Galaxy Cluster Mass with Deep Learning and Cosmological Hydrodynamic Simulation

Ziang Yan

@

GCCL Seminar

Syllabus

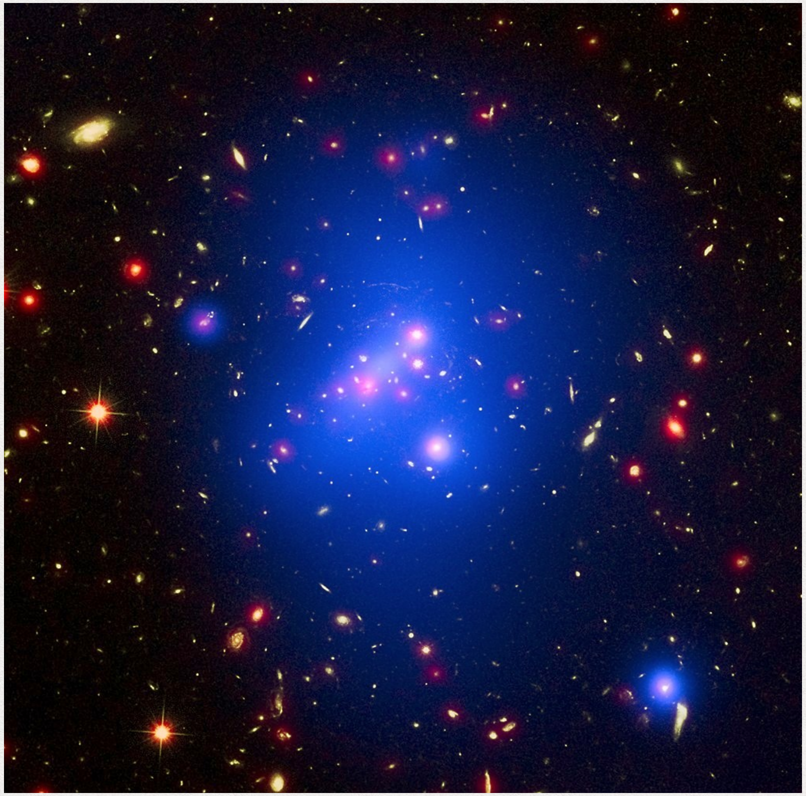
- Background and motivation
- An Introduction to Artificial Neural Network (ANN)
- Estimating galaxy cluster masses with Convolutional Neural Network (CNN)
- An attempt to interpret the CNN
- Summary

Background and motivation

Galaxy clusters

- Galaxy clusters: the most massive collapsed objects

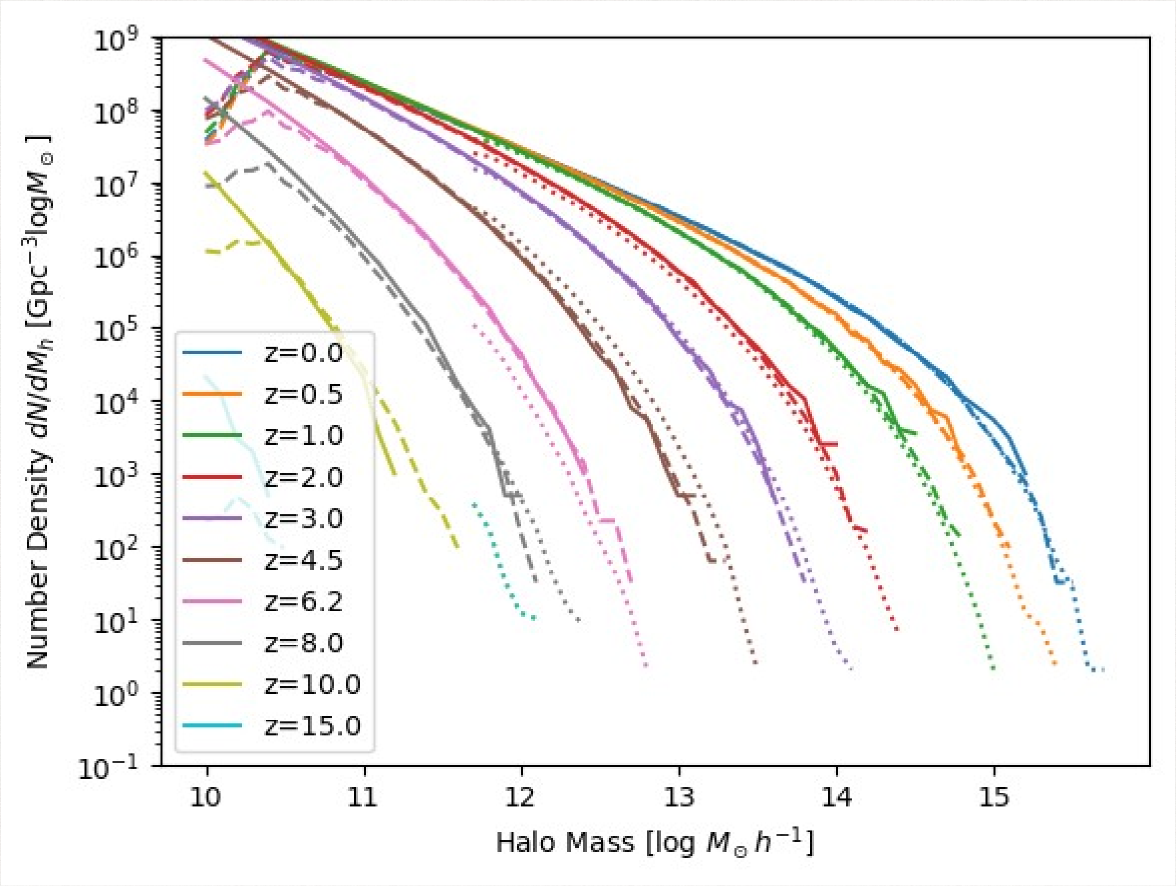
Mass budget:
~1% stars;
~7-13% intergalactic gas;
The rest: dark matter



Galaxy cluster IDCS J1426 , image from Wikipedia

Cluster masses

Why is cluster mass important?
- telling us the evolution of LSS (through mass function);
- depends on cosmological parameters



Halo mass function at different redshift.

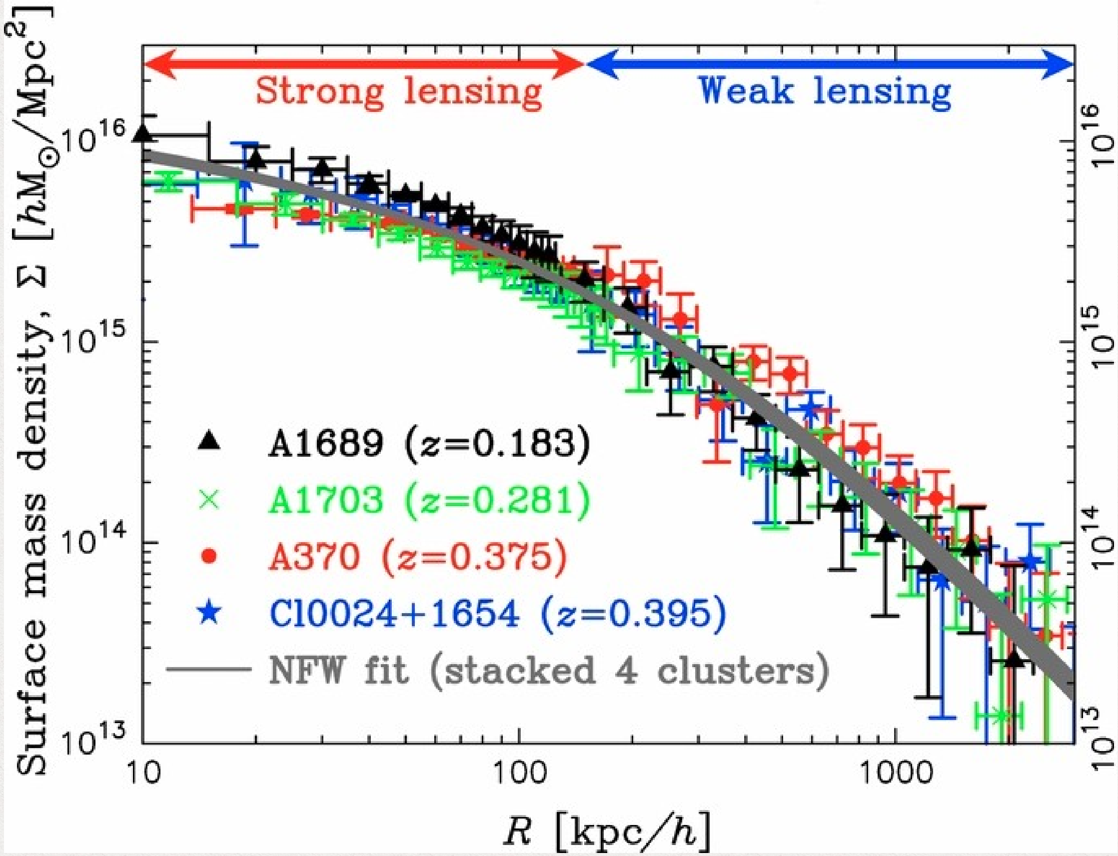
Image from https://github.com/JohannesBuchner/hmf

Traditional mass estimation methods

Traditional estimation of cluster mass: profile fittings

$$\rho(r) = \frac{\rho_0}{\frac{r}{R_s}\left(1 + \frac{r}{R_s}\right)^2}$$

$$M = \int_0^{R_{\max}} 4\pi r^2 \rho(r) \mathrm{d}r$$



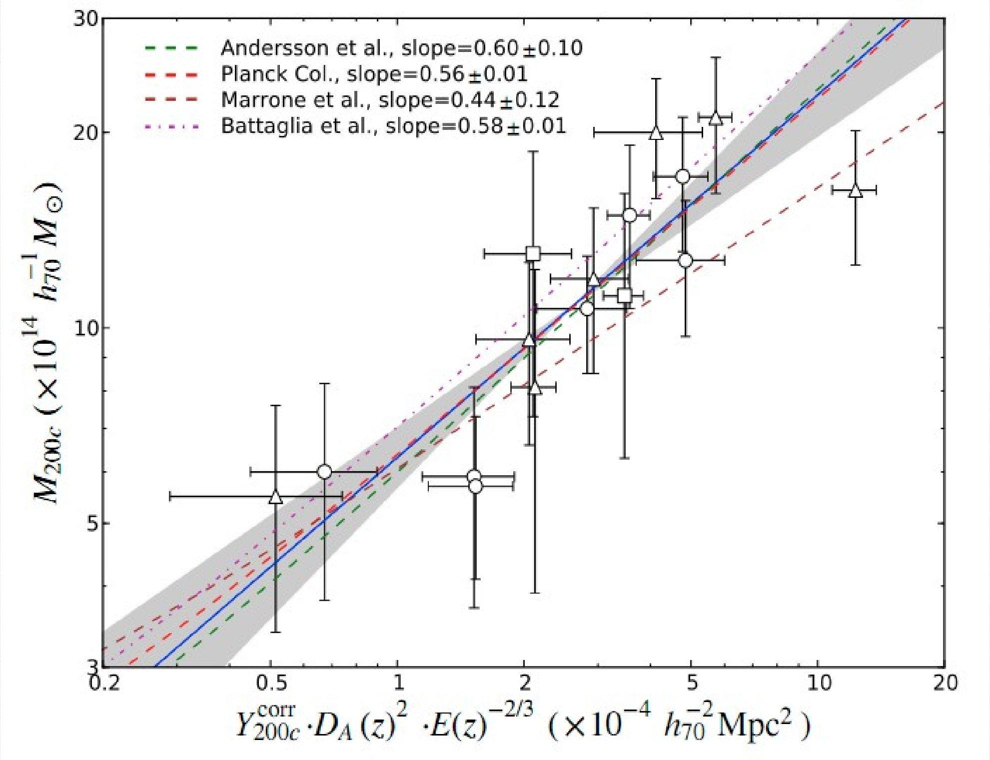
(Postman et al. 2012)

Limitations:
-S/N not enough to measure profile for single cluster
-Substructure information lost!

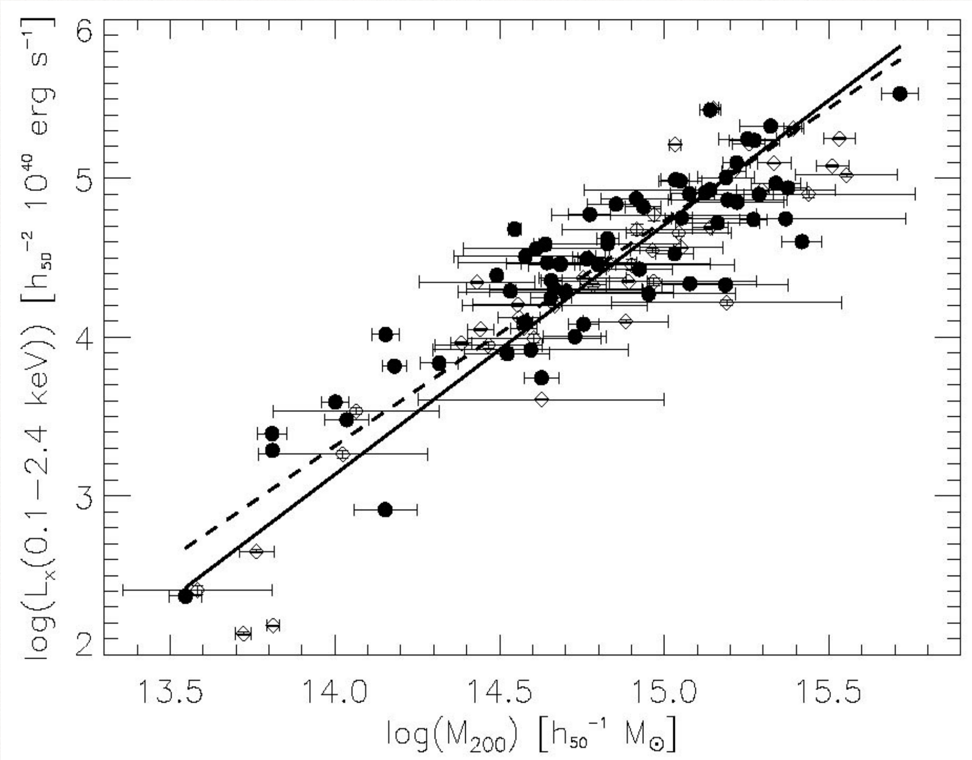
Traditional mass estimation methods

Traditional estimation of cluster mass: self-similar scale relation

$$M \propto L^{\alpha}$$



TSZ effect, Sifon et al. 2013



X-ray luminosity, P. Rosati et al.

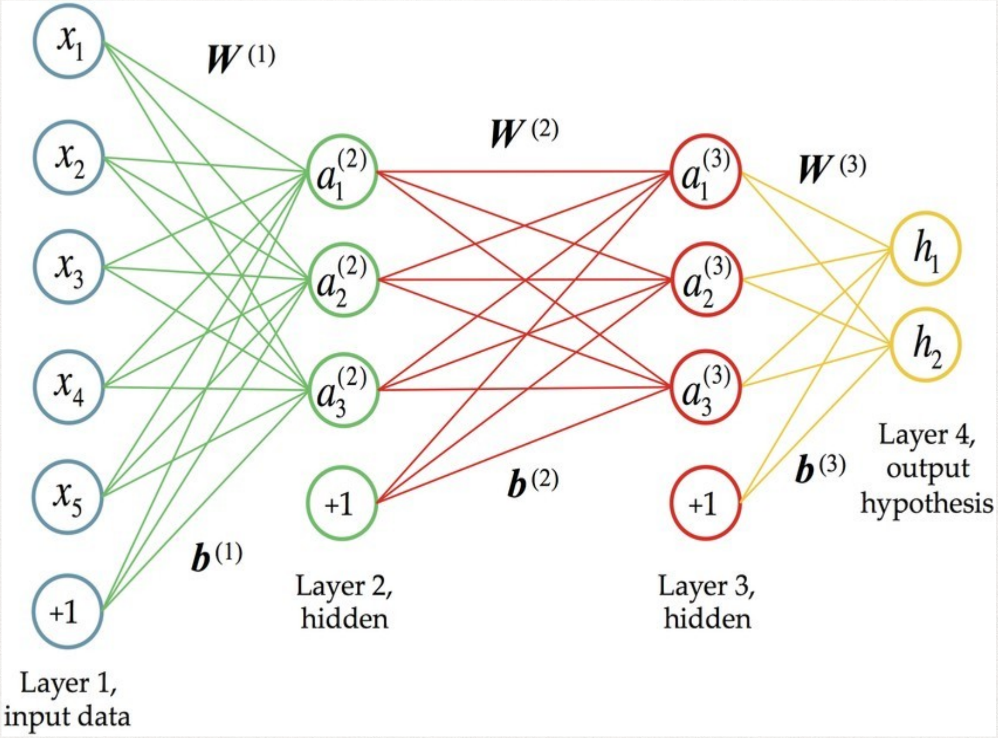
Limitation: bias; no structure information

Can we estimate galaxy cluster mass directly from their 2-D images?

A promising tool: convolutional neural network

An Introduction to Artificial Neural Network (ANN)

ABC of Artificial Neural Network



- ANN is a highly non-linear "function" as a hierarchy of simple "activation functions"

weights

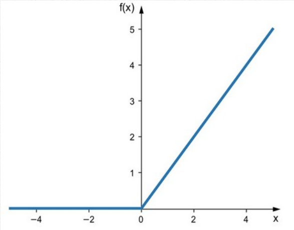
bias

$$a_k^{(l)} = f(W_{jk}^{(l)} a_j^{(l-1)} + b_k^{(l)})$$

$$f(x) = \max(0, x)$$

Rectified Linear Unit, ReLU

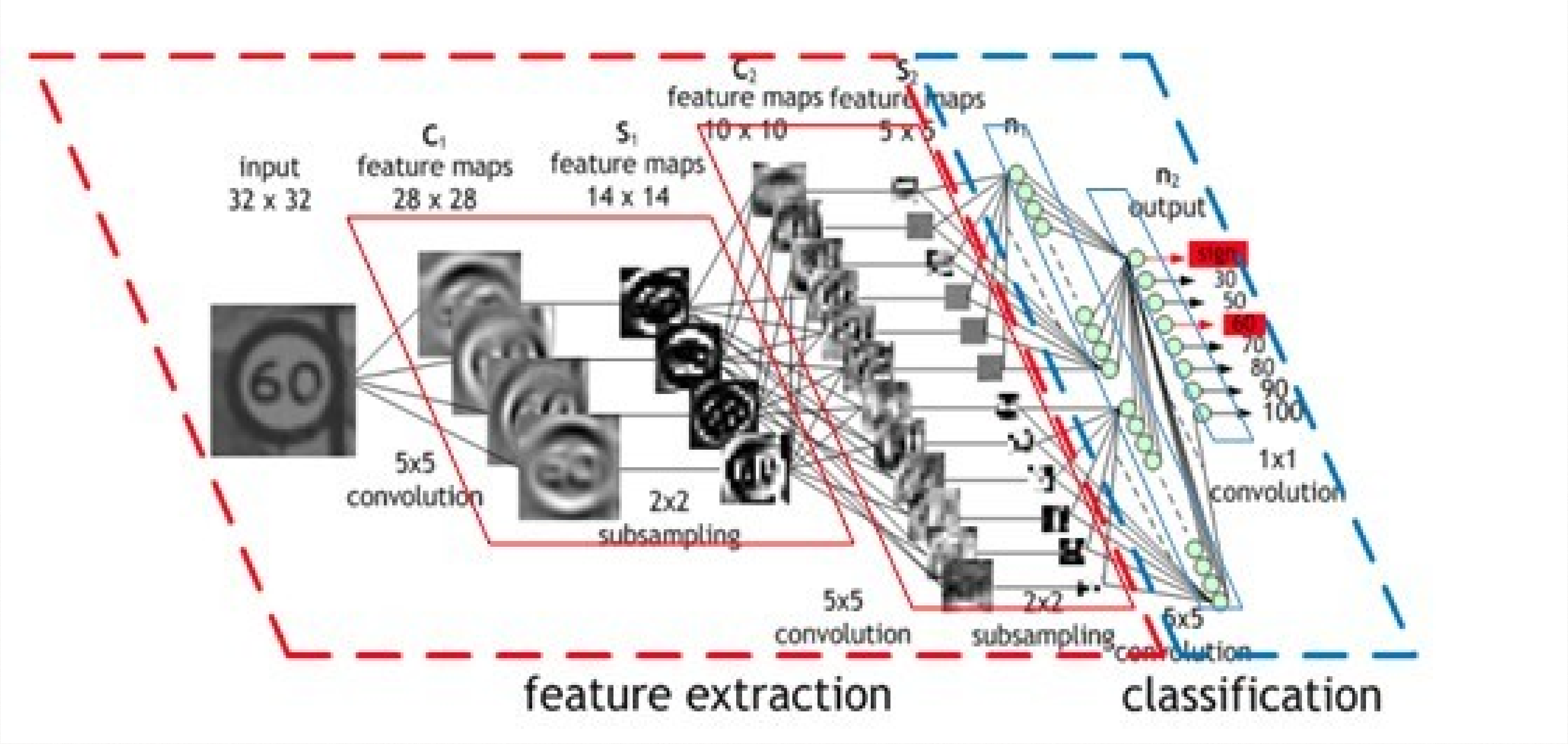
Activation function



- It is parameterized by "weights" and "biases" as model parameters

- Training an ANN is to optimize the weights and biases so that the output matches the label of training data (minimizing 'loss function')

Convolutional Neural Network

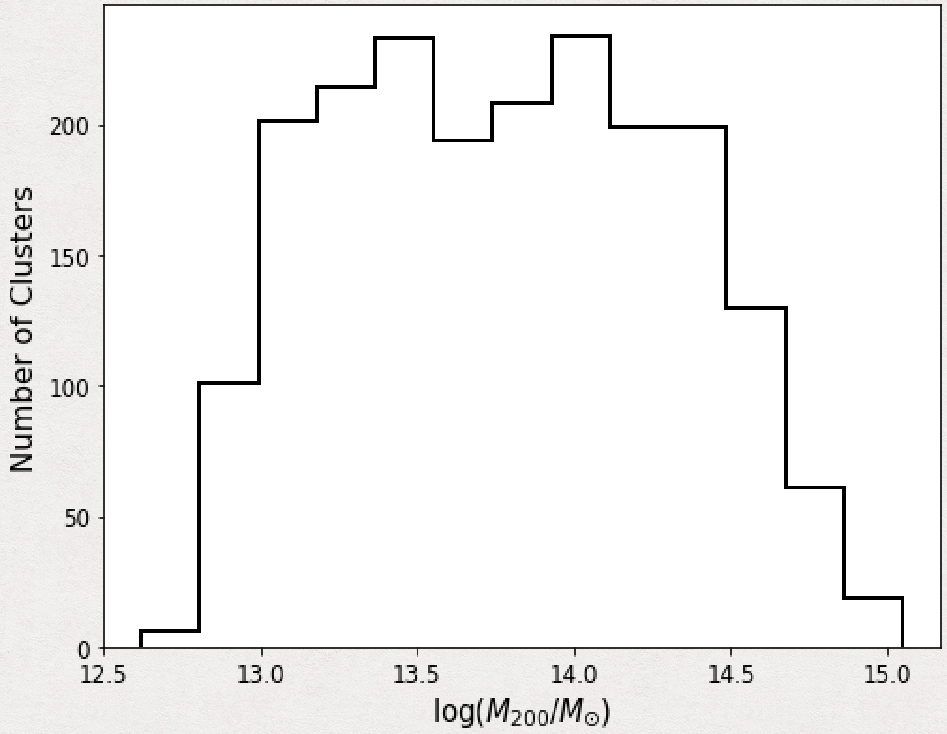


https://www.cs.ryerson.ca/~aharley/vis/conv/

Estimating galaxy cluster masses with Convolutional Neural Network (CNN)

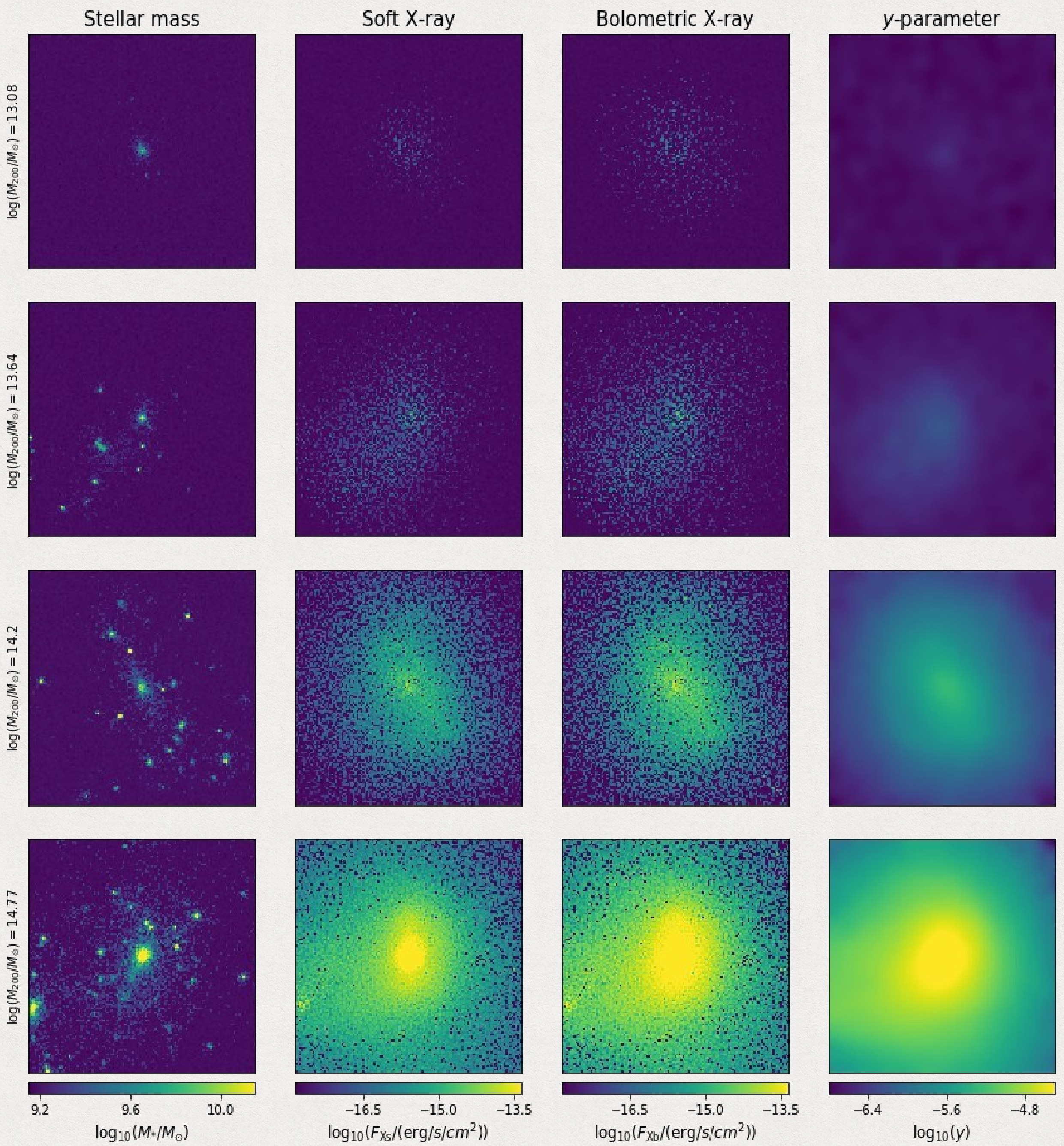
Data overview

- Data: galaxy clusters from BAHAMAS simulation with known mass M_{200}
- Stellar mass; soft and bolometric X-ray flux; SZ y-parameter images are derived
- Cluster Redshift uniformly distributed within 0.03~0.07 (no evolution)



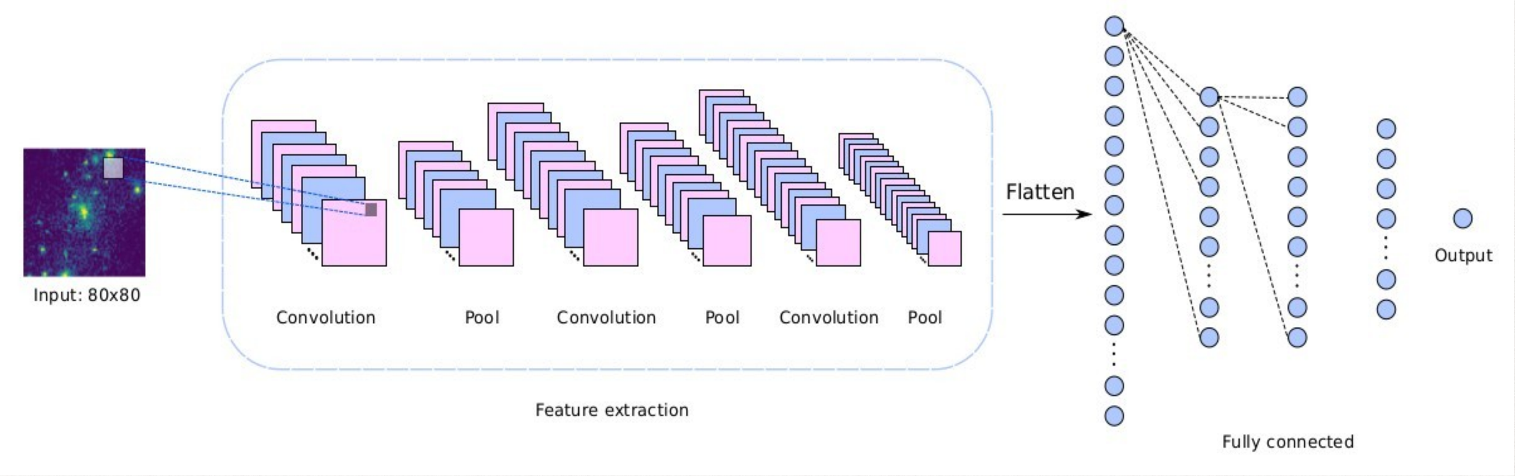
- Image pixelisation: 120 pix x 120 pix corresponding to 20’ x 20’ in the sky

- Interlope included; Gaussian random noise added; smoothed

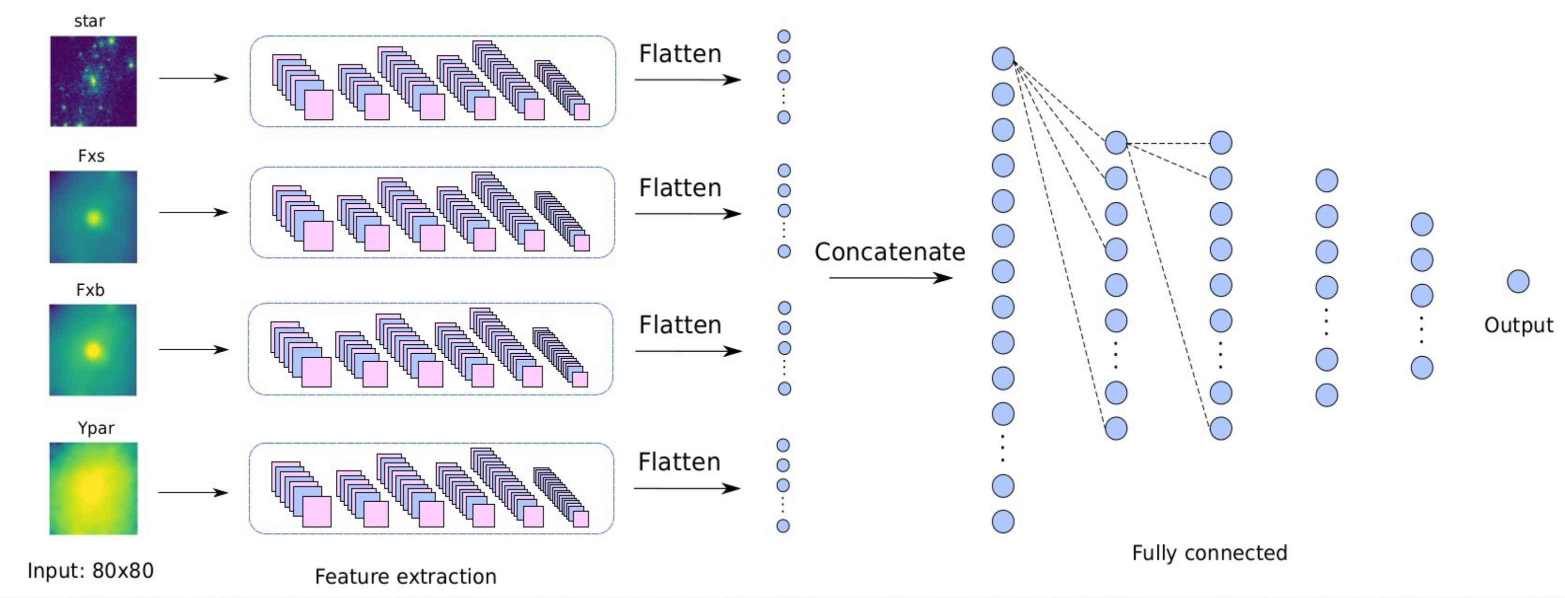


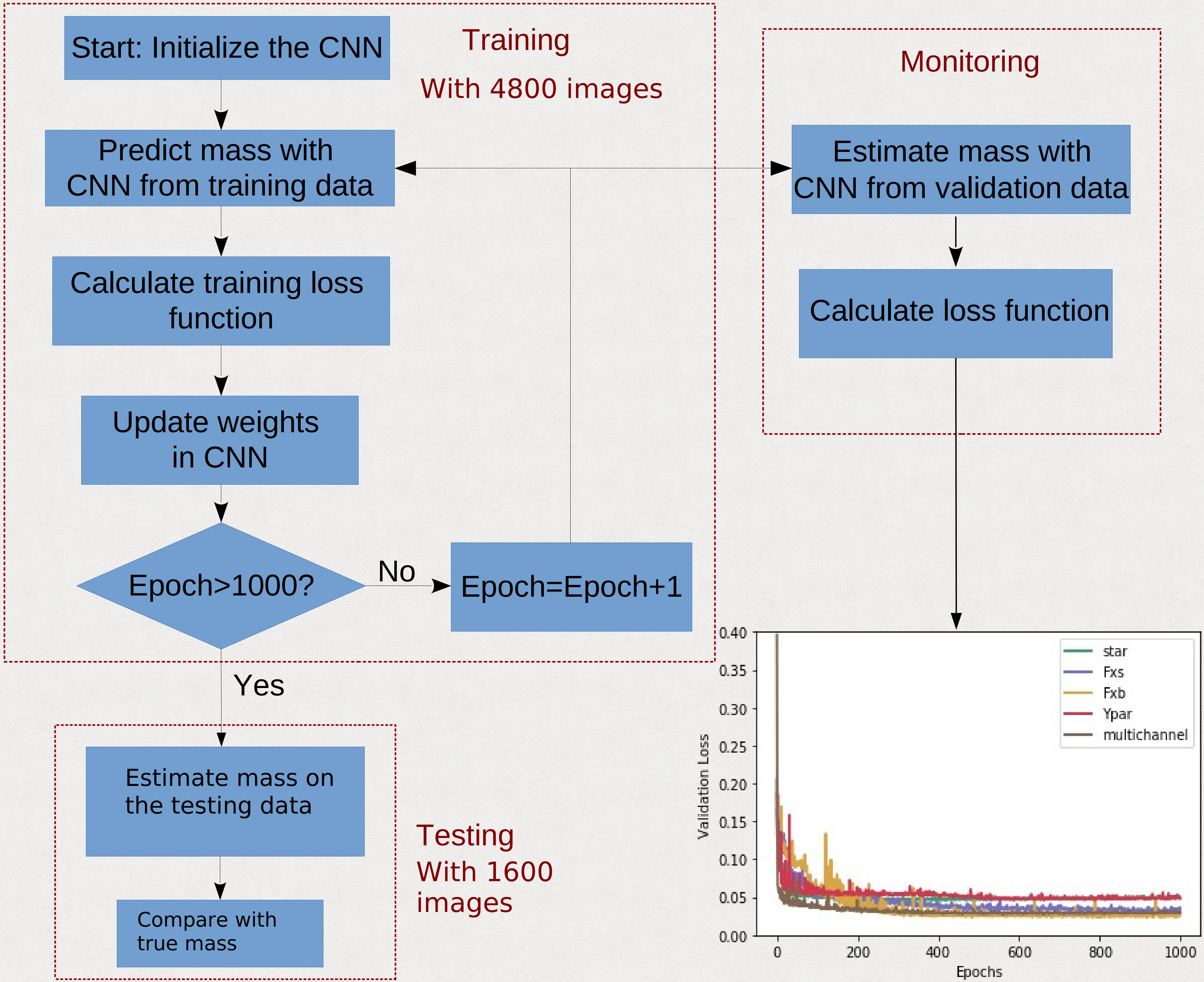
Convolutional Neural Network: Architecture

Single Channel (x4)



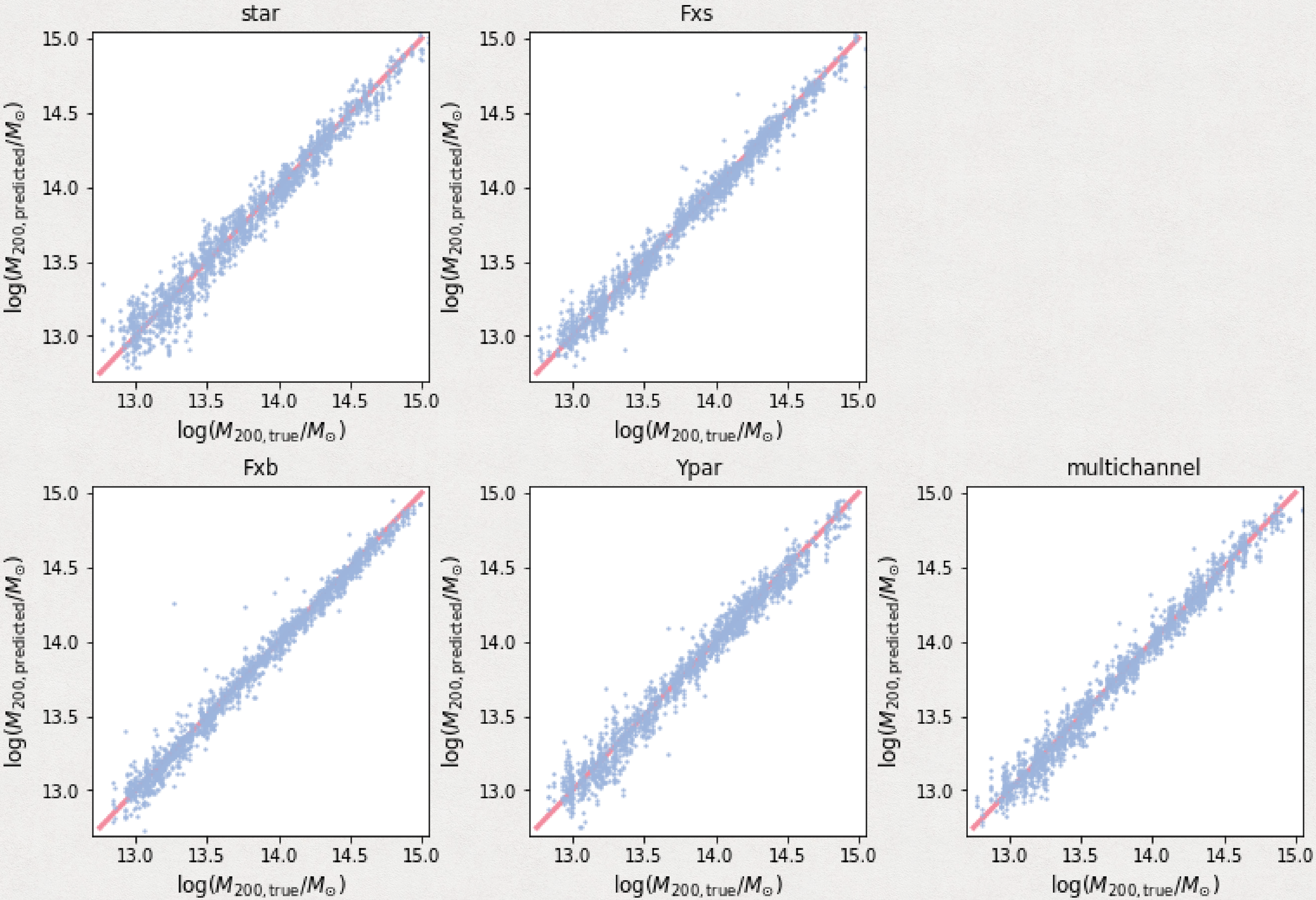
Multi-Channel



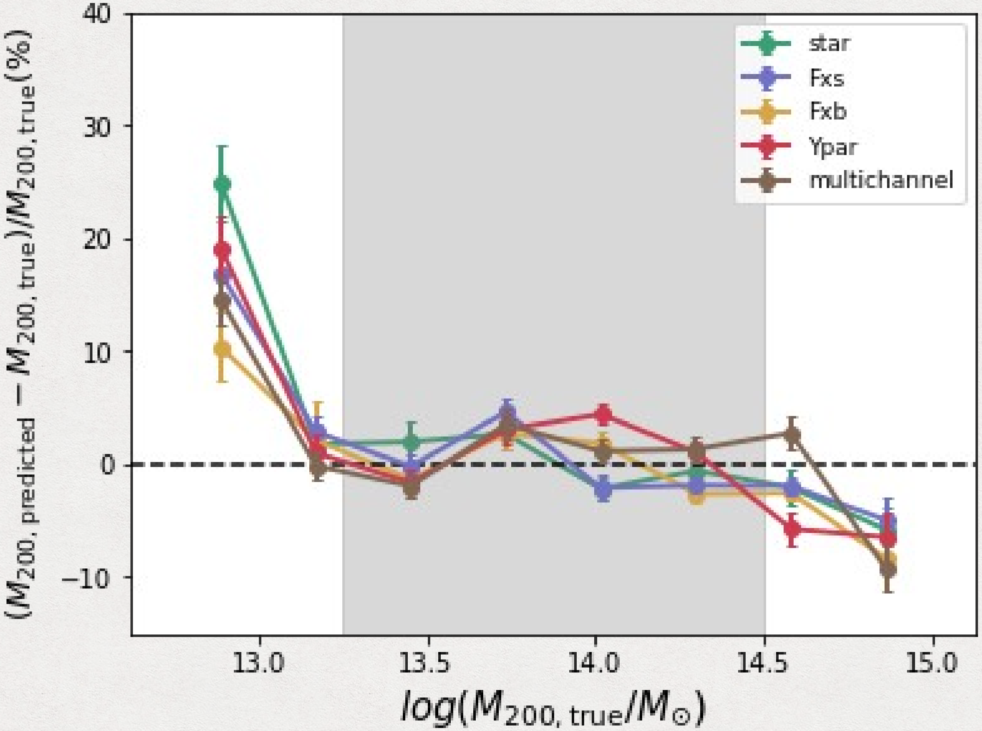


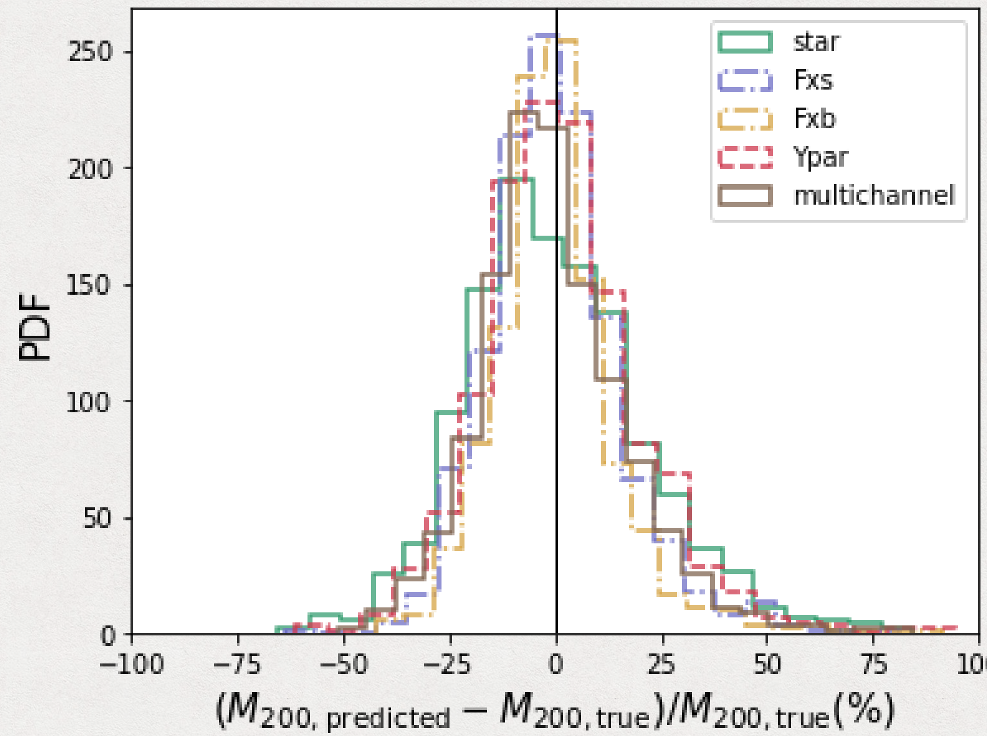
Training
With 4800 images
Start: Initialize the CNN
Predict mass with CNN from training data
Calculate training loss function
Update weights in CNN
Epoch>1000?
No
Epoch=Epoch+1
Yes
Monitoring
Estimate mass with CNN from validation data
Calculate loss function
Estimate mass on the testing data
Compare with true mass
Testing
With 1600 images
Validation Loss
Epochs
star
Fxs
Fxb
Ypar
multichannel
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.40
0
200
400
600
800
1000

Mass Prediction



Mass Bias





Dataset	$\left\langle \log \frac{M_{\mathrm{predict}}}{M_{\mathrm{true}}} \right\rangle$	$\left\langle \frac{M_{\mathrm{predict}}}{M_{\mathrm{true}}} \right\rangle - 1(\%)$	$\langle \mathrm{RMS} \rangle$
star	-0.01±0.003	-0.516±0.621	19.028
Fxs	-0.007±0.002	-0.349±0.517	16.49
Fxb	-0.004±0.002	0.094±0.524	16.036
Ypar	0.002±0.002	1.814±0.559	17.662
multichannel	-0.001±0.002	1.075±0.575	17.693

Table 2. The mean mass bias ($\Delta M \equiv M_{\mathrm{pred}} - M_{\mathrm{true}}$) and scatter obtained from the test set for $13.25 < \log(M_{200,\mathrm{true}}/M_{\odot}) < 14.5$.

- Another machine learning estimation: 7% scatter (Armitage et al. 2019, no observational effect added)
- NFW profile fitting with same set of clusters: -6.4 +- 0.3 % (Hensel et al. 2016 no observational effect added)
- X ray observation: 30% scatter (Zhang et al. 2008);
- tSZ observation: 24% scatter (Bleem et al. 2015)
- Weak lensing: 20 % (Hoekstra et al. 2015)

An attempt to interpret the CNN

An Attempt to Interpret our CNN

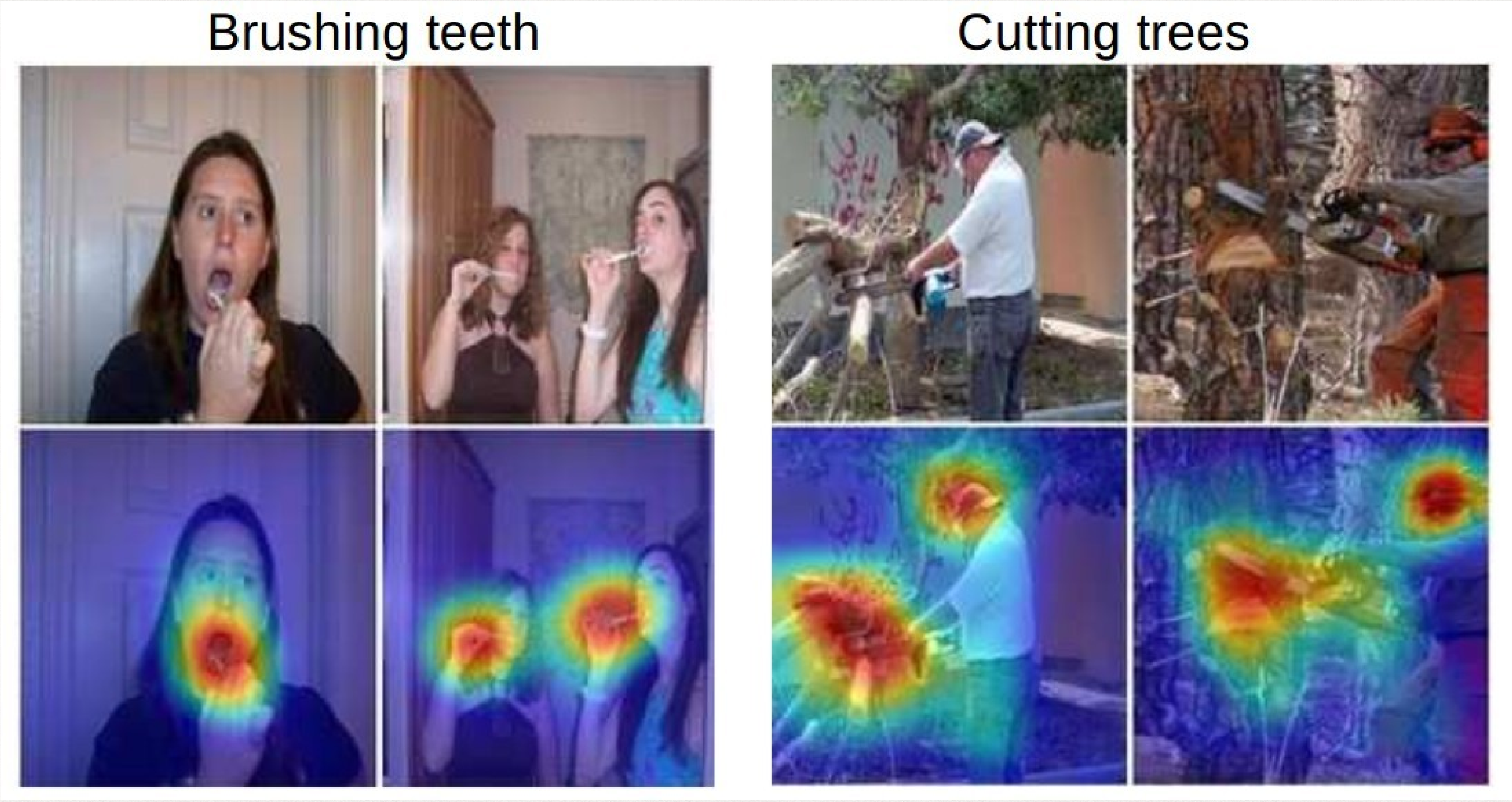
Why?

Because we are astrophysicists. We care not only about the results, but also the underlying physics.

-Try to find the features that ‘trigger’ the neural network make measurements.

An Attempt to Interpret our CNN

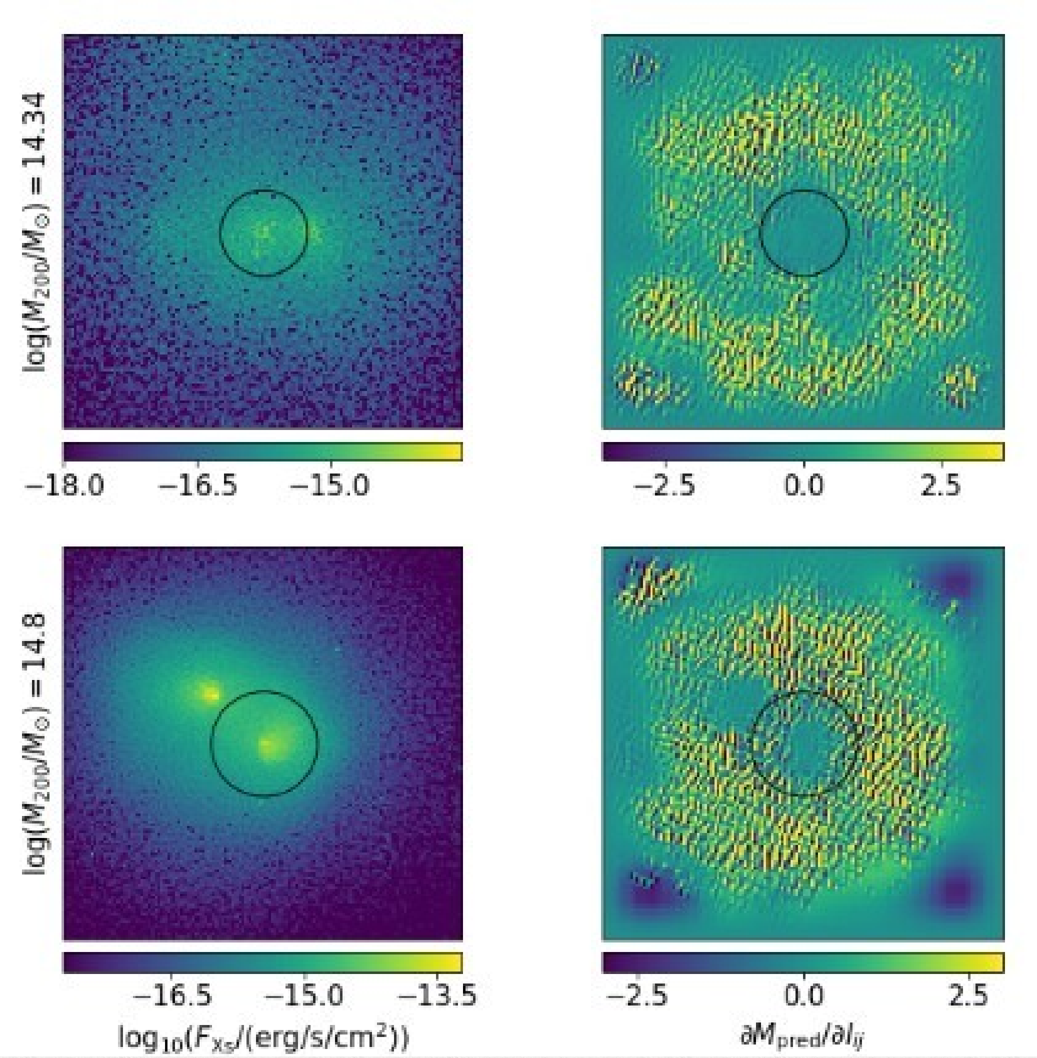
- Try to understand which parts of the image that are important for the CNN to make prediction.



Soft X-ray flux

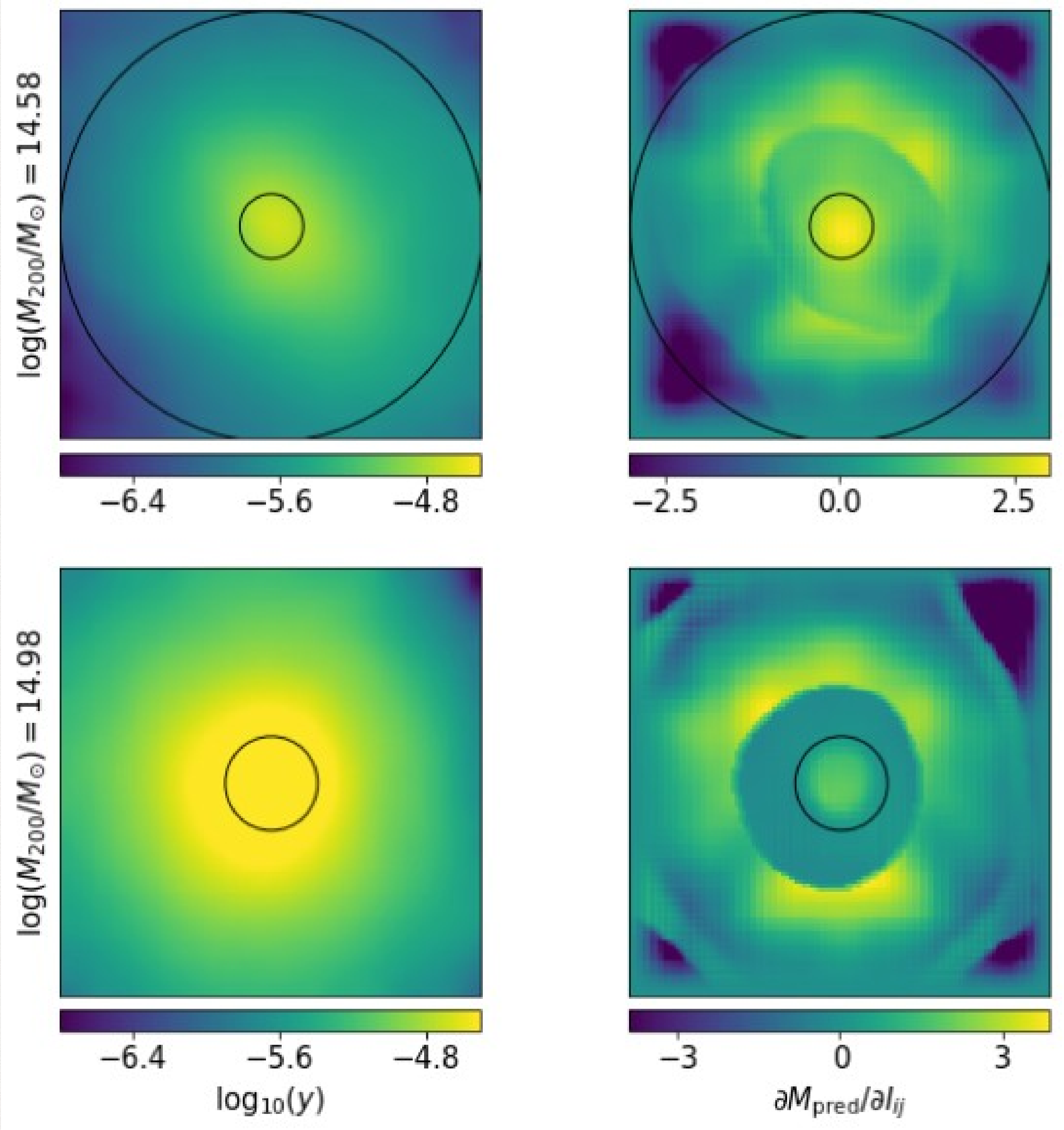
The neural network ‘cares’ more about regions around the cluster rather than cluster center.

- consistent with conclusions in (Mantz et al. 2018; Maughan 2007)



Ypar

The neural network ‘sketches’ the outline of the clusters.



Summary

- Convolutional neural network can estimate cluster mass from images to a high accuracy
- Possible questions of interest: training separately with different cluster types? Adding cluster evolution?
- Possible improvements: wider mass range; more realistic simulations; more realistic systematics, etc
- Outline shape of clusters is taken as relevant to mass estimation (rather than central signal)