Dark Substructure Sensitivity in the Euclid Survey with Machine Learning

Conor O'Riordan

IAP Colloquium | 22.10.2021

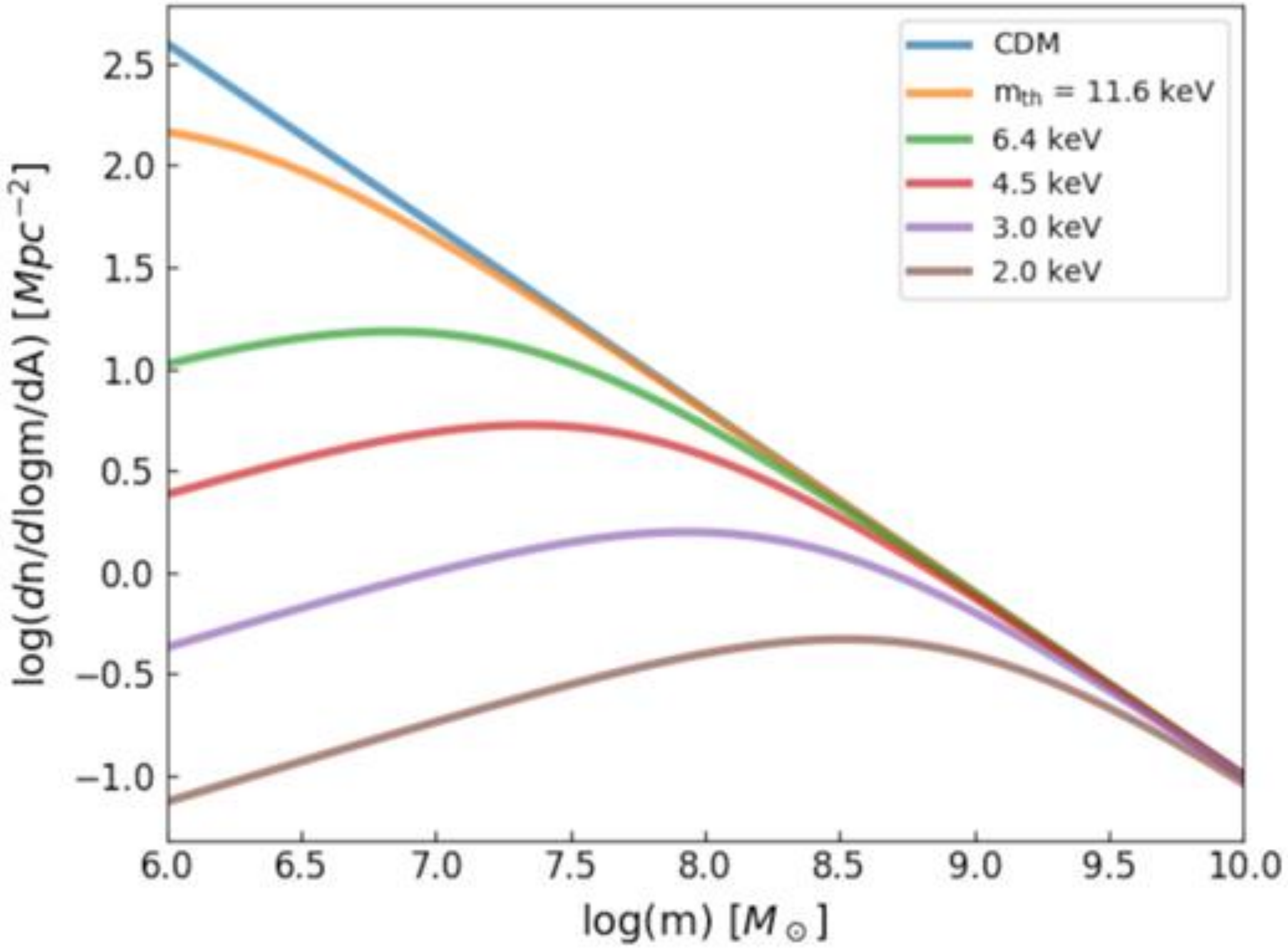
MAX PLANCK INSTITUTE
FOR ASTROPHYSICS

Background: Dark Matter

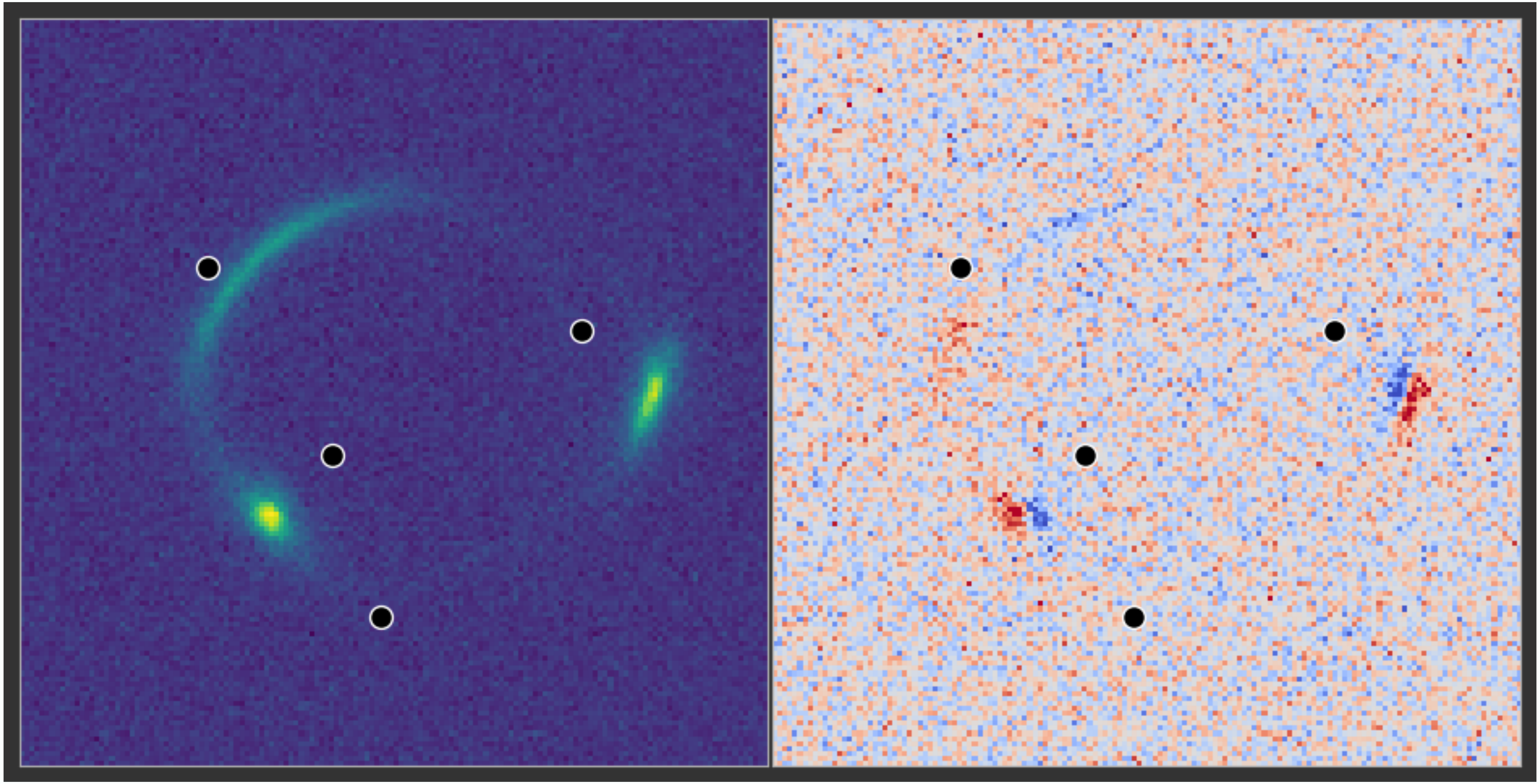
- The subhalo mass function depends on the temperature of DM
- Warmer models suppress structure formation below a certain mass

Hsueh et al (2019)

Expected number



Subhalo mass



Motivation: Upcoming Surveys

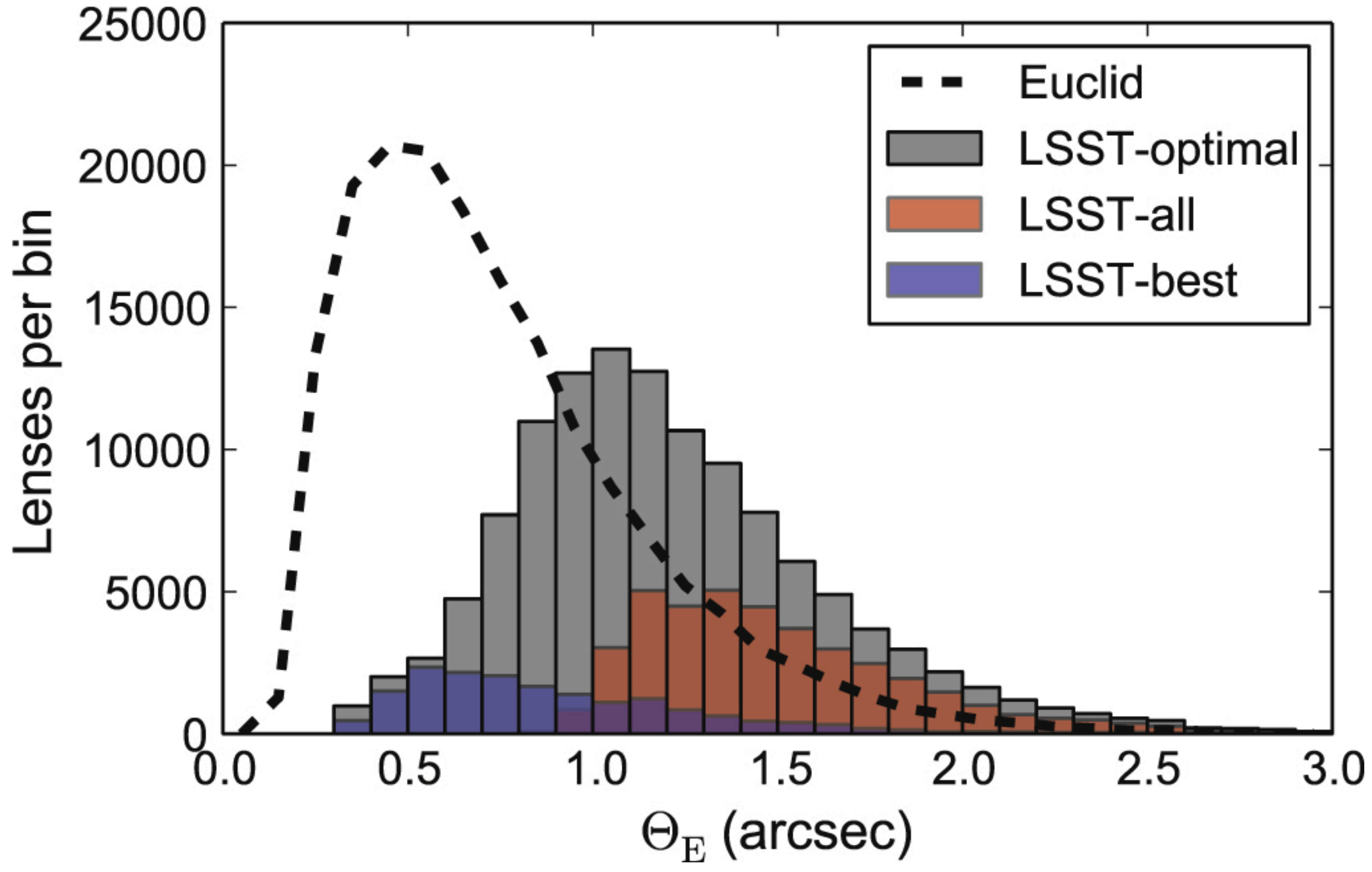
Currently known strong lenses

~10^2

Euclid, *DES* and *LSST* will increase this to

~10^5

Collett (2015)



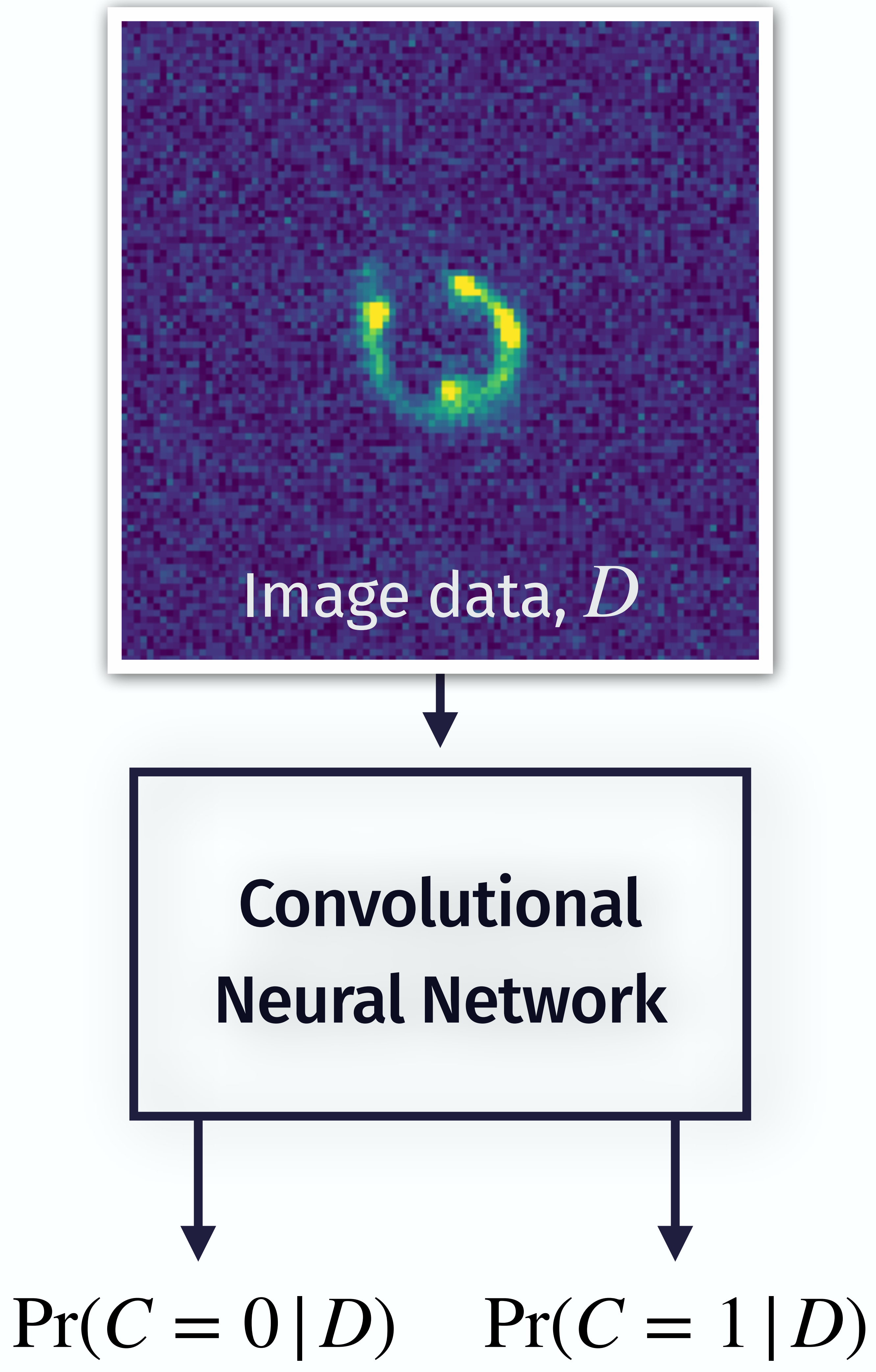
Two possible routes for machine learning to improve gravitational imaging in the era of large numbers of lenses...

Rank lenses prior to modelling to make the best use of our time

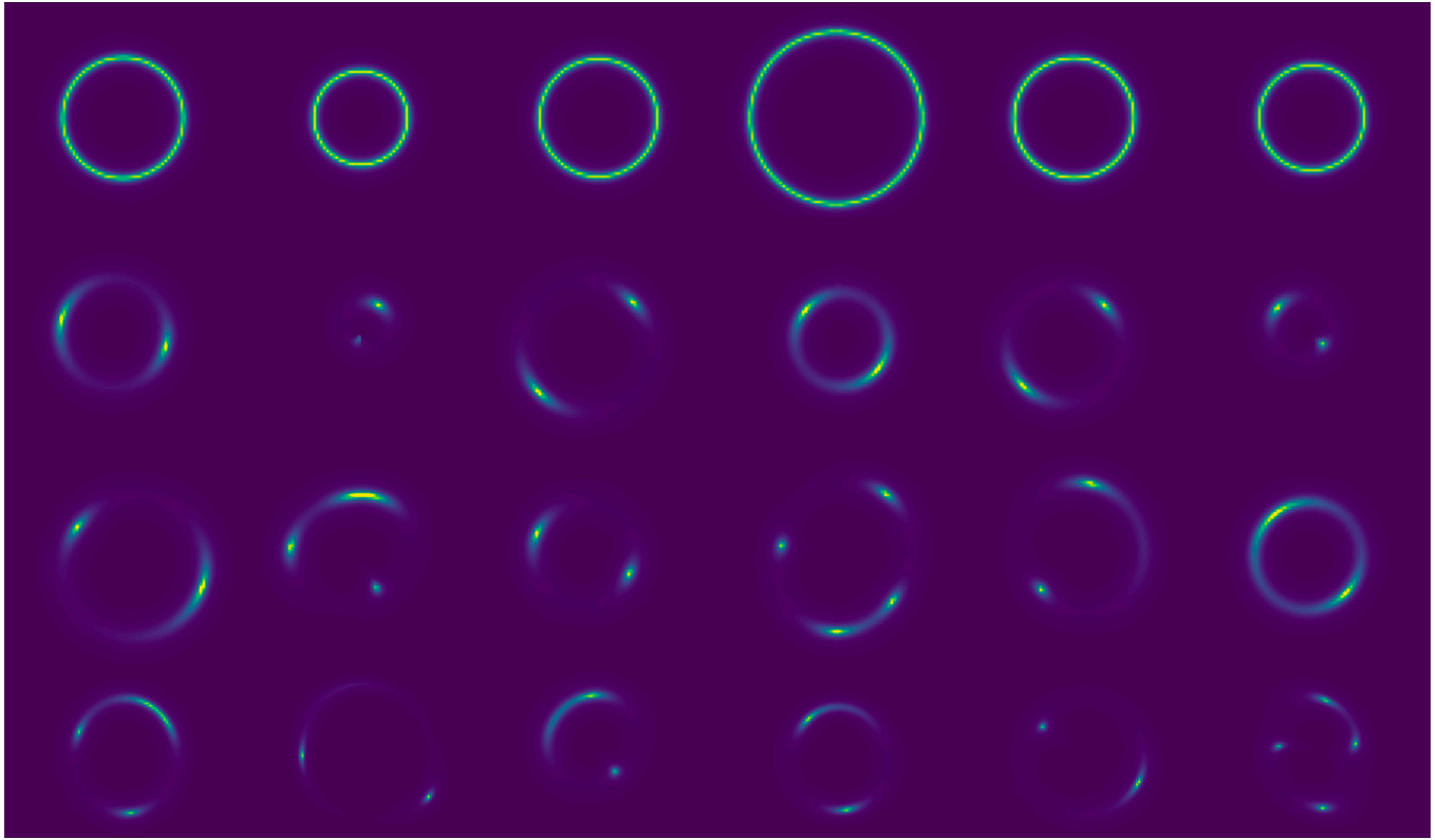
Approximate the modelling with ML to get through more data

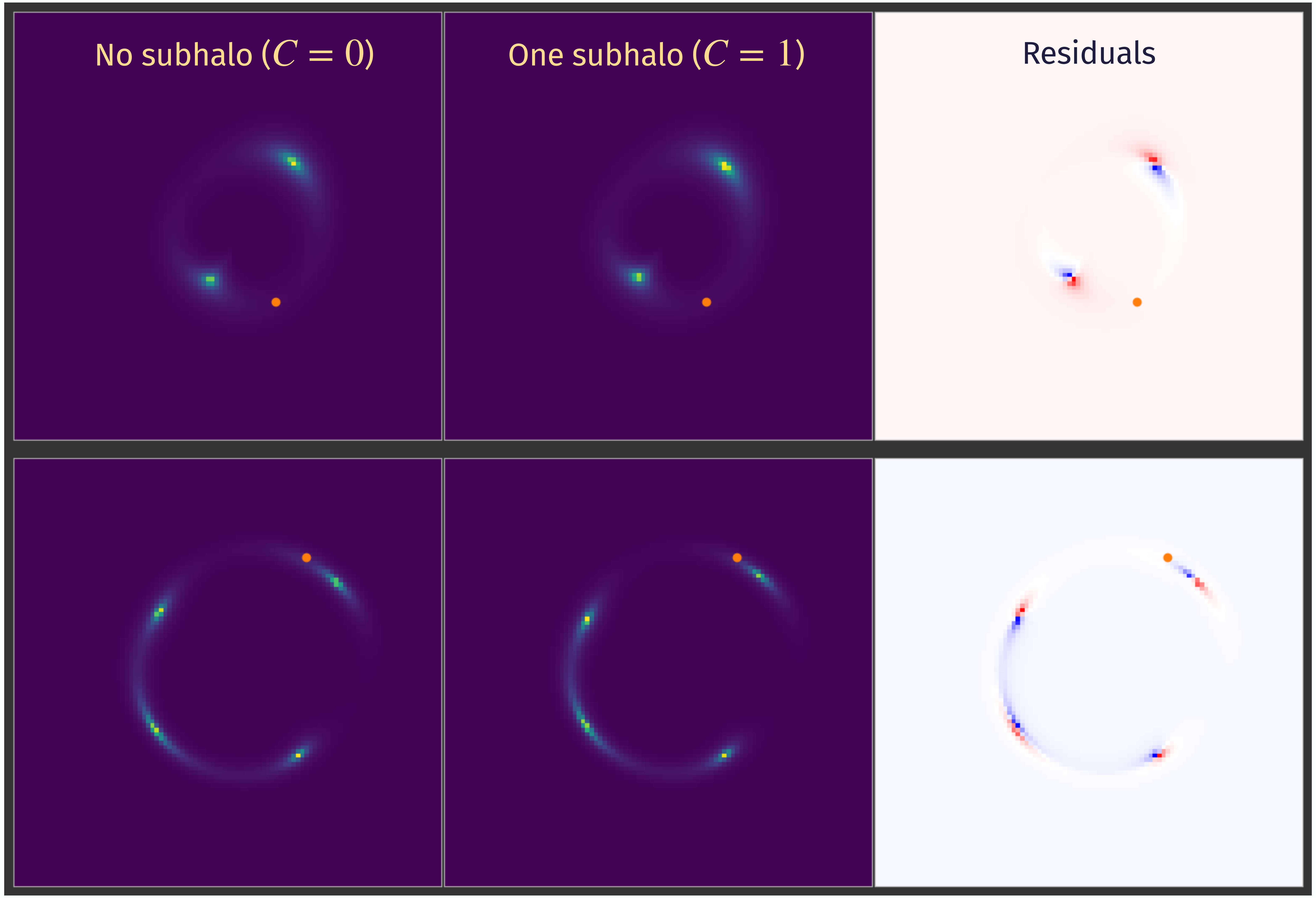
Machine Learning

- We use ResNet to predict the binary presence or absence of substructure in an image
- We start by testing the architecture on increasingly complex levels of (noiseless) data



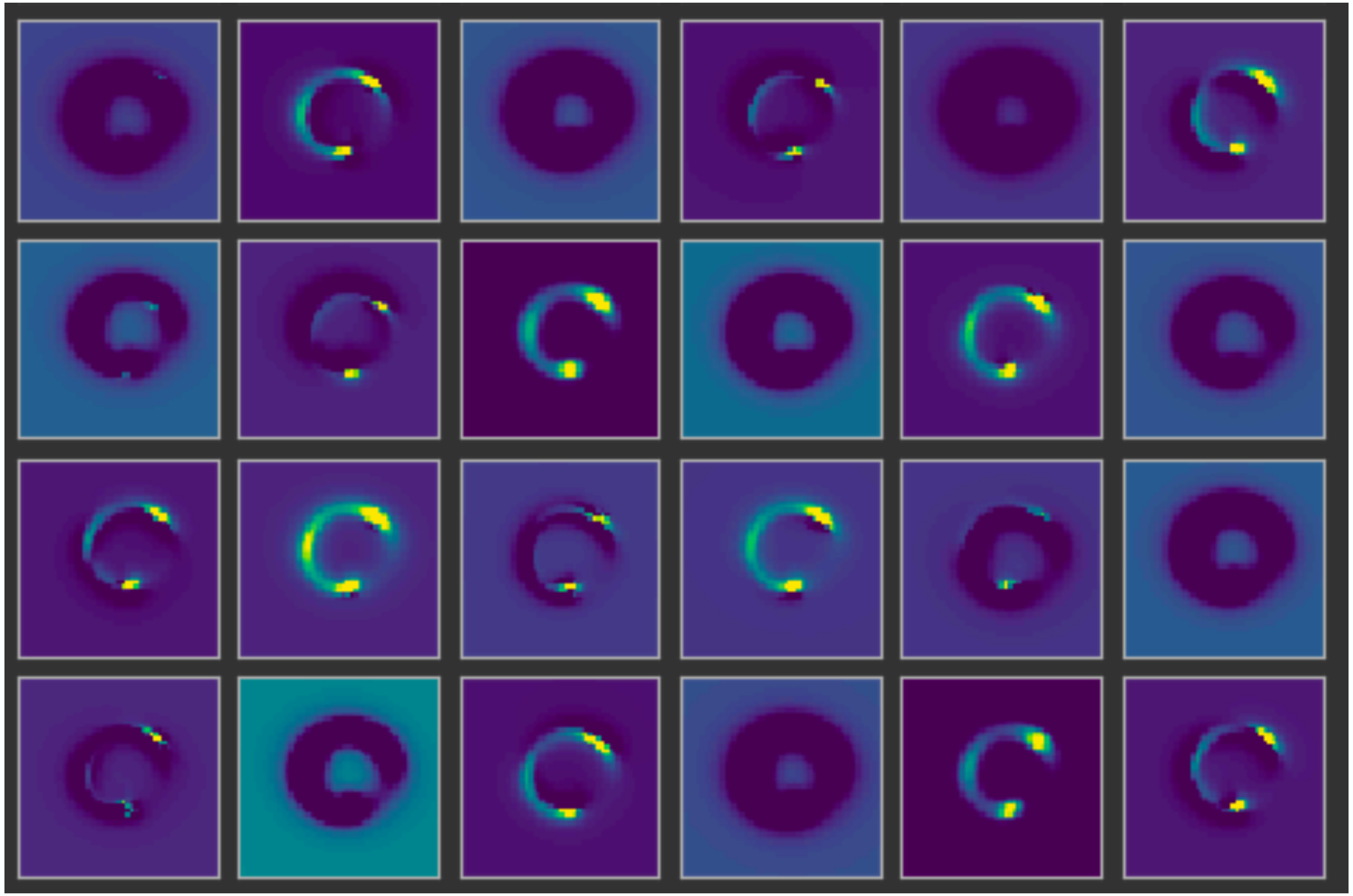
Level	Images	Accuracy
0		>99%
1		99%
2		99%
3		97%





Substructure causes a slight variation in flux across the extended image compared with the smooth model

With noise and a PSF, a source with the right structure can mimic this signal

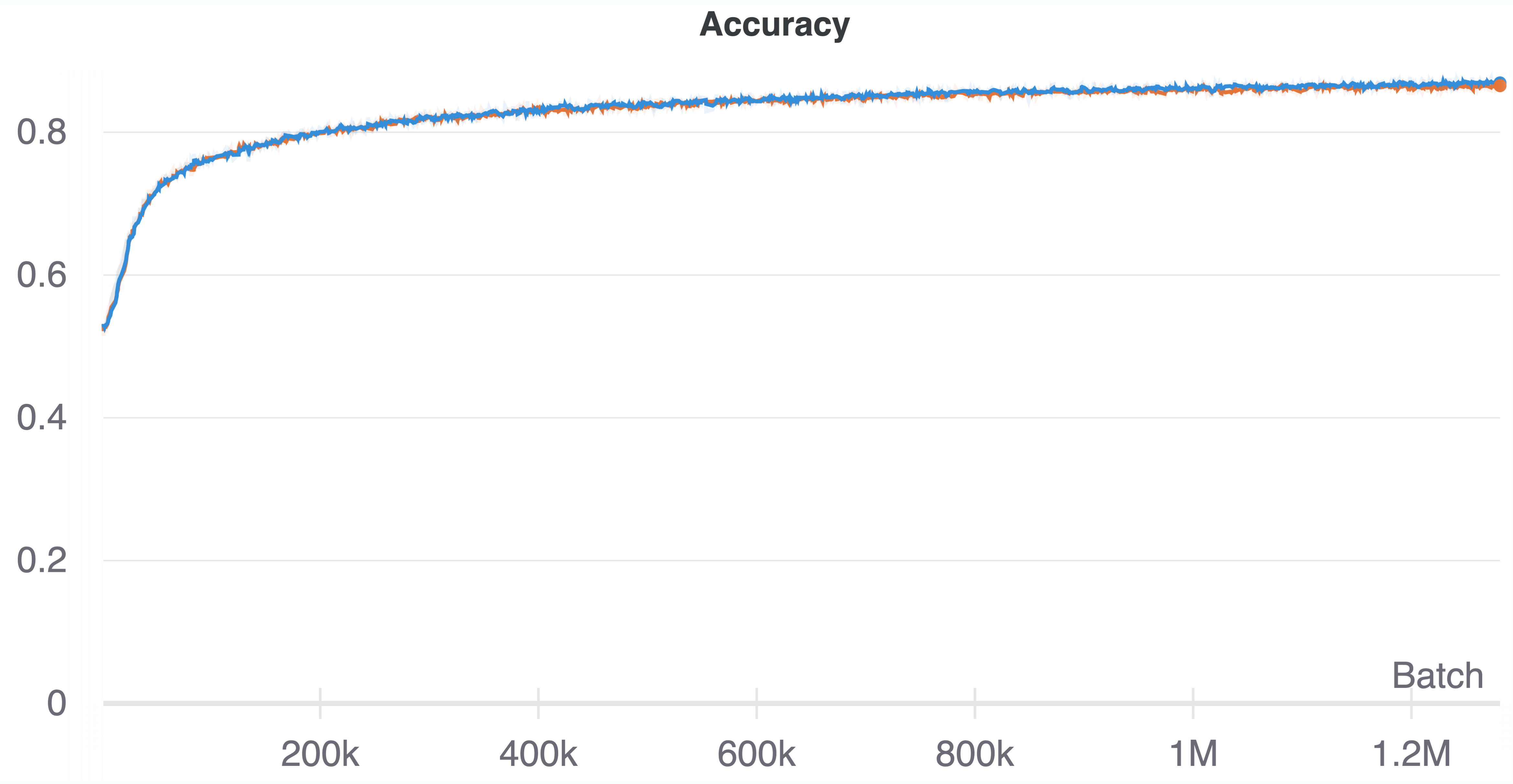


Interpretability?
Activation maps from the first convolution layer shows that the network picks up on this signal

Simulated Euclid data with S/N > 20

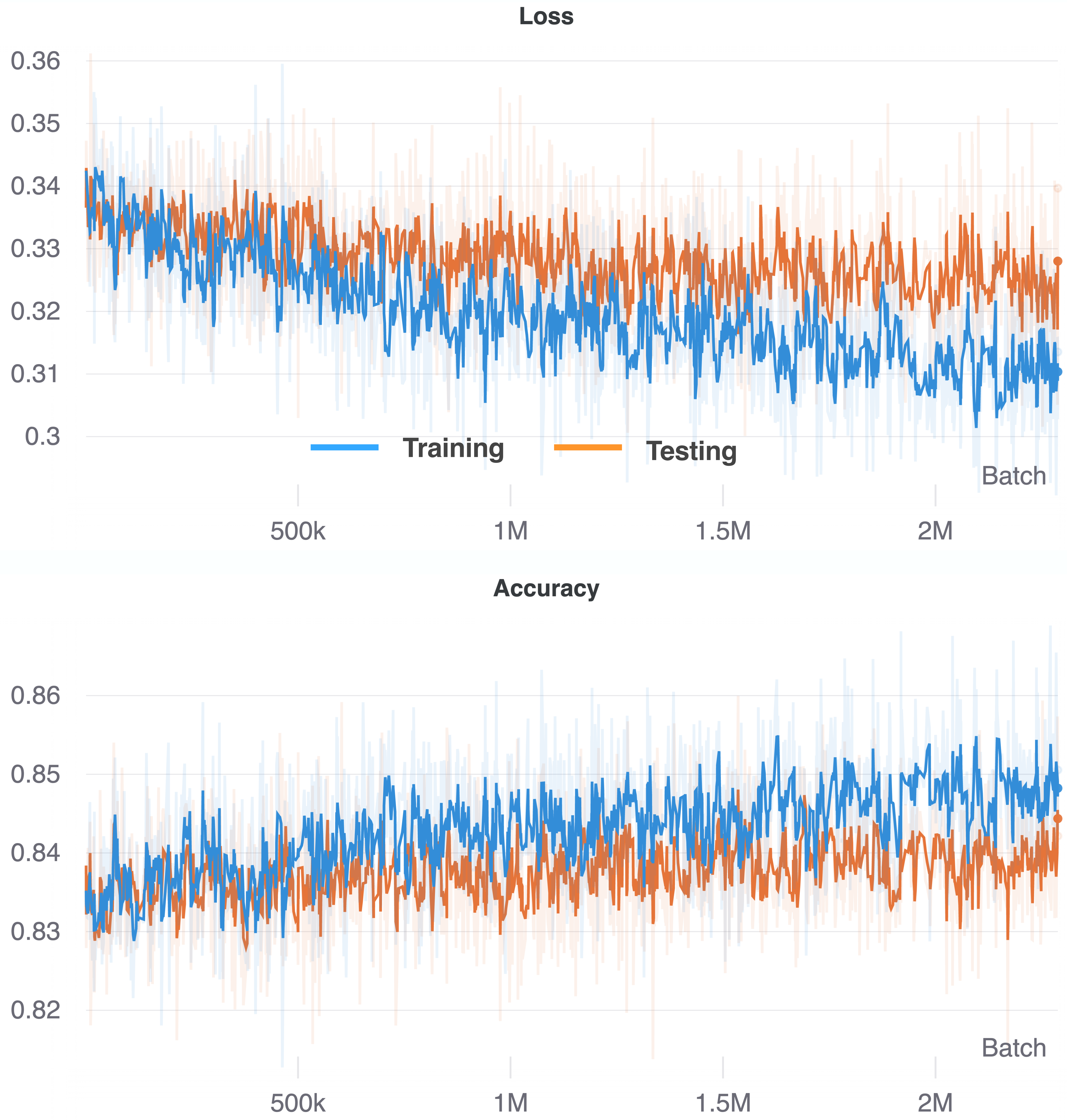
Euclid Initial Results

- We generate approx 1M Euclid images as described
- We add subhaloes from a log-uniform with $10^9 < M_{\text{sub}}/M_{\odot} < 10^{11}$
- Training paused after 70 epochs
- Final accuracy:
  - 86% (training)
  - 85% (testing)

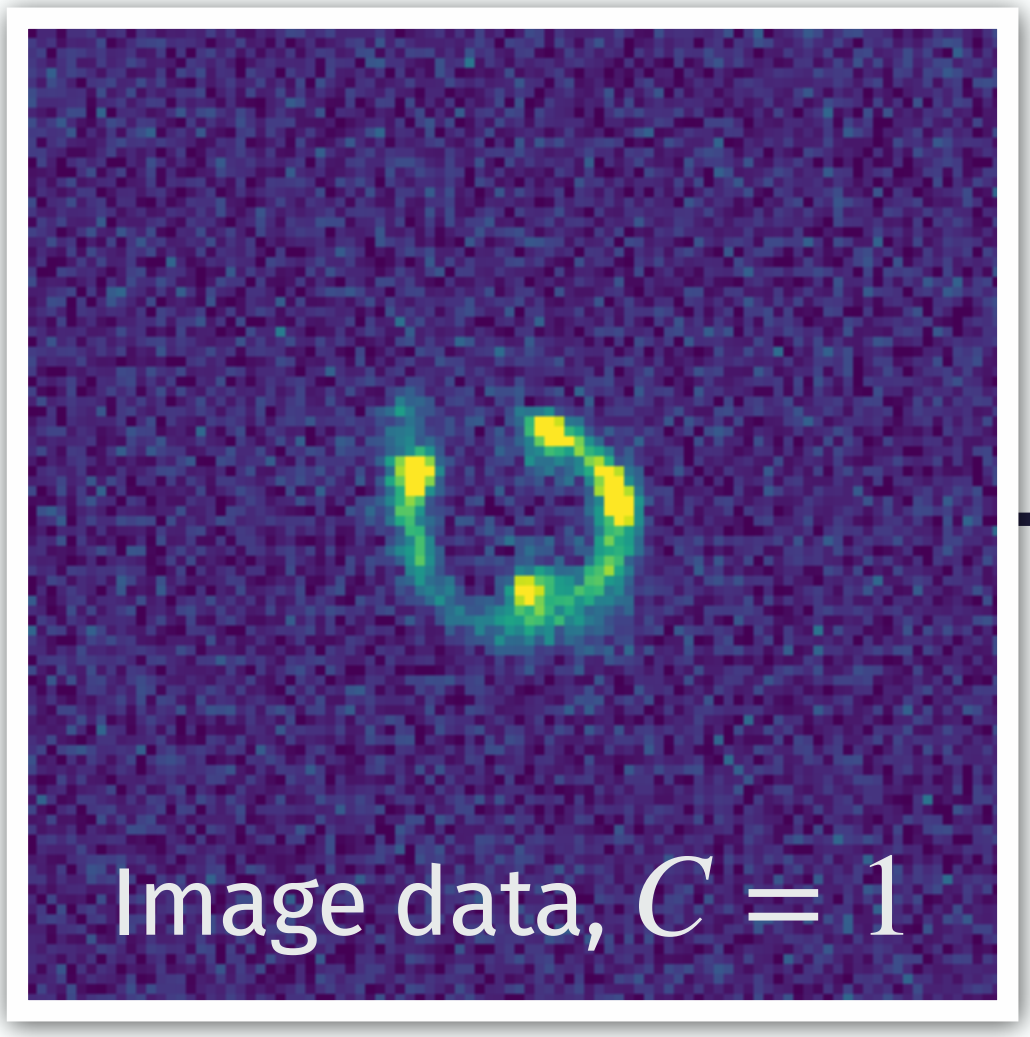


Euclid Initial Results

- After the initial training we expand the range of masses to $10^{8.5} < M_{\rm sub}/M_{\odot} < 10^{11.5}$
- Restart training on the new data but start from previously trained weights
- Final accuracy:
  - 85% (training)
  - 84% (testing)



Recall the network output...



Convolutional Neural Network

$\Pr(C = 0 \,|\, D)$

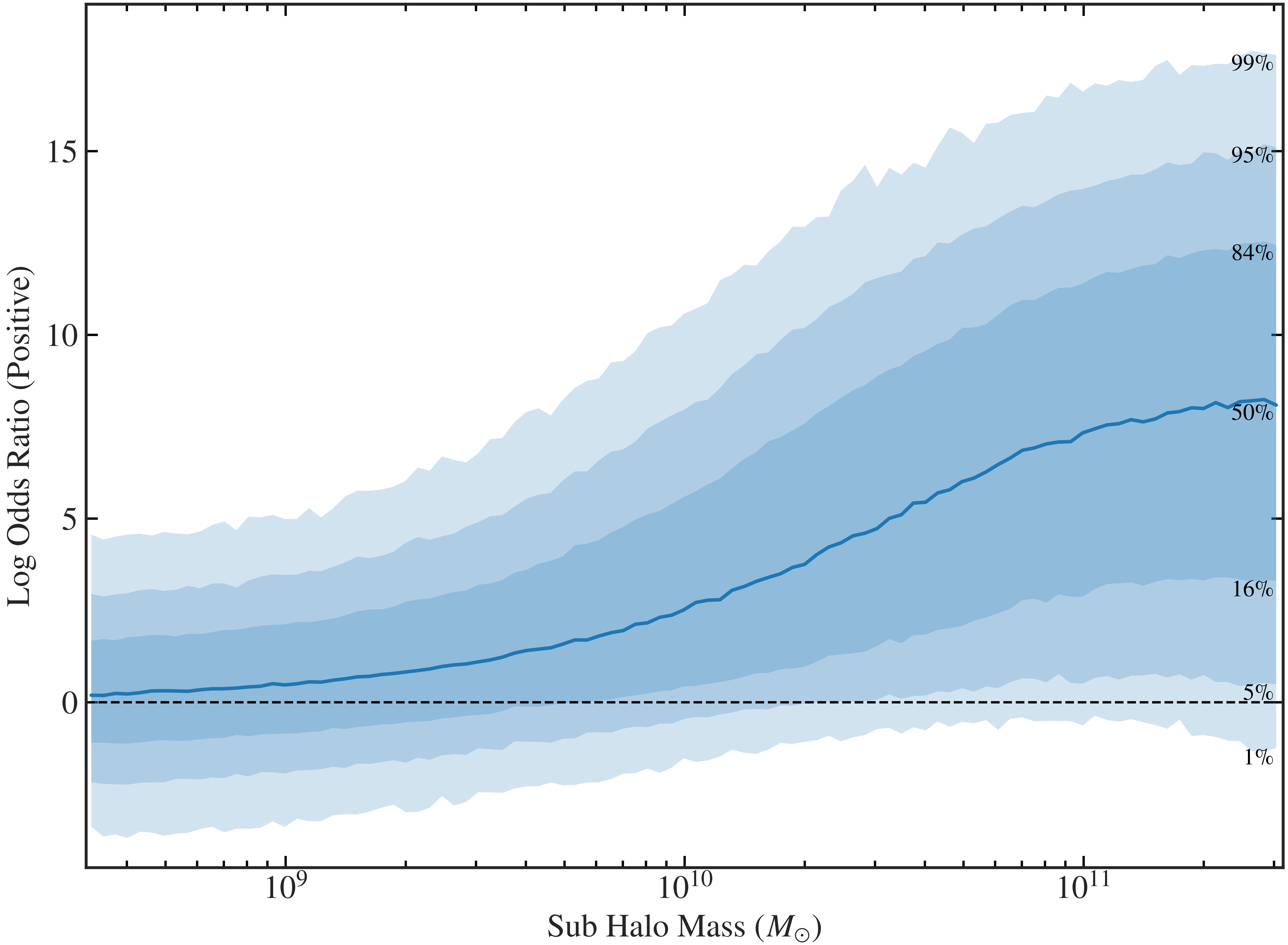
$\Pr(C = 1 \,|\, D)$

Images **that contain subhaloes** with $\log\left(\frac{\Pr(C = 1 \,|\, D)}{\Pr(C = 0 \,|\, D)}\right) > 0$ are correct predictions

We compute the *positive* log odds ratio for all *positive* images in the **testing set**.

We bin the images by subhalo mass and plot here the distribution of odds per bin

For example, a subhalo of $10^{10} M_\odot$ is correctly observed ~90% of the time



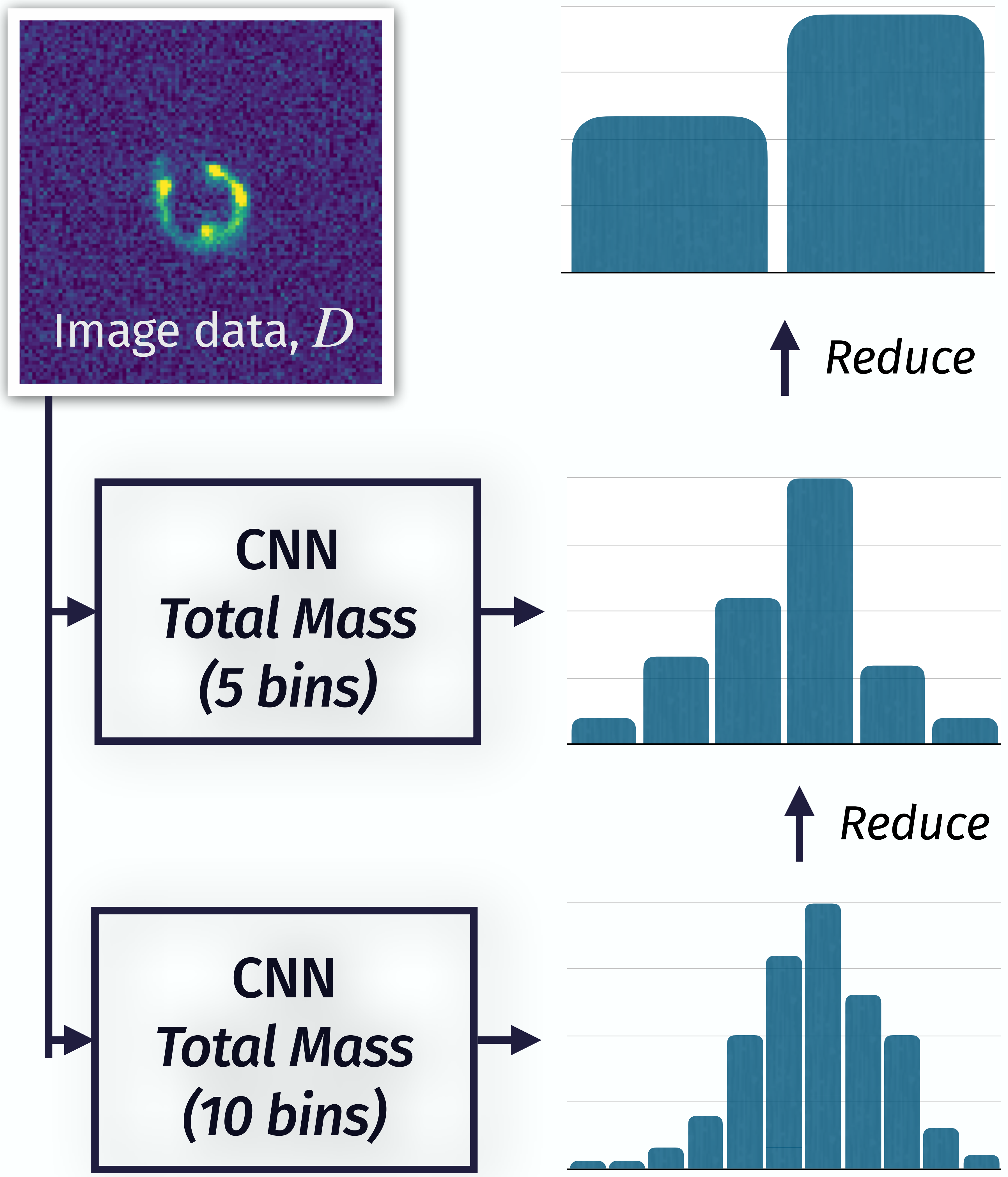
Recovering the Mass Function

- We now fix the range of subhalo mass to $10^9 < M_{\rm sub}/M_\odot < 10^{11}$
- We extend the method to deal with multiple subhaloes per image
- Images are labelled by the **total mass in substructure**

Label	Total Mass ($M_\odot$)
0	No Mass
1	$10^9 \leq M_{\rm sub} < 10^{9.5}$
2	$10^{9.5} \leq M_{\rm sub} < 10^{10}$
3	$10^{10} \leq M_{\rm sub} < 10^{10.5}$
4	$10^{10.5} \leq M_{\rm sub} < 10^{11}$
5	$10^{11} \leq M_{\rm sub}$

Recovering the Mass Function

- Can we add complexity without losing performance on the simpler task?
- Can these networks perform just as well on the even simpler binary task from earlier?
- How complex does the *model* need to be to reach the sensitivity limit of the data?



Different model accuracies with 10 mass bins

Model	**Top 1** Correct class has highest probability	**Top 3** Correct class in top 3 highest probabilities	**Binary** Presence or absence of any substructure
ResNet 50	36.1%	68.8%	84.0%
ResNet 101	37.2%	70.1%	84.2%

A model twice as large makes little difference in performance

ResNet 50 accuracies with different binning

Model	**Top 1** Correct class has highest probability	**Top 3** Correct class in top 3 highest probabilities	**Binary** Presence or absence of any substructure
5 mass bins	56.5%	86.9%	83.4%
10 mass bins	36.1%	68.8%	84.0%
10 bin model with 5 bin data	55.9%	87.7%	84.0%

Model trained on 10 bin data performs as well on 5 bin data as 5 bin model
Both models match binary accuracy of earlier 18 layer model

Conclusions/outlook

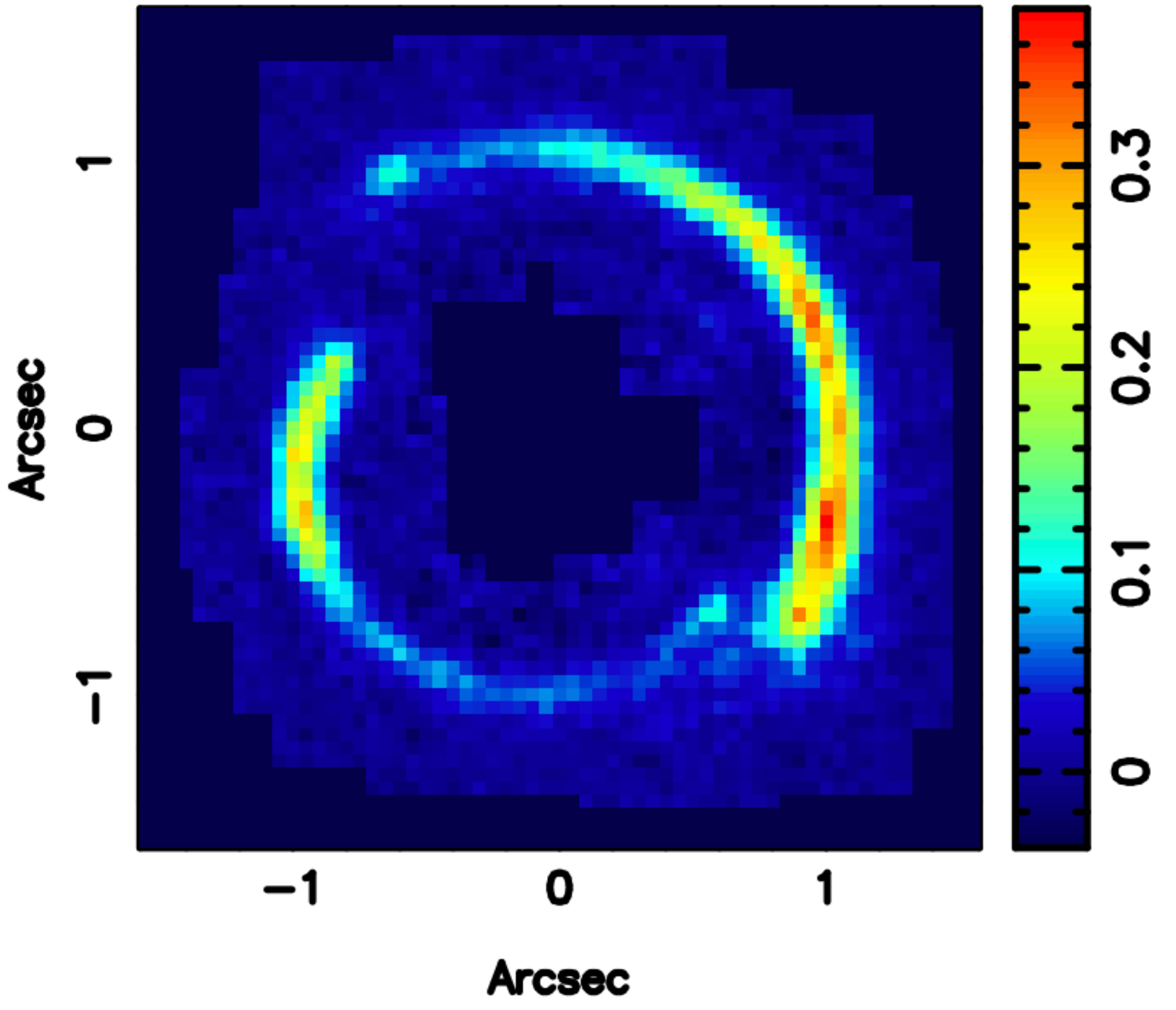
- Euclid can reliably detect subhaloes of $M \sim 10^{10} M_{\odot}$
- Lower masses are not impossible - depends on your definition of detectable
- We can also reliably predict the total mass in multiple substructures with similar limits
- Running predictions for different DM models will tell us the available constraints on e.g. M_{HM} from N Euclid strong lenses
- Complex models can perform as well as their counterparts on simpler tasks

Extra slides...

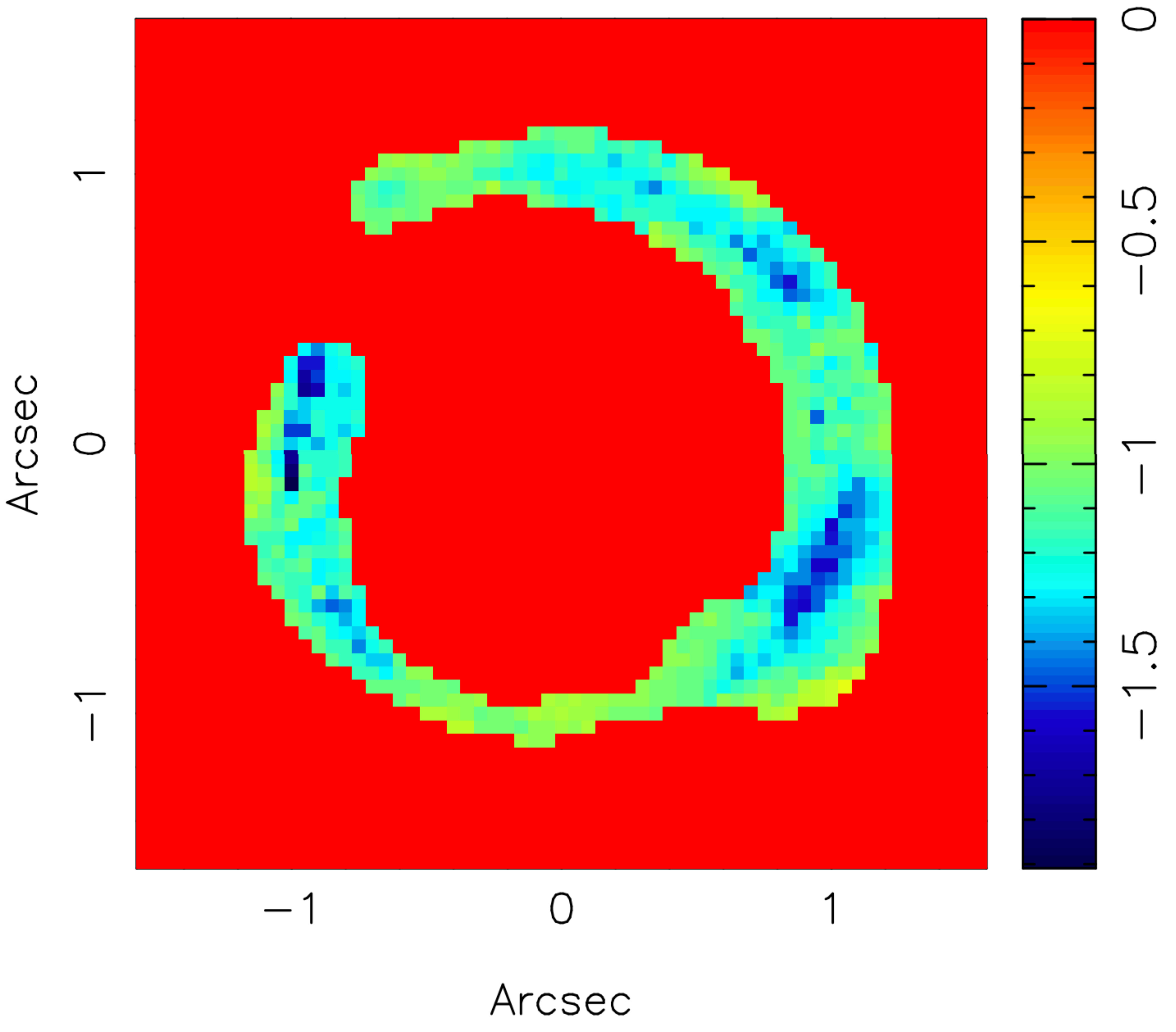
Gravitational imaging in J0252+0039

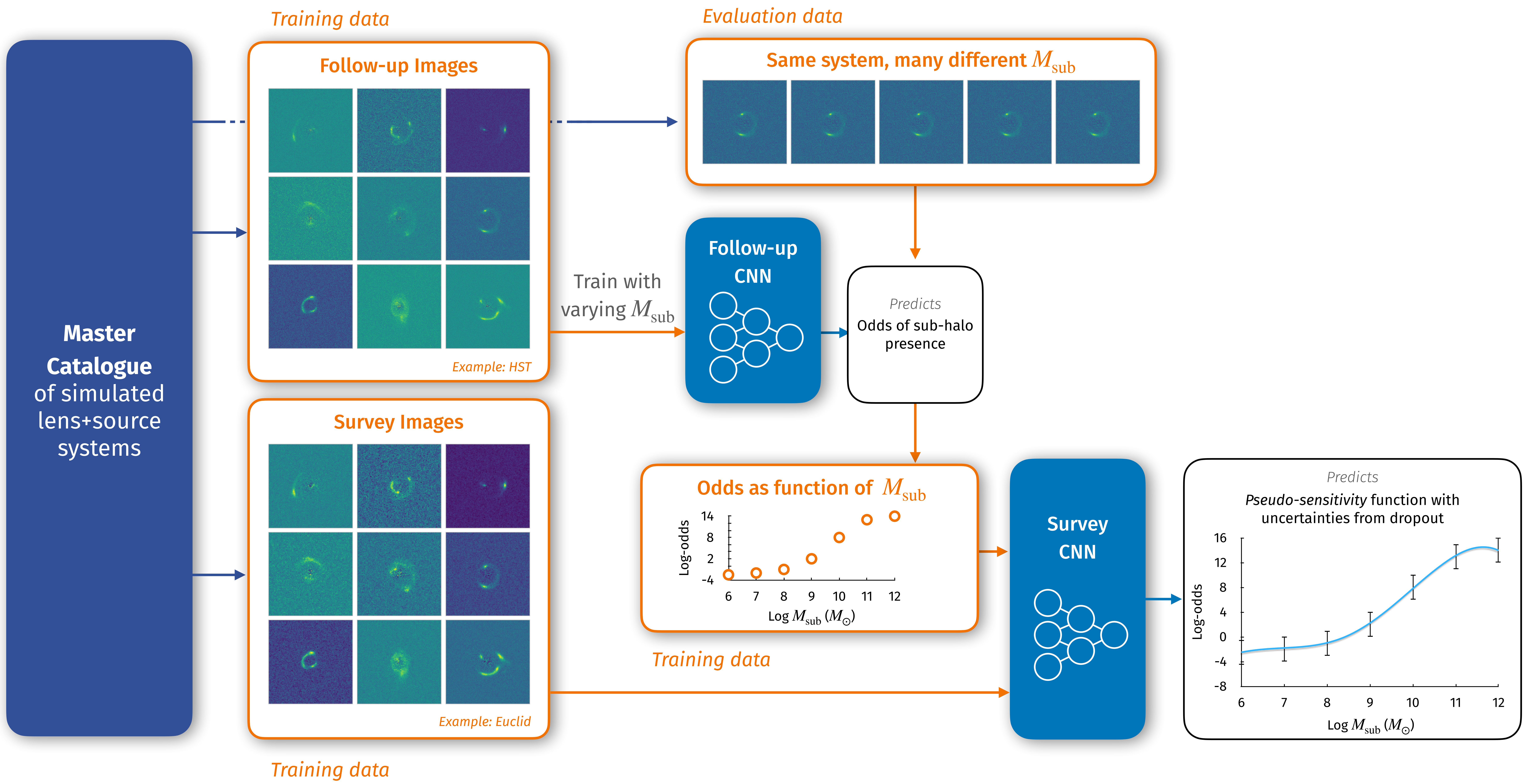
Vegetti et al (2014)

Image



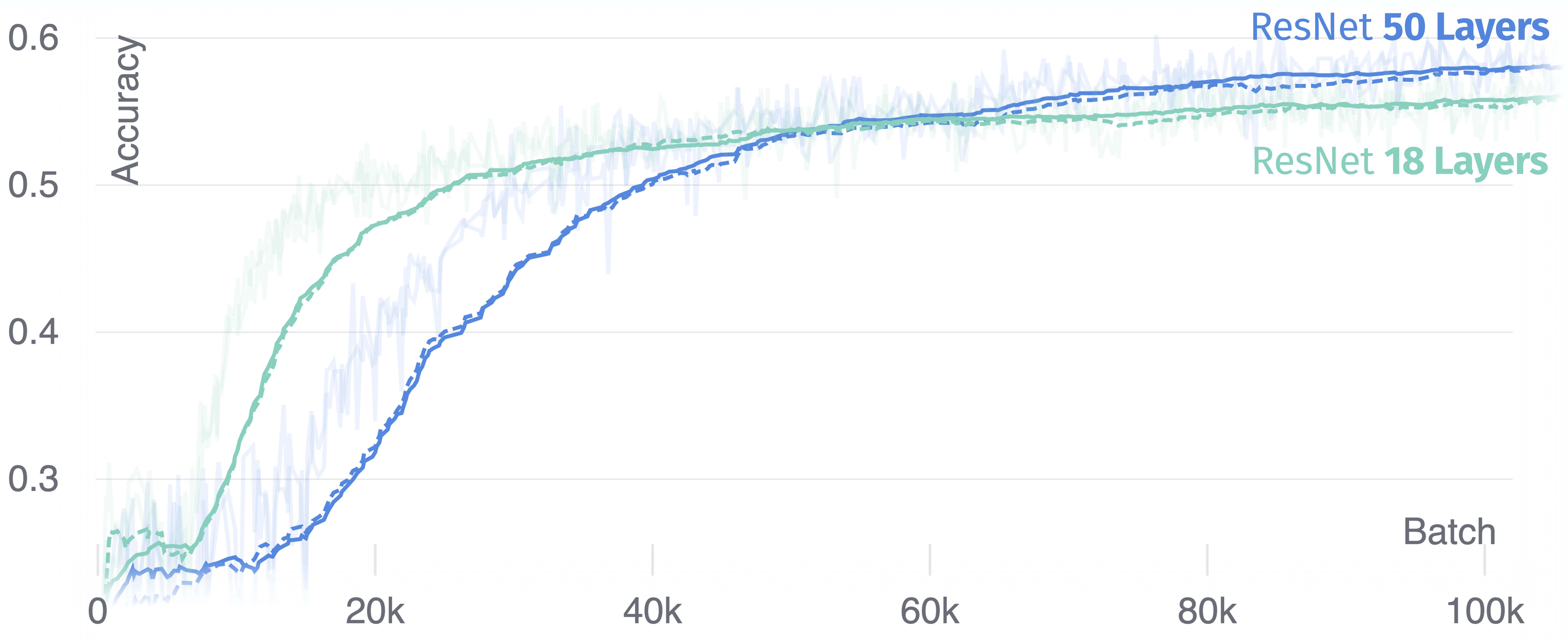
Sensitivity $(M/10^{10}M_\odot)$

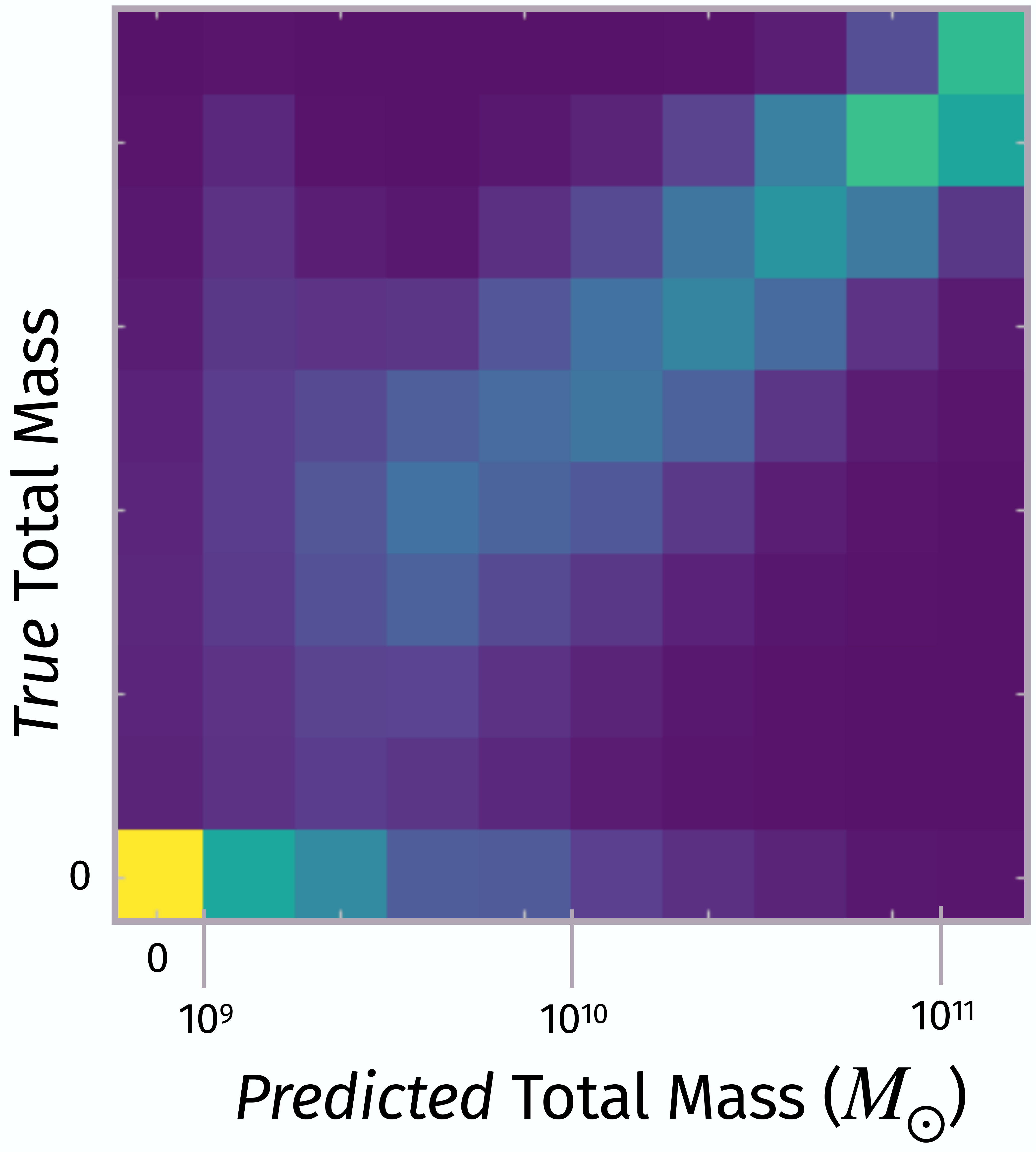




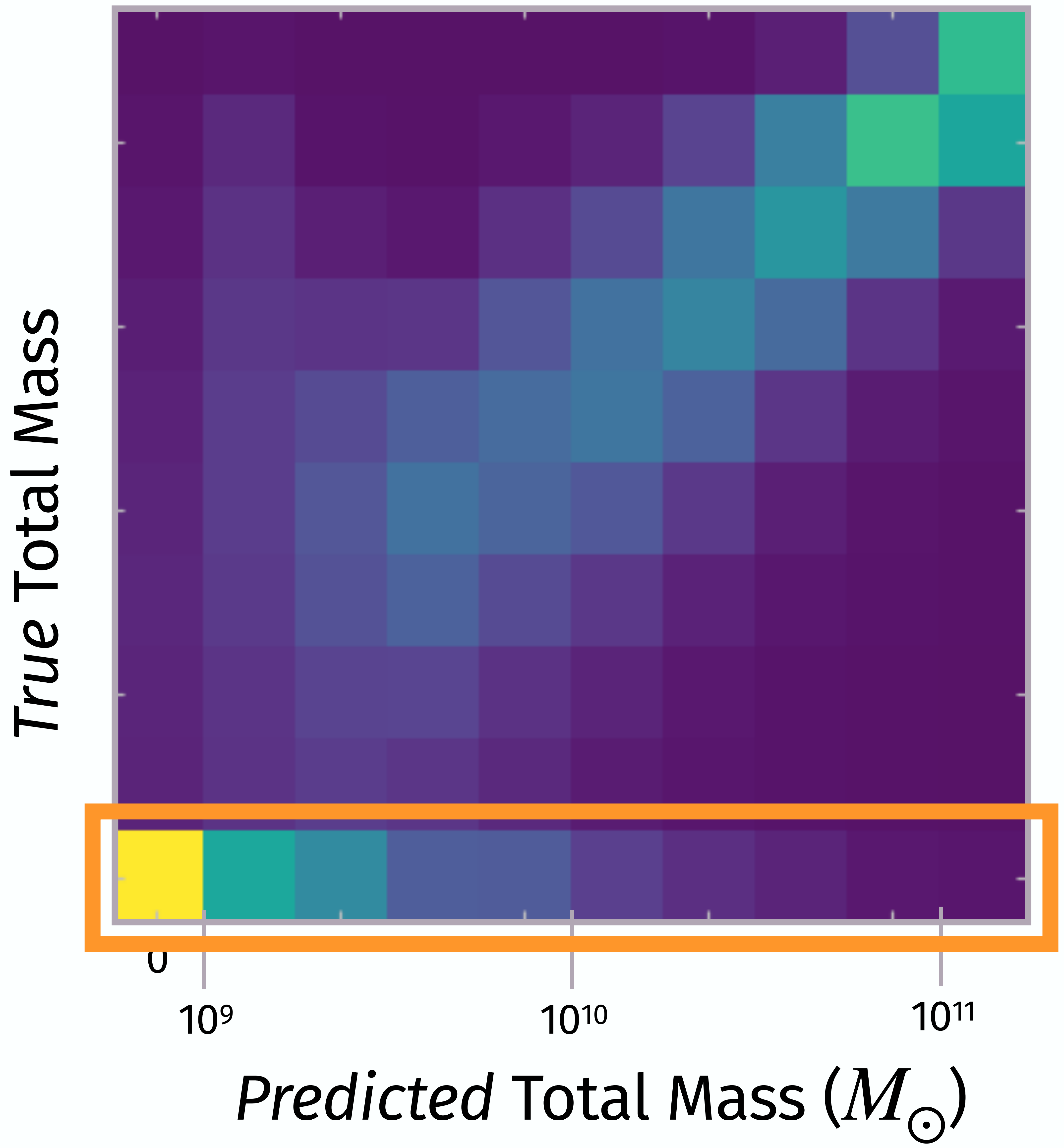
Training data
Evaluation data
Master Catalogue of simulated lens+source systems
Follow-up Images
Example: HST
Same system, many different $M_{\rm sub}$
Train with varying $M_{\rm sub}$
Follow-up CNN
Predicts
Odds of sub-halo presence
Survey Images
Example: Euclid
Training data
Odds as function of $M_{\rm sub}$
Log-odds
Log $M_{\rm sub}$ ($M_\odot$)
Training data
Survey CNN
Predicts
Pseudo-sensitivity function with uncertainties from dropout
Log-odds
Log $M_{\rm sub}$ ($M_\odot$)

Effect of model complexity

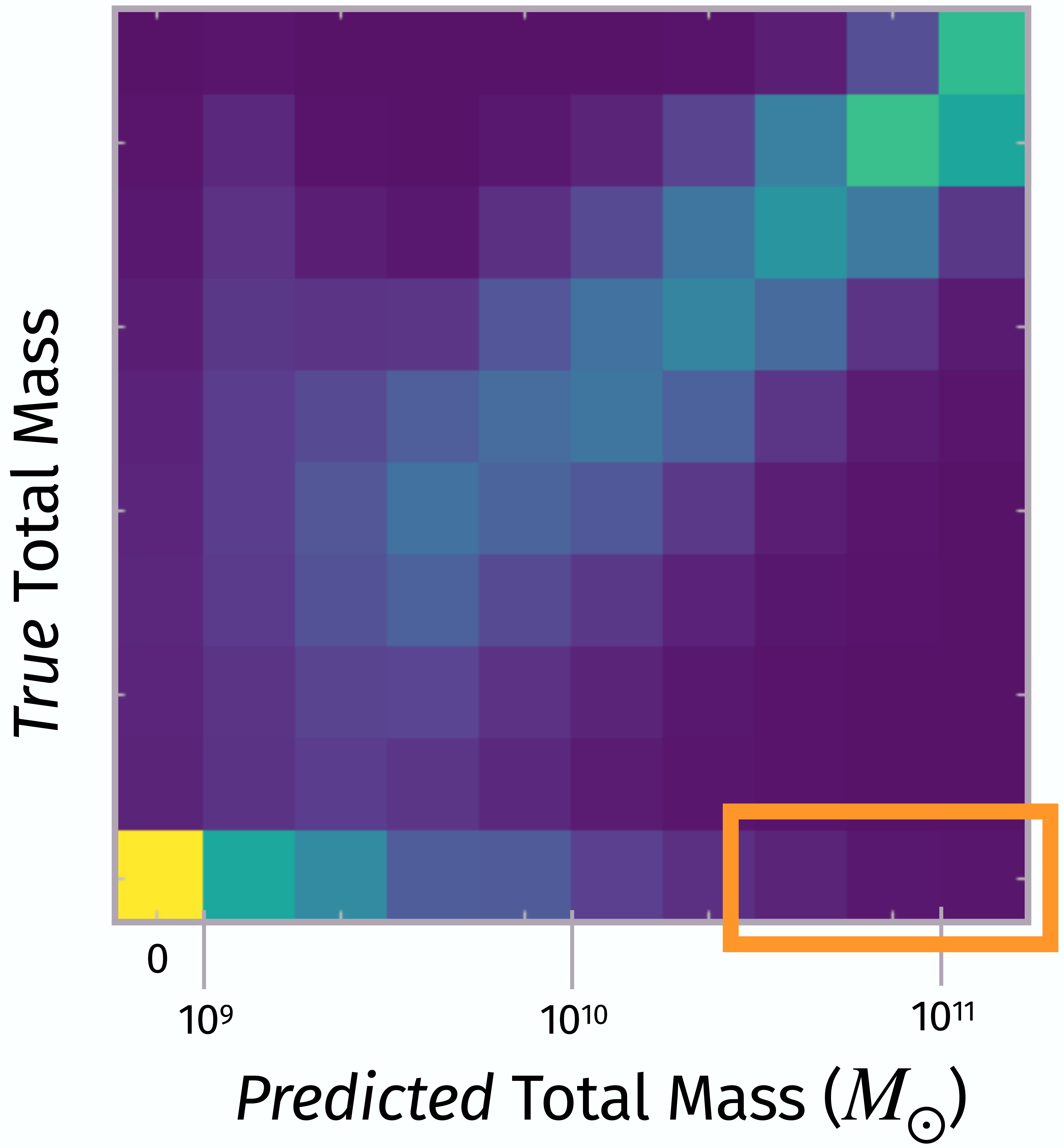




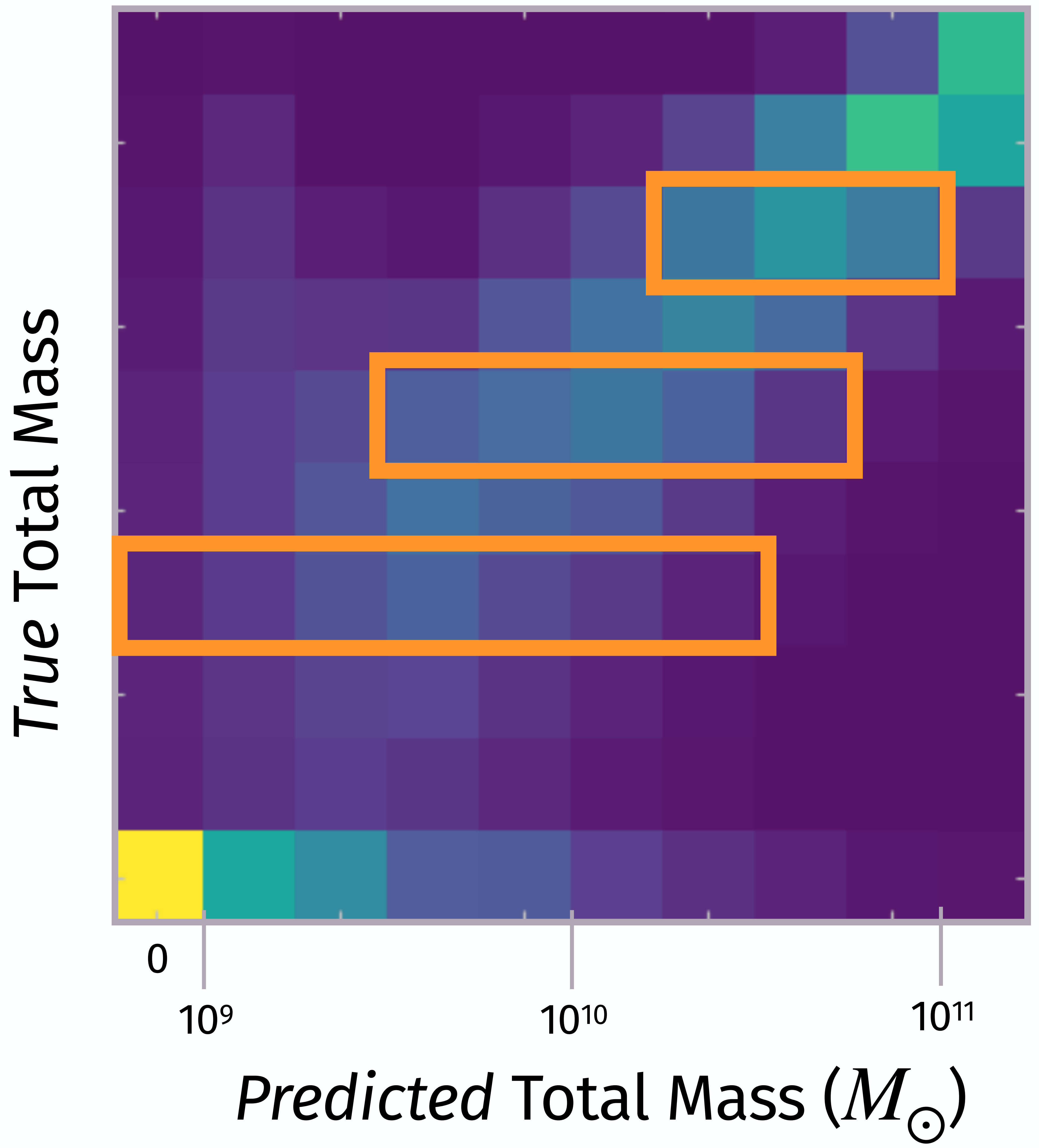
The *confusion matrix* shows the network's top-one classifications versus their true values in the data



Images with no subhaloes are most often classified as such, but are confused for having low masses most often

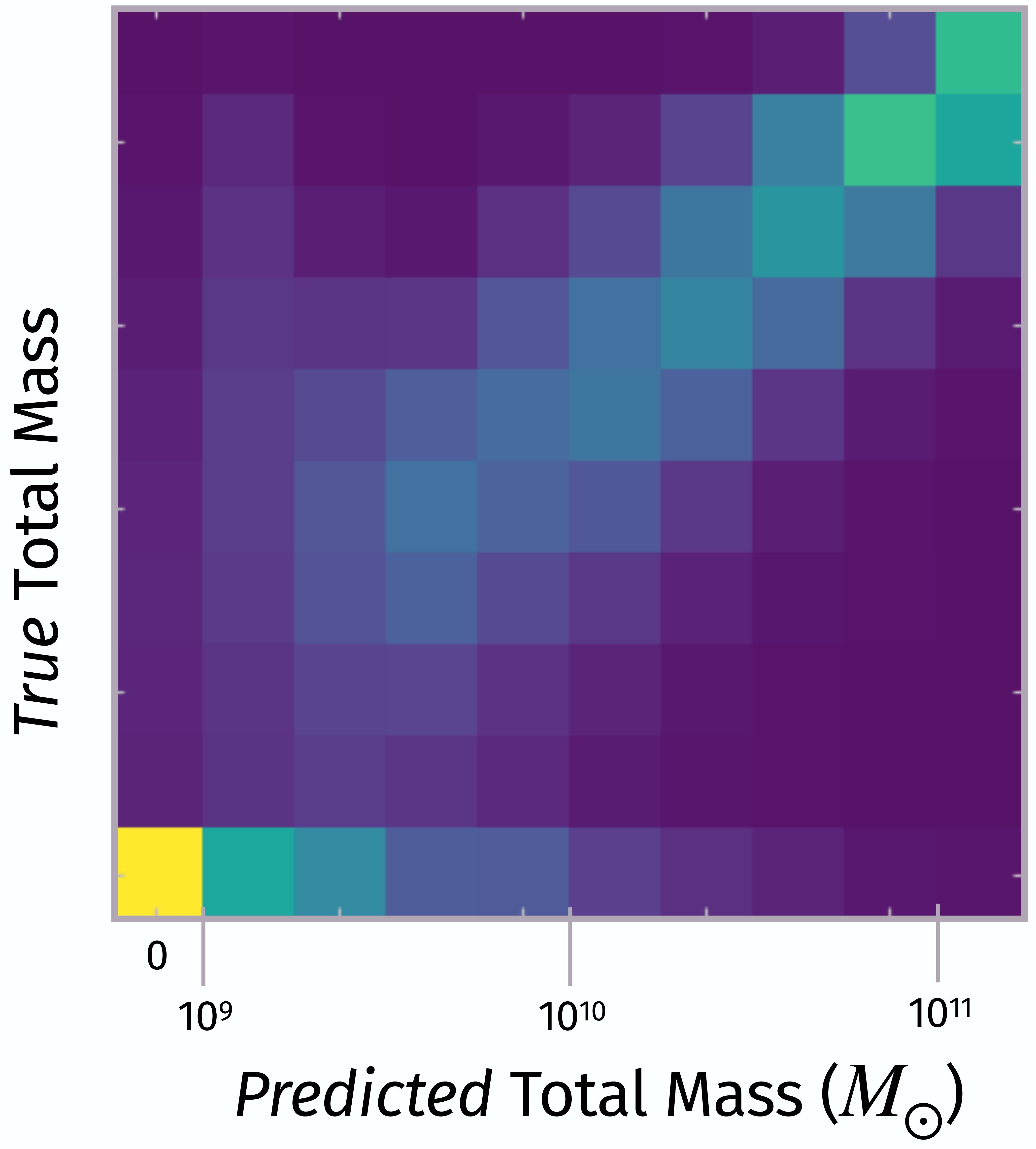


They are also never confused for having large masses



For images with mass the correct bin is the most popular bin in all but two cases

Incorrect predictions are most typically in the neighbouring bin and distributed symmetrically



The precision of predictions increases with subhalo total mass

Sensitivity also depends on S/N in a straightforward way