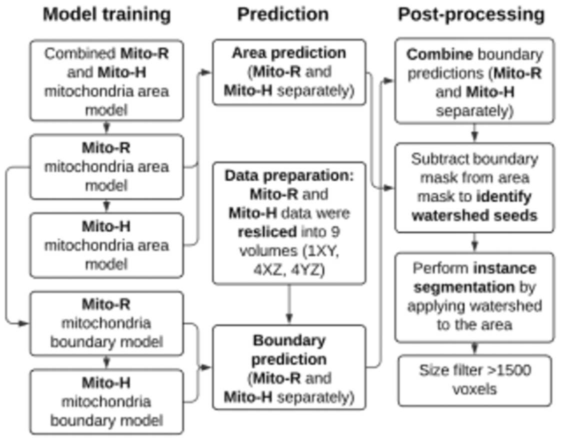
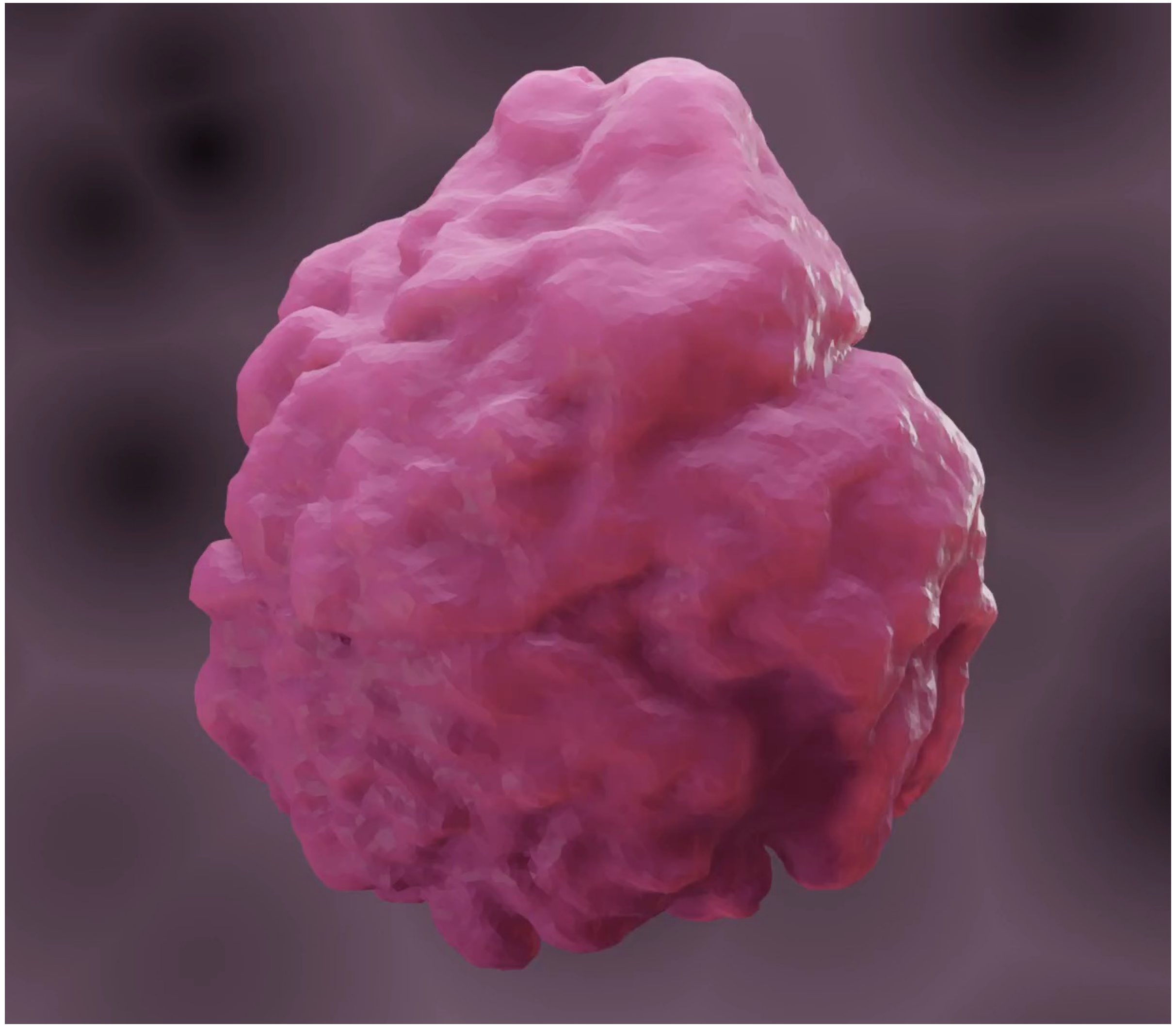


Fig. 1. Analytical pipeline: model training, prediction and post-processing

Next steps and Outlook

THE FRANCIS CRICK INSTITUTE

- Convert slice-by-slice segmentations to meshes & optimise
- Perform measurements on multiple structures
- Automate pipeline to produce statistically significant studies *(no more N=1)*
- Simulation!



Electron Microscopy STP at the Crick

www.crick.ac.uk/research/platforms-and-facilities/electron-microscopy

THE FRANCIS CRICK INSTITUTE

Etch a Cell: bit.ly/etchiverse

Labs: bit.ly/emstp_labs

Lucy Collinson
Raffa Carzaniga
Chris Peddie
Lizzy Brama
Alana Burrell
MC Domart
Wei Guan
Daniel Krentzel
Catherine Maclachlan
Adam Pestana Motala
Matt Russell
Aaron Sait
Harry Songhurst
Anne Weston
Azumi Yoshimura
Yuxin Zhang

Machine learning team
Amy Strange
Joost de Folter
Luke Nightingale

Zooniverse: Helen Spiers
Chris Lintott

MRC Medical Research Council
UCL
Imperial College London
King's College London
wellcome
CANCER RESEARCH UK