

## Full length article

Galaxy detection and identification using deep learning and data augmentation<sup>☆</sup>R.E. González<sup>a,b,\*</sup>, R.P. Muñoz<sup>a</sup>, C.A. Hernández<sup>b</sup><sup>a</sup> Centro I+D MetricArts, Santiago, Chile<sup>b</sup> Centro de Astro-Ingeniería, Pontificia Universidad Católica, Av. Vicuña Mackenna 4860, Santiago, Chile

## ARTICLE INFO

## Article history:

Received 25 March 2018

Received in revised form 17 August 2018

Accepted 4 September 2018

Available online 15 September 2018

## Keywords:

Galaxies

General

Techniques

Image processing

Computing methodologies

Machine learning

## ABSTRACT

We present a method for automatic detection and classification of galaxies which includes a novel data-augmentation procedure to make trained models more robust against the data taken from different instruments and contrast-stretching functions. This method is shown as part of AstroCV, a growing open source computer vision repository for processing and analyzing big astronomical datasets, including high performance Python and C++ algorithms used in the areas of image processing and computer vision.

The underlying models were trained using convolutional neural networks and deep learning techniques, which provide better results than methods based on manual feature engineering and SVMs in most of the cases where training datasets are large. The detection and classification methods were trained end-to-end using public datasets such as the Sloan Digital Sky Survey (SDSS), the Galaxy Zoo, and private datasets such as the Next Generation Virgo (NGVS) and Fornax (NGFS) surveys.

Training results are strongly bound to the conversion method from raw FITS data for each band into a 3-channel color image. Therefore, we propose data augmentation for the training using 5 conversion methods. This greatly improves the overall galaxy detection and classification for images produced from different instruments, bands and data reduction procedures.

The detection and classification methods were trained using the deep learning framework DARKNET and the real-time object detection system YOLO. These methods are implemented in C language and CUDA platform, and makes intensive use of graphical processing units (GPU). Using a single high-end Nvidia GPU card, it can process a SDSS image in 50 ms and a DECam image in less than 3 s.

We provide the open source code, documentation, pre-trained networks, python tutorials, and how to train your own datasets, which can be found in the AstroCV repository. <https://github.com/astroCV/astroCV>.

© 2018 Elsevier B.V. All rights reserved.

## 1. Introduction

Astronomical datasets are constantly increasing in size and complexity. The modern generation of integral field units (IFUs) are generating about 60 GB of data per night while imaging instruments are generating 300 GB per night. The Large Synoptic Survey Telescope (LSST; Ivezic et al., 2008) is under construction in Chile and it is expected to start full operations in 2022. With a wide 9.6 square degree field of view 3.2 Gigapixel camera, LSST will generate about 20 TB of data per night and will detect more than 20 million of galaxies.

Machine learning techniques have been increasingly employed in data-rich areas of science. They have been used in genomics,

high-energy physics and astronomy. Some examples in astronomy are the detection of weird galaxies using Random Forests on Sloan data (Baron and Poznanski, 2017), Gravity Spy (Zevin et al., 2017) for LIGO detections and using convolutional neural network (CNN) in identifying strong lenses in imaging data (Jacobs et al., 2017).

Computer Vision is an interdisciplinary field that focuses on how machines can emulate the way in which human's brains and eyes work together to visually process the world around them. For many years, the detection of objects was computed using manual feature engineering and descriptors such as SIFT and HOG (Dalal and Triggs, 2005). Thanks to the advent of large annotated datasets and gains in computing power, deep learning methods have become the favorite for doing detection and classification of objects.

The classification of optical galaxy morphologies is based on a few simple rules that make them suitable for machine learning and computer vision techniques. The Kaggle Galaxy Zoo (Willett et al., 2013) was a competition based on a citizen science project where the aim was to predict the probability distribution of people's

<sup>☆</sup> This code is registered at the ASCL with the code entry 1804.004.

\* Corresponding author at: Centro de Astro-Ingeniería, Pontificia Universidad Católica, Av. Vicuña Mackenna 4860, Santiago, Chile.

E-mail address: [regonzar@astro.puc.cl](mailto:regonzar@astro.puc.cl) (R.E. González).

**Table 1**  
Annotations in different subsamples.

Dataset	Elliptical	Spiral	Edge-on	DK	Merge	Total	Number images
Training S1	10 366	4535	4598	223	381	20 103	6 458
Validation S1	1 261	714	723	27	45	2 770	921
Training S2	18 030	7828	7910	350	648	34 766	11 010
Validation S2	2 119	856	873	36	82	3 966	1 161
Custom	705	401	462	474	135	2 177	87

responses about the morphology of a galaxy using optical image data, and the winning solution used CNNs (Dieleman et al., 2015).

We present a method for galaxy classification and identification with a novel data augmentation procedure which is part of AstroCV, a computer vision library for processing and analyzing big astronomical datasets. The goal of AstroCV is to provide a community repository for fast Python and C++ implementations of common tools and routines used in the areas of image processing and computer vision. In particular, it is focused on the task of object detection, segmentation and classification applied to astronomical sources.

In this paper we will focus on the automatic detection and classification of galaxies. The detection and classification methods were trained end-to-end using public datasets from the Sloan Digital Sky Survey (SDSS), (Alam et al., 2015), and Galaxy Zoo (Lintott et al., 2008, 2011) explained in Section 2.1. We use YOLO method, (Redmon et al., 2015), for object detection which is explained in Section 2.2. Training process is described in Section 2.3, and results are shown in Section 3.

The open source code, training datasets, documentation and python notebooks of AstroCV are freely available in a Github repository.<sup>1</sup>

## 2. Data and training

### 2.1. Dataset

Galaxy Zoo<sup>2</sup> (Lintott et al., 2008, 2011) is the most successful citizen project in Astronomy. It consists of a web platform for doing visual inspection of astronomical images and morphological classification of galaxies. Hundreds of thousands of volunteers classified images of nearly 900,000 galaxies drawn from the SDSS survey. The Galaxy Zoo classification consists of six categories: elliptical, clockwise spiral, anticlockwise spiral, edge-on, star/do not know, or merger.

We extracted the galaxy classification for a sub-sample of 38,732 galaxies and downloaded their respective gri-band images from the SDSS fields. Sub-sample S1 is produced from 20 000 field images and sub-sample S2 is produced from 32 000 field images.

For each field image, we select the galaxies with a size larger than 22 pixels box side. This galaxy size is computed as 2.1 times the  $r$ -band petrosian radius. After this size filter we stay with two samples of 7397 images and 12 171 images. Then, we split each of these two samples into training and validation sub-sets, resulting in S1 and S2 datasets. In addition, we include a small custom dataset (Custom hereafter) with manually annotated galaxies from Hubble Deep Field image gallery,<sup>3</sup> CFHT,<sup>4</sup> and others images randomly taken from public databases. See Table 1 with details of the different samples.

### 2.2. YOLO

You only look once (YOLO) method (Redmon et al., 2015; Redmon and Farhadi, 2016), is a Single Shot Detector (SSD), it means

**Table 2**  
Training sets.

Name	Dataset	Filters	Images
T1	S1	L	6 458
T2	S1	LH	6 458
T3	S2	L	11 010
T4	S2	L+LH+S+SH+Q	55 050
T5	S2	LH+SH	22 020
T6	S2	L+LH+S+SH+Q	32 290

(L = Lupton, LH = Lupton high contrast, S = sinh, SH = sinh high contrast, Q = sqrt; C = custom Hubble sample.)

it computes in a single network the region proposal and classifier. This method runs the image on a Convolutional Neural Network (CNN) model and gets the detection on a single pass. The network is composed of 23 convolution layers and 5 maxpool layers shown in Fig. 1, and it is programmed on Darknet, an open source neural network framework in C and CUDA. It is very fast and takes full advantage of graphical processing units (GPU). This method formerly developed for person/object detection, is configured for the training and detection of galaxies.

### 2.3. Training and data augmentation

YOLO method is designed to work on 3 channel color images, usually RGB color scale. In astronomy images are taken for each filter in FITS format with raw CCD data for each pixel. Data conversion from FITS to RGB images (or contrast stretching) is not unique and depends on the telescope's camera, band filters, reduction schema, and most important, it depends on the conversion method used to scale photon counts to color scale.

There are several conversion methods, however to emphasize galaxies with strong luminosity gradients, linear scaling is not suitable, i.e. a spiral galaxy radial luminosity profile can be modeled as a power law (de Vaucouleurs profile) for the bulge plus an exponential for the disk. In those cases, the scaling methods commonly used are *sinh*, *asinh*, *sqrt* functions. SDSS uses (Lupton et al., 2004) as standard conversion method from FITS in *igr* bands to RGB an image (Lupton method hereafter).

In general, to train neural networks using images, the data augmentation is fundamental, it means increasing the training dataset with different transformations of the same data (scaling, rotations, crop, warp, etc.); In YOLO this data augmentation is already implemented in the training procedure, however we need to produce a color-scale/filter conversion augmentation as well (filter hereafter), to build a training robust against RGB images coming from different filters, bands and instruments.

In the top performance deep learning methods for object detection, we have also Faster R-CNN, Single shot detectors (SSD), Deconvolutional DSSD, Focal Loss, Feature Pyramid Networks (FPN). In Lin et al. (2017) there is a complete review and comparison on current methods. Most of these methods present similar mean average precision when compared to YOLO, however we stay with the latter since it is the fastest and implemented in C with CUDA acceleration.

In Table 2 we show 5 different training sets with RGB images produced to explore dataset size and filter augmentation, we use

<sup>1</sup> <https://github.com/astroCV>.

<sup>2</sup> <https://www.galaxyzoo.org/>.

<sup>3</sup> [http://www.spacetelescope.org/science/deep\\_fields](http://www.spacetelescope.org/science/deep_fields).

<sup>4</sup> <http://www.cfht.hawaii.edu>.

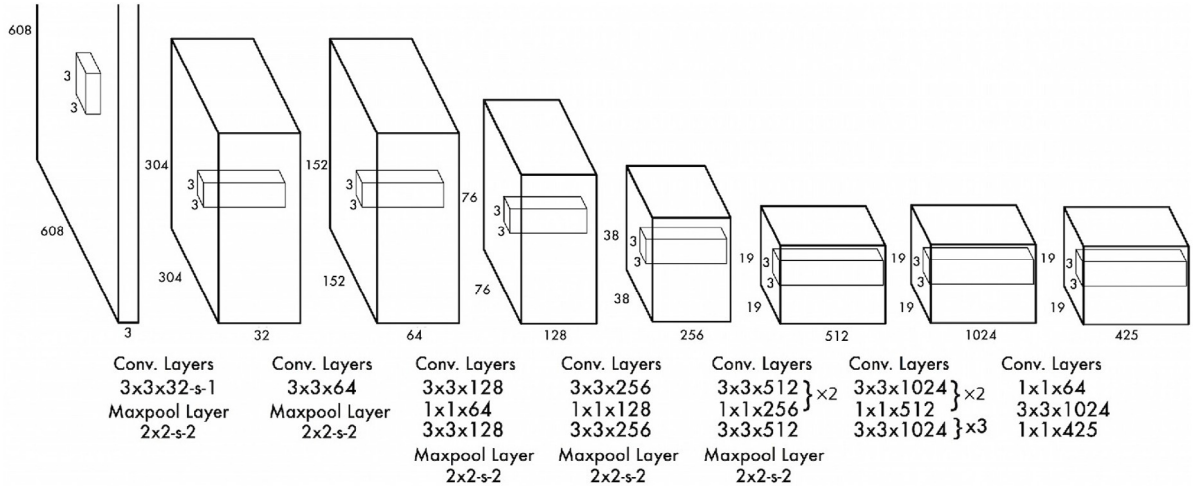


Fig. 1. YOLO network schema.

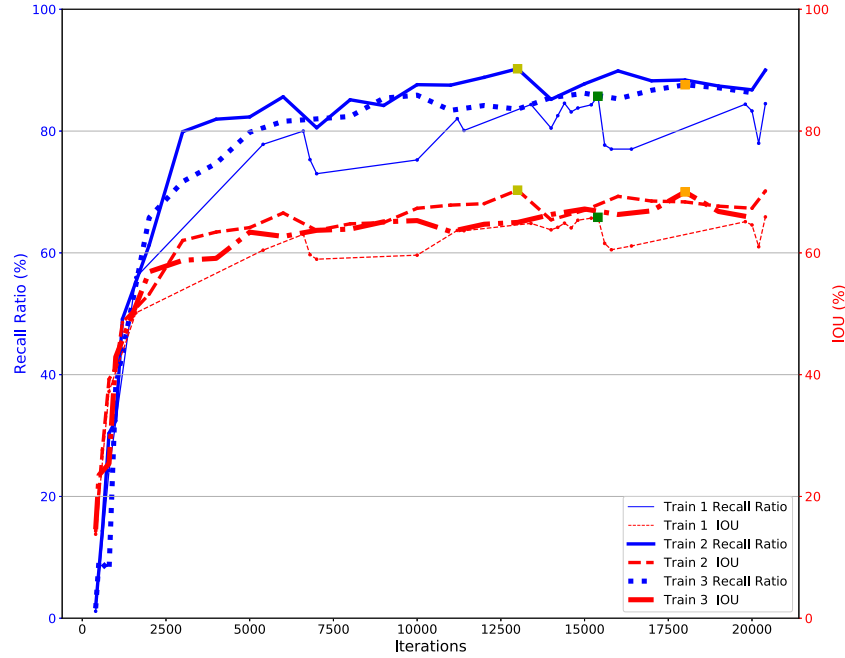


Fig. 2. Convergence of T1, T2 and T3 training sets. Green, yellow and orange squares indicate optimal recall ratio and IOU for the three trainings. Comparison between traditional (T1) and high contrast (T2) Lupton filters shows a detection improvement of 5% for T2. Increasing the dataset size (T3) for Lupton filter (T1) shows a detection improvement of 2%. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Lupton, Lupton high contrast,<sup>5</sup> *sinh*, *sinh* high contrast, and *sqrt* conversion functions.

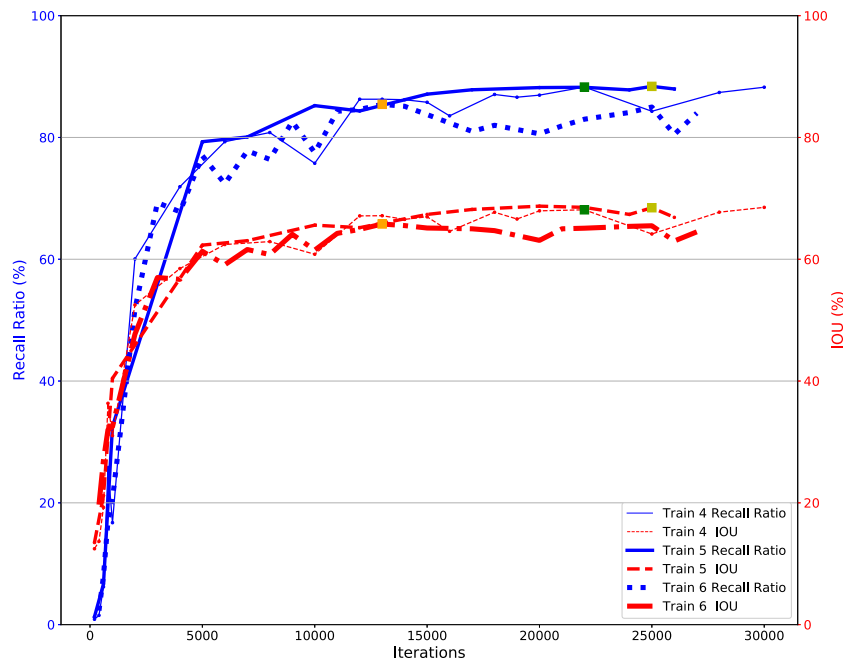
The purpose of these 5 training sets (T1, T2, T3, T4, T5 hereafter) is to test color scale conversion effects, dataset size, and produce a training robust against any astronomical FITS image converted to RGB. T1 and T2 are intended to show training differences produced between two similar contrast stretching functions; T3 will show the effect of increasing dataset size when compared to T1; The effect of including additional contrast stretching filters for data augmentation will be obtained from comparison between T3, T5 and T6. Finally T6, will be used to check the effect of dataset size in the case of augmented trainings when compared with T4.

### 3. Results

In the training process, we look for the optimal number of iterations where the network accuracy converges. We look for maximum recall ratio (fraction of detections recovered from the ground truth), and IOU (Intersect over union, overlap fraction between the detected and ground truth bounding box) running the training over the corresponding validation sets.

We can explore the effect of brightness and contrast using Lupton filter by comparing T1 and T2, results are shown in Fig. 2, where we get for T1, a maximum recall ratio and IOU of 85.7% and 65.8% respectively. In the case of T2, maximum recall ratio and IOU rise to 90.23% and 70.3% respectively. T2 returns better results since higher contrast and brightness images built from FITS images gives more information to the training, in particular for disk galaxies where disk can be seen more clearly with a higher contrast

<sup>5</sup> Contrast and Brightness enhanced by 2 using ImageEnhance python library.



**Fig. 3.** Convergence of the T4, T5 and T6 training sets where data is augmented by several conversion filters. Recall and IOU remain similar to Fig. 2, however this time detection and classification is more robust against different contrast stretching functions, instruments or reduction processes.

**Table 3**

Confusion matrix for galaxy classification using T2. Predicted values (columns) vs. actual values (rows).

n = 2756	Elliptical	Spiral	Edge-on	DK	Merge	Recall
Elliptical	1172	33	57	1	2	0.92
Spiral	176	469	69	0	3	0.65
Edge-on	96	60	554	0	3	0.78
DK	6	9	3	3	0	0.14
Merge	26	2	3	0	9	0.23
Precision	0.79	0.82	0.81	0.75	0.53	Accuracy = 0.80

**Table 4**

Confusion matrix for galaxy classification using T4. Predicted values (columns) vs. actual values (rows).

n = 19246	Elliptical	Spiral	Edge-on	DK	Merge	Recall
Elliptical	9525	476	257	0	3	0.93
Spiral	940	2991	305	0	6	0.71
Edge-on	688	499	3063	0	4	0.72
DK	39	36	24	6	4	0.06
Merge	170	129	44	0	37	0.10
Precision	0.84	0.72	0.83	1.0	0.69	Accuracy = 0.81

and brightness. However, excessive brightness may add too much noise to the background and excessive contrast tends to erase central bulge profile. In Fig. 2 we also show T3 where we get 87.6% and 69.8% maximum recall and IOU, when compared with T1, we have that increasing the dataset size for a fixed conversion filter improves results accuracy, however this improvement is smaller than the case of using a different high contrast conversion filter.

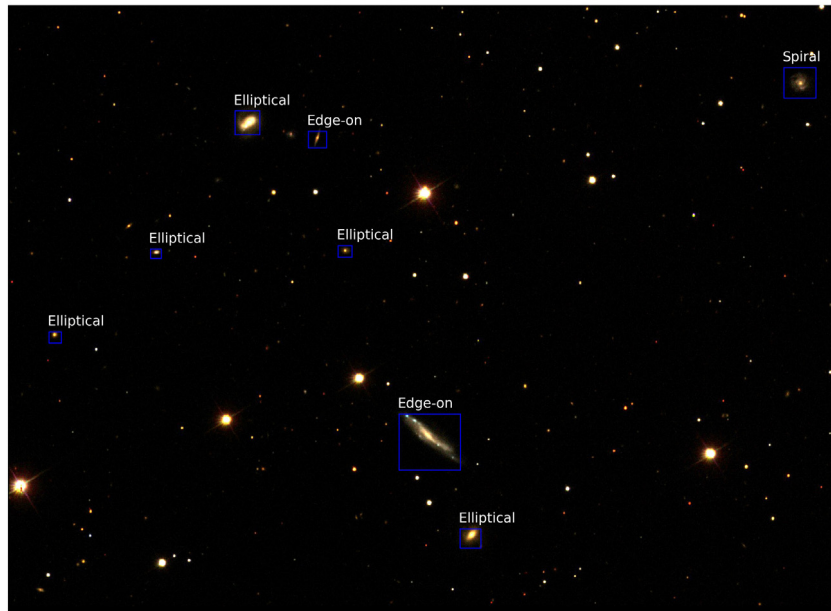
We also explore data augmentation using additional color conversion filters. As shown in Table 2 T4 and T6 are augmented using 5 filters and T5 using 2 filters. Fig. 3 shows results for T4, T5 and T6, where we get for T4, maximum recall ratio and IOU of 88.28% and 68.12% respectively, in the case of T5, maximum recall ratio and IOU are 88.24% and 68.52% respectively, and in the case of T6, maximum recall ratio and IOU are 85.44% and 65.79% respectively. Detection and classification accuracy remains close to 90% regardless of whether we are using data augmentation with 2 or 5 filters, or if we are using the S1 or S2 sample. This allows

the method to be robust against a wide range of color scale and filter conversions from FITS to RGB images, without losing detector accuracy.

Regarding data sample size, we explore dataset size effect on training using S1 and S2. As shown in Table 1 dataset S2 is 75% larger than dataset S1, in addition, when we compare recall ratio and IOU converges between S1 and S2 based trainings (for instance, T1 and T3, or T4 and T6), we have that S2 based trainings show a slightly better accuracy and smoother convergence compared to S1 based training sets. From this result, we can expect that using a larger training set should not improve substantially method accuracy/convergence and we are close to YOLO network maximum accuracy for this particular model.

We explore in more detail the two most relevant training sets T2 and T4. T2 has the highest recall ratio using a single conversion filter and T4 includes the largest data augmentation.

To study the accuracy of the method we compute the confusion matrix for T2 and T4 in Tables 3 and 4 respectively. The confusion



**Fig. 4.** Detection and classification of galaxies in a  $2048 \times 1489$  image from SDSS using T2 training; i, r, and g filters converted to rgb color space using Lupton filter.

matrix is constructed using the corresponding validation sets for T2 and T4 with  $2770$  and  $3966 \times 5$  galaxies respectively. We use a detection threshold of 15%. Columns indicate predicted values and row the ground truth.<sup>6</sup> In the top left corner we show the number of detected galaxies, notice that numbers do not match exactly with the validation set from Table 1, this is because of many reasons, such as, merge objects detected as two objects; other reason is the detection threshold value; and a third reason is that detection process includes non-maximum suppression (NMS) issues, this corresponds to the post-processing algorithm responsible for merging all detections that belong to the same object or overlapping objects. However, the sum of all these effects is small, in both training sets we miss less than 3% detections.

Confusion matrix for both trainings shows very little confusion between the three major classes elliptical, spiral and edge-on. In the case of DK and Merge classes the number of annotation is very small compared to the other classes resulting in poor performance for those classes. In addition, merge class in galaxy zoo is very ambiguous since the difficulty separating a blended galaxy pair and two independent close galaxies.

The total classification accuracy is 80% for T2 and 81% for T4, which is high taking into account the nature of Galaxy Zoo classification based on votes, where some objects have divided votes but are assigned to a single classification (i.e. Dwarf spheroidal or S0 galaxies). This shows that conversion filter data augmentation does not degrade method accuracy, and makes the training more robust against different image sources coming from different instruments.

In Fig. 4, we show results of the method using T2 on an image from the SDSS converted using Lupton filter. In the github repository there are examples to reproduce this figure.

In Fig. 5 we show result over a Hubble Deep field image (taken from Custom dataset) using T2 and T4. This image uses infrared bands and uses a different color scale conversion, not comparable with SDSS filters and conversions. However, we can see how filter data augmentation from T4 produces a  $3\times$  improvement in the detection of galaxies compared with T2. Then T4 training is more robust to images produced from different filters and instruments.

<sup>6</sup> Notice that ground truth in the context of galaxy zoo defines the galaxy class based on the number of votes.

**Table 5**

Detection accuracy for different training sets validated with the custom dataset.

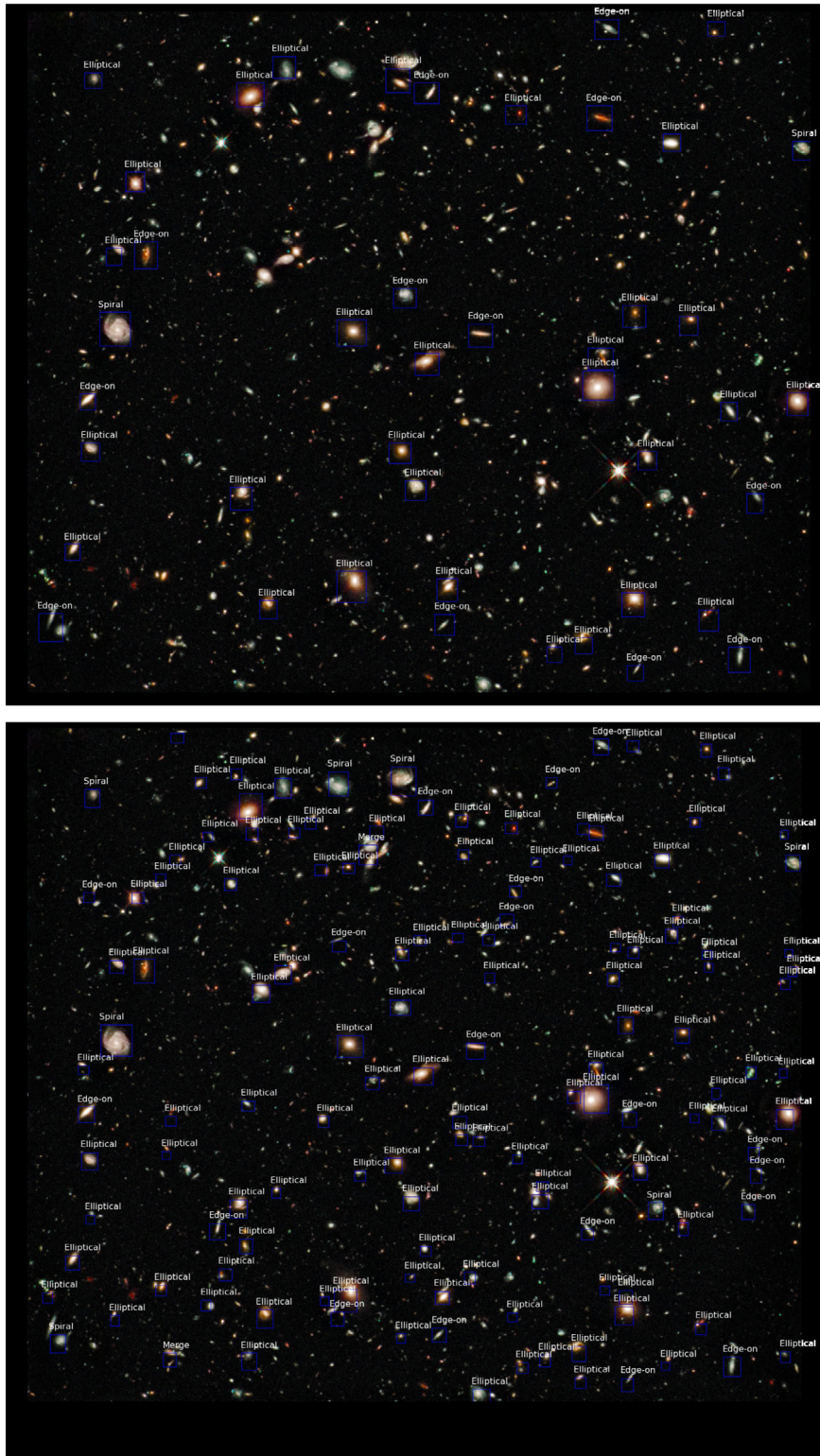
Training set	Recall ratio	IOU ratio
T1	0.120	0.236
T2	0.183	0.296
T3	0.192	0.292
T4	0.404	0.423
T5	0.271	0.336
T6	0.364	0.403

To quantify the improvement of data augmentation for the different trainings, we validate them using the custom dataset which includes images randomly taken from different telescopes and filters, and manually annotated. Results for the recall ratio and IOU are shown in Table 5. Performance comparison between (T1, T2, T3) and (T4, T6), shows that data augmentation using 5 contrast stretching functions, improves detection and identification of galaxies in random images by a factor of 2–3. One to one comparison between trainings indicates: (i) T1–T2, LH gives more information to the training compared to standard Lupton conversion; (ii) T1–T3 or T4–T6, increasing the dataset may improve detector accuracy, but just a little when compared to data augmentation; (iii) T3–T5–T4, there is a consistent increasing improvement in detection accuracy when including more data augmentation of the same data.

#### 4. Discussion

The current classification model was created using a sample of SDSS galaxies from the Galaxy Zoo Project. We propose a data augmentation based on different additional filters to convert FITS data from different bands into RGB color scale, and this show an important improvement in comparison with traditional training using a single color scale conversion, specially with images produced with bands and conversion filters totally different from the one used in the former training (see Fig. 5 and Table 5). We show a factor of  $\sim 3$  improvement in the detection and classification of random galaxy images (Custom dataset) taken from different telescopes, bands, and contrast stretching functions. This makes the training set more robust and generic to be applied to images





**Fig. 5.** Galaxy detection results on a Hubble deep field image. Top panel: Using T2(44 detections). Bottom panel: Using augmented training T4(140 detections).

produced from any telescope and band, and shows that increasing the dataset size for training is not necessarily the most important target to produce a good galaxy detector/classifier, while a well based data augmentation may lead to more significant improvements.

This method is part of the AstroCV library intended to provide computer vision and image processing tools for processing big

astronomical datasets. These tools are focused on the task of object detection, segmentation and classification of astronomical objects using deep learning and neural network frameworks. This method is implemented on DARKNET which takes advantage of GPUs capabilities, improving performance for real-time applications such as data reduction pipelines. As an example, running galaxy detection using a Titan-X Nvidia GPU card on a typical full HD image takes

nearly 50 ms, and for a DECam 500 Mpx images it takes less than three seconds.

There are several other interesting datasets to be trained such as Galaxy Zoo 2, Willett et al. (2013) and the Next Generation Fornax Survey (NGFS; Muñoz et al., 2015). The Galaxy Zoo 2 extends the classification for nearly 300,000 of the largest Galaxy Zoo 1 galaxies, including detailed features such as discs, bars, and arms. The NGFS survey extends the classification to low surface brightness galaxies down to  $\mu_i = 28$  mag arcsec<sup>2</sup>.

The next generation of astronomical observatories, such as the Large Synoptic Survey Telescope (LSST) and the Extreme Large Telescope (ELT), will observe hundreds of thousands of galaxies and will generate Terabytes of data every night. We are planning to improve the scalability of AstroCV to do real-time processing of tomorrow's big astronomical datasets. Another interesting application we are exploring is the identification and classification of transient objects, and detection of high redshift candidates for spectroscopy.

In this paper we limit the training procedure for three color bands only, to produce the RGB contrast stretching, since this is the classic approach to manually find features and classify galaxies (i.e. Galaxy Zoo). However, training may be strongly improved if we use more color bands (i.e. ugriz), or an optimal contrast stretching, and this will be part of further work.

AstroCV is open source, the github repository at <https://github.com/astroCV/astroCV> includes code, documentation, pre-trained networks, python tutorials for using the library and training your own datasets.

## Acknowledgments

The authors acknowledge support from the NVIDIA developer program. The Centro I+D MetricArts acknowledges support from the NVIDIA Inception program and the Microsoft BizSpark Program. The authors are grateful to Patricio Cofre, Wilson Pais and Adrian Fernandez for helpful discussions. RG was supported by Proyecto Financiamiento Basal PFB-06 'Centro de Astronomía y Tecnologías Afines'.

Funding for SDSS-III has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, and the U.S. Department of Energy Office of Science. The SDSS-III web site is <http://www.sdss3.org/>.

## References

Alam, S., Albareti, F.D., Allende Prieto, C., Anders, F., Anderson, S.F., Anderton, T., Andrews, B.H., Armengaud, E., Aubourg, É., Bailey, S., et al., 2015. The eleventh and twelfth data releases of the Sloan Digital Sky Survey: final data from sdss-iii. *Astrophys. J. Suppl.* 219, 12. doi:10.1088/0067-0049/219/1/12. arXiv:1501.00963.

Baron, D., Poznanski, D., 2017. The weirdest SDSS galaxies: results from an outlier detection algorithm. *Mon. Not. R. Astron. Soc.* 465, 4530–4555. doi:10.1093/mnras/stw3021. arXiv:1611.07526.

Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Vol. 1. pp. 886–893 vol. 1. doi:10.1109/CVPR.2005.177.

Dieleman, S., Willett, K.W., Dambre, J., 2015. Rotation-invariant convolutional neural networks for galaxy morphology prediction. *Mon. Not. R. Astron. Soc.* 450, 1441–1459. doi:10.1093/mnras/stv632. arXiv:1503.07077.

Ivezic, Z., Tyson, J.A., Abel, B., Acosta, E., Allsman, R., Alsayyad, Y., Anderson, S.F., Andrew, J., Angel, R., Angeli, G., Ansari, R., Antilogus, P., Arndt, K.T., Astier, P.,

Aubourg, E., Axelrod, T., Bard, D.J., Barr, J.D., Barrau, A., Bartlett, J.G., Bauman, B.J., Beaumont, S., Becker, A.C., Becla, J., Beldica, C., Bellavia, S., Blanc, G., Blandford, R.D., Bloom, J.S., Bogart, J., Borne, K., Bosch, J.F., Boutigny, D., Brandt, W.N., Brown, M.E., Bullock, J.S., Burchat, P., Burke, D.L., Cagnoli, G., Calabrese, D., Chandrasekharan, S., Chesley, S., Cheu, E.C., Chiang, J., Claver, C.F., Connolly, A.J., Cook, K.H., Cooray, A., Covey, K.R., Cribbs, C., Cui, W., Cutri, R., Daubard, G., Daves, G., Delgado, F., Digel, S., Doherty, P., Dubois, R., Dubois-Felsmann, G.P., Durech, J., Eracleous, M., Ferguson, H., Frank, J., Freemon, M., Gangler, E., Gawiser, E., Geary, J.C., Gee, P., Geha, M., Gibson, R.R., Gilmore, D.K., Glanzman, T., Goodenow, I., Gressler, W.J., Gris, P., Guyonnet, A., Hascall, P.A., Haupt, J., Hernandez, F., Hogan, C., Huang, D., Huffer, M.E., Innes, W.R., Jacoby, S.H., Jain, B., Jee, J., Jernigan, J.G., Jevremovic, D., Johns, K., Jones, R.L., Juramy-Gilles, C., Juric, M., Kahn, S.M., Kalirai, J.S., Kallivayalil, N., Kalmbach, B., Kantor, J.P., Kasliwal, M.M., Kessler, R., Kirkby, D., Knox, L., Kotov, I., Krabbenand, V.L., Krughoff, S., Kubanek, P., Kuczewski, J., Kulkarni, S., Lambert, R., Le Guillou, L., Levine, D., Liang, M., Lim, K., Lintott, C., Lupton, R.H., Mahabal, A., Marshall, P., Marshall, S., May, M., McKecher, R., Migliore, M., Miller, M., Mills, D.J., Monet, D.G., Moniez, M., Neill, D.R., Nief, J., Nomerotski, A., Nordby, M., O'Connor, P., Oliver, J., Olivier, S.S., Olsen, K., Ortiz, S., Owen, R.E., Pain, R., Peterson, J.R., Petry, C.E., Pierfederici, F., Pietrowicz, S., Pike, R., Pinto, P.A., Plante, R., Plate, S., Price, P.A., Prouza, M., Radeka, V., Rajagopal, J., Rasmussen, A., Regnault, N., Ridgway, S.T., Ritz, S., Rosing, W., Roucelle, C., Rumore, M.R., Russo, S., Saha, A., Sassolas, B., Schalk, T.L., Schindler, R.H., Schneider, D.P., Schumacher, G., Seabag, J., Sembroski, G.H., Seppala, L.G., Shipsey, I., Silvestri, N., Smith, J.A., Smith, R.C., Strauss, M.A., Stubbs, C.W., Sweeney, D., Szalay, A., Takacs, P., Thaler, J.J., Van Berg, R., Vanden Berk, D., Vetter, K., Virieux, F., Xin, B., Walkowicz, L., Walter, C.W., Wang, D.L., Warner, M., Willman, B., Wittman, D., Wolff, S.C., Wood-Vasey, W.M., Yoachim, P., Zhan, H., for the LSST Collaboration, 2008. LSST: from Science Drivers to Reference Design and Anticipated Data Products. ArXiv e-prints arXiv:0805.2366.

Jacobs, C., Glazebrook, K., Collett, T., More, A., McCarthy, C., 2017. Finding strong lenses in CFHTLS using convolutional neural networks. *Mon. Not. R. Astron. Soc.* 471, 167–181. doi:10.1093/mnras/stx1492. arXiv:1704.02744.

Lin, T.Y., Goyal, P., Girshick, R., He, K., Dollár, P., 2017. Focal loss for dense object detection. ArXiv e-prints. arXiv:1708.02002.

Lintott, C., Schawinski, K., Bamford, S., Slosar, A., Land, K., Thomas, D., Edmondson, E., Masters, K., Nichol, R.C., Raddick, M.J., Szalay, A., Andreescu, D., Murray, P., Vandenberg, J., 2011. Galaxy Zoo 1: data release of morphological classifications for nearly 900 000 galaxies. *Mon. Not. R. Astron. Soc.* 410, 166–178. doi:10.1111/j.1365-2966.2010.17432.x. arXiv:1007.3265.

Lintott, C.J., Schawinski, K., Slosar, A., Land, K., Bamford, S., Thomas, D., Raddick, M.J., Nichol, R.C., Szalay, A., Andreescu, D., Murray, P., Vandenberg, J., 2008. Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. *Mon. Not. R. Astron. Soc.* 389, 1179–1189. doi:10.1111/j.1365-2966.2008.13689.x. arXiv:0804.4483.

Lupton, R., Blanton, M.R., Fekete, G., Hogg, D.W., O'Mullane, W., Szalay, A., Wherry, N., 2004. Preparing red-green-blue images from CCD data. *Publ. Astron. Soc. Pac.* 116, 133–137. doi:10.1086/382245. arXiv:astro-ph/0312483.

Muñoz, R.P., Eigenthaler, P., Puzia, T.H., Taylor, M.A., Ordenes-Briceño, Y., Alamo-Martínez, K., Ribbeck, K.X., Ángel, S., Capaccioli, M., Côté, P., Ferrarese, L., Galaz, G., Hempel, M., Hilker, M., Jordán, A., Lançon, A., Mieske, S., Paolillo, M., Richtler, T., Sánchez-Janssen, R., Zhang, H., 2015. Unveiling a rich system of faint dwarf galaxies in the next generation fornax survey. *Astrophys. J. Lett.* 813, L15. doi:10.1088/2041-8205/813/1/L15. arXiv:1510.02475.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2015. You only look once: Unified, real-time object detection. ArXiv e-prints. arXiv:1506.02640.

Redmon, J., Farhadi, A., 2016. YOLO9000: Better, Faster, Stronger. ArXiv e-prints arXiv:1612.08242.

Willett, K.W., Lintott, C.J., Bamford, S.P., Masters, K.L., Simmons, B.D., Casteels, K.R.V., Edmondson, E.M., Fortson, L.F., Kaviraj, S., Keel, W.C., Melvin, T., Nichol, R.C., Raddick, M.J., Schawinski, K., Simpson, R.J., Skibba, R.A., Smith, A.M., Thomas, D., 2013. Galaxy Zoo 2: detailed morphological classifications for 304 122 galaxies from the Sloan Digital Sky Survey. *Mon. Not. R. Astron. Soc.* 435, 2835–2860. doi:10.1093/mnras/stt1458. arXiv:1308.3496.

Zevin, M., Coughlin, S., Bahaadini, S., Besler, E., Rohani, N., Allen, S., Cabero, M., Crowston, K., Katsaggelos, A.K., Larson, S.L., Lee, T.K., Lintott, C., Littenberg, T.B., Lundgren, A., sterlund, C., Smith, J.R., Trouille, L., Kalogera, V., 2017. Gravity Spy: integrating advanced LIGO detector characterization, machine learning, and citizen science. *Classical Quantum Gravity* 34 (6), 064003. URL <http://stacks.iop.org/0264-9381/34/i=6/a=064003>.