

Large Scale Study of Binary Galaxy Image Classification and the Impact of Image Augmentation Techniques

Tomas Mūžas(✉), Andrius Vytautas Misiukas Misiūnas, and Tadas Meškauskas

Institute of Computer Science, Faculty of Mathematics and Informatics, Vilnius University, Didlaukio 47, 08303 Vilnius, Lithuania
tomas.muzas@mif.stud.vu.lt

Abstract. Current galaxy classification studies are usually conducted on small, expert-classified datasets, constrained within a low redshift (z) range. Lower redshift implies better image quality – the lower the z value, the closer the object is, thus more features and details can be observed. Additionally, various augmentation methods are used to further improve classification accuracy, however, there is a lack of studies measuring their impact on other metrics. Therefore, we study the impact of augmentation techniques using the largest dataset that covers a broad redshift range (315,942 galaxies, $0 < z \leqslant 0.28$). We provide comparable evidence that for binary galaxy image classification, common image augmentation techniques – rotation, zoom and flipping – increase accuracy and F1 score, unless the model is underfitting. The most significant increase was observed on the ResNet50 model, for which the accuracy increased from 93.53% to 95.21%, and the F1 score – from 88.66% to 91.82%. Additionally, combining the aforementioned techniques with random noise stabilises a model by significantly decreasing the spread of metrics – for ResNet50, the standard deviation of the F1 score decreased by more than 9 times.

Keywords: Galaxy morphology · GalaxyZoo · Image augmentation · Deep learning · Convolutional neural networks

1 Introduction

Based on their morphology (visual appearance), galaxies are grouped into three major categories – spiral, elliptical and irregular. Spiral galaxies have arms, which extend from the centre of the galaxy, while elliptical galaxies are round and smooth, with a continuously declining brightness distribution [3]. The galaxies that do not fit the aforementioned criteria are known as irregular.

It was shown that the morphological classification of the galaxy also correlates with its properties, such as star formation and merger history, gas accretion or nuclear activity [3]. Thus, it is important to classify galaxies to better understand their properties.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
O. Gervasi et al. (Eds.): ICCSA 2023, LNCS 13957, pp. 402–412, 2023.
https://doi.org/10.1007/978-3-031-36808-0_27

The classification process can be automated by utilising machine learning. Studies have shown that convolutional neural networks (hereinafter CNNs) in particular excel at such task [2,4,8]. While the results achieved by other authors are great (over 95%) [5,8], or sometimes even near-perfect (over 98%) [2,4], the authors tend to carry out their experiments using datasets of varying size and image quality. Usually, a small, expert-classified dataset is taken (such as [4,5,8]), which implies a better image quality. Additionally, authors limit the range of the galaxy redshift (see Sect. 3.4 for comparison), which is another proxy to the image quality – the higher the redshift value, the less features of an object can be observed, making it more difficult to classify.

Moreover, the majority of the authors use image augmentation techniques to improve classification accuracy, though there is a lack of studies quantifiably and comparably evaluating whether those techniques have any meaningful impact on metrics other than accuracy.

Hence, in this paper, we combine data from two "citizen science" projects – Galaxy Zoo 1 and 2 (hereinafter GZ1 and GZ2) – to produce the largest dataset that covers a broad redshift range and consists only of reliably classified galaxies (see Sect. 3 for details). Using this dataset, we aim to provide comparable and reliable binary classification results of different CNN models, as well as measure the impact of currently used image augmentation techniques on various metrics. Furthermore, we introduce two novel, noise-based means of augmentation and study their impact in combination with other techniques.

We conclude that commonly used image augmentation techniques – rotation, horizontal and vertical flipping and zoom – do in fact increase accuracy and F1 score by a statistically significant amount. The most significant increase was observed on the ResNet50 model, for which the accuracy improved from 93.5% to 95.2% and the F1 score rose from 88.66% to 91.82%. However, single-class metrics – precision, recall and TNR – have improved by less than one standard deviation. Nonetheless, we also show that combining the aforementioned techniques with random noise allows reducing the standard deviation of metrics by up to 9 times – for the same ResNet50 model, the σ of F1 reduced from 1.14 to 0.14, and σ of precision reduced from 2.27 to 0.57.

2 Related Work

A significant amount of research has been conducted to accurately classify galaxies into spiral and elliptical with the help of various machine learning techniques. In particular, convolutional neural networks have produced better results compared to other machine learning methods such as decision trees, K-nearest neighbours or support vector machines [2,5,8,12]. This paper analyses galaxy classification into spiral and elliptical, as these are the main classes.

Some authors claim to have achieved near-perfect classification results in terms of binary classification. One example would be Barchi et al. [2], who managed to achieve 99.5% overall accuracy for the brightest galaxies from the GZ1 dataset, and 98.7% on a wider range of GZ1 galaxies. Similar results were

achieved by Cavanagh et al. [4], who classified galaxies with 98% accuracy. Those results are closely followed by Jimenez et al. [8] scoring 96.43% accuracy and Cheng et al. achieving [5] 95.1%.

Another prominent feature of the aforementioned studies is that the majority of authors use image augmentation techniques to achieve better results. The most frequently used augmentations are rotating and flipping image both horizontally and vertically. Such transformations were used by [4–6,12]. Two additional noteworthy techniques are Gaussian noise used by Cheng et al. [5] and zoom utilised by [6,9,16].

However, there are several shortcomings in the aforementioned papers. The first one is that the accuracy results are not comparable to each other, as datasets of different sizes are used. Additionally, all authors use either expert-classified [4,8] or low-redshift galaxies, which implies better image quality (the dataset diversity in terms of redshift is discussed in detail in Sect. 3.4).

The second issue is that different verification metrics and methods are provided by the authors. With respect to verification methods, [2,8] use cross-validation and average the metrics across the folds, while [4,5] use training-validation split technique and take the best results. In terms of metrics, [2,4,5] provide additional metrics such as precision, recall or F1 score, while [8] use only accuracy as their key metric.

The third shortcoming concerns augmentation techniques. While multiple authors use image augmentation techniques, only Cheng et al. study its impact on the metrics [5]. Other authors simply present results with augmentations already applied. Hence, it is not known whether image augmentation has any significant impact on the metrics.

Therefore, the aim of this paper is twofold: to provide comparable baseline binary classification metrics (accuracy, precision, recall, TNR and F1) using a large, reliable and diverse dataset, and to study the impact of frequently used augmentation techniques. Two novel, noise-based augmentation techniques are studied as well.

3 Data

This section describes the process of combining GZ1 and GZ2 data into one large, reliable and diverse dataset.

3.1 Galaxy Zoo 1

GZ1 project provides classification for 667,944 galaxies, where each is assigned one of the three classes – spiral, elliptical or uncertain [11]. The final, author-proposed classification is based on the debiased votes – a galaxy is considered either spiral or elliptical only if the debiased vote fraction is above the threshold of 0.8, otherwise, the "uncertain" class is attributed.

Thus, all the galaxies that have an `UNCERTAIN` flag in the dataset were filtered out. There were no galaxies that have both `SPIRAL` and `ELLIPTICAL` flags. Therefore, only 252,415 GZ1 galaxies were considered reliable.

3.2 Galaxy Zoo 2

As for the GZ2 dataset, the distinction whether the galaxy is spiral or elliptical was done based on the first question, in which participants were asked whether the galaxy is smooth (elliptical), or has features or a disk (spiral). After the votes for each answer option were counted, they also underwent the debiasing process [15]. Based on the debiased vote fraction, two different flags are assigned independently only if the fraction for the particular answer option exceeds 0.8 - `smooth_flag` and `features_or_disk_flag`. Thus, only the galaxies that have either one of these flags are considered reliable.

However, it is important to note that some of the galaxies had both of those flags. In this paper, such galaxies were considered ambiguous and subsequently excluded. All entries that had a `gz2_class` value of "A" were considered artefacts (not galaxies) and were excluded as well. After applying all of the aforementioned criteria, 154,887 galaxies were kept.

3.3 Combined Dataset

The resulting datasets from Sects. 3.1 and 3.2 were merged by their SDSS object identifiers (`OBJID` for GZ1, `dr7objid` for GZ2). The GZ1 data was used as a base. This has left a total of 319,991 unique galaxies.

If a galaxy was present in both datasets, and the final classifications did not agree, it was considered ambiguous as well. 4,049 such galaxies were identified and subsequently removed. Therefore, the resulting combined dataset contained 315,942 unique, reliably classified and unambiguous galaxies (224,775 spiral (71%) and 91,167 elliptical (29%)). This dataset will be referred to as `GZC` throughout the paper.

3.4 Diversity in Terms of Redshift

To satisfy the criterion of diversity, galaxy redshift (denoted as z) was chosen as a proxy for image quality. Redshift is the decrease in frequency of the electromagnetic radiation (such as light) emitted by the galaxy as it travels throughout space [7]. This means that the further the galaxy is from the observer (in this case - the telescope), the greater the redshift, hence the fainter it gets, corresponding to poorer quality.

GZ2 project was constrained to contain galaxies within the redshift range of $0.0005 < z < 0.25$, while GZ1 did not impose any redshift restrictions. This lack of restrictions introduced outliers as well as expanded the redshift range of the dataset. Using the 3-σ method, it was calculated that the galaxies outside of the range of $z \in (-0.01; 0.28)$ are outliers. Nonetheless, it was decided to keep all the initial galaxies, as they have been unambiguously classified by the GZ1 volunteers. As there were only 4 galaxies with negative z, it will be considered that the redshift range of the dataset is $0 < z \leqslant 0.28$.

The redshift range of the `GZC` dataset was compared to the ranges of other authors' datasets. The comparison of dataset ranges and sizes is visualised in

the Table 1. It can be concluded that the GZC dataset not only is 3 times larger in terms of size, but also has the widest redshift range.

Table 1. Comparison of the dataset sizes of various authors that performed binary classification.

Author	Size	Redshift range	Comment
Cheng et al.	2,862	$0 < z < 0.25$	DES and GZ1 data
Cavanagh et al.	14,304	$0.01 < z < 0.1$	Nair and Abraham dataset used
Jimenez et al.	41,424	$0.05 < z < 0.1$	Bigger redshift range taken among MOSES and Longo subsets
Barchi et al.	104,787	$0.05 < z < 0.1$	Size of GZ1 dataset used for binary classification. Redshift range not explicitly specified, largest stated range taken
This work	315,942	$0 < z \leqslant 0.28$	Galaxies that have $z > 0.28$ are considered outliers, though still kept in the dataset

3.5 Image Gathering Process

An image for each galaxy in the GZC dataset was generated using the SkyServer's ImgCutout service[1] [13] and SciServer script[2]. Each image had a dimension of 424×424 pixels, generated using the same parameters as described by GZ1 and GZ2 projects [11,15].

After that, a square of 212×212 pixels was cropped from the centre of the initial image, removing part of the background. Finally, the image was rescaled to 128×128 pixels and all the pixel values were normalised to be between 0 and 1 by diving them by 255.

4 Models and Their Modifications

In total, four different architectures were chosen to carry out the experiments. Three of them were used by other researchers to classify galaxies with high accuracy – ResNet50, used by Jimenez et al. [8], and Dieleman and C2 architectures, studied by Cavanagh et al. [4]. The fourth architecture was introduced by the authors of this paper to evaluate the effectiveness of augmentations on a model with significantly fewer parameters (approximately 100k), which would potentially underfit. The architecture itself is a simplification of the C2 and Dieleman models. It is visualised in Table 2 and will be referred to as the simplified model.

[1] Accessible at http://skyservice.pha.jhu.edu/DR7/ImgCutout/getjpeg.aspx.

[2] https://github.com/sciserver/SciScript-Python/blob/master/py3/SciServer/SkyServer.py.

Table 2. Architecture of the simplified model.

Convolution, 5x5, 64 channels, ($ReLU$)
Max pool, 2x2
Convolution, 5x5, 32 channels, ($ReLU$)
Max pool, 2x2
Convolution, 5x5, 16 channels, ($ReLU$)
Dropout, probability - 0.5
Flatten
Fully connected layer, 1 node ($Sigmoid$)

Each architecture was modified in such a way that it accepts a batch of images of size 128×128, and produces an output of 1 node in the range of $(0, 1)$. Then, a threshold of 0.5 is applied – if the output value is strictly greater than 0.5, then the galaxy is treated as elliptical, spiral otherwise.

5 Training Process

5.1 Parameters and Process

In total, 5 experiments were carried out with each model using the `GZC` dataset – baseline (no augmentations applied), standard augmentations only (rotation, horizontal and vertical flips and zoom), and standard augmentations together with random, random centred, random background noises.

To ensure that all the results are as comparable and reliable as possible, fixed initial model weights were used for all experiments. Additionally, to eliminate any bias inherent in the dataset split selection, 10-fold cross-validation method was employed, and all the results were averaged.

Each experiment had exactly the same configuration – training batch size was 1024, test batch size was 16. Adam optimiser with parameters of learning rate of 10^{-4}, $\beta_1 = 0.9, \beta_2 = 0.999, \epsilon = 1 \times 10^{-8}$ [10] was used to optimise the loss function of the binary cross-entropy. Each experiment was continued for maximum of 500 epochs, although if loss has not improved for 10 consecutive epochs, the training process was stopped early.

Some constrains were introduced by the fact that the model training was conducted on the tensor processing units (TPUs) provided by Google Collaboratory[3]. To ensure that all the dataset batches are full, two different approaches were used. For the training data, additional galaxies were sampled to complete the last batch. For the test data, such approach would skew the results, therefore it was decided to lower batch size to 16, and discard the last underfull batch. However, this was done only during the training process – after all the experiments were completed, the final results were reevaluated using the full test split. Only the aforementioned reevaluated results are presented in this paper.

[3] https://colab.research.google.com/.

5.2 Augmentation Techniques

For all the experiments, each augmentation technique had an independent probability of 0.5 to be applied on each image individually.

1. **Rotation**. This augmentation technique is widely used by multiple authors [4–6,9,12,16]. An image is rotated by $\{90, 180, 270\}$ degrees to avoid interpolation and introduction of new pixels.
2. **Horizontal and vertical flip.** Another highly utilised augmentation technique, used by [4,6,9,12,16]. This technique simply flips image pixels either horizontally or vertically.
3. **Zoom.** This technique, or its variations, were used by [6,9,16]. In this paper, 10 pixels are cropped from each side of an image. The resulting image is rescaled to the initial size of 128×128 pixels.
4. **Random noise.** This method is intended to simulate the noise that may be introduced by the observing telescope itself and was used by Cheng et al. [5]. As this is still an ongoing research, only a single noise amplitude of 0.2 was explored. To produce the noise, a random value between -0.2 and 0.2 is taken from a uniform distribution and applied to each pixel, in each channel, individually. The resulting value is clipped to the original range of 0 to 1.
5. **Random centre and background noises.** This novel technique aims to apply targeted noise only on a part of an image. The centre noise effect is achieved by first creating a Sersic surface brightness profile [14] using the Astropy library [1] with the following parameters: `amplitude` = 3, `r_eff` = 5, `n` = 1, and `ellip` = 0 (fully round). The generated values are then scaled between 0 and 1 using log_{10} and 90-per-cent percentile interval scaling. The resulting profile is then multiplied with random noise (generated as mentioned above) and applied on the image. As for the background noise profile, the exact same process is repeated, except that the Sersic profile is inverted. Both noise profiles are visualised in Fig. 1.

6 Results and Discussion

This section outlines and compares the results of all the experiments. In this paper, five model performance metrics will be analysed – accuracy, precision (also known as true positive rate), recall, true negative rate (hereinafter TNR) and F1 score.

Accuracy, precision, recall and F1 were chosen as those metrics are quite often used by other authors [2,4–6], which would allow for better comparability of studies. Along with precision, which is also known as true positive rate, another metric is taken into account – true negative rate. This is the same metric as precision, except it measures model performance on the negative class (in this case spiral galaxies), instead of positive (elliptical galaxies). Having both of those metrics allows to evaluate model performance for both morphological classes.

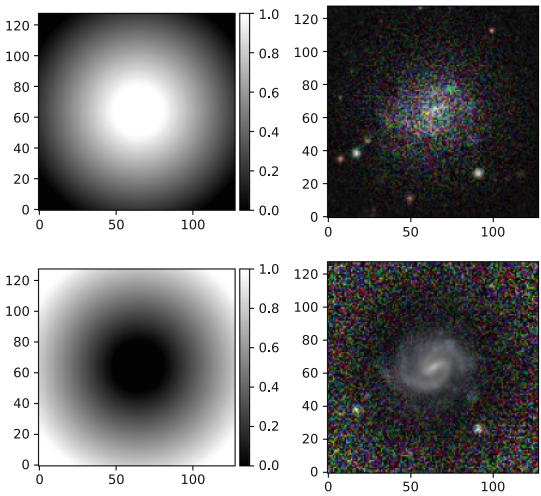

Fig. 1. A visualisation of the novel noise profiles. Top row depicts the Sersic profile and resulting centre noise, the bottom row – inverted Sersic profile and resulting background noise.

Throughout all the experiments, zoom, rotation and flipping were combined into a single group and will be referred to as the standard augmentations. The results of the experiments are visualised in Tables 3a–3e.

First of all, a few important observations can be made from Tables 3a and 3b. First of all, none of the metrics has improved for the simplified model. This is most likely due to the fact that the model is too simple and underfits as initially predicted. As for the C2, Dieleman and ResNet50 models, applying standard augmentations has improved average accuracy and F1 score by a statistically significant amount. However, other metrics had just a negligible improvement which falls within one standard deviation. Nonetheless, applying standard augmentations has helped the ResNet50 model to reduce the standard deviation of all 5 metrics by 2, or in the case of recall and F1 score, 4 times, thus stabilising the model. A similar effect can be observed in the C2 model, though the improvement is not as significant.

According to Tables 3c–3e, it is evident that there is no statistically significant difference between the experiments with only standard augmentations. It is worth mentioning that since this is an ongoing research, only one value of noise amplitude was explored. A wider variety of noise amplitudes should be considered in the future, which could potentially alter the results.

Nevertheless, even though the metrics have not improved, there is one important observation to be made in terms of the random noise profile. Using this technique, the ResNet50 model was stabilised even further, by once again halving σ compared to standard augmentations alone.

Table 3. Average metrics of the experiments across 10 folds. Value in parentheses denotes standard deviation, all values are given in percentages.

(a) None of the augmentations applied.

Model	Acc. (σ)	Precision (σ)	Recall (σ)	TNR (σ)	F1 (σ)
Simplified	94.72 (0.24)	90.15 (0.51)	91.72 (0.63)	96.62 (0.25)	90.93 (0.41)
C2	94.71 (0.12)	89.80 (1.18)	92.19 (1.47)	96.80 (0.56)	90.96 (0.23)
Dieleman	94.70 (0.12)	89.68 (0.79)	92.28 (0.93)	96.83 (0.36)	90.95 (0.21)
ResNet50	93.53 (0.48)	89,54 (2.27)	88.01 (3.93)	95.21 (1.44)	88.66 (1.14)

(b) Standard augmentations only.

Model	Acc. (σ)	Precision (σ)	Recall (σ)	TNR (σ)	F1 (σ)
Simplified	94.56 (0.17)	89.84 (0.61)	91.52 (0.76)	96.54 (0.29)	90.67 (0.30)
C2	95.36 (0.14)	90.73 (0.68)	93.49 (0.95)	97.33 (0.37)	92.09 (0.26)
Dieleman	95.42 (0.17)	91.18 (0.81)	93.17 (0.72)	97.20 (0.28)	92.16 (0.28)
ResNet50	95.21 (0.23)	90.50 (1.11)	93.21 (1.02)	97.22 (0.39)	91.82 (0.36)

(c) Standard augmentations + random noise.

Model	Acc. (σ)	Precision (σ)	Recall (σ)	TNR (σ)	F1 (σ)
Simplified	94.47 (0.16)	89.80 (0.52)	91.21 (0.73)	96.41 (0.28)	90.49 (0.30)
C2	95.34 (0.10)	90.73 (0.55)	93.42 (0.54)	97.30 (0.21)	92.05 (0.16)
Dieleman	95.31 (0.08)	90.74 (0.46)	93.28 (0.53)	97.24 (0.20)	91.99 (0.14)
ResNet50	95.38 (0.09)	91.32 (0.57)	92.80 (0.53)	97.06 (0.20)	92.05 (0.14)

(d) Standard augmentations + centre noise.

Model	Acc. (σ)	Precision (σ)	Recall (σ)	TNR (σ)	F1 (σ)
Simplified	94.63 (0.20)	89.76 (0.55)	91.89 (0.34)	96.68 (0.13)	90.81 (0.32)
C2	95.29 (0.19)	90.67 (0.51)	93.27 (0.47)	97.24 (0.19)	91.95 (0.32)
Dieleman	95.34 (0.13)	91.04 (0.78)	93.02 (1.06)	97.15 (0.41)	92.01 (0.26)
ResNet50	95.23 (0.28)	90.75 (0.91)	92.97 (1.07)	97.12 (0.42)	91.84 (0.49)

(e) Standard augmentations + background noise.

Model	Acc. (σ)	Precision (σ)	Recall (σ)	TNR (σ)	F1 (σ)
Simplified	94.72 (0.23)	89.93 (0.42)	92.01 (0.84)	96.73 (0.33)	90.95 (0.42)
C2	95.39 (0.17)	90.86 (0.60)	93.43 (0.75)	97.30 (0.29)	92.12 (0.30)
Dieleman	95.52 (0.11)	91.44 (0.56)	93.22 (0.43)	97.23 (0.17)	92.32 (0.17)
ResNet50	95.07 (0.37)	90.33 (1.15)	92.88 (0.98)	97.08 (0.38)	91.58 (0.61)

To summarise, this study shows that standard augmentation techniques do in fact increase accuracy and F1 score. Additionally, standard augmentations,

especially in combination with random noise, greatly improve model stability by decreasing σ values across different runs.

Nonetheless, there was no statistically significant evidence that the metrics that represent single-class performance (namely precision, recall and TNR) improve with the use of augmentations. Moreover, neither background nor centre noise profiles had any significant impact on the metrics. Although those techniques might still prove useful with different noise amplitudes which are to be explored with future research.

It is crucial to emphasise that contrary to other similar studies, this research provides not only all of the most commonly used performance metrics, but also uses the largest dataset of only reliably classified galaxies that have a great diversity in terms of redshift values. Furthermore, the use of 10-fold cross-validation allows to remove the bias of dataset splits by exploring all split combinations. Thus, the aforementioned methods together make all the metrics reliable, comparable, and more representative of models' expected performance with larger amounts of data than the metrics provided by other authors.

7 Conclusions

By conducting multiple experiments using the largest dataset of reliably classified galaxies covering the range of $0 < z \leqslant 0.28$ and verifying results using 10-fold cross-validation and models of four different CNN architectures, we conclude that the usage of standard augmentation techniques (rotation, horizontal and vertical flip and zoom) improves accuracy and F1 score by a statistically significant amount. The most significant improvement was observed for the ResNet50 model, for which accuracy improved from 93.53% to 95.21%, and F1 score from 88.66% to 91.82%.

Moreover, adding the random noise profile greatly stabilises the performance of the model by reducing the spread of all the metrics across the folds. The most significant improvement was observed for the ResNet50 model – the standard deviation (σ) improved from the 2.27% to 0.57% for the precision metric, from 1.44% to 0.20% for the TNR score, and from 1.14% to 0.14% for the F1 score.

However, none of the augmentation techniques improve single-class metrics (precision, recall and TNR) by a statistically significant amount, nor have any effect if the model is underfitting (as in the case of the simplified model).

References

1. Astropy Collaboration: The astropy project: sustaining and growing a community-oriented open-source project and the latest major release (v5.0) of the core package. Astrophys. J. **935**(2), 167 (2022). https://doi.org/10.3847/1538-4357/ac7c74
2. Barchi, P., et al.: Machine and deep learning applied to galaxy morphology - a comparative study. Astron. Comput. **30** (2020). https://doi.org/10.1016/j.ascom.2019.100334
3. Buta, R.J.: Planets, Stars and Stellar Systems: Volume 6: Extragalactic Astronomy and Cosmology. Springer, Netherlands (2013)

4. Cavanagh, M.K., Bekki, K., Groves, B.A.: Morphological classification of galaxies with deep learning: comparing 3-way and 4-way CNNs. Mon. Not. R. Astron. Soc. **506**(1), 659–676 (2021). https://doi.org/10.1093/mnras/stab1552
5. Cheng, T.Y., et al.: Optimizing automatic morphological classification of galaxies with machine learning and deep learning using Dark Energy Survey imaging. Mon. Not. R. Astron. Soc. **493**, 4209–4228 (2020). https://doi.org/10.1093/mnras/staa501
6. Dieleman, S., et al.: Rotation-invariant convolutional neural networks for galaxy morphology prediction. Mon. Not. R. Astron. Soc. **450**(2), 1441–1459 (2015). https://doi.org/10.1093/mnras/stv632
7. Gray, R., Dunning-Davies, J.: A review of redshift and its interpretation in cosmology and astrophysics. arXiv (2008)
8. Jiménez, M., et al.: Galaxy image classification based on citizen science data: a comparative study. IEEE Access **8**, 47232–47246 (2020). https://doi.org/10.1109/ACCESS.2020.2978804
9. Khalifa, N.E., et al.: Deep galaxy V2: robust deep convolutional neural networks for galaxy morphology classifications. In: 2018 International Conference on Computing Sciences and Engineering (ICCSE), pp. 1–6. IEEE (2018). https://doi.org/10.1109/ICCSE1.2018.8374210
10. Kingma, D., Ba, J.: Adam: a method for stochastic optimization. In: International Conference on Learning Representations (2014)
11. Lintott, C.J., et al.: Galaxy Zoo: morphologies derived from visual inspection of galaxies from the Sloan Digital Sky Survey. Mon. Not. R. Astron. Soc. **389**(3), 1179–1189 (2008). https://doi.org/10.1111/j.1365-2966.2008.13689.x
12. Mittal, A., Soorya, A., Nagrath, P., Hemanth, D.J.: Data augmentation based morphological classification of galaxies using deep convolutional neural network. Earth Sci. Inf. **13**(3), 601–617 (2019). https://doi.org/10.1007/s12145-019-00434-8
13. Nieto-Santisteban, M., Szalay, A., Gray, J.: ImgCutout, an engine of instantaneous astronomical discovery. In: Astronomical Data Analysis Software and Systems (ADASS) XIII, vol. 314, p. 666 (2004)
14. Sérsic, J.L.: Influence of the atmospheric and instrumental dispersion on the brightness distribution in a galaxy. Boletin de la Asociacion Argentina de Astronomia **6**, 41 (1963)
15. Willett, K.W., et al.: Galaxy Zoo 2: detailed morphological classifications for 304,122 galaxies from the Sloan Digital Sky Survey. Mon. Not. R. Astron. Soc. **435**(4), 2835–2860 (2013). https://doi.org/10.1093/mnras/stt1458
16. Zhu, X.-P., Dai, J.-M., Bian, C.-J., Chen, Yu., Chen, S., Hu, C.: Galaxy morphology classification with deep convolutional neural networks. Astrophys. Space Sci. **364**(4), 1–15 (2019). https://doi.org/10.1007/s10509-019-3540-1