YOLO-CL: Galaxy cluster detection in the SDSS with deep machine learning

Kirill Grishin1, Simona Mei1,2, and Stéphane Illić3,1,4,5

1 Université Paris Cité, CNRS/IN2P3, Astroparticule et Cosmologie, 10 rue Alice Domon et Léonie Duquet, 75013 Paris, France
e-mail: grishin@apc.in2p3.fr; mei@apc.in2p3.fr
2 Jet Propulsion Laboratory, Caltech Center for Astronomy & Astrophysics, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, California, USA
3 Université PSL, Observatoire de Paris, Sorbonne Université, CNRS, LERMA, 75 Avenue Denfert-Rochereau, 75014 Paris, France
4 CNES, Centre National d’Études Spatiales, 18 Av. Edouard Belin, Toulouse, France
5 IJCLab, Université Paris-Saclay, CNRS/IN2P3, IJCLab, rue André Ampère, Campus de la Faculté des Sciences, 94405 Orsay, France

Received 23 January 2023 / Accepted 10 May 2023

ABSTRACT

Galaxy clusters are powerful probes for cosmological models. Next-generation, large-scale optical and infrared surveys are poised to reach unprecedented depths and, thus, they require highly complete and pure cluster catalogs, with a well-defined selection function. We have developed a new cluster detection algorithm named YOLO for Cluster detection (YOLO-CL), which is a modified version of the state-of-the-art object detection deep convolutional network named You only look once (YOLO) that has been optimized for the detection of galaxy clusters. We trained YOLO-CL on the red-sequence Matched-filter Probabilistic Percolation (redMaPPer) cluster catalog, based on Sloan Digital Sky Survey (SDSS) color images. We find that YOLO-CL detects 95–98% of the redMaPPer clusters, with a purity of 95–98%, that is calculated by applying the network to SDSS blank fields. When compared to the Meta-Catalog of X-Ray Detected Clusters of Galaxies 2021 (MCXC2021) X-ray catalog in the SDSS footprint, YOLO-CL recovers all clusters at \( L_X \gtrsim 2 \times 10^{44} \text{ erg s}^{-1} \), \( M_{500} \gtrsim 2 \times 10^{14} \text{ M}_\odot \), \( R_{500} \gtrsim 0.75 \)–0.8 Mpc and \( 0.4 \leq z \leq 0.6 \). When compared to the redMaPPer detection of the same MCXC2021 clusters, we find that YOLO-CL is more complete than redMaPPer, which means that the neural network has indeed improved the cluster detection efficiency of its training sample. In fact, YOLO-CL detects \( \gtrsim 98\% \) of the MCXC2021 clusters with an X-ray surface brightness of \( I_{500} \gtrsim 2 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ s}^{-1} \text{ cm}^{-2} \text{ arcmin}^{-2} \) at \( 0.2 \leq z \leq 0.6 \) and \( \sim 100\% \) of the MCXC2021 clusters with \( I_{500} \gtrsim 3 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcmin}^{-2} \) at \( 0.3 \leq z \leq 0.6 \). While redMaPPer detects \( \sim 98\% \) of the MCXC2021 clusters with \( I_{500} \gtrsim 55 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcmin}^{-2} \) at \( 0.2 \leq z \leq 0.6 \) and \( \sim 100\% \) of the MCXC2021 clusters with \( I_{500} \gtrsim 2 \times 10^{-15} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcmin}^{-2} \) at \( 0.5 \leq z \leq 0.6 \). The YOLO-CL selection function is approximately constant with redshift, with respect to the MCXC2021 cluster X-ray surface brightness. YOLO-CL exhibits a high level of performance when compared to traditional detection algorithms applied to SDSS. Deep learning networks display a strong advantage over traditional galaxy cluster detection techniques because they do not require the galaxy’s photometric and photometric redshift catalogs. This eliminates systematic uncertainties that may be introduced during source detections and photometry, as well as photometric redshift measurements. Our results show that YOLO-CL is an efficient alternative to traditional cluster detection methods. In general, this work shows that it is worth exploring the performance of deep convolutional networks for future cosmological cluster surveys, such as the Rubin/Legacy Survey of Space and Time (Rubin/LST), Euclid, and Roman Space Telescope surveys.

Key words. large-scale structure of Universe -- galaxies: clusters: general -- catalogs

1. Introduction

Clusters of galaxies are powerful probes for constraining cosmological models. In fact, since they are the largest and most massive gravitationally bound systems in the Universe, their abundances probe the growth history of structures (e.g., Allen et al. 2011). Future large-scale surveys, such as the Dark Energy Survey1 (Abbott et al. 2018), the Dark Energy Spectroscopic Instrument2 (Dey et al. 2019), the Vera C. Rubin Observatory3 (formerly Large Synoptic Survey Telescope, Kahn 2018), the Euclid satellite4 (Laureijs et al. 2011), and the Nancy Grace Roman Space Telescope (Eifler et al. 2021), will use large cluster samples as cosmological probes. Thus, they will require the development of fast and efficient cluster detection algorithms. These surveys will reach unprecedented depths and will require highly complete and pure cluster catalogs, with a well-defined selection function.

Cluster detection algorithms have been developed by the astronomical community at different wavelengths. The detection in optical and near-infrared bandpasses is mainly based on the search of spatial overdensities of a given class (quiescent, line-emitters, massive, etc.) of galaxies (e.g., Gladders & Yee 2005; Knobel et al. 2009; Muzzin et al. 2012; Sobral et al. 2010; Bayliss et al. 2011; Rykoff et al. 2014; Wylezalek et al. 2013, 2014), while detections in the X-rays and submillimeter rely on the assumption of model profiles that fit the data, for instance, a characteristic galaxy cluster luminosity and a radial profile (e.g., Olsen et al. 2007; Grove et al. 2009; Planck Collaboration XXVII 2016; Böhringer et al. 2004; Marriage et al. 2011; Predehl et al. 2021).

---

1 https://www.darkenergiesurvey.org/
2 https://www.desi.lbl.gov/
3 https://www.vro.org/
4 https://www.euclid-ec.org/
In the local Universe and up to \( z \sim 1 \), the reference cluster catalog has been the all-sky ROSAT (Röntgensatellit) X-ray catalog (Ebeling et al. 1998; Voges et al. 1999; Böhringer et al. 2004) over recent decades. However, over the past ten years, optical and millimeter-wave surveys provided large cluster samples that have been used to constrain our cosmological model parameters (Rozo et al. 2010; Allen et al. 2011; Hasselfield et al. 2013; Bleem et al. 2015; Planck Collaboration Int. XXVI 2015; de Haan et al. 2016; Costanzi et al. 2021; Chiu et al. 2023). Cluster surveys from the X-ray survey eRosita (extended Röntgen Survey with an Imaging Telescope Array; Merloni et al. 2012), the next-generation cosmic microwave background (CMB) surveys (e.g., Simons Observatory Ade et al. 2019), and Euclid (e.g.,Ascaso et al. 2015; Euclid Collaboration 2019), as well as the Nancy Grace Roman Space Telescope (Eifler et al. 2021) and the Rubin Legacy Survey Legacy Survey of Space and Time (LSST; e.g., Ivezic et al. 2019) will extend these cluster samples at lower mass and higher redshifts. In the high redshift Universe \(( z \gtrsim 1.5 \)) , cluster detection will be mainly performed by optical and infrared surveys, combined with radio and far-infrared observations. At these epochs, clusters have been predicted to be less massive (Chiang et al. 2013). This means that the X-ray and the SZ (Sunyaev-Zel’dovich) effect signals are fainter (e.g., Ascaso et al. 2017), limiting the performance of X-ray and SZ cluster detection algorithms.

The use of deep machine learning (ML) algorithms in various areas of astrophysics has been rising for the past decade, with applications ranging from the analysis of galaxy surveys (see, e.g., Huertas-Company & Lanusse 2023, for a recent review) and photometric redshift estimations (see, e.g., Henghes et al. 2019, for a recent review) to dark matter map reconstructions (e.g., Jeffrey et al. 2020). In particular, convolutional neural networks (CNN) have proved especially useful in object detection and characterization (e.g., Huertas-Company et al. 2015, 2018; Dimauro et al. 2018; Pasquet et al. 2019; Zanisi et al. 2021; Euclid Collaboration 2022, 2023a,b; Davidzon et al. 2022), as well as galaxy cluster detections (e.g., Chan & Stott 2019; Bonjean 2020; Hurier et al. 2021; Lin et al. 2021).

Many object detection algorithm have been developed in the field of deep ML (see the recent reviews of Zhou et al. 2019 and Zaidi et al. 2021), most of which have not been applied in the field of astrophysics. In this paper, we use as a basis the architecture of the well-known detection-oriented deep machine learning neural network “You only look once” (YOLO, Redmon et al. 2015; Redmon & Farhadi 2016) to detect clusters of galaxies in the Sloan Digital Sky Survey (SDSS) and assess its efficiency. The YOLO algorithm has been developed for a very wide range of real-life situations, for example, for face detection, the analysis of medical images, and self-driving cars. Its last implementations, among which YOLOv3 developed by Redmon & Farhadi (2018), are particularly efficient for multiple object detection and well-adapted to cluster detection.

Our results show that our YOLO network, which we called YOLO for C Luster detection (YOLO–CL), adapted for the detection of galaxy clusters, shows a high performance with respect to traditional cluster detection algorithms in obtaining dependable cluster catalogs with high levels of completeness and purity. Our results show that our YOLO–CL cluster catalogs have a purity of 95–98% on blank fields and a completeness of \( \sim 98\% \) for X-ray detected clusters with \( L_X,500 \gtrsim 2 \times 10^{15} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcmin}^{-2} \) at 0.2 \( \leq z \leq 0.6 \), and of \( \sim 100\% \) for clusters with \( L_X,500 \gtrsim 30 \times 10^{15} \text{ erg s}^{-1} \text{ cm}^{-2} \text{ arcmin}^{-2} \) at 0.3 \( \leq z \leq 0.6 \). Our selection function is flat as a function of redshift, when considering X-ray mean surface brightness.

In Sect. 2, we describe the data and the catalog used for the training and validation of our network. Section 3 presents our network implementation and how we build our cluster catalog. In Sect. 4, we compare our results with the training cluster catalog and a X-ray cluster catalog. We discuss and summarize our results in Sects. 5 and 6, respectively.

2. Observational dataset

For the past two decades, the SDSS has been the largest imaging and spectroscopic survey of the local Universe (York et al. 2000). It uses a dedicated 2.5-m wide-field-of-view optical telescope, located at the Apache Point Observatory, and has provided astronomers with a tremendous amount of data. This wealth of data has consistently yielded cosmological constraints via the various SDSS Data Releases (DR), so far culminating in the 17th data release (DR17, Abdurro’uf et al. 2022).

To train and test the application of YOLO to cluster detection, we focus on the most complete and pure SDSS cluster catalog (see also Sect. 5), namely, the red-sequence Matched-filter Probabilistic Percolation (redMaPPer) DR8 (Data Release 8) catalog from Rykoff et al. (2014). The redMaPPer algorithm is a red sequence cluster finder specifically designed for large photometric surveys. The redMaPPer algorithm was applied to the \( \sim 10000 \) square degrees of the SDSS DR8 data release, yielding a catalog \( ^7 \) of 26 111 clusters over the redshift range \( z \in [0.08, 0.55] \). With respect to the Meta-Catalog of X-Ray Detected Clusters of Galaxies X-ray detection catalog (MCXC, Piffaretti et al. 2011), the redMaPPer catalog was found to be 100% complete up to \( z \leq 0.35 \), above the X-ray temperature \( T_X \geq 3.5 \text{keV} \), and \( L_X \gtrsim 2 \times 10^{44} \text{ erg s}^{-1} \), decreasing to 90% completeness at \( L_X \leq 10^{43} \text{ erg s}^{-1} \). Overall, 86% of the redMaPPer clusters are correctly centered with respect to their X-ray centers (Rykoff et al. 2014). All redMaPPer rich clusters \( (l > 100) \) are detected in the X-ray ROSAT All Sky Survey (Voges et al. 1999).

To train and validate our network, we excluded clusters with redshifts \( z < 0.2 \) that cover regions in the sky that are larger than the cutout images that we consider in this work because of calculation time efficiency. Ultimately, we worked with a final sample of 24 406 clusters, whose distribution is shown in Fig. 1. For each cluster, the algorithm provides its position, the richness \( X \) as a proxy for cluster mass and a list of cluster members (Rozo & Rykoff 2014).

For the network training and validation, we retrieved JPEG versions of the original SDSS DR16 raw images for each redMaPPer cluster, and constructed color images using the ImgaCutout web service \(^8\) by querying the SDSS Catalog Archive Server databases. These images are based on the \( g, r, i \)-band FITS corrected frame files from the Science Archive Server, and the color images were built using the conversion algorithm \(^9\) based on Lupton et al. (2004). These three bandpasses are

\[ 5 \] https://lsst.slac.stanford.edu/
\[ 6 \] https://classic.sdss.org/
\[ 7 \] Version 6.3 of the catalog, from the VizieR archive: https://vizier.cds.unistra.fr
\[ 8 \] The cluster richness is defined as the number of cluster members above a given luminosity. For redMaPPer, it is defined as a sum of the probability of being a cluster member over all galaxies in a cluster field (Rozo et al. 2009).
\[ 9 \] For the SDSS Data Release 16, help docs on Image service are available at the following link: http://skyserver.sdss.org/dr16/en/help/docs/api.aspx#imgcutout
\[ 10 \] Detailed here: https://www.sdss.org/dr16/imaging/jpg-images-on-skyserver
Fig. 1. redMaPPer sample of 24406 clusters used to train and validate our network. Top: sky map of the positions of the redMaPPer clusters in celestial coordinates, where the color indicates the photometric redshift of the cluster as estimated by the redMaPPer algorithm. Bottom: training and validation redMaPPer sample redshift (left) and richness (right) distribution.

sufficient to identify passive early-type galaxies in clusters at $z < 1$. Figure 2 shows an example of such cut-out images.

3. YOLO–CL: our YOLO network for galaxy cluster detections

3.1. The YOLO network

The state-of-the-art deep convolutional network You only look once (YOLO, Redmon et al. 2015) offers real-time object detection. Competing architectures in the ML literature tend to apply first a “localizer” network on a given image, at multiple locations and on multiple scales, and assign a detection probability. The high-probability regions of the image are considered as detections, and are then classified using a separate network. The YOLO architecture uses a different approach: it applies a single neural network to the full image, combining the detection and classification into a single process. This gives the network several advantages over classifier-based systems, because its predictions take into account the global context of the image. It also has the advantage of making predictions with a single network evaluation, unlike systems such as Region Based Convolutional Neural Networks (R-CNN, Girshick et al. 2013, and following iterations Fast and Faster R-CNN), which require thousands of evaluations for a single image. This may result in several orders of magnitude faster YOLO execution times, compared to R-CNN and Fast R-CNN.

In practice, the network divides the image into a $S \times S$ grid of regions (or cells), within which the detection and

Fig. 2. SDSS image cutout of a redMaPPer cluster in our sample. The yellow box corresponds to the minimal rectangle encompassing all redMaPPer cluster members, which is the box used to train YOLO–CL. In cyan, the box detected by our network YOLO–CL, with the associated confidence level in the top left corner. The image size is $13.5 \times 13.5$ arcmin$^2$ and the pixel size is 0.396 arcsec.
classification are performed. YOLO predicts $B$ bounding boxes per region, with their associated “objectness” probability (i.e., how confident we are that there is an object in the box) and “class probabilities” (i.e., for a set of classes, the respective probabilities that the potential object belongs to them). Both $B$ and $S$ are hyperparameters of the network and can be adjusted by the user.

The predicted bounding boxes and their associated probabilities are returned by the network in the following format:

$$(x, y, w, h, C, p(c_1), \ldots, p(c_n)),$$

where $(x, y)$ are the coordinates of the box center, $w$ and $h$ its width and height, $C$ the objectness, and $p(c_1), \ldots, p(c_n)$ are the probabilities (summing to one) that the object in the box belongs respectively to the class $c_1, \ldots, c_n$.

When training a YOLO network using a set of images (with their associated “true” bounding boxes), we optimize the following multi-part loss function $L$ (Redmon et al. 2015):

$$L = L_{\text{bbox}} + L_{\text{obj}} + L_{\text{class}}.$$  

(2)

The first term of Eq. (2) is the “bounding box loss” as follows:

$$L_{\text{bbox}} = \alpha_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{B} \sum_{i=0}^{\text{obj}_{ij}} \left[ (x_i - \hat{x}_i)^2 + (y_i - \hat{y}_i)^2 \right] + \alpha_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^{\text{obj}_{ij}} \left[ (\sqrt{w_i} - \sqrt{\hat{w}_i})^2 + (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \right],$$  

(3)

where the $(x, y)$ coordinates represent the center of the box relative to the bounds of the grid cell and $w$ and $h$ are the width and height of the box. The symbol $\text{obj}_{ij}$ denotes if an object appears in cell $i$ and $\text{obj}_{ij}$ denotes that the $j$th bounding box predictor in cell $i$ is “responsible” for that prediction. In these equations (and those below), the variables with a hat over them are the “true values” that the network is learning.

The second term is the “objectness loss” as follows:

$$L_{\text{obj}} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} \sum_{i=0}^{\text{obj}_{ij}} (C_i - \hat{C}_i)^2 + \alpha_{\text{noobj}} \sum_{i=0}^{S^2} \sum_{j=0}^{\text{noobj}_{ij}} (C_i - \hat{C}_i)^2,$$

(4)

where $C$ represents the conditional class probability. Finally, the last term represents the “classification loss” as follows:

$$L_{\text{class}} = \sum_{i=0}^{S^2} \sum_{j=0}^{B} \sum_{c=0}^{\text{classes}} (p(c) - \hat{p}(c))^2,$$

(5)

where the $p(c)$ correspond to the probabilities to belong to a certain class, $i$. The $\alpha_{\text{coord}}$ and $\alpha_{\text{noobj}}$ coefficients appearing in the previous formulas can be changed to give more weight to certain components of the total loss. We chose to set $\alpha_{\text{coord}} = \alpha_{\text{noobj}} = 1$.

The first version of the YOLO network used a Darknet-19 neural network architecture, which contains 19 layers, as the feature extractor. Its second version (YOLO9000, Redmon & Farhadi 2016) added 11 more layers to Darknet-19, reaching a total of 30. These first architectures had difficulties to detect small objects due to the coarseness of their $S \times S$ grid.

We base our cluster detection network on the third iteration of YOLO (Redmon & Farhadi 2018), YOLOv3, which represents a significant improvement over the first two versions, and while several other YOLO versions were developed afterwards, we will consider their application only in future work. The feature extractor was replaced by Darknet-53 and residual networks. The new extractor uses 53 convolution layers, with consecutive $3 \times 3$ and $1 \times 1$ convolution layers followed by a skip connection (introduced by ResNet, He et al. 2015) to help the activations to propagate through deeper layers without gradient diminishing. With 53 additional layers for the detection, YOLO totals 106 fully convolutional layers. Its larger size makes it slower as compared to previous iterations, but significantly enhances its accuracy.

Moreover, YOLOv3 introduces a multi-scale feature in the detection process. In practice, instead of producing a single “feature map” (to be fed to the detection part of the network) at a single $S \times S$ resolution, Darknet-53 produces three different maps at three different levels of resolution. The underlying idea here is to provide the detection network with feature maps at three different scales: the coarser the map, the bigger the objects that it will detect.

YOLO networks have been used in applications in astrophysics to detect galaxies and other sources (González et al. 2018; He et al. 2021), and astrophysical transients (Li et al. 2022).

3.2. Network optimization for galaxy cluster detection

3.2.1. Modifications to the original YOLOv3 network

To optimize YOLOv3 for galaxy cluster detection, we applied several modifications to its standard architecture, and we call our new implementation YOLO-CL (YOLO for CLuster detection). YOLO–CL is based on a TensorFlow implementation of the YOLO architecture. The network was trained on a NVIDIA Tesla V100-SXM2-32GB GPU, equipped with 32 GB of memory.

Our first modification is the definition of a single (instead of multiple classes in the original network) object class, clusters, and we therefore removed the $L_{\text{class}}$ term, defined in Eq. (5), from the original loss function of Eq. (2). This results in fewer network weights, which leads to a lighter, faster, and easier-to-train network.

Then, we replaced the standard YOLO bounding box loss of Eq. (3) with the so-called generalized Intersection over Union (gIoU) loss of Reza et al. (2019). In fact, the traditional IoU metric has among its main weaknesses the fact that it has a plateau (equal to 0) when the true and predicted bounding boxes are non-overlapping, making it impossible to optimize the corresponding loss term because of the vanishing gradient. The gIoU addresses this weakness by amending the IoU as follows:

$$\text{gIoU} = \text{IoU} + \frac{U}{A_i} - 1.$$  

(6)

The IoU the ratio of the area of the intersection area to the area of their union. A value of 1 corresponds to a perfect agreement, while a value that tends towards 0 indicates increasingly
disjointed boxes and/or significantly different sizes. $\mathcal{U}$ and $\mathcal{A}_k$ are respectively the areas of the union of the two boxes and the smallest box enclosing both boxes. This allows the glIoU to extend smoothly into negative values for boxes that are disjointed, and to tend towards $-1$ for more and more distant boxes (as $\text{IoU} = 0$ and $\mathcal{U} \ll \mathcal{A}_k$). The glIoU was shown to yield better performance for multiple object detection compared to the standard bounding box metrics.

3.2.2. Hyperparameter optimization

Our hyperparameters were tested for best performance in optimizing the completeness and purity of our final cluster catalog. Completeness and purity are the two parameters that characterize a galaxy cluster catalog and the algorithm selection function as a function of the cluster redshift and physical properties (such as mass, richness, X-ray luminosity, etc.). The catalog completeness quantifies the fraction of true clusters that are detected by the algorithm. The catalog purity quantifies the fraction of detected clusters that are true, instead of false detections. In machine learning literature, the completeness corresponds to the recall, and the purity to the precision.

The dimension of the first network layer sets an upper limit for the amount of information that is used by the network, while also defining the size and complexity of the architecture. We start with SDSS images with size $2048 \times 2048$ pixels, which corresponds to $\sim 13.5 \times 13.5$ arcmin$^2$ images and to twice a typical cluster virial radius of 1 Mpc at $z = 0.3$, the SDSS average redshift. We resized each image to the dimensions of this network first layer by average pooling. In order to explore a trade-off between performance and precision, we consider two different input layer sizes, namely, $1024 \times 1024$ and $512 \times 512$ pixels. We keep the same stride parameters as in the original YOLOv3 publication, namely $8, 16, \text{and } 32$.

Multiple detections of the same object are discarded by applying a glIoU threshold of 0.5, which is similar to the Intersection-Over-Union (IoU) threshold in the original YOLO and YOLOv3 publications (Redmon et al. 2015; Redmon & Farhadi 2018). The IoU and glIoU measure the region of overlap of the bounding boxes that define two different detections (see their definition in Sect. 3.2.1). The choice of a threshold of 0.5 means that when two bounding boxes overlap more than 50%, we consider that they define the same object. If they do, we keep the highest probability detection and discard the other. If the glIoU is smaller than 50% we consider that the two objects are different. Table 1 shows the settings used for the network training.

3.3. Training and validation

We trained YOLO–CL with about half ($\sim 12,000$) of our selected redMaPPer cluster images and the same number of random SDSS field images of the same size. The training sample is split into subsets (batches) that are simultaneously processed by the network. Our original SDSS color image cutouts are centered on the redMaPPer cluster positions. This original centering does not have an impact on cluster detection, which does not depend on the position of the cluster in the image. In fact, when we train the network, we apply data augmentation to the original cutouts, including translation, flipping, and cropping, which change the initial cluster position in the image. This prevents the network from learning that an object at the center of the image belongs to the “cluster” class and, instead, forces it to focus on the relevant features associated with clusters.

Our validation sample consists of the remaining cluster images and an equivalent number of random field images. We calculate the validation loss at the end of each training epoch. We started by setting a learning rate of $10^{-4}$, which grows slowly to $10^{-3}$ during the four warm-up epochs and then slowly decreases to $10^{-6}$.

Figure 3 shows the loss function for the training and validation sets for initial input images of size $512 \times 512$ pixels and $1024 \times 1024$ pixels. In both cases, there is a good agreement between the training and the validation loss function, excluding significant overfitting and confirming our network stability.

At each epoch, the network output is a catalog of detection positions, bounding boxes, and their probability of being a cluster (hereafter, the detection probability). To build a sample of detected cluster candidates, we had to apply a probability cut by choosing only detections with the probability of belonging to the class “cluster” (see Eq. (5)) that are higher than a given threshold. To define which epoch and which detection probability threshold to apply for our final cluster sample, we evaluated the performance of our network with respect to our YOLO–CL catalog completeness and purity for each epoch and then chose the epoch in which we obtain the largest completeness and purity. In this optimization, the redMaPPer cluster catalog is the true one.

Figure 4 shows the completeness and purity of the detected cluster candidate sample as a function of a given detection threshold when using the $512 \times 512$ and the $1024 \times 1024$ resampled SDSS images, for each in the epoch for which we obtain the highest completeness and purity. This leads to the selection of different cluster samples for different resolutions. With the $512 \times 512$ resolution, we miss 982 clusters found with the $1024 \times 1024$ ($\sim 8\%$ of the validation sample) and with the $1024 \times 1024$ resolution we miss 280 clusters found with the $512 \times 512$ resolution ($\sim 2\%$ of the validation sample). In both cases, we build our final detection catalog by choosing the threshold that optimizes both completeness and purity at the same time and that corresponds to the intersection of the completeness and purity curves in the figure. When using the $512 \times 512$ and the $1024 \times 1024$ resampled SDSS images, we obtain a completeness and purity of 95 and 98%, respectively, at the optimal threshold of $93$ and $60\%$, respectively. Our YOLO–CL final detection catalog (hereafter, the cluster catalog, with the caveat that the final detections are not confirmed galaxy clusters but cluster candidates with a given probability of being a cluster) includes the detections obtained when applying the above optimal thresholds.

<table>
<thead>
<tr>
<th>Image resolution</th>
<th>Batch size</th>
<th>Number of training epochs</th>
<th>Data augmentation technique</th>
<th>Augmentation frequency per technique</th>
<th>glIoU threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>1024 × 1024</td>
<td>2</td>
<td>100</td>
<td>horizontal flip, vertical flip, transpose</td>
<td>50%</td>
<td>50%</td>
</tr>
<tr>
<td>512 × 512</td>
<td>8</td>
<td>100</td>
<td>horizontal flip, vertical flip, transpose</td>
<td>50%</td>
<td>50%</td>
</tr>
</tbody>
</table>

Table 1. Settings used for the YOLO–CL training.
4. Comparison to redMaPPer and X-ray cluster detections

In this section, we compare our YOLO-CL cluster catalog with our original redMaPPer catalog and X-ray cluster detections in the SDSS footprint.

4.1. Comparison to redMaPPer detections

Figures 5 and 6 show the fraction of redMaPPer clusters detected by YOLO-CL as a function of redshift and richness (i.e., the YOLO-CL completeness with respect to redMaPPer), respectively, and when using the 512 × 512 and the 1024 × 1024 resampled SDSS images. When using 512 × 512 resampled SDSS images, YOLO-CL detects ∼98% of the redMaPPer clusters at z ≥ 0.3 and λ ≥ 40 (corresponding to $M_{200} \sim 3 \times 10^{14} M_\odot$ from Simet et al. 2017). For lower redshift (0.2 < z ≤ 0.3) and richness (20 < λ ≤ 40), the completeness is still very high at ∼92–93%. When using 1024 × 1024 resampled SDSS images, the performance of the network is similar to that of a function of richness, with a completeness of ∼98%; however, the completeness as a function of redshift is different. At z ≤ 0.5, YOLO-CL detects > 98% of the redMaPPer clusters, and at higher redshift the completeness drops to ∼92%. This is synthesized in Fig. 7, which shows the YOLO-CL sample completeness as a function of both redshift and richness.

Figure 8 shows the angular distance between the YOLO-CL cluster centers and redMaPPer cluster centers. The median angular distance between our and redMaPPer’s cluster centers is of ~5.6 ± 2.9 kpc, which is a very accurate recovery of redMaPPer cluster centers.

When checking if undetected clusters were found within detected cluster bounding boxes, we found that only ~1% of the redMaPPer cluster detections are not detected by YOLO-CL because they lie within the bounding box of another YOLO-CL detections (i.e., because they are superposed on the line of sight). This outlines the efficiency of the network in object separation. The remaining 99% are detections either at $\lambda \leq 40$ or $z \geq 0.5$, respectively, when using the 512 × 512 (left), and 1024 × 1024 (right) resampled SDSS images. When using the 512 × 512 and the 1024 × 1024 resampled SDSS images, we obtain a completeness and purity of 95 and 98%, respectively, at the optimal (see text) threshold of 93 and 60%, respectively. Overall, YOLO-CL has a very good performance in the redshift range covered by redMaPPer.
Fig. 5. YOLO–CL completeness with respect to the redMaPPer cluster catalog as a function of redshift, when using the $512 \times 512$ (left) and the $1024 \times 1024$ (right) resampled SDSS images. When using $512 \times 512$ and $1024 \times 1024$ resampled SDSS images, YOLO–CL reaches a completeness of $\gtrsim 98\%$ for $z \gtrsim 0.3$ and $z \lesssim 0.4$, respectively. In the other redshift ranges, the completeness is of $\sim 92$–94%. Overall, YOLO–CL has a very good performance in the redshift range covered by redMaPPer.

Fig. 6. YOLO–CL completeness with respect to the redMaPPer cluster catalog as a function of the redMaPPer richness, when using the $512 \times 512$ (left) and the $1024 \times 1024$ (right) resampled SDSS images. When using $512 \times 512$ and $1024 \times 1024$ resampled SDSS images, YOLO–CL reaches a completeness of $\sim 98\%$ for $\lambda \gtrsim 40$. At lower richness the completeness is of $\sim 92$–94%. Overall, YOLO–CL has a very good performance in the richness range covered by redMaPPer.

Fig. 7. YOLO–CL cluster sample completeness for redMaPPer detections for the images resampled to $512 \times 512$ (left) and to $1024 \times 1024$ (right) as a function of both richness and redshift. This figure synthesizes the conclusions of Figs. 4 and 5 as a function of both variables. On the right of each figure is the completeness scale.
in the regime where redMaPPer detections are less complete because of the SDSS depth (Rozo & Rykoff 2014).

Concerning the false positive detections in the random blank fields (2–5% of the detections), a visual inspection shows galaxy groups, in some cases, large and crowded star fields, as well as, in one case: a globular cluster. Since the redMaPPer catalog completeness decreases rapidly for low redMaPPer richness and low X-ray temperature and luminosity, as well as for clusters at $z \gtrsim 0.4$ (Rozo & Rykoff 2014), some of our false positive detections might be clusters or groups with low masses that are not detected by redMaPPer. This is further explored in the next section.

4.2. Comparison to the MCXC2021 X-ray catalog

To validate the performance of YOLO-CL on a cluster sample independent of redMaPPer, we applied our network to galaxy clusters detected by X-ray emission and published in the MCXC2021 catalog, the updated version of the MCXC catalog (Piffaretti et al. 2011). X-ray detections confirm cluster detections as virialized dark matter haloes by their hot gas emission. We expect that this is not a mass selected sample because X-ray selected samples are biased towards relaxed cool core clusters (e.g., Rossetti et al. 2016).

We built SDSS $g$, $r$, and $i$-band color images for all MCXC2021 clusters in the SDSS footprint (927/1841 clusters) by generating 2048x2048 pixel (13.5 $\times$ 13.5 arcmin) cutouts with SkyServer14, following the same procedure as for redMaPPer clusters. From this sample, we excluded the clusters that are only partially covered by the SDSS footprint. As a comparison, we show the efficiency of redMaPPer on the same cluster sample. To minimize selection bias, for this comparison we use the complete redMaPPer catalog (i.e., without any cut in richness; E. Rykoff & E. Rozo, priv. comm., and Rykoff et al. 2016). We crossmatched the MCXC2021 and redMaPPer catalogs within a range of 0.1 in redshift and a radius of 5 arcmin in position.

Figure 9 shows the YOLO-CL and redMaPPer cluster catalog completeness in redshift bins with respect to the MCXC2021 catalog as a function of X-ray luminosity, $L_X$, derived cluster mass, $M_{500}$, and radius, $R_{500}$. YOLO-CL recovers all clusters at $L_X \gtrsim 1 \times 10^{44}$ erg s$^{-1}$, $M_{500} \gtrsim 2 \times 10^{14} M_\odot$, $R_{500} \gtrsim 0.75$–0.8 Mpc and $z \gtrsim 0.4$. At lower luminosity, mass, radius, and redshift, its performance worsens. redMaPPer recovers all clusters at $L_X \gtrsim 3 \times 9 \times 10^{44}$ erg s$^{-1}$, $M_{500} \gtrsim 2 \times 6 \times 10^{14} M_\odot$, $R_{500} \gtrsim 0.8$–1.2 Mpc and $z \gtrsim 0.4$.

To better compare YOLO-CL to redMaPPer, Fig. 10 shows the MCXC2021 cluster detection completeness as a function of surface brightness $I_{X,500}$16 and redshift. The X-ray surface brightness quantifies the X-ray luminosity in a given area, and depends on the cluster luminosity and compactness, combining the information from the cluster $L_X$ and $R_{500}$ shown in Fig. 9.

Our YOLO-CL network detects $\sim$98% of the MCXC2021 clusters with $I_{X,500} \gtrsim 2 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$ at $0.2 \lesssim z \lesssim 0.6$; $\sim$100% of the MCXC2021 clusters with $I_{X,500} \gtrsim 3 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$; $z \gtrsim 0.3$. redMaPPer detects $\sim$98% of the MCXC2021 clusters with $I_{X,500} \gtrsim 5 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$ at $0.2 \lesssim z \lesssim 0.6$; and $\sim$100% of the MCXC2021 clusters with $I_{X,500} \gtrsim 2 \times 10^{-15}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$ at $0.5 \lesssim z \lesssim 0.6$.

Figure 8 shows the angular distance between YOLO-CL and MCXC2021 cluster centers. The median angular distance is of 261 ± 327 kpc. This large dispersion is consistent with the precision on the position of most of the MCXC2021 clusters that were detected by ROSAT with a 2.3 arcmin angular resolution, which corresponds to $\sim$600 kpc at the median MCXC2021 cluster redshift of $z \sim 0.3$, for our sample within the SDSS footprint.

From this comparison, YOLO-CL is more efficient than redMaPPer in detecting MCXC2021 clusters, which means that when optimizing YOLO-CL in terms of completeness and purity we improved cluster detection with respect to our training sample. This means that the network recovers cluster features that redMaPPer does not recover in clusters with lower X-ray surface brightness. It is also interesting that the YOLO-CL selection function is approximately constant with redshift, with respect to the X-ray surface brightness.

5. Discussion and conclusions

Our deep convolutional network YOLO-CL shows high completeness and purity in detecting galaxy clusters in the SDSS

---

14 https://www.galaxyclusterdb.eu/m2c/
15 http://skyserver.sdss.org/dr16/en/help/docs/api.aspx#imgcutout
16 $M_{500}$ is defined as the mass within the circular region of radius $R_{500}$ containing a mean mass density equal to five hundred times the critical density of the Universe at a given redshift. The luminosity $L_X$ is the luminosity $L_{500}$ in the same region.
17 Defined as a mean X-ray flux within the region containing a mass density equal to five hundred times the critical density of the Universe at a given redshift divided by its angular circular area.
footprint. When compared to the existing redMaPPer catalog, we obtained cluster catalogs with a completeness and purity of 95–98% for our optimal thresholds when using $512 \times 512$ and the $1024 \times 1024$ resampled SDSS images. The X-ray parameter that we found more interesting is the X-Ray surface brightness, which defines a clear threshold after which YOLO-CL and redMaPPer are 100% complete. When compared to the MCXC2021 X-ray detected clusters, YOLO-CL detects ∼98% of the MCXC2021 clusters with $L_x, 0.0 \geq 20 \times 10^{44}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$ at $0.2 \leq z \leq 0.6$ and ∼100% of the MCXC2021 clusters with $L_x, 0.0 \geq 30 \times 10^{41}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$ and $z \geq 0.3$. The lower detection rates for clusters at lower redshift could be explained by their large angular size that exceeds the size of image cutouts.

Several other SDSS cluster catalogs have been published, using different methods: the MaxBCG (Koester et al. 2007), WHL09/12 (Wen et al. 2009, 2012), GMBCG (Hao et al. 2010), and AMF (Adaptive Matched Filter cluster finder, Szabo et al. 2011) catalogs. We used the same methodology described for YOLO-CL and redMaPPer in Sect. 4.2 to assess the performance of these methods on the MCXC2021 cluster catalogs and we present our results in Figs. 11 and 12. MaxBCG recovers clusters with a ∼70% completeness at $0.2 < z < 0.3$ in the entire range of $L_x, M_{500}$, and $R_{500}$ that we cover here, with a 100% recovery only for $R_{500} \geq 1.4$ Mpc. At higher redshift, the completeness drops to ≤10%. GMBCG recovers clusters with a ∼100% completeness only at $0.2 < z < 0.3$ and $L_x \geq 10 \times 10^{44}$ erg s$^{-1}$, $M_{500} \geq 8.5 \times 10^{14} M_{\odot}$, $R_{500} \geq 1.3$ Mpc. At a redshift of $0.3 < z < 0.5$, its average completeness is of 50–70%. AMF recovers clusters with a ∼100% completeness at $0.2 < z < 0.3$ at $L_x \geq 2 \times 10^{44}$ erg s$^{-1}$, $M_{500} \geq 4 \times 10^{14} M_{\odot}$, $R_{500} \geq 1$ Mpc and $0.2 < z < 0.3$. At higher redshift, it recovers clusters with a ∼100% at $L_x \geq 10 \times 10^{44}$ erg s$^{-1}$, $M_{500} \geq 6 \times 8 \times 10^{14} M_{\odot}$, $R_{500} \geq 1–1.3$ Mpc. WHL09 has a performance similar to AMF, and WHL12 is the most complete of those traditional methods, with a completeness of ∼80% for in the entire $L_x, M_{500}$, and $R_{500}$.

Fig. 9. YOLO-CL (continuous lines) and redMaPPer (dashed lines) cluster detection completeness above a given X-Ray luminosity, $L_x, L_x$ (left panel), $M_{500}$ (middle panel), and $R_{500}$ (right panel). YOLO-CL recovers all clusters at $L_x \geq 1–3 \times 10^{44}$ erg s$^{-1}$, $M_{500} \geq 2–3 \times 10^{14} M_{\odot}$, $R_{500} \geq 0.75–0.8$ Mpc and $z \geq 0.4$. At lower luminosity, mass, radius and redshift, its performance worsens. The redMaPPer algorithm recovers all clusters at $L_x \geq 30–90 \times 10^{44}$ erg s$^{-1}$, $M_{500} \geq 2–6 \times 10^{14} M_{\odot}$, $R_{500} \geq 0.8–1.2$ Mpc and $z \geq 0.4$. At high redshifts both YOLO-CL and redMaPPer demonstrate a similar performance that is limited by the SDSS depth.

Fig. 10. YOLO-CL and redMaPPer MCXC2021 cluster detection completeness as a function of redshift and X-ray surface brightness. Left: YOLO-CL detects ∼98% of the MCXC2021 clusters with $L_x, 0.0 \geq 20 \times 10^{44}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$ at $0.2 \leq z \leq 0.6$ and ∼100% of the MCXC2021 clusters with $L_x, 0.0 \geq 30 \times 10^{41}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$ and $z \geq 0.3$. Right: redMaPPer detects ∼98% of the MCXC2021 clusters with $L_x, 0.0 \geq 55 \times 10^{43}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$ at $0.2 \leq z \leq 0.6$ and ∼100% of the MCXC2021 clusters with $L_x, 0.0 \geq 20 \times 10^{44}$ erg s$^{-1}$ cm$^{-2}$ arcmin$^{-2}$ at $0.5 \leq z \leq 0.6$. On the right of each figure, we give the completeness scale. From this comparison, YOLO-CL is more complete than redMaPPer in detecting MCXC2021 clusters.
Fig. 11. Fraction of MCXC2021 clusters recovered by tradition cluster detection methods in the SDSS (see text) from to bottom: MaxBCG, GMBCG, AMF, WHL09, and WHL12. The details of cluster recovery in each case are detailed in the text. In all cases, except WHL12, their completeness is worse than that reached by redMaPPer and YOLO-CL. These results, compared with Fig. 8 outline the high performance of our YOLO-CL with respect to traditional cluster detection methods in optical images.

ranges at $0.2 \lesssim z \lesssim 0.6$. This last method reaches a completeness of $\sim 100\%$ for $L_X \gtrsim 0.8 - 3 \times 10^{44}$ erg s$^{-1}$, $M_{500} \gtrsim 2 - 3 \times 10^{14} M_{\odot}$, $R_{500} \gtrsim 0.8 - 0.9$ Mpc for $z < 0.5$, and for $L_X \gtrsim 10 \times 10^{44}$ erg s$^{-1}$, it is $M_{500} \gtrsim 7 \times 10^{14} M_{\odot}$, $R_{500} \gtrsim 1.2$ Mpc for $0.5 < z < 0.6$.

Figure 12 summarizes the performance of these three algorithms. While the MaxBCG, GMBCG, and WHL09 cluster catalogs are much less complete than the redMaPPer and YOLO-CL catalogs, the AMF cluster catalog results are very similar to redMaPPer’s when considering the X-ray surface brightness. WHL12 shows a completeness very similar to YOLO-CL. In fact, 82% of the clusters found by YOLO–CL and not found by redMaPPer are also found in WHL12.
Fig. 12. Completeness of the MaxBCG, GMBCG and AMF galaxy cluster catalogs as a function of redshift and mean X-Ray surface brightness, $I_{500}$. On the right of each figure is the completeness scale. All traditional cluster detection algorithms applied to SDSS are less complete that redMaPPer and YOLO-CL (see text), except AMF, which has a performance similar to redMaPPer, and WHL12, which has a performance similar to YOLO-CL.

This confirms that the redMaPPer catalog that we have used to train our network is very good in terms of the recovery of X-ray cluster detections and outlines the high performance of our YOLO-CL with respect to traditional cluster detection methods applied to SDSS. Unfortunately, we cannot complete our comparison using both completeness and purity because the estimates of purity for each method are not homogeneous.

Together with its high performance in terms of completeness and purity, a strong advantage of galaxy cluster detection by deep learning networks is that clusters can be found without the need of measuring galaxy photometry and photometric redshifts. In fact, the direct use of color images allows us to skip the step of photometric and photometric redshift catalog preparation, and eliminates the systematic uncertainties that can be introduced during this process. This advantage has been pointed out in Chan & Stott (2019), where the authors introduced, for the first time, the use of deep learning for cluster detection in SDSS with the development of Deep-CEE (Deep Learning for Galaxy Cluster Extraction and Evaluation). Deep-CEE is based on Faster region-based convolutional neural networks, trained on the cluster catalog of Wen et al. (2012) in the redshift range $0.05 < z < 0.8$. As a proof of concept, they obtained completeness and purity of ~75% and ~80% when optimizing both on their validation sample and the redMaPPer catalog, respectively.
The use of convolutional networks with color images also allows us to equally focus on the two main aspects of galaxy clusters that are used for detection: (i) the fact that they are galaxy overdensities and (ii) the same distance and redshift (and therefore similar colors) of cluster members. The vast majority of existing detection algorithms focus on one of these aspect more than on the other. Color images preserve both the information about galaxy positions and colors, without assumptions on the significance of the overdensity or galaxy colors.

We conclude that our YOLO-CL network exhibits a higher level of performance in terms of completeness and purity in detecting galaxy clusters when compared to traditional cluster detection algorithm applied to SDSS images and catalogs. A strong advantage of deep learning networks is that clusters can be found without any need for measuring galaxy photometry and photometric redshifts or biases inherent to galaxy detection as well as these two measurements.

6. Summary

We apply the YOLO object detection deep convolutional network to the detection of galaxy clusters in the SDSS survey. Our network implementation, YOLO–CL, is a modification of the original YOLO v3 implementation to optimize galaxy cluster detection.

YOLO–CL was trained and validated using three color images (in the $g$, $r$, $i$ bandpasses) of 24,406 detections from redMaPPer cluster catalog and the equivalent number of SDSS blank field images. In the validation, we obtained an estimation of our network cluster catalog completeness and purity. To assess our sample completeness with respect to X-ray detected clusters, we compared our YOLO–CL detections to the X-ray-detections MCXC2021 clusters. YOLO–CL also shows a very high level of purity, measured using SDSS blank field images. As pointed out in the first implementation of a deep convolutional network for galaxy cluster detection (Chan & Stott 2019), deep learning networks have a strong advantage of galaxy cluster detection with respect to traditional techniques because they do not need galaxy photometric and photometric redshift catalogs. This eliminates the systematic uncertainties that can be introduced during the source detection, as well the measurements of photometry and photometric measurements, and focuses the detection method on the two main aspects of galaxy cluster detection, namely: the search for overdensities of galaxies that have similar colors because they are at the same redshift, without assumptions on the significance of the overdensity or galaxy colors. Our results highlight another advantage: a higher cluster catalog completeness than traditional cluster detection algorithms applied SDSS and a very high purity. Interestingly, the YOLO–CL selection function is approximately constant with redshift, with respect to the X-ray surface brightness.

We conclude that deep convolutional network for galaxy cluster detection are an efficient alternative to traditional cluster detection methods, and it is worth exploring their performance for future cosmological cluster catalogs for large-scale surveys, such as Rubin/Lsst, Euclid, and Roman Space Telescope.

Acknowledgements. We thank Université Paris Cité (UPC), which founded KG’s Ph.D. research, and University Paris Science & Letters (PSL), which founded SI postdoctoral research. We thank our PSL and UPC colleagues, and collaborators from the LightOn (https://lighton.ai/) company and the École Normale Supérieure (ENS), Laurent Daudet, Florent Krzakala, and Amelie Chatelain, for fruitful discussions. We thank Jean-Baptiste Melin and James Bartlett for useful discussions and help in choosing the X-ray catalog that we use in this paper. We thank Alex Saro for a insightful question that improved our comparison to the redMaPPer algorithm. We thank Eli Rykoff for proving the complete redMaPPer catalog. We gratefully acknowledge support from the CNRS/IN2P3 Computing Center (Lyon - France) for providing computing and data-processing resources needed for this work. This research has made use of the M5C Galaxy Cluster Database, constructed as part of the ERC project M2C (The Most Massive Clusters across cosmic time, ERC-Adv grant No. 340519). This work was supported by the French Space Agency (CNES). We thank the anonymous referee for her/his careful reading of the manuscript and useful suggestions that helped to improve the paper.

References

