University of Crete
Department of Physics

Searching for a Signature of Turnaround in Velocity Profiles of Galaxy Clusters with Machine Learning

by
Nikolaos Triantafyllou

Supervised by:
PhD Candidate Giorgos Korkidis
Professor Vasiliki Pavlidou
Dr. Paolo Bonfini

A thesis submitted in partial fulfillment for the
Master of Science degree

Heraklion, June 2023

Contents

Abstract

Context: ΛCDM is currently recognized as the standard model in cosmology, but recent tensions between different types of high-accuracy data indicate potential inconsistencies. This generates an urgent need for a new examination of this model using novel observables. Pavlidou et al. have recently illustrated that the turnaround density, the average density within the turnaround radius of galaxy clusters, can be utilized as a probe for this purpose.
Aims: We search for an efficient way of measuring the turnaround radius of galaxy clusters on the plane of the sky.
Methods: We use the MultiDark Planck 2 and Virgo simulations to acquire projections of galaxy clusters and calculate their turnaround radius based on their velocity profiles. We use these data to train (and test) a Convolutional Neural Network (CNN), involving several layers of filters, normalization and regularization techniques.
Results: We find that: (a) the turnaround radius is correlated to the central mass of a galaxy cluster and that the mass distribution around the cluster is almost an irrelevant feature to the model's prediction. (b) The velocity dispersion of the galaxies also contains information related to the turnaround radius. The model's accuracy is not significantly affected by the absence of information inside the R_{200} of the central overdensity in each cluster.

*A*cknowledgements

I wish to express my deep gratitude to my advisor, Prof. Vasiliki Pavlidou, whose trust, guidance, and support over the last two years have been invaluable, restoring my passion for the field. It has been an honor to learn and grow under her tutelage, and I am deeply thankful for the opportunity. I would also like to thank Giorgos Korkidis for his precious advisory and for providing the simulation data used for this thesis. Last but not least, I would like to thank Paolo Bonfini for his vital insights on machine learning. Without their help, this work would have never been realized.

The CosmoSim database used in this work is a service by the Leibniz-Institute for Astrophysics Potsdam (AIP). The MultiDark database was developed in cooperation with the Spanish MultiDark Consolider Project CSD2009-00064. The rest of the simulations in this work were carried out by the Virgo Supercomputing Consortium using computers based at Computing Centre of the Max-Planck Society in Garching and at the Edinburgh Parallel Computing Centre. The data are publicly available at www.mpa-garching.mpg.de/galform/virgo/int_sims.

List of Figures

Chapter 1

Introduction

Detailed cosmological observations of the large scale structure in the Universe validate the ΛCDM model, with the exception of some apparent contradictions, referred to as "tensions", particularly in relation to observations of the Hubble Constant and the clumpiness factor [see e.g.,(Bernal, Verde, & Riess, 2016; Zhao et al., 2017; Joudaki et al., 2017; Hildebrandt et al., 2017; Riess et al., 2018; Riess, Casertano, Yuan, Macri, & Scolnic, 2019; Motloch & Hu, 2018; Miville-Deschênes et al., 2020; Raveri & Hu, 2019; Adhikari & Huterer, 2019; Handley, 2021; Di Valentino, Melchiorri, & Silk, 2020, 2021; Shah, Lemos, & Lahav, 2021)]. Furthermore, current evidence for the existence of a cosmological constant is based on the relation between the present-day values of the cosmological density parameters of matter, Ω_m and dark energy, Ω_Λ, probed either by the cosmic microwave background or by observations of distant supernovae. While some cosmological datasets (cluster abundances, baryon acoustic oscillations) also induce constrains that are sensitive to Ω_m, this is not true for Ω_Λ, leading to several revisions of the evidence that $\Omega_\Lambda \neq 0$. Pavlidou, Korkidis, Tomaras, & Tanoglidis, 2020; Pavlidou & Tomaras, 2014; Korkidis et al., 2020 and Tanoglidis, Pavlidou, & Tomaras, 2015 have demonstrated how the turnaround density ($\rho_{\rm ta}$) and hence, the turnaround radius ($R_{\rm ta}$) can be used as a novel way to probe cosmology. If a measurement of $\rho_{\rm ta}$ is made in a sufficient number of clusters at a range of low redshifts, constraints very sensitive to Ω_Λ can be obtained.

The concept of the turnaround radius appears naturally in the context of the spherical collapse model (Gunn & Gott III, 1972; Fillmore & Goldreich, 1984; Bertschinger, 1985). If we assume the structures in the Universe emerged from spherical perturbations in an otherwise homogeneous and isotropic early Universe (ΛCDM), according to General relativity, we can evolve these inhomogeneities in time as an isolated closed Universe. These sub-Universes obey the Friedmann equation, thus their collapse is isotropic. Here, the turnaround radius is defined as the boundary where the expansion of the structure is halted and the structure starts to collapse, detaching from the Hubble Flow. To avoid any confusion, turnaround is not an event that appears once in the history of a structure. Rather, it can be found around all structures at all times. Around a virialized galaxy cluster, for instance, there is an accretion region of infalling material, but at sufficiently large distance, other structures are carried away by the expansion of the Universe. Hence, there is some point in space where the accretion meets the Hubble flow, a zero-shell velocity, which is

located at the turnaround radius. Therefore, turnaround radius is a well defined boundary in analytic calculations, simulations and observations.

According to the spherical collapse model, the turnaround density ($\rho_{\rm ta}$) as a function of cosmological parameters and redshift is given by (Pavlidou et al., 2020):

$$\int_0^1 \frac{da_s}{\sqrt{\Omega_{ta}(a_s^{-1}-1)+\Omega_{\Lambda,0}(a_s^2-1)}} = \int_0^{a_{ta}} \frac{da}{\sqrt{\Omega_{m,0}a^{-1}+\Omega_{\Lambda,0}a^2+(1-\Omega_{m,0}-\Omega_{\Lambda,0})}} \quad (1.1)$$

where $\Omega_{\rm ta} = \frac{\rho_{\rm ta}}{\rho_{\rm critical}}$, a and a_s are the scale factors of the background Universe and the shell respectively, and the "0" index represents the present cosmic epoch. Thus, spherical collapse predicts that a single $\rho_{\rm ta}$ can characterize all structures in the same redshift given the cosmological parameters. This density is related to the turnaround radius, $R_{\rm ta}$ and the mass $M_{\rm ta}$ of all matter (baryonic and dark) within a sphere of radius $R_{\rm ta}$ through:

$$\rho_{\rm ta} = \frac{M_{\rm ta}}{(4/3)\pi R_{\rm ta}^3} \quad (1.2)$$

In principle, then, observations of $R_{\rm ta}$ and $M_{\rm ta}$ can be combined to constrain Ω_m and Ω_Λ.

It has been shown by (Korkidis et al., 2020) that a single turnaround radius can indeed accurately describe simulated halos, even in the presence of significant asphericities in their mass distribution. The value of the turnaround radius as measured from radial velocity profiles is in good agreement with spherical collapse and all of the aforementioned characteristics are persistent.

Presently there exists no method of measuring the turnaround radius of galaxy clusters observationally from readily available information. We can only observe structures projected on the plane of the sky and so, it is not possible to compute the turnaround radius by considering the velocity profiles of the shells surrounding the perturbation. Even if a method was available, we would still need to have projections of clusters on a 2D plane with known turnaround radii to test the model's accuracy.

Consequently, the aim of this work is to investigate what type of observational datasets may encode information on the $R_{\rm ta}$ of galaxy clusters. In particular, we wish to establish whether there exists, in principle, a way of measuring the turnaround radius on the plane of the sky, using information on the distribution of galaxies and their redshifts around galaxy clusters. We used N-body dark matter simulations to simulate projections of dark matter halo clusters (baryonic effects are negligible at these scales therefore we use the terms halo and galaxy interchangeably) and measure their turnaround radii from their velocity profiles. To this end, we implemented machine learning methods to find a relation between these projections and the accurate, numerically computed turnaround radii from the 3D information available in simulations.

The outline of this thesis is as follows. In Sect. 2 we describe the set of N-body simulations used in this work and the characteristics of the resulting turnaround radii. In Sect. 3 we describe the methods used for the creation of the model that tries to predict the turnaround radius. In Sect. 4 we present the results of our analysis. We discuss these findings in Sect. 5.

Chapter 2

Data

2.1 N-body Simulations

In this work we use data from: (a) The MuldiDark Planck 2 (MDPL2) Simulation (Riebe et al., 2013). The MDPL2 simulation was performed with (1 $Gpc/h)^3$ box size, with about 57 billion particles in a cosmology consistent with the Planck data (2013). It was performed using the L-Gadget2 code. (b) The Virgo Intermediate Scale Simulations simulations (ΛCDM, OCDM, SCDM) (Frenk et al., 2000). These simulations contain 256^3 particles each in a box of (239.5 $Mpc/h)^3$ and represent different cosmological models. The calculations were done using the external AP3M N-body code. For the data extraction we used the methods presented on the respective project websites[1,2]. The basic properties of the simulations can be seen in Table 2.1. Table 2.2 shows the cosmological parameters adopted in each simulation.

For the halo catalog production MDPL2 uses the ROCKSTAR algorithm (Behroozi, Wechsler, & Wu, 2012) while the Virgo simulations use the Friends-of-Friends slgorithm (FOF; (Davis, Efstathiou, Frenk, & White, 1985; Knebe et al., 2011)). In the latter, particles are linked together only if their distance is below a linking length and create groups that do not interfere with one another. ROCKSTAR is based on FOF but uses six phase-space dimensions and one time dimension for adaptive hierarchical refinement.

Table 2.1: Simulation box size in h^{-1} Mpc, number of particles, particle mass in $h^{-1}M_\odot$, and Force resolution in units h^{-1}kpc. All simulations are from dark-matter–only runs.

Simulation	Box	Particles	M_p	ϵ
MDPL2	1000	3840^3	1.51×10^9	13-5
Virgo ΛCDM	239.5	256^3	6.86×10^{10}	
Virgo SCDM	239.5	256^3	22.7×10^{10}	
Virgo OCDM	239.5	256^3	6.86×10^{10}	

[1] https://www.cosmosim.org/metadata/mdpl2/
[2] https://wwwmpa.mpa-garching.mpg.de/galform/virgo/int_sims/index.shtml

Table 2.2: Cosmological parameters used in simulations

Simulation	Ω_m	Ω_b	Ω_Λ	h_{100}
MDPL2	0.307115	0.048206	0.692885	0.6777
Virgo ΛCDM	0.3	-	0.7	0.70
Virgo SCDM	1.0	-	0.0	0.50
Virgo OCDM	0.3	-	0.0	0.70

2.2 Acquired data

From the simulations we located the largest halos (3248 from MDPL2 and 328 from the Virgo simulations) of the catalogues which provide us with the positions, virial masses[3] and velocities of the halos.

We derived the turnaround radius kinematically from the velocity profiles of the surrounding halos by measuring the mean velocity of spherical shells as described by Korkidis et al., 2020. The distribution of the turnaround radius values for each simulation can be seen in Fig. 2.1.

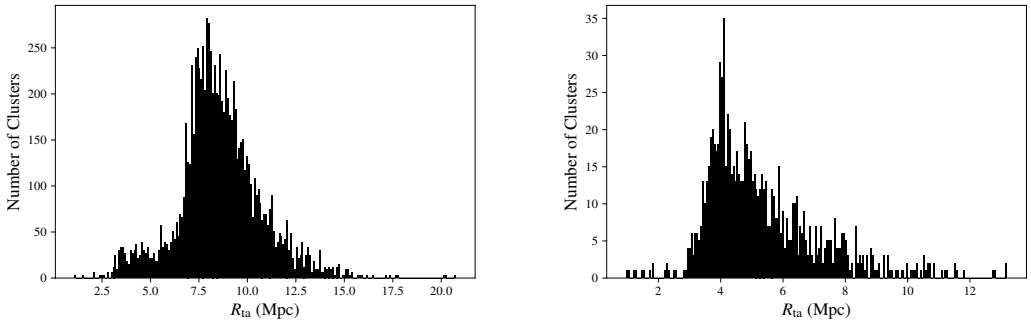

Figure 2.1: Histogram of R_{ta} values from MDPL2 (left panel) and Virgo ΛCDM (right panel). *Left panel:* MDPL2 R_{ta} values from 9744 halos; min:1.08 ; max: 20.76 ; mean: 8.58; std: 2.15. *Right panel:* Virgo ΛCDM R_{ta} values from 962 halos; min:0.95 ; max: 13.20 ; mean: 5.32; std: 1.83.

We considered large enough regions (radius of 25Mpc) around the central halos so that they contain the largest turnaround radius found. Subsequently, we acquired all of projections of these regions (9744 from MDPL2 and 984 from the Virgo simulations), along with the virial masses, positions and line-of-sight velocities of the collapsed halos with respect to the central one.

In this procedure we only considered the halos with masses $> 10^{12} M_\odot$. The reason why we applied the limit is that we might not be able to distinguish small galaxies embedded in halos less massive than $10^{12} M_\odot$ observationally (especially if in the future we apply our method to higher redshift observations) and so this makes our simulated data more aligned with such a realistic scenario.

For the same reason, even though we can determine through redshift the distance of

[3]The mass within the radius R_{200} which encloses a mean density 200 times the background mass density of the Universe

such objects, there is always an error and it is dependent on the method used for this determination e.g. spectroscopic or photometric techniques. Consequently, we considered a velocity cut of $0.3 \cdot 10^4$km/s, meaning that in each cluster we added all of the halos that were located in front or behind of the projections for which their line-of-sight velocities (with respect to the central halo) lied below the designated limit. This simulates the effect of redshift errors on the inclusion of individual galaxies in the analyzed sample. Our approach simulates redshift errors one would get by spectroscopic techniques. All of the aforementioned steps were carried out at $z = 0$.

For the MDPL2 data we additionally considered another 2 redshift snapshots, namely $z = 0.49$ and $z = 1.03$. In Fig. 2.2 we can see the distribution of the $R_{\rm ta}$ values from each of the additional MDPL2 redshift snapshots (left panel) and from each of the non-ΛCDM Virgo simulations (right panel). The different location of peaks in different redshifts and different cosmologies is a direct indication that $\rho_{\rm ta}$ is a function of the redshift and the cosmological parameters as described by Equation 1.1.

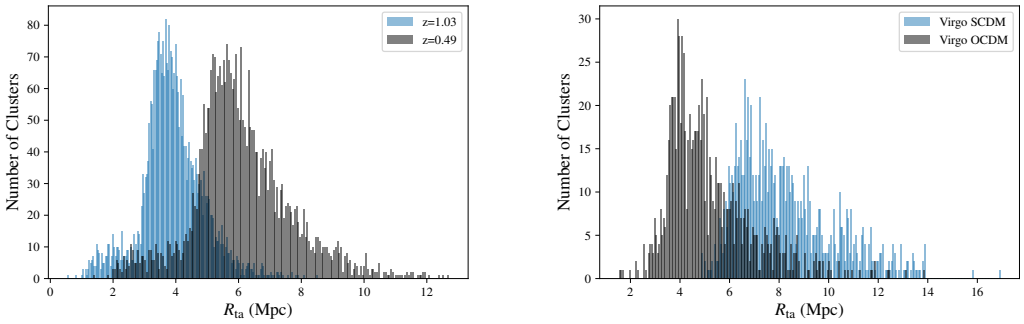

Figure 2.2: Histogram of R_{ta} values from MDPL2 form $z = 0.49$ and $z = 1.03$ (left panel) and Virgo from OCDM and ΛCDM cosmology (right panel). *Left panel:* MDPL2 ($z = 0.49$) R_{ta} values from 3340 halos; min: 1.35 ; max: 12.7 ; mean: 6.09; std: 1.56. MDPL2 ($z = 1.03$) R_{ta} values from 3356 halos; min: 0.54 ; max: 8.51 ; mean: 3.87; std: 0.99. *Right panel:* Virgo SCDM R_{ta} values from 971 halos; min:4.85 ; max: 16.93 ; mean: 8.25; std: 1.96. Virgo OCDM R_{ta} values from 958 halos; min:1.53 ; max: 13.89 ; mean: 5.20; std: 1.73

For the MDPL2 data a velocity cut of $2 \cdot 10^4$ km/s was also considered, corresponding to the accuracy of photometric redshifts. A projection of a region around a random cluster from each velocity cut (upper and lower panel) is shown in Fig. 2.3. It is clear that the high velocity cut data have more contamination due to the thicker redshift slice considered in this case. As expected, the larger velocity cut includes more halos that will not be a part of the initially considered region of 25Mpc.

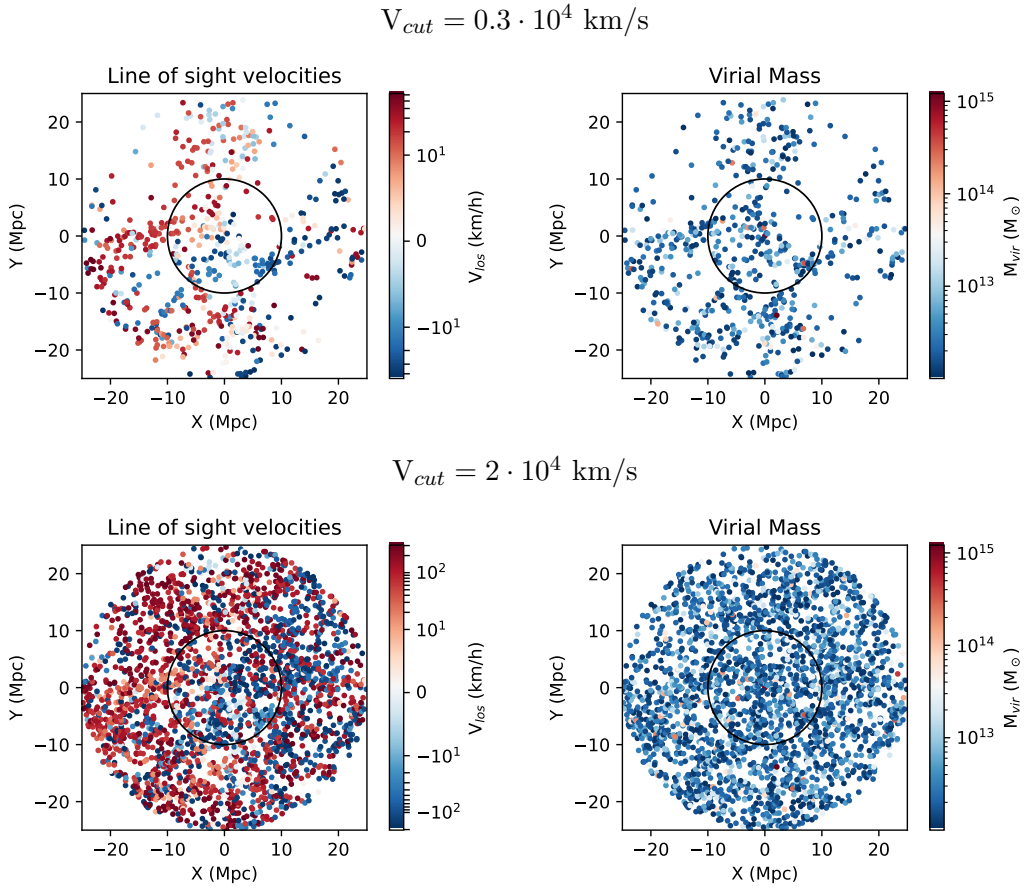

Figure 2.3: Random cluster projection on a 50×50 Mpc area. The black circles represent the turnaround radius. The upper and bottom panels show the 2 cases of the velocity cuts. We can see that the higher velocity cut data include more halos and ergo have more contamination. In the left side the colors represent the line of sight velocities of the halos with respect to the central. In the right side the colors represent the virial masses of the halos. In every case the colorbars are logarithmically scaled. From the left side we can see that there is not a preferable direction in the line of sight for the surrounding halos. The right column seems to be containing only small masses (blue colored) with respect to the color bar but the central halo has a higher mass by two orders of magnitude from almost all of the rest.

Chapter 3

Methods

3.1 Artificial Neural Networks

Over the course of the last years machine learning has gained considerable attention in the field of astrophysics and cosmology [see (Lahav, 2023; Baron, 2019; Carleo et al., 2019; Huertas-Company & Lanusse, 2022)] with several discussions on the effectiveness of these data driven methods (Lin et al., 2022). The subcategory of deep learning in particular, has been proven to be a powerful tool, with feature extraction capabilities embedded in the internal structure of its models.

In this paper we used Artificial Neural Networks (NNs) to find a relation between observables, such as virial masses of the line-of-sight velocities, and the turnaround radius. NNs can be viewed as multi-variable functions, transforming inputs into target outputs via a series of calculations related to a structure reminiscent of the biological neural networks. This structure can be best resembled by the form of a flowchart with the neurons and the weights forming its fundamental components. Neurons are variables appearing as nodes in the flowchart, and are organized in groups which constitute the layers of a neural network. These layers are successively connected and the strength of these connections is determined by a set of synaptic weights. The procedure of training involves the iterative adjustment of these weights until the system achieves its purpose as successfully as possible, in this case predicting the $R_{\rm ta}$.

In conventional NNs, the layers are fully connected, designed primarily for one-dimensional inputs. For this reason we use Convolutional Neural Networks (CNNs), which are best suited for single and multi-channel images. CNNs use convolutional layers which connect the neurons of one layer only to a small region of the neurons of the previous layer (e.g. pixels of the image), allowing for the recognition of low level spacial features. The main structure of our NN can be seen in Table A.1 . This structure was the best candidate in almost all of the trials performed. The focus on point estimates is due to the restrictive computational power.

3.2 Metrics

To assess the quality of the predicted R_{ta}, we adopted the following commonly used statistics:

3.2.1 Coefficient of Determination (R^2 score)

The coefficient of determination, also known as the R^2 score, is a statistical metric that is often used to quantify the quality of a regression model. Formally, it measures the proportion of the variance in the dependent variable that is predictable from the independent variables.

An R^2 score of 1.0 indicates that the model perfectly predicts the dependent variable using the independent variables. On the other hand, an R^2 score of 0 indicates that the model does not predict the dependent variable at all, i.e., the model is no better than simply taking the mean of the dependent variable. The R^2 score is defined as:

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{3.1}$$

where SS_{res} is the residual sum of squares defined as $\sum_{i=1}^{n}(y_i - f(x_i))^2$, y_i are the actual values, $f(x_i)$ are the predicted values, and SS_{tot} is the total sum of squares defined as $\sum_{i=1}^{n}(y_i - \overline{y})^2$, with $\overline{y}$ being the mean value of the dependent variable.

Note that the R^2 score can become negative when the chosen model fits the data worse than a horizontal hyperplane.

3.2.2 Shapley Values

Shapley values is a concept that originates from cooperative game theory. In the context of machine learning, particularly in the field of interpretability, Shapley values provide a method to allocate the contribution of each feature to the prediction for each individual instance. Formally, the Shapley value of a feature i is defined as:

$$\phi_i(v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N| - |S| - 1)!}{|N|!}(v(S \cup \{i\}) - v(S)) \tag{3.2}$$

where N is the set of all features, S is a subset of features, and $v(S)$ is the value function that assigns a value to a set of features S.

The Shapley value of a feature can be interpreted as its average marginal contribution to the prediction, across all possible combinations of features. This property, along with others such as efficiency (the sum of the Shapley values equals the total effect of all features) and symmetry (identical features receive the same Shapley value), makes Shapley values a powerful tool for our model interpretation.

3.3 Preprocessing

As mentioned in Sect. 2 our data consists of projected regions around dark matter halo clusters, incorporating the details of each halo's virial mass and line-of-sight velocity. In order to feed the data to the algorithm we converted them into images as 2D histograms. For the histograms we used the appropriate weights to create 3 types of images: mass density, number density and mean line-of-sight velocity in each bin, in our case pixel[1]. This transformation does not affect the outcome since by dividing by the area of each pixel (constant) to acquire the densities we don't change the relation between the images. Additionally, by taking the mean velocity some of the pixels values may change relative to one another but we discovered through several tests that this does not affect the outcome and sometimes improves it.

Different resolution of the images were tested but for more accurate results and due to computational restrictions 25×25 images were chosen as the best candidate for the best results. Surprisingly, it seems that the binning of the halos with the histogram technique improves the algorithm's performance. This is partially due to the fact that with a higher resolution there are too many 0-valued pixels in the image which seems to confuse the algorithm even with all the regularizations (l1 and l2 regularizations were considered). Images of 11, 21, 25, 31, 41 and 51 side pixel length were tested for this purpose. The number of the pixels is odd in order for the central halo to be located exactly in the center. For each case, 80% of the simulation data were used for the training and validation and 20% was used for testing the trained model.

We can see from Fig. 2.1 that the turnaround radius data are not uniform but rather they present a Gaussian distribution-like form. This could potentially result in the algorithm only predicting values close to the peak, resulting in a problem usually referred to as heteroscedasticity (Andreon & Hurn, 2013; Feigelson, De Souza, Ishida, & Babu, 2021) or class dependent residuals in classification problems (Lin et al., 2022). Several normalizations were tested at this point like a uniform transformation but the original data were more successful.

[1]We have not considered observational errors in each "observed" quantity, although binning simulates to some extent observational uncertainties. The reason for not adding noise is that we want to provide the model with the most favorable conditions possible

Chapter 4

Results

4.1 General

In this section we only consider the MDPL2 data in redshift $z = 0$, in both the low and the high velocity cuts, to get a general idea of the potential of our methods. We investigated all of the possible combinations of the mass, number, and velocity data in a 2D or 3D data cube. The resulting R^2 scores can be seen in Fig. 4.1.

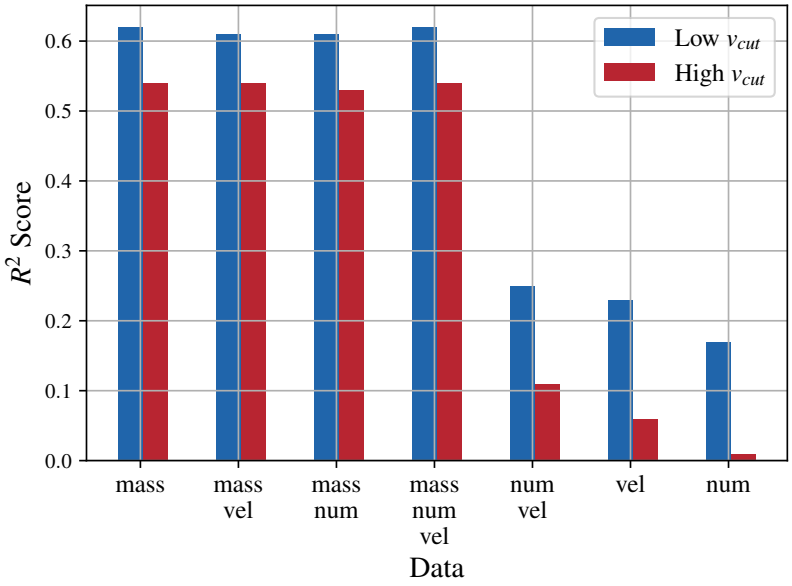

Figure 4.1: Comparison between low and high velocity models performance. This bar plot shows the different R^2 values for each combination of the 3 features that we considered: the mass density, number density and mean line of sight velocity (indicated as "mass", "num", and "vel" respectively). The blue bars represent the scores on the low velocity cut while the red ones on the high velocity cut data. We can see that the best model is for the low velocity cut (after all, there is lower contamination) and for all of the data containing the mass.

The highest accuracy was achieved with the low velocity cut data containing the mass information. This is a quite logical result since the gravitational strength of the mass of the cluster should correlate to how far from the center of mass the Hubble flow becomes

dominant. Surprisingly, the velocity data show no such strong correlation. We proceed to investigate these 2 cases in more detail in Sect. 4.2 and 4.3. We note in passing, that the results of a Gaussian prediction around the mean of the true $R_{\rm ta}$ values produced similar results to our unsuccessful candidates. We continue with only considering the low velocity cut data as an attempt to study the best possible case.

4.2 Mass distribution

The significantly better performance of the models based on data containing the mass, lead us to investigate what would this model's performance be without information on the central, more massive, halo (i.e. without information on the mass of the actual central collapsed galaxy cluster). Consequently, we removed it completely from all of the regions and re-trained our model. The results can be seen in the Fig. 4.2 that were acquired from the low velocity cut data that performed better in the previous section. We can see that without the central halo the models' performances containing the mass drop significantly. This implies that their better performance was due to the central halo's mass information. All of the other models predict the turnaround with about the same score.

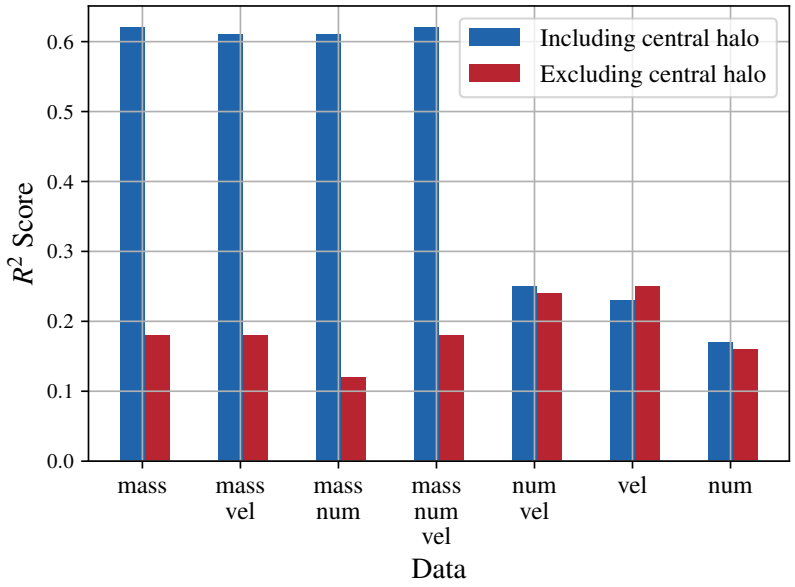

Figure 4.2: Comparison between 2 cases: including or excluding the central halo. This bar plot shows the different R^2 values for each combination of the 3 features that we considered, the mass density, number density and mean line of sight velocity (indicated as "mass", "num", and "vel" respectively). The blue bars represent the models' performance with the central halo while the red ones represent the scores of the models without it. We can see that without the central halo the model's performance drops significantly and all of the models predict the turnaround with about the same score.

A more accurate representation of these results is shown in Fig. 4.3 with scatter plots between the predicted values of the turnaround radii and their real values. If the model was perfect then all the points ($R_{\rm ta,predicted}$, $R_{\rm ta,true}$) would be located on the $y = x$ line. It is evident that in the case where the central halo is missing (lower panel) there seems to be

a little to no correlation whatsoever.

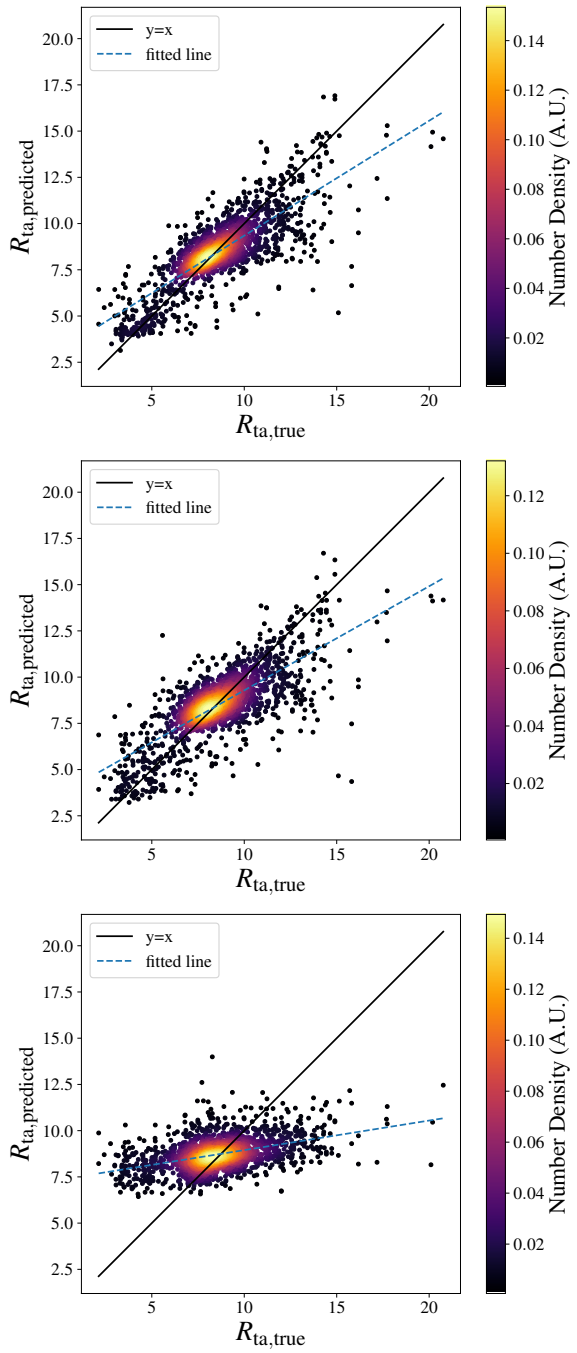

Figure 4.3: Comparison between predicted and true values of the R_{ta} for different CNN models using mass information. *Upper panel:* The low velocity cut data were used with the model having an R^2 Score of 0.62. *Middle panel:* The high velocity cut data were used with the model having an R^2 Score of 0.54. *Bottom column:* For this model the low velocity cut data were used but this time exluding the central halo. R^2 Score =0.18.While the other models show some correlation between the treu and the predicted values, this one seems that its predictions are focused around the mean value of the R_{ta} values distribution.

The results of the R^2 score from different redshifts from MDPL2 and from different

cosmologies from both MDPL2 and the Virgo Simulations can be seen in Fig. 4.4 (upper and lower panel respectively). Best models seem to be the ones containing the data from the MDPL2 possibly due to the fact of the larger number of training instances. Additionally, it seems that a model trained at one redshift is not successful when tested at another one. This is expected because of the relation between $\rho_{\rm ta}$ and redshift discussed in Chap. 1.

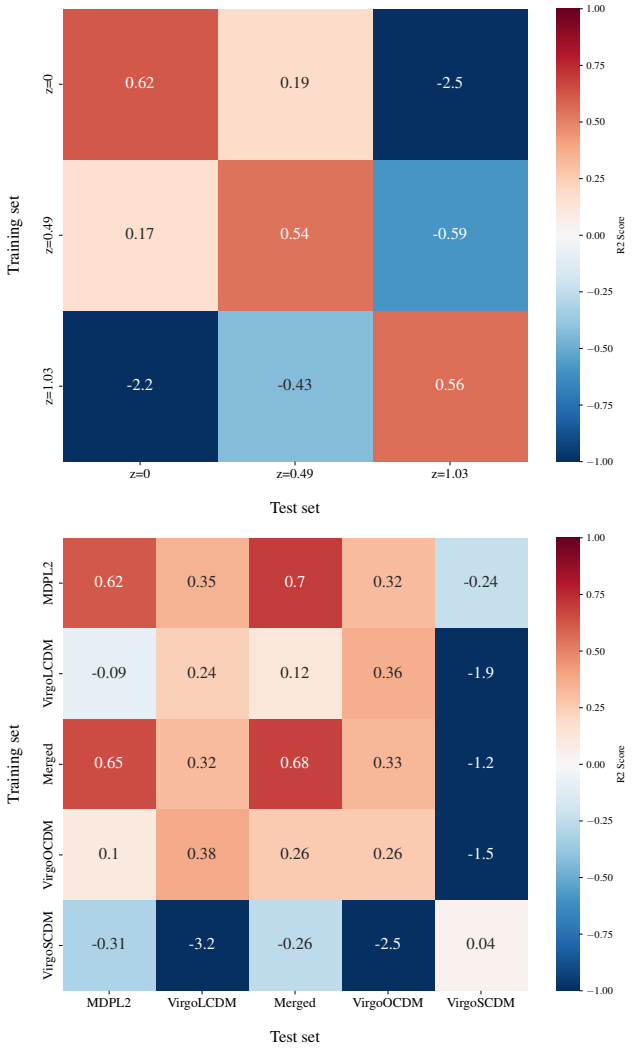

Figure 4.4: R^2 scores of different combinations of training and testing data using mass information. 80% of the simulation data were used for training-validation and 20% was used for the testing in each case. *Upper panel:* Training and testing data from different MDPL2 redshifts (z=0,0.49,1.03). As expected, for each training set the R^2 score is the highest when evaluated on data from the same redshift. Since MDPL2 is a ΛCDM concordance cosmological simulation, the $\rho_{\rm ta}$ depends on redshift. Therefore, for different redshifts the model cannot find an apparent correlation between $M_{\rm vir}$ and $R_{\rm ta}$ because $\rho_{\rm ta}$ varies with redshift. *Bottom panel:* Training and testing data from different simulations and cosmological models (MDPL2, VirgoΛCDM, VirgoOCDM, VirgoSCDM and the merged data from MDPL2 and VirgoΛCDM). Best models seem to be the ones containing the data from the MDPL2 possibly due to the fact of the larger number of training instances.

4.3 Velocity distribution

The result that the velocity data were not containing any information is surprising. Since the mass of the central halo is correlated with $R_{\rm ta}$ one would assume that the velocity dispersion of the halos surrounding it would be correlated to that mass and hence the turnaround radius. Consequently, we attempt to do a manual feature extraction in this case, since our models could be lacking complexity due to computational limitations or the architectures tested could have lead us to a local minimum since we did not explore the whole probability space.

We suppose that the velocity data are sparse to find a relation between them and the $R_{\rm ta}$. If this is the case, we attempt to "stack" our acquired projections based on their $R_{\rm ta}$ and therefore merge some data together in order to have more velocity information. This "stacking" was performed by separating the available $R_{\rm ta}$ into bins and merging the data from the corresponding clusters. A new $R_{\rm ta}$ was then considered as the mean of the $R_{\rm ta}$ values in each bin. In principle, similar mass, or similar $R_{\rm ta}$ clusters should exhibit similar characteristics. We then again transform these data into images but this time computing the standard deviation of the velocities in each bin.

In the left panel of Fig. 4.5 we can see the True vs Predicted values of the model with 3000 R_{ta} bins (resulting in about 1200 cluster images since some of the bins are empty): It appears to be a small correlation since it performs slightly better that our previous attempt (see Fig. 4.1) but not a very apparent one.

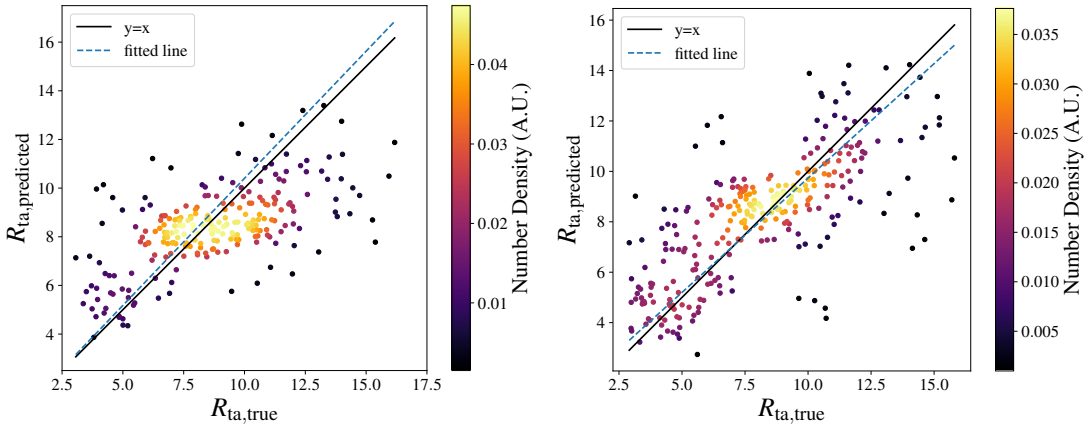

Figure 4.5: Comparison between predicted and true values of of the R_{ta} for models using velocity information."Stacked" images were used from 3000 bins of R_{ta} taken from: the MDPL2 data (*left panel*) and the merged MDPL2 with Virgo ΛCDM data (*right panel*). R^2 scores are 0.36 and 0.53. The merging of the data from the 2 simulations seems to prevent the predictions of the model focusing around the mean value of the MDPL2 $R_{\rm ta}$ values distribution.

We continue by merging the data from MDPL2 and VirgoΛCDM in an attempt to deal with the issue of heteroscedasticity. Although the simulation parameters slightly differ, in principle the turnaround should be probed independently of these minor differences. The results can be seen in the right panel of Fig. 4.5. An R^2 Score of 0.53 together with the visual representation of the predicted vs true values is a strong indication that there is correlation.

For reference, a Gaussian prediction around the mean value would give $R^2 \approx -1$.

In order to interpret these results we plotted the Shapley values of the pixels some of the test images shown in Fig. 4.6. We can see that in the case where the central pixel has a lower velocity standard deviation the algorithm interprets this as an indication to decrease the value of the predicted R_{ta} while the reverse holds true under inverse circumstances. This seems to validate the assumption that the velocity dispersion is correlated with the central halo's mass. However, these are 2 specific examples and Shapley values can only provide an indication. Nonetheless, the apparent focus in the central area of the cluster seems to appear in $\sim 90\%$ of the inspected test images.

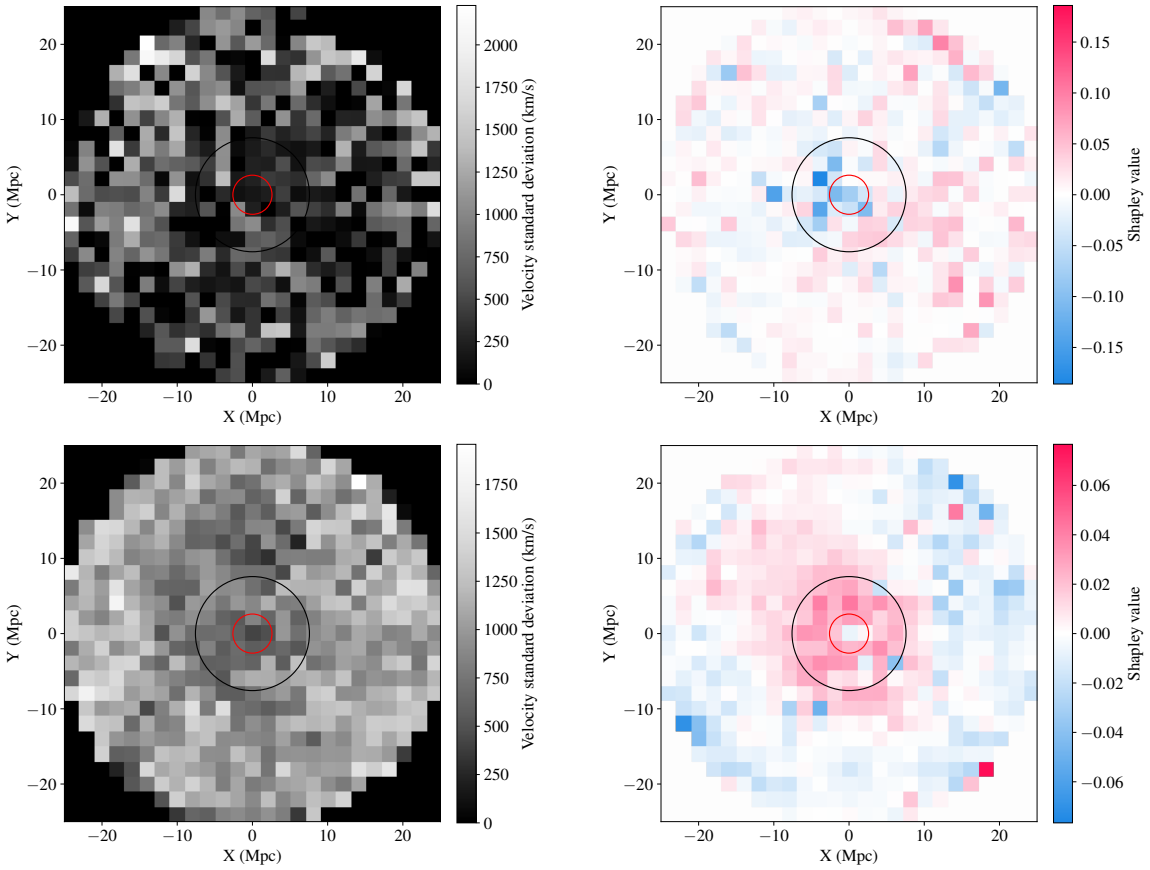

Figure 4.6: Shapley values of 2 random test clusters (upper and lower panel). *Left panel:* Standard deviation of the halo velocity in each pixel. *Right panel:* Shapley values in each pixel. The black and red circles correspond to the turnaround radius and the R_{200} respectively. In the first cluster, the center of the structure seems to have low dispersion compared to the second one. The low dispersion seems to indicate to the model that the turnaround radius should be lower while the higher dispersion indicates the opposite.

This observable emphasis of the CNN model in the central area of the structures needs to be investigated. Consequently, we removed all of the halos located inside the R_{200} of the central overdensity in each cluster and re-trained the model in order to probe the importance of this information. The resulting model's predicted vs true values plot can be seen in the right panel of Fig. 4.7. The model had an R^2 score of 0.60 which is surprisingly higher than the one containing the information inside R_{200} ($R^2 = 0.53$). This could be a result of

the latter reaching a local minimum (during the gradient decent algorithm) that we could not have escaped with the current architecture and network parameters. We argue that the information of the halos projected to the central area of our images, confuses the model due to the contamination we introduce with the consideration of velocity cuts. We deduce that the information inside the R_{200} of the central overdensity does not have particular importance.

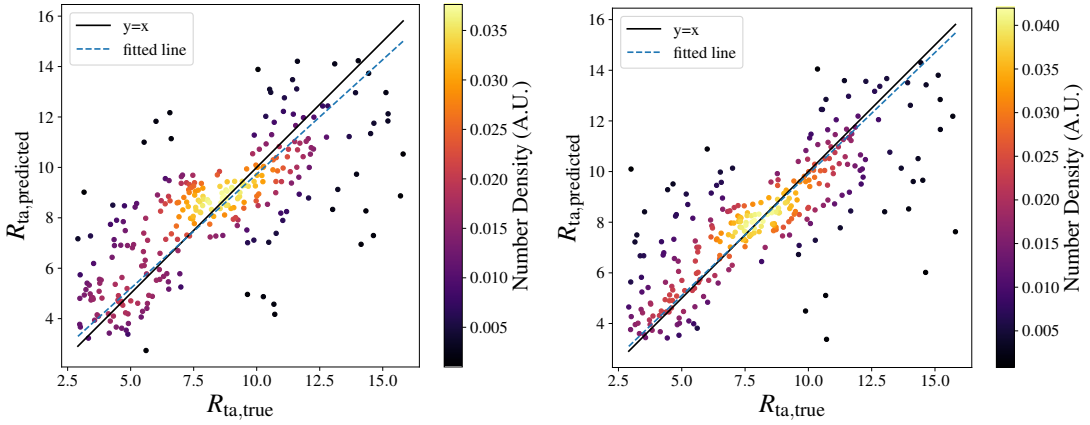

Figure 4.7: Comparison between predicted and true values of of the R_{ta} for models using velocity information including (*left panel*) or excluding (*right panel*) the information inside the central halo's R_{200}."Stacked" images were used from 3000 bins of R_{ta} taken from the merged MDPL2 with Virgo ΛCDM data. R^2 scores are 0.53 and 0.60 respectively.

A number of training runs was again performed with different combinations of training and test sets. The results of the R^2 score from different redshifts from MDPL2 and from different cosmologies from both MDPL2 and the Virgo Simulations can be seen in Fig. 4.8 (upper and lower panel respectively). It seems that the models containing the MDPL2 in the training data were more successful, as expected, due to the larger amount of training instances. The R^2 scores from different redshifts follow the same pattern as the ones in the upper panel of Fig. 4.4 resulting from the CNN trained with the mass information. Additionally, there seems to be no successful prediction of a model trained on a specific cosmology to when tested to another. All of the aforementioned features indicate that the velocity dispersion information used by the CNN to predict the $R_{\rm ta}$ correlates to the mass of the central overdensity. Therefore, the model cannot predict the $R_{\rm ta}$ successfully in other redshifts and cosmologies for the same reason the mass is unable to do so mentioned in Sect 4.2.

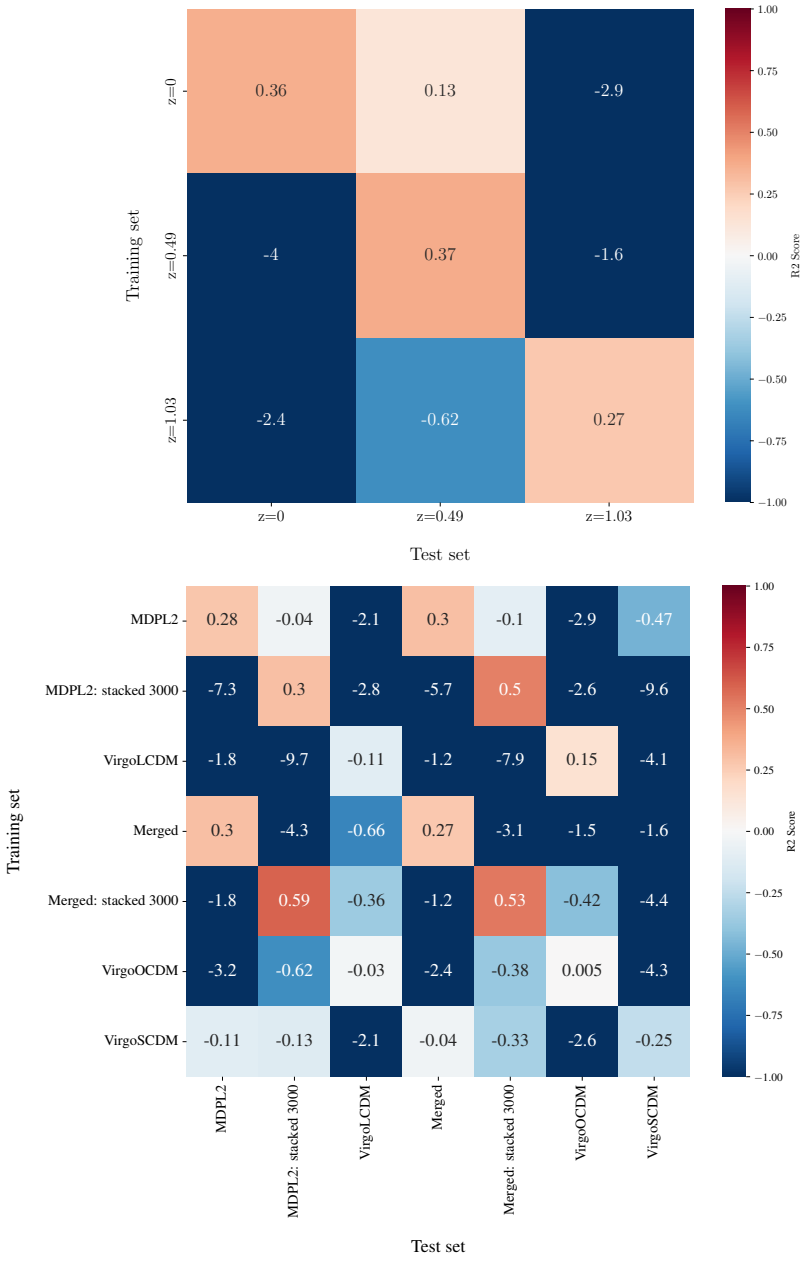

Figure 4.8: R^2 scores of different combinations of training and testing data using velocity information. 80% of the simulation data were used for training-validation and 20% was used for the testing in each case. *Upper panel:* Training and testing data from different MDPL2 redshifts (z=0,0.49,1.03). For each training set the R^2 score is the highest when evaluated on data from the same redshift. *Bottom panel:* Training and testing data from different simulations and cosmological models (MDPL2, VirgoΛCDM, VirgoOCDM, VirgoSCDM and the merged data from MDPL2 and VirgoΛCDM). Best models seem to be the ones containing the data from the MDPL2 possibly due to the fact of the larger number of training instances.

Chapter 5

Conclusions

It is established by the work of Pavlidou & Tomaras, 2014; Pavlidou et al., 2020; Korkidis et al., 2020 that the turnaround density of galaxy clusters can be used as a novel way to probe cosmology and further test the current cosmological model that in the recent years is faced with 'tensions' regarding the observations of the Hubble Constant and the clumpiness factor.

Here, we searched for a way to probe the turnaround radius of galaxy clusters based on their projections on the plane of the sky. We use cosmological N-body simulations to acquire idealized observational data (containing mass and velocity information, free of uncertainties) inside 2 "slices" with thickness corresponding to what would be achievable in spectroscopic and in photometric surveys. We infer the "true" values of the turnaround radii from the velocity profiles surrounding the central overdensities of the galaxy clusters in the simulations. We use Convolutional Neural Networks to train a model that predicts the turnaround radius using different combinations of training and test data from different simulations, redshifts and cosmologies.

We find that the mass is the most important information we have for the prediction of the turnaround radius. This at first may appear counter-intuitive as the turnaround is defined and computed by the velocity profiles of the dark matter particles around the central halo. However, it is quite logical that the mass of the cluster contributes to the gravitational dynamics defining the turnaround radius. Furthermore, we find that the mass of the central, more massive halo, is almost entirely contributing to this prediction.

The velocity dispersion of the galaxies around the cluster also can be used as an important feature since it is correlated with the mass of the central overdensity. However, the small number of halos large enough to be observed inside a single region around a cluster is not enough to accurately and easily describe this dispersion, since we used "stacked" images to find the aforementioned correlation combining several projections. Potentially, the velocity dispersion of the halos around a cluster hides additional information.

As expected, the contamination of the photometric redshifts is fairly significant, especially regarding the velocity information but also in the case of the mass information since by possibly including halos as massive as the central one in the final projection, the model is affected negatively. Consequently potential spectroscopic observations, that are related to the low velocity cut, would result in a more accurate prediction of the turnaround radius.

All of the models, when tested on data that were acquired from different redshifts and different cosmologies were not successful in their predictions. For the models trained with the mass information this is expected since the turnaround radius scales with the mass for a specific redshift and cosmology. For the velocity trained models, unfortunately, the same pattern appears, indicating that the model, indeed, simply correlates the mass of the cluster with the velocity dispersion of the halos around it.

Several improvements can be made in future investigations for this matter. Perhaps more "stacked" images with higher resolution may be able to reveal information independent of the redshift and cosmological model. The use of bayesian deep learning or bias-reducing techniques such as the one described in Lin et al., 2022 could become particularly useful. Furthermore, more complex models e.g. with the use of inception layers in the architecture of the neural network might be able to unravel additional information hidden in the mass and velocity information of galaxy clusters.

References

Adhikari, S., & Huterer, D. (2019). A new measure of tension between experiments. *Journal of Cosmology and Astroparticle Physics*, *2019*(01), 036.

Andreon, S., & Hurn, M. (2013). Measurement errors and scaling relations in astrophysics: a review. *Statistical Analysis and Data Mining: The ASA Data Science Journal*, *6*(1), 15–33.

Baron, D. (2019). Machine learning in astronomy: A practical overview. *arXiv preprint arXiv:1904.07248*.

Behroozi, P. S., Wechsler, R. H., & Wu, H.-Y. (2012). The rockstar phase-space temporal halo finder and the velocity offsets of cluster cores. *The Astrophysical Journal*, *762*(2), 109.

Bernal, J. L., Verde, L., & Riess, A. G. (2016). The trouble with h0. *Journal of Cosmology and Astroparticle Physics*, *2016*(10), 019.

Bertschinger, E. (1985). Self-similar secondary infall and accretion in an einstein-de sitter universe. *The Astrophysical Journal Supplement Series*, *58*, 39–65.

Carleo, G., Cirac, I., Cranmer, K., Daudet, L., Schuld, M., Tishby, N., ... Zdeborová, L. (2019). Machine learning and the physical sciences. *Reviews of Modern Physics*, *91*(4), 045002.

Davis, M., Efstathiou, G., Frenk, C. S., & White, S. D. M. (1985). The evolution of large-scale structure in a universe dominated by cold dark matter. *The Astrophysical Journal*, *292*, 371. doi: 10.1086/163168

Di Valentino, E., Melchiorri, A., & Silk, J. (2020). Planck evidence for a closed universe and a possible crisis for cosmology. *Nature Astronomy*, *4*(2), 196–203.

Di Valentino, E., Melchiorri, A., & Silk, J. (2021). Investigating cosmic discordance. *The Astrophysical journal letters*, *908*(1), L9.

Feigelson, E. D., De Souza, R. S., Ishida, E. E., & Babu, G. J. (2021). Twenty-first-century statistical and computational challenges in astrophysics. *Annual Review of Statistics and Its Application*, *8*, 493–517.

Fillmore, J. A., & Goldreich, P. (1984). Self-similar gravitational collapse in an expanding universe. *Astrophysical Journal*, *281*(1), 1–8.

Frenk, C., Colberg, J., Couchman, H., Efstathiou, G., Evrard, A., Jenkins, A., ... others (2000). Public release of n-body simulation and related data by the virgo consortium. *arXiv preprint astro-ph/0007362*.

Gunn, J. E., & Gott III, J. R. (1972). On the infall of matter into clusters of galaxies and some effects on their evolution. *The Astrophysical Journal*, *176*, 1.

Handley, W. (2021). Curvature tension: evidence for a closed universe. *Physical Review D*, *103*(4), L041301.

Hildebrandt, H., Viola, M., Heymans, C., Joudaki, S., Kuijken, K., Blake, C., ... others (2017). Kids-450: Cosmological parameter constraints from tomographic weak gravitational lensing. *Monthly Notices of the Royal Astronomical Society*, *465*(2), 1454–1498.

Huertas-Company, M., & Lanusse, F. (2022). The dawes review 10: The impact of deep learning for the analysis of galaxy surveys. *arXiv preprint arXiv:2210.01813*.

Joudaki, S., Mead, A., Blake, C., Choi, A., de Jong, J., Erben, T., ... others (2017). Kids-450: Testing extensions to the standard cosmological model. *Monthly Notices of the Royal Astronomical Society*, *471*(2), 1259–1279.

Knebe, A., Knollmann, S. R., Muldrew, S. I., Pearce, F. R., Aragon-Calvo, M. A., Ascasibar, Y., ... Zemp, M. (2011). Haloes gone MAD: The Halo-Finder Comparison Project. *Monthly Notices of the Royal Astronomical Society*, *415*(3), 2293–2318. doi: 10.1111/j.1365-2966.2011.18858.x

Korkidis, G., Pavlidou, V., Tassis, K., Ntormousi, E., Tomaras, T. N., & Kovlakas, K. (2020). Turnaround radius of galaxy clusters in N-body simulations. *Astronomy & Astrophysics*, *639*, A122. doi: 10.1051/0004-6361/201937337

Lahav, O. (2023). Deep machine learning in cosmology: Evolution or revolution? *arXiv preprint arXiv:2302.04324*.

Lin, Q., Fouchez, D., Pasquet, J., Treyer, M., Ouahmed, R. A., Arnouts, S., & Ilbert, O. (2022). Photometric redshift estimation with convolutional neural networks and galaxy images: A case study of resolving biases in data-driven methods. *arXiv preprint arXiv:2202.09964*.

Miville-Deschênes, M., Pettorino, V., Bucher, M., Delabrouille, J., Ganga, K., Le Jeune, M., ... others (2020). Planck 2018 results: Vi. cosmological parameters. *Astronomy and Astrophysics*, *641*, A6–A6.

Motloch, P., & Hu, W. (2018). Tensions between direct measurements of the lens power spectrum from planck data. *Physical Review D*, *97*(10), 103536.

Pavlidou, V., Korkidis, G., Tomaras, T. N., & Tanoglidis, D. (2020). Turnaround density as a probe of the cosmological constant. *Astronomy & Astrophysics*, *638*, L8. doi: 10.1051/0004-6361/201937358

Pavlidou, V., & Tomaras, T. (2014). Where the world stands still: turnaround as a strong test of ΛCDM cosmology. *Journal of Cosmology and Astroparticle Physics*, *2014*(09), 020–020. doi: 10.1088/1475-7516/2014/09/020

Raveri, M., & Hu, W. (2019). Concordance and discordance in cosmology. *Physical Review D*, *99*(4), 043506.

Riebe, K., Partl, A. M., Enke, H., Forero-Romero, J., Gottlöber, S., Klypin, A., ... others (2013). The multidark database: release of the bolshoi and multidark cosmological simulations. *Astronomische Nachrichten*, *334*(7), 691–708.

Riess, A. G., Casertano, S., Yuan, W., Macri, L., Anderson, J., MacKenty, J. W., ... others (2018). New parallaxes of galactic cepheids from spatially scanning the hubble space telescope: Implications for the hubble constant. *The Astrophysical Journal*, *855*(2),

136.

Riess, A. G., Casertano, S., Yuan, W., Macri, L. M., & Scolnic, D. (2019). Large magellanic cloud cepheid standards provide a 1% foundation for the determination of the hubble constant and stronger evidence for physics beyond λcdm. *The Astrophysical Journal*, *876*(1), 85.

Shah, P., Lemos, P., & Lahav, O. (2021). A buyer's guide to the hubble constant. *The Astronomy and Astrophysics Review*, *29*, 1–69.

Tanoglidis, D., Pavlidou, V., & Tomaras, T. (2015). Testing λcdm cosmology at turnaround: where to look for violations of the bound? *Journal of Cosmology and Astroparticle Physics*, *2015*(12), 060.

Zhao, G.-B., Raveri, M., Pogosian, L., Wang, Y., Crittenden, R. G., Handley, W. J., ... others (2017). Dynamical dark energy in light of the latest observations. *Nature Astronomy*, *1*(9), 627–632.

Appendix A

Neural Network architecture

The NN architecture used can be seen in Table. A.1. During training, the Mean Absolute Error was used to evaluate the model's performance on the validation set. In addition to this basic architecture, the hyperparameters were tuned in each case to obtain the best result. In most of the cases the following additional elements were used: Activation function: ReLU, Padding: same, Kernel initializer: He initialization, Learning rate: 10^{-3}. In some of the cases l1 and l2 regularizations were also implemented.

Table A.1: Characteristics of each layer of the CNN architecture. Columns are: name of the layer, input layer, size of the convolution kernel (in pixels), size (height × width in pixels) and number of the resulting feature maps.

Layer	Inputs	Kernel size	$h \times w$	#feature maps
Convolution (C1)	Input image	5×5	25×25	32
Activation (A1)	C1	-	25×25	32
Batch Normalization (BN1)	A1	-	25×25	32
Pooling (P1)	BN1	2×2 (stride 1 pix)	12×12	32
Dropout (D1)	P1	-	12×12	32
Convolution (C2)	D1	5×5	25×25	64
Activation (A2)	C2	-	25×25	64
Batch Normalization (BN2)	A2	-	25×25	64
Pooling (P2)	BN2	2×2 (stride 1 pix)	12×12	64
Dropout (D2)	P2	-	12×12	64
Flatten (F1)	D2	-	2304	-
Fully Connected (FC1)	F1	-	32	-
Activation (A3)	FC1	-	32	-
Batch Normalization (BN3)	A3	-	32	-
Dropout (D3)	BN3	-	32	-
Fully Connected (FC2)	D3	-	16	-
Activation (A4)	FC2	-	16	-
Dropout (D4)	A4	-	16	-
Fully Connected (FC3)	D4	-	1	-