# Deep Learning with Convolutional Neural Networks for Galaxy Cluster Mass Calculations

Bryant Dean, Yizhou He, Matthew Ho, Hy Trac
NSF AI Physics Summer Undergraduate Research Program (SURP)

## Introduction

To this day, dark matter has been somewhat of a mystery since its discovery all the way back in 1933. Dark matter halos are the invisible dark matter that surrounds one or a cluster of galaxies. The objective of this project is to estimate dark matter halo masses using simulated particle data from a hydro-particle-mesh simulation from a code name HYPER. That is to be done by using a Convolutional Neural Network, a deep learning method for classification.

(Left) Image of Coma Galaxy Cluster

## Method

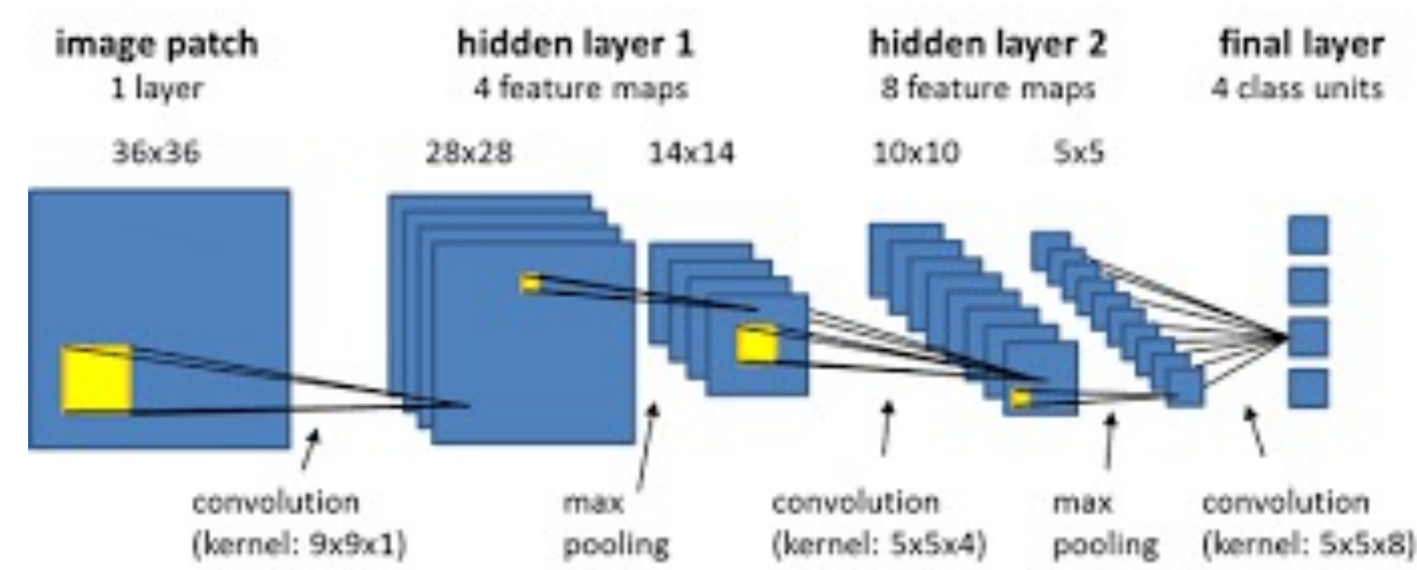
Before being moving to deep learning, first images were made of the halos from the simulated particle data from HYPER. The simulated data was processed to project its distribution in 2D using kernel density estimation and its kernel function embedded inside this expression for density distribution shown below,

$$\hat{f}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \qquad (1)$$

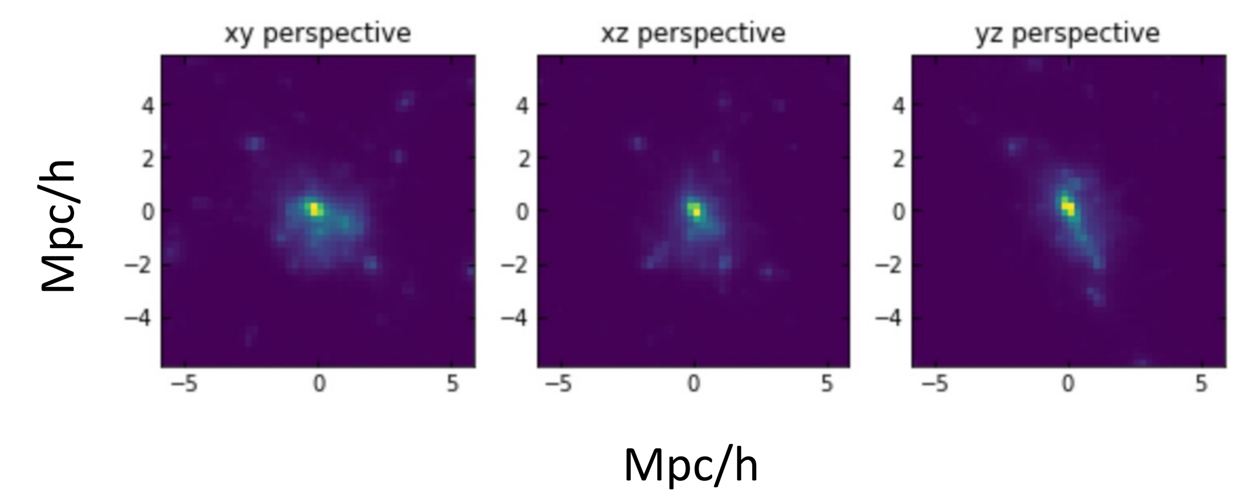
where K is the kernel function, h is the kernel bandwidth, and n is the number of particles. This was done by integrating along the axis in which the projection was along as well. This created a density field for the dark matter particles displayed in the visualization of the dark matter halo and its surrounding filament.

## Convolutional Neural Network (CNN)

A CNN is a feed-forward neural network that takes multidimensional arrays of data and uses various layers in order to extract key features for classifications and predictions.

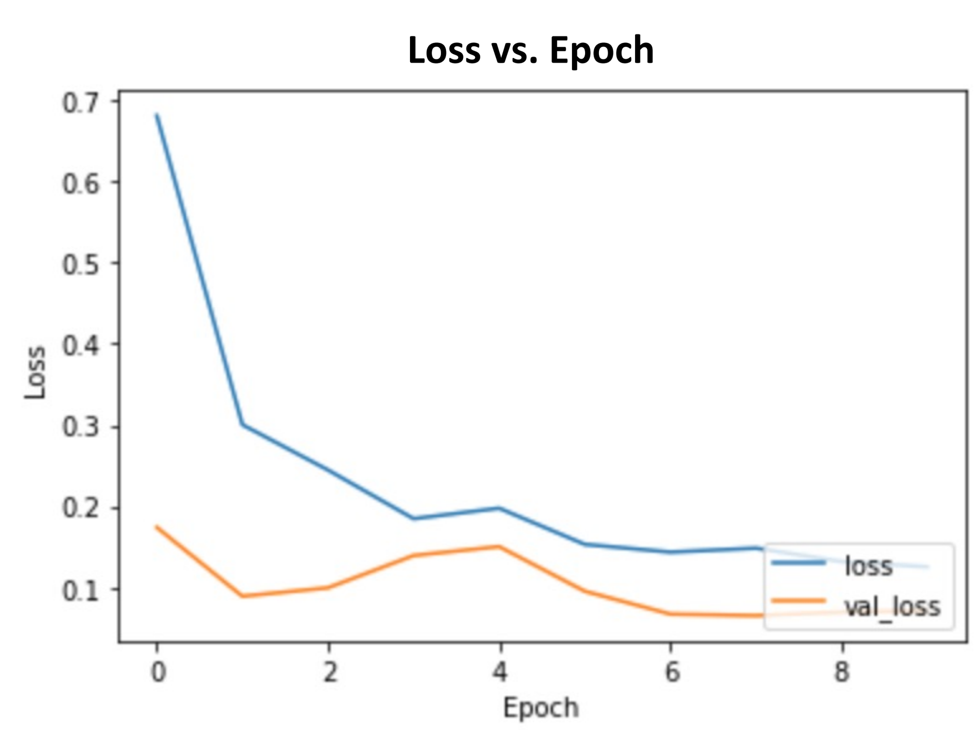


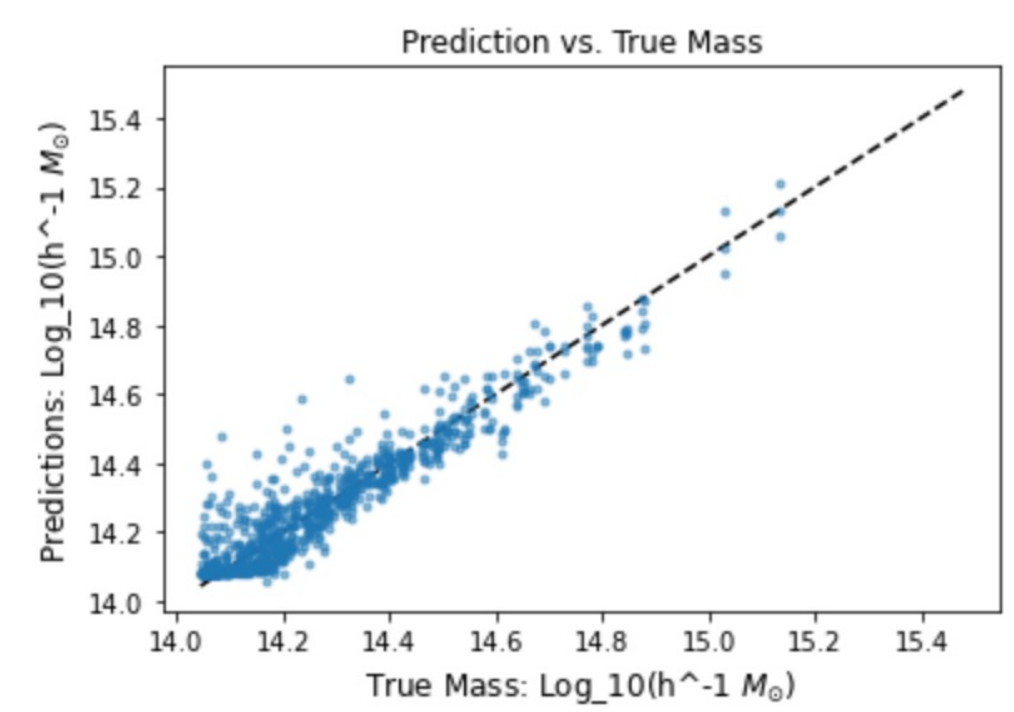

Arrays of simulated 3d dimensional coordinates for 3000 halos that create images like the ones shown are the input for the CNN which will train to calculate and predict the total mass of the halos based on its density.

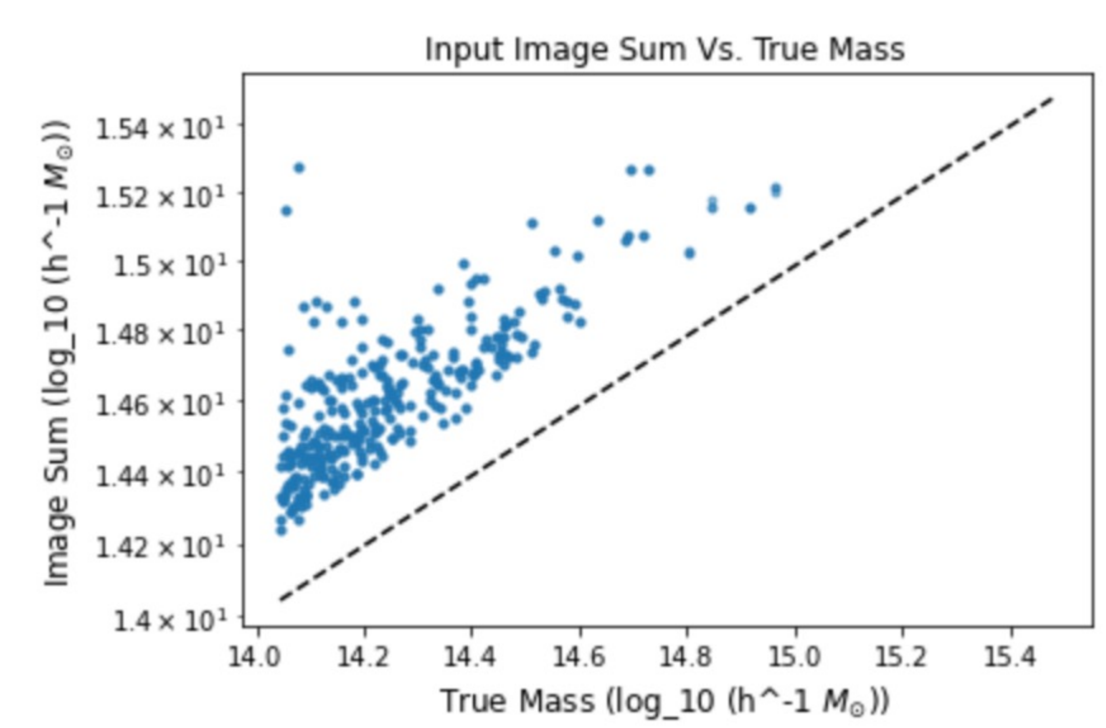



## Results

Over a total of ten epochs of training and testing, the CNN displayed an average loss of 0.2494 and an average validation loss of 0.2144 showing that the deep learning method is an accurate way to predict the masses of halos. By plotting the predictions and image sum by the total mass, we can demonstrate how the CNN predicts masses better than a simple summation.



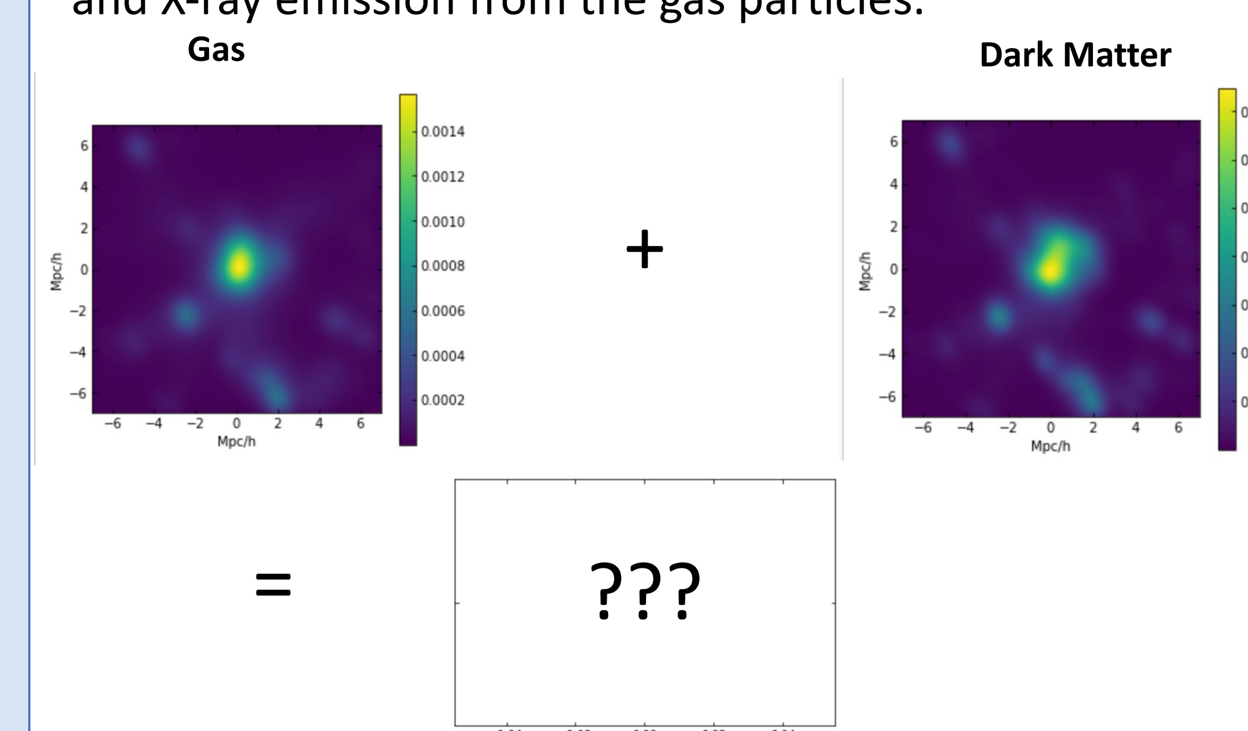




## Conclusion

Using the deep learning method of a convolutional neural network, total masses of galaxy clusters can be estimated the most accurately on average than ever before using density field images of the halos. With an average loss and validation loss under 0.25 we can tell that over each epoch, the loss begins to converge, demonstrating the machine learning models success in training and testing on data more accurately over time.

## What's Next?

The next idea for the application of deep learning methods are to use the convolutional neural network to train on the density of gas particles within the halos to predict the dark matter density or density of the total matter of the halo. Also, it is a goal to do this by training on other features of the particles such as gas pressure and X-ray emission from the gas particles.




## References

[1] Ho et al, A Robust and Efficient Deep Learning Method for Dynamical Mass Measurements of Galaxy Clusters, 2019
[2] Ho et al, Approximate Bayesian Uncertainties on Deep Learning Dynamical Mass Estimates of Galaxy Clusters, 2021
[3] He et al, A Hydro-Particle-Mesh Code for Efficient and Rapid Simulations of the Intracluster Medium, 2021

## Acknowledgments

A special thank you to the NSF AI Institute: Physics of the Future, NSF PHY-202095, Professor Trac and his entire research group for this summer experience, especially Yizhou He and Mathew Ho for their help and guidance throughout this project.