

Image restoration with point-spread function regularization and active learning

Peng Jia^{1,2,3}, Jiameng Lv^{1*}, Runyu Ning¹, Yu Song¹, Nan Li^{4*}, Kaifan Ji^{5*}, Chenzhou Cui³ and Shanshan Li³

¹College of Electronic Information and Optical Engineering, Taiyuan 030024, China

²Peng Cheng Lab, Shenzhen 518066, China

³Department of Physics, Durham University, Durham DH1 3LE, UK

⁴National Astronomical Observatories, Beijing 100101, China

⁵Yunnan Observatories, Kunming, Yunnan, China

Accepted 2023 October 30. Received 2023 October 30; in original form 2023 April 24

ABSTRACT

Large-scale astronomical surveys can capture numerous images of celestial objects, including galaxies and nebulae. Analysing and processing these images can reveal the intricate internal structures of these objects, allowing researchers to conduct comprehensive studies on their morphology, evolution, and physical properties. However, varying noise levels and point-spread functions can hamper the accuracy and efficiency of information extraction from these images. To mitigate these effects, we propose a novel image restoration algorithm that connects a deep-learning-based restoration algorithm with a high-fidelity telescope simulator. During the training stage, the simulator generates images with different levels of blur and noise to train the neural network based on the quality of restored images. After training, the neural network can restore images obtained by the telescope directly, as represented by the simulator. We have tested the algorithm using real and simulated observation data and have found that it effectively enhances fine structures in blurry images and increases the quality of observation images. This algorithm can be applied to large-scale sky survey data, such as data obtained by the Large Synoptic Survey Telescope (LSST), *Euclid*, and the *Chinese Space Station Telescope (CSST)*, to further improve the accuracy and efficiency of information extraction, promoting advances in the field of astronomical research.

Key words: methods: numerical – techniques: image processing – software: data analysis.

1 INTRODUCTION

Sky survey projects obtain numerous observational images containing celestial objects with extended structures, including nebulae and galaxies, which are of utmost interest and importance for astronomical research. Scientists select these targets for in-depth analysis. However, aberrations and noise in astronomical observations can degrade image quality, resulting in blurred and distorted images of celestial objects. These factors significantly affect the precision of scientific information extracted from such images. Thus, it is crucial to develop and utilize image restoration algorithms to improve image quality and enable further scientific exploration. Recent developments in deep neural networks have led to the emergence of several new image restoration algorithms, which take advantage of the powerful representation capabilities of deep neural networks and have demonstrated impressive results. Two primary types of image restoration algorithms have been proposed based on different representation strategies: algorithms based on the properties of astronomical images and those based on the properties

of the degradation process, such as the point-spread function (PSF) of telescopes and the characteristics of noise. These algorithms can enhance the quality of images obtained from astronomical observations.

Image restoration algorithms based on image properties are often developed with generative neural networks, such as Auto-Encoding Variational Bayes (VAE) (Kingma & Welling 2013), U-Net (Ronneberger, Fischer & Brox 2015), or Generative Adversarial Networks (GANs) (Goodfellow et al. 2020). These neural networks are trained on a large number of high signal-to-noise ratio (SNR) and high spatial resolution real observation images, from which they learn features. The learned features are then used to restore low SNR or low spatial resolution images. Properly selected training data can lead to effective results for galaxy images (Schawinski et al. 2017; Arcelin et al. 2021; Gan, Bekki & Hashemizadeh 2021; Jia et al. 2021; Li et al. 2022) or solar images (Jia et al. 2019). However, manual interventions are often required to obtain appropriate training data, and overfitting or training on improper data can result in the generation of fake structures by the neural network (Jia et al. 2021).

Image restoration algorithms based on the properties of PSFs aim to learn the generalized inverse function of PSFs. The method first constructs a PSF model using wavefront decomposition (Beltramo-

* E-mail: lvjiameng1224@link.tyut.edu.cn (JL); nan.li@nao.cas.cn (NL); jkf@ynao.ac.cn (KJ)

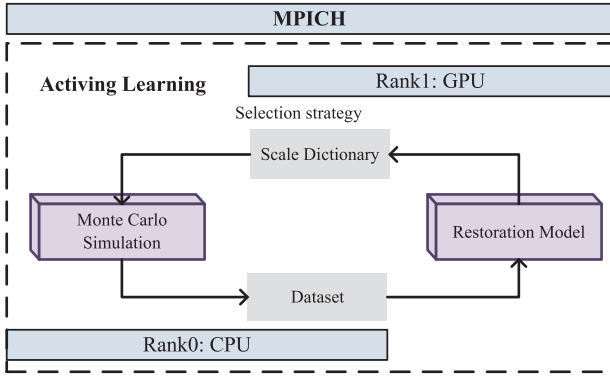


Figure 1. The image restoration framework is illustrated in the diagram, which is composed of three main parts: the Monte Carlo simulation part, the image restoration part, and the parameter selection part. In the Monte Carlo simulation part, PSFs are generated under different observation conditions and sample images are convolved to generate blurred images as training data. The image restoration part restores these blurred images. Meanwhile, the parameter selection part evaluates the quality of the restored images and generates parameter dictionaries for the data generation part. To speed up the training process, we employ the MPI technology to run the data generation and image restoration parts in separate processors or computers.

Martin et al. 2019; Fétick et al. 2019; Fusco et al. 2020; Jia et al. 2020) or PSF basis (Jia et al. 2017; Sun et al. 2020), and then uses this model as prior knowledge to restore images through supervised or unsupervised learning. In unsupervised learning algorithms, the PSFs are decomposed into parameters, and the image restoration algorithm fits these parameters through training (Qi et al. 2014; Gao et al. 2017; Sureau, Lechat & Starck 2020). However, these algorithms also face the parameter-tuning problem encountered by classical image restoration algorithms. In supervised learning algorithms, PSFs are used to generate blurred images as the training set (Jia et al. 2020; Li & Alexander 2023; Wang et al. 2022). However, two issues limit the application of supervised algorithms. First, PSFs are necessary as prior information for supervised learning algorithms. Although some methods have been proposed to extract star images as PSF templates (Terry et al. 2023), obtaining appropriate PSFs is complicated due to the presence of sky background noise or read-out noise. Furthermore, obtaining star images from extended target images (such as galaxies or nebulae) is challenging (Long et al. 2019). Second, the neural network’s generalization ability is dependent on the training set. Because PSFs can change significantly, it is difficult to obtain PSFs and develop datasets that can reflect different states of real observation data. Improper PSF datasets may introduce epistemic uncertainty, which can limit the performance of trained image restoration algorithms (Hüllermeier & Waegeman 2021).

Therefore, we present a novel framework for processing images from large-scale astronomical sky surveys. Our framework integrates a high-fidelity simulator of a specific telescope and a deep neural network based image restoration algorithm with the active learning strategy. The simulator is customized for the telescope used in the sky survey and can generate simulated images with various PSFs and noise levels typically encountered in real observations. The image restoration algorithm restores these images, and we evaluate the quality of the restored images by computing the mean square error (MSE) between the original and restored images. The active learning strategy governs the simulator to generate more images with PSFs or noise levels that the image restoration neural network could not restore well, thereby further training the neural network.

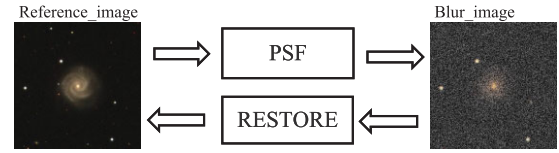


Figure 2. This figure shows the structure of PSF-NET. It includes an image restoration neural network (RESTORE) and a PSF generation neural network (PSF).

This approach enables us to obtain an effective neural network for image restoration. We discuss our framework in detail in Section 2 and demonstrate its ability to restore simulated and real observation images with varying levels of blur in Section 3. Finally, we draw our conclusions and outline future work in Section 4.

2 THE FRAMEWORK

Large-scale astronomical surveys produce a vast number of blurred images with varying quality, making it impractical to process them manually one by one using human intervention-based algorithms. Therefore, an image restoration framework capable of producing stable results for images captured by the sky survey project is necessary. This work proposes a framework that addresses this need, which is illustrated in Fig. 1. The framework consists of three parts: the Monte Carlo simulation part, the image restoration part, and the parameter selection part. To expedite the training process of the framework, we propose using parallel computing technology to run the data generation and image restoration parts on different computers or processors. We will discuss the details of our framework in the following subsections.

2.1 The Monte Carlo simulation part

The Monte Carlo simulation part generates PSFs of the telescope for various observation conditions and convolves sample images to create blurred images as the training data. The imaging process of an optical telescope can be modelled with equation (2):

$$Img(x, y) = [Obj(x, y) * PSF(x, y)]_{\text{pixel}(x, y)} + Noise(x, y), \quad (1)$$

where $Obj(x, y)$ and $Img(x, y)$ are original and observed images. $PSF(x, y)$ is the point-spread function of the telescope, $[\]_{\text{pixel}(x, y)}$ stands for the pixel response function of the detector, and $Noise(x, y)$ stands for the noise from the background and the detector.

This work focuses primarily on studying long-exposure images captured by ground-based optical telescopes, but our framework can be applied to images collected by any telescope, provided that we have an appropriate simulator. In the case of ground-based long-exposure observations, atmospheric turbulence has a significant impact on the PSFs, which can be accurately modelled by using the Moffat model with β equal to 4.765 (Moffat 1969). We set the full width at half-maximum (FWHM) of the PSF as the first free parameter in the simulator. Additionally, we use Gaussian random numbers to model various levels of noise caused by the detector and background, and we set the standard deviation as the second free parameter in the simulator. By using these parameters, we can generate images with realistic PSFs and noise, which represent actual observation conditions for image restoration algorithms better. Finally, we convolve several high-resolution

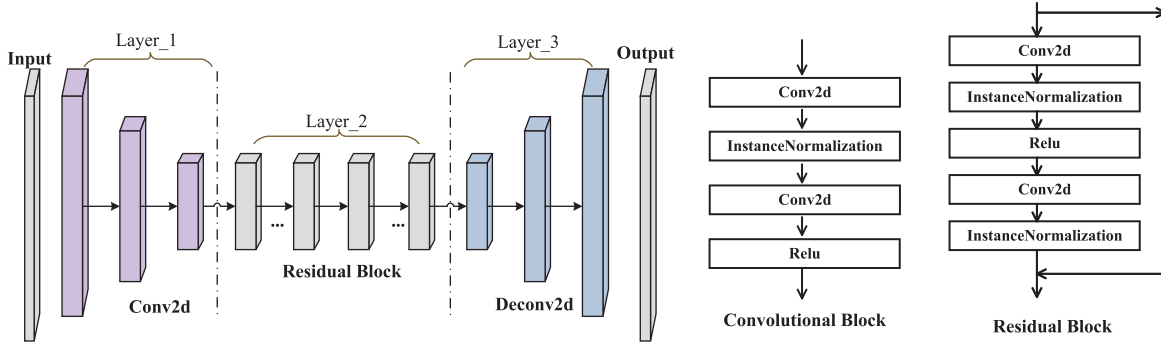


Figure 3. The left panel of the figure displays the architecture of the PSF neural network, which is similar to that of the RESTORE neural network. The middle panel illustrates the convolutional block, while the right panel depicts the residual block.

template images with PSFs and add noise to generate the training data.

2.2 The image restoration part

The image restoration neural network we use in this study is PSF-NET, which was proposed by Jia et al. (2020). PSF-NET consists of two neural networks, namely the PSF network (PSF) and the RESTORE network (RESTORE), as shown in Fig. 2. Here is a detailed description of their roles and functionalities. Firstly, the primary task of the PSF network is to model the image's blurring process. It achieves this by learning and representing the characteristics of the PSF and noise, transforming a high-resolution image into a blurred version. The objective of this step is to capture the fundamental information regarding image blurring, thereby enhancing the accuracy and efficiency of the subsequent restoration process. Secondly, the RESTORE network plays the role of a deconvolution algorithm, tasked with generating high-resolution, clear images from the blurred images. After being trained, the RESTORE network is capable of effectively restoring blurred images to high-quality images, recovering the fine details and information in the image. It is noteworthy that both the PSF and RESTORE networks comprise six residual blocks, along with several convolutional and transposed convolutional blocks. The structures of these blocks are depicted in Fig. 3. During the training phase, we adopt a joint training approach to train both the PSF and RESTORE networks, thereby improving training efficiency and mitigating overfitting risks. After training, the RESTORE network can be deployed for the task of restoring blurred images.

An important point to consider is that both the PSF network and the RESTORE network are trained together, with the PSF network acting as a constraint for the RESTORE network. Unlike a simple deconvolution algorithm, the RESTORE network handles deconvolution and noise reduction. Therefore the PSF network learns impacts brought by the PSF and the noise. This structural design not only infuses the restoration process with enhanced physical realism but also bolsters the model's ability to generalize, enabling it to address various blurring scenarios effectively. The PSF neural network in this study generates blurred images from high-resolution images, while the RESTORE neural network generates high-resolution images from blurred images. To reflect the functions of these two neural networks, we have designed the loss function as presented in

equation (2):

$$Loss = L_{idt} + L_{rec} + L_{ff}, \quad (2)$$

where L_{idt} is the identity loss function, L_{rec} is the cycle loss function, and L_{ff} is the focal frequency loss function. They are defined in equation (3):

$$\begin{aligned} L_{idt} &= \|PSF(Img_{org}) - Img_{blur}\|_2 \\ &\quad + \|Restore(Img_{blur}) - Img_{org}\|_2, \\ L_{rec} &= \|Restore(PSF(Img_{org})) - Img_{org}\|_2 \\ &\quad + \|PSF(Restore(Img_{blur})) - Img_{blur}\|_2, \\ L_{ff} &= W \times L_{ff}, \end{aligned} \quad (3)$$

where PSF and $Restore$ stand for operators carried out by the PSF neural network and the RESTORE neural network, while Img_{blur} and Img_{org} stand for blurred images generated by the simulation method and original high-resolution images. We use L_{idt} and L_{rec} to set the MSE between the original images and the images restored by the neural network to be as small as possible.

Both the PSF neural network and the RESTORE neural network adopt an encoder-decoder structure. However, since the decoder contains several upsampling deconvolution layers that may introduce gaps in restored images, using only the aforementioned loss function may pose a problem. Our empirical observations suggest that gaps affect the structure of restored galaxy images and are more noticeable in the spatial frequency domain. To address this issue, we propose using the focal frequency loss L_{ff} to further enhance the performance of our neural network. L_{ff} was originally introduced by Jiang et al. (2021), and involves the use of regularized weights W to modulate the power spectral density, with L_{ff} representing the mean squared error (MSE) between the original and restored images in the spatial frequency domain, as defined in equation (4):

$$\begin{aligned} W &= |FFT(Img_{org}) - FFT(Restore(Img_{blur}))|^\alpha, \\ L_{ff} &= \|FFT(Img_{org}) - FFT(Restore(Img_{blur}))\|^2, \end{aligned} \quad (4)$$

The fast Fourier transform is represented by FFT and the regularized parameter α is used in the focal frequency loss L_{ff} , which we defined as 1 in this study.

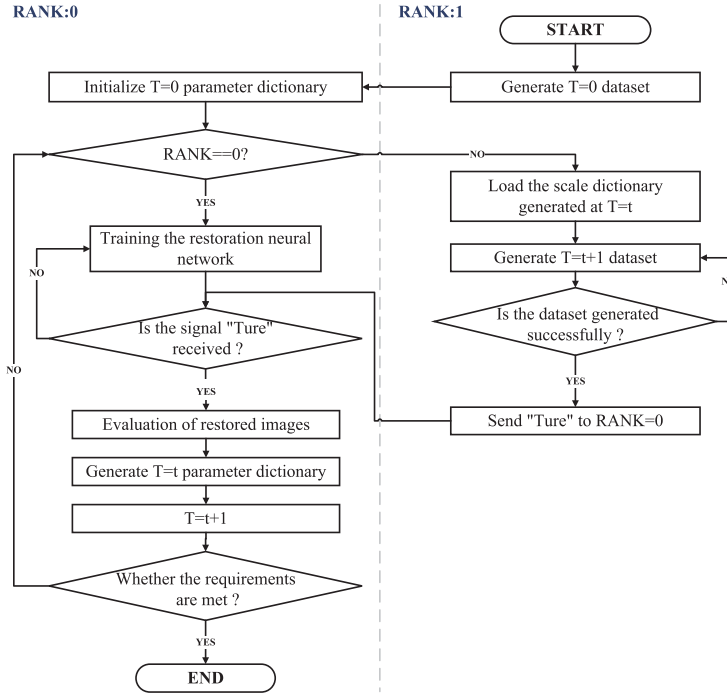


Figure 4. The flow chart of the parameter selection part is shown in this figure. The parameter selection part comprises two ranks, each with its own distinct role. Rank0 is responsible for training and testing the image restoration component, and runs primarily on a computer equipped with an RTX 3090 GPU. On the other hand, Rank1 is responsible for generating blurred images using the simulation component and runs on several other computers equipped with CPUs. For this purpose, we utilize older laptops or desktops that are equipped with i5 or i7 processors to run the simulation code.

2.3 The parameter selection part

The parameter selection part is the central part of our framework, responsible for adjusting the proportion of images with various blur or noise levels used for training the neural network. Additionally, this part oversees the execution of the simulation and image restoration parts in parallel using the (MPICH) Message Passing Interface Implementations for Computing Heterogeneous clusters. (Gropp 2002). We will discuss the parameter selection component below.

To begin, our parameter selection part involves defining a dictionary to control the distribution of images with varying blur or noise levels in the training set. The keys of this dictionary are determined by the FWHMs of PSFs and noise levels, while the values represent the mean loss function of all images with the respective FWHM or noise level in the test set. These values are updated at the end of each epoch. Subsequently, we generate a new set of images based on these updated key values using equation (5):

$$\begin{aligned} test_{count}(r_i) &= (total(r_i) - 10) \times (1 - list(r_i)) + 10, \\ train_{count}(r_i) &= (total(r_i) - 10) \times list(r_i). \end{aligned} \quad (5)$$

Here, $test_{count}(r_i)$ and $train_{count}(r_i)$ represent the total number of images with a particular FWHM and noise level (r_i) in the test set and the training set, respectively, $total(r_i)$ represents the total number of images with the same FWHM or noise level, and $list$ represents the parameter dictionary used to store the normalized mean value of the loss function for all images with predefined FWHMs of PSFs or noise levels. The values in $list$ are used to calculate the percentage of images at a particular FWHM and noise level. By generating more blurred images that cannot be restored well by the restoration neural

network in the previous epoch, the network is trained to gain better performance for these images in the next epoch because we have more images of that particular blur or noise level. With the parameter selection part, our neural network can effectively sample the space of blur levels or noise levels for a particular sky survey project and obtain a stable generalization ability for all images obtained by the project.

To accelerate the computationally expensive simulation and image restoration parts, we propose dynamically adjusting the computation load on different computers during the training stage. The parameter selection part controls the data generation and image restoration parts and is illustrated in Fig. 4. The parameter selection part has two ranks. Rank0 trains and tests the image restoration part, which requires significant GPU resources on a single computer. Rank1 generates blur images with the data generation part on several other computers, which require CPU resources. The training data are set with an appropriate number based on the data generation and image restoration speeds. During the training stage, Rank1 generates blur images and sends a ‘TRUE’ signal to Rank0 when all images in the training set of one epoch are generated. Rank1 continues to generate blurred images even after sending the ‘TRUE’ signal. Meanwhile, Rank0 restores these blur images before receiving the ‘TRUE’ signal from Rank1. Upon receiving the ‘TRUE’ signal, Rank0 stops training and evaluates the qualities of all restored images. The parameter selection part calculates the MSE of all restored images in the test set and updates the parameter dictionary. The framework runs continuously with the parameter dictionary as input parameters until the set number of iterations is reached or the loss does not decrease for ten consecutive epochs.

3 PERFORMANCE EVALUATION WITH SIMULATED AND REAL OBSERVATION IMAGES

To assess the effectiveness of our framework, we will test it on both simulated and real observational data. In Section 3.1, we will introduce several criteria for evaluating its performance. In Section 3.2, we will use simulated data to test the framework's performance. Since we have control over the blur and noise levels, we can effectively evaluate the neural network's generalization ability using simulated data. In Section 3.3, we will evaluate the framework's performance using real observational data obtained from the Sloan Digital Sky Survey (SDSS) project (Abazajian et al. 2009). We send images of galaxies directly to the neural network and the results demonstrate the effectiveness of our framework. Further details about our framework will be discussed below.

3.1 Performance evaluation criteria

We employ the peak signal-to-noise ratio (PSNR) as defined by Xu et al. (2014) and the structural similarity (SSIM) as defined by Wang et al. (2004) to assess the quality of the images quantitatively. The PSNR can be found using equation (6):

$$PSNR(img, img_{ref}) = 20 \log_{10} \left(\frac{MAX(img_{ref})}{\sqrt{MSE}} \right),$$

$$MSE = \frac{1}{\sqrt{N \times M}} \sum_{i=1}^N \sum_{j=1}^M (img_{(i,j)} - img_{ref(i,j)}), \quad (6)$$

where $N \times M$ is the size of the image, and img and img_{ref} are restored images and original images. For simulated images, we could obtain img and img_{ref} , and for real observation images, img_{ref} is the mean value of img . The PSNR can reflect the similarity between restored images and blurred images directly. Images with larger PSNR will have better quality. Meanwhile, the SSIM is defined in equation (7):

$$SSIM(img, img_{ref}) = l(img, img_{ref}) \cdot c(img, img_{ref}) \cdot s(img, img_{ref}),$$

$$l(img, img_{ref}) = \frac{2\mu_{img}\mu_{ref} + C_1}{\mu_{img}^2 + \mu_{ref}^2 + C_1},$$

$$c(img, img_{ref}) = \frac{2\delta_{img}\delta_{ref} + C_2}{\delta_{img}^2 + \delta_{ref}^2 + C_2},$$

$$s(img, img_{ref}) = \frac{\delta_{img,ref} + C_3}{\delta_{img}\delta_{ref} + C_3},$$

$$C_1 = (K_1 L)^2,$$

$$C_2 = (K_2 L)^2,$$

$$C_3 = C_2/2, \quad (7)$$

where μ and δ are the average and standard deviation of the greyscale values, respectively, $\delta_{img,ref}$ is the covariance of the greyscale values, L is the dynamic range of the image, and K_1 and K_2 are small arbitrary values (0.001 in this work). The SSIM is a perceptual model that considers the brightness, contrast, and intensity scale of two images simultaneously. The quality of an image is considered better if its SSIM is larger.

3.2 Performance evaluation with simulated images

In order to evaluate the effectiveness of our framework on simulated images, we first simulate an observation scenario carried out by a ground-based telescope with long exposure. We obtain galaxy images

in the r band from SDSS DR 7 and generate blurred images through simulation. Our simulator assumes that the PSF follows the Moffat model, with a FWHM distribution ranging from 2.0–8.0 pixels (equivalent to 0.792–3.168 arcsec) across five different levels and Gaussian noise distributed equally across five different levels with σ ranging from 1.0–15.0. This results in a total of 25 different levels of blurred images. We train our framework using the images generated by the simulator with the aforementioned parameters. Additionally, we generate blurred images with higher levels of blur and noise to evaluate the performance of our algorithm. We divide the FWHM of the PSFs into 13 bins ranging from 2–14 pixels (equivalent to 0.792–5.544 arcsec), with the FWHMs of the PSFs in each bin considered as random variables. We also divide the noise levels into 24 bins ranging from 2–25, with σ considered as a random variable in each bin.

In this section, we apply the evaluation criteria outlined in Section 3.1 to assess the effectiveness of our framework. We utilize box plots to illustrate the performance of our algorithm and the Richardson–Lucy method (RL) with the Gaussian denoising method across varying noise levels and different PSFs. In Fig. 5, results obtained from our algorithm are denoted in green, RL algorithm results in red, and the original blurred data in blue. Upon analysing these box plots, it becomes apparent that our algorithm consistently outperforms the RL method under diverse conditions, encompassing various noise levels and distinct PSFs. Notably, our algorithm exhibits a significant advantage when dealing with datasets characterized by fluctuating noise levels, demonstrating its robustness and stability in the presence of noisy input data.

Furthermore, we conduct a t -test to determine whether a significant performance difference exists between the RL algorithm and our model. Under the null hypothesis (H_0), we assume there is no performance difference between the RL algorithm and our model. The alternative hypothesis (H_1) posits a performance difference between the RL algorithm and our model. In this study, we conduct statistical significance analyses for different scenarios and calculate p -values, all of which are less than 0.05, as shown in Tables 1, 2, 3, and 4. In all these cases, we reject the null hypothesis, signifying a significant performance difference between the RL algorithm and our model. These findings collectively underscore the remarkable image restoration quality achieved by our algorithm. When evaluated using PSNR or SSIM metrics, our algorithm consistently outperforms the traditional RL method. In summary, our test results unequivocally establish the superior performance of our image restoration framework across various observation conditions.

To assess the performance of our algorithm in a qualitative manner, we utilize simulated images to test the effectiveness of our framework. Specifically, we apply the trained RESTORE neural network to restore simulated images that have PSFs and noise levels within the range defined by the training set. The resulting images are presented in Fig. 6. The figure illustrates that the trained RESTORE neural network can enhance the quality of the images significantly. The fine details, such as the spirals and bars of galaxies, can be clearly observed in the restored images. Additionally, we test the performance of our framework on simulated images that have PSFs or noise levels that exceed the range defined in the training set. The results are shown in Fig. 7, and it is evident that the trained RESTORE neural network exhibits a strong generalization ability, even for images with larger FWHM or higher noise levels.

3.3 Performance evaluation with real observation images

First, we demonstrate the efficacy of our framework with real observation data obtained from SDSS Data Release 7 (York et al.

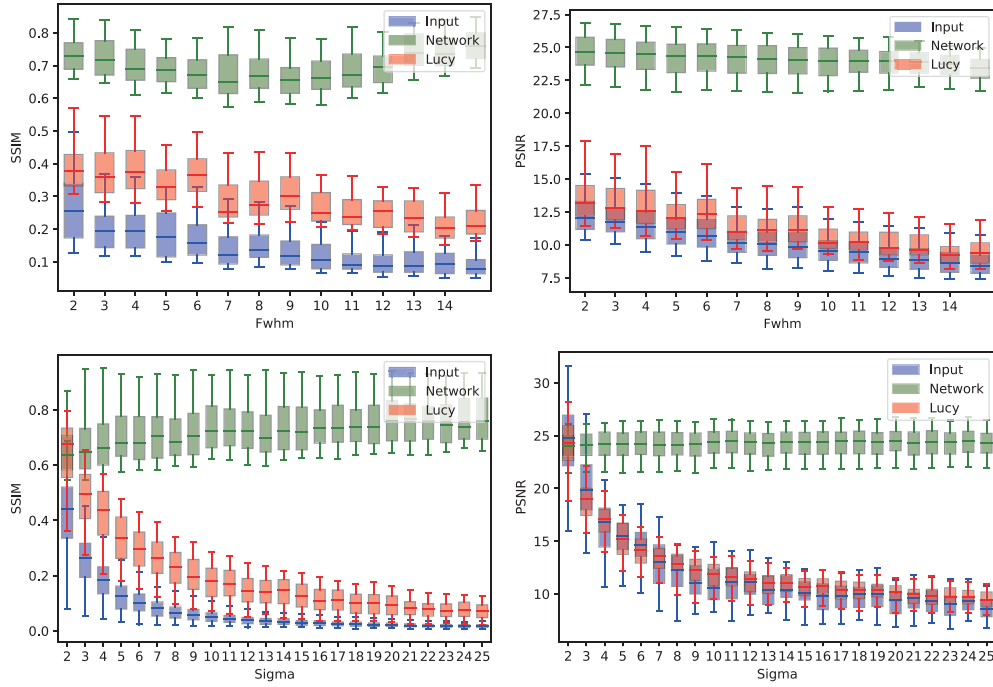


Figure 5. Comparison of different image restoration algorithms under different noise and PSF conditions. The box plot shows the differences between our restoration algorithm (green) and the RL (red) method under different PSF conditions (top panel) and different noise levels (bottom panel).

Table 1. Statistical analysis of SSIM performance for RL and our algorithm under different blur levels (PSFs).

FWHM	2	4	6	8	10	12	14
<i>t</i> -statistic	17.97	20.46	17.57	20.27	20.76	18.79	24.56
<i>p</i> -value	5.96e−38	1.23e−43	5.08e−37	3.40e−43	2.72e−44	7.41e−40	3.21e−52

Table 2. Statistical analysis of PSNR performance for RL and our algorithm under different blur levels (PSFs).

FWHM	2	4	6	8	10	12	14
<i>t</i> -statistic	49.69	73.94	56.31	75.21	76.49	68.05	110.96
<i>p</i> -value	6.02e−90	6.82e−113	4.29e−97	6.88e−114	7.09e−115	4.77e−108	8.31e−137

Table 3. Statistical analysis of SSIM performance for RL and our algorithm under different noise levels.

Sigma	2	9	12	15	18	21	23	25
<i>t</i> -statistic	1.70	17.81	26.46	32.40	32.12	40.43	40.60	43.98
<i>p</i> -value	4.88e−02	3.25e−29	1.05e−40	5.35e−47	1.00e−46	4.40e−54	3.24e−54	1.63e−54

Table 4. Statistical analysis of PSNR performance for RL and our algorithm under different noise levels.

Sigma	2	9	12	15	18	21	23	25
<i>t</i> -statistic	6.66	34.76	42.52	39.62	43.38	50.09	47.96	59.78
<i>p</i> -value	3.52e−09	3.11e−49	1.02e−55	1.98e−53	2.27e−56	4.52e−61	1.19e−59	6.52e−67

2000; Abazajian et al. 2009). We employ our method to restore images of low surface brightness galaxies (LSBGs) as detected by Yi et al. (2022). LSBGs are a class of galaxies with central surface brightness fainter than the sky background, often exhibiting high gas content and believed to be in the early stages of galaxy formation or to have recently undergone a burst of star formation

(Impey & Bothun 1997). Due to their low luminosity, studying LSBGs is challenging (Du et al. 2015), making image restoration algorithms necessary to enhance their image quality before any scientific analysis can be performed. However, it is difficult to use traditional deconvolution-based image restoration algorithms,

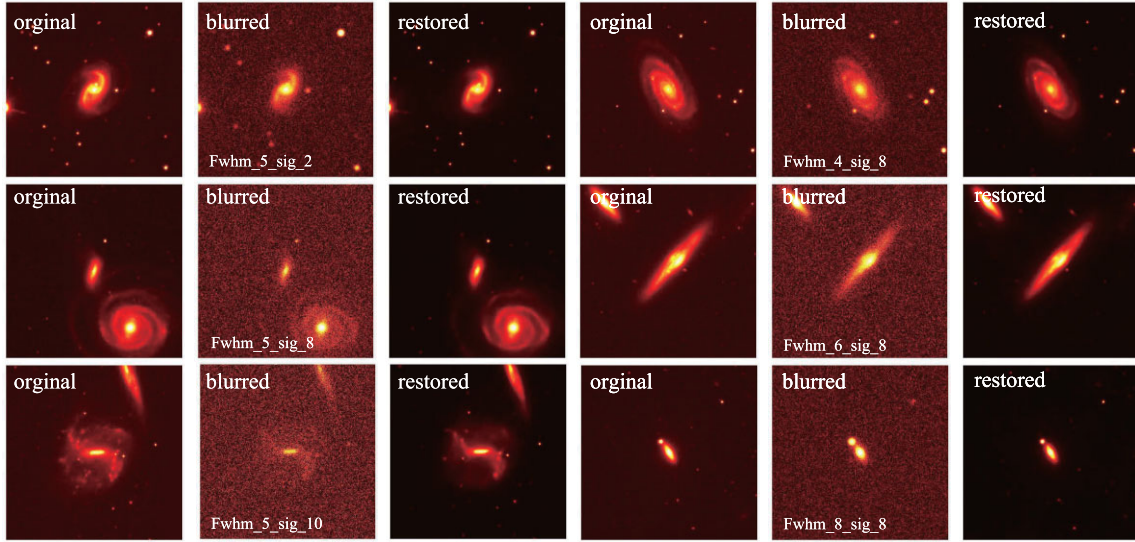


Figure 6. The performance of our framework in the restoration of images that are generated within FWHMs or noise levels defined by the training set.

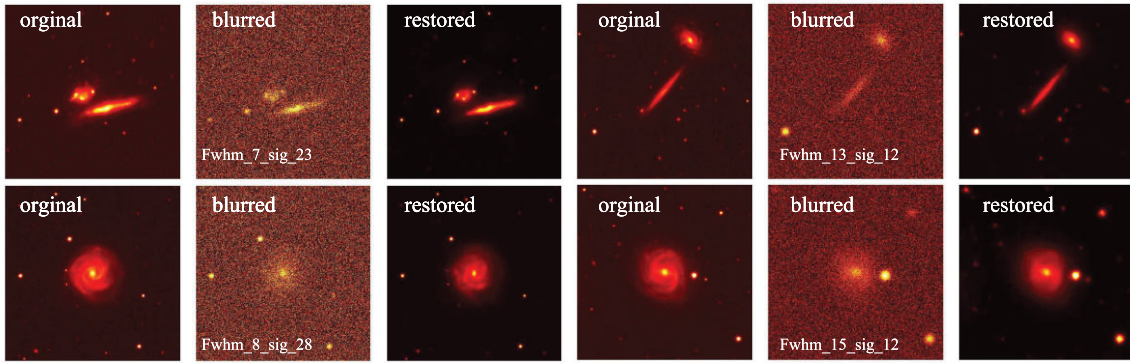


Figure 7. The performance of our framework in the restoration of images that are generated with larger FWHM or higher noise level defined by the training set.

which require bright stars as references, and content-based image restoration methods are ineffective due to the very small number of photons in LSBGs. Thus, we use our framework to restore LSBG images.

These images in the SDSS project are captured by a 2.5-m telescope and have a pixel scale of 0.396 arcsec with an exposure time of 53.9 s. As a result, the PSF and noise levels used in the simulation model in Section 2.2 can reflect the properties of real observation data. Therefore, we train our framework with the simulator and the aforementioned parameters to obtain the weights of the RESTORE neural network. Subsequently, we apply the RESTORE neural network directly to process the LSBG images captured by the SDSS project in the r band. Fig. 8 exhibits both the original images and the images restored by our neural network. Additionally, we also employ the RL deconvolution algorithm (Fish et al. 1995) and select PSF references according to Infante-Sainz, Trujillo & Román (2020) to restore these images for comparison. Furthermore, we present the LSBG images acquired by the DESI Legacy Imaging Surveys (Dey et al. 2019) in this figure to provide a reliable reference. Since larger telescopes are used to execute the DESI Legacy Imaging Surveys, the data obtained from them can improve our evaluation of the effectiveness of our method.

Firstly, as demonstrated in the upper left corner of each figure, our framework can efficiently improve the PSNR of LSBGs in less time (around 10 times faster than the RL method with appropriate prior PSFs). Next, we examine the restored images in detail. Although the RL method could provide effective results, thanks to the accurate PSF model, images restored by the RL method are still affected by strong noise. Our framework effectively restores fine structures of these galaxies, such as spirals, discs, and filaments, and also reduces effects brought by noise. When comparing the restored images with those obtained by the DESI Legacy Imaging Surveys, we find that structures restored by our framework are true. It is worth noting that our framework builds PSFs and the deconvolution procedures for image restoration, so in principle it does not generate artificial features. Moreover, our framework can effectively suppress noise in these images, resulting in some images with even better quality than those obtained by the DESI Legacy Imaging Surveys. Overall, our framework can assist scientists in studying the properties and morphological structures of LSBGs in greater detail.

In our further investigation, our primary focus was on evaluating the algorithm's performance concerning the enhancement of photometry accuracy and detection efficiency. To do this, we randomly select SDSS R -band images, each with dimensions of 1024×1024

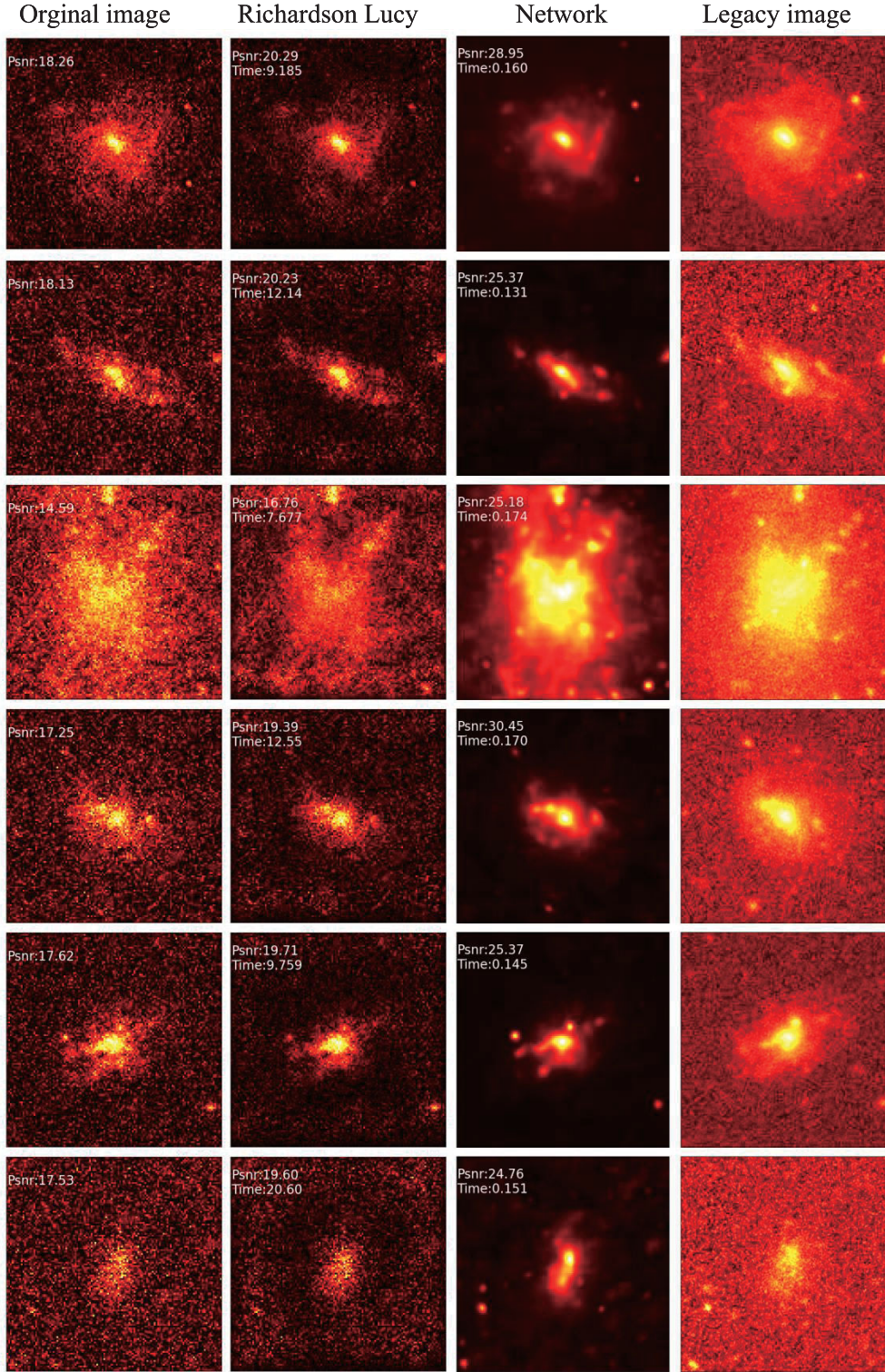


Figure 8. This figure shows original images obtained by the SDSS project in the r band, images restored by the RL method, with references of PSFs provided by Infante-Sainz, Trujillo & Román (2020), images restored by our framework, and images of the same LSBG obtained by the DESI Legacy Imaging Surveys. In the upper left corner of each figure, we show the PSNR of restored images and calculation speed of different methods in restoration of different images. As shown in this figure, our framework could restore blurred images effectively.

pixels. Subsequently, we apply both our image restoration algorithm and the RL algorithm to restore these images. Following restoration, we utilize SEXTRACTOR for detection and photometry (Bertin & Arnouts 1996). The results of this study are presented visually in

Fig. 9. We conduct this analysis for each magnitude, considering a dataset of 2000 images for evaluation. Subsequently, we assess these results through a rigorous statistical analysis. Our findings indicate that the processed data, particularly the data treated with

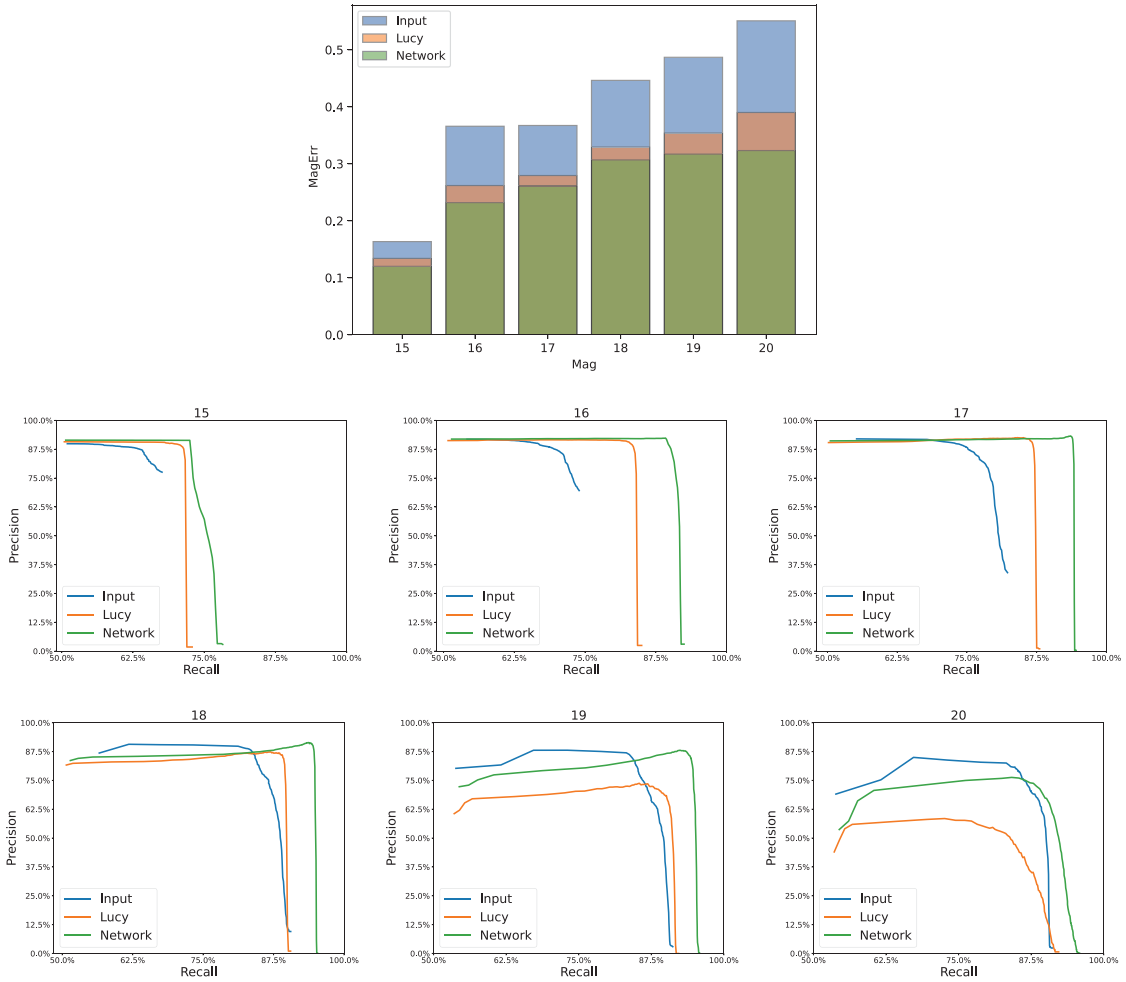


Figure 9. Comparison of photometry and detection results for stars with different magnitudes. MagErr: the average percentage error between photometric measurements and true values. P–R curve: the precision-recall (P–R) curve is a graphical representation of the trade-off between precision and recall for a binary classification model. Precision is the ratio of true positive predictions to the total number of positive predictions, while recall is the ratio of true positive predictions to the total number of actual positive instances.

our algorithm, show a higher recall rate and precision rate. This outcome demonstrates that our algorithm significantly improves the efficiency of celestial object detection, especially for stars with low signal-to-noise ratio. Moreover, our method has proven effective in enhancing photometry accuracy, as vividly depicted in the top panel of Fig. 9. These results provide strong support for the practical applicability of our algorithm in the field of astronomy.

4 CONCLUSIONS AND FUTURE WORK

We have introduced a novel framework in this article for restoring blurred astronomical images by merging deep learning techniques with simulation algorithms. Our framework actively trains the restoration neural network using a simulation algorithm that represents a specific telescope. Once trained, the restoration neural network can produce restored images more efficiently compared with the traditional RL deconvolution algorithm. We have tested our method on both simulated and real observational data and have found that it effectively minimizes the impact of noise

and PSFs, making previously unseen fine structures of galaxies visible.

We have also identified two areas for future improvement. Firstly, we have shown that the physical parameters and prior model used to represent the image degradation process are crucial for our framework. To extend our approach to data from different sky survey projects, we need to develop an adequate parametric PSF model and an adequate telescope simulator. Therefore, we will introduce physics-informed machine learning algorithms as necessary tools to build PSF models. Meanwhile, the digital twin technology is a promising method for generating simulation data according to telemetry data and high-fidelity simulators, and we are currently developing a digital twin as a telescope simulator (Jia et al. 2022; Zhan et al. 2022). Secondly, we use the L_2 norm and FFL to train our restoration neural network, and we could investigate further new regularization methods based on human attention and big data obtained from previous sky survey projects to improve the performance of our framework.

Overall, our proposed framework is suitable for restoring images obtained from future sky survey projects, such as the LSST, *Euclid*,

and the *CSST*. Our framework could help scientists to recognize the morphology of galaxies better, which could increase outcomes from citizen science platforms. Also, our framework could increase the accuracy of shape measurements for galaxies with low signal-to-noise ratio. We are now using our framework to process data obtained by Dark Energy Camera Legacy Survey (DECaLS) Dark Energy Camera Legacy Survey. (Dark Energy Survey Collaboration et al. 2016) for further scientific research. We will deploy our method for data obtained by the *CSST* and *Euclid* in the future.

ACKNOWLEDGEMENTS

This work is supported by the National Natural Science Foundation of China (NSFC) with funding numbers 12173027 and 12173062. We acknowledge science research grants from the China Manned Space Project, No. CMS-CSST-2021-A01. We acknowledge science research grants from the Square Kilometer Array (SKA) Project, No. 2020SKA0110102. Peng Jia acknowledges support from the Civil Aerospace Technology Research Project (D050105) and the Major Key Project of PCL. Jiameng Lv acknowledges support from the Shanxi Graduate Innovation Project (2022Y274).

Funding for the SDSS and SDSS-II has been provided by the Alfred P. Sloan Foundation, the Participating Institutions, the National Science Foundation, the US Department of Energy, the National Aeronautics and Space Administration, the Japanese Monbukagakusho, the Max Planck Society, and the Higher Education Funding Council for England. The SDSS website is <http://www.sdss.org/>.

The SDSS is managed by the Astrophysical Research Consortium for the Participating Institutions. The Participating Institutions are the American Museum of Natural History, Astrophysical Institute Potsdam, University of Basel, University of Cambridge, Case Western Reserve University, University of Chicago, Drexel University, Fermilab, the Institute for Advanced Study, the Japan Participation Group, Johns Hopkins University, the Joint Institute for Nuclear Astrophysics, the Kavli Institute for Particle Astrophysics and Cosmology, the Korean Scientist Group, the Chinese Academy of Sciences (LAMOST), Los Alamos National Laboratory, the Max-Planck-Institute for Astronomy (MPIA), the Max-Planck-Institute for Astrophysics (MPA), New Mexico State University, Ohio State University, University of Pittsburgh, University of Portsmouth, Princeton University, the United States Naval Observatory, and the University of Washington.

DATA AVAILABILITY

We have released the relevant code in the PaperData repository, supported by the China Virtual Observatory (China-VO), and assigned a DOI number (10.12149/101315) to the code.

REFERENCES

- Abazajian K. N. et al., 2009, *ApJS*, 182, 543
 Arcelin B., Doux C., Aubourg E., Roucelle C., *LDES Collaboration*, 2021, *MNRAS*, 500, 531
 Beltramo-Martin O., Correia C., Ragland S., Jolissaint L., Neichel B., Fusco T., Wizinowich P., 2019, *MNRAS*, 487, 5450

- Bertin E., Arnouts S., 1996, *A&AS*, 117, 393
 Dark Energy Survey Collaboration et al., 2016, *MNRAS*, 460, 1270
 Dey A. et al., 2019, *AJ*, 157, 168
 Du W., Wu H., Lam M. I., Zhu Y., Lei F., Zhou Z., 2015, *AJ*, 149, 199
 Fétick R. et al., 2019, *A&A*, 628, A99
 Fish D., Brinicombe A., Pike E., Walker J., 1995, *J. Opt. Soc. America A*, 12, 58
 Fusco T. et al., 2020, *A&A*, 635, A208
 Gan F. K., Bekki K., Hashemizadeh A., 2021, preprint (arXiv:2103.09711)
 Gao W., Zhao X., Zou J., Yang Y., Xu R., Zhang R., Xuebin X., 2017, *Opt. Rev.*, 24, 278
 Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair S., Courville A., Bengio Y., 2020, *Communications of the ACM*, 63, 139
 Gropp W., 2002, in Kranzlmüller D., Volkert J., Kacsuk P., Dongarra J. eds, *Recent Advances in Parallel Virtual Machine and Message Passing Interface*. Springer, Berlin, Heidelberg, p. 7
 Hüllermeier E., Waegeman W., 2021, *Machine Learning*, 110, 457
 Impey C., Bothun G., 1997, *ARA&A*, 35, 267
 Infante-Sainz R., Trujillo I., Román J., 2020, *MNRAS*, 491, 5317
 Jia P., Sun R., Wang W., Cai D., Liu H., 2017, *MNRAS*, 470, 1950
 Jia P., Huang Y., Cai B., Cai D., 2019, *ApJL*, 881, L30
 Jia P., Wu X., Yi H., Cai B., Cai D., 2020, *AJ*, 159, 183
 Jia P., Ning R., Sun R., Yang X., Cai D., 2021, *MNRAS*, 501, 291
 Jia P., Wang W., Ning R., Xue X., 2022, *Opt. Express*, 30, 21362
 Jiang L., Dai B., Wu W., Loy C. C., 2021, in *Proceedings of the IEEE/CVF International Conference on Computer Vision*. p. 13919, <https://ui.adsabs.harvard.edu/abs/2020arXiv212821J/abstract>
 Kingma D. P., Welling M., 2013, preprint (arXiv:1312.6114)
 Li T., Alexander E., 2023, *MNRAS*, 522, L35
 Li Y., Niu Z., Sun Q., Xiao H., Li H., 2022, *Remote Sensing*, 14, 4852
 Long M., Soubo Y., Weiping N., Feng X., Jun Y., 2019, *ApJ*, 888, 20
 Moffat A. F. J., 1969, *A&A*, 3, 455
 Qi G., Kedong W., Hong Z., Guibin L., Chao D., 2014, *Infrared and Laser Engineering*, 43, 1327
 Ronneberger O., Fischer P., Brox T., 2015, in *International Conference on Medical Image Computing and Computer-Assisted Intervention*. p. 234 <https://ui.adsabs.harvard.edu/abs/2015arXiv150504597R/abstract>
 Schawinski K., Zhang C., Zhang H., Fowler L., Santhanam G. K., 2017, *MNRAS Lett.*, 467, L110
 Sun R., Yu S., Jia P., Zhao C., 2020, *MNRAS*, 497, 4000
 Sureau F., Lechat A., Starck J.-L., 2020, *A&A*, 641, A67
 Terry S. K. et al., 2023, *Astron. Telescopes*, 9, 018003
 Wang Z., Bovik A. C., Sheikh H. R., Simoncelli E. P., 2004, *IEEE Trans. Image Processing*, 13, 600
 Wang H., Sreejith S., Lin Y., Ramachandra N., Slosar A., Yoo S., 2022, preprint (arXiv:2210.01666)
 Xu L., Ren J. S., Liu C., Jia J., 2014, *Advances in Neural Information Processing Systems*, 27
 Yi Z. et al., 2022, *MNRAS*, 513, 3972
 York D. G. et al., 2000, *AJ*, 120, 1579
 Zhan Y., Jia P., Xiang W., Li Z., 2022, in *Modeling, Systems Engineering, and Project Management for Astronomy X*. p. 719, <https://ui.adsabs.harvard.edu/abs/2022SPIE12187E..1SZ/abstract>

This paper has been typeset from a \LaTeX file prepared by the author.