MASARYKOVA UNIVERZITA
Přírodovědecká fakulta
Ústav teoretické fyziky a astrofyziky

DIPLOMOVÁ PRÁCE

Brno 2019

Matej Kosiba

MASARYKOVA
UNIVERZITA

Přírodovědecká fakulta
Ústav teoretické fyziky a astrofyziky

Klasifikace kup galaxií pomocí konvolučních neuronových sítí

Diplomová práce

Matej Kosiba

Vedoucí diplomové práce:
Mgr. Filip Hroch, Ph.D.

Konzultant:
Mgr. Maggie Lieu, Ph.D.

Brno 2019

Bibliografický záznam

Autor:	Bc. Matej Kosiba Přírodovědecká fakulta, Masarykova univerzita Ústav teoretické fyziky a astrofyziky
Název práce:	Klasifikace kup galaxií pomocí konvolučních neuronových sítí
Studijní program:	Teoretická fyzika a astrofyzika
Směr:	Astrofyzika
Vedoucí práce:	Mgr. Filip Hroch, Ph.D.
Konzultant:	Mgr. Maggie Lieu, Ph.D.
Akademický rok:	2018/2019
Počet stran:	viii + 62
Klíčová slova:	kupy galaxií, galaxie, morfologie, klasifikace, neuronová síť

Bibliographic record

Author:	Bc. Matej Kosiba Faculty of Science, Masaryk University Department of Theoretical Physics and Astrophysics
Title of Thesis:	Classification of galaxy clusters with convolutional neural networks
Degree Programme:	Theoretical physics and astrophysics
Field of study:	Astrophysics
Supervisor:	Mgr. Filip Hroch, PhD
Advisor:	Mgr. Maggie Lieu, PhD
Academic Year:	2018/2019
Number of Pages:	viii + 62
Keywords:	galaxy clusters, galaxy, morphology, classification, neural network

Abstrakt

V současné době je klasifikace kandidátů na galaktické kupy obvykle prováděna vizuálně, což je pomalý a neefektivní způsob, navíc spojený s menší reprodukovatelností výsledků. K řešení tohoto problému může napomoci moderní přístup založený na metodě konvolučních neuronových sítí s vestavěnými supermoderními klasifikátory obrázkových dat, které v této studii aplikujeme na případ automatické klasifikace kup galaxií. Využili jsme techniky zvané transfer-learning, nejlepších výsledků jsme ovšem dosáhli pomocí vlastní neuronové sítě. Použité neuronové sítě byly trénovány na kombinaci rentgenových obrázků z družice XMM-Newton a jejich optických protějšků ve druhé Palomarské fotografické přehlídce oblohy (DSS2). Kvalita sítě byla testována na množině 85 spektroskopicky potvrzených kup galaxií a 85 objektů odborníky naopak nepovažovaných za galaktické kupy. Při binární klasifikaci, na kupy galaxií a ostatní objekty, sestavená síť správně rozpoznala 77 kup galaxií a 83 objektů, jež nejsou kupami galaxií, čímž dosáhla přesnosti 94%.

Dále v práci představujeme Zooniverse projekt — The Hunt for Galaxy Clusters — určený pro širokou veřejnost ve kterém dobrovolníci online ručně klasifikovali kandidáty na kupy galaxií. Projekt obsahuje 1 600 objektů, každý s vyžadovaným minimem 30 klasifikací. Do projektu se zapojilo 1 227 dobrovolníků. Porovnáním jejich odpovědí s odpovědmi od profesionálů na vzorku 404 kandidátů na kupy galaxií se zjistilo, že dobrovolníci tíhli ke klasifikování objektů jako by nešlo o kupy galaxií, avšak vytvořili extrémně čistý vzorek kup galaxií, s vynikající přesností klasifikace 99% po porovnání s experty.

Abstract

Galaxy cluster candidates are usually classified by hand, making it a slow and inefficient process with biases impossible to model. We tackle this problem with a novel approach, using convolutional neural networks, state-of-the-art image classification tool, for automatic galaxy cluster classification. We train the networks on combined XMM-Newton's X-ray observations with their optical counterparts from DSS2. The networks were tested on a sample of 85 spectroscopically confirmed clusters and 85 objects classified as non-clusters by the experts. We have created our own custom network and also used the transfer learning approach. The network we hand-made from scratch achieved the best performance in binary classification to classes *cluster* and *non-cluster*, correctly classifying 77 clusters and 83 non-clusters acquiring accuracy 94.1 %.

We also introduce an online official Zooniverse citizen science project, The Hunt for Galaxy Clusters, in which the citizen volunteers were asked to classify candidate clusters by hand. The project contained 1 600 cluster candidates, each with a threshold of 30 classifications. 1 227 volunteers participated in the project. After testing their agreement with the experts on a sample of 404 objects, they turned to be biased towards classifying candidate clusters as non-clusters but they created a very pure sample with precision 99 % after comparing with the experts. The Zooniverse volunteers found 506 clusters in the whole data set.

MASARYKOVA UNIVERZITA
Přírodovědecká fakulta

ZADÁNÍ DIPLOMOVÉ PRÁCE

Akademický rok: 2018/2019

Ústav: Ústav teoretické fyziky a astrofyziky

Student: Bc. Matej Kosiba

Program: Fyzika

Obor: Teoretická fyzika a astrofyzika

Směr: Astrofyzika

Ředitel *Ústavu teoretické fyziky a astrofyziky* PřF MU Vám ve smyslu Studijního a zkušebního řádu MU určuje diplomovou práci s názvem:

Název práce: Klasifikace galaktických kup pomocí neuronových sítí

Název práce anglicky: Classification of galaxy clusters by neural networks

Oficiální zadání:
Large volumes of high quality imaging data are being acquired by ongoing large astronomical surveys such as Sloan Digital Sky Survey or Dark Energy Survey. Even more, there are some upcoming research projects, like Large Synoptic Sky Survey, observing in optical wavebands, or eROSITA, the future X-ray satellite, which are expected to detect a lot of currently unknown objects, and also galaxy clusters among them; and galaxy clusters are a subject of matter of the proposed project.

The acquired data will provide an invaluable information for deep understanding of the Universe. Future projects should study galaxy clusters themselves, their evolution, and subsequently, giving constrains on some cosmological parameters. First step towards such tasks is to find and to classify galaxy clusters, moreover without any human touch. To accomplish these aims, solid skills in advanced machine learning techniques are required.

The student should familiarise with available multi-wavelength observations, sky surveys and methods commonly used in astronomy. The student will develop a method based on convolutional neural networks, to recognise galaxy clusters by their appearance on regular sky frames in distinct spectral bands. Some advanced techniques, like transfer learning, or citizen science project frameworks, are supposed to be utilised.

Thesis adviser will be Mgr. Maggie Lieu, PhD.

Jazyk závěrečné práce: angličtina

Vedoucí práce: Mgr. Filip Hroch, Ph.D.

Datum zadání práce: 3. 4. 2019

V Brně dne: 3. 4. 2019

Souhlasím se zadáním (podpis, datum):

.....	.....	.....
Bc. Matej Kosiba	Mgr. Filip Hroch, Ph.D.	prof. Rikard von Unge, Ph.D.
student	**vedoucí práce**	**ředitel Ústavu teoretické fyziky a astrofyziky**

Poďakovanie

V prvom rade sa chcem poďakovať mojim vedúcim, Filipovi Hrochovi a Maggie Lieu, za ich veľkú pomoc s písaním práce, častými diskusiami o napredovaní projektu a vynikajúcimi odbornými radami. Ich vždy priateľský prístup vo mne podnecoval moju hlbokú motiváciu a prinášal mi radosť pri práci na mojom projekte. Tiež sa chcem poďakovať kolaborácii X-CLASS, menovite Brunovi Altierimu, Ivanovi Valtchanovovi, Nicolasovi Clercovi, Lorenzovi Facciolimu a Tatyane Sadibekovej, za ich rozsiahlu pomoc s prípravou dát k trénovaniu neurónových sietí. Poďakovanie patrí aj Norbertovi Wernerovi, ktorý ma ako vedúci mojej bakalárskej práce priviedol k využívaniu umelej inteligencii v astrofyzike a Lukášovi Burgetovi, ktorý mi venoval počas môjho bakalárskeho štúdia nesmierne množstvo času a naučil ma od základov programovať umelú inteligenciu. Obom ďakujem taktiež za ich odborné rady k mojej diplomovej práci.

Ďakujem svojej priateľke Eve Šikutovej za rešpektovanie môjho pracovného režimu, keď ma celé týždne nevidela chodiť spávať ani vstávať a že dokázala nájsť pochopenie pre moju častú duševnú neprítomnosť počas našich rozhovorov, keď som myšlienkami nedobrovoľne odbiehal k práci a nedokázal jej venovať pozornosť, ktorú by si zaslúžila. Rád by som sa poďakoval aj Xene a Íris Hamasovým, verným sprievodkyňám môjho štúdia, ktoré ma vždy dokázali vypočuť a pochopiť. V neposlednom rade sa chcem poďakovať mojej rodine, na čele s mojím bratom Tomášom a otcom Petrom, ktorá ma vždy podporovala a vkladala neochvejnú dôveru v moje úspechy.

Prohlášení

Prohlašuji, že jsem svoji diplomovou práci vypracoval samostatně s využitím informačních zdrojů, které jsou v práci citovány.

Brno 16. května 2019

Podpis autora

Contents

1 Galaxy clusters **1**
- 1.1 Introduction . . . 1
- 1.2 Formation of galaxy clusters . . . 2
- 1.3 Optical observations . . . 3
  - 1.3.1 Optical morphological classifications of galaxy clusters . . . 4
  - 1.3.2 Optical automatic machine detection of galaxy clusters . . . 7
- 1.4 X-ray observations . . . 8
  - 1.4.1 X-ray properties of galaxy clusters . . . 9
  - 1.4.2 X-ray automatic machine detection of galaxy clusters . . . 10

2 Method **12**
- 2.1 Neural networks . . . 12
- 2.2 Convolutional neural networks . . . 14
- 2.3 Transfer learning . . . 16
- 2.4 Performance measurements . . . 16

3 Data and The Hunt for Galaxy Clusters **18**
- 3.1 XAmine pipeline . . . 18
- 3.2 Data preprocessing . . . 19
- 3.3 Data augmentation . . . 20
- 3.4 The Hunt for Galaxy Clusters . . . 20
- 3.5 Weighting volunteers classifications . . . 21
- 3.6 Classification schemes . . . 21

4 Results and Discussion **25**
- 4.1 The Hunt for Galaxy Clusters results . . . 25
- 4.2 Training and performance . . . 26
- 4.3 Images of classified objects . . . 42
- 4.4 Visualisation of filter's activations . . . 42

5 Conclusion and future plans **49**
- 5.1 The Hunt for Galaxy Clusters . . . 49
- 5.2 Neural networks . . . 50
- 5.3 Future plans . . . 51

Bibliography **53**

Chapter 1

Galaxy clusters

1.1 Introduction

Galaxy clusters are the largest virialised systems in the observable Universe. An ordinary galaxy cluster has a diameter about a few Mpc. Clusters typically contain 10's – 100's of galaxies above a magnitude limit. Main source of clusters radiation in optical wavelengths is an integrated stellar light coming from clusters galaxies, which are however only about $\sim 2\,\%$ of clusters mass. Clusters X-ray emission is dominated by an intra cluster gas, containing $\sim 12\,\%$ of clusters mass, with temperatures of $10^7 - 10^8\,\mathrm{K}$ and a typical particle number density of $10^{-1} - 10^{-4}\,\mathrm{cm}^{-3}$. This fully ionized plasma is not associated with individual cluster's galaxies, but creates the intra-cluster medium (ICM). Temperature of the ICM is consistent with velocities of galaxies, indicating that the gas and galaxies are in thermodynamic equilibrium inside of the common gravitational potential well. Combined masses of galaxies and ICM can not explain large cluster's gravitational potential. Nowadays leading theory to explain this problem is the dark matter, contributing to $\sim 86\,\%$ of cluster mass. Total mass of a typical galaxy cluster is about $10^{13} - 10^{15}\,\mathrm{M}_\odot$.

Being the densest regions in the Universe makes galaxy clusters ideal astrophysical laboratories for studying interactions of galaxies with themselves and also with the ICM. Effects like ram pressure stripping, tidal forces on large scales, enrichment of ICM by elements of star production of galaxies via supernovae explosions, AGN feedback with ICM or temperature transfer into ICM are being intensively studied.

Energy of photons of the cosmic microwave background (CMB) has been shifted to very low energies since they traveled from the epoch of reionization (Penzias and Wilson (1965), Dicke et al. (1965)). CMB shows remarkable isotropy of its radiation, with perturbations on the level of 10^{-5} discovered by the COBE satellite. As the photons of the CMB pass through ICM, they substract energy from ICM electrons via inverse Compton scattering, creating small distortions in the black-body spectrum of the CMB, giving rise to Sunyaev-Zeldovich effect (SZE, Sunyaev and Zeldovich (1972)), when observed in the direction towards galaxy cluster. SZE is a powerful tool to detect clusters into large cosmological distances given that it is basically just a measurement of the intensity of CMB.

Gravitational lensing can be used to directly probe the distribution of all of the matter of galaxy clusters. This effect is strong in the inner regions of clusters, nicely visible in the distorted images of foreground galaxies appearing as curved long thin arcs around the cluster centre. Images of foreground galaxies at larger radii from the centre of cluster are also affected by gravitational lensing. Their shapes are slightly distorted, measurable when statistically compared to the distribution of their shapes expected from an isotropic distribution. This effect is called weak gravitational lensing, which points out that the dark matter extends well beyond the regions filled by the X-ray emitting hot intra-cluster gas. Both, strong and weak gravitational lensing provide great tools to compare other clusters mass measuring methods and make galaxy clusters perfect dark matter probes.

Galaxy clusters are one of important cosmology probes. With increasing redshift we probe larger volume of space, expecting to find more clusters. However, further away the cluster is, less extended and fainter its emission appears, resulting in a bias towards larger and more massive clusters. Another uncertainity comes from the manual confirmation of a cluster's detection, which is typically done by few scientists, resulting in biases impossible to model, making cluster selection function very messy.

Development of an automatic cluster classification software is inevitable to obtain reliable cluster selection function, with which it will be possible to make tighter constrains on cosmological parameters. Most of cluster catalogues have been done using only single wavelength observations, typically optical (e.g. (Abell, 1958), (Miller et al., 2005), Gal et al. (2003) or (Gilbank et al., 2011)) or X-ray (e.g. Jones and Forman (1992) or Clerc et al. (2012b)). Both of those have different advantaged and biases. Combinaton of the two wavelengths will increase the performance of automatic selection method.

1.2 Formation of galaxy clusters

The standard hierarchical structure formation scenario describes formation of structures in the Universe via gravitational collapse of fluctuations in the initial primordial density field. Galaxy clusters arise from the largest of those fluctuations, being formed by a hierarchical sequence of mergers and by accretion of smaller systems, mostly driven by the gravity of the dark matter, dominating the gravitational field. Properties of the initial field depend on specific processes occurring during the early inflationary stage of evolution of the Universe (Bardeen et al. (1983), Guth and Pi (1982), Starobinsky (1982)) and conditions prior to recombination (Bardeen et al. (1986), Bond and Efstathiou (1984), Eisenstein and Hu (1999), Peebles (1982)).

The simplest model of a non-linear collapse of initial density fluctuations assumes them to be spherically-symmetric perturbations of radius R with constant density. It provides useful approximations of the time scale of a non-linear halo collapse and is often used for statistical studies of halo populations.

The complexities of many important details which this model misses are explored with 3D numerical cosmological simulations (Evrard (1988) made one of first simulations which followed dynamics of both, baryons and dark matter. Examples of more recent simulations are Habib et al. (2016) or Pillepich et al. (2018)). The results of those simulations show that the initial density fluctuations do not collapse uniformly. On the contrary, different regions of the density fluctuation collapse with various time scales, because real fluctuations do not have constant density and well defined boundary. Simulations show that fluctuations have a radial profile and a curvature (Bardeen et al. (1986), Dalal et al. (2008)). Different regions of the fluctuation collapse at different times, so the fluctuation has multiple collapse timescales, making it extended in time (Diemand et al., 2007).

Moreover, density fluctuations in the primordial field are surrounded by other ones and different density inhomogeneities. The matter surrounding density fluctuations has been sculptured into filamentary structure by tidal forces of the most massive density fluctuations, forming connections between them (Bond et al., 1996). Galaxy clusters continue accreting matter from those filaments, which becomes the main matter accretion process in the late epochs of clusters lives.

Evolution of clusters depends on the matter density parameter Ω_m (e.g. (Oukbir and Blanchard, 1992), Eke et al. (1998)). Observed space density of distant galaxy clusters can be used as a powerful cosmological probe, because universes with different density show very different clusters evolutionary patterns. Cosmological simulations are used to predict galaxy clusters number density of a given mass across different redshifts for specific initial cosmological parameters. Comparing results of simulations with observations helps to constrain cosmological parameters, giving information of the Universe we live in.

1.3 Optical observations

First galaxy cluster catalogue (Abell, 1958) was created by George Ogden Abell in 1958. He looked for galaxy clusters as clustering of galaxies in optical wavelengths on photographic plates of National Geographic Society — Palomar Observatory Sky Survey (Minkowski and Abell, 1963). Abell's first criterion for cluster selection uses magnitude of third brightest galaxy in a cluster, m_3, clusters of galaxies were systems which contained at least 50 galaxies between magnitude range m_3 and m_3+2, estimating that third brightest galaxy should have similar absolute magnitude in most clusters. Second Abell's criterium was, that those galaxies should be contained in a circle of radius $R_A = 1.7/z$ arcmin, where z is an estimated redshift and Abell's third criterion was, that clusters estimated redshift should be in the range $0.02 \leq z \leq 0.2$.

Inner parts of an older, already virialised and regular galaxy clusters are dominated by elliptical and lenticular galaxies, however the opposite is true in case of newly forming irregular clusters and in the field, where dominates

population of spiral galaxies. The morphology of galaxies is obviously altered by the environment they are born in. The two main important processes which influence the evolution of galaxies in clusters are galaxy mergers and interaction of galaxies with ICM. Tremaine et al. (1975), Gunn and Tinsley (1976) and Tovmassian and Andernach (2012) suggest cannibalism scenario of formation of cD galaxies, in which dynamical friction decays orbits of massive cluster galaxies, making them eventually merge into one central dominant (cD) galaxy, defined by Matthews et al. (1964) as a very luminous ellipticals surrounded by a large diffuse optical halo of a low surface brightness. Gallagher and Ostriker (1972), Richstone (1975) and Richstone (1976) hypothesised that the halo is created by stripped material of clusters galaxies during their passing through clusters core. The stripped material would condense into the centre of cluster, settle down around the cD galaxy. Postman and R. Lauer (1995) argue that cD galaxies can be used as a standard candles for large cosmological distances because of their narrow colour and photometric distribution, unique rich cluster environment and lack of second-parameter effects.

1.3.1 Optical morphological classifications of galaxy clusters

Several different properties of galaxy clusters have been used to create their morphological classifications based on optical observations. The primary factor is projected 2D distribution of galaxies. Properties such as shape, richness, clumpiness, galactic content, etc. are often considered.

The first galaxy cluster catalogue (Abell, 1958) divided clusters into *rich* vs. *poor* and *regular* vs. *irregular*, shortly followed by classification of Zwicky et al. (1961) who separated clusters into *compact*, *medium-compact* and *open*.

Bautz and Morgan (1970) based their classification system on the dominance of clusters brightest galaxies. Bautz-Morgan *type I* clusters optical emission is dominated by a single central dominant (cD) galaxy. Bautz-Morgan *Type II* clusters have their brightest galaxies intermediate between cD and normal giant ellipticals. Bautz-Morgan *type III* clusters do not have dominant galaxies at all, while *type I-II* and *type II-III* are intermediate types.

Rood and Sastry (1971) created the "tuning fork" to classify rich clusters of galaxies (Fig. 1.1). They used the distribution and nature of the ten brightest cluster galaxies to classify clusters into 6 classes. Cluster of Rood-Sastry *cD* class is dominated by a central cD galaxy, which are often found in centres of clusters and groups of galaxies. Class *B* (binary) clusters are dominated by two very luminous galaxies. Class *L* (line) contains clusters which have at least 3 of the brightest galaxies appearing to be aligned in one line. Class *C* (core), are clusters with a core formed by at least four of the ten brightest clusters galaxies of comparable galaxy separation. Class *F* (flat) are clusters which galaxies form a flattened distribution on the sky. Class *I* (irregular) clusters do not have any obvious core or centre, galaxies appear to be distributed irregularly.

Based on numerical N-body simulations of collapse of galaxy cluster (White (1976), Carnevali et al. (1981), Farouki et al. (1983)), and on the content of

galaxy clusters and trends in the galaxy distribution, Struble and Rood (1984) revised Rood-Sastry classification, proposing a "split-linear" scheme describing the evolution of clusters from irregular to regular clusters (Fig.1.1).

Morgan (1961) and Oemler (1974) have used galactic content, fraction of spiral galaxies (S), lenticular galaxies (S0) and elliptical galaxies (E) of galaxy clusters for their classification. Morgan (1961) classified clusters into two types, type *i* for galaxy clusters containing large number of spiral galaxies and type *ii* for clusters with a small number of spiral galaxies. Oemler (1974) augmented Morgan's system, creating three classes of clusters. *Spiral-rich* class of clusters in which the majority of galaxies are spiral, *spiral-poor* clusters, in which the majority of galaxies are lenticular S0 galaxies with spiral galaxies being less common and *cD* class of clusters, in which the majority of galaxies are ellipticals, also having central dominant cD galaxy.

Rostagni et al. (2011) used the SDSS C4 cluster catalogue (Miller et al., 2005) to classify clusters according to their regularity into 5 categories: *relaxed clusters*, *bimodal major mergers*, *minor mergers*, *multiple mergers* and *clusters in virialisation phase*.

Those classification schemes are found to be highly correlated and can be represented as a linear sequence ranging from regular, through intermediate to irregular clusters as shown in Tab. (1.1). Regular clusters have symmetric shape with a central core to which galaxies tend to cluster, on the contrary to irregular clusters, which do not show symmetric shape, but usually come with strong sub-clustering without obvious central denser region, pointing out that regular clusters are dynamically more evolved systems than irregular clusters.

Property	Regular	Intermediate	Irregular
Zwicky type	compact	medium-compact	open
Bautz-Morgan type	I, I-II, II	II, II-III	II-III, III
Rood Sastry type	cD, B, L, C	L, C, F	F, I
Galactic Content	elliptical rich	spiral poor	spiral rich
E:SO:S	3:4:2	1:4:2	1:2:3
Morgan type	*ii*	*i-ii*	*i*
Oemler type	cD, spiral poor	spiral poor	spiral rich
Symmetry	spherical	intermediate	irregular
Central concentration	high	moderate	low
Subclustering	absent	moderate	significant
Richness	rich	rich-moderate	rich-poor

Table 1.1: Optical classifications schemes of galaxy clusters mapped into a linear sequence.

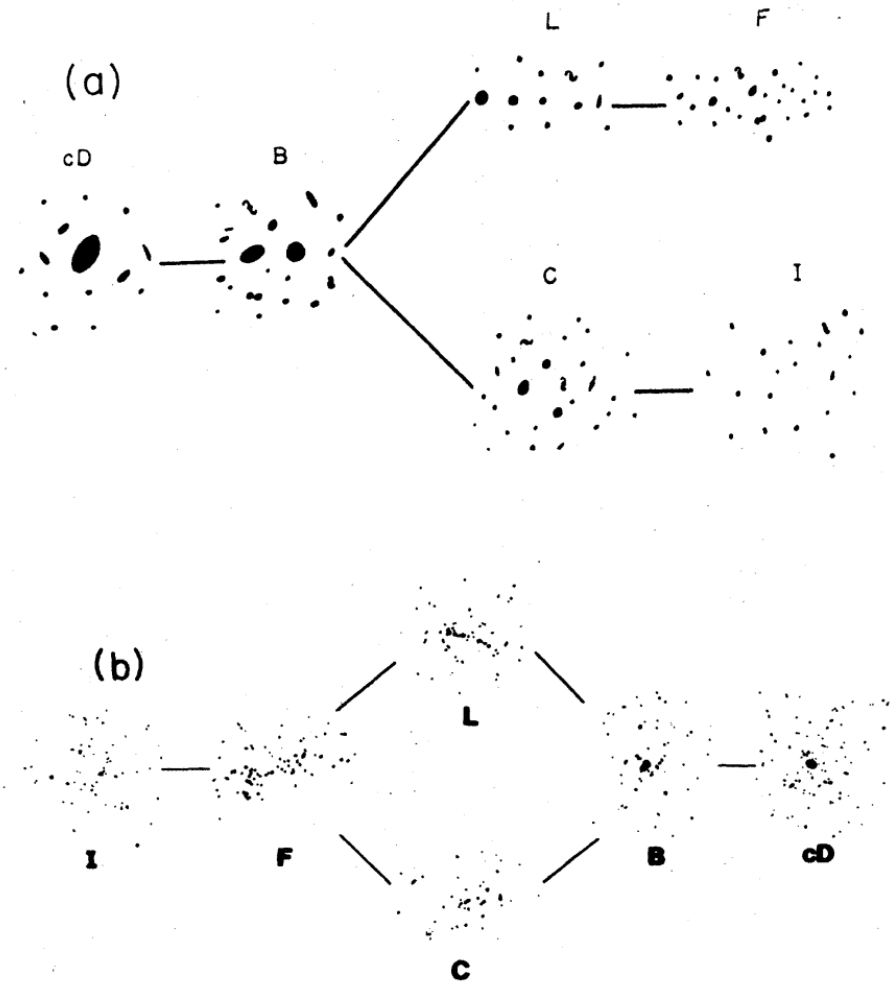

Figure 1.1: (a) Rood-Sastry "tuning fork" diagram for classification of rich clusters of galaxies, correlated to show clusters from regular (left) to irregular (right). (b) Struble-Rood "split-linear" revision of Rood-Sastry "tuning fork" describing the evolution of clusters from irregular (left) to regular (right).

1.3.2 Optical automatic machine detection of galaxy clusters

The first cluster catalogue created by software was the The Edinburgh-Durham Southern Galaxy Catalog (Lumsden et al., 1992) consisting of 737 clusters or groups of different richnesses. It was based on Abell-like criteria, using optical observations around the southern galactic pole taken from the COSMOS scans of the UK Schmidt plates.

The Automatic Plate Measuring machine galaxy catalog (Dalton et al., 1997) was done in a similar way, using Abell-like criteria and the data from the APM galaxy survey (Maddox et al., 1990), which covers the southern galactic pole, it resulted in a sample of $\sim 1\,000$ clusters.

Palomar Distant Cluster Survey (Postman et al., 1996) was the first which used V and I bands for cluster search. It probed over 5 square degrees, compiling a sample of 79 cluster candidates using matched-filter algorithm, which is a maximum-likelihood algorithm analysing the galaxy distribution, using prior knowledge of some model profiles to fit the data such as density distribution profile or luminosity function. Similar algorithm has been used to create a sample of 35 distant cluster candidates conducted upon ESO Imaging Survey I-band data, covering about 3 square degrees (Olsen et al., 1999). Another survey using I-band data was conducted by (Lidman and Peterson, 1996), probing 13 square degrees, selecting a sample of 104 clusters, but this time using counts-in-cell method, equally effective to matched-filter algorithm.

Sloan Digital Sky Survey (SDSS) (Abazajian et al., 2004), benefiting from the CCD detectors, created revolutionary sky survey of the local Universe. Miller et al. (2005) developed the C4 cluster finder algorithm, which they used to probe the SDSS second data release (DR2), detecting 748 clusters of galaxies in a ~ 2600 square degrees of sky with redshifts ranging from 0.02 to 0.17. The C4 algorithm looks for overdensities in 7 dimensional position and colour space, minimising projection effects in optical wavelengths.

Adaptive smoothing kernel algorithm (Silverman, 1986) uses a two-step process to produce a density map. Firstly, estimate of galaxy density is done for each point on the map, followed by applying a smoothing kernel on this map, which size changes as a function of the local density, with a larger kernel at lower density and smaller at higher density. Northern Sky Optical Cluster Survey (NoSOPC) (Gal et al. (2000), Gal et al. (2003)) used the adaptive kernel algorithm together with colour selection criterion against field galaxies to create galaxy overdensity maps, which they used as an input to the SExtractor (Bertin and Arnouts, 1996) object detection algorithm, detecting galaxy overdensities, which they identified as clusters of galaxies. NoSOPC catalogue contains 8155 candidate clusters selected from $\sim 5\,800$ square degrees of the northern galactic cap region.

Voronoi Tesselation method is a complete separation of a studied region into smaller volumes. Objects in the image are characterised by these volumes, with radius being the mean object separation. Ramella et al. (2001) and Kim et al. (2002) applied Voronoi Tesselation to galaxy cluster detection, using galaxy

positions and magnitudes. They identified clusters as significant fluctuations from the beckground, together with their main properties such as size, richness and contrast above the background. The method still requires users input in a form of setting significance threshold, but it is non-parametric and does not smooth the data, resulting in an identification of clusters irrespective to their shape, however being slightly affected by border effects and holes in the galaxy distribution on the sky.

Brightest clusters galaxies have predictable colours and magnitudes. The maxBCG algorithm (Annis et al. (2002), Koester et al. (2007)) uses this information to calculate a likelihood for each of galaxy being a BCG as a function of redshift based on its colours and the presence of a red sequence from the surrounding objects in the SDSS (York et al. (2000), Abazajian et al. (2004)) data, without creating density maps. It is primary developed for detection of massive galaxy clusters characterised by an overdensity of bright, uniformly red galaxies.

Gladders and Yee (2000) created cluster detection method using two-band optical/near-IR imaging data. The method is based on the red sequence of early-type galaxies commonly appearing in all rich clusters, the empirical information that cluster's elliptical galaxies are redder than spiral galaxies of same redshift, which is a strong indicator of a cluster's presence. This technique was used to construct the first (Gladders and Yee, 2005) and the second (Gilbank et al., 2011) red sequence cluster surveys.

1.4 X-ray observations

The X-ray emission of galaxy clusters is mainly created by the thermal free-free collisions of intra-cluster medium (ICM) electrons with other charged particles, like atomic nuclei (Felten et al., 1966), also called bremsstrahlung radiation with typical luminosity of $10^{43} - 10^{45}$ erg/s. Electron of highly ionized ICM plasma of temperatures $10^7 - 10^8$ K radiates X-ray photons when deflected by charged particles. Other sources of clusters X-ray emission are: free-bound collisions, radiation caused by recombination of ions, capture of electron by an ion during close encounter followed by ionization and bound-bound collisions, radiation caused by deexcitation of an electron when changing the quantum level in an ion. The deexcitation process causes line emission, while the first two contribute to continuum emission.

Observational discovery of line emission of highly ionized iron line in clusters X-ray spectra, firstly discovered in the spectrum of Perseus cluster (Mitchell et al., 1976) and shortly after in the spectra of Perseus and Coma clusters (Serlemitsos et al., 1977), finally established the mechanism of clusters X-ray emission to be thermal, meaning that the hot intra-cluster gas had to be a significant source of clusters X-rays, which must have been ejected or enriched by stars at some point, pointing out to the interaction of galaxies with ICM. This detected Fe line was actually a blend of lines from iron ions, mainly Fe^{24+} and Fe^{25+} and

some weaker lines from nickel ions. The main photon energies of those lines were between 6.5 and 7.0 keV.

Merging events strongly affect dynamical properties of clusters as has been shown by both, observations and simulations. Hydrodynamical simulations showed that it takes about 2 Gyr to 4 Gyr for a cluster to re-establish its equilibrium after a major merger event (Lacey and Cole (1993), Nelson et al. (2012)). Moreover, cluster mergers lead to relativistic particle accelerations, creating large radio halos, which are found to be correlated with X-ray emission (Buote (2001), Hallman and Jeltema (2011)).

1.4.1 X-ray properties of galaxy clusters

Early X-ray observations revealed that the cooling time of the ICM gas in the central regions of many clusters is much shorter than the Hubble time (e.g. Lea et al. (1973), Mathews and Bregman (1978)). This discovery led to development of the cooling flow model, in which the ICM of dense cores clusters hydrostatically cools. The cool gas is being compressed by the hot gas of the outer regions. Hot gas flows to the centre, replacing the compressed gas and generating a cooling flow. However, optical observations (McNamara and O'Connell (1989), Edge (2001)) did not detect expected rates of star formation, CO and molecular gas predicted by the cooling flow model. Moreover, spectra (e.g. Peterson et al. (2001) or Sanders et al. (2008)) from the XMM-Newton satellite (Jansen, 1999) find out that the cooling rate of the gas in central parts does not match with the cooling flow model, but the cooling flow model assumes no significant heating mechanism, so this results gave rise to search for heating models that could explain current observations.

The cooling flow model was ultimately replaced by classification of clusters as cool-core (CC) and non-cool-core (NCC) (Molendi and Pizzolato, 2001). Condensed regions of the cool gas in CC clusters demonstrate sharply peaked X-ray emission, because of their higher brightness than that of the surrounding material (Fabian (1994), Donahue and Voit (2004)). CC clusters seem to be relaxed, achieving relatively steady state of cooling and accretion of the gas, making them an attractive tool for cosmological studies, because being dynamically relaxed makes the fitting of a density profile that provides the gas mass to dark matter ratio more accurate (Allen et al. (2004), Allen et al. (2008), Vikhlinin et al. (2009)). Active galactic nuclei (AGN) feedback is the current best leading model to explain the dampening of poor cooling flows (e.g. Gitti et al. (2012), Li et al. (2015), Pinto et al. (2018)). Cooling gas creates clumps and filamentary structure, feeding star formation and super-massive black hole, setting on the AGN outbursts, which increase gas entropy, decreasing its cooling time, reducing the cool gas inflow to the AGN, so the AGN eventually briefly shuts off, making place for ICM to cool and redevelop multiphase gas again, triggering another star formation and AGN activity, making this process continue in a loop. Another forms of heating could also contribute, such as

thermal conduction and cosmic rays (see McNamara and Nulsen (2007) and McNamara and Nulsen (2012)).

X-ray observations of galaxy clusters also provide a powerful tool to explore clusters regularity. Based on X-ray observations, clusters can be classified into two main groups, regular (relaxed objects) and disturbed (dynamically active systems). Jones and Forman (1992) used X-ray observations to create cluster classification into classes *single*, *double*, *primary with small secondary*, *complex*, *elliptical* (based on the X-ray contours), *off-center* (displacement of center in X-ray and optical or an X-ray tail extending only to a single sector off the X-ray peak) and *galaxy* (for clusters with X-ray emission dominated by the central galaxy). Other measurements of the disturbance of clusters X-ray emission are power-law ratios (Buote and Tsai, 1995), according to clusters projected morphology or axial ratios (Mohr et al., 1993), based on the X-ray surface brightness distribution.

Dynamically active, disturbed clusters are either young newly forming clusters or cluster mergers. While regular clusters are important for cosmology, dynamically active clusters are very interesting for astrophysical studies of processes like interactions of galaxies, interaction of ICM with galaxies, AGN feedback, galaxy cluster mergers, ram pressure stripping, gas evaporation from galaxies, turbulence, sloshing, cold fronts formed from sloshing, particle re-acceleration, magneto-astrophysics or dark matter studies.

1.4.2 X-ray automatic machine detection of galaxy clusters

Great advantage of X-ray detection of galaxy clusters comes from the strong ICM X-ray emission, allowing their observations up to large cosmological distances. X-ray cluster observations are also less sensitive to projection effects, which optical detections suffer from. Detectability of clusters in X-rays depends mainly on their distance and size, larger and closer the cluster is, easier it is to detect it. With higher redshift, larger volumes of space are probed. Detection methods will be biased towards spotting larger clusters in higher redshifts, which has to be accounted for in modelling of selecton function.

The XAmin pipeline (Pacaud et al., 2006) is used to automatically detect galaxy clusters in X-ray images. It operates on a combined MOS1+MOS2+PN X-ray image of an XMM-Newton observations. A dedicated wavelet smoothing program called `mr_filter` (Starck et al. (1998), Starck and Pierre (1998)) firstly smooths the combined image, so the X-ray structures, characterized by the low number of photons, would be effectively recovered.

Second step is an analysis of the waveleth smoothed image by a source extraction software, SExtractor (Bertin and Arnouts, 1996). The software creates a list of candidate sources for a further analysis. Each source has an estimation of a position and a flux. SExtractor was developed to be used on optical images, which contain significantly more photons than X-ray images, which is why it is necessary to smooth the X-ray image, otherwise the SExtractor would not be able to work upon raw X-ray images. Smoothing can be done

in various ways, the wavelet smoothing XAmine is just one of possibilities. Valtchanov et al. (2001) showed how it gives base results for diffuse sources, like galaxy clusters, in X-ray images.

The final step is a characterisation of the source found by the SExtractor. Each of found sources is fitted by a point source model given by the XMM-Newton's PSF, which is computed for the source at its position. Extended β model (Cavaliere and Fusco-Femiano, 1976), better describing galaxy clusters, is computed together with the PSF. If the PSF model fits better, the source is declared as a point source (typically an AGN), if the β model fits better, the source is declared as an extended source (galaxy cluster candidate). Details of selection criteria for definition of an (almost) pure sample of galaxy clusters explains Pacaud et al. (2006).

Clerc et al. (2012b) used the Xamine pipeline to process 2774 XMM-Newton archival observations with high galactic latitude, extracting fortuitous catalogue of ~ 850 clusters, based on a pure X-ray criteria, in concordance of methodology developed for the XMM-LSS survey. They also performed cosmological analysis on the highest signal-to-noise objects (347 clusters), consisting of modelling the observed colour-magnitude diagram conducted on instrumental count-rates measured in the $(0.5-2.0)$ keV, $(1.0-2.0)$ keV and $(0.5-1.0)$ keV bands.

Chapter 2

Method

2.1 Neural networks

Neural networks (Mcculloch and Pitts, 1943), a class of machine learning algorithms, consist of layers, which are made of mathematical neurons. Convolutional layers are made of kernels, consisting of a stack of learnable filters. Layers input is scanned across by layer's filters, applying the convolution to produce the layer's output. The filters spacial dimension is usually very small, typically 3×3 pixels or 5×5 pixels. Filters learn to extract representing features from the training data, e.g. colour patches or differently oriented edges. Deep learning neural networks are networks with more than one hidden layer, the layers bounded by the input and the output layer. Layer's output is a non-linear transformation of its input. Layers are usually connected in a sequence, however, there are architectures like ResNet (He et al., 2015), GoogleNet (Szegedy et al., 2014) or recurrent neural networks Bengio et al. (1994), where this is not always true.

An example of a neural network with layers stacked in a sequence is a fully connected feed-forward neural network (Fig. 2.1). Each of the neurons of the fully connected layer connects to all the neurons of the adjacent layers, but with none of its own layer. Individual connections have their importance, the weights, which are being learned (Fig. 2.2). Fully connected architecture results in a very large number of parameters, making it very big and computationally expensive to train. The output of a layer j is a vector,

$$\mathbf{x}_j = \boldsymbol{f}_j(\mathbf{W}_j\mathbf{x}_{j-1} + \mathbf{b}_j), \tag{2.1}$$

where $\mathbf{x}_{j-1}$ is the input to layer j with matrix of weights $\mathbf{W}_j$, vector of biases $\mathbf{b}_j$ and layer's non-linear activation function $\boldsymbol{f}_j$.

Back-propagation is used to train the network. In case of a supervised learning, the predicted output $\mathbf{y}$ for a given input $\mathbf{x}$ is compared to the true label $\hat{\mathbf{y}}$ applying a loss function $\boldsymbol{L}(\mathbf{y}, \hat{\mathbf{y}})$. Gradient descent (LeCun et al., 1999) computes the gradient for each weight connection. Those gradients are used to determine the direction along which neural networks parameters, weights and biases, should be renewed to decrease the loss. The gradients give the direction of the fastest descent of the loss function.

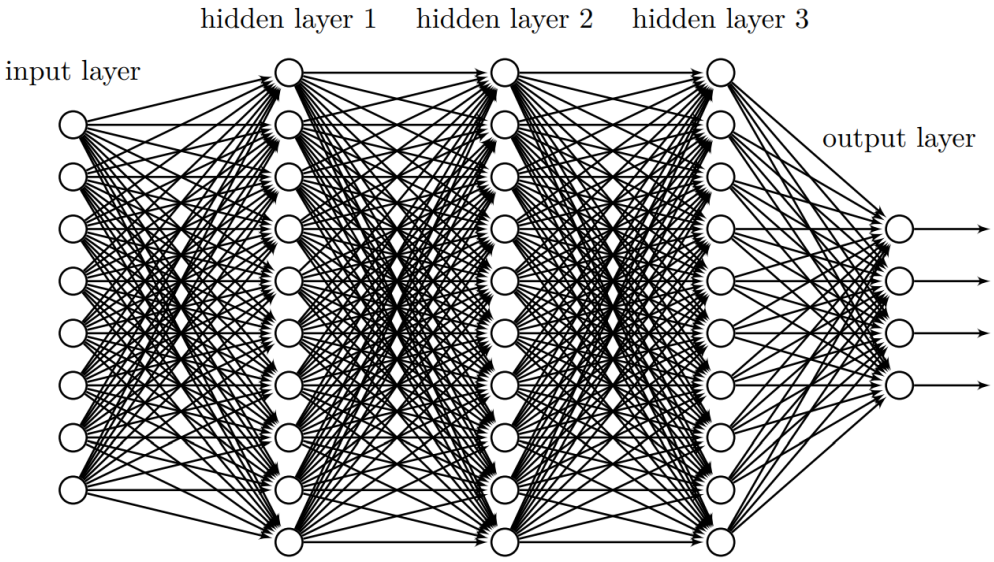

Figure 2.1: Example of an architecture of a feed-forward neural network consisting of one input layer, three hidden layers and one output layer. All layers are fully connected.

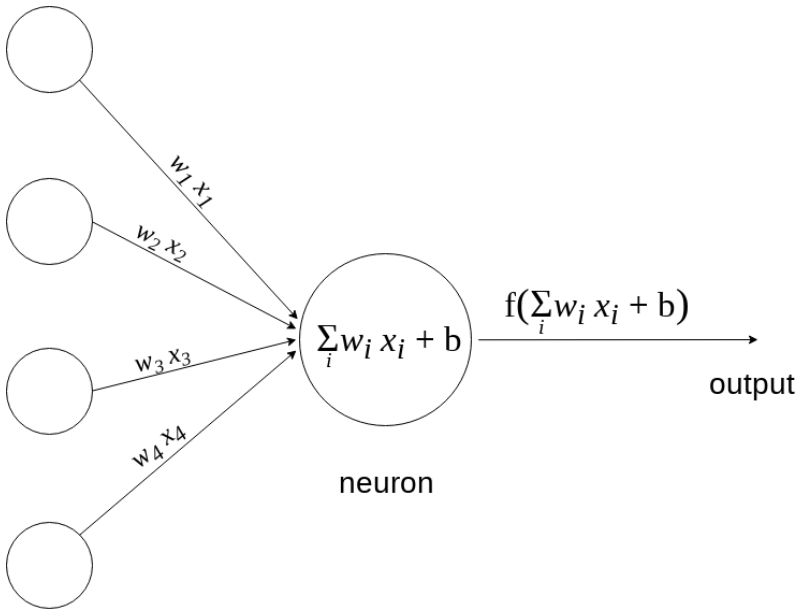

Figure 2.2: Example of a neuron connecting to 4 neurons and producing its output.

$$w^{(i+1)} = w^{(i)} - \alpha \frac{\partial L}{\partial w^{(i)}} \tag{2.2}$$

$$b^{(i+1)} = b^{(i)} - \alpha \frac{\partial L}{\partial b^{(i)}} \tag{2.3}$$

The update of weights and biases is performed simultaneously for all of their values. The learning rate, parameter α, is used to manually change the weight update. It is used to prevent rapid changes of weights, which would make them diverge.

The original gradient descent is an optimisation method updating all of the weights at once after iterating through all of the training samples once. Stochastic gradient descent, its alternative, updates the weights progressively after every mini-batch, a subset of the training data presented to the neural network. The size of the mini-batch is a hyperparameter, like the learning rate, it has to be manually chosen.

2.2 Convolutional neural networks

Convolutional neural networks (CNNs) consist of convolutional layers. Each convolutional layer is made of a stack of learnable filters. A convolutional filter is a 3D matrix, which pixel values are being learned during the training. The connectivity of neurons makes the difference between convolutional and dense layers. Convolutional layers connect only locally, on the contrary to the fully connected dense layers, which neurons are connected to all of the adjacent layers neurons. This results in a huge decrease in computations and parameters. Convolutional layer convolves its input with its filters, creating an output (Fig. 2.3) . The input of a convolutional layer is a stack of feature maps, which can be represented as a set of K matrices $\mathbf{X}_{j-1}^{k}$ with $k = 1 \ldots K$. The output of a convolutional layer is also a set of feature maps $\mathbf{X}_{j}^{s}$ with $s = 1 \ldots S$, which can be represented by following formula (Dieleman et al., 2015)

$$\mathbf{X}_{j}^{s} = f_j\left(\sum_{k=1}^{K} \mathbf{W}_{j}^{k,s} * \mathbf{X}_{j-1}^{k} + \mathbf{b}_{j}^{s}\right). \tag{2.4}$$

Filters of layer j are expressed by matrices $\mathbf{W}_{j}^{k,s}$, the bias of a feature map s is characterised by $\mathbf{b}_{j}^{s}$ and * represents the 2-dimensional convolution, which is defined as

$$\mathbf{X} * \mathbf{H} = \sum_{k} \sum_{s} \mathbf{X}_{k,s} \mathbf{H}_{n-k,m-s}, \tag{2.5}$$

where n and m are filter width and height, $\mathbf{X}$ is the input and $\mathbf{H}$ is the filter. The depth dimension of each of the filters of a convolutional layer has to be the same as the depth dimension of layers input, each depth slice of a filter thus operates over one of the depth slices of layers input. Each of the convolutional

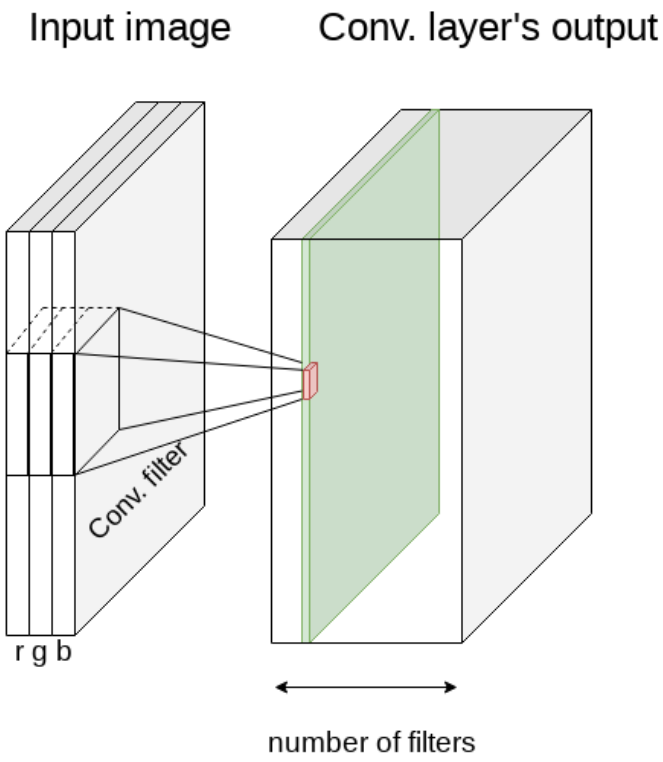

Figure 2.3: One of network's first convolutional layer's filters scanning its input.

filters outputs a separate 2 dimensional depth slice of its input, which is a 2D feature map. Output of a convolutional layer is created by stacking feature maps of all of its filters along the depth dimension, with spatial size defined by the layer's input image size, layer's filters size, number of pixels to move the filter across the input image during convolution (stride) and the number of added zero value pixel shells to the input in order to obtain resulted output size (padding). Individual filters learn different features.

Pooling layer usually follows a convolutional layer. The pooling layer reduces its input size, preserving only the most important features. This process reduces the input information while maintaining the most important features. The information reduction also reduces computation time, because smaller input requires fewer connections in the network. CNNs are usually ended by a stack of dense layers. While convolutional layers mainly lear to extract most important features from the training data set, dense layer stores information about which of those features are active and ultimately predicts the classification class.

Dropout technique is usually used for the fully connected dense layers at the end of the CNN. Neurons of a layer with the dropout are given probability to be inactive during training, which is recalculated after each training iteration. The network can not rely on information stored in connections of close neurons,

forcing it to better generalisation, so it decreases the over-fitting, tendency to learn features specific to the training data set. Dropout is used only during the training of the CNN, which is often monitored by a number of mini-batches yielded to the network or by a number of epochs, iterations through the whole training data set.

2.3 Transfer learning

Transfer learning (PRATT and JENNINGS (1996), Do and Ng (2005)) is based on using a deep neural network, trained on a large data set (e.g. VGG19 (Simonyan and Zisserman, 2014), InceptionV3 (Szegedy et al., 2015) or MobileNet (Howard et al., 2017)) and partially retraining it on a new data set. If the network is used for classification of different classes than it was trained on, its top dense layers are not used at all, because they encapsulate the information about the individual classes the network is trained to classify. Features of the convolutional layers are very simple, e.g. small (3×3 pixels) edges, curves or colour patches, which can be applied to various sorts of image classes, so the rest of the network is used as a well-trained feature extractor and it is partially fine-trained, with a newly initialised dense layer on top of it, on the data set of interest. During the fine-training, the transferred convolutional layers are usually frozen, fixing weights and biases, which are no longer being free-parameters to be trained.

2.4 Performance measurements

Accuracy is a ratio of all correct predictions to all predictions, defined as,

$$A = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}}, \tag{2.6}$$

where N_{TP} declares true positive rate, in our case of the binary classification it would be the number of clusters correctly classified as clusters, N_{TN} is a true negative rate (number of non-clusters correctly classified as non-clusters), N_{FP} refers to the false positive rate (number of non-cluster incorrectly classified as clusters) and N_{FN} denotes the false negative rate (number of clusters incorrectly classified as non-clusters).

Precision, ratio of correctly predicted positive observations over all predicted positive observations, is defined as,

$$P = \frac{N_{TP}}{N_{TP} + N_{FP}}. \tag{2.7}$$

Recall is the ratio of correct predictions for a certain class overall predictions for the same class, e.g. number of correctly predicted non-clusters to a number of all non-clusters in the test sample. It is defined as,

$$R = \frac{N_{TP}}{N_{TP} + N_{FN}}. \tag{2.8}$$

F_1 score is mainly useful in case of unbalanced data sets or for classification with differently weighted data sets, but it might be less intuitive than the accuracy. It is measured as a weighted average of precision and a recall accounting for both, N_{FP} and N_{FN} and is defined as,

$$F_1 = 2\frac{RP}{R+P}. \tag{2.9}$$

ROC, rectified operational curve, is a common way how to visualise classifier's performance. Threshold T defines whether the object belongs to first or second class, in a binary classification. ROC is obtained by plotting N_{TP} rate against the N_{FP} rate with T as a varying parameter.

The capability of a classifier to distinguish between the two classification classes describes the area under the ROC, AUC. Plotting ROC and AUC for the case of a multi-class classification problem would be done for each of the classification classes separately as one versus all, reducing the problem to the binary case.

Chapter 3

Data and The Hunt for Galaxy Clusters

3.1 XAmine pipeline

The `XAmin` pipeline (Pacaud et al., 2006) was used to create images of galaxy cluster candidates. It operates in the following way.

Firstly a combined MOS1+MOS2+PN image of XMM observations is smoothed with a dedicated wavelet smoothing program called `mr_filter` (Starck et al., 1998). Starck and Pierre (1998) shown it to be effective for recovering structures in X-ray images, which are characterised by low numbers of photons (Starck et al., 1998).

Secondly, source extraction software `SExtractor` (Bertin and Arnouts, 1996) is used to analyse the wavelet smoothed image. The software returns a lost of candidate sources to be further analysed, together with the estimate of their position and flux.

`SExtractor` was developed for optical images, which contain many more photons than X-ray images. To make the `SExtractor` work properly, smoothing of the X-ray image is a necessity as the `SExtractor` would not be able to work with raw X-ray images otherwise. While the smoothing can be performed in several ways, the wavelet smoothing used by `XAmin` was shown by Valtchanov et al. (2001) to give the best results for X-ray images of diffuse sources like galaxy clusters.

The final step characterises candidate clusters found by the `SExtractor` by fitting to each of those both a point source model is given by the XMM-Newton PSF, which has to be computed at the source position and an extended β model (Cavaliere and Fusco-Femiano, 1976), which is used to describe galaxy clusters. The source is declared as a point source (AGN) or an extended source (cluster) based on which of these two models best fits the candidate source. The details of the procedure, including relevant formulas and the selection criteria defining an (almost) pure sample of galaxy clusters, are described in Pacaud et al. (2006).

Considering that the galaxy cluster candidates in the project are candidate clusters picked by the `XAmin` pipeline, the point-sources in this data set are those which the pipeline classified as extended X-ray emission. This necessarily results in a bias of the point source sample towards those, which the `XAmin` misclassified as extended X-ray emission.

3.2 Data preprocessing

Neural networks were trained using only images without contours and marks. Each cluster candidate comes with a pair of X-ray and optical images, those were combined to create one Portable Network Graphics[1] image. To create one RGB PNG image, X-ray and optical images were grayscaled and concatenated together as individual channels.

Images need to be converted into one RGB PNG image for transfer learning approach because the transferred network was originally trained on 3 colour channel images, so it requires the same input channel size. Training on a concatenated PNG X-ray and PNG optical images was also conducted, but only with our custom network. Regardless of the image channel size, before fed to the network, the image values were always normalised to the interval $0-1$.

Grayscaling was done using OpenCV python library (Bradski, 2000) function `cv2.cvtColor(image, cv2.COLOR_RGB2GRAY)`, where image is an RGB PNG image selected for grayscaling (Fig. 3.1).

The neural network used for transfer learning can exhibit various sensitivity in different colour channels if the initial data set used for its training contained more information in one of the colour channels. This effect was explored by training transfer learning models on 3 different channel configurations, shown in Fig. 3.2. Neural networks were also trained using only RGB PNG X-ray images (input channel shape 3), only RGB PNG optical images (input channel shape 3) and of concatenated RGB PNG X-ray and optical images (input channel shape 6).

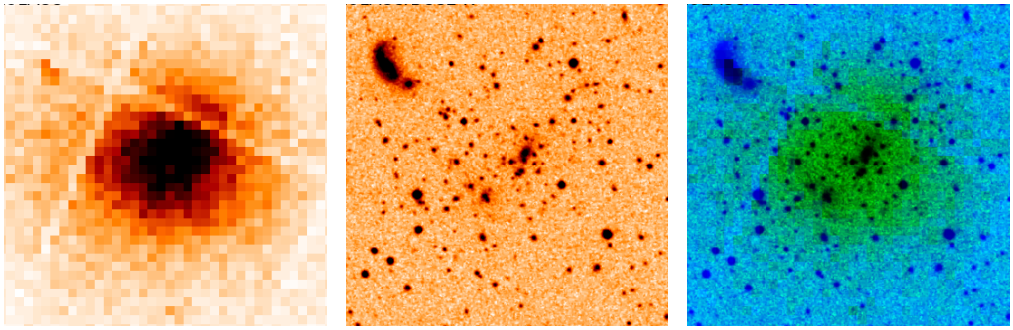

Figure 3.1: Left is the 510×510 X-ray RGB PNG image, the middle is its 510×510 optical PNG counterpart and right is concatenated RGB PNG image made of a grayscaled optical image as a blue channel, grayscaled X-ray image as green channel and the red channel was filled with zeros. The image was then cropped into 356×356 sizes.

[1]See e.g. `https://en.wikipedia.org/wiki/Portable_Network_Graphics` (PNG)

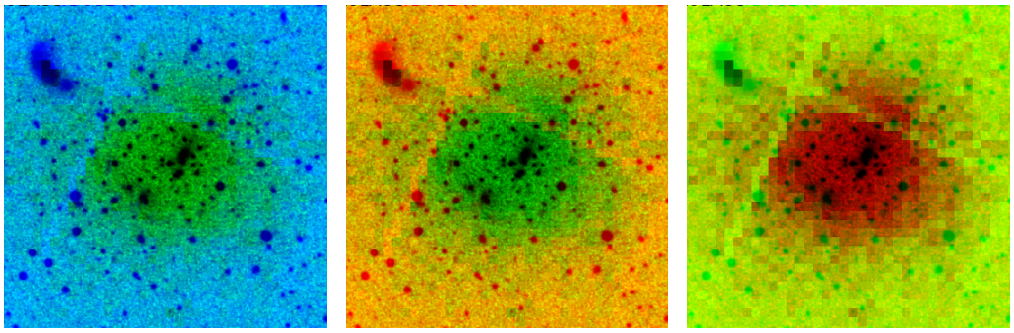

Figure 3.2: Different channel configurations of the training data set. Left: X-ray used as the green, optical as the blue, while the red channel is left empty. Middle: X-ray green, optical red, empty blue. Right: X-ray red, optical green and empty is red.

3.3 Data augmentation

Risk of over-fitting, poor generalisation, increases with smaller data sets. This can be prevented using data augmentations, making sure, that the network will never see exactly the same image twice. The data augmentation was essential to avoid the over-fitting and to increase the performance of classification. Each time an image is generated to the network during training, it is randomly scaled, rotated and translated. Augmentation parameters are described in Tab. 3.1.

Augmentation	value
Scaling	uniform distribution [1.3, 1/1.3]
Rotation	uniform distribution [0, 360] deg
Translation	x, y random pixel shift [-4, 4] pix

Table 3.1: Values of augmentation parameters.

3.4 The Hunt for Galaxy Clusters

The Hunt for Galaxy Clusters is an online citizen science project built via The Zooniverse[2] project builder platform. It contains a data set of 1 600 galaxy cluster candidates, detected as extended X-ray sources by the `XAmin` wavelet-based pipeline (Pacaud et al., 2006). Each of project's cluster candidates was classified by 30 volunteers, who were asked questions to classify candidate clusters into various categories according to the decision tree (Fig. 3.3), see Sec. 3.5.

Each object is represented by 4 images, each covering the exact same area of the sky, 7×7 arcmins. Two X-ray images, one with contours visualising the areas of constant X-ray brightness and a central cyan cross, which marks the object

[2]http://zooniverse.org

selected as a cluster candidate. The second X-ray image is the same but without central cross and contours. The other 2 images are optical counterparts, one with superimposed X-ray contours and the central cyan cross marking object selected for classification, the other one is without contours the central cross. Fig. 3.4. shows all 4 images as shown in The Hunt for Galaxy Clusters. The data in the project come from the XMM CLuster Archive Super Survey (X-CLASS) (Clerc et al., 2012a), an X-ray galaxy cluster search in the archival data of the European Space Agency's X-ray observatory, XMM-Newton (Jansen, 1999), optical counterparts come from the Digitized Sky Survey POSS-II (DSS2).

3.5 Weighting volunteers classifications

The Hunt for Galaxy Clusters decision tree (Fig. 3.3) encompasses 6 questions, each having 2 or 3 possible answers. Nevertheless, due to the structure of the decision tree, it may be possible that only a subset of the questions was answered. The agreement of volunteers with experts will vary between individuals. Moreover, it does not need to be the same for each of the questions. Someone may be good at some questions while struggling with others. It may also happen that some of the users were purposely creating malicious classifications. Users classifications were weighted question-wise to mitigate those effects. Weighting was conducted based on the agreement of individuals with the majority, defining the accuracy of an individual for each of classification questions as,

$$G_i = \frac{C_i}{Q_i}, \tag{3.1}$$

where G_i is the accuracy of an individual for question i, C_i is the number of individual's answers in agreement with the majority for question i and Q_i is the number of total individual's answers for question i. The accuracy describes the abbility of an individual to classify as the majority of volunteers would. When an accuracy for each of the 6 questions is computed for each of the volunteers, weighted classifications are computed. The final Zooniverse classification of an object is the classification ending answer with the highest weighted score.

3.6 Classification schemes

Neural networks were trained for two different classification schemes, binary and multiclass. Training of the networks was done using two different data sets, images classified by the Zooniverse volunteers in The Hunt for Galaxy Clusters and images classified by the experts in the field. Tab. 3.2 describes binary and multi-class classification schemes created from the Zooniverse classifications. Tab. 3.3 explains construction of classification schemes using classifications of experts. Categories without classification in the scheme were not used. Nearby or low-z galaxy clusters are those, which galaxies are still visibbe in the DSS2 optical images, so they have a redshift z lower than ~ 0.35.

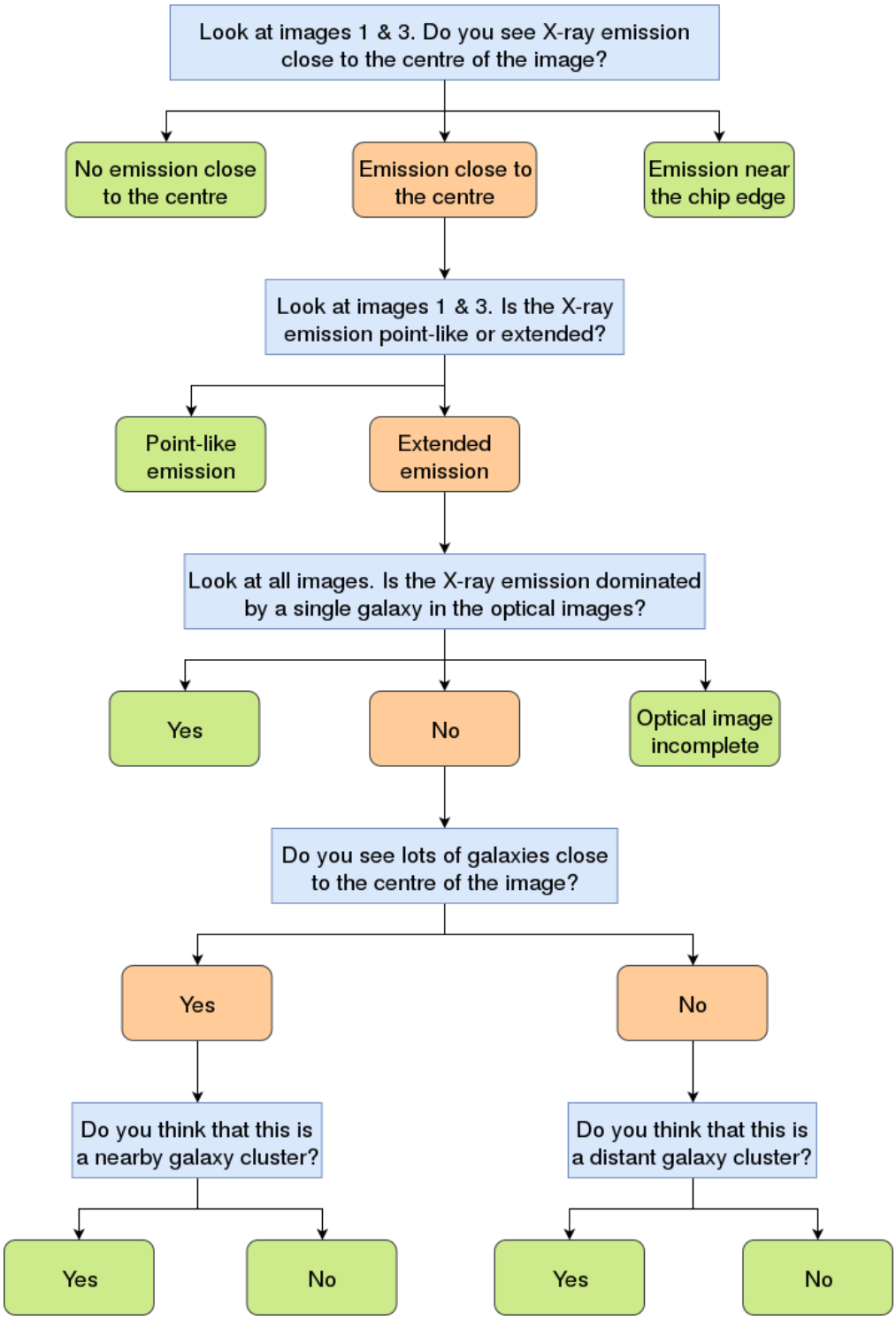

Figure 3.3: Flowchart of the decision tree of The Hunt for Galaxy Clusters. Blue cells represent questions, red are next question leading answers and green are classification ending answers.

Task	Answer	Binary	Multi-class
1	Emission close to the centre.	–	–
	No mission close to the centre.	non-cluster	other
	Emission near the chip edge.	non-cluster	other
2	Point-like.	non-cluster	point source
	Extended.	–	–
3	Yes.	non-cluster	nearby galaxy
	No.	–	–
	Optical image incomplete.	non-cluster	other
4	Yes.	–	–
	No.	–	–
5	Yes.	cluster	low-z cluster
	No.	non-cluster	other
6	Yes.	cluster	high-z cluster
	No.	non-cluster	other

Table 3.2: Description of binary and multi-class classification schemes using Zooniverse classifications.

Classification	Binary	Multi-class
low-z cluster	cluster	low-z cluster
high-z cluster	cluster	high-z cluster
point	non-cluster	point source
star or AGN	non-cluster	point source
double	non-cluster	point source
fossile group	–	–
nearby galaxy	non-cluster	nearby galaxy
artefact	non-cluster	other
bkg too high	no cluster	other
no optical image	non-cluster	other
edge	non-cluster	other
dubious	–	–

Table 3.3: Description of binary and multi-class classification schemes constructed from classifications of experts.

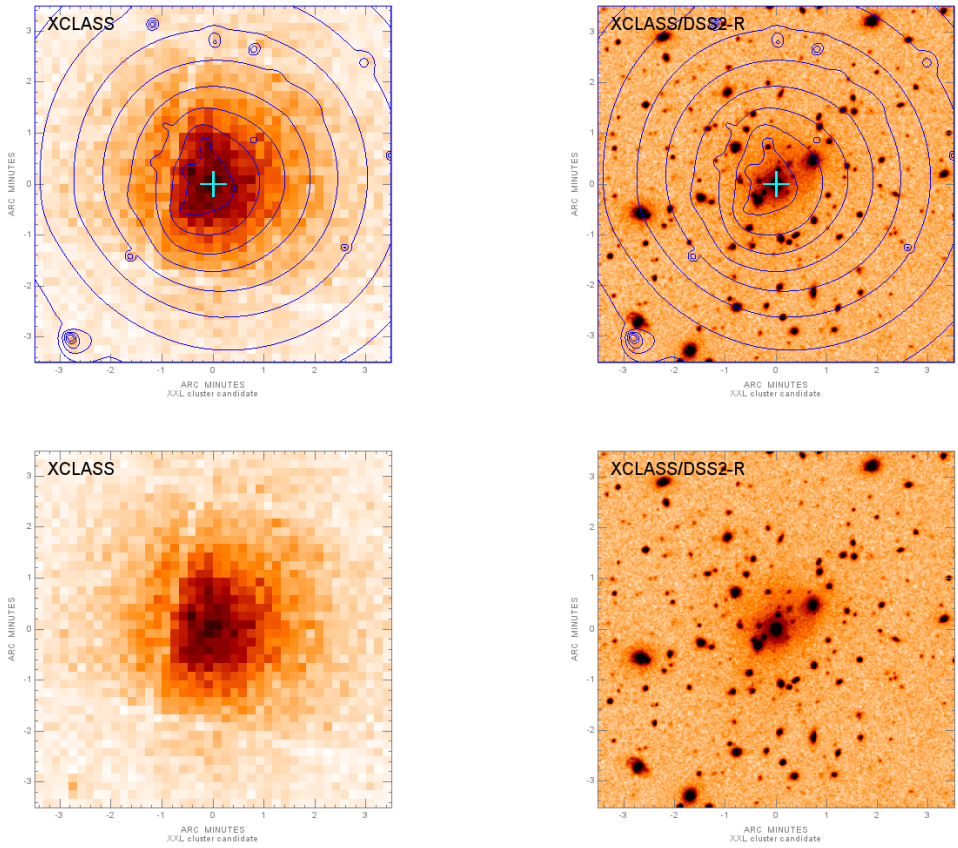

Figure 3.4: Top left is an X-ray raw image with contours showing the areas of constant X-ray brightness and the central cyan cross marking the object selected for classification. Top right is the optical counterpart image with superimposed X-ray contours together with the central cross. Bottom are the same images, but without contours and the central cross.

Chapter 4

Results and Discussion

4.1 The Hunt for Galaxy Clusters results

On the 24th October 2018, the project was launched as an official Zooniverse[1] project. It contained 1 600 cluster candidates, each with a threshold classification by a 30 different volunteers. Classifications were finished on 29th April 2019. 1 227 volunteers participated in the project. Classifications of users who were not logged in The Zooniverse page were not considered. 7 volunteers classified all of the 1 600 cluster candidates. Some of those continued classifying, making multiple classifications for the same objects, those and classifications of users who were not logged in were not used for further analyses.

Keeping communication with volunteers in the forum was absolutely inevitable to maintain the project alive. Discussions started for more than 1 280 project's cluster candidates. All of the volunteers questions in the forum were addressed within ~ 7 days, emerging into a challenging public outreach. Promoting the project in social media, Twitter and Facebook also helped to attract attention and new volunteers. We have promoted the project heavily at the beginning, with a pause during November 2018, starting again on Christmas 2018 until the end of January 2019. Classification rate in the project had a tendency to continuously drop, which might be connected with our activity on social media. We have also experienced a significant decrease in classification activity when our project dropped from The Zooniverse project's front page at the beginning of April 2019.

404 cluster candidates from the sample of 1 600 cluster candidates in The Hunt for Galaxy Clusters were previously classified by experts. This sample has been used to conduct agreement study between citizen volunteers and experts in the field (Tab. 4.1).

Fig. 4.1 shows the fraction of the Zooniverse answers in agreement with experts to all Zooniverse answers for classification ending answers, except for *not a nearby cluster* and *not a distant cluster*, which do not have direct counterpart in classification of experts. Zooniverse volunteers experienced a few biases. In general, they were biased to classify objects as $non - clusters$. The highest disagreement with experts lies in distinguishing point from extended X-ray

[1] http://zooniverse.org

Classifications	unweighted	weighted
N_{TP} (cluster – agreement)	69	102
N_{FN} (cluster – non-agreement)	185	152
N_{TN} (non-cluster – agreement)	150	149
N_{FP} (non-cluster – non-agreement)	0	1
Accuracy	54.2 %	61.1 %
Precision	100.0 %	99.0 %
Recall	27.2 %	40.1 %
F_1 score	42.7 %	57.1 %

Table 4.1: Weighted Zooniverse volunteers results in comparison to classifications of experts for a binary classification into clusters and non-clusters.

emission. Many more objects were classified by the Zooniverse volunteers as *no emission* than there were. This bias probably comes from the inability of recognising faint X-ray emission. The *edge* category also received many more classifications from volunteers than from experts. From the communication with volunteers in the project's forum, we find out, that they often misunderstood edge of the XMM-Newton's field of view with small chip gaps between individual XMM-Newton's detectors. Those small chip gaps were often present in the area of the X-ray emission, which is most probably a reason behind this bias. The fourth most difficult category to classify for the Zooniverse volunteers was the *nearby galaxy*, related to the third question in the project's decision tree. Volunteer's feedback in the forum enlightened reasons behind this bias, they have often classified nearby galaxy clusters with prominent BCG galaxy as a nearby galaxy, which is how a lot of nearby galaxy clusters were missed. On the other way, the Zooniverse volunteers created an extremely pure sample of clusters with just 1 non-cluster classified as a cluster after the weighting.

4.2 Training and performance

Neural networks were always trained on the same data set classified by experts, except for the neural networks trained on the Zooniverse classifications, data set from The Hunt for Galaxy Clusters. Training image data format was the concatenated image of the grayscaled X-ray, grayscaled optical and an empty channel used as the green, blue and red channels respectively, if not stated otherwise. Zooniverse data set does not contain the same objects as the data set classified by experts. Despite the training data set, all networks were tested on the same data set of 85 spectroscopically confirmed clusters and 85 `XAmin` selected cluster candidates classified as non-clusters by experts (Tab. 4.2).

In the beginning, the trainig dataset is shuffled, a balanced batch is selected and used for training (one iteration), the trainig data set is than re-shuffled, next, a balanced batch is selected and used to train the network. Batches presented to

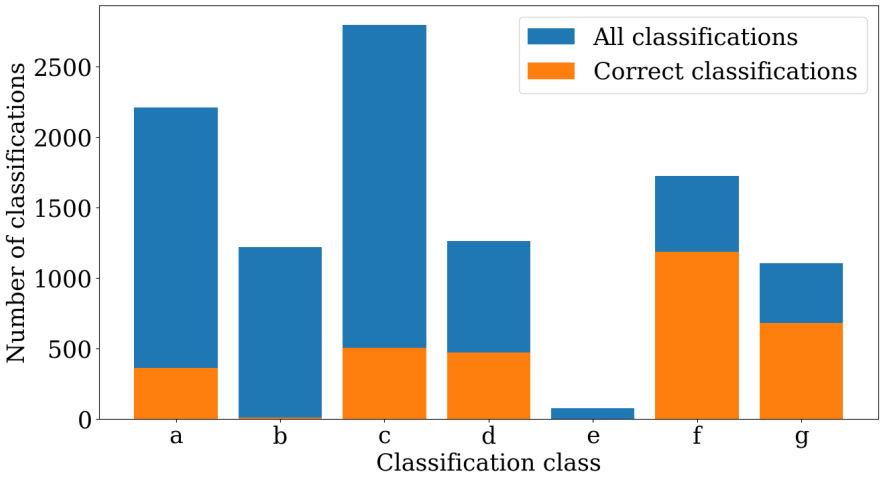

Figure 4.1: Histogram comparing classificationsi in agreement with experts to all classifications of the Zooniverse volunteers. Coding of answers is following: a) no emission, b) edge, c) point, d) nearby galaxy, e) no optical image, f) nearby galaxy cluster, g) distant galaxy cluster.

Class	Number of objects
low-z cluster	62
high-z cluster	23
point (point source)	20
double (point source)	9
star or AGN (point source)	6
nearby galaxy	28
artefact	21
edge	1

Table 4.2: Numbers of objects in test dataset.

the network were always balanced, having the same number of classes, 5 clusters and 5 non-clusters in case of the binary cassification or 2 classes of each for the multi-class classification. This was done to make sure that the network will not be biased, because our training and validation data sets contained unbalanced numbers of classes. Tab. 4.3 and Tab. 4.4 show numbers of objects of different classes in training and validation data sets, classified by experts and Zooniverse volunteers, respectively.

Class	Number of objects	
	Training data set	Validation data set
low-z cluster	465	100
high-z cluster	384	100
point (point source)	74	25
star or AGN (point source)	6	2
double (point source)	15	10
nearby galaxy	180	35
artefact	111	30
no optical image	9	0
edge	2	2
Clusters	845	200
Non-clusters	397	104
Total	1246	304

Table 4.3: Numbers of objects in the training and validation data sets classified by the experts.

Tab. 4.5 describes hyperparameters of our custom network and our transfer learning approach with MobileNet pre-trained neural network. Our custom network, trained from scratch, exhibit most sensitivity to the choice of the initial learning rate, number of convolutional layers, and of their filters and filter sizes, number of dense layers and their neurons.

Fig. 4.2 and Tab. 4.7 display the performance of our custom network (CN), when trained on a different data formats, together with the transfer learning (TL) approach. Grayscaled in the legend refers to the combination of X-ray and optical used as green and blue channels respectively, with red channel left to zeroes. The best performance is achieved with our custom network. Training on concatenated PNG X-ray and optical images slightly decreased performance. Training on solo PNG X-ray or solo PNG optical images resulted in significantly lower performance, which is expected because the network was not given information from the other wavelength, which is crucial for galaxy cluster classification. Nearby galaxies appear nearly indistinguishable from clusters in X-ray, while optical observations suffer mainly from projection effects, e.g. looking at a galaxy filament aligned in a line of sight would create apparent overdensity of galaxies, undifferentiable from a galaxy cluster without information of a presence of the hot-intracluster

Class	Number of objects	
	Training data set	Validation data set
nearby galaxy cluster	235	100
distant galaxy cluster	85	30
point-like X-ray emission	392	30
nearby galaxy	187	30
no optical image	2	0
emission near the chip edge	49	10
no emission close to the centre	249	30
not a nearby galaxy cluster	0	0
not a distant galaxy cluster	1	0
Clusters	320	130
Non-clusters	880	100
Total	1200	230

Table 4.4: Numbers of objects in the training and validation data sets classified by the Zooniverse volunteers.

Hyperparameters	Custom net	MobileNet
Batch size	10	10
Iterations	158 000	7 900
Optimizer	SGD	Adadelta
Nest. Momentum	0.9	-
Rho	-	0.95
Initial lr.	0.0001	1.0
lr. decay	10^{-6}	0.95
Minimal lr.	10^{-4}	0.01
lr. red. patience	14	4
lr. red. factor	0.75	0.85
Conv. layers	Fig. 4.15	MobileNet
Activations	ReLU	MobileNet
Dense layer	1×256	1×256
Dense dropout	0.65	0.65
Output layer	2	2
Activation	softmax	softmax
Loss function	cat. crossentropy	cat. crossentropy
Input image size	356×356	224×224

Table 4.5: Hyperparameters of used networks.

gas visible in the X-ray wavelengths. Projection overdensities of field galaxies also contaminate optical images. Combination of the two wavelengths is very beneficial, resulting in an automatic galaxy cluster classifier, which achieves 94.1 % accuracy in binary classification of galaxy cluster candidates to clusters and no-clusters, tested on a sample of 85 spectroscopically confirmed clusters and 85 cluster candidates selected by the `XAmin` pipeline, which were classified as no-clusters by experts.

Three different pre-trained neural networks, VGG19 (Simonyan and Zisserman, 2014), InceptionV3 (Szegedy et al., 2015) and MobileNet (Howard et al., 2017), were tested for transfer learning. Each of them was originally trained on the ImageNet (Deng et al., 2009) data set composed of approximately 1.2 million images of 1 000 classes. MobileNet always outperformed the other two, so we further discuss only it. MobileNet was found to achieve the best performance in transfer learning approach also by Lieu et al. (2018) in case of solar system objects detection.

The most sensitive hyperparameter for transfer learning was the number of transferred layers to froze (Fig. 4.3, Tab. 4.9). MobileNet did not achieve good performance with all of its convolutional layers frozen. The performance was not increased when deeper convolutional layers were unfrozen. However, unfreezing the first 50 layers immediately increased the performance. Unfreezing more of the first layers continuously increased the performance.

To further examine this we trained the MobileNet network, without transferring weights, using the MobileNet architecture randomly initialised from scratch. It achieved similar results as with its weights transferred and unfrozen (Fig. 4.4 and Tab. 4.11). Training only the newly initialised top dense layer on top of the frozen MobileNet's transferred convolutional layers achieved AUC = 72.9 %. Initialising MobileNet's layers from scratch and freezing them produces, as expected, diagonal ROC curve of AUC = 50.0 %. This experiment shows, that the MobileNet weights are better than random initialisation in our study case, but freezing them does not lead to satisfying performance, so they need to be unfrozen and fine-tuned.

Fig. 4.5 and Tab. 4.12 depict ROC curves and performance of our custom network trained for multi-class classification. The network performed best in classifying nearby galaxies. This was expected, because single galaxy, large on the optical image, is hard to be misclassified by something else from classification classes.

Transfer learning MobileNet network performed a bit better than our custom network in case of multi-class classification (Fig. 4.6, Tab. 4.13).

Fig. 4.7 and Tab. 4.15 show the performance of our custom network and the MobileNet transfer learning with and without training augmentation. Those networks were trained on the grayscaled X-ray green, optical blue, red empty channel images on classification done by experts. The augmentation during the training was crucial to obtain good performance, to reduce over-fitting.

All of the randomly initialised layers used the glorot-uniform initialisation (Glorot and Bengio, 2010). Impact of initialisation's randomness on network's

performance was explored for transfer learning MobileNet architecture, where only the top dense layer was randomly initialised (Fig. 4.8, Tab. 4.17). It is visible that the performance changes when the network is re-trained multiple times. The network always had the exact same hyperparameters, training data were grayscaled images with the green channel: X-ray, blue channel: optical and red channel: zeroes, labelled by experts.

Fig. 4.9, Tab 4.18 and Tab. 4.19 show that the transfer learning with the MobileNet network was not influenced by the channel configuration of the grayscaled images used to create one RGB PNG image for the training (Fig. 3.1).

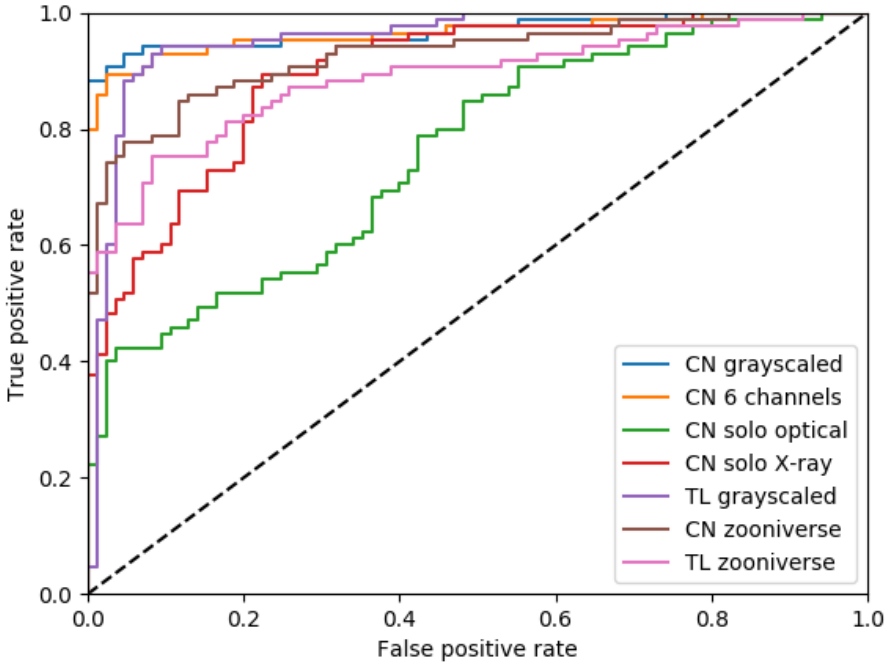

Figure 4.2: ROC curves of best performing networks when trained on different data formats.

Network	Clusters		Non-clusters	
	N_{TP}	N_{FN}	N_{TN}	N_{FP}
CN grayscaled	77	8	83	2
CN 6 channels	79	6	76	9
CN solo optical	69	16	50	35
CN solo X-ray	78	7	60	25
TL grayscaled	75	10	81	4
CN zooniverse	58	27	83	2
TL zooniverse	50	35	84	1

Table 4.6: Classification results of the best performing networks when trained on different data formats and labels.

Network	Accuracy	Precision	Recall	F_1 score	AUC
CN grayscaled	94.1 %	97.5 %	90.6 %	93.9 %	96.9 %
CN 6 channels	91.8 %	91.8 %	91.8 %	91.8 %	94.5 %
CN solo optical	68.2,%	65.0 %	78.8 %	71.3 %	87.3 %
CN solo X-ray	81.2 %	75.7 %	91.8 %	83.0 %	86.2 %
TL grayscaled	91.8 %	94.9 %	88.2 %	91.5 %	94.0 %
CN zooniverse	83.0 %	96.7 %	68.2 %	80.0 %	92.7 %
TL zooniverse	78.8 %	98.0 %	58.8 %	73.5 %	88.7 %

Table 4.7: Performance of the best performing networks when trained on a different data formats and labels.

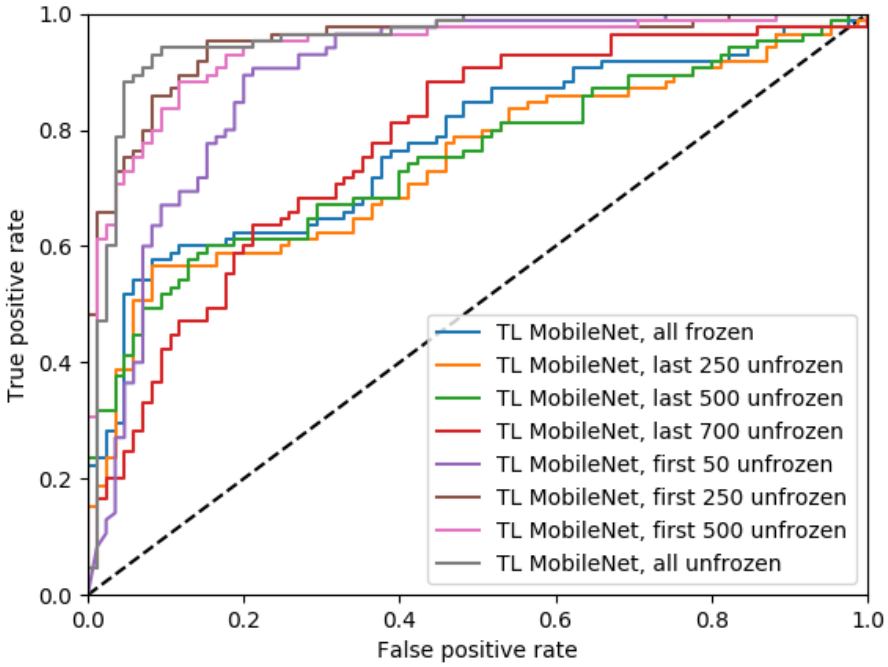

Figure 4.3: ROC curves of transfer learning with MobileNet for differently frozen transferred convolutional layers.

Network	Clusters		Non-clusters	
	N_{TP}	N_{FN}	N_{TN}	N_{FP}
TL, all frozen	80	5	12	72
TL, last 250 unfrozen	82	3	5	80
TL, last 500 unfrozen	74	11	27	58
TL, last 700 unfrozen	73	12	48	37
TL, first 50 unfrozen	82	3	53	32
TL, first 250 unfrozen	81	4	72	13
TL, first 500 unfrozen	75	10	74	11
TL, all unfrozen	75	10	81	4

Table 4.8: Classification results of transfer learning with the MobileNet network having the same hyperparameters, but differently frozen transferred convolutional layers.

Network	Accuracy	Precision	Recall	F_1 score	AUC
TL, all frozen	54.1 %	52.3 %	94.1 %	67.2 %	77.4 %
TL, last 250 unfrozen	51.1 %	50.6 %	96.5 %	66.4 %	74.1 %
TL, last 500 unfrozen	59.4 %	56.1 %	87.1 %	68.2 %	74.9 %
TL, last 700 unfrozen	71.2 %	66.4 %	85.9 %	74.9 %	77.6 %
TL, first 50 unfrozen	79.4 %	71.9 %	96.5 %	82.4 %	89.1 %
TL, first 250 unfrozen	90.0 %	86.2 %	95.3 %	90.5 %	94.8 %
TL, first 500 unfrozen	87.6 %	87.2 %	88.2 %	87.7 %	93.9 %
TL, all unfrozen	91.8 %	94.9 %	88.2 %	91.5 %	95.5 %

Table 4.9: Performance of the transferred MobileNet network with the same hyperparameters, but differently frozen transferred convolutional layers.

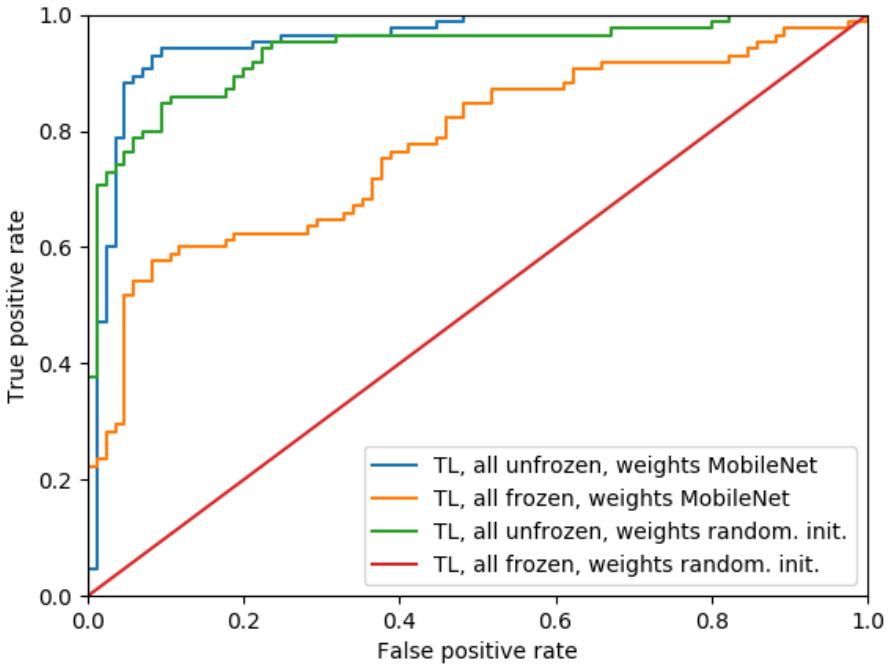

Figure 4.4: ROC curves of the MobileNet network with transferred weights of its convolutional layers, compared to training with its weights randomly initialised.

Network	Clusters		Non-clusters	
	N_{TP}	N_{FN}	N_{TN}	N_{FP}
TL, all unfrozen	75	10	81	4
TL, all frozen	80	5	12	72
TL, all unfrozen, weights random. init.	73	12	76	9
TL, all frozen, weights random. init	85	0	0	85

Table 4.10: Classification results of the MobileNet network with transferred weights of its convolutional layers, compared to training with its weights randomly initialised.

TL MobileNet	Accuracy	Precision	Recall	F_1 score	AUC
All unfrozen, MobileNet	91.8 %	94.9 %	88.2 %	91.4 %	95.5 %
All frozen, MobileNet	54.1 %	52.3 %	94.1 %	67.2 %	77.4 %
All unfrozen, random. init.	87.6 %	89.0 %	85.9 %	87.4 %	93.6 %
All frozen, random. init.	50.0 %	50.0 %	100.0 %	66.6 %	50.0 %

Table 4.11: Performance of the MobileNet network with transferred weights of its convolutional layers, compared to training with its weights randomly initialised.

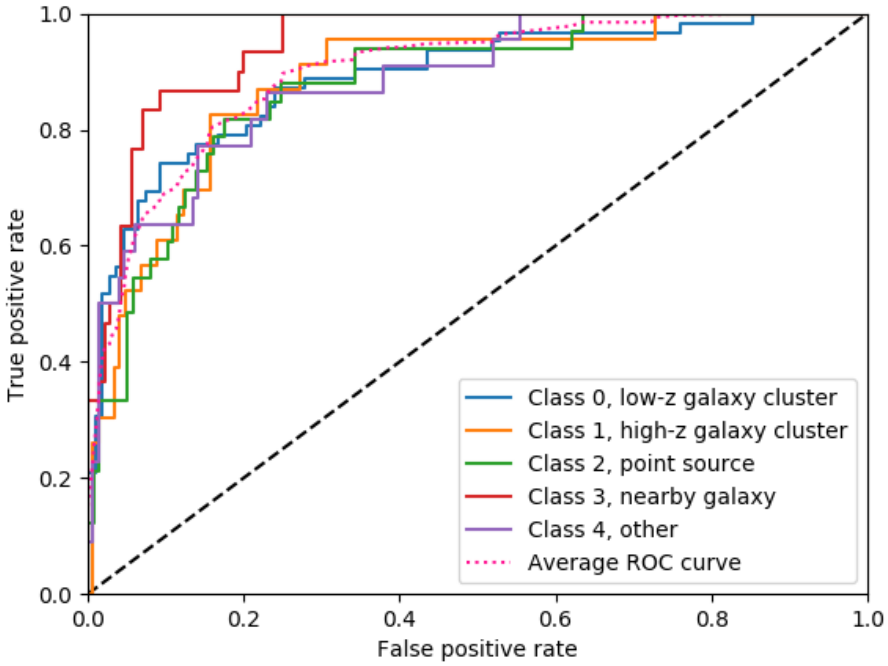

Figure 4.5: ROC curves of multi-class classification of our custom network.

Class	Accuracy	Precision	Recall	F_1 score	AUC
Low-z cluster	78.8 %	67.6 %	80.6 %	73.5 %	89.0 %
High-z cluster	84.1 %	44.1 %	65.2 %	52.6 %	88.5 %
Point source	84.1 %	75.0 %	27.3 %	40.0 %	87.8 %
Nearby galaxy	88.8 %	73.9 %	56.7 %	64.2 %	84.5 %
Other	87.6 %	51.8 %	63.6 %	57.1 %	88.4 %

Table 4.12: Performance of our custom network for multi-class classificaton.

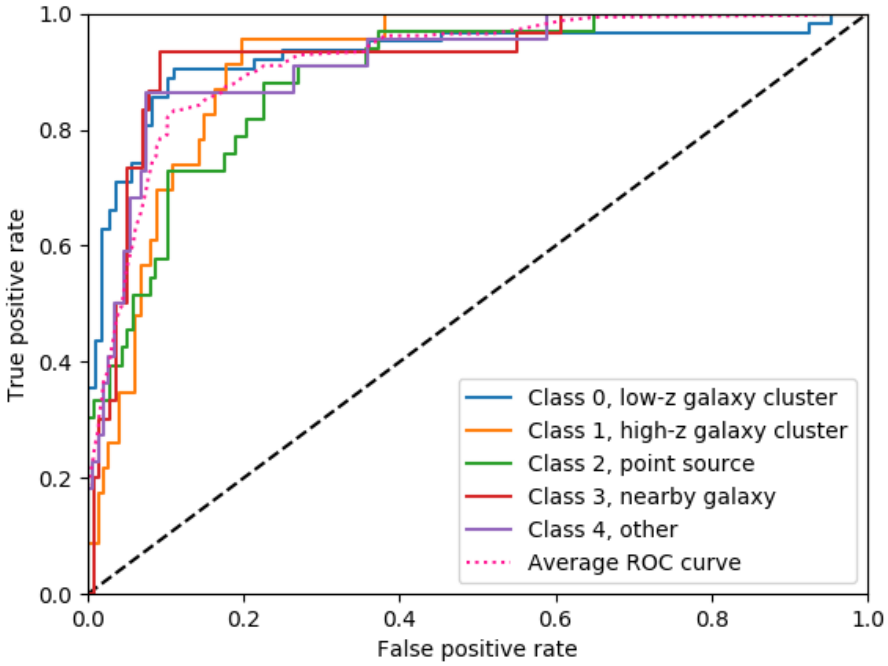

Figure 4.6: ROC curves of multi-class classification performed by transfer learning with MobileNet's convolutional layers.

Class	Accuracy	Precision	Recall	F_1 score	AUC
Low-z cluster	77.1 %	62.3 %	93.5 %	74.8 %	92.6 %
High-z cluster	87.1 %	55.6 %	21.7 %	31.2 %	91.1 %
Point source	86.5 %	85.7 %	36.3 %	51.0 %	89.0 %
Nearby galaxy	90.0 %	70.1 %	73.3 %	72.1 %	92.4 %
Other	91.2 %	65.2 %	68.2 %	66.7 %	91.6 %

Table 4.13: Performance of the transfer learning with the MobileNet for multi-class classificaton.

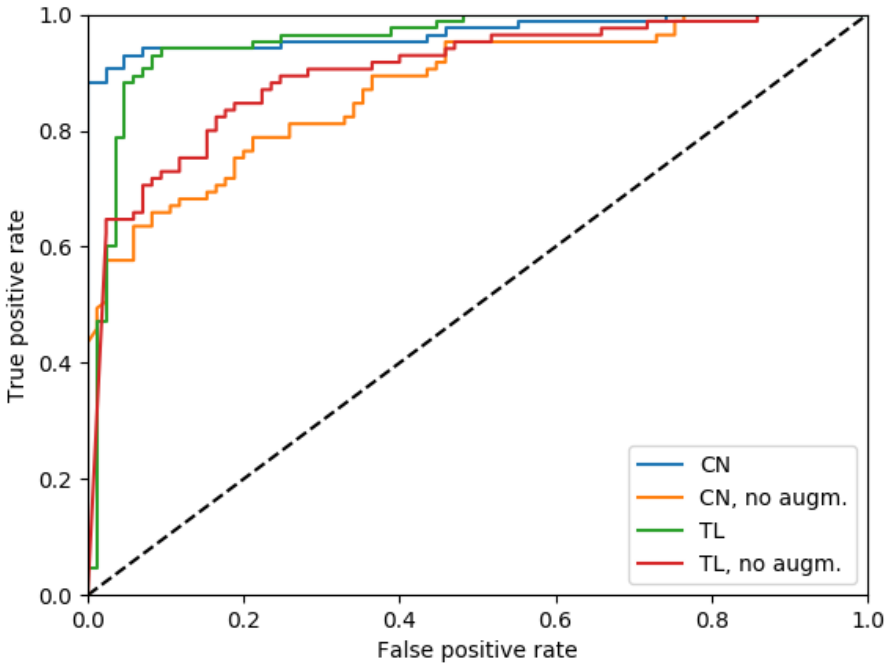

Figure 4.7: ROC curves of our custom network and the transferred MobileNet network with and without training augmentation.

Network	Clusters		Non-clusters	
	N_{TP}	N_{FN}	N_{TN}	N_{FP}
CN	77	8	83	2
CN, no augm.	77	6	46	39
TL	75	10	81	4
TL, no augm.	77	8	55	30

Table 4.14: Classification results for our custom network and the transfer learning with the MobileNet when training augmentation is and is not applied.

Network	Accuracy	Precision	Recall	F_1 score	AUC
CN	94.1 %	97.5 %	90.6 %	93.9 %	96.9 %
CN, no augm.	73.5 %	66.9 %	92.9 %	77.8 %	87.31 %
TL	91.8 %	94.9 %	88.2 %	91.5 %	95.5 %
TL, no augm.	77.6 %	72.0 %	90.6 %	80.2 %	90.1 %

Table 4.15: Performance of our custom network and the transfer learning with the MobileNet with and without augmentation.

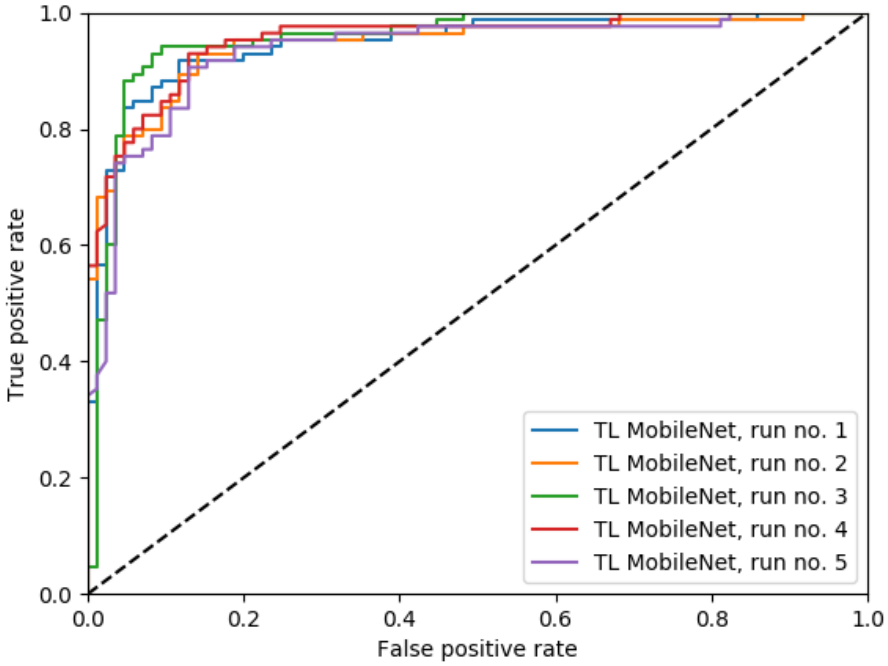

Figure 4.8: ROC curves of the same transfer learning MobileNet network, trained multiple times with exactly same hyperparameters, to explore effects of final layers random initialisation on the network's performance.

Network	Clusters		Non-clusters	
	N_{TP}	N_{FN}	N_{TN}	N_{FP}
TL MobileNet, run no. 1	75	10	76	9
TL MobileNet, run no. 2	79	6	71	14
TL MobileNet, run no. 3	75	10	81	4
TL MobileNet, run no. 4	81	4	70	15
TL MobileNet, run no. 5	77	8	74	11

Table 4.16: Classification results of the transfer learning MobileNet with exactly same hyperparameters, to examine the effects of the top dense layer's random initialisation on networks results.

Network	Accuracy	Precision	Recall	F_1 score	AUC
TL MobileNet, run no. 1	88.8 %	89.2 %	88.2 %	88.8 %	94.7 %
TL MobileNet, run no. 2	88.2 %	84.9 %	92.9 %	88.7 %	94.5 %
TL MobileNet, run no. 3	91.8 %	94.9 %	88.2 %	91.4 %	95.5 %
TL MobileNet, run no. 4	88.8 %	84.4 %	95.3 %	89.5 %	95.4 %
TL MobileNet, run no. 5	88.8 %	84.4 %	95.3 %	89.5 %	93.4 %

Table 4.17: Performance of the transferred MobileNet network with the same hyperparameters trained multiple times to probe the randomness effect of the top dense layer's initialisation on networks performance.

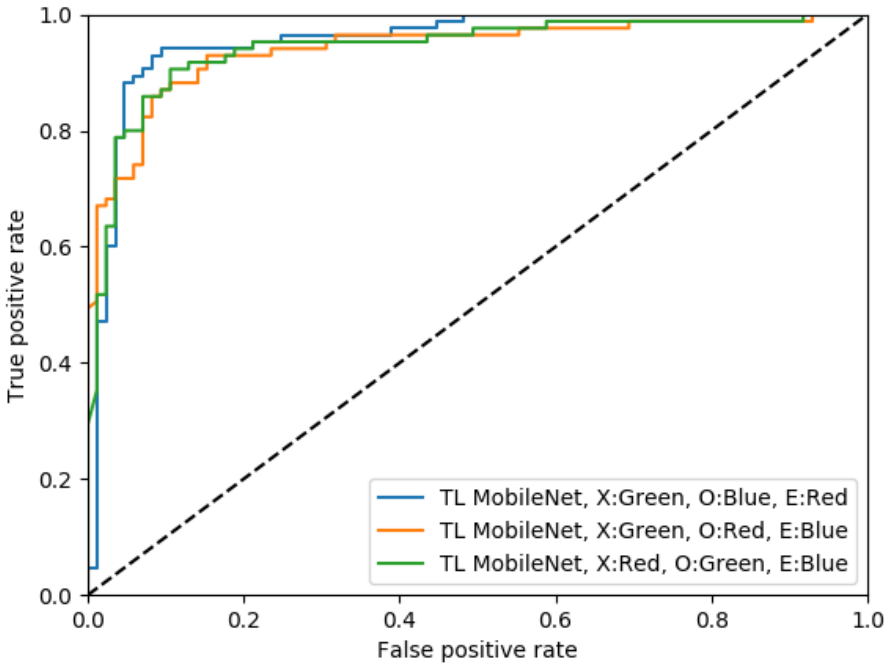

Figure 4.9: ROC curves of the transfer learning with the MobileNet trained on the data sets with different channel order. X refers to X-ray, O to optical and E to empty channel.

Network	Clusters		Non-clusters	
	N_{TP}	N_{FN}	N_{TN}	N_{FP}
TL, X:Green, O:Blue, E:Red	75	10	76	9
TL, X:Green, O:Red, E:Blue	79	6	70	15
TL, X:Red, O:Green, E:Blue	78	7	72	13

Table 4.18: Classification results of the transferred MobileNet network with the same hyperparameters, trained on data with different channel order. X refers to X-ray, O to optical and E to empty channel.

Network	Accuracy	Precision	Recall	F_1 score	AUC
TL, X:Green, O:Blue, E:Red	91.8 %	94.9 %	88.2 %	91.4 %	95.5 %
TL, X:Green, O:Red, E:Blue	87.6 %	84.0 %	92.9 %	88.2 %	94.1 %
TL, X:Red, O:Green, E:Blue	88.2 %	85.7 %	91.7 %	88.6 %	94.3 %

Table 4.19: Performance of the transferred MobileNet network with the same hyperparameters, trained on data with different channel order. X refers to X-ray, O to optical and E to empty channel.

4.3 Images of classified objects

Fig. 4.10 shows some of the spectroscopically confirmed galaxy clusters which have been classified with a probability higher than 99 % being galaxy clusters by our custom network. The network identified with this very high probability nearby and also distant galaxy clusters. The first 4 objects in the figure are examples of the nearby galaxy clusters, the last 2 are examples of distant galaxy clusters. All of them have very prominent extended X-ray emission.

Fig. 4.11, Fig. 4.12 and Fig. 4.13 show images of objects which have been incorrectly classified by our custom network. The non-galaxy clusters were probably classified as clusters due to a faint X-ray emission present et the centre of those objects. Most of the clusters which have been missed by our network are very similar to nearby galaxies, having at the centre one elliptical galaxy of a similar size in optical as the size of clusters X-ray emission.

4.4 Visualisation of filter's activations

Visualising outputs of individual filters in a convolutional layer, when presented an input image selected for classification, helps to conclude if the network learned to extract relevant features. Fig. 4.14 shows activations of three selected filters from each convolutional layer of our custom network when the network classifies a spectroscopically confirmed galaxy cluster. Filters learned to look for features such as colour patches of X-ray and optical emission, edges of X-ray and optical light and their combinations. Filters learned to look for a combination of the X-ray and optical emission more commonly than only one of the two. Different filters learned to search for the emission of various intensity, helping the network to tell apart different objects. Size of filter's output decreases with deeper convolutional layers, because of the max-pooling operation applied after each convolutional layer, making filter outputs of layers progressively smaller. Output sizes of individual filters for a specific layer are described in Fig. 4.15.

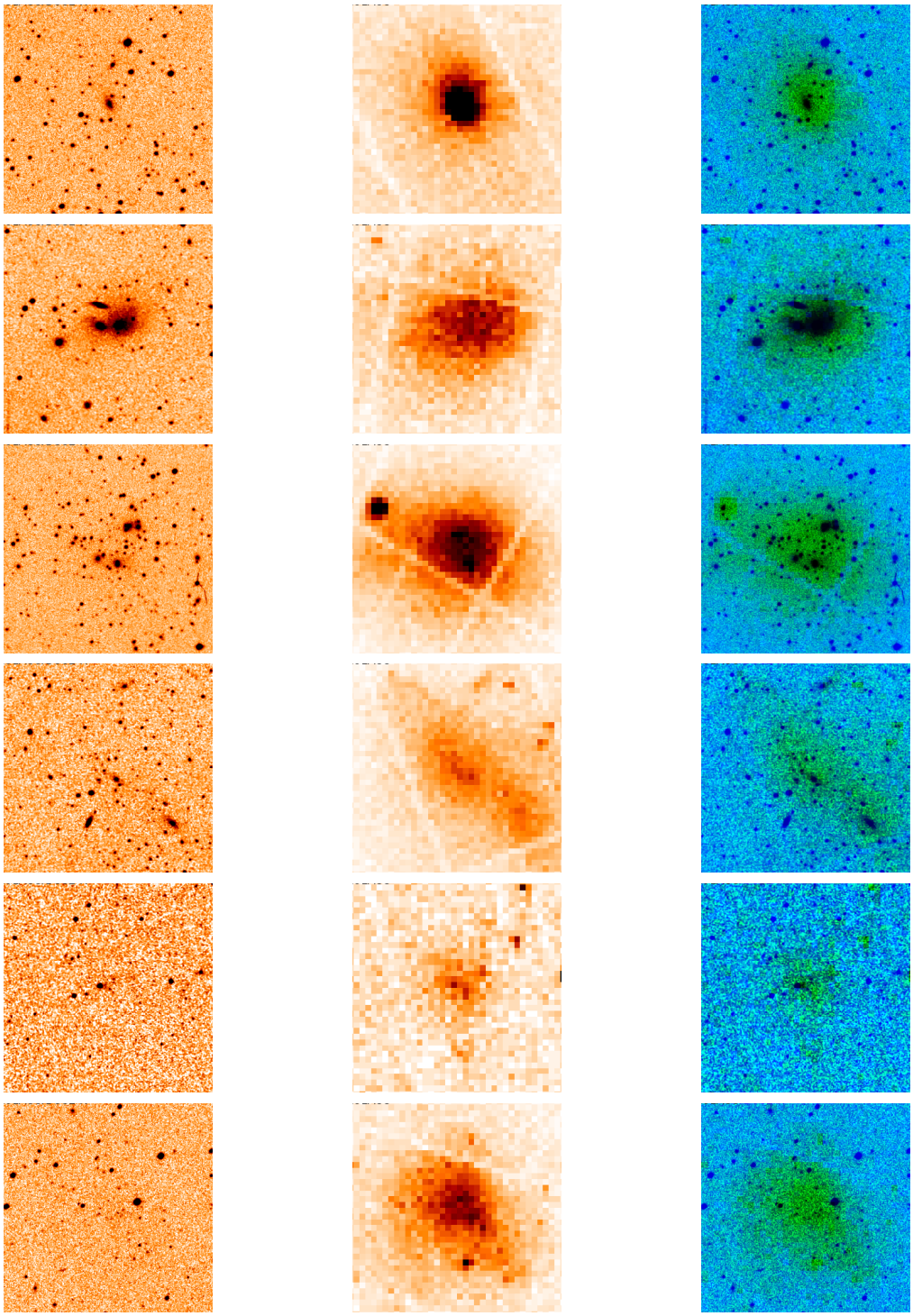

Figure 4.10: Examples of spectroscopically confirmed galaxy clusters which our custom network correctly classified with a probability higher than 99 %. Each row shows one object. On the left is an optical image, the middle is the X-ray image and right is a concatenation of the two with the red channel left with zeroes.

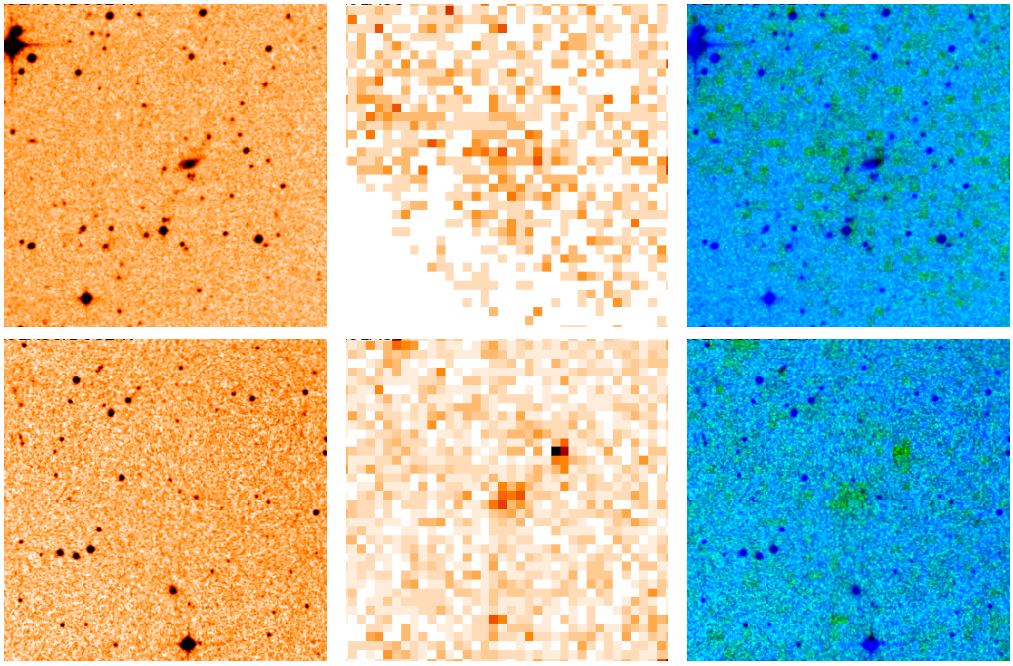

Figure 4.11: Non-galaxy clusters incorrectly classified as galaxy clusters N_{FP} by our custom network (2 objects of 85 non-clusters in the test sample). Each row shows one object, left: optical image, middle: X-ray image, right: a concatenated image of optical (blue channel), X-ray (green channel) and red channel left with zeroes.

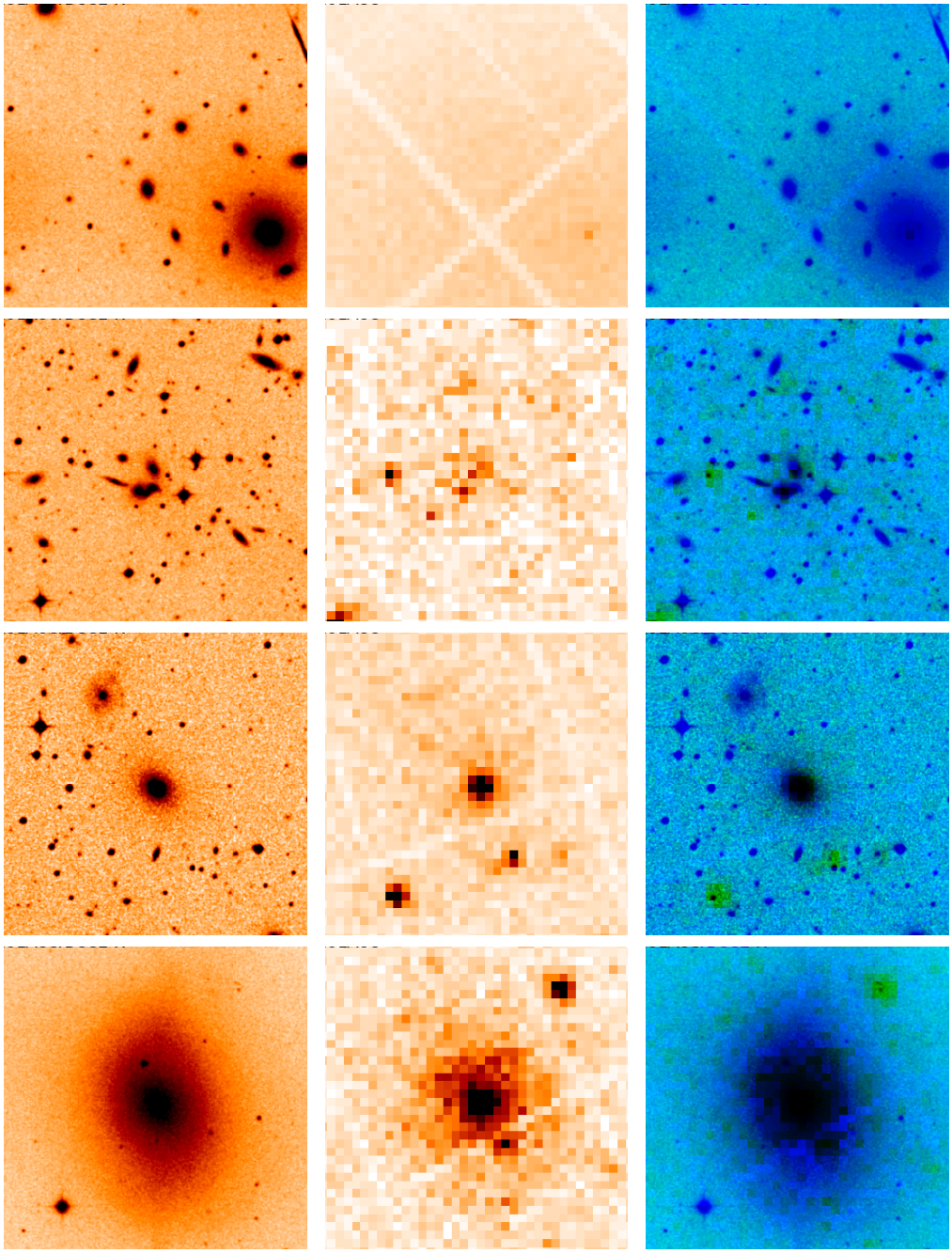

Figure 4.12: The first 4 of the 8 galaxy clusters incorrectly classified as non-galaxy clusters N_{FN} by our custom network. Each row shows one object, left: optical image, middle: X-ray image, right: a concatenated image of optical (blue channel), X-ray (green channel) and red channel left with zeroes.

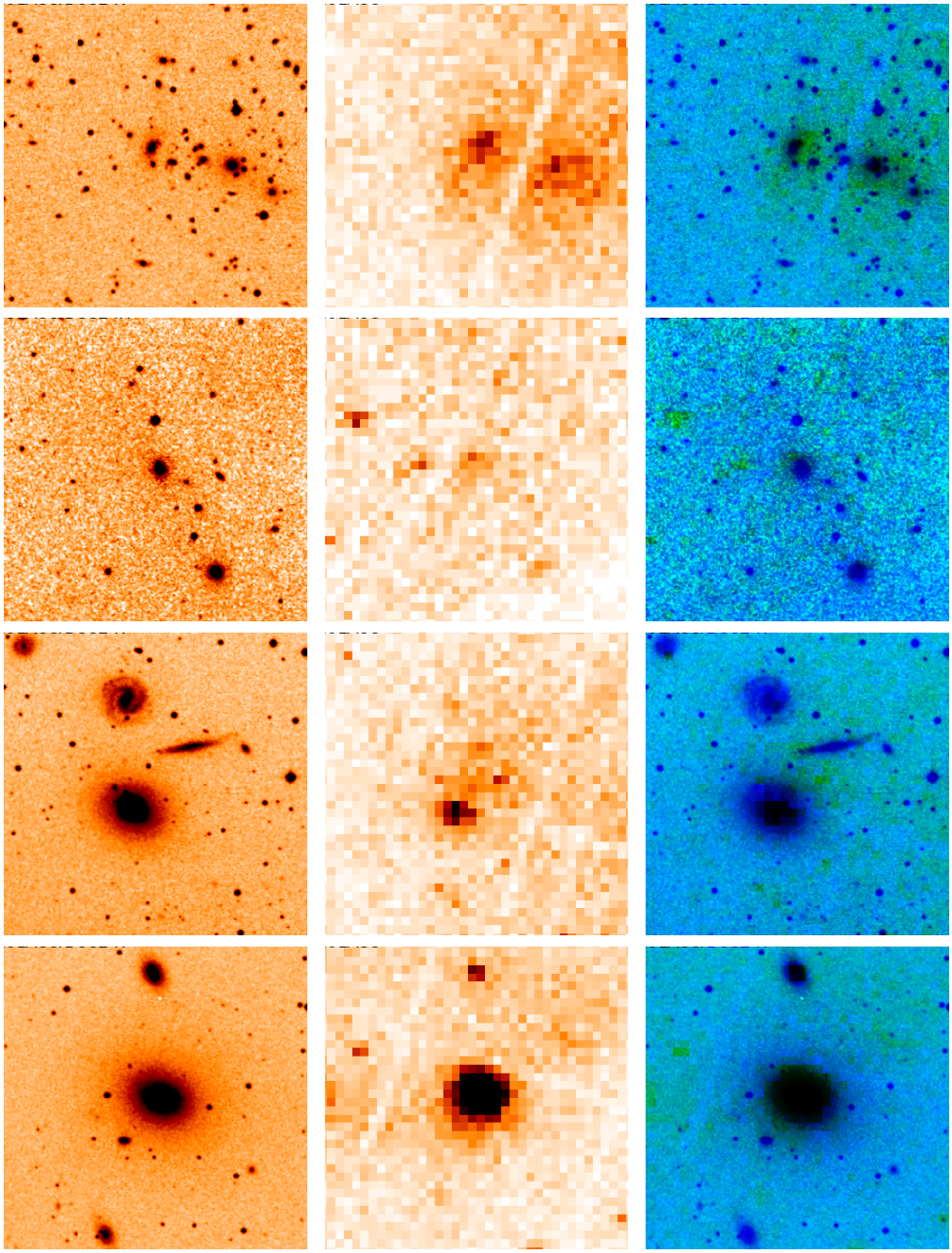

Figure 4.13: The other 4 galaxy clusters incorrectly classified as non-galaxy clusters (N_{FN}) by our custom network (8 objects in a total of 85 spectroscopically confirmed clusters in the test sample). Each row shows one object, left: optical image, middle: X-ray image, right: a concatenated image of optical (blue channel), X-ray (green channel) and red channel left with zeroes.

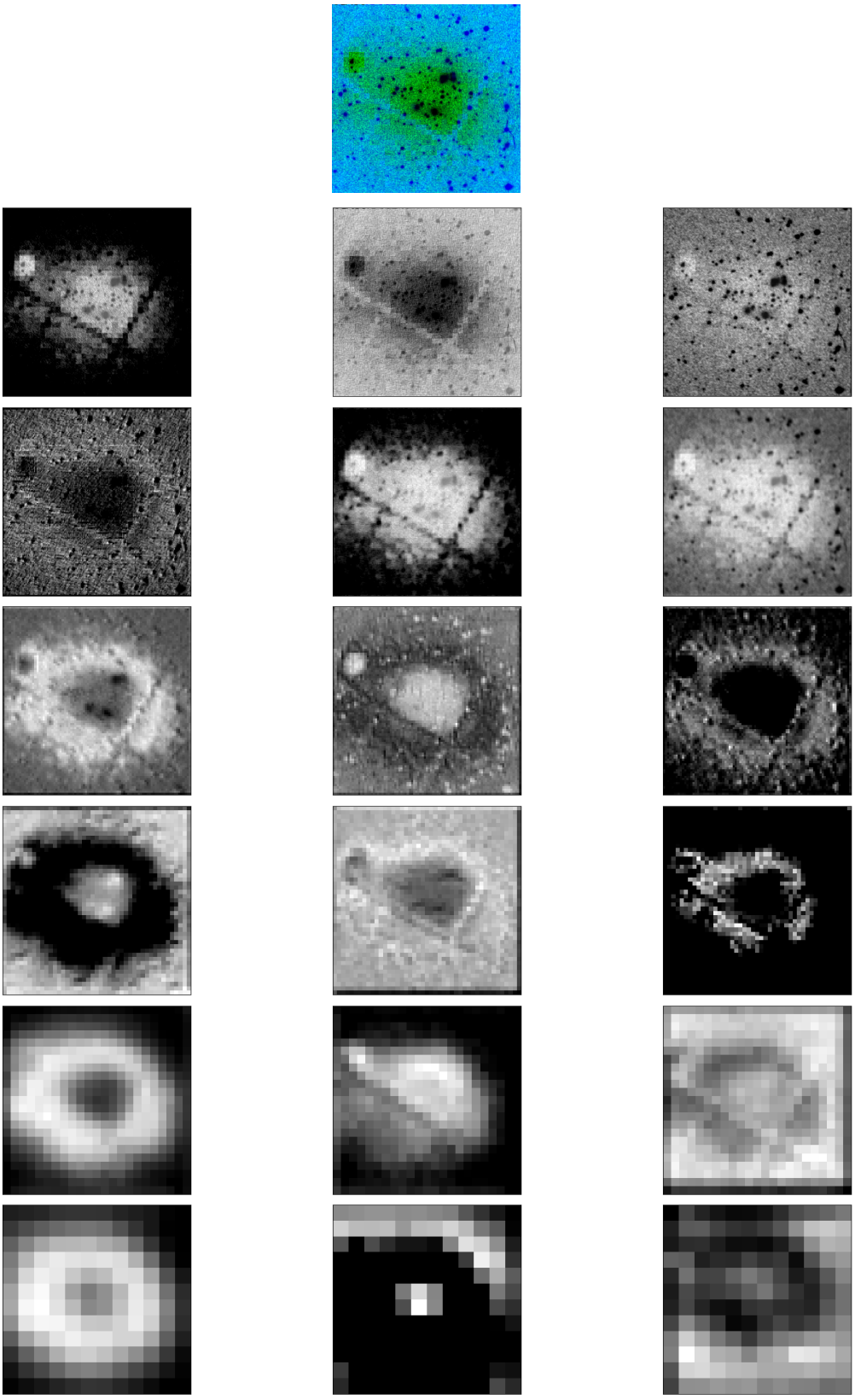

Figure 4.14: Activation maps of a few selected kernels from each convolutional layer of our custom network. Top: the concatenation of grayscaled optical and X-ray images used as single channels, blue and green respectively. The red channel is left with zeroes. Each row from second to seventh shows activation maps of three chosen filters from the first to the sixth layer of our custom network, respectively.

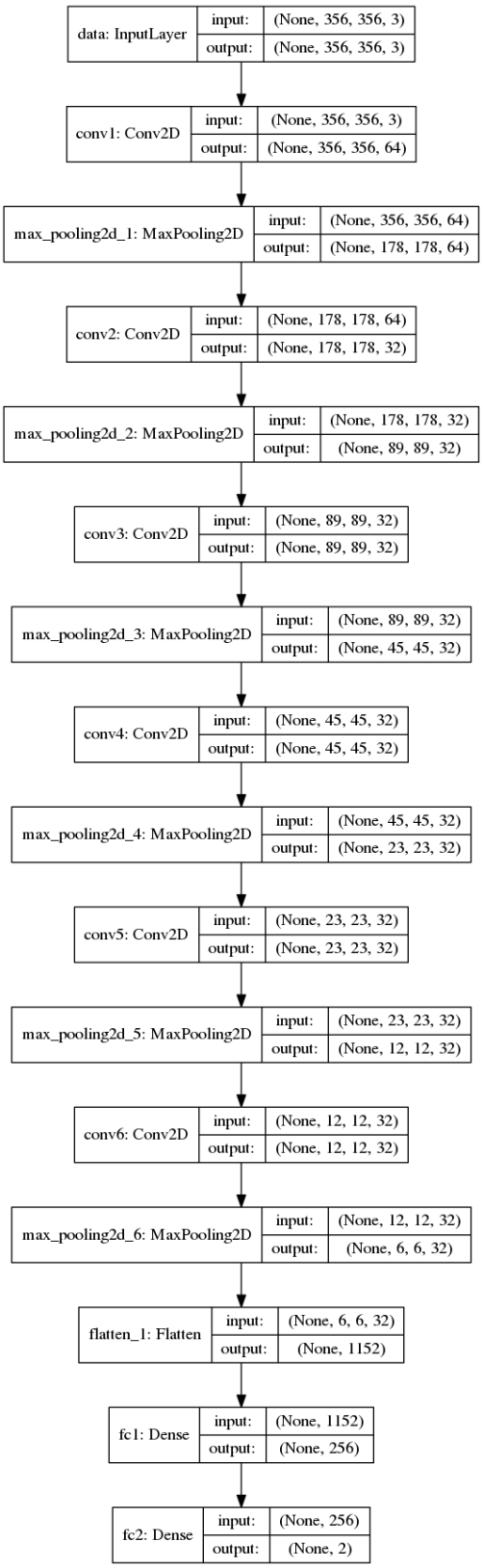

Figure 4.15: The architecture of our custom network which achieved the best performance.

Chapter 5

Conclusion and future plans

5.1 The Hunt for Galaxy Clusters

Our crowd-sourcing Zooniverse citizen science project, The Hunt for Galaxy Clusters incorporated 1 600 cluster candidates picked by the `XAmin` pipeline. Each cluster candidate had a threshold 30 classifications to be retired. It took about 6 months to all of the 1 227 Zooniverse volunteers who participated to finish the classifications.

The volunteers achieved 62 % agreement with experts calculated on a sample of 404 cluster candidates classified by both, volunteers and experts. The experts classified 254 of those as clusters and 150 as non-clusters. The Zooniverse volunteers agreed with the experts on 102 clusters and 149 non-clusters achieving 61.1 % accuracy (Tab. 4.1). However, this sample contained 254 objects classified as clusters by experts, making it biased towards clusters. Given that the Zooniverse volunteers struggled to classify clusters, performance measurement on such a sample necessarily decreases volunteer's preciseness. When this test sample was reduced to be balanced, having 150 clusters and 150 non-clusters (classifications by experts), the Zooniverse volunteers achieved 72.2 % agreement with the experts. On the other hand, the Zooniverse volunteers created an extremely clear sample of clusters, classifying only 1 object as a *cluster* which has been claimed as *non-cluster* by experts. Such a data set might be very useful for cosmological studies, which require cluster catalogue with high purity.

Comparing the Zooniverse volunteers classifications of objects classified as clusters by the experts reveals, that the volunteers experienced several biases (Fig. 4.1). The Zooniverse volunteers were mostly biased to identify extended X-ray emission as *point-like*, *no emission*, emission close to the *edge* and as a *nearby galaxy*. The bias of the *point-like* category points out that the Zooniverse volunteers battled to identify the size of the X-ray emission. Many of their classifications ended in the *edge* category, which was intended for the X-ray emission close to the chip edge of the XMM-Newton's field of view. This bias has been linked with the misconception of the small chip gaps between individual XMM-Newton's detectors with the edge. The bias to classify as *no emission* might have been caused by struggling to identify fainter X-ray emission.

Distant galaxy clusters might have been missed by the Zooniverse volunteers

due to their biases of classifying extended X-ray emission as a point-like and no emission. The Zooniverse volunteers were also prone to classify nearby galaxy clusters with prominent brightest central cluster's galaxy as a nearby galaxy. This is a way how many nearby clusters were missed.

Adding example images focused on clarifying the biases of the Zooniverse volunteers could bolster the agreement between the volunteers and the experts. Re-writing some of the questions in the project's decision tree (Fig. 3.3) might also help. For example, the third question, asking if the X-ray emission is dominated by a single galaxy or not. The second question asking about the extension of the X-ray emission, which was subjected to the highest bias, would need more example subjects in its help note, to better describe how to distinguish the point-like and extended X-ray emission.

Keeping communication with the volunteers in the project's talk section was undeniably critical for maintaining the project alive. We have answered all of the volunteer's questions, explaining them the physics behind clusters together with instrumental effects behind various artefacts in the images. We believe that the volunteers might be able to find even more clusters if a more detailed explanation of the biases they experienced had been given in the help notes. Adding more example images to directly put a light on those biases would also increase their performance, but we conclude the Zooniverse project being successful because as discussed, the volunteers showed an ability to create a very pure sample of clusters.

5.2 Neural networks

We have proved convolutional neural networks to be highly accurate automatic galaxy cluster classifiers. Our hand-made custom network achieved the best performance for the automatic classification of galaxy cluster candidates in two classes: *cluster* and *non-cluster* (Fig. 4.2). It was trained using multiwavelength, X-ray and optical, observations and tested on a sample of 85 spectroscopically confirmed clusters and 85 objects classified as non-clusters by the experts. The best network correctly classified 77 clusters and 83 non-clusters (Tab. 4.6), achieving 94.1 % accuracy (Tab. 4.7). It classified only 2 non-clusters as clusters, resulting in a very pure sample of clusters, which is especially useful for cosmological studies.

The transfer learning approach with the MobileNet network achieved nearly the same performance, correctly classifying 75 clusters and 81 non-clusters, obtaining accuracy 91.8 %.

Our custom network, trained on the Zooniverse classifications, achieved 80.0 % accuracy, correctly classifying 60 clusters and 76 non-clusters. It cought the same bias as the Zooniverse volunteers had, making more classifications for *non-cluster* class.

Neural networks trained for multi-class classification of classes *low-z cluster*, *high-z cluster*, *point source*, *nearby galaxy* and *other*, resulted in an overall

accuracy 84.7 % (Tab. 4.12) for our custom network (Fig. 4.5) and 86.4 % accuracy (Tab. 4.13) for the transfer learning with the MobileNet network (Fig. 4.6).

Training of our custom network only on the X-ray images, without their optical counterparts (Fig. 4.2), achieved accuracy 81.2 % (Tab. 4.7). The network correctly classified 78 clusters and 60 non-clusters (Tab. 4.6), being biased to classify objects as non-clusters. The custom network trained on solo optical observations properly classified only 67 clusters and 49 non-clusters, achieving accuracy 68.2 %. It points out that the X-ray data are more suited to search for galaxy clusters, they do not suffer from that many projection effects as optical data do.

The transfer learning with the MobileNet architecture did not work very well with all of its transferred convolutional layers frozen (Fig. 4.3, Tab. 4.8 and Tab. 4.9). Unfreezing its transferred layers was undoubtedly necessary to achieve competitive results with the best performing, our custom hand-made network.

Training augmentation (Tab. 3.1) was absolutely inevitable in order to achieve good performance (Fig. 4.7, Tab. 4.14 and Tab. 4.15). The accuracy of our custom network would drop from 94.4 % to 73.5 % without the training augmentation. Both networks, our custom network and the transfer learning with the MobileNet experienced strong bias towards classyfing objects as clusters whithout training augmentation (Tab. 4.14).

The `XAmin` pipeline is used mainly to filter X-ray point sources. Considering that this pipeline created the training data set for the networks and all of the objects in it are classified as extended sources by the pipeline, the sample of point sources in our data set is biased. We can not consider the neural network trained for multiclass classification on this data set as a reliable point source classifier, due to this bias. Further examination of its performance on a sample of point sources with general appearance would have to be done in order to explore the network's point source classification efficiency.

5.3 Future plans

The upcoming large sky surveys, such as Athena or eROSITA, observing in X-ray, LSST in optical, or JWST and EUCLID in infrared bands, will encompass tens of thousands new galaxy clusters in their vast data sets. Those surveys will create hundreds of gigabytes of data per night, so the development of highly accurate, automatic software for classification of astrophysical objects is the next logical and inevitable step forward.

We have created a neural network capable of a very accurate automatic galaxy cluster classification. The ultimate vision of our project is to use the network to create a catalogue of galaxy clusters, which will be finally used to conduct cosmological studies and will open doors to the astrophysical investigation of a very large galaxy cluster sample.

We are currently utilising a data set of $(2.0 - 10.0)$ keV band X-ray observations, which is going to be added to the training of the networks. Hardness

ratio maps of the $(0.5-2.0)$ keV and $(2.0-10.0)$ keV bands might be also used for training of the networks. This additional information should enable the networks to achieve better performance.

Galaxy clusters halos exhibit various properties in radio observations. Training our custom network on radio observations only or adding those as an additional channel to the data we currently use is another step we are planning to carry out. This project has a great potential to probe the interesting physics behind the cluster's radio halos.

another direction we are considering to go lies in the Sunyaev–Zeldovich effect (SZ), shadows of the cosmic microwave background created when its photons pass through the intra-cluster medium. SZ is a great tool for search of galaxy clusters, mainly because of its redshift independence. Applying our methods for SZ observations might result in the finding of many new galaxy clusters. Galaxy clusters with high redshift are exceptionally interesting for constraining cosmological parameters, and SZ automatic galaxy cluster search with our methods could result in finding many new of those, seriously impacting nowadays cosmological knowledge.

We believe that the methods we have developed are going to be extremely useful for the future of astrophysics and science in general, which are being overwhelmed by ever-increasing data volumes.

Bibliography

Kevork Abazajian, Jennifer K. Adelman-McCarthy, Marcel A. Agüeros, Sahar S. Allam, Kurt Anderson, Scott F. Anderson, James Annis, Neta A. Bahcall, Ivan K. Baldry, Steven Bastian, Andreas Berlind, Mariangela Bernardi, Michael R. Blanton, Jr. Bochanski, John J., William N. Boroski, John W. Briggs, J. Brinkmann, Robert J. Brunner, Tamás Budavári, Larry N. Carey, Samuel Carliles, Francisco J. Castander, A. J. Connolly, István Csabai, Mamoru Doi, Feng Dong, Daniel J. Eisenstein, Michael L. Evans, Xiaohui Fan, Douglas P. Finkbeiner, Scott D. Friedman, Joshua A. Frieman, Masataka Fukugita, Roy R. Gal, Bruce Gillespie, Karl Glazebrook, Jim Gray, Eva K. Grebel, James E. Gunn, Vijay K. Gurbani, Patrick B. Hall, Masaru Hamabe, Frederick H. Harris, Hugh C. Harris, Michael Harvanek, Timothy M. Heckman, John S. Hendry, Gregory S. Hennessy, Robert B. Hindsley, Craig J. Hogan, David W. Hogg, Donald J. Holmgren, Shin-ichi Ichikawa, Takashi Ichikawa, Željko Ivezić, Sebastian Jester, David E. Johnston, Anders M. Jorgensen, Stephen M. Kent, S. J. Kleinman, G. R. Knapp, Alexei Yu. Kniazev, Richard G. Kron, Jurek Krzesinski, Peter Z. Kunszt, Nickolai Kuropatkin, Donald Q. Lamb, Hubert Lampeitl, Brian C. Lee, R. French Leger, Nolan Li, Huan Lin, Yeong-Shang Loh, Daniel C. Long, Jon Loveday, Robert H. Lupton, Tanu Malik, Bruce Margon, Takahiko Matsubara, Peregrine M. McGehee, Timothy A. McKay, Avery Meiksin, Jeffrey A. Munn, Reiko Nakajima, Thomas Nash, Jr. Neilsen, Eric H., Heidi Jo Newberg, Peter R. Newman, Robert C. Nichol, Tom Nicinski, Maria Nieto-Santisteban, Atsuko Nitta, Sadanori Okamura, William O'Mullane, Jeremiah P. Ostriker, Russell Owen, Nikhil Padmanabhan, John Peoples, Jeffrey R. Pier, Adrian C. Pope, Thomas R. Quinn, Gordon T. Richards, Michael W. Richmond, Hans-Walter Rix, Constance M. Rockosi, David J. Schlegel, Donald P. Schneider, Ryan Scranton, Maki Sekiguchi, Uros Seljak, Gary Sergey, Branimir Sesar, Erin Sheldon, Kazu Shimasaku, Walter A. Siegmund, Nicole M. Silvestri, J. Allyn Smith, Vernesa Smolčić, Stephanie A. Snedden, Albert Stebbins, Chris Stoughton, Michael A. Strauss, Mark SubbaRao, Alexander S. Szalay, István Szapudi, Paula Szkody, Gyula P. Szokoly, Max Tegmark, Luis Teodoro, Aniruddha R. Thakar, Christy Tremonti, Douglas L. Tucker, Alan Uomoto, Daniel E. Vanden Berk, Jan Vandenberg, Michael S. Vogeley, Wolfgang Voges, Nicole P. Vogt, Lucianne M. Walkowicz, Shu-i. Wang, David H. Weinberg, Andrew A. West, Simon D. M. White,

Brian C. Wilhite, Yongzhong Xu, Brian Yanny, Naoki Yasuda, Ching-Wa Yip, D. R. Yocum, Donald G. York, Idit Zehavi, Stefano Zibetti, and Daniel B. Zucker. The Second Data Release of the Sloan Digital Sky Survey. AJ, 128: 502–512, Jul 2004. doi: 10.1086/421365.

G. O. Abell. The Distribution of Rich Clusters of Galaxies. ApJS, 3:211, May 1958. doi: 10.1086/190036.

S. W. Allen, R. W. Schmidt, H. Ebeling, A. C. Fabian, and L. van Speybroeck. Constraints on dark energy from Chandra observations of the largest relaxed galaxy clusters. MNRAS, 353:457–467, Sep 2004. doi: 10.1111/j.1365-2966.2004.08080.x.

S. W. Allen, D. A. Rapetti, R. W. Schmidt, H. Ebeling, R. G. Morris, and A. C. Fabian. Improved constraints on dark energy from Chandra X-ray observations of the largest relaxed galaxy clusters. MNRAS, 383:879–896, Jan 2008. doi: 10.1111/j.1365-2966.2007.12610.x.

J. Annis, M. Makler, S. Kent, S. Dodelson, J. Frieman, E. Sheldon, T. McKay, N. Bahcall, and SDSS Collaboration. SDSS Cluster Abundance and Cosmological Implications. In *American Astronomical Society Meeting Abstracts #200*, volume 200 of *American Astronomical Society Meeting Abstracts*, page 77.06, May 2002.

J. M. Bardeen, P. J. Steinhardt, and M. S. Turner. Spontaneous creation of almost scale-free density perturbations in an inflationary universe. Phys. Rev. D, 28: 679–693, August 1983. doi: 10.1103/PhysRevD.28.679.

J. M. Bardeen, J. R. Bond, N. Kaiser, and A. S. Szalay. The statistics of peaks of Gaussian random fields. ApJ, 304:15–61, May 1986. doi: 10.1086/164143.

L. P. Bautz and W. W. Morgan. On the Classification of the Forms of Clusters of Galaxies. ApJ, 162:L149, December 1970. doi: 10.1086/180643.

Y. Bengio, P. Simard, and P. Frasconi. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5(2): 157–166, 1994.

E. Bertin and S. Arnouts. SExtractor: Software for source extraction. A&AS, 117:393–404, June 1996. doi: 10.1051/aas:1996164.

J. R. Bond and G. Efstathiou. Cosmic background radiation anisotropies in universes dominated by nonbaryonic dark matter. ApJ, 285:L45–L48, October 1984. doi: 10.1086/184362.

J. R. Bond, L. Kofman, and D. Pogosyan. How filaments of galaxies are woven into the cosmic web. Nature, 380:603–606, April 1996. doi: 10.1038/380603a0.

G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000.

David A. Buote. On the Origin of Radio Halos in Galaxy Clusters. ApJ, 553: L15–L18, May 2001. doi: 10.1086/320500.

David A. Buote and John C. Tsai. Quantifying the Morphologies and Dynamical Evolution of Galaxy Clusters. I. The Method. ApJ, 452:522, Oct 1995. doi: 10.1086/176326.

P. Carnevali, A. Cavaliere, and P. Santangelo. Merging instability in groups of galaxies. ApJ, 249:449–461, October 1981. doi: 10.1086/159305.

A. Cavaliere and R. Fusco-Femiano. X-rays from hot plasma in clusters of galaxies. A&A, 49:137–144, May 1976.

N. Clerc, T. Sadibekova, M. Pierre, F. Pacaud, J.-P. Le Fevre, C. Adami, B. Altieri, and I. Valtchanov. VizieR Online Data Catalog: X-ray clusters from XMM (Clerc+, 2012). *VizieR Online Data Catalog*, 742, September 2012a.

N. Clerc, T. Sadibekova, M. Pierre, F. Pacaud, J.-P. Le Fèvre, C. Adami, B. Altieri, and I. Valtchanov. The cosmological analysis of X-ray cluster surveys - II. Application of the CR-HR method to the XMM archive. MNRAS, 423:3561–3583, July 2012b. doi: 10.1111/j.1365-2966.2012.21153.x.

Neal Dalal, Olivier Doré, Dragan Huterer, and Alexander Shirokov. Imprints of primordial non-Gaussianities on large-scale structure: Scale-dependent bias and abundance of virialized objects. Phys. Rev. D, 77:123514, Jun 2008. doi: 10.1103/PhysRevD.77.123514.

G. B. Dalton, S. J. Maddox, W. J. Sutherland, and G. Efstathiou. The APM Galaxy Survey - V. Catalogues of galaxy clusters. MNRAS, 289:263–284, Aug 1997. doi: 10.1093/mnras/289.2.263.

J. Deng, W. Dong, R. Socher, L. Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, June 2009. doi: 10.1109/CVPR.2009.5206848.

R. H. Dicke, P. J. E. Peebles, P. G. Roll, and D. T. Wilkinson. Cosmic Black-Body Radiation. ApJ, 142:414–419, Jul 1965. doi: 10.1086/148306.

S. Dieleman, K. W. Willett, and J. Dambre. Rotation-invariant convolutional neural networks for galaxy morphology prediction. MNRAS, 450:1441–1459, June 2015. doi: 10.1093/mnras/stv632.

Jürg Diemand, Michael Kuhlen, and Piero Madau. Formation and Evolution of Galaxy Dark Matter Halos and Their Substructure. ApJ, 667:859–877, Oct 2007. doi: 10.1086/520573.

Chuong B. Do and Andrew Y. Ng. Transfer learning for text classification. In *Proceedings of the 18th International Conference on Neural Information Processing Systems*, NIPS'05, pages 299–306, Cambridge, MA, USA, 2005. MIT Press. URL `http://dl.acm.org/citation.cfm?id=2976248.2976286`.

Megan Donahue and G. Mark Voit. Cool Gas in Clusters of Galaxies. In J. S. Mulchaey, A. Dressler, and A. Oemler, editors, *Clusters of Galaxies: Probes of Cosmological Structure and Galaxy Evolution*, page 143, Jan 2004.

A. C. Edge. The detection of molecular gas in the central galaxies of cooling flow clusters. MNRAS, 328:762–782, Dec 2001. doi: 10.1046/j.1365-8711.2001.04802.x.

D. J. Eisenstein and W. Hu. Power Spectra for Cold Dark Matter and Its Variants. ApJ, 511:5–15, January 1999. doi: 10.1086/306640.

Vincent R. Eke, Shaun Cole, Carlos S. Frenk, and J. Patrick Henry. Measuring Omega_0 using cluster evolution. MNRAS, 298:1145–1158, Aug 1998. doi: 10.1046/j.1365-8711.1998.01713.x.

A. E. Evrard. Beyond N-body - 3D cosmological gas dynamics. MNRAS, 235: 911–934, December 1988. doi: 10.1093/mnras/235.3.911.

A. C. Fabian. Cooling Flows in Clusters of Galaxies. *Annual Review of Astronomy and Astrophysics*, 32:277–318, Jan 1994. doi: 10.1146/annurev.aa.32.090194.001425.

R. T. Farouki, G. L. Hoffman, and E. E. Salpeter. The collapse and violent relaxation of N-body systems - Mass segregation and the secondary maximum. ApJ, 271:11–21, August 1983. doi: 10.1086/161171.

J. E. Felten, R. J. Gould, W. A. Stein, and N. J. Woolf. X-Rays from the Coma Cluster of Galaxies. ApJ, 146:955–958, Dec 1966. doi: 10.1086/148972.

R. R. Gal, R. R. de Carvalho, S. C. Odewahn, S. G. Djorgovski, and V. E. Margoniner. The Northern Sky Optical Cluster Survey. I. Detection of Galaxy Clusters in DPOSS. AJ, 119:12–20, January 2000. doi: 10.1086/301185.

R. R. Gal, R. R. de Carvalho, P. A. A. Lopes, S. G. Djorgovski, R. J. Brunner, A. Mahabal, and S. C. Odewahn. The Northern Sky Optical Cluster Survey. II. An Objective Cluster Catalog for 5800 Square Degrees. AJ, 125:2064–2084, Apr 2003. doi: 10.1086/368240.

III Gallagher, John S. and Jeremiah P. Ostriker. A Note on Mass Loss during Collisions between Galaxies and the Formation of Giant Systems. AJ, 77:288, May 1972. doi: 10.1086/111280.

David G. Gilbank, M. D. Gladders, H. K. C. Yee, and B. C. Hsieh. The Red-sequence Cluster Survey-2 (RCS-2): Survey Details and Photometric Catalog Construction. AJ, 141:94, Mar 2011. doi: 10.1088/0004-6256/141/3/94.

Myriam Gitti, Fabrizio Brighenti, and Brian R. McNamara. Evidence for AGN Feedback in Galaxy Clusters and Groups. *Advances in Astronomy*, 2012: 950641, Jan 2012. doi: 10.1155/2012/950641.

Michael D. Gladders and H. K. C. Yee. A New Method For Galaxy Cluster Detection. I. The Algorithm. AJ, 120:2148–2162, Oct 2000. doi: 10.1086/301557.

Michael D. Gladders and H. K. C. Yee. The Red-Sequence Cluster Survey. I. The Survey and Cluster Catalogs for Patches RCS 0926+37 and RCS 1327+29. *The Astrophysical Journal Supplement Series*, 157:1–29, Mar 2005. doi: 10.1086/427327.

Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In Yee Whye Teh and Mike Titterington, editors, *Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics*, volume 9 of *Proceedings of Machine Learning Research*, pages 249–256, Chia Laguna Resort, Sardinia, Italy, 13–15 May 2010. PMLR. URL `http://proceedings.mlr.press/v9/glorot10a.html`.

J. E. Gunn and B. M. Tinsley. Dynamical friction: the Hubble diagram as a cosmological test. ApJ, 210:1–6, Nov 1976. doi: 10.1086/154797.

Alan H. Guth and So-Young Pi. Fluctuations in the new inflationary universe. *Phys. Rev. Lett.*, 49:1110–1113, Oct 1982. doi: 10.1103/PhysRevLett.49.1110. URL `https://link.aps.org/doi/10.1103/PhysRevLett.49.1110`.

Salman Habib, Adrian Pope, Hal Finkel, Nicholas Frontiere, Katrin Heitmann, David Daniel, Patricia Fasel, Vitali Morozov, George Zagaris, Tom Peterka, Venkatram Vishwanath, Zarija Lukić, Saba Sehrish, and Wei-keng Liao. HACC: Simulating sky surveys on state-of-the-art supercomputing architectures. New A, 42:49–65, Jan 2016. doi: 10.1016/j.newast.2015.06.003.

Eric J. Hallman and Tesla E. Jeltema. Structure and turbulence in simulated galaxy clusters and the implications for the formation of radio haloes. MNRAS, 418:2467–2480, Dec 2011. doi: 10.1111/j.1365-2966.2011.19637.x.

K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image Recognition. *ArXiv e-prints*, December 2015.

A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *ArXiv e-prints*, April 2017.

F. A. Jansen. XMM: advancing science with the high-throughput X-ray spectroscopy mission. *ESA Bulletin*, 100:9–12, December 1999.

C. Jones and W. Forman. Imaging the Hot Intra Cluster Medium. In A. C. Fabian, editor, *NATO Advanced Science Institutes (ASI) Series C*, volume 366, page 49, Jan 1992.

Rita Seung Jung Kim, Jeremy V. Kepner, Marc Postman, Michael A. Strauss, Neta A. Bahcall, James E. Gunn, Robert H. Lupton, James Annis, Robert C. Nichol, Francisco J. Castander, J. Brinkmann, Robert J. Brunner, Andrew Connolly, Istvan Csabai, Robert B. Hindsley, Željko Ivezić, Michael S. Vogeley, and Donald G. York. Detecting Clusters of Galaxies in the Sloan Digital Sky Survey. I. Monte Carlo Comparison of Cluster Detection Algorithms. AJ, 123: 20–36, Jan 2002. doi: 10.1086/324727.

Benjamin P. Koester, Timothy A. McKay, James Annis, Risa H. Wechsler, August E. Evrard, Eduardo Rozo, Lindsey Bleem, Erin S. Sheldon, and David Johnston. MaxBCG: A Red-Sequence Galaxy Cluster Finder. ApJ, 660: 221–238, May 2007. doi: 10.1086/512092.

Cedric Lacey and Shaun Cole. Merger rates in hierarchical models of galaxy formation. MNRAS, 262:627–649, Jun 1993. doi: 10.1093/mnras/262.3.627.

Susan M. Lea, Joseph Silk, E. Kellogg, and S. Murray. Thermal-Bremsstrahlung Interpretation of Cluster X-Ray Sources. ApJ, 184:L105, Sep 1973. doi: 10.1086/181300.

Yann LeCun, Patrick Haffner, Léon Bottou, and Yoshua Bengio. Object recognition with gradient-based learning. In *Shape, Contour and Grouping in Computer Vision*, pages 319–, London, UK, UK, 1999. Springer-Verlag. ISBN 3-540-66722-9.

Yuan Li, Greg L. Bryan, Mateusz Ruszkowski, G. Mark Voit, Brian W. O'Shea, and Megan Donahue. Cooling, AGN Feedback, and Star Formation in Simulated Cool-core Galaxy Clusters. ApJ, 811:73, Oct 2015. doi: 10.1088/0004-637X/811/2/73.

Christopher E. Lidman and Bruce A. Peterson. An Optically Based Search for Distant Galaxy Clusters. AJ, 112:2454, Dec 1996. doi: 10.1086/118195.

M. Lieu, L. Conversi, B. Altieri, and B. Carry. Detecting solar system objects with convolutional neural networks. *ArXiv e-prints*, July 2018.

S. L. Lumsden, R. C. Nichol, C. A. Collins, and L. Guzzo. The Edinburgh-Durham southern galaxy catalogue. IV. Thecluster catalogue. MNRAS, 258: 1–22, Sep 1992. doi: 10.1093/mnras/258.1.1.

S. J. Maddox, G. Efstathiou, W. J. Sutherland, and J. Loveday. Galaxy correlations on large scales. MNRAS, 242:43, Jan 1990. doi: 10.1093/mnras/242.1.43P.

W. G. Mathews and J. N. Bregman. Radiative accretion flow onto giant galaxies in clusters. ApJ, 224:308–319, Sep 1978. doi: 10.1086/156379.

T. A. Matthews, W. W. Morgan, and M. Schmidt. A Discussion of Galaxies Indentified with Radio Sources. ApJ, 140:35, July 1964. doi: 10.1086/147890.

Warren Mcculloch and Walter Pitts. A logical calculus of ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5:127–147, 1943.

B. R. McNamara and P. E. J. Nulsen. Heating Hot Atmospheres with Active Galactic Nuclei. *Annual Review of Astronomy and Astrophysics*, 45:117–175, Sep 2007. doi: 10.1146/annurev.astro.45.051806.110625.

B. R. McNamara and P. E. J. Nulsen. Mechanical feedback from active galactic nuclei in galaxies, groups and clusters. *New Journal of Physics*, 14:055023, May 2012. doi: 10.1088/1367-2630/14/5/055023.

Brian R. McNamara and Robert W. O'Connell. Star formation in cooling flows in clusters of galaxies. AJ, 98:2018–2043, Dec 1989. doi: 10.1086/115275.

Christopher J. Miller, Robert C. Nichol, Daniel Reichart, Risa H. Wechsler, August E. Evrard, James Annis, Timothy A. McKay, Neta A. Bahcall, Mariangela Bernardi, Hans Boehringer, Andrew J. Connolly, Tomotsugu Goto, Alexie Kniazev, Donald Lamb, Marc Postman, Donald P. Schneider, Ravi K. Sheth, and Wolfgang Voges. The C4 Clustering Algorithm: Clusters of Galaxies in the Sloan Digital Sky Survey. AJ, 130:968–1001, Sep 2005. doi: 10.1086/431357.

R. L. Minkowski and G. O. Abell. *The National Geographic Society-Palomar Observatory Sky Survey*, page 481. the University of Chicago Press, 1963.

R. J. Mitchell, J. L. Culhane, P. J. N. Davison, and J. C. Ives. Ariel 5 observations of the X-ray spectrum of the Perseus cluster. MNRAS, 175:29P–34P, May 1976. doi: 10.1093/mnras/175.1.29P.

Joseph J. Mohr, Daniel G. Fabricant, and Margaret J. Geller. An X-Ray Method for Detecting Substructure in Galaxy Clusters: Application to Perseus, A2256, Centaurus, Coma, and Sersic 40/6. ApJ, 413:492, Aug 1993. doi: 10.1086/173019.

Silvano Molendi and Fabio Pizzolato. Is the Gas in Cooling Flows Multiphase? ApJ, 560:194–200, Oct 2001. doi: 10.1086/322387.

W. W. Morgan. The Classification of Clusters of Galaxies. *Proceedings of the National Academy of Science*, 47:905–906, July 1961. doi: 10.1073/pnas.47.7.905.

Kaylea Nelson, Douglas H. Rudd, Laurie Shaw, and Daisuke Nagai. Evolution of the Merger-induced Hydrostatic Mass Bias in Galaxy Clusters. ApJ, 751:121, Jun 2012. doi: 10.1088/0004-637X/751/2/121.

A. Oemler, Jr. The Systematic Properties of Clusters of Galaxies. Photometry of 15 Clusters. ApJ, 194:1–20, November 1974. doi: 10.1086/153216.

L. F. Olsen, M. Scodeggio, L. da Costa, C. Benoist, E. Bertin, E. Deul, T. Erben, M. D. Guarnieri, R. Hook, M. Nonino, I. Prandoni, R. Slijkhuis, A. Wicenec, and R. Wichmann. ESO imaging survey. II. Searching for distant clusters of galaxies. A&A, 345:681–690, May 1999.

J. Oukbir and A. Blanchard. X-ray clusters in open universes. A&A, 262: L21–L24, Sep 1992.

F. Pacaud, M. Pierre, A. Refregier, A. Gueguen, J.-L. Starck, I. Valtchanov, A. M. Read, B. Altieri, L. Chiappetti, P. Gandhi, O. Garcet, E. Gosset, T. J. Ponman, and J. Surdej. The XMM Large-Scale Structure survey: the X-ray pipeline and survey selection function. MNRAS, 372:578–590, October 2006. doi: 10.1111/j.1365-2966.2006.10881.x.

P. J. E. Peebles. Large-scale background temperature and mass fluctuations due to scale-invariant primeval perturbations. ApJ, 263:L1–L5, December 1982. doi: 10.1086/183911.

A. A. Penzias and R. W. Wilson. A Measurement of Excess Antenna Temperature at 4080 Mc/s. ApJ, 142:419–421, Jul 1965. doi: 10.1086/148307.

J. R. Peterson, F. B. S. Paerels, J. S. Kaastra, M. Arnaud, T. H. Reiprich, A. C. Fabian, R. F. Mushotzky, J. G. Jernigan, and I. Sakelliou. X-ray imaging-spectroscopy of Abell 1835. A&A, 365:L104–L109, Jan 2001. doi: 10.1051/0004-6361:20000021.

Annalisa Pillepich, Volker Springel, Dylan Nelson, Shy Genel, Jill Naiman, Rüdiger Pakmor, Lars Hernquist, Paul Torrey, Mark Vogelsberger, Rainer Weinberger, and Federico Marinacci. Simulating galaxy formation with the IllustrisTNG model. MNRAS, 473:4077–4106, Jan 2018. doi: 10.1093/mnras/stx2656.

C. Pinto, C. J. Bambic, J. S. Sanders, A. C. Fabian, M. McDonald, H. R. Russell, H. Liu, and C. S. Reynolds. AGN feedback in the Phoenix cluster. MNRAS, 480:4113–4123, Nov 2018. doi: 10.1093/mnras/sty2185.

Marc Postman and Tod R. Lauer. Brightest cluster galaxies as standard candles. *ApJ*, 440, 03 1995. doi: 10.1086/175245.

Marc Postman, Lori M. Lubin, James E. Gunn, J. B. Oke, John G. Hoessel, Donald P. Schneider, and Jennifer A. Christensen. The Palomar Distant Clusters Survey. I. The Cluster Catalog. AJ, 111:615, Feb 1996. doi: 10.1086/117811.

LORIEN PRATT and BARBARA JENNINGS. A survey of transfer between connectionist networks. *Connection Science*, 8(2):163–184, 1996. doi: 10.1080/095400996116866. URL https://doi.org/10.1080/095400996116866.

M. Ramella, W. Boschin, D. Fadda, and M. Nonino. Finding galaxy clusters using Voronoi tessellations. A&A, 368:776–786, Mar 2001. doi: 10.1051/0004-6361:20010071.

D. O. Richstone. Collisions of galaxies in dense clusters. II. Dynamical evolution of cluster galaxies. ApJ, 204:642–648, Mar 1976. doi: 10.1086/154213.

Douglas O. Richstone. Collisions of Galaxies in Dense Clusters. I. Dynamics of Collisions of Two Galaxies. ApJ, 200:535–547, Sep 1975. doi: 10.1086/153820.

H. J. Rood and G. N. Sastry. "Tuning Fork" Classification of Rich Clusters of Galaxies. PASP, 83:313, June 1971. doi: 10.1086/129128.

F. Rostagni, C. Benoist, and S. Maurogordato. Morphological classification of galaxy clusters. In G. Alecian, K. Belkacem, R. Samadi, and D. Valls-Gabaud, editors, *SF2A-2011: Proceedings of the Annual meeting of the French Society of Astronomy and Astrophysics*, pages 185–189, Dec 2011.

J. S. Sanders, A. C. Fabian, S. W. Allen, R. G. Morris, J. Graham, and R. M. Johnstone. Cool X-ray emitting gas in the core of the Centaurus cluster of galaxies. MNRAS, 385:1186–1200, Apr 2008. doi: 10.1111/j.1365-2966.2008.12952.x.

P. J. Serlemitsos, B. W. Smith, E. A. Boldt, S. S. Holt, and J. H. Swank. X-radiation from clusters of galaxies: spectral evidence for a hot evolved gas. ApJ, 211:L63–L66, Jan 1977. doi: 10.1086/182342.

B. W. Silverman. *Density estimation for statistics and data analysis*. 1986.

K. Simonyan and A. Zisserman. Very Deep Convolutional Networks for Large-Scale Image Recognition. *ArXiv e-prints*, September 2014.

J.-L. Starck and M. Pierre. Structure detection in low intensity X-ray images. A&AS, 128:397–407, March 1998. doi: 10.1051/aas:1998150.

J.-L. Starck, F. D. Murtagh, and A. Bijaoui. *Image Processing and Data Analysis*. July 1998.

A.A. Starobinsky. Dynamics of phase transition in the new inflationary universe scenario and generation of perturbations. *Physics Letters B*, 117(3):175 – 178, 1982. ISSN 0370-2693. doi: https://doi.org/10.1016/0370-2693(82)90541-X. URL http://www.sciencedirect.com/science/article/pii/037026938290541X.

M. F. Struble and H. J. Rood. Morphological classification (revised RS) of Abell clusters in D = 5 and an analysis of observed correlations. AJ, 89:1487–1513, October 1984. doi: 10.1086/113650.

R. A. Sunyaev and Y. B. Zeldovich. The Observations of Relic Radiation as a Test of the Nature of X-Ray Radiation from the Clusters of Galaxies. *Comments on Astrophysics and Space Physics*, 4:173, November 1972.

C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. Going Deeper with Convolutions. *ArXiv e-prints*, September 2014.

C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the Inception Architecture for Computer Vision. *ArXiv e-prints*, December 2015.

Hrant M. Tovmassian and Heinz Andernach. On the formation of cD galaxies and their parent clusters. MNRAS, 427:2047–2056, Dec 2012. doi: 10.1111/j.1365-2966.2012.22044.x.

S. D. Tremaine, J. P. Ostriker, and Jr. Spitzer, L. The formation of the nuclei of galaxies. I. M31. ApJ, 196:407–411, Mar 1975. doi: 10.1086/153422.

I. Valtchanov, M. Pierre, and R. Gastaud. Comparison of source detection procedures for XMM-Newton images. A&A, 370:689–706, May 2001. doi: 10.1051/0004-6361:20010264.

A. Vikhlinin, A. V. Kravtsov, R. A. Burenin, H. Ebeling, W. R. Forman, A. Hornstrup, C. Jones, S. S. Murray, D. Nagai, H. Quintana, and A. Voevodkin. Chandra Cluster Cosmology Project III: Cosmological Parameter Constraints. ApJ, 692:1060–1074, Feb 2009. doi: 10.1088/0004-637X/692/2/1060.

S. D. M. White. The dynamics of rich clusters of galaxies. MNRAS, 177:717–733, December 1976. doi: 10.1093/mnras/177.3.717.

D. G. York, J. Adelman, J. E. Anderson, Jr., S. F. Anderson, and Annis et al. The Sloan Digital Sky Survey: Technical Summary. AJ, 120:1579–1587, September 2000. doi: 10.1086/301513.

F. Zwicky, E. Herzog, P. Wild, M. Karpowicz, and C. T. Kowal. *Catalogue of galaxies and of clusters of galaxies, Vol. I.* 1961.