NEXT PROJECT

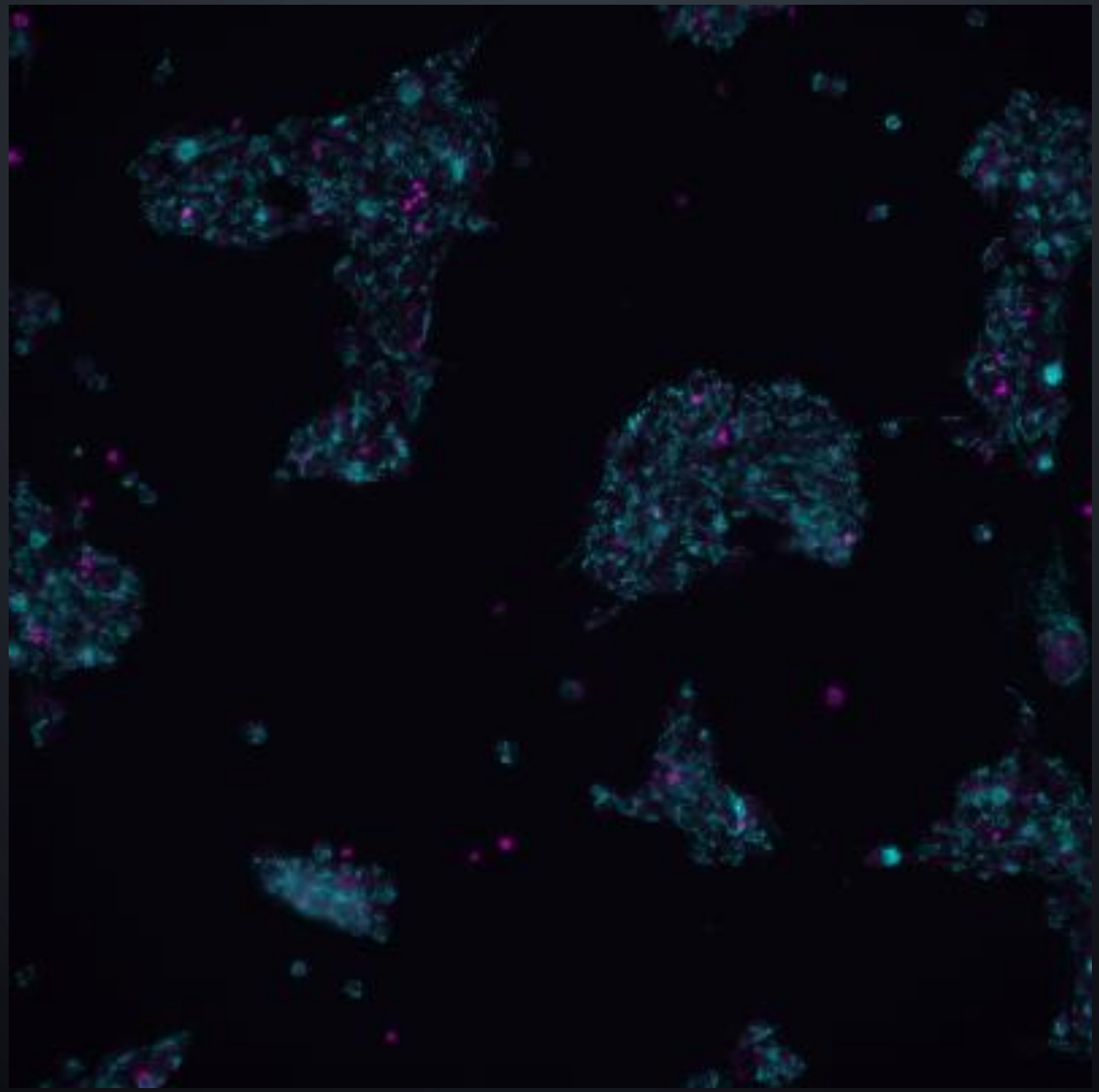
WRITING A THESIS …

„EXPLAINABLE AI FOR STEM CELL MODELS WITH LEIGH DISEASE“

Previous Project:

Classifying Leigh Disease Images

Combined image:

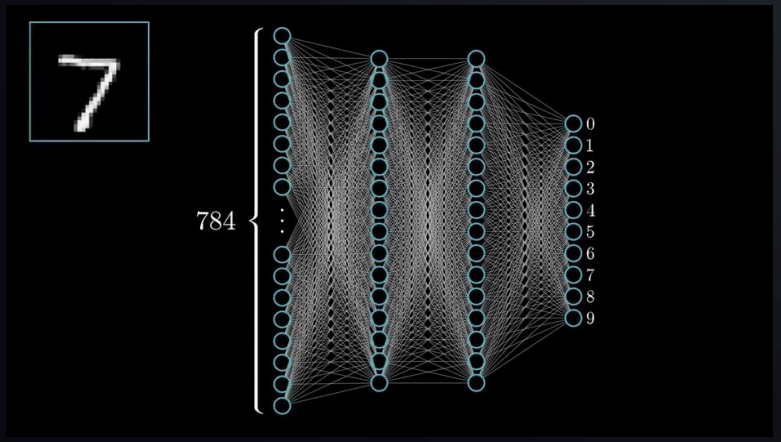


CLASSIFICATION (DS VS. TT)

Got 540/540 with accuracy 100.000%

	precision	recall	f1-score	support
DS	1.00	1.00	1.00	264
TT	1.00	1.00	1.00	276
accuracy			1.00	540

Z'-Factor: 0.924



WHAT ARE WE GOING TO DO WITH THIS?

> Help to find the best component/drug for treating Leigh Disease

  -> scan for cell culture images appearing to be wildtypes after treatment

> **Try to understand the difference between the classes**

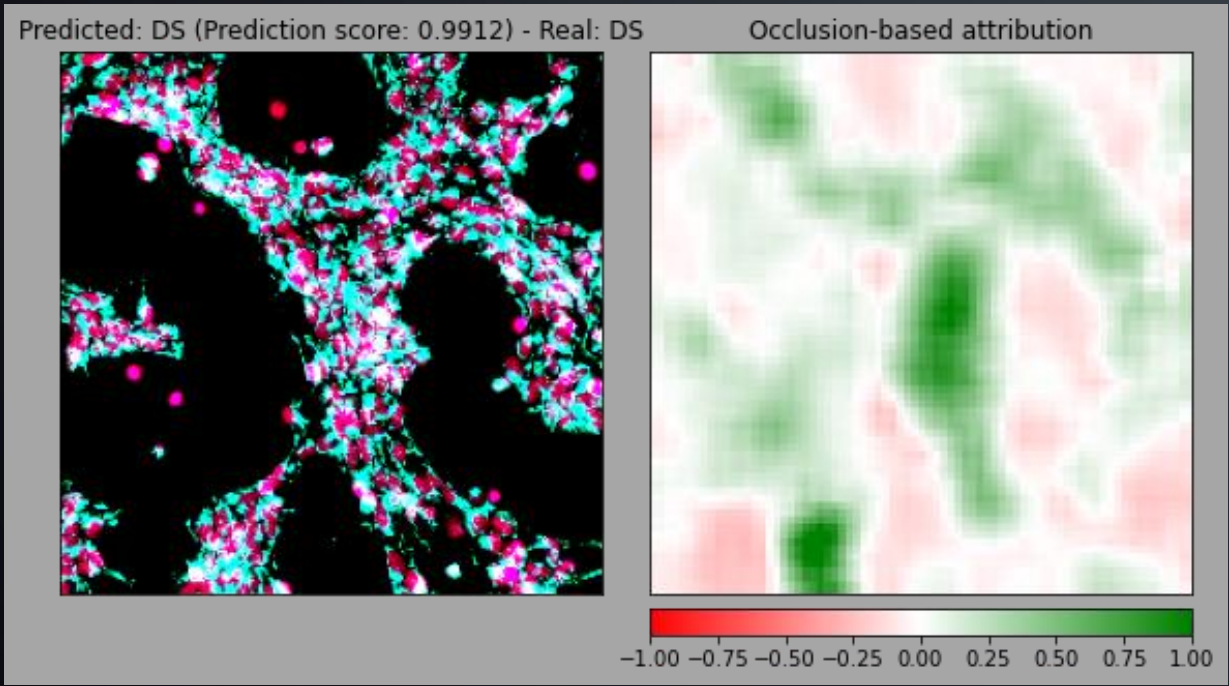
HEAT-MAPS CAN BE SUFFICIENT, SOMETIMES

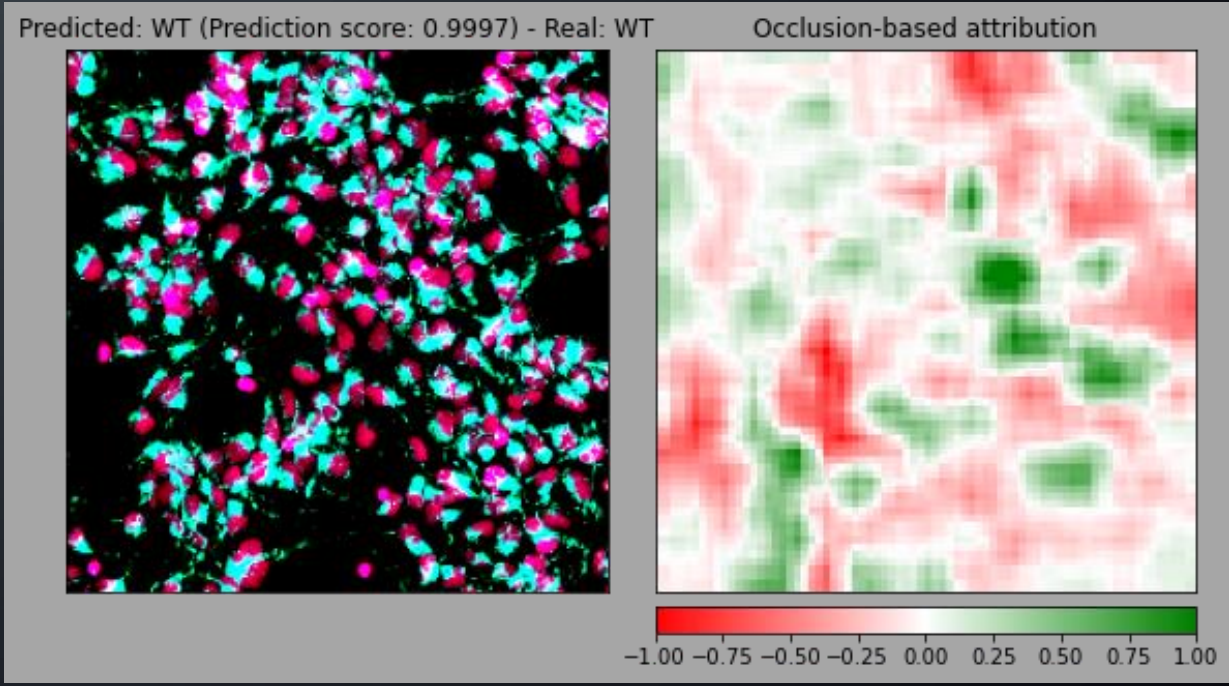
Figure 2: Example of a saliency map explanation of a True Positive (TP) image for the label *"train"*. It highlights the contours of the lines below the train. A possible interpretation is that the CNN has learned to recognise trains when rails are present.

Figure 3: Example of a saliency map explanation of a False Positive (FP) image for the label *"train"*. A possible interpretation is that edges in the lower part appeared similar to rails, which could explain this error.

OCCLUSION BASED ATTRIBUTIONS

Sanity check: It can detect our cells





Problematic: In some images, empty areas are more important

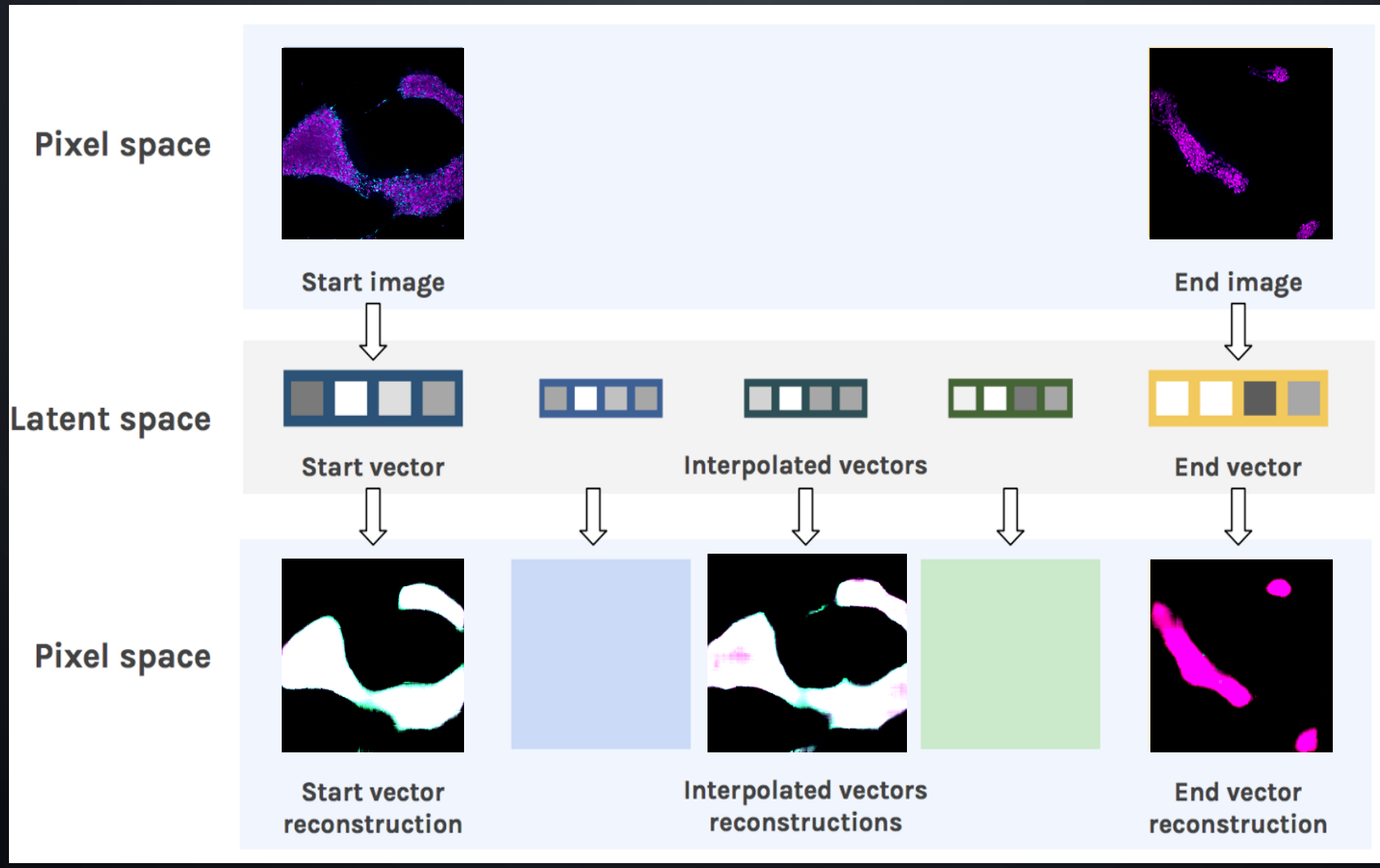
NEW PLAN

> Set up a neural network that can classify & generate images

> Modulate an image to contain the information of another class

> How does it change? Can we spot a difference?

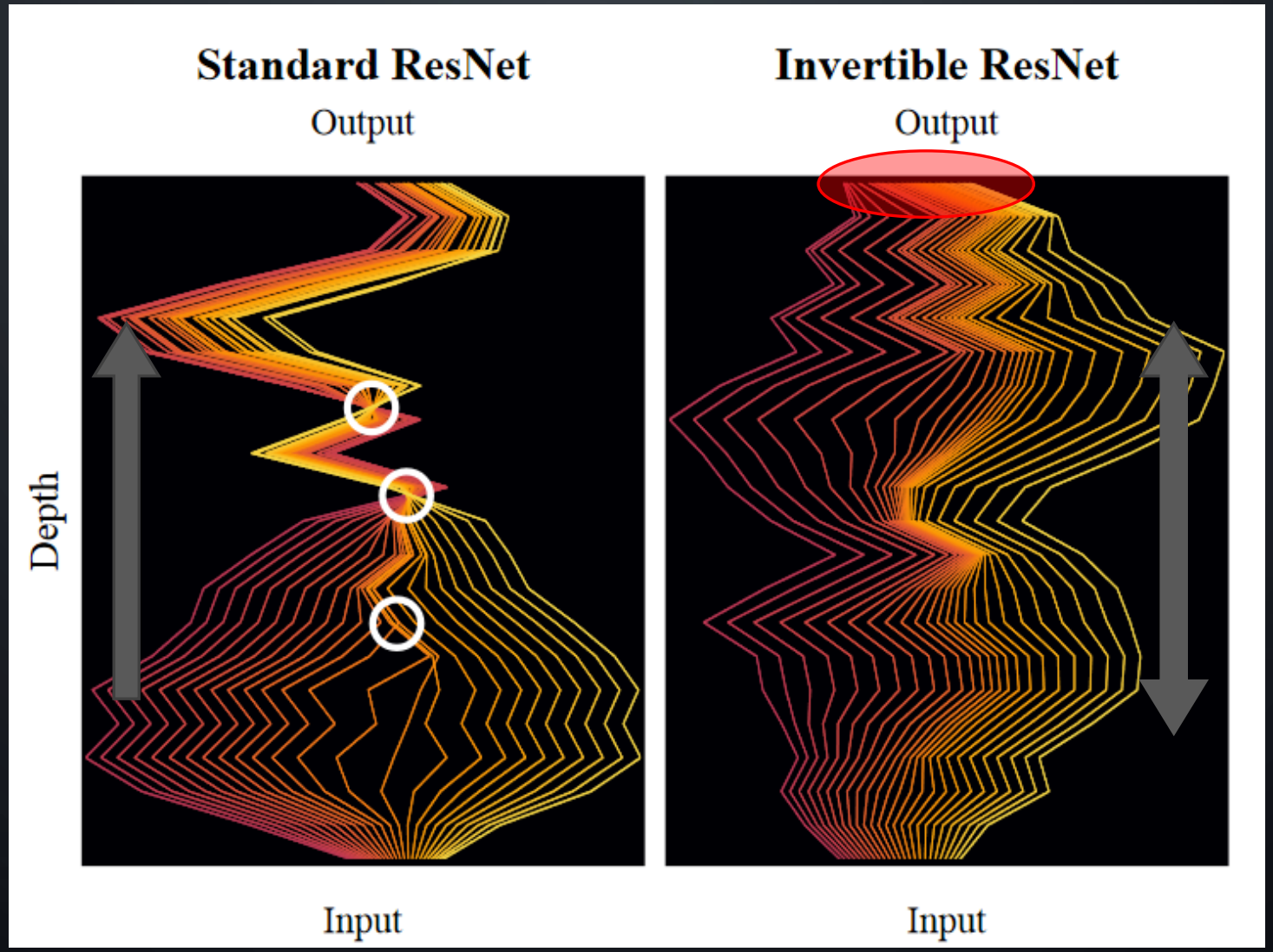
GENERATE NEW IMAGES (AUTOENCODER)



GENERATE NEW IMAGES (AUTOENCODER)

GENERATE NEW IMAGES (AUTOENCODER)

INVERTIBLE NEURAL NETWORK

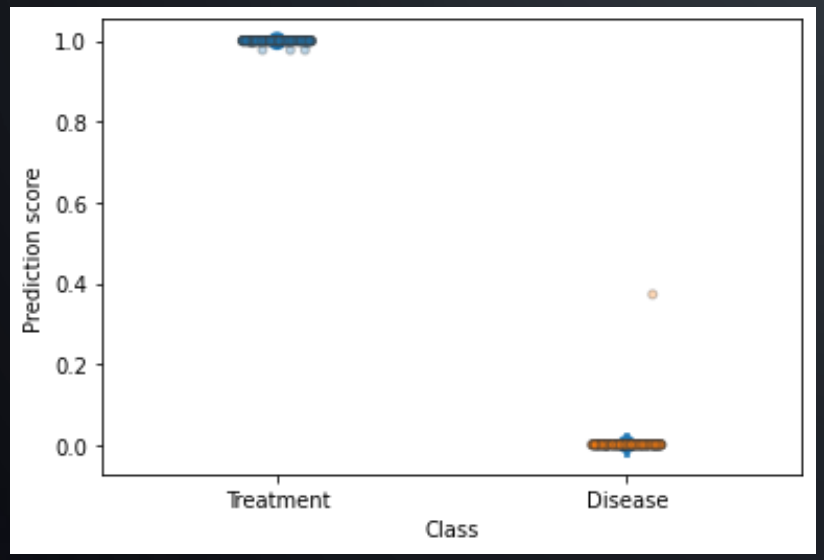


CLASSIFICATION (DS VS. TT)

ResNext50:

Got 540/540 with accuracy 100.000%

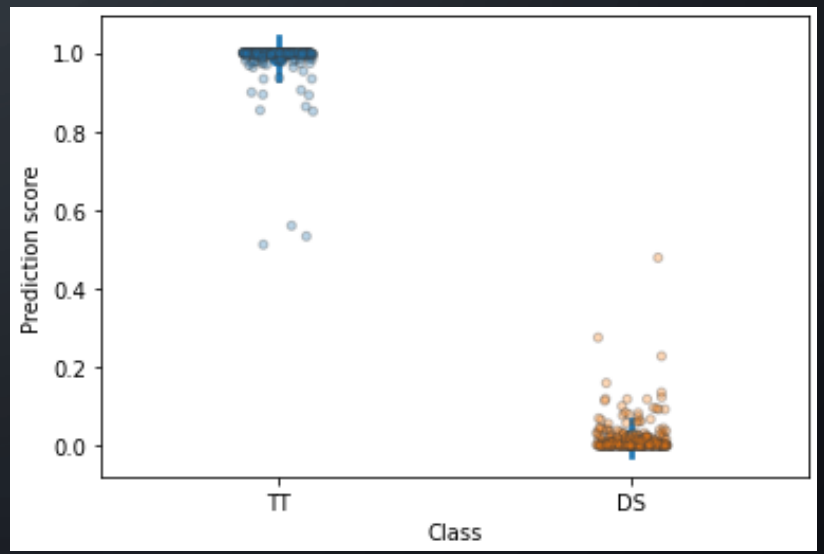
Z'-Factor: 0.924



i-ResNet50:

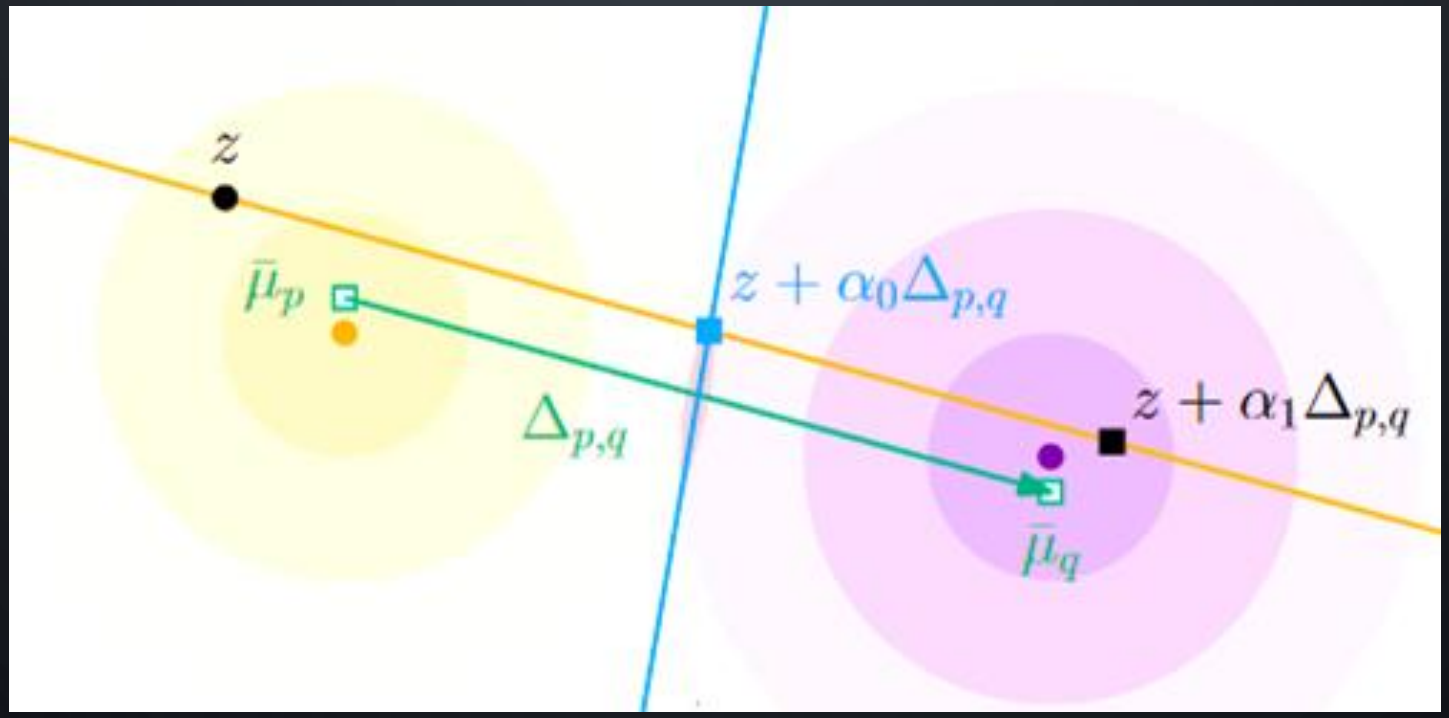
Got 540/540 with accuracy 100.000%

Z'-Factor: 0.701

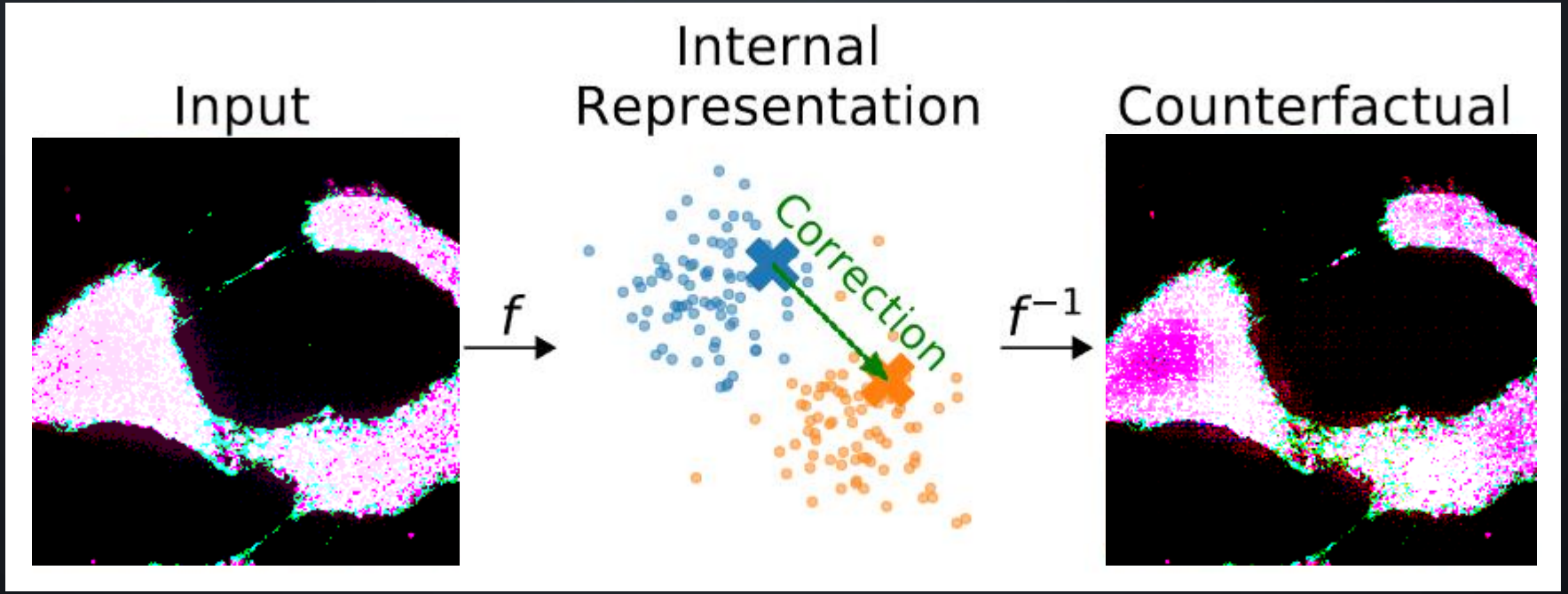


COUNTERFACTUALS

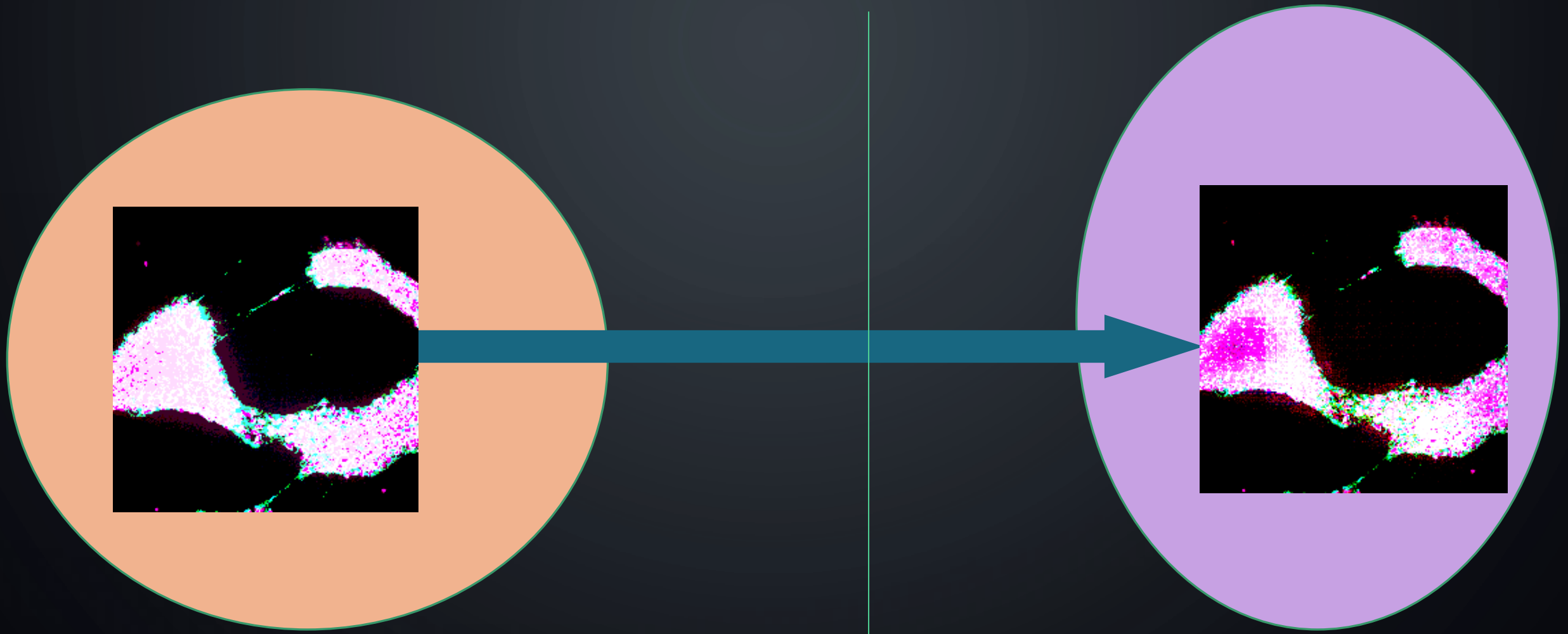
Disease FCCP-Treated



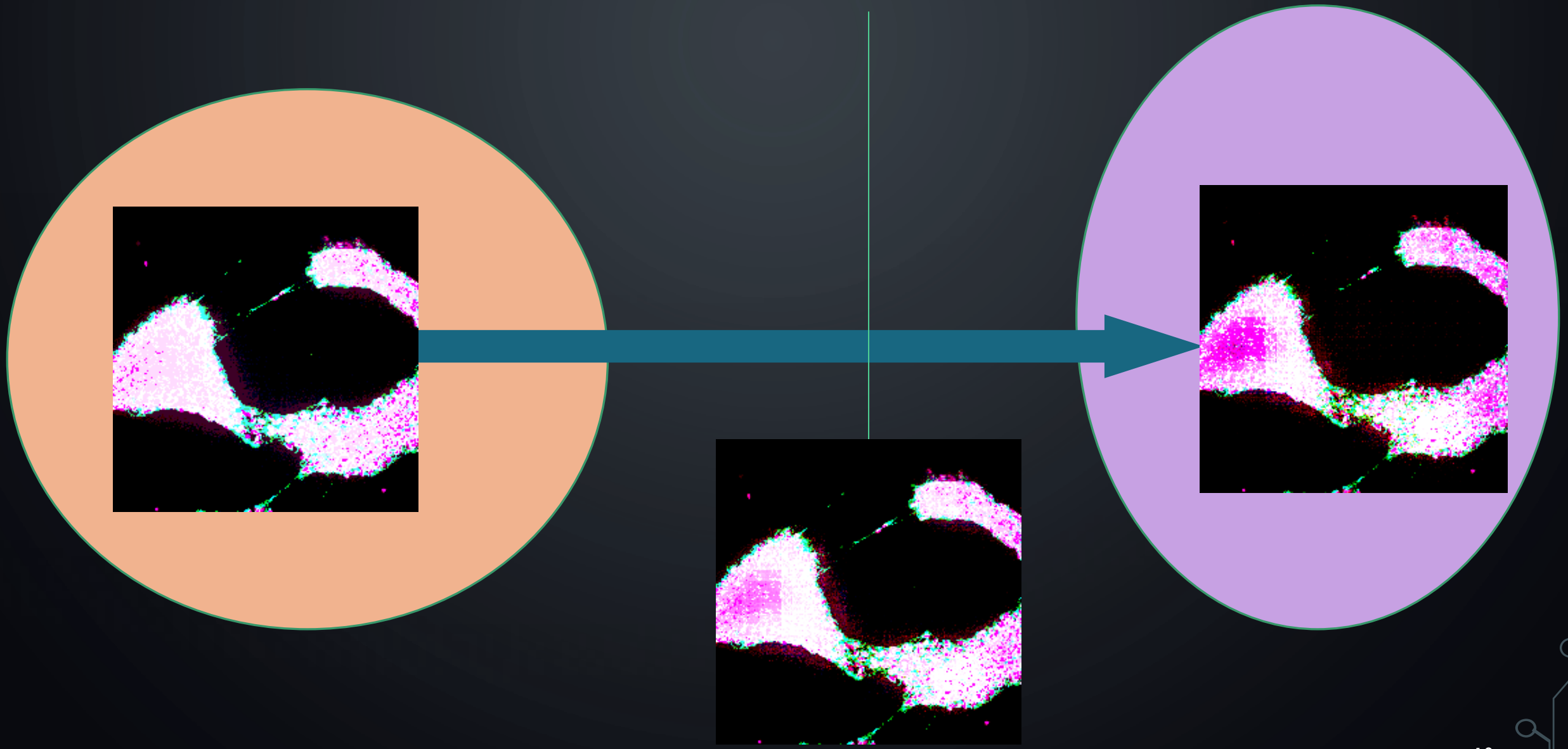
COUNTERFACTUALS



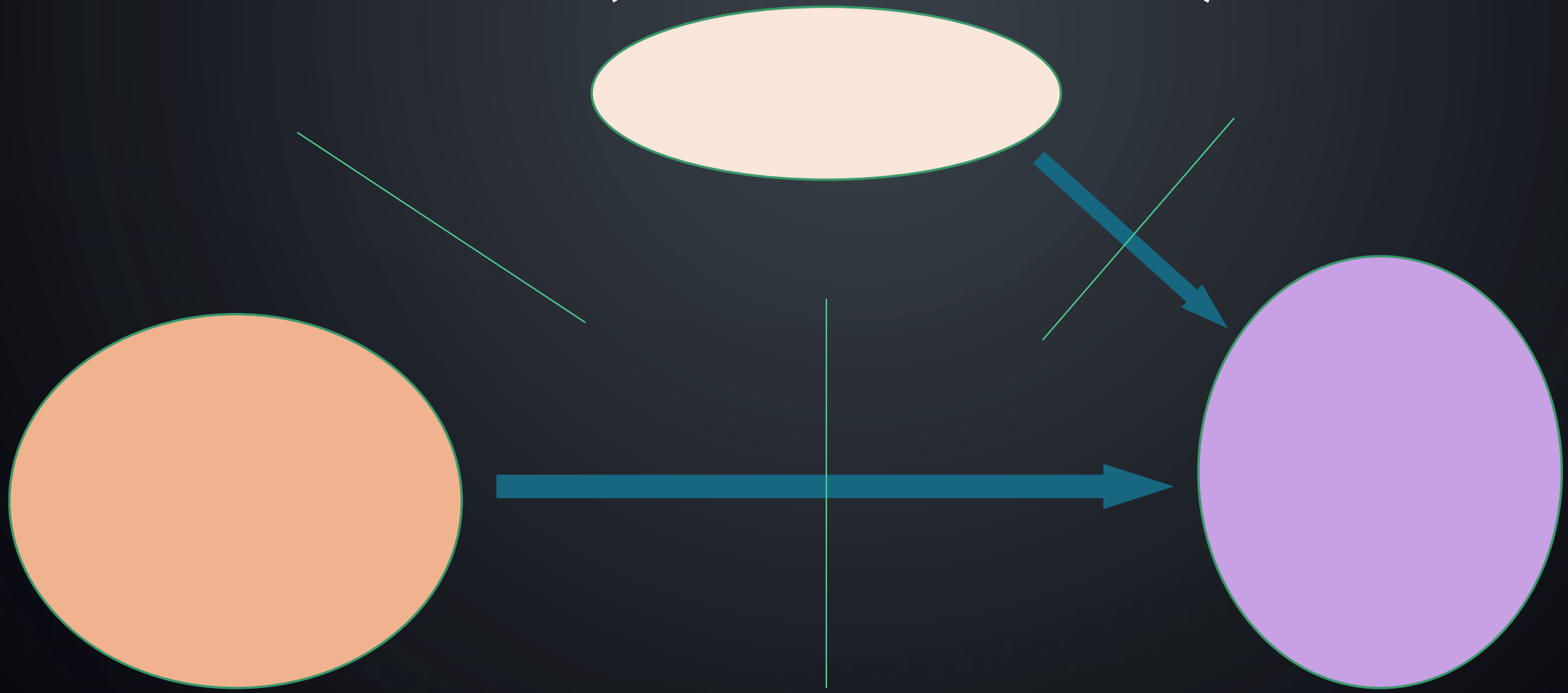
DECISION BOUNDERY



DECISION BOUNDERY



FUTURE WORK: MULTI-CLASS APPLICATION (DIFFERENT TREATMENTS)



> Set up a neural network that can classify & generate images

→ Found one

> Modulate an image to contain the information of another class

→ Convincing images

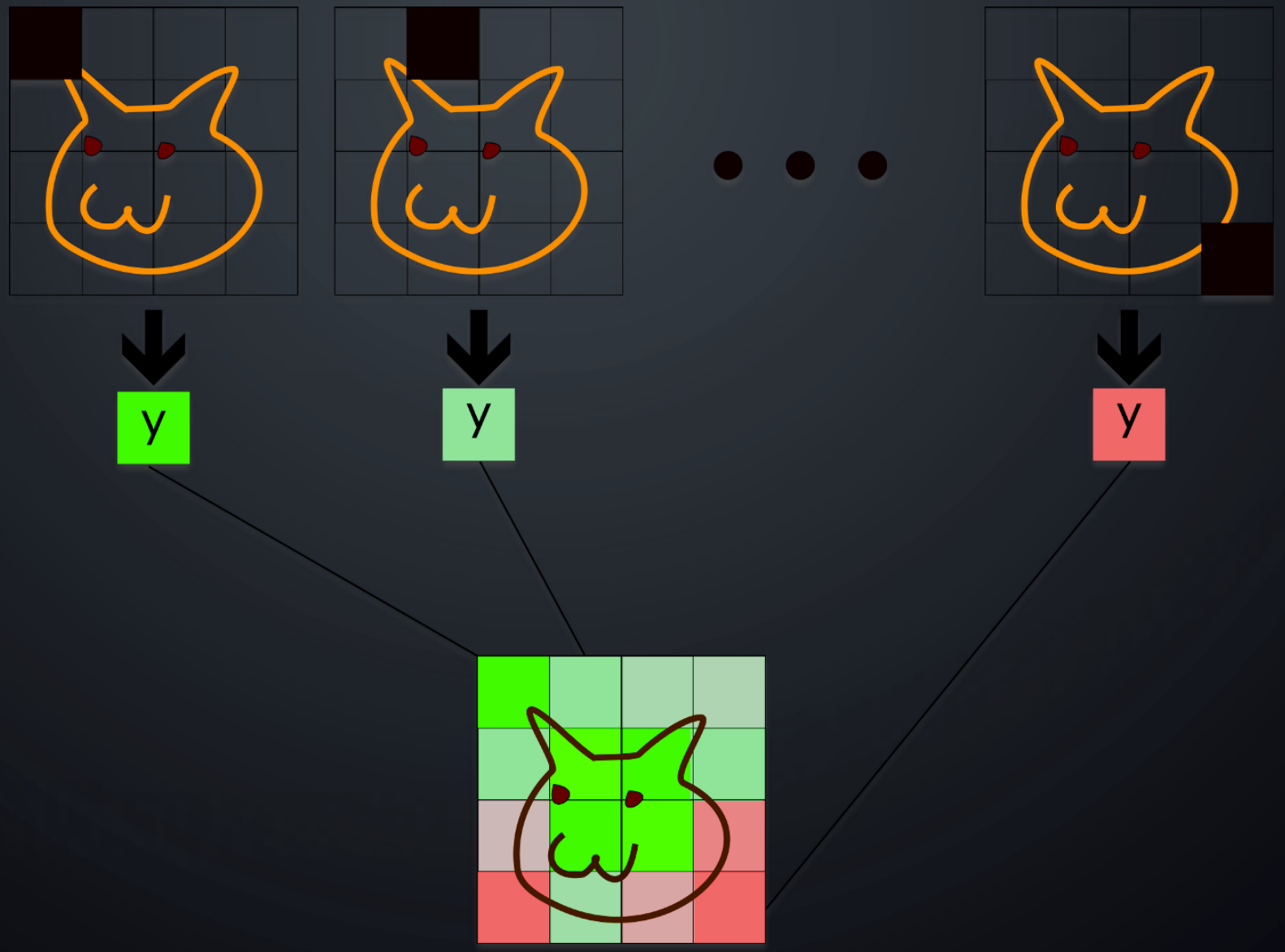
> How does it change? Can we spot a difference?

→ The treated cell cultures have a higher mitochondrial membrane activity (red)

→ Application for DS/WT is still ongoing

ATTRIBUTIONS

Occlusion Example:



AUTOENCODER