Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

forward together
sonke siya phambili
saam vorentoe

Image Processing

From Virtual Reality to Deep Learning

Dr Rensu Theart

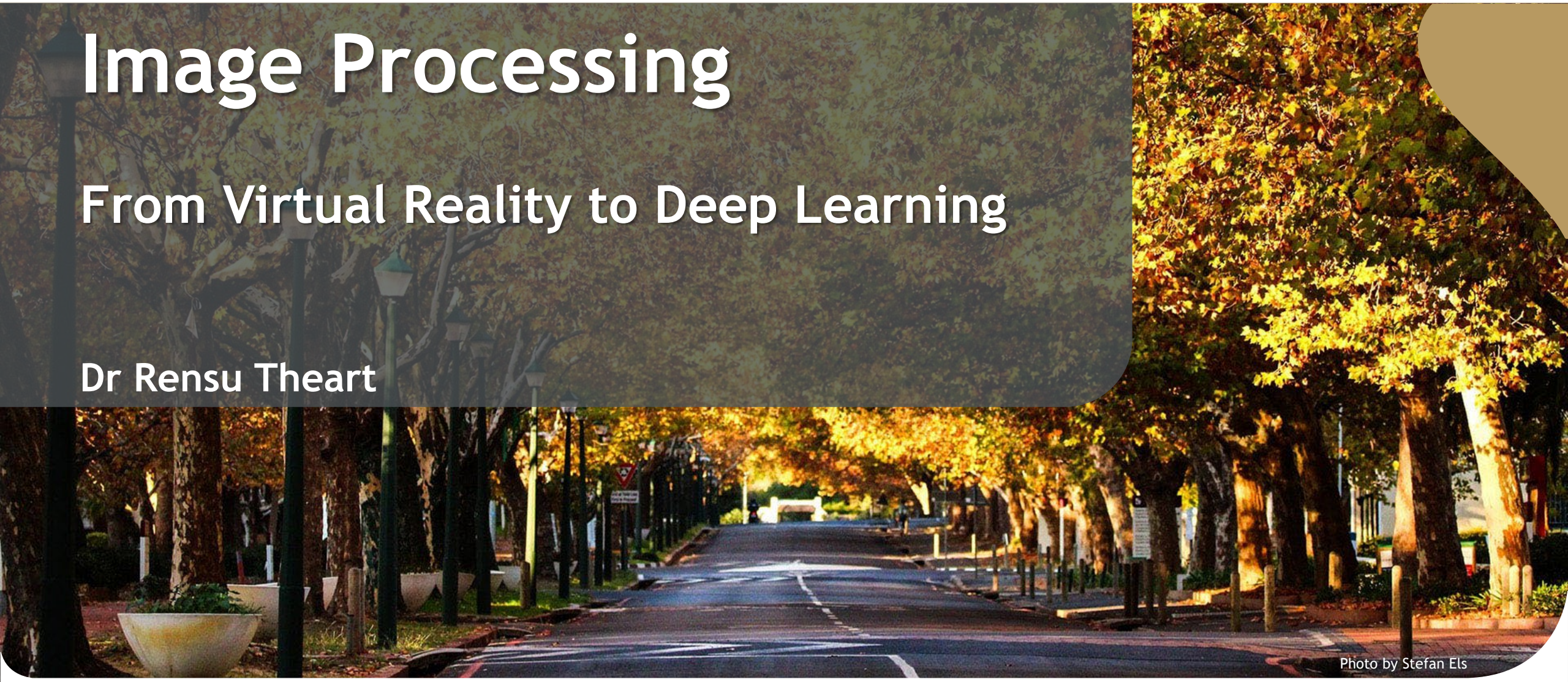
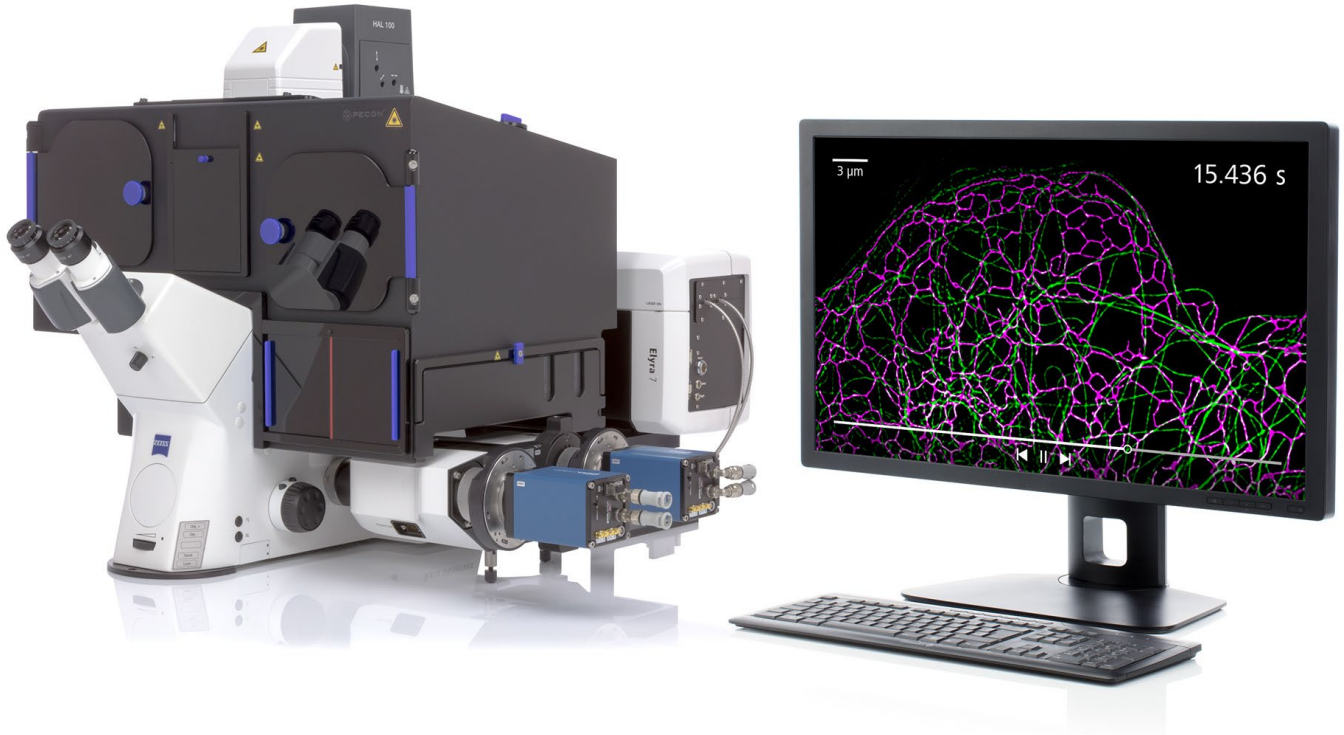


Photo by Stefan Els

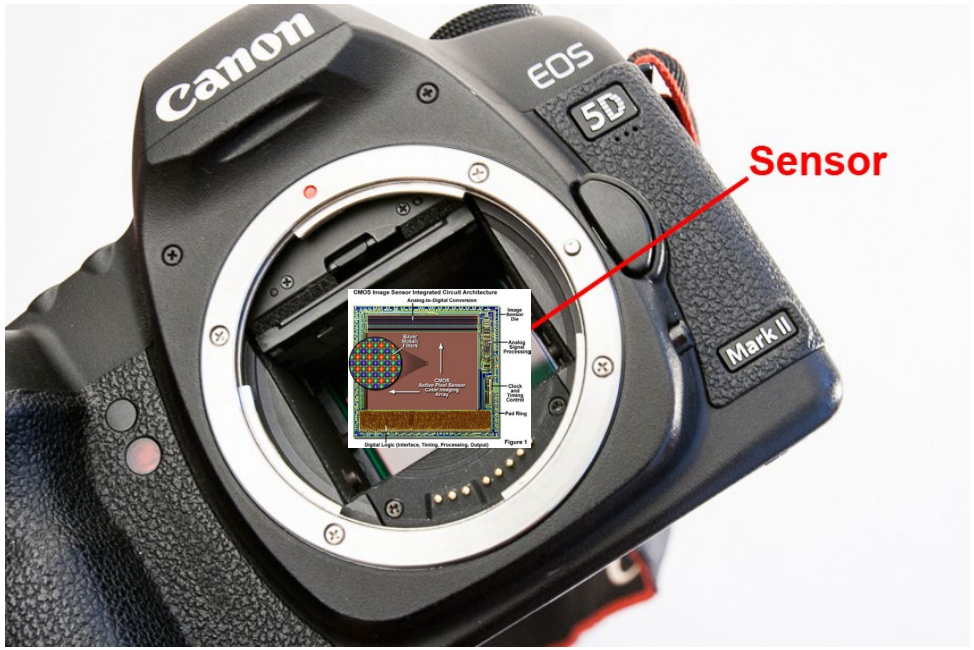
Image Processing

- Is the use of a computer to process **digital images** through an **algorithm**.
- A camera is simply a light measuring device (CCD or CMOS sensor).



Images only contain three colour components

- Camera sensors are designed to capture RGB.



RGB Example

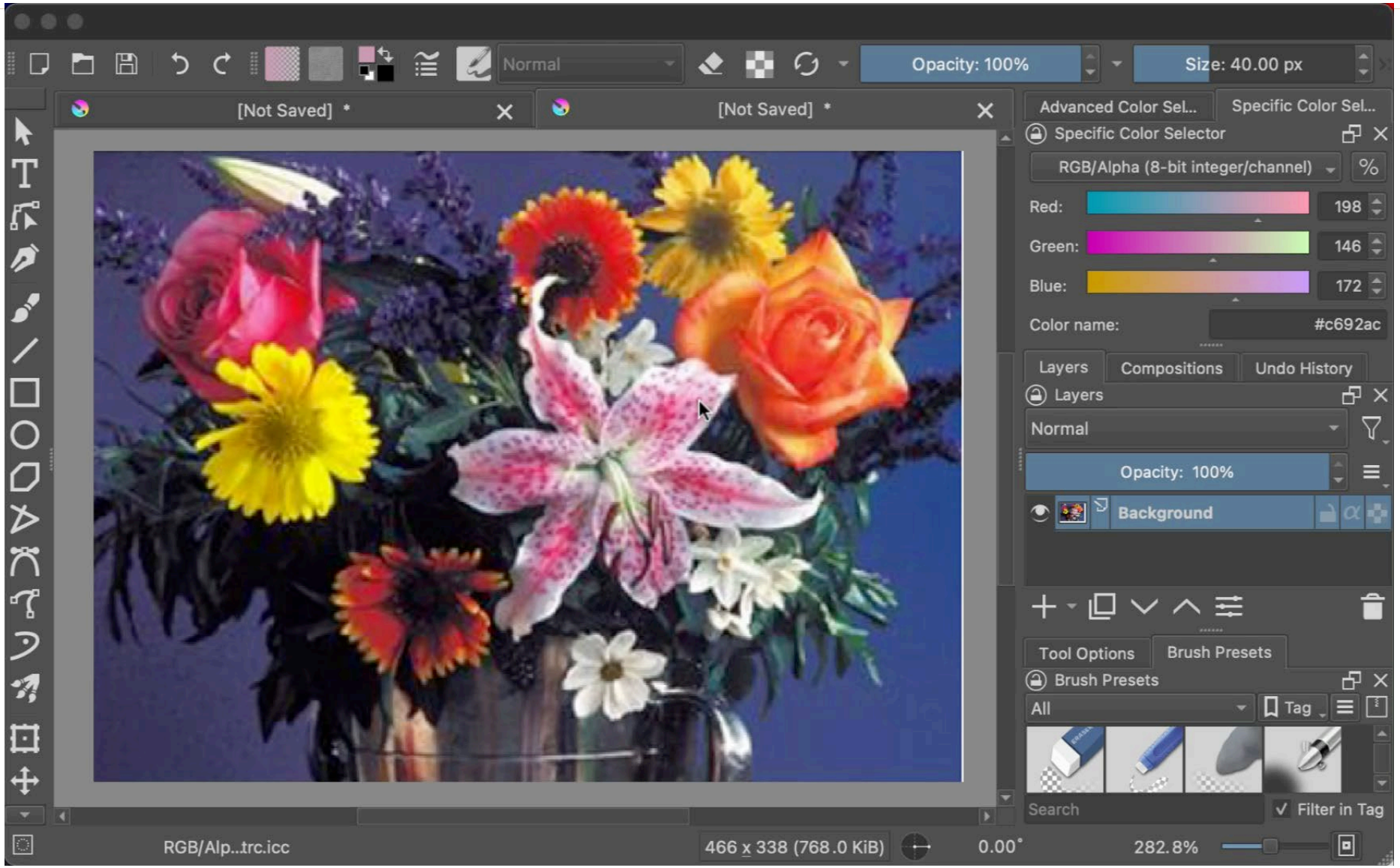
Combined (RGB)

Red Component

Green Component

Blue Component

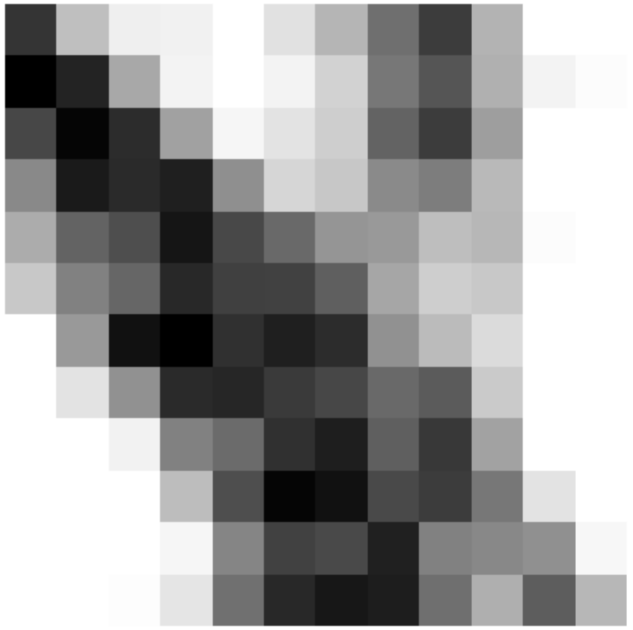
Pixels and RGB values



Pixels and RGB values

(A) Original image

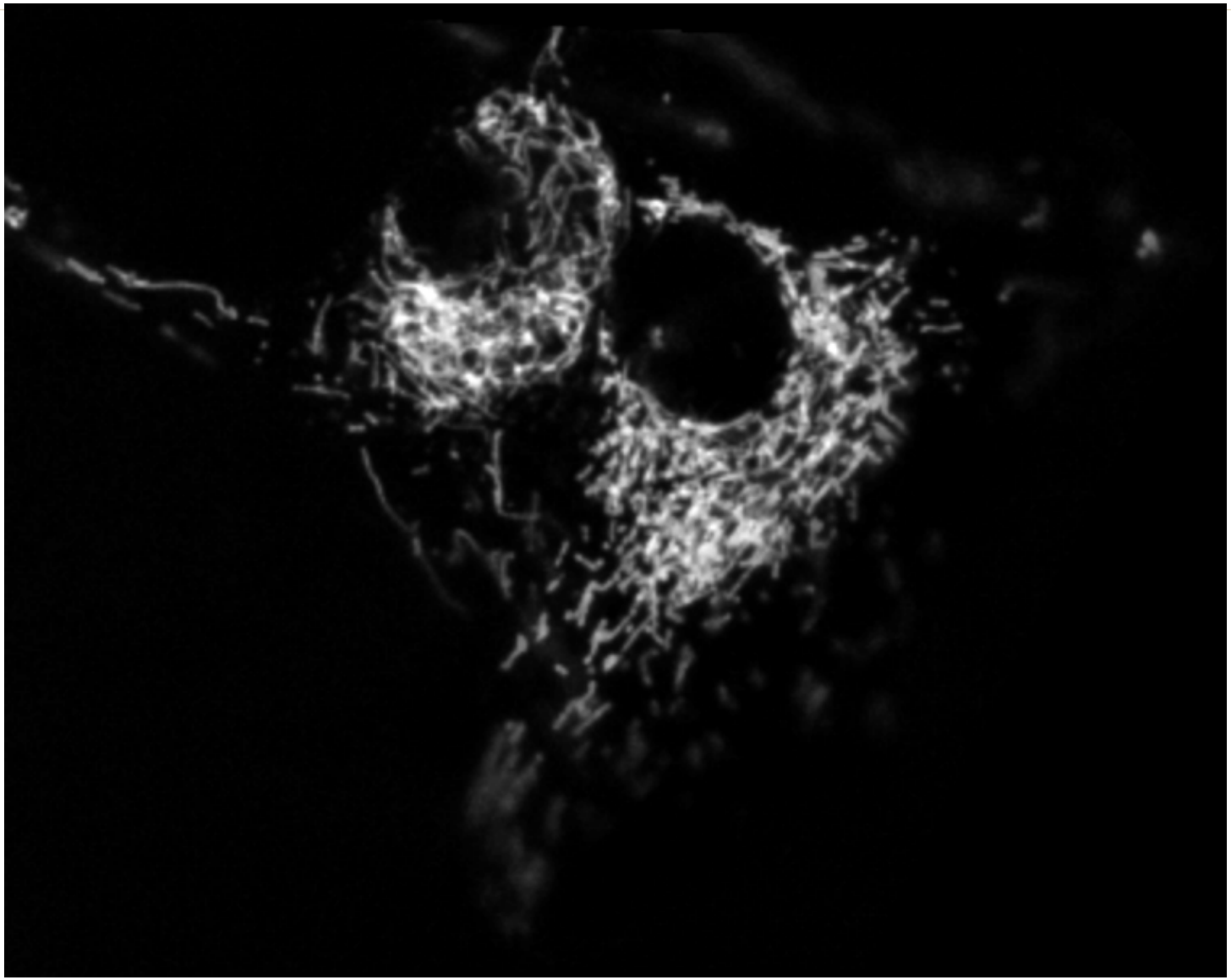
(B) Enlarged view from (A)



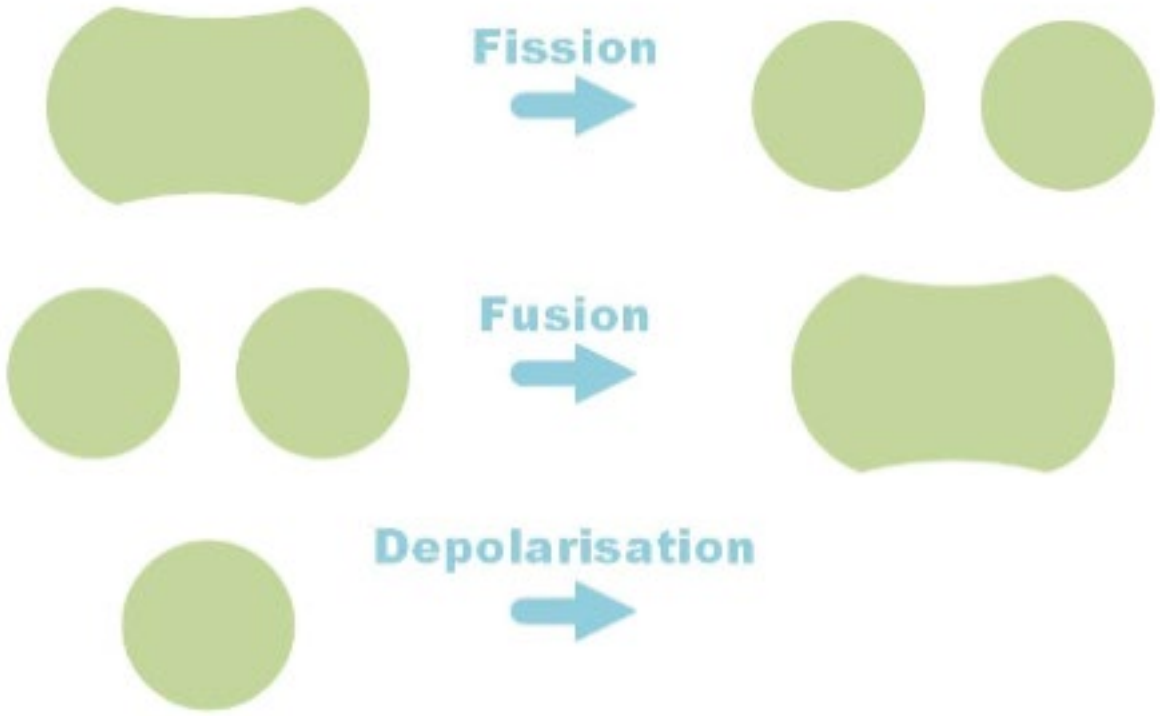
(C) Pixel values from (B)

53	191	239	241	255	225	181	111	61	180	255	255
	35	168	244	255	243	210	119	85	176	244	252
71		45	161	246	227	206	99	60	158	255	255
137	26	42	31	143	214	199	138	125	185	255	255
172	99	78	21	72	106	149	153	190	183	252	255
200	129	102	41	64	65	95	166	206	200	255	255
255	153	17		49	31	44	145	187	219	255	255
255	227	145	42	38	58	71	106	91	202	255	255
255	255	242	129	107	48	30	95	57	162	255	255
255	255	255	189	78		17	74	60	119	228	255
255	255	255	246	133	65	73	32	129	136	144	247
255	255	253	229	112	40	23	29	111	175	93	183

Mitochondria

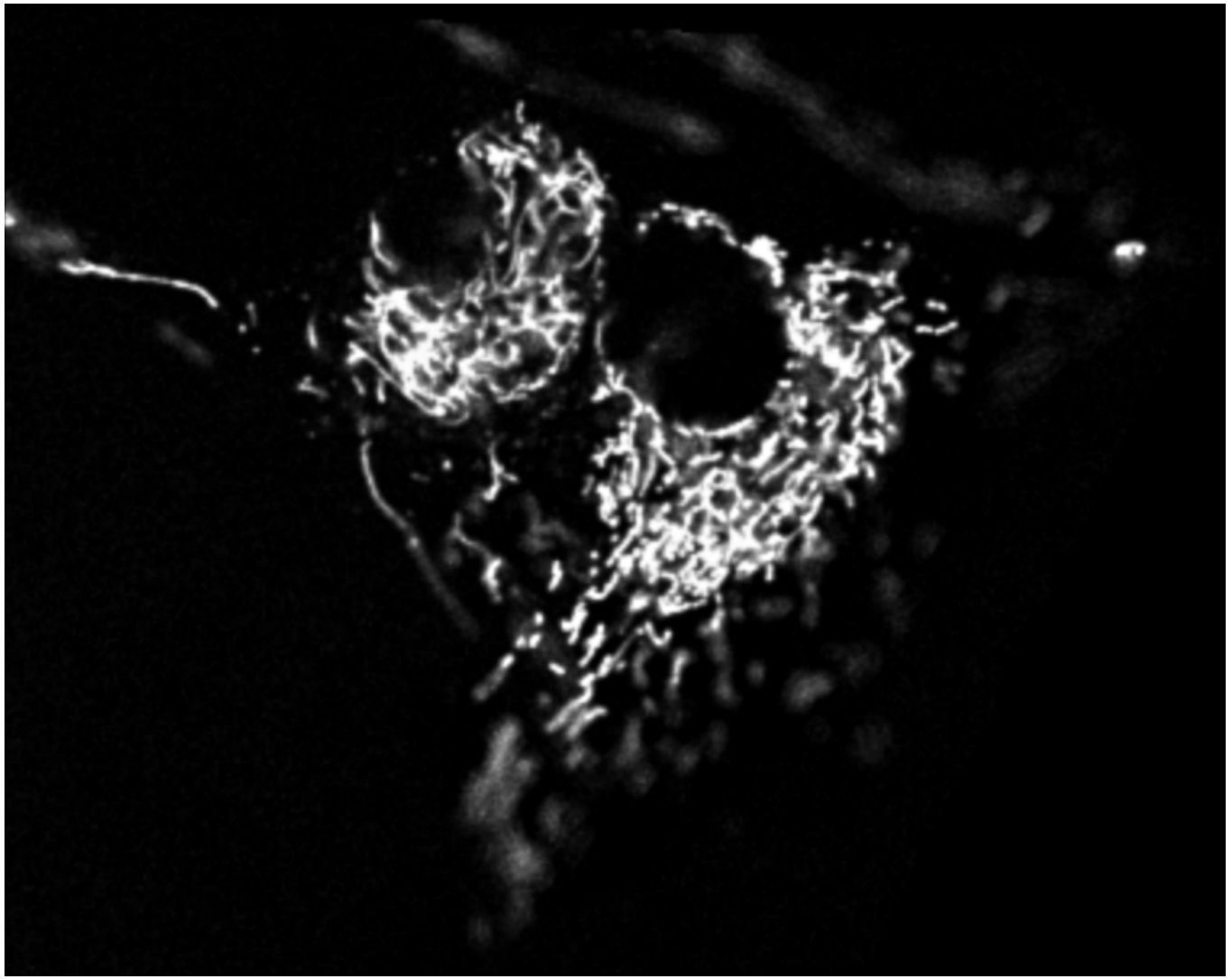


Mitochondrial events



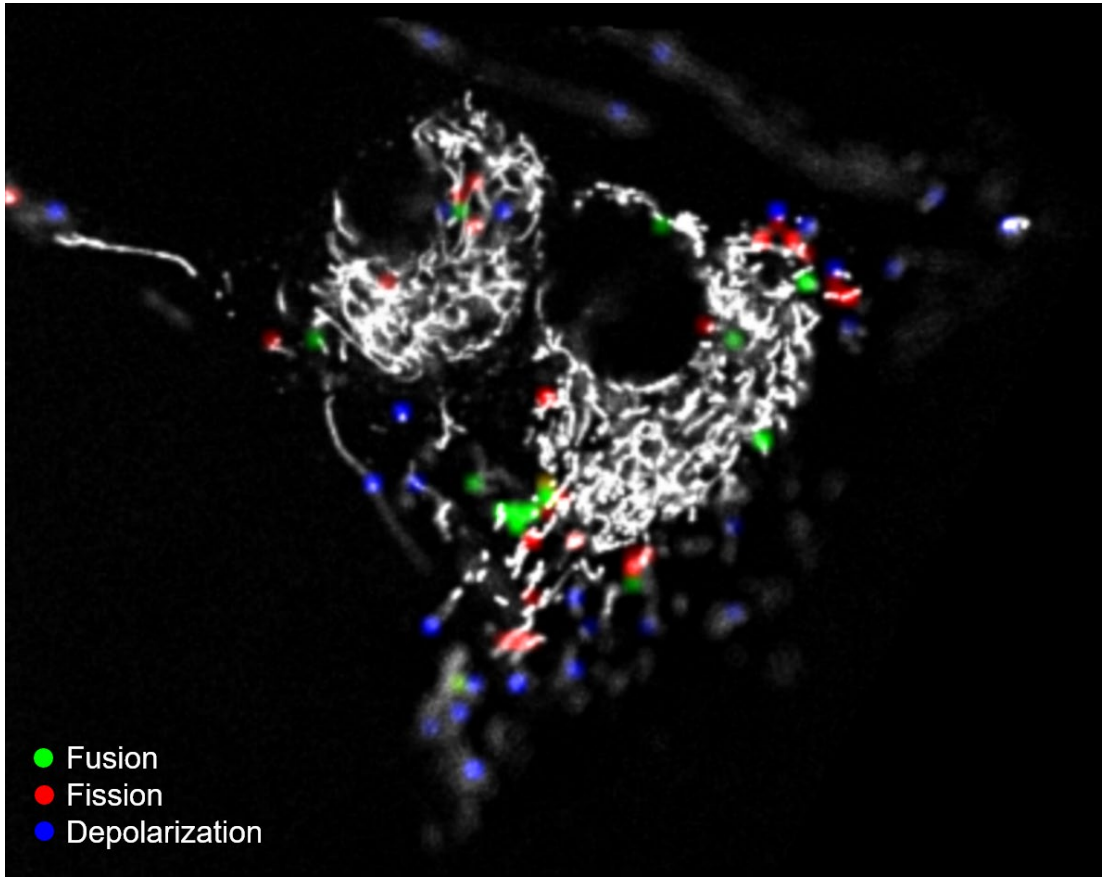
An illustration of the different types of mitochondrial events

What we wanted to achieve

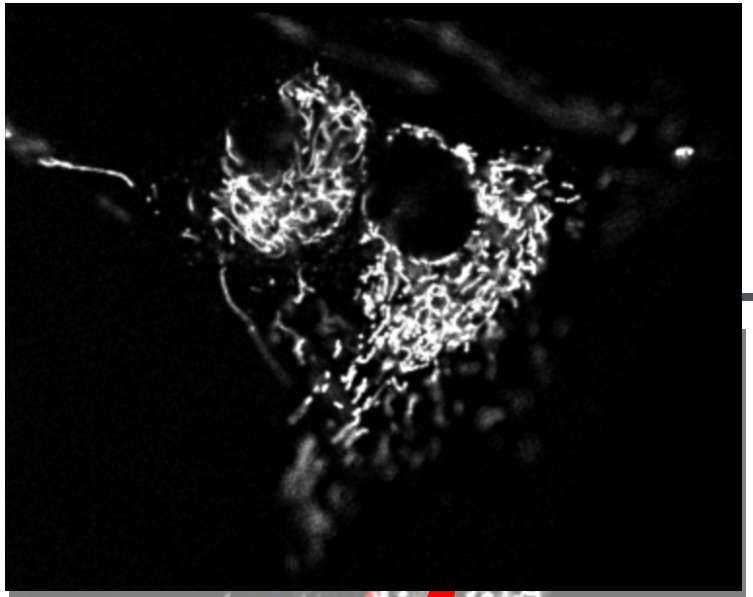


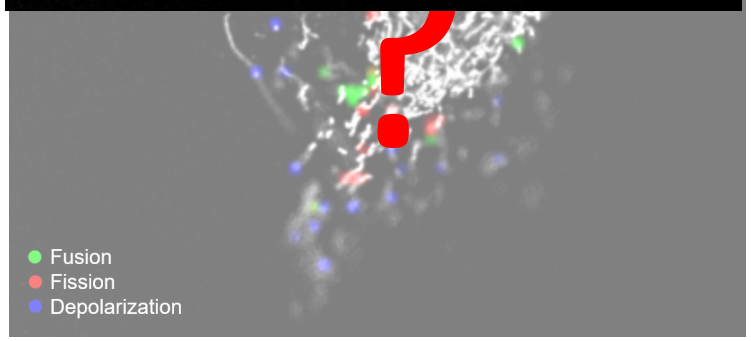
Why?

- Mitochondrial disorders are often presented as neurological disorders such as Alzheimer's disease.
- Fission/fusion events play a critical role in maintaining functional mitochondria when cells experience metabolic or environmental stresses.
- Having a system that can automatically predict the **number of mitochondrial events** as well as their **location** will help researchers gain insights mitochondrial function which in turn could lead to improved treatment of neurological disorders.



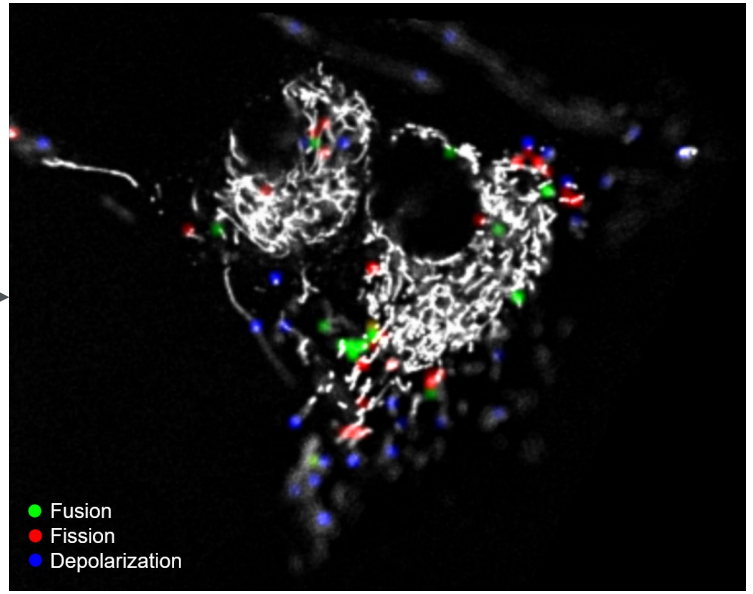
What we want to achieve





cGAN /
Pix2Pix /
Vox2Vox

Neural net

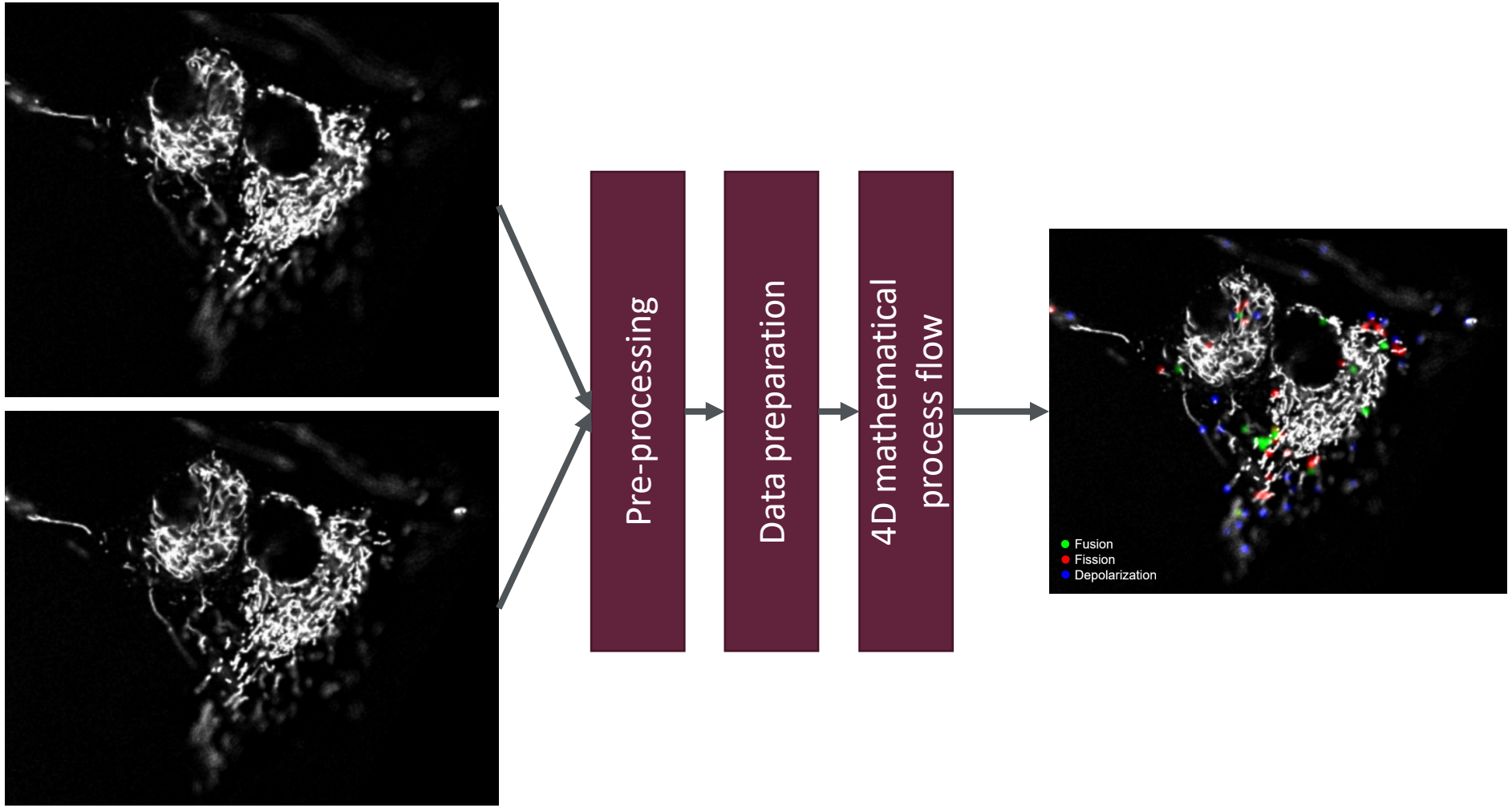


We need the ground truth!

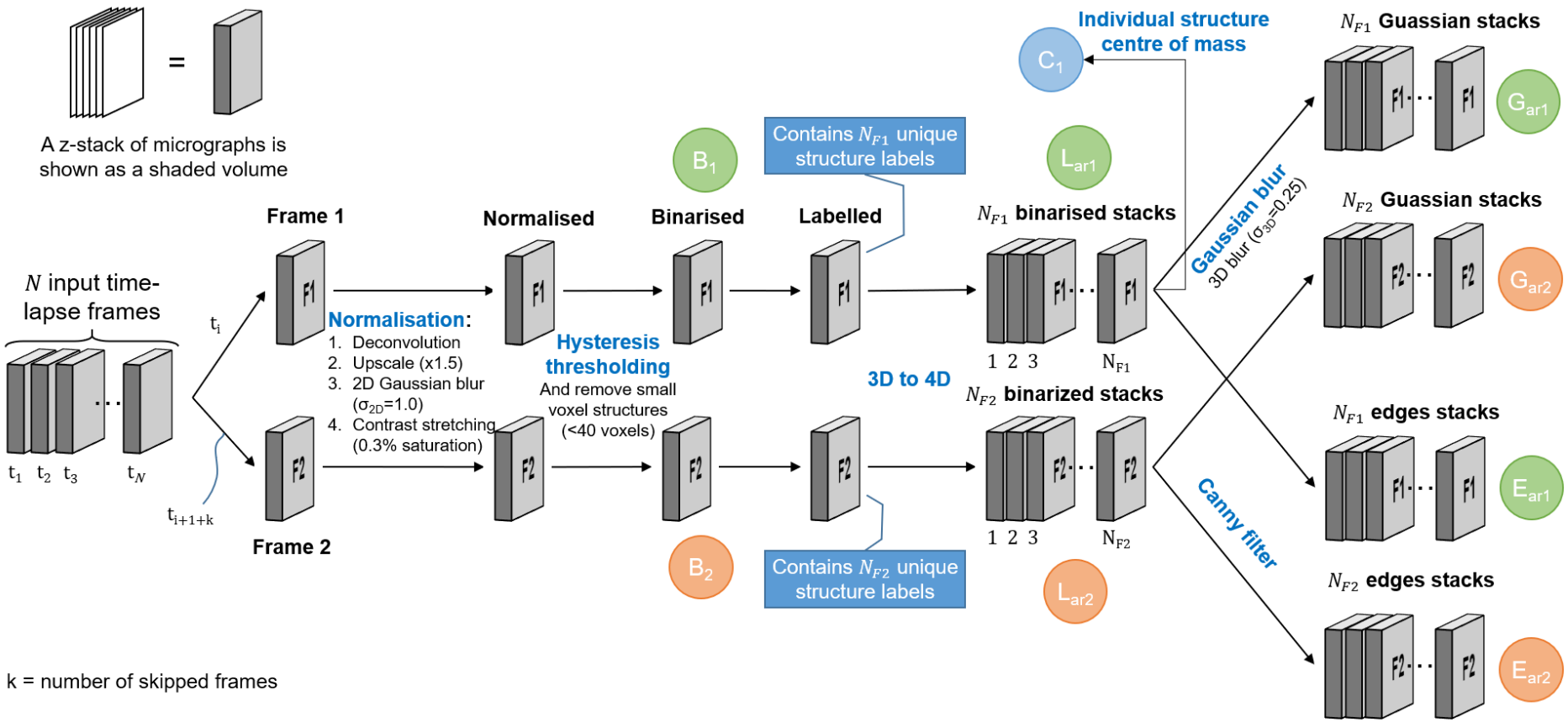
But we can't generate it by hand…

We need ground truth

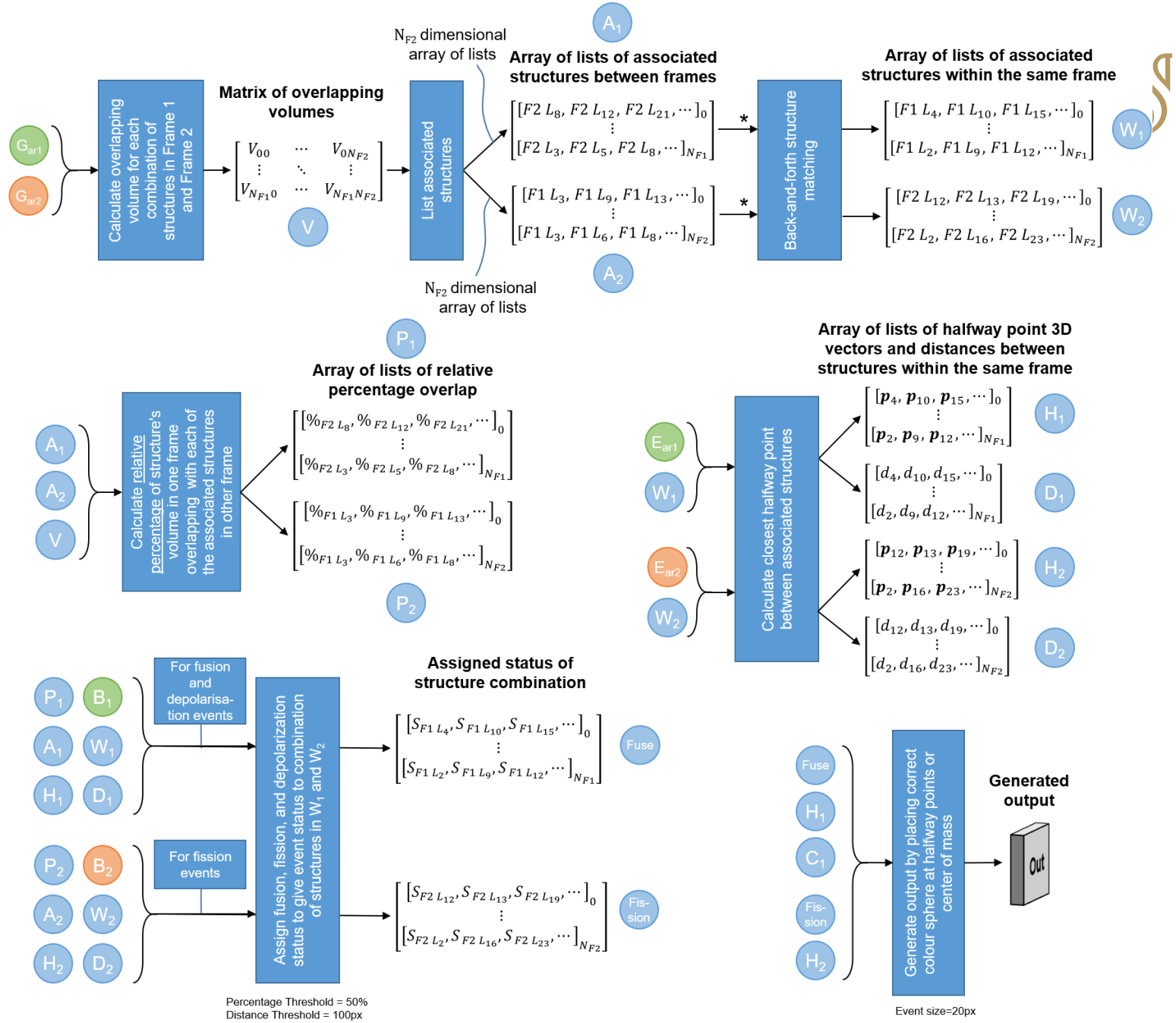
- Use time-lapse sequence to generate ground truth



Mitochondrial event localiser



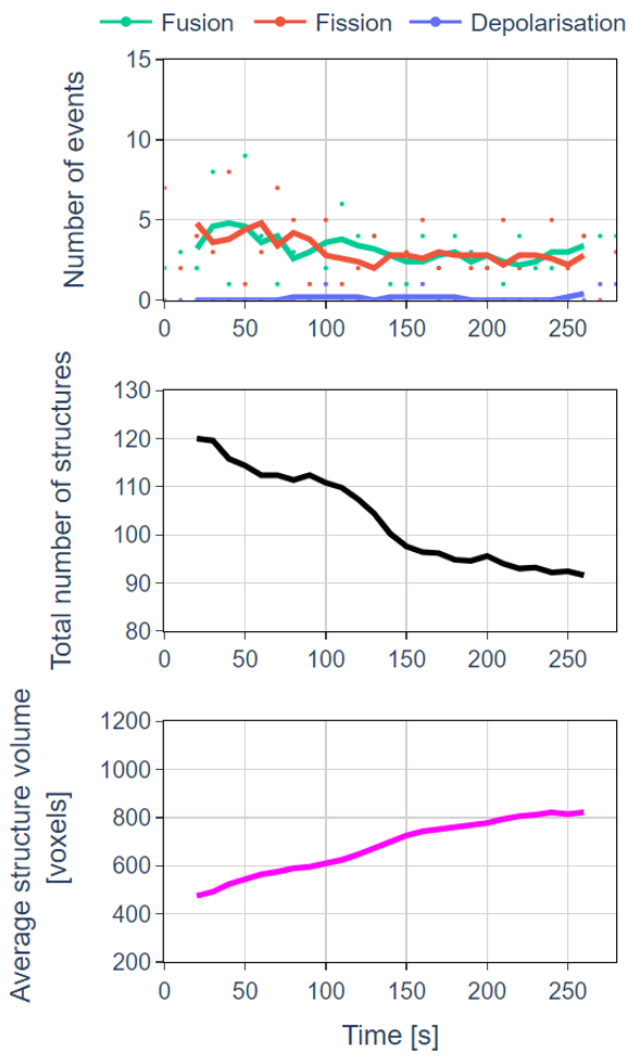
k = number of skipped frames



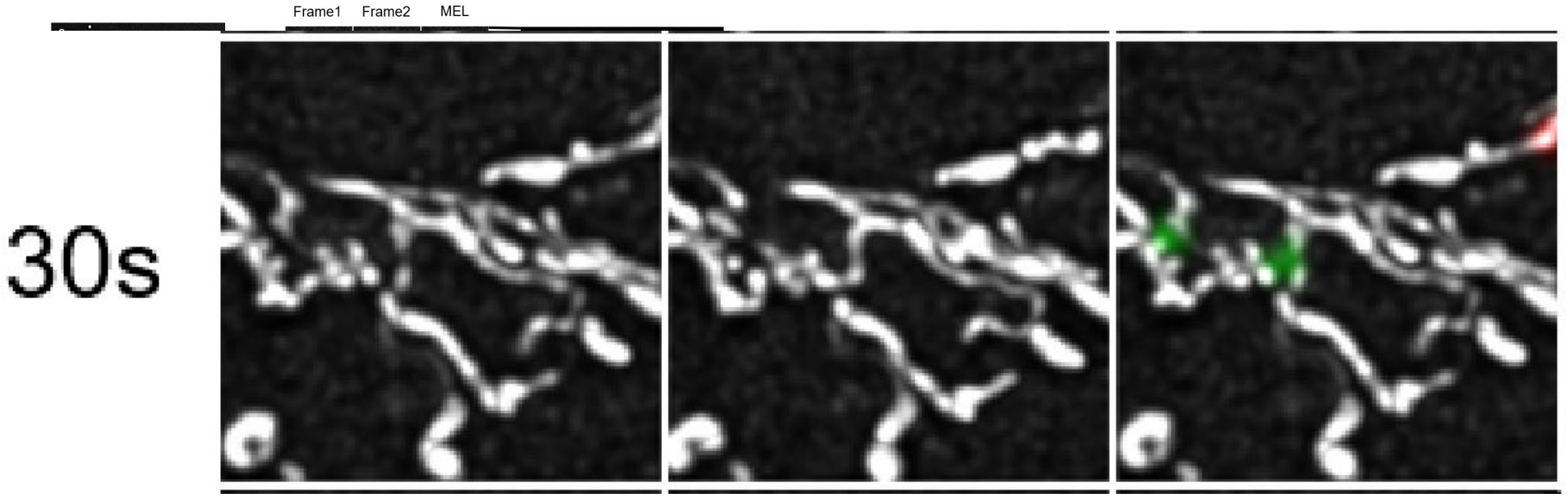
*If two arrows pass into a box, the function is called twice

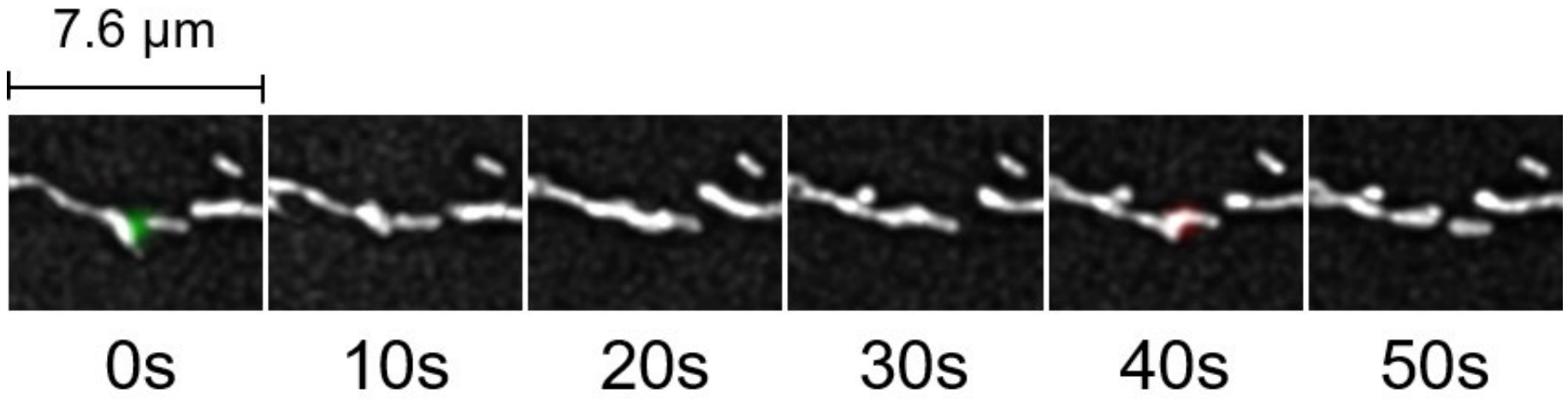
Example output

D. Graphs

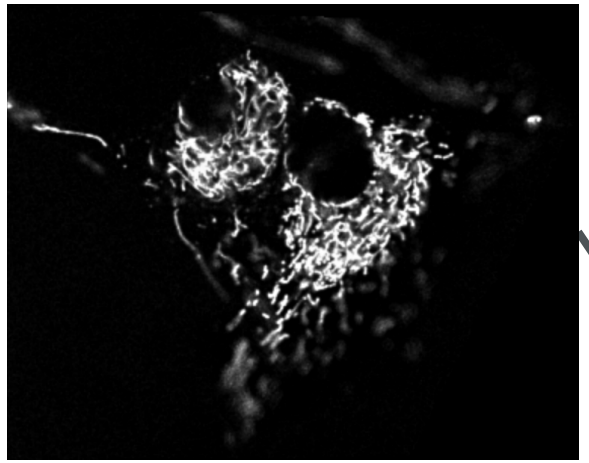


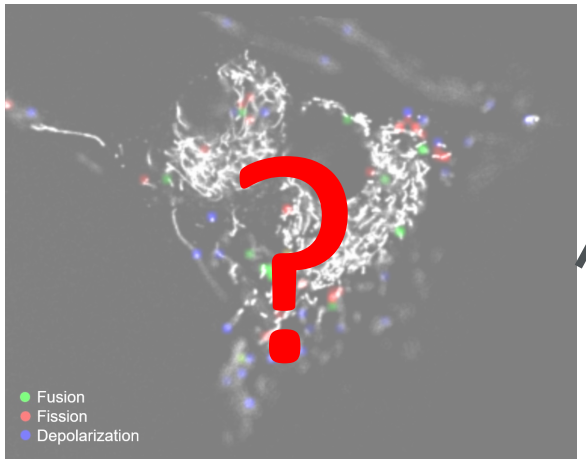
A. Full sample MIP **B.** ROI MIP **C.** MEL 3D ROI **D.** Graphs





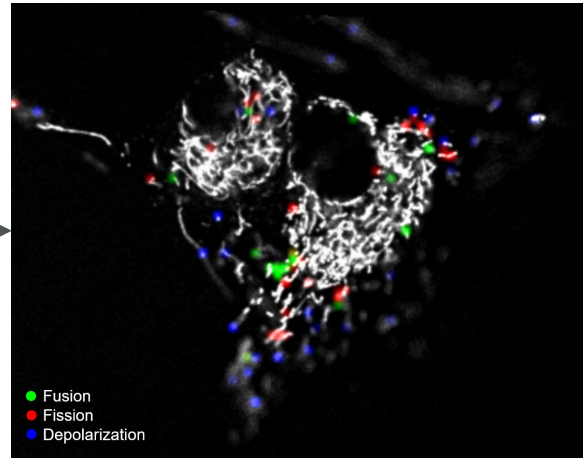
What we want to achieve





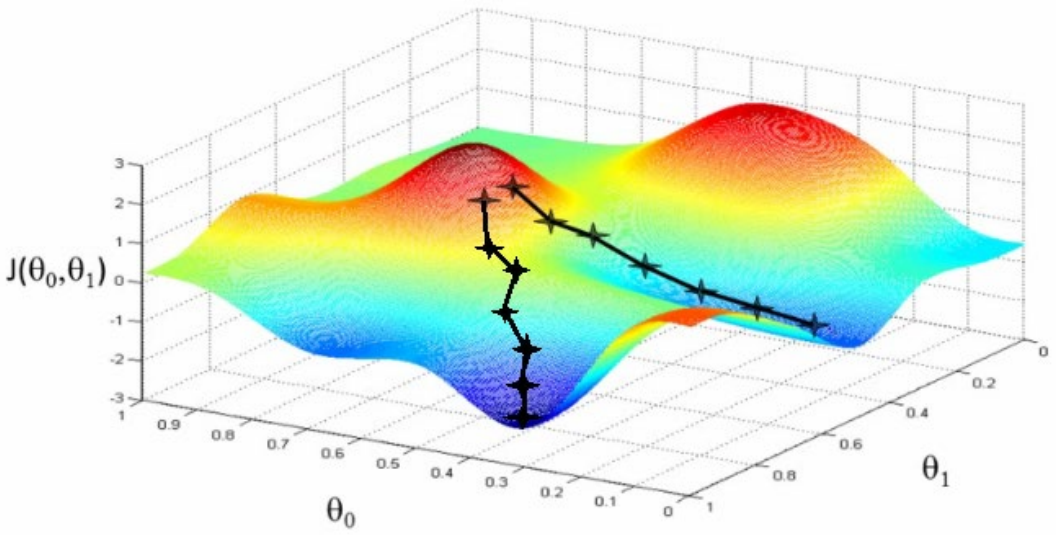
cGAN /
Pix2Pix /
Vox2Vox

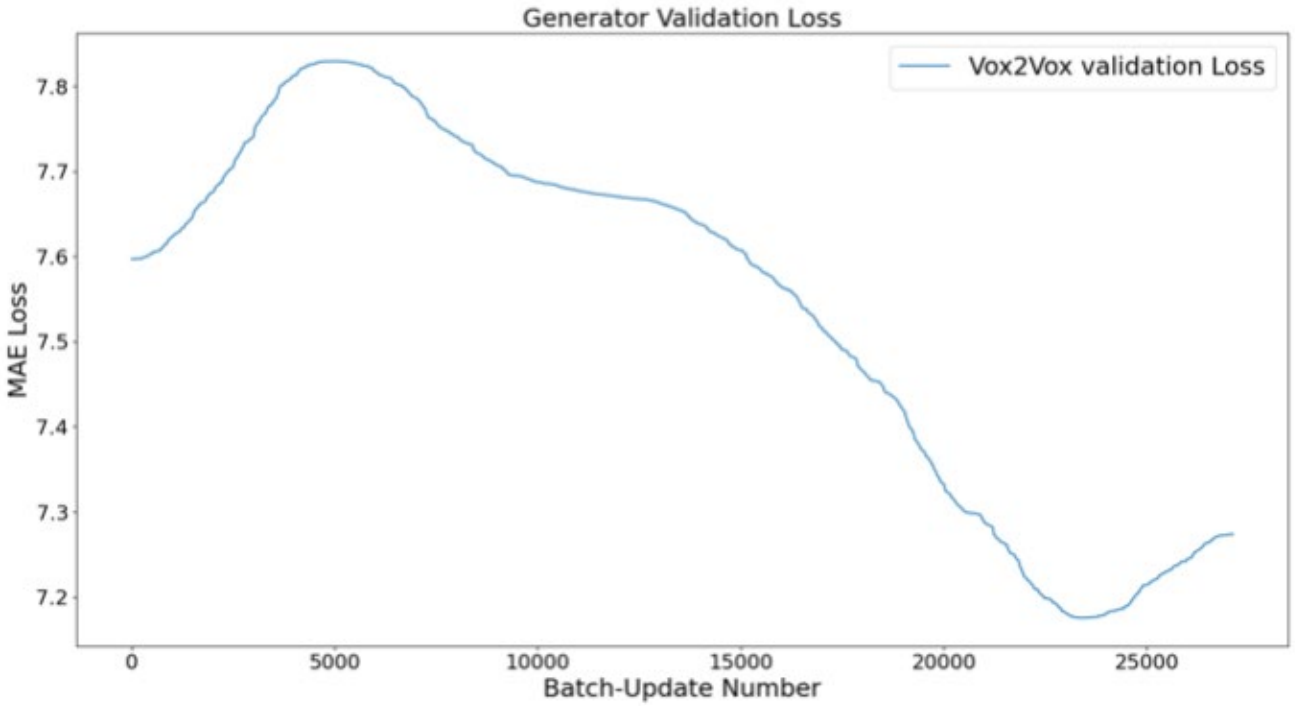
Neural net



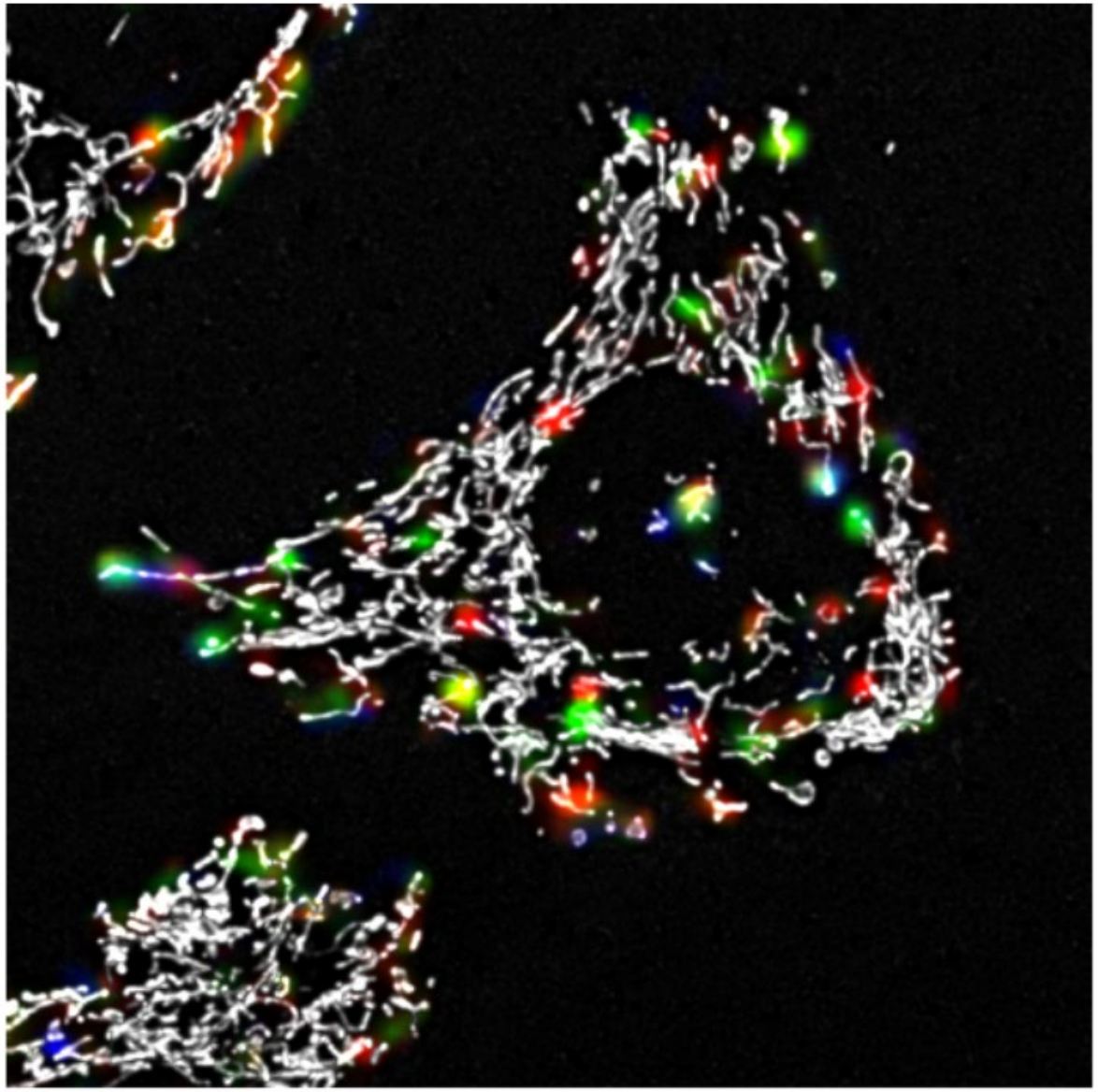
Gradient Descent

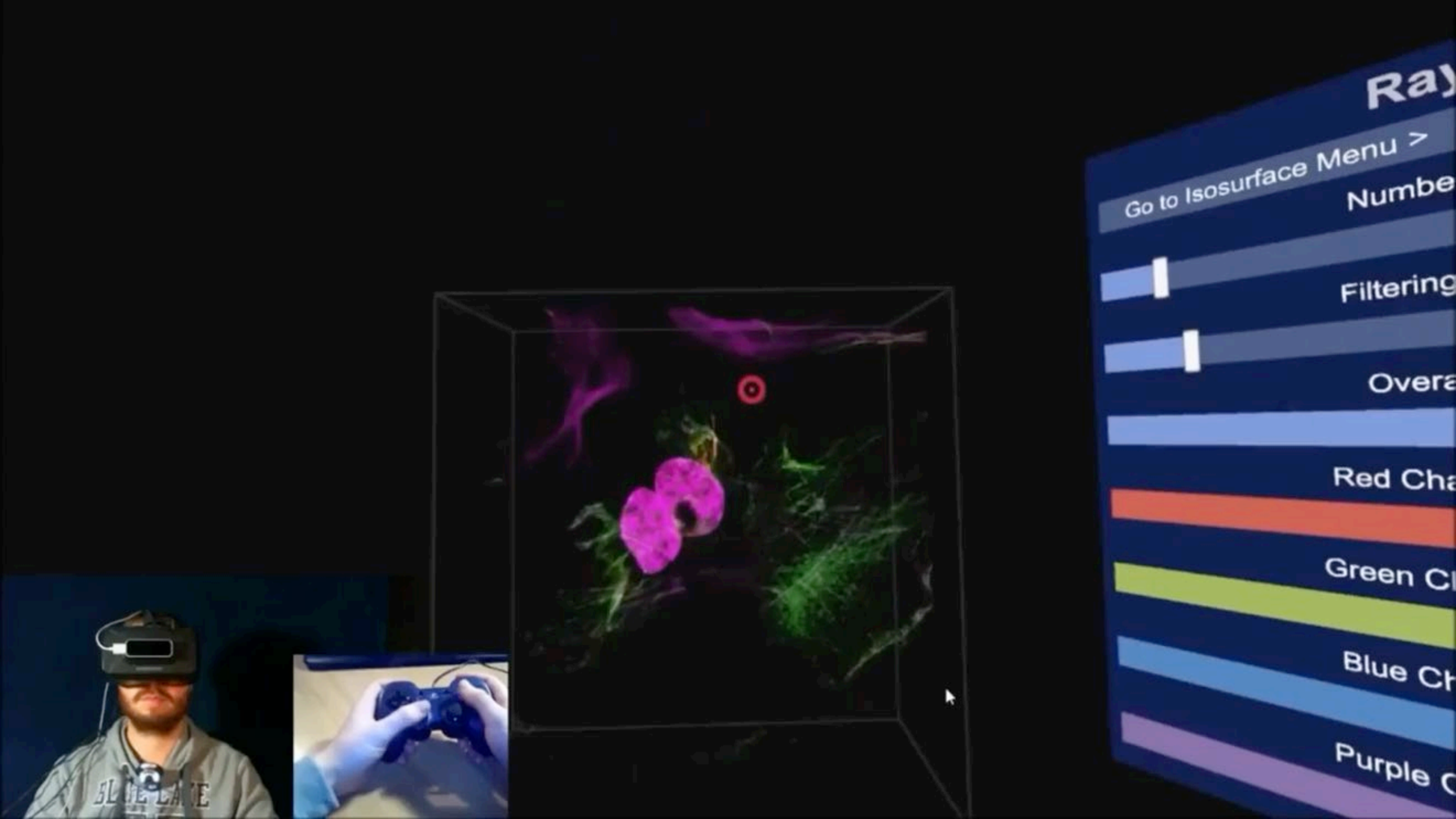
- Sometimes the objective function/loss function has more than one minima.





Results




Ray
Go to Isosurface Menu >
Numbe
Filtering
Overa
Red Cha
Green Ch
Blue Ch
Purple C

Virtual Learning Experience improves learning

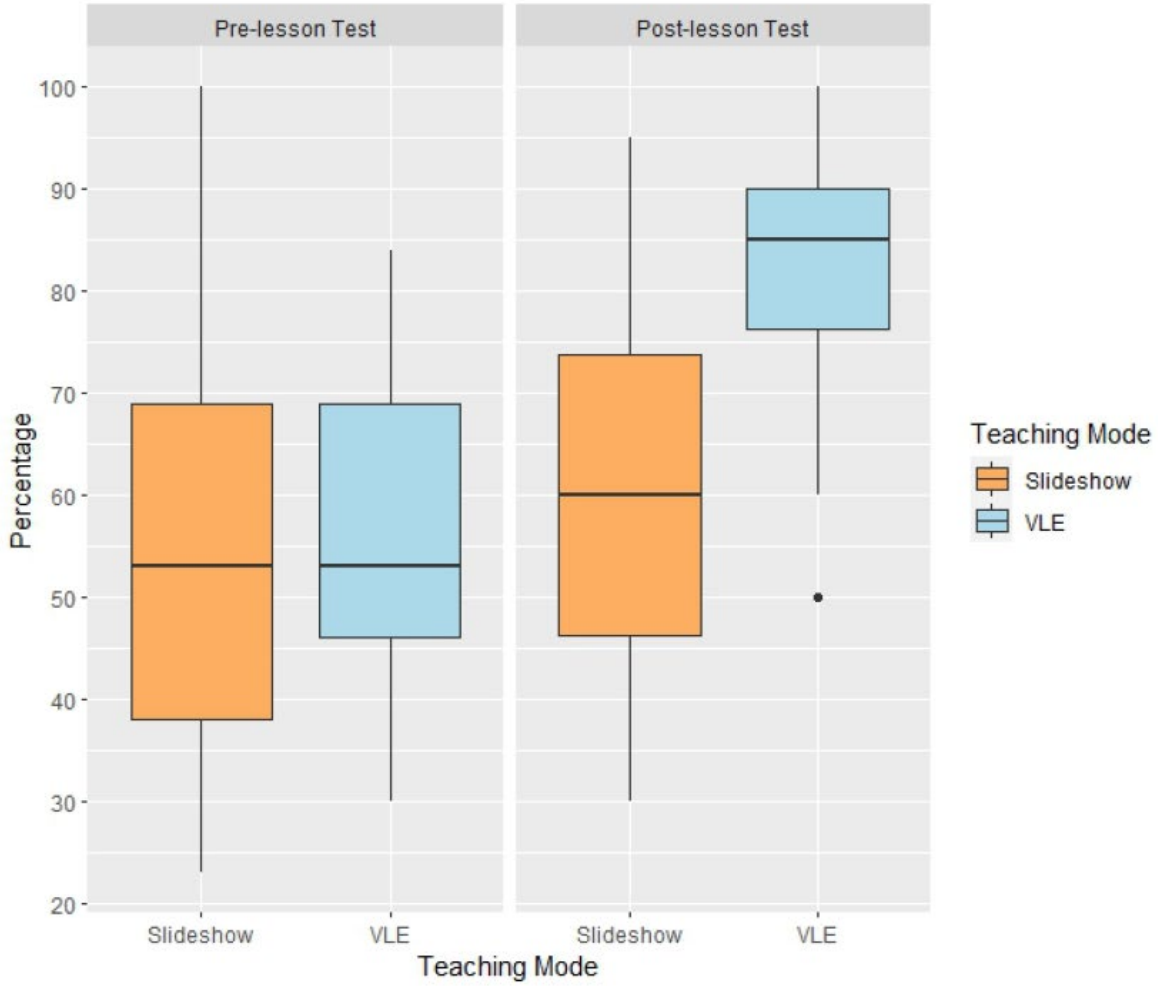


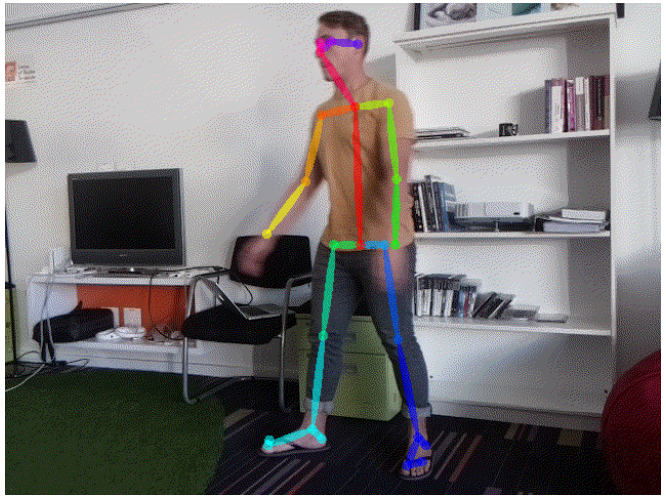
Figure 6.8: Box plots of pre- and post-lesson test scores for the VLE and Slideshow Group.

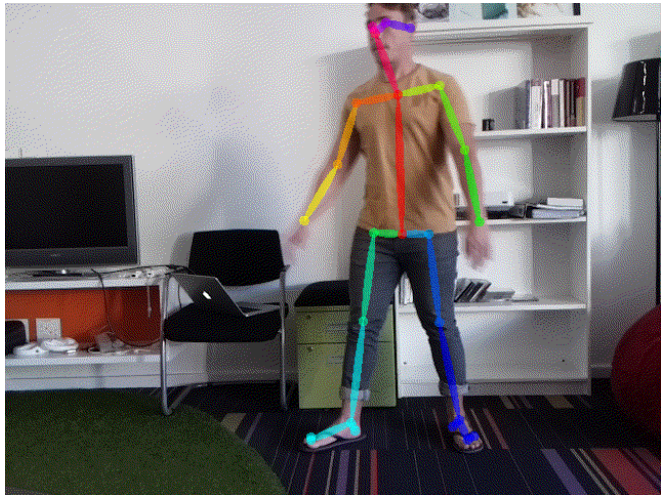
3D pose estimation

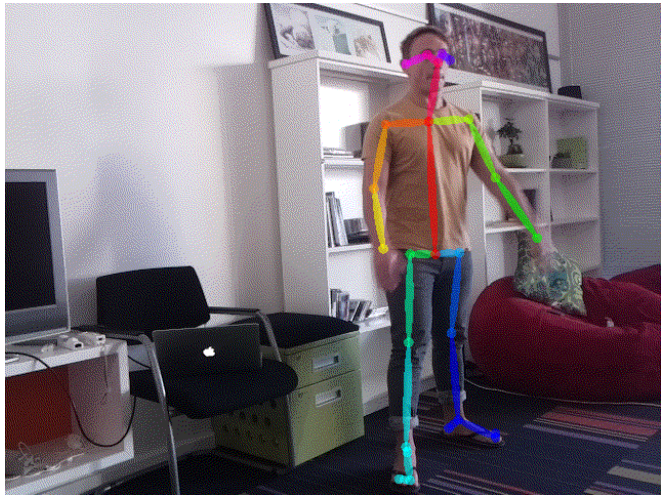
Stellenbosch UNIVERSITY IYUNIVESITHI UNIVERSITEIT

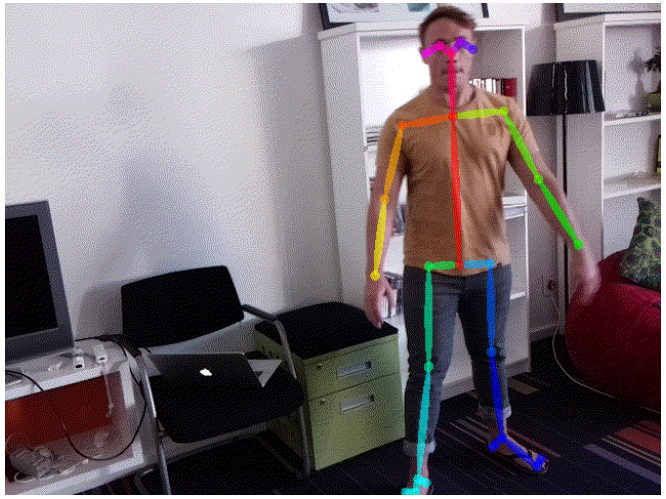
forward together
sonke siya phambili
saam vorentoe

4 camera views (new system)

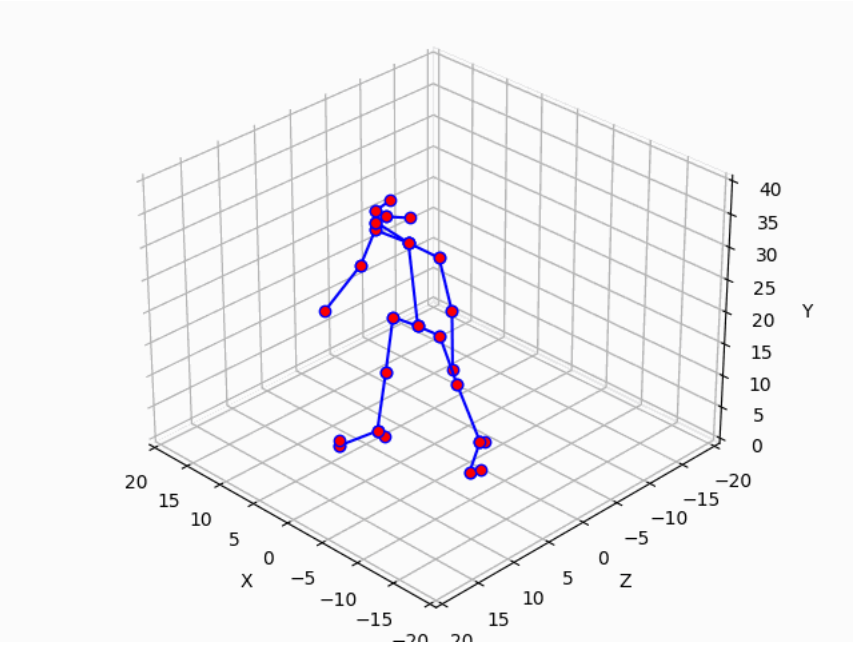




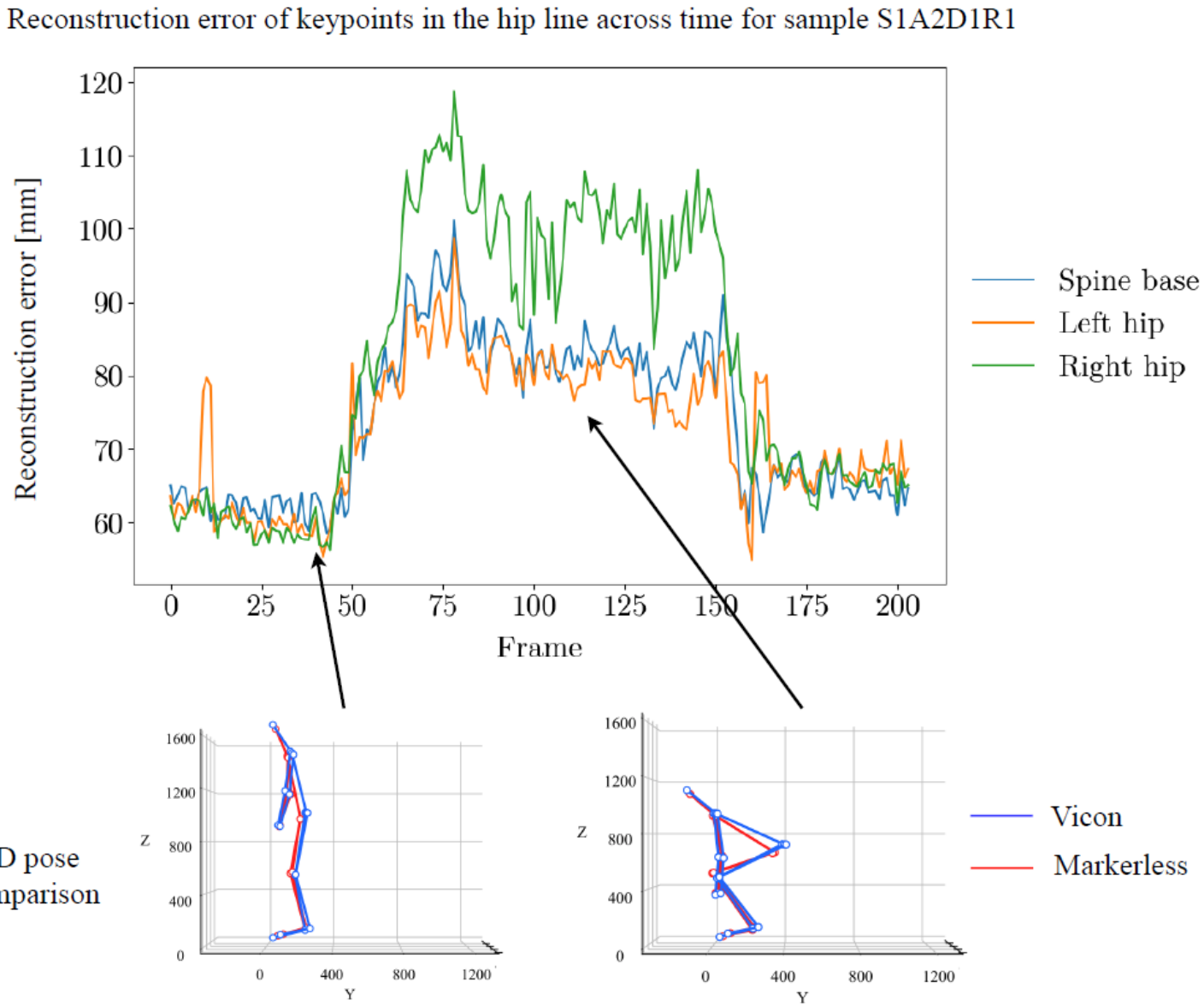




Reconstructed 3D pose



Reconstruction error for markerless system



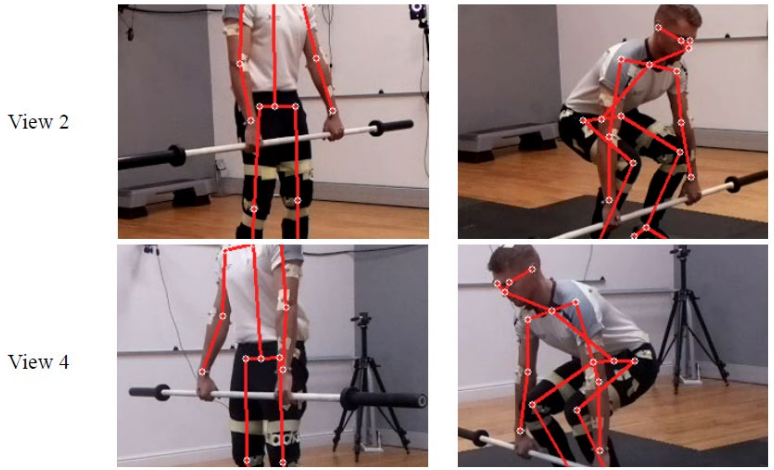


Figure 5.4: *Deadlift* action exhibiting inconsistent localisation of the hip joints. Ideally, the error should remain constant through the different phases of the action.

Spatial-Temporal Graph Convolutional Networks (ST-GCN) for metric learning

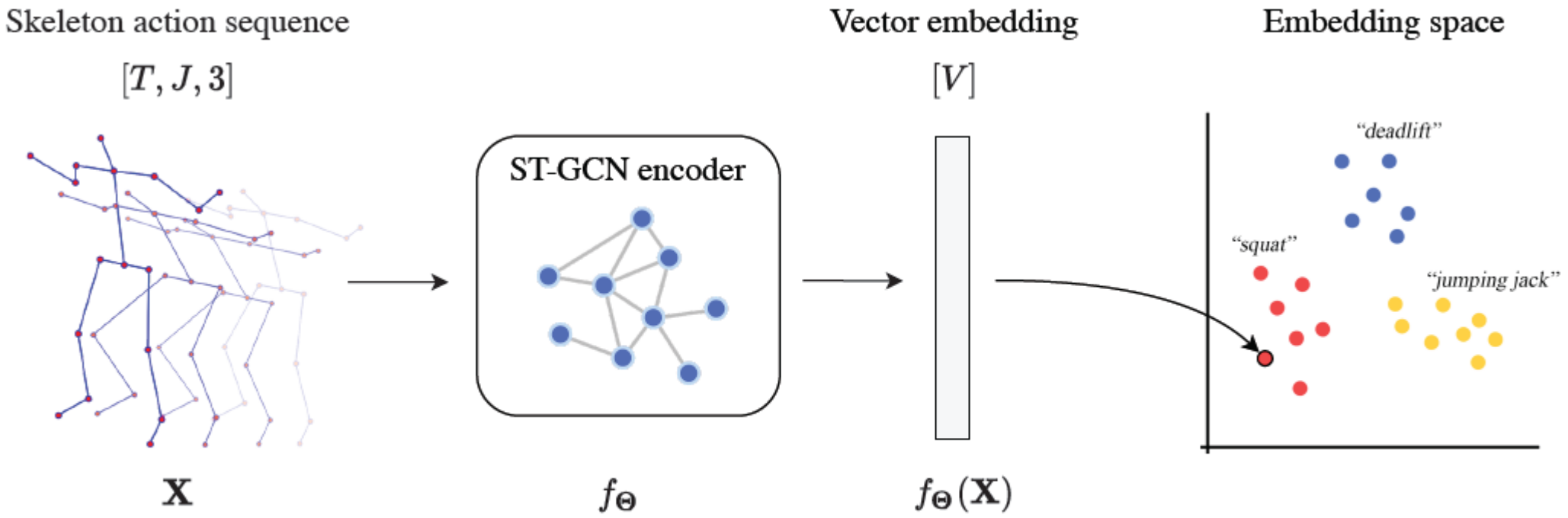
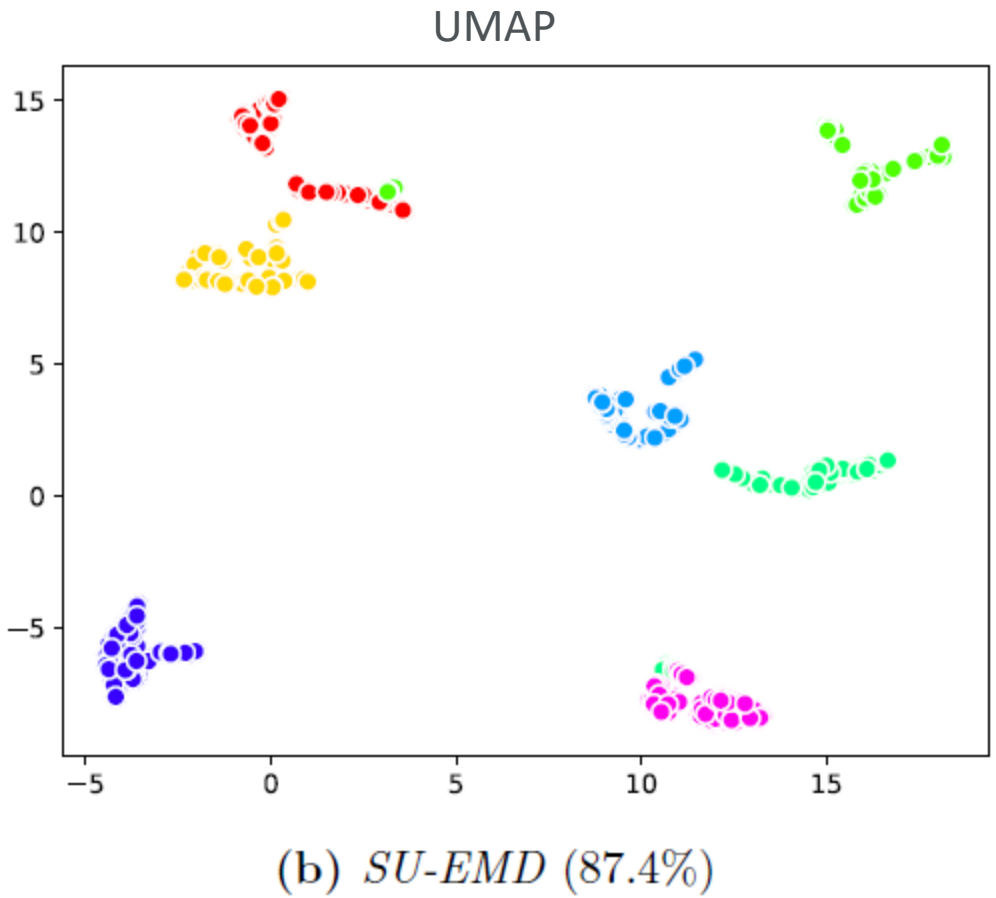


Figure 6.4: The metric GCN model. Illustration of the GCN feature extractor as a feature encoder in a metric learning paradigm.

One-shot action recognition on 7 never before seen classes



(b) *SU-EMD* (87.4%)

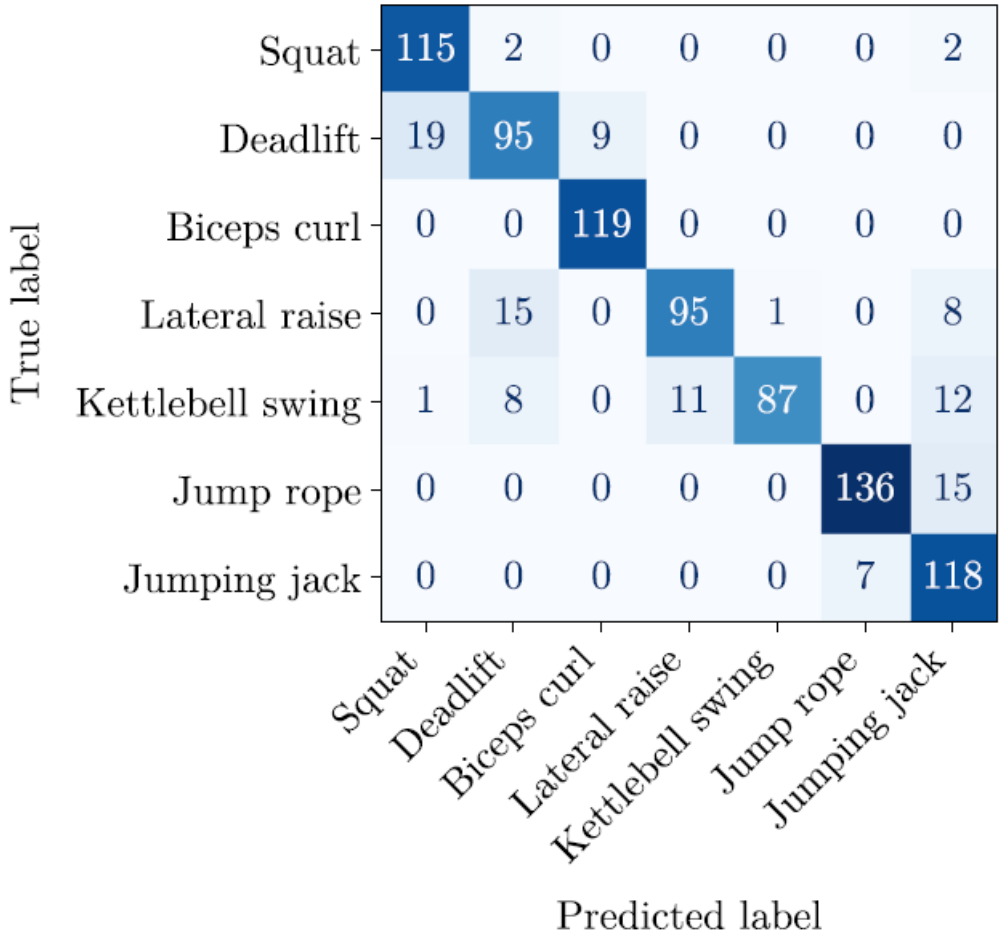


Figure 6.6: Confusion matrix for the final one-shot tests in the seven *SU-EMD* classes.

Improving Face Recognition of Individuals with Highly Pigmented skin

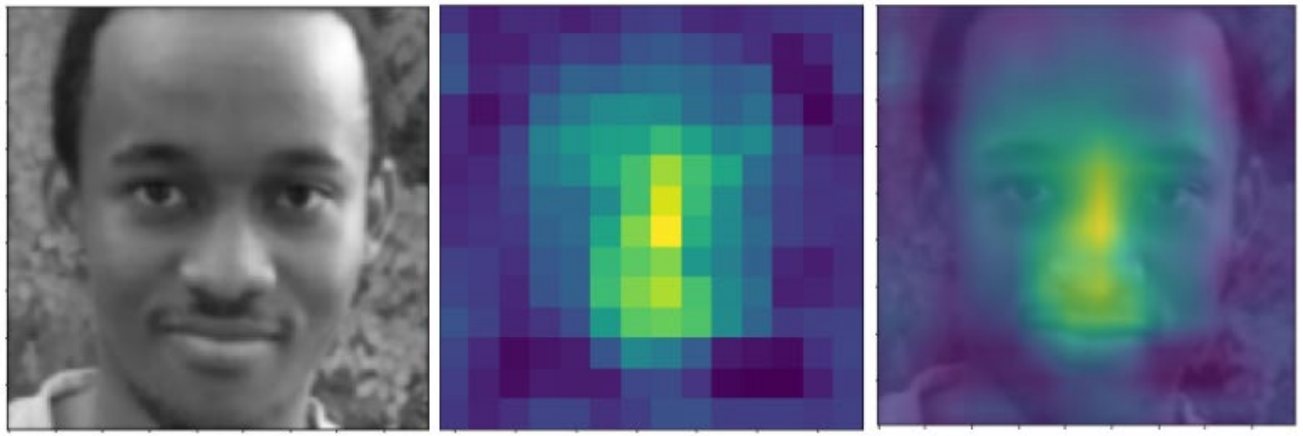
Stellenbosch UNIVERSITY IYUNIVESITHI UNIVERSITEIT
forward together sonke siya phambili saam vorentoe

- 545 individuals
- 5 different orientations

Table 4: Accuracy for a model with fine-tuned weights.

VGG16						
Optimiser	**Visible**		**Infrared**		**Full Spectrum**	
	Accuracy	*AUC*	*Accuracy*	*AUC*	*Accuracy*	*AUC*
Adam	97.3	0.986	99.7	0.993	99.1	1.000
SGD	97.6	0.985	99.7	0.986	99.4	1.000
AdaGrad	97.3	0.986	99.7	1.000	99.1	1.000
ResNet50						
Optimiser	**Visible**		**Infrared**		**Full Spectrum**	
	Accuracy	*AUC*	*Accuracy*	*AUC*	*Accuracy*	*AUC*
Adam	0.0	-	0.3	-	0.0	-
SGD	97.9	0.991	99.7	0.998	99.1	1.000
AdaGrad	97.9	0.994	98.4	0.988	99.1	0.990

Conclusion: Using infrared light improves the accuracy of face detection algorithms



(a) Original image. (b) Activation map. (c) Overlaid activation map.

Figure 4: Example of activation map produced from a CNN model, and how it can be overlaid on the original image.

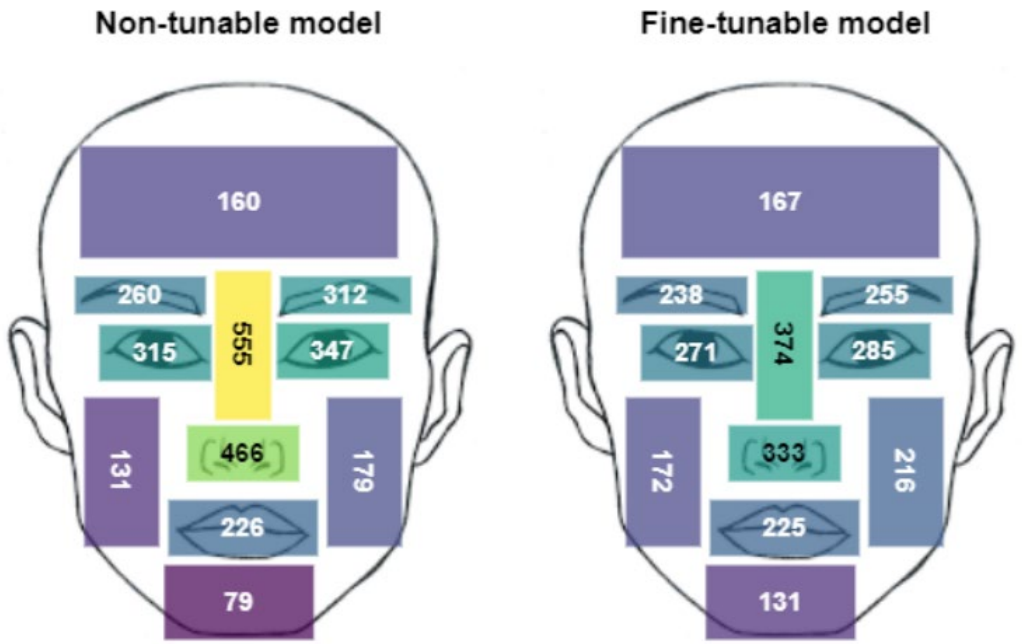


Figure 13: Comparison of average activation intensity values over facial regions.

Acknowledgements

Stellenbosch UNIVERSITY IYUNIVESITHI UNIVERSITEIT
forward together
sonke siya phambili
saam vorentoe

Alex Muthua

James de Villiers

Michael Deyzel

Heino Laubscher

Stellenbosch
UNIVERSITY
IYUNIVESITHI
UNIVERSITEIT

forward together
sonke siya phambili
saam vorentoe

Thank you
Enkosi
Dankie

Photo by Stefan Els